Koushik Maharatna · Silvio Bonfiglio
Editors

Systems Design for Remote Healthcare

Springer

Systems Design for Remote Healthcare

Koushik Maharatna • Silvio Bonfiglio
Editors

Systems Design for Remote Healthcare

Springer

Editors
Koushik Maharatna
School of Electronics
and Computer Science
University of Southampton
Southampton, UK

Silvio Bonfiglio
FIMI S.r.l.
Saronno, Italy

ISBN 978-1-4614-8841-5 ISBN 978-1-4614-8842-2 (eBook)
DOI 10.1007/978-1-4614-8842-2
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013952911

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

We would like to dedicate this book to the millions of people suffering from chronic conditions. We will be particularly thrilled if this book paves the way for helping them in future with the development of universally acceptable next-generation remote healthcare solution.

Preface

The prevalence of aging population and chronic diseases worldwide has put the conventional model of healthcare delivery under serious strain. The productivity loss coupled with serious degradation of quality of life has resulted in significant socio-economic challenge that every national government is finding difficult to cope with. This leads to the discussion of radically changing the conventional "reactive" care delivery model to "proactive" and "care delivery anytime anywhere"—the next-generation remote healthcare. The boom of Information and Communication Technology (ICT) together with advances of digital Ultra-Large Scale Integrated circuit design has been projected as the enabling technology to achieve this goal. Although the technological advances are there, development of next-generation remote healthcare system is a massively complex task owing to the integration of heterogeneous technologies into integrated service structure. Numerous research projects have been initiated bringing together expertise from industry and academia over multiple domains to develop the complete next-generation remote healthcare system.

When we started the work towards such a development for next-generation remote healthcare system under the Artemis Joint Undertaking funded pan-European project CHIRON, we also stumbled upon the difficulties of integrating knowledge from different disciplines ranging from pure clinical down to the actual technology innovation/adaptation. However one striking fact we noticed that although there are wealth of information seemingly relevant to such a task is available, it is very difficult to understand which information is actually relevant and which is not. Tons of documents are available on the Internet and unfortunately most of them try to project their own work as the best solution. Also a substantial number of literatures is based on only theoretical aspects and may not be completely applicable in real-life scenario. Therefore the main question for us was which type of solution should we adapt and what part/parts need further innovation. It was a difficult question to answer and the only way to find it out was through trial-and-error method. As a consequence, substantial time was spent on this process to filter out the actual applicable solution.

Another problem we faced was the multidisciplinary nature of the area. We often found that clinicians and technologists do not communicate with the same "language" leaving a huge gap in understanding. The lack of domain knowledge is significant factor of misunderstanding leading to unnecessary and time-consuming iterations. Technology enthusiasts try to innovate solutions which in effect are overengineering from the clinical point of view. On the other hand, the lack of knowledge of the clinicians regarding what is possible with the modern technology forces them to think in more conservative way leading to insufficient functionality of the system. As we got more and more projects in this domain, we observed that the abovementioned phenomenon is quite general and hinders the development of the system a lot.

At that point we conceived the idea of this book to bring together the most useful information pertinent to the development of the next-generation remote healthcare system. The main objective of this book is to give a complete overview of the developmental phase of the system starting from what the clinicians actually expect from such a system. We took the stand that every technological adaptation/innovation needs to be guided by this expectation as the clinicians are the main users (in one sense although there are other stakeholders) of such a system. Therefore we started from that point and tried to incorporate our own experiences into giving a step-by-step guideline on the overall structure and components for such a system development. A big emphasis is also placed on the existing standards that the system needs to comply with. This is a major point which normally is not considered at the beginning in quite a few cases resulting in ineffectiveness of the system. To alleviate all these stumbling blocks, we involved experts from clinic, industry, academia and research communities to write different chapters of this book sharing their knowledge and first-hand experience for such a complex system development.

We start in Chap. 1 describing the vision of the clinicians for the next-generation healthcare which discusses about the problems of the existing remote healthcare, their performance indices and finally the fine points that the developers need to understand from the clinical expectation point of view. Following this, Chap. 2 describes the high-level architecture for the next-generation remote healthcare system and describes the functional and non-functional requirements of the system that may satisfy the clinicians' vision. After that it enters into discussion about the essential components of the system and how they could be integrated together for a functional system. Chapter 3 is dedicated to the discussion on the vital sign monitoring sensors as they are the fundamental components required for monitoring the physiological conditions of patients. Apart from giving a general overview of different types of sensors, it also describes the commercially available sensor products and their specifications. Once the vital sign data is captured it needs to be processed to extract relevant clinical information. A number of signal processing and machine learning algorithms could be invoked for extracting such information and aiding the clinical decisions. Chapters 4 and 5 are particularly dedicated for addressing these issues. One major point we wanted to emphasise in these chapters is that although there are several high-efficiency algorithms available in literature,

they may not be applicable for the target system development owing to the fact that each of the sensor nodes are very much constrained in terms of computational resources and energy—being battery-powered. Therefore new way of implementing such algorithms, and innovation in this field is necessary. Therefore detailed analysis on computational complexity is highlighted in those chapters which may be useful for the developers. A major component, the Patient Health Record (PHR) system is described in Chap. 6. PHR acts as the central part for information storage and exchange amongst the stakeholders of the community and therefore its efficient design is a "must" condition. This chapter describes the standards for PHR systems along with example of how the PHR system could be developed. A main enabling technology for next-generation remote healthcare system implementation—the Wireless Sensor Network (WSN)—is described in Chap. 7. Again we took the stand of conveying the practicality of the implementation process of a WSN in the context of healthcare. Chapter 8 discusses the integration issues of the overall system where main emphasis is placed on the existing standards, test methodology and effective integration process of such a complex system. We feel that no book on next-generation remote healthcare system is complete without a detailed business perspective as it illustrates the main roadblocks for commercial acceptability of the system. Chapter 9 is dedicated for that aspect clearly identifying the main reasons of why the remote healthcare is still not a universal reality and shed light on how to make them universally acceptable. Finally in Chap. 10 we produced two case studies resulting from two European Commission funded projects—one regarding the development of a remote cardiovascular disease monitoring (CHIRON) and another for ICT assisted tele-rehabilitation of autistic children (MICHELANGELO). In particular we described step-by-step solution for the CHIRON system which may act as a guideline for similar types of system development.

As mentioned earlier, the main idea of this book is to allow the designer to get a complete overview of the next-generation remote healthcare system which may enable them to have a clear idea of what is needed and what is not. We intentionally did not get into all the intricacies of each of the components of the system as each of these areas are eligible to be discussed by multiple books and therefore it is not possible to capture everything in one book. However, we provided enough references that one may dig into for more detailed understanding of the individual areas. At the end of each chapters we also provide a list of the open challenges that may stimulate further research/development as we believe, from our practical experience, there are lot of rooms for further developments and innovations.

If the present book enables the designers to think in the effective way for developing the next-generation remote healthcare system, we will consider our objectives are fulfilled.

Southampton, UK — Koushik Maharatna
Saronno, Italy — Silvio Bonfiglio

Acknowledgements

We are first grateful to all the authors for their tireless work in sharing their practical experiences and knowledge in the form of different chapters of this book. They managed to take time out from their busy industrial and academic schedule to make this book a reality. We are also grateful to the European Commission and Artemis Joint Undertaking for providing funds for different projects from which we gathered our experiences and tried to translate those in the form of this book. We also acknowledge the tireless assistance of Dr. Saptarshi Das and Dr. Evangelos Mazomenos with the editing of this book.

Contents

Contributors

Amit Acharyya Indian Institute of Technology, Hyderabad, India

Carlos Cavero Barca ATOS Research and Innovation (ARI), ATOS, Madrid, Spain

Andoni Beriain Department of Electronics and Communication, Centre of Studies and Technical Research (CEIT) and School of Engineering of the University of Navarra (Tecnun), Madrid, Spain

Silvio Bonfiglio FIMI-BARCO, Saronno, Italy

Marcin Brzozowski IHP GmbH—Innovations for High-Performance Microelectronics, Frankfurt (Oder), Germany

Emilio Centaro Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy

Andy Cranny School of Electronics and Computer Science, University of Southampton, Southampton, UK

Božidara Cvetković Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

Alessandra D'Ambrosi Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy

Saptarshi Das School of Electronics and Computer Science, University of Southampton, Southampton, UK

Francesco Fedele Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy

Gian Franco Gensini Faculty of Medicine, University of Florence, Florence, Italy

Hristijan Gjoreski Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

Simon Kozina Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

Artur Krukowski Intracom Telecom, Athens, Greece

Peter Langendoerfer IHP GmbH—Innovations for High-Performance Microelectronics, Frankfurt (Oder), Germany

Mitja Lustrek Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

Jure Lampe Mobili d.o.o, Ljubljana, Slovenia

Koushik Maharatna School of Electronics and Computer Science, University of Southampton, Southampton, UK

Domen Marincic Mobili d.o.o, Ljubljana, Slovenia

Evangelos B. Mazomenos School of Electronics and Computer Science, University of Southampton, Southampton, UK

Steffen Ortmann IHP GmbH—Innovations for High-Performance Microelectronics, Frankfurt (Oder), Germany

Cristiano Paggetti Connected Health Unit, I+ S.r.l., Florence, Italy

Giovanni Pioggia Institute of Clinic Physiology, National Research Council (CNR), Pisa, Italy

Krzysztof Piotrowski IHP GmbH—Innovations for High-Performance Microelectronics, Frankfurt (Oder), Germany

Paolo Emilio Puddu Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy

Stefano Puglia WLAB S.r.L., Rome, Italy

Juan Mario Rodríguez ATOS Research and Innovation (ARI), ATOS, Madrid, Spain

Paola Scarparo Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy

Michele Schiariti Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy

Hector Solar CEIT, Parque Tecnológico de San Sebastián Paseo Mikeletegi, Donostia/San Sebastián, Spain

Gennaro Tartarisco Institute of Clinic Physiology, National Research Council (CNR), Pisa, Italy

Concetta Torromeo CEIT, Parque Tecnológico de San Sebastián Paseo Mikeletegi, Donostia/San Sebastián, Spain

Andrea Vitaletti DIAG, "SAPIENZA" Università di Roma, Rome, Italy

WLAB S.r.L., Rome, Italy

Emmanouela Vogiatzaki RFSAT Ltd., Sheffield, UK

Chapter 1
A Clinician's View of Next-Generation Remote Healthcare System

Paolo Emilio Puddu, Alessandra D'Ambrosi, Paola Scarparo, Emilio Centaro, Concetta Torromeo, Michele Schiariti, Francesco Fedele, and Gian Franco Gensini

Abstract There are few critical questions to answer in relation to remote healthcare systems and the pertinent technologies implied in transferring surveillance and care next to the patients: (1) will they help reducing and/or at least rationalizing care and lower hospitalizations and connected high costs?; (2) will this be really at hand in the near future?; and (3) what are the relevant parameters to obtain and follow in order to make remote healthcare an impacting reality on health systems? "There will be titanic changes ahead—Medicine can and will be rebooted and reinvented one individual at a time": cultural, technological and scientific features will have a stronger relevance and impact on orienting the future revolution and decision than administrative and merely technological processes. However, the future remote healthcare model should be based not only on Darwin's selection of a more and more powerful market (hundreds of billions in Europe), or cumbersome, restrictive, national (or even local) rules. It should start from strong ethical and scientific bases. The great potential benefits should be addressed on humans and on processes of care. Continuity of care, prevention and healthcare throughout the person's life are an excellent beginning. The future will not depend on our technological capacities but on our decision-making capacity of creating systems based on a global approach, in order to cope with the complexity of healthcare environment. Remote healthcare should involve physicians, patients, academics, health service organizations and industries, integrating all their different points of views, not

P.E. Puddu (✉) • A. D'Ambrosi • P. Scarparo • E. Centaro • C. Torromeo
M. Schiariti • F. Fedele
Department of Cardiovascular, Respiratory, Nephrologic, Anesthesiological and Geriatric Sciences, "Sapienza" University, Policlinico "Umberto I", Rome, Italy
e-mail: paoloemilio.puddu@uniroma1.it; alessandra.dambrosi@gmail.com; pl.scarparo@gmail.com; emiliazu83@tiscali.it; concetta.torromeo@uniroma1.it; michele.schiariti@uniroma1.it; francesco.fedele@uniroma1.it

G.F. Gensini
Faculty of Medicine, University of Florence, Florence, Italy
e-mail: gianfranco.gensini@unifi.it

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_1,

neglecting the final user's needs. Use case scenarios are reviewed here, from pre- to post-hospitalization and from the management of congestive heart failure to chronic obstructive pulmonary disease, hypertension and diabetes and renal diseases. It seems that future remote health system should provide the parameters to evaluate trends and short-term predictive indexes and may thus help paving the road for a great revolution in Medicine, based on decision-logic module that computes in an automated way the potential risk of an impending episode by information fusion of heterogeneous sensor, demographic data and evidence-based clinical diagnosis process.

1.1 The Vision of the Future: Remote Healthcare

With the increasingly aging population and prevalence of chronic diseases leading to dramatically escalating health costs and productivity loss the world is facing a significant socio-economic challenge. The conventional reactive care and management model is inadequate in addressing this challenge and new models for delivering cost-effective healthcare services both in-hospital and nomadic environment providing a continuum of care with optimal clinical outcome is much sought at this moment. Advances in Information and Communication Technology (ICT) in its various forms (new miniaturized body-worn sensors, enormous increase in capability for information processing and data management) may be the key ingredient in developing an effective model that may provide proactive care and patient management anytime anywhere linking the in-hospital and nomadic environments through seamless exchange of information making a step change in the care delivery model. More conventionally, such model of healthcare can be defined as remote healthcare or electronic Healthcare (e-Health).

The exchange of information is a key element of the continuity of care. It should be a structured process able to support the interaction among all the stakeholders involved in the care pathway, both in the self-management and lifestyle promotion domain, and acute diseases management. Moreover the exchange of information should ensure "*smooth coordination and efficient transition*" (Dy et al. 2013; Gallagher et al. 2013) among the different phases of a care pathway, in order to promote the continuity of the process.

The impact of a proper exchange of information is not always evaluable in terms of clinical outcomes, especially in the acute diseases management, but there is no doubt that the exchange of information between General Practitioners (GP) and emergency or specialized services can improve the treatment's efficiency; effectiveness could be also improved thanks to the introduction of the technological innovation in the clinical pathway, the reduction of medical errors (McMillan et al. 2013) and the promotion of the patents' interaction with the health care service provider.

In the cardiovascular diseases and chronic metabolic disorders domains there is a strong evidence of the effectiveness and improvement of the real and perceived needs and/or adherence to the treatment in the following domains: management of

diabetes (Chen et al. 2013; Hong and Kang 2013) education of patients with chronic disorders (Kearns et al. 2012) and management of heart failure (McBride et al. 2013) whose treatment is also supported by simple tools of Telemedicine or Telematics Medicine (Gabbrielli 2013).

Innovation of healthcare scenarios (Siddiqui and Pollack 2013), radical changes in the local communities and continuous updating of technologies even according to the different cultural issues and movements of patients or family groups (i.e. migrants people) could further limit the effectiveness of the continuity of care, a suitable exchange of information and proper coordination of workflow activities in the care pathway. We think only Institutional choices in healthcare could facilitate the accessibility to social-care services (Davies et al. 2011); these choices should promote a global approach from the organizational point of view and a national administration level, also taking into account of the multi-sectorial nature of these services. The remote healthcare can play a strategic role in facilitating the acquisition, management and exchange of clinical and personal information, and in making easier the foreseen controls established by the care protocols without moving the patient and/or the physician or the remaining medical staff.

The correct application of ICT innovations will be probably the keystone to promote new sustainable assistance methodologies, also in a remote way. We could remark that "to assist" (*ad sistere* = stop next door, stay close) seems to be in contradiction with the remote concept, but it is not that remote healthcare making spatial or cultural distances; on the contrary it arises as a key element to reduce them or tear them down. The most recent Action Plan of the European Union goes in this direction (2012-2020 eHealth AP 2012) and it promotes the resolution of the most urgent healthcare problems, including:

- "To improve chronic disease and multi-morbidity (multiple concurrent disease) and management and to strengthen effective prevention and health promotion practices";
- "To increase sustainability and efficiency of health systems by unlocking innovation, enhancing patient/citizen-centric care and citizen empowerment and encouraging organisational changes";
- "To foster cross-border healthcare, health security, solidarity, universality and equity";
- "To improve legal and market conditions for developing e-Health products and services".

Anyway a strong cultural, technological and scientific change is needed in order to reach the aforementioned requirements and remote healthcare may play a strong part in these aspects.

Cultural: a continuous "collaborative elaboration and negotiation of e-Health law" is necessary, so as a proper sharing of e-Health strategy by everybody, not only by Institutions but also citizens, researchers and clinicians. This is a real cultural revolution that is only at the beginning. The example of some very recent "population model", like Geneve's one, confirms this issue and highlights some

difficulties (Geissbuhler 2013). There are a lot of lessons to be learned to avoid mistakes, like that one of the United Kingdom National Health Service (NHS) which reported "delay and frustration" due to the implementation of not suitable models (Robertson et al. 2010; Sheikh et al. 2011) many years ago.

Unlike of the opinion of the most part of people (citizens and rulers), quite all the current remote healthcare applications (devices, software, networks, Patients Summary and Electronic Health Record) weren't created to provide objective answers to the real and perceived needs of citizens. Purposes of the remote healthcare applications are not categorized based on relevance or ethics criteria (Catwell et al. 2009). The market is fragmented, not structured and strongly influenced by "policymakers and techno-enthusiasts" (Black et al. 2011). Policy makers are asked to carry out a deeper and more careful analysis on actions plans to be proposed (Greenhalgh et al. 2011); this is because "The complexity of contemporary healthcare, combined with the multiple stakeholders in large technology initiatives, means that national e-Health programs require considerably more thinking through than has sometimes. . .". According to the state of the art outcomes, a continuous and systematic review of results is strongly recommended (Catwell and Sheikh 2009), moreover it seems there is still a gap between postulated and empirically proved benefits of the remote healthcare (Black et al. 2011; Greenhalgh et al. 2011). Public institutions role in the health infrastructure control (also in remote healthcare) should be enhanced (Bloland et al. 2012).

Technological: a technological revolution is necessary too, in order to make available to everybody (and really used by everybody) an infrastructure for sharing and moving on real information, not only *.pdf (or other format) documents that can just bring about a high number of errors and misunderstandings. A platform for cooperation and federation of different services and actors is the most likely proposal. This cooperative platform should be totally interoperable and every interested stakeholder (citizen, patient, medical professional, researcher, student) should insert the real necessary information. This information is citizens' property and should be protected, but also well described and characterised by attributes able to allow the current and future use. Literature is wide and complete on this matter (Blobel 2013; Giovanni et al. 2013; de la Torre-Díez et al. 2013; Wang et al. 2013; White et al. 2013).

Using a sharing interoperable platform on one hand will decrease single and not coordinated experiences and on the other will spontaneously bring to convergence and homogeneity of different languages and facilitate the definition of common contents. Advantages of this approach for medical research are evident in various domains: public health, clinical epidemiology, research on drugs, surveillance systems of the population health status and clinical research.

Scientific: we are facing an evolving scenario in the modern System Medicine. More than 40 years ago Sacket (1969) brilliantly realised the Evidence Based Medicine which was "enslaved" to the medical practice in the recent past (Sackett 2002; Sackett et al. 1985). Nowadays research is strongly oriented to the interpretation of the complexity, abandon of reductionism (Ahn et al. 2006a) and implementation of "Creative Destruction of Medicine", as Topol wrote in a recent book

(Topol 2012): "there will be titanic changes ahead—Medicine can and will be rebooted and reinvented one individual at a time". The Topol's perspective is towards a clinic individual approach that is scientifically effective by itself, without the need of evidences from similar cases (that are usually referred to the analysis of the average behaviour of parameters affecting similar diseases in different persons). This approach is now possible (i.e. it will be possible) not because we have a great quantity of information (from big "analytics" projects that decompose humans in elementary components up to molecules), but thanks to the capability of design models of complex phenomena allowing an integrated and uniform vision of the single patient.

Nowadays this systemic approach is possible not because it is merely derived or derivable from the system biology (Ahn et al. 2006a, b; Gaddi et al. 2013; Topol 2012); it will be achieved only if method could be suggested by a multidisciplinary team of clinicians, experts in computational and theoretic domains working together (Clermont et al. 2009). This approach could be more effective if it is proposed from the beginning, leaving the traditional interdisciplinary vision (not enough to take into account the real complexity), and redesigning the study matter and the models to be implemented. New ICT research instruments (including computational ones and the new emerging models on interactions between intelligent software and researchers' abilities to synthesize information) will simplify the definition of new research paths which could take into account the real human complexity. Medicine as human science has to be invented at all levels and remote healthcare platform is one of the possible instruments to do this.

These features (cultural, technological and scientific) are deeply joined together. Each of them should have a stronger relevance and impact on decision than administrative and merely technological processes. These last requirements instead characterize current pseudo-progress of ICT applications in the health domain of industrialised and not industrialised countries.

The future remote healthcare model should be based not only on Darwin's selection of a more and more powerful market (hundreds of billions in Europe), or cumbersome, restrictive, national (or even local) rules. It should start from strong ethical and scientific bases. The great potential benefits should be addressed on humans and on processes of care. From this point of view continuity of care, prevention and healthcare throughout the person's life are, by definition, an excellent beginning and reference point for every health politics including the remote healthcare. The greatest guarantee of success will be the capability of keeping linked to the person, little and wide technological revolutions (not always useful) and more properly scientific innovations creating new interpretation models (Clermont et al. 2009).

The future of remote healthcare does not depend on our technological capacities, which are already widely available although need to be adapted properly according to the actual need, but on our decision-making capacity of creating systems based on a global approach, in order to cope with the complexity of healthcare environment (Kuhn 1996). Remote healthcare involves physicians, patients, academics, health service organizations and industries. All these stakeholders have different

expectations about remote healthcare development. Governments want to cut costs, industries want to make business, patients want a better healthcare, and physicians want an easier and faster way to get and provide information. A good approach for development of remote healthcare should integrate all these different points of views, and it should especially not neglect the final user's needs.

1.2 Remote Healthcare: A General Overview

There are at present three simple yet crucial questions to answer: (a) does remote healthcare systems and the pertinent technologies implied in transferring surveillance and care next to the patients, in a new patient-centric approach not only at home (Bonfiglio 2012), help reducing and/or at least rationalizing care and lower hospitalizations and connected high costs?; (b) will this be really at hand in the near future?; and (c) what are the relevant parameters to obtain and follow (Puddu et al. 2012) in order to make remote healthcare an impacting reality on health systems?

Under the above mentioned questions, definition of remote healthcare, the metrics to judge its effectiveness and the currently used telemedicine structures for facilitating implementation of the overall concept of remote healthcare in health systems are strictly interrelated prerequisites and they should be clearly defined: our views are therefore addressed to these points at first glance. Indeed, it is not possible to separate technology from the crucial parameters to measure and follow, although the time-varying nature of risk may present some peculiarities whereby adaptation and new approaches may be needed. There has been an ambiguous insight from earlier investigations not only due to power issues but also frequently to the few standard, static parameters taken into consideration (Abraham et al. 2011; Chaudhry et al. 2007, 2010; Clark et al. 2007; Conraads et al. 2011; Eapen et al. 2011; Inglis et al. 2010; Klersy et al. 2009; Koehler et al. 2011; Riezebos 2011; Swedberg et al. 2011; Tompkins and Orwat 2009; Wade et al. 2011). On the other hand, since as many of 70 % of inpatient beds in the UK NHS are occupied by people with chronic conditions, the idea to use remote healthcare system to try to manage part of those certainly makes sense (Eapen et al. 2011) and should concentrate all possible efforts not only from Governments and administrative agents but also from technicians, industries and doctors.

1.2.1 Definitions of Telemedicine

Telemedicine is the ability to provide interactive healthcare utilizing modern technology and telecommunications. The terms "telemedicine" or "remote healthcare" and "e-Health" encompass both "telemonitoring" and telephone support. With telemonitoring, patients transmit data on their vital signs for real time monitoring

via a communication link or by store and forward systems. With telephone support, healthcare providers support patients or carers via the standard telephone system, which may involve monitoring of vital signs reported by patients. Methods involving standard telephone are relatively old-fashion and present several disadvantages. It is important to imagine new methods such as mobile phone communication (Bonfiglio 2012), also to enter a more nomadic era whereby monitoring is not just home confined.

On the other hand, telemedicine allows patients to visit with physicians live over video for immediate care or capture video/still images and patient data are stored and sent to physicians for diagnosis and follow-up treatment at a later time. Whether a patient lives in the center of London or Rome or deep in Sahara, telemedicine is an invaluable tool in healthcare. Instead of traveling to the nearest specialist, which depending where the patient live could be anywhere between a 45-min drive and an 18-h car ride up sanded roads, the patient's service provider connects the patient directly to any medical specialist via telemedicine. The specialist then may hear the medical history and current condition directly from the patient instead of reading a written account dictated by the first health provider. There might be medical peripherals (such as electrocardiographic, echocardiographic signals or nasopharyngoscope probes) that might be handled by the first-line health provider to allow the specialist receiving important direct elements for diagnosis. There might be direct questions from the specialist and immediate replies from the patient. At the end of the teleconsultation the specialist can diagnose and recommend treatment immediately.

In general, the existing remote healthcare or telemedicine systems can be classified in three broad categories: real-time telemedicine, asynchronous telemedicine and home care. The real-time telemedicine is the most common type of remote healthcare approach. Like the example above, live video allows the provider, patient and specialist to all communicate together to achieve the optimal outcome for the patient. It might be used in outpatient specialty consultation, for physician supervision of non-MD first-line health providers. It requires large bandwidths (>256 Kb). Asynchronous telemedicine is used when both the clinician and health service provider are not available or not required at the same time. The provider's voice or text dictation on the patient's history, current affliction including pictures and/or video, radiology images, electro- and echocardiograms, are attached for diagnosis. This record is either emailed or placed on a server for the clinician's access to follow up with his/her diagnosis and treatment plan. On the other hand, it is with home care technologies that the future approaches. When a patient is in the hospital and he/she is placed under general observation after a surgery or other medical procedure, the hospital is usually losing a valuable bed and the patient would rather not be there as well. Home health allows the remote observation and care of a patient. Home health equipment consists of vital signs capture, video conferencing capabilities, and patient stats can be reviewed and alarms can be set from the hospital nurse's station, depending on the specific home health device. There are newer systems to support higher bandwidth capabilities.

1.2.2 Benefits of Telemedicine

There are benefits in the telemedicine system for both the spoke sites and the patients. The formers receive education from the providers and the specialists in particular. There is a better health outcome for the patients around the spoke site. The community around the spoke site enhances confidence in the local healthcare since they know that there are opportunities of continuing medical education for the local health providers via the telemedicine system.

There are important advantages for the patients as well. The loved ones remain in their community with close family support and cost savings from not having to travel extensively. When urgent care is needed there is the possibility to look for immediate consultation. Confidentiality is increased since the patients receive consultation from the specialist without anyone else knowing apart from the GP with whom the patient looks for telehealth. Finally, early diagnosis prior to escalated medical episodes is more frequently obtained and in urgent situations the patient is more adequately stabilized prior to transport. All this has economic consequences: if patients remain in their communities they also keep their money there and expenses are not afforded in other places where medical advice is looked for in absence of telemedicine.

1.2.3 Metrics for Evaluating Effectiveness of Remote Healthcare

Governments look at remote healthcare/telemedicine as a tool for reducing health-related costs, which is far from being ascertained today, notwithstanding decisors' driving force and important investments in this field (Eapen et al. 2011; Department of Health 2011; Henderson et al. 2013; Steventon et al. 2012, 2013). Very few large and powered random studies were performed in the past 10 years (Klersy et al. 2009) and this is quite surprising if one considers the consequences of a positive demonstration of effectiveness from the public health perspective and from the costs involved and the potential economies. It will be then essential to run clinical investigations by pathology class, publically funded, to not only examine effectiveness but also the association between outcomes and costs.

The economics of telemedicine have not been fully investigated (Henderson et al. 2013; Steventon et al. 2012, 2013), despite accumulating evidence on the effectiveness of several monitored measures, spanning from heart rate, to body weight and, more recently to non-invasive measures of fluid retention or pulmonary pressure (Abraham et al. 2011; Chaudhry et al. 2007, 2010; Clark et al. 2007; Conraads et al. 2011; Eapen et al. 2011; Inglis et al. 2010; Klersy et al. 2009; Koehler et al. 2011; Riezebos 2011; Swedberg et al. 2011; Tompkins and Orwat 2009; Wade et al. 2011). A great effort should be paid on metrics: hospitalizations, satisfaction (also from the family side), overall mortality, Quality Adjusted Life

Year (QALY) and Disability-Adjusted Life Year (DALY). QALY is a measure of disease burden; including both the quality and the quantity of life lived (Prieto et al. 2003). It is used in assessing the value for money of a medical intervention. The QALY model requires utility independent, risk-neutral and constant proportional tradeoff behavior. The QALY is based on the number of years of life that would be added by the intervention. Each year in perfect health is assigned the value of 1.0 down to a value of 0.0 for being dead. If the extra years would not be lived in full health, for example if the patient would lose a limb, or be blind or have a myocardial infarction or to use a wheelchair, then the extra life-years are given a value between 0 and 1 to account for this.

On the other hand, DALY reflect years of life lost from premature death and years of life lived in less than full health and can be used to define the overall disease burden. DALY is becoming an increasingly common term in the field of public health impact assessment. Indeed it extends the idea of potential years of life lost due to premature death and it includes equivalent years of healthy life lost by being in poor health or disability. Therefore by DALY, mortality and morbidity are combined into a unified metric. By this, DALY across a given population, translates the important concept of the burden of disease and might be thought about as a measurement of the gap between current health status and an ideal health situation where the entire population lives to an advanced age, free of disease and disability. It may be useful to consider:

$$\mathrm{DALY} = \mathrm{YLL} + \mathrm{YLD}$$

where YLL = years of life lost and YLD = years lived with disability, since this underlines that time is the most important measure to score the effects of chronic illness, either due to lost time as a consequence of premature death or time spent disabled by disease. Under these assumptions, 1 DALY is equal to 1 year of healthy life lost. If DALY is measured, there are surprises in the rank order of disease burden in the population and this may be used to design or decide about priorities in public health and the needed funding. For example in Paris (France), the six leading causes of 242,061 DALYs in 1999 were: alcoholic psychosis and dependence (6.5 % of total), lung cancer (5.7 %), ischemic heart disease (4.8 %), depression (4.4 %), dementias (4.2 %) and arthritis (3.9 %) and for the 4 of the 6 causes, the majority of DALYs came from years lived with disability, rather than mortality, mainly contributing to DALYs in lung cancer and ischemic heart disease (Granados et al. 2005).

The meaning and usefulness of the QALY is debated (Mortimer and Segal 2008). Perfect health is hard, if not impossible, to define. Some argue that there are health states worse than being dead (in fact some particular QALY indices may have a negative value). Perhaps the largest limit is that the effects of a patient's health on the quality of life of others (e.g. caregivers or family) do not figure into these calculations.

If a metrical tool is developed as a step for objective evaluation of remote healthcare protocols and their usefulness, it is essential to consider a closer

time-varying monitoring of physiological parameter changes (Klersy et al. 2009) and to try to include them among the risk indices for risk-assessment at short- or very-short term (Puddu et al. 2012). Although this may be useful in relation to Congestive Heart Failure (CHF) (Puddu et al. 2012), the situation may well extend to other chronic illnesses as Chronic Obstructive Pulmonary Disease (COPD) and Diabetes. It will be of paramount importance to establish in the future whether by remotely monitoring multiple parameters one may catch specific time-varying new dimensions that may highlight risk of impending clinical deterioration, thus calling for prompt interventions aimed at interrupting vicious cycles and so maintain health stability.

Finally, since the vast majority of studies on telemedicine with few exceptions were undertaken in the United States, the local applicability to European care systems and life-philosophy should be accurately investigated and the interference with diet, habits and social relationships should be closely taken into account.

1.3 The Reality of Effectiveness of Current Remote Healthcare Systems

The largest economic investigation ever funded in UK went wrong with telemedicine (Henderson et al. 2013): 3,230 people with a long-term condition (CHF, COPD, and Diabetes) were recruited into the Whole Systems Demonstrator (WSD) telehealth trial between May 2008 and December 2009 and 845 were randomized to telehealth and 728 to usual care. Intervention participants received a package of telehealth equipment and monitoring services for 12 months, in addition to the standard health and social care services available in their area. Controls received usual health and social care. Primary outcome for the cost effectiveness analysis was incremental cost per QALY gained. In the end result, the QALY gain by patients using telehealth in addition to usual care was similar to that by patients receiving usual care only, and the total costs associated with the telehealth intervention were higher.

Cost effectiveness acceptability curves were generated to examine decision uncertainty in the analysis surrounding the value of the cost effectiveness threshold. The incremental cost per QALY of telehealth, when added to usual care, was £92,000. With this amount, the probability of cost effectiveness was low (11 % at willingness to pay the National Institute for Health and Clinical Excellence, NICE, threshold of £30,000; >50 % only if the threshold exceeded about £90,000). In sensitivity analyses, telehealth costs remained slightly (non-significantly) higher than usual care costs, even after assuming that equipment prices fell by 80 % or telehealth services operated at maximum capacity. However, the most optimistic scenario (combining reduced equipment prices with maximum operating capacity) eliminated this group difference (cost effectiveness ratio £12,000 per QALY: Fig. 1.1). The investigation was concluded by stating that telemedicine does not seem to be a cost effective addition to standard support and treatment.

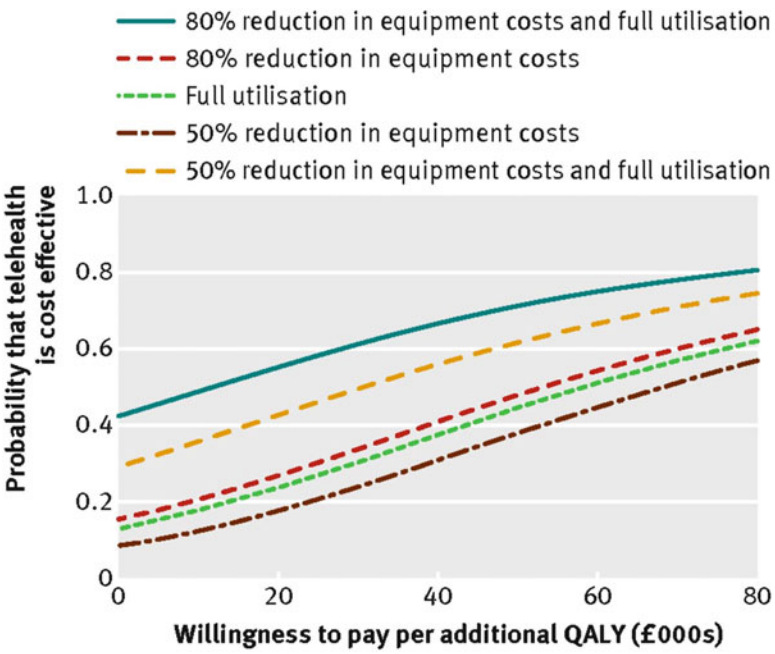


Fig. 1.1 Willingness to pay per additional QALY and the probability that telehealth is cost effective (reproduced from (Henderson et al. 2013) with permission from BMJ Publishing Group Ltd.

Slightly different results, using the same randomized study and a different analysis (Steventon et al. 2012), were published few days earlier: 46.8 % of intervention participants were admitted to hospital, compared with 49.2 % of controls. Unadjusted differences were not statistically significant (odds ratio: 0.90, 95 % CI: 0.75–1.07, P = 0.211). However, they reached statistical significance after adjusting for baseline covariates, but this was not replicated when adjusting for the predictive risk score. A significant reduction in service use, at least in terms of results assessed over 12 months, was not concluded in favor of telemedicine (Department of Health 2011).

Nevertheless, in the UK government's view the abovementioned telemedicine programs already provided evidence of the benefits of remote healthcare and the Department of Health announced its plan to collaborate with industry to introduce the technology on a wide scale (Steventon et al. 2012). This is clear proof of the importance given by administrators to the potential of remote healthcare to at least take care of elderly people outside the hospital.

It was pointed out importantly (Fig. 1.1) that on the assumption of reduced equipment costs and increased working capacity, the probability that telehealth is cost effective would be about 61 %, assuming a willingness to pay threshold of £30,000 NICE trade-off per QALY (Henderson et al. 2013). Thus, industry and

the Governments should work on the first arm of the balance, which may however be an impossible task if the second arm, increasing working capacity, is not concurrently ameliorated.

1.4 Use Case Scenarios of Remote Healthcare

It is already evident that amidst the inconclusive evidences on the benefit of the current form of remote healthcare systems, there is a possibility that advances in technology will provide a sophisticated platform for mass information collection which could be exploited to develop radically new diagnostics leading to further optimization of the system in terms of care delivery and cost and hence the overall benefits. Technological advances over the past decades have allowed improved diagnosis and monitoring of patients with acute coronary syndromes as well as patients with advanced CHF (Fedele and Scalvini 2009). High-quality digital recordings transmitted wirelessly by cellular telephone networks have increased the pre-hospital use of transportable electrocardiogram machines as well as implantable devices for arrhythmia monitoring and therapy (Fazio et al. 2010; Fedele et al. 2010). Based on these facts, there are already several attempts made by different national governments which see more technological utilization and restructuring of care delivery model.

A sophisticated example of telemedicine in Europe is the French model where there are two networks that have developed the best advantages of electronic medicine: the perinatal network, with the use of videoconferencing for staff meetings and the network of oncology, though the use of a regional instrument for a prescription distribution of chemotherapy. The doctors employ software installed on a single regional platform. This allows a patient to begin chemotherapy, then go to the place of reference and return at the first one without any interruption of his care. In Italy an application example is the Lombardy Region, by DGR N. VII/20592 of 11/02/2005 and N. DGR VIII/1375 of 14/12/2005, as identified two routes, then be allowed to structures with DGR VIII/2471 of 11/05/2006:

1. Path post-surgery domiciliary hospitalization;
2. Remote monitoring of home health care path for medium-severe CHF.

However it is important to understand the potential use case scenarios where remote healthcare may have maximal impact. In the following we present a few of such cases. At this point it is pragmatic to make a distinction between a use case and use case scenario. A use case represents the actions that are required to meet a goal. It may have multiple “paths” that can be taken by any user at any one time. A use case scenario is a single path through the use case; it is a description that illustrates, step by step, how a user is intending to use the system, essentially capturing its behavior from the user’s point of view. Such use case scenarios, in the context of remote healthcare could be broadly categorized in three classes: pre-hospitalization, hospitalization and post-hospitalization and patient management.

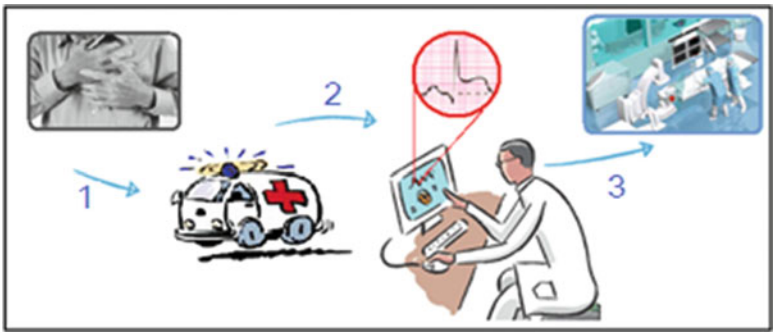


Fig. 1.2 Acute myocardial infarction with ST segment elevation patient management. (1) Ambulance call, (2) the ECG is recorded in the ambulance and sent wirelessly to the cardiologist, (3) the cardiologist send the patient to the catheterization laboratory

1.4.1 Pre-hospitalization

A particularly attractive application for remote healthcare in pre-hospitalization phase is in treating patients with Acute Coronary Syndrome (ACS). Patients with ST-elevation myocardial infarction (STEMI) should receive reperfusion therapy as soon as possible to reduce mortality and morbidity. Guidelines recommend primary Percutaneous Coronary Intervention (PCI) as the preferable strategy, if this can be performed in a timely manner. Thus, to treat STEMI patients within the recommended time limits, early diagnosis is essential in order for patients to be transported to a PCI center in due time. This has led to an increased interest in pre-hospital diagnosis. The diagnostic modality of choice has been the electrocardiogram (ECG). The ECG is recorded in the ambulance and transmitted to a PCI center. The ECG recording is interpreted by the on-call cardiologist and after consulting the patient and/or paramedic a decision is made to either direct the patient for primary PCI, in case of STEMI, or carry him/her to the nearest local hospital for further diagnosis and initial pharmacological therapy. The entire philosophy is depicted in Fig. 1.2.

Increasing the treatment efficacy in ST elevation myocardial infarction and reducing the door-to-balloon time, defined as the time between the arrival at the hospital and the first balloon inflation during PCI, are major targets of contemporary patient management (Bradley et al. 2006). Only a minority of hospitals treats patients with STEMI within 90 min after their arrival and hospital performance has not improved substantially in recent years (Jencks et al. 2003; McNamara et al. 2006). The implementation of telemedicine technology reduce this crucial time period. A simple approach is to transmit out-of-hospital ECG by means of cellular phone cameras, alerting catheterization teams prior to patient arrival at the hospital. In addition, there are devices that transmit 12-lead ECG tracings via regular and mobile telephones and even some that have the capability of interpreting the ECG findings (Hsieh et al. 2009). In order to reduce the time to

catheterization, the pre-hospital transmission of the ECG of patients suffering from chest pain to the cardiologist's mobile telephone is indispensable. The impact of pre-hospital ECG recording and interpretation in patients suspected of acute myocardial infarction should not be underestimated. It enables a more widespread access to rapid reperfusion therapy, thereby reducing treatment delay, morbidity and mortality (Sable 2001).

Telemetry may also enable the use of pre-hospital thrombolytic treatment in STEMI, thus reducing the call to treatment times, especially in a rural setting. However, this benefit must be balanced against the very small proportion of eligible patients identified as suitable for pre-hospital thrombolysis.

Telemedicine could be an effective approach for primary prevention. As an example, telecardiology can be used as a screening tool for cardiovascular disease in large populations. Effective prevention in the population has the potential to reduce the prevalence of the cardiovascular disorder and lower the costs of these disorders to the community. Vantage telecardiology enables rapid access to early diagnosis and treatment of patients with potentially life-threatening cardiac conditions, reducing unnecessary costs and achieving significant cost savings. In Italy a telecardiology service is in use for preventing sudden cardiac death in the young people. More than 10,000 high school students were screened in their schools with 12-lead ECG, which were then wirelessly transmitted to the telemedicine center where a cardiologist interprets the ECG. Then the reports with the instructions for the care or other exams were sent to the students. In a population of these apparently healthy 10,000 students about 21 % was found with an abnormal ECG and was subjected to second level tests. In this case, telemedicine has allowed identification of the young people potentially at-risk of sudden cardiac death within a large population.

1.4.2 Hospitalization

The telemedicine in the hospital scenario is daily exploited between different practitioners and different hospitals. The intra-hospital telecardiology services mainly aim to connect the peripheral hospitals and the hospitals of excellence or district. In this case the system can rationalize access to the operating units of hemodynamic, cardiac critical care and cardiovascular radiology.

Utilization of telemedicine has increased significantly over the last 10 years, with radiology and cardiology being the biggest users. Telecardiology is growing faster than any other field in telemedicine with echocardiography being the primary form of telecardiology. Evolving computer technology and data transfer modalities are modifying the way cardiologists perform, interpret and store echocardiograms. Other telecardiology applications include digital medical record transfer, radiography, stethoscopy and angiography. Evolving telemedicine technology has the potential to improve access to echocardiography diagnoses in the intensive care unit, emergency room and newborn nursery. The two primary modes of

telemedicine practice are "store and forward" and "real-time" videoconferencing. A digital echocardiogram can be stored at one site and forwarded across a telemedicine network to a receiving station for review at a later time. Pediatric cardiologists often favor real-time telemedicine because of the ability to guide sonographers with limited experience in congenital heart disease.

1.4.3 Post-hospitalization

Probably the most important benefit of adoption of population-wide remote healthcare, in particular the remote monitoring, is in the effectiveness of patient management in the post-hospitalization phase. It is known that many patients encounter a variety of problems in the first weeks after they have been discharged from hospital to home. In recent years many projects have addressed discharge planning, with the aim of reducing problems after discharge. Chronic illnesses, such as COPD, Diabetes, CHF and hypertension represent a significant burden of disease that needs effective management.

Management of CHF: Effective use of telemedicine in managing CHF patients has been expanding considerably worldwide. Many studies have demonstrated a substantial reduction in hospital admission, in the length of hospitalization and an increase in survival among CHF patients who were managed by telemedicine (Eisner et al. 2006). The remote patient monitoring significantly reduced the risk of death and the length of hospitalization for any cause, including CHF, compared with usual care (Diemberger et al. 2008). Typically a device is installed in the patient's home and another at the CHF clinic. Preferably the patient's device is portable and capable of both videoconferencing and data transmission, so it can be managed by the healthcare professional using a remote control. According to individual arrangements patients could be instructed to record blood pressure, heart rhythm, and weight and give a subjective description of the symptoms every morning, then transmit data daily, just prior to or during a scheduled remote follow-up. The patients' vital signs and weight are transmitted daily to the system's monitor center; the monitor center's nurse telephone the patient to assess a number of parameters, including well-being, frequency of specific symptoms, adherence to treatment and to remind him/her of the regular use of medication. The nurse instructs patients with "red alarms" (e.g., increase of weight, modification of diastolic and/or systolic blood pressure) and advises the doctor (Fig. 1.3). Data can then be reviewed before the videoconference with the patient, so treatment can be changed if needed. It is important that the healthcare professional be empowered to change medication regimens so that delay is diminished and the correct treatment is established in a timely manner. In addition to objective technological monitoring, most telemonitoring programs include subjective questionnaire regarding the patient's health and comfort. This questioning can take place automatically over the phone, or telemonitoring software can help to keep the patient in touch with the healthcare provider. The provider can then make decisions about the patient's

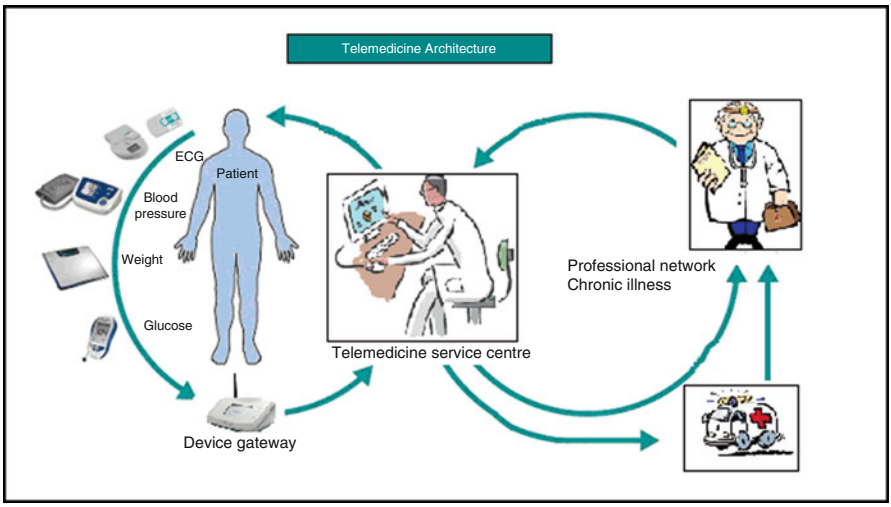


Fig. 1.3 The typical philosophy of remote monitoring

treatment based on a combination of subjective and objective information similar to what would be shown during an on-site appointment. Several advantages of the use of telehealthcare in CHF have been reported (Marschollek 2012). Patients get a better understanding of their disease by being responsible for monitoring and being involved in some self-managing, which enhances patient education and empowerment (Spinsante et al. 2012). This supports adherence to the planned treatment, but also ensures preventive care with early detection of disease exacerbations and timely management (Khan et al. 2006; Varma et al. 2005). Travel and waiting time is reduced or even abolished for the patient, so quality of life improves. Additionally, patients are at home in comfortable and safe surroundings and still in close contact with the hospital. An advantage for the hospital is that the technology leaves room for new and unstable patients. The results provide evidence that remote monitoring combined with discharge planning could reduce family caregiver burden, improve stress mastery and family function during the first 30 days at home after CHF patients are discharged from the hospital.

Other studies have investigated the management of CHF patients using data from implantable devices (Varma et al. 2010; Yu et al. 2005). The efficacy of an implantable monitor lies in the fact that it continuously measures and stores hemodynamic information of CHF patients that can be reviewed remotely. An increase in pulmonary vascular congestion is reflected by increased impedance, thus intra-thoracic impedance monitoring may be a valuable tool in the management of CHF in patients, such as those with an Implantable Cardioverter defibrillator (ICD) (Catanzariti et al. 2009; Small et al. 2009). This is supported by a study that found significantly greater sensitivity to predict worsening heart failure events for intra-thoracic impedance monitoring compared to acute weight increase (Conraads et al. 2011). Remote monitoring and remote follow-up by the use of

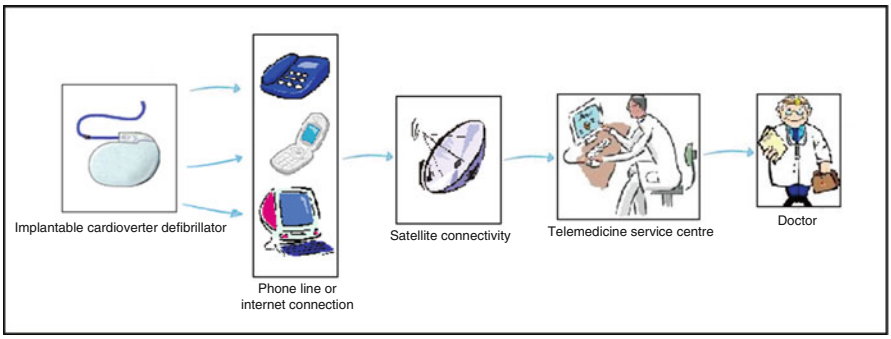


Fig. 1.4 Data transmission through remote monitoring systems: the data is transmitted through the device via telephone or mobile phone or computer to the central server and then to the hospital telemedicine center via telephone or internet line

telehealthcare offers an alternative management strategy (Poole et al. 2008). Preferably, the system should be able to automatically transmit data stored in the device to the outpatient clinic using the wireless global system for mobile communication network or a landline. Automatic wireless data transmission requires a pacemaker or ICD equipped with a micro-antenna for communication with a transmitter located close to the patient as shown in Fig. 1.4. This setup diminishes the need for patient compliance and increases the frequency of transmissions. Several companies provide systems for automatic wireless data transmission. Data stored in the device's memory (e.g., battery voltage, lead characteristics, arrhythmias, alerts) can be transmitted. The alerts may be based on a change in device performance (battery status, lead impedance), programming (disabling of ventricular fibrillation therapy, insufficient safety margins for sensing or capture) or medical data (arrhythmias, indication of lung fluid accumulation). Data are usually transmitted to a central database, where they are processed and made available to the physician on a secure webpage. Additional notification by e-mail, SMS, fax or phone messages may be valuable when critical data are available for consultation. Many pacemakers and ICDs can automatically run tests like battery status, lead impedances, sensing and capture thresholds that used to be manually performed at the outpatient clinic. Both scheduled follow-up, with transmission of pre-defined data, and unscheduled monitoring can then be performed. Daily monitoring allows transmission of data with any predefined alerts to a physician. Many studies demonstrate that a robust telemedicine program not only reduces costs and increases revenue but also saves lives and prevents complications improving patient safety and quality of care (Inglis 2010; Lazarus 2007; Sack et al. 2011; Whellan et al. 2010).

Management of COPD: Patients with COPD experience benefits from telehomecare and remote monitoring. In the past two decades, despite medical advances and gradual decline in mortality with other chronic diseases (e.g., cardiovascular diseases and stroke), mortality from COPD has been rising exponentially. Exacerbations—acute events of worsening of symptoms—are important events in

the progression of COPD, such that monitoring patients in a home setting to detect exacerbation onset may be warranted. COPD exacerbations can lead to severe clinical consequences for the patients and generate high expenditures for the health care system. More precisely, frequent exacerbations are associated with impaired quality of life and a rapid decline in lung function (Seemungal et al. 1998). COPD can be successfully monitored in the home with a high quality of clinical service and no increase in cost. Generally there are four functions of home telehealth interventions for COPD:

– To monitor vital signs or biological health data (e.g., oxygen saturation);
– To monitor symptoms, medication, or other non-biologic endpoints (e.g., exercise adherence);
– To provide information and/or other support services (such as reminders to exercise or positive reinforcement);
– To establish a communication link between patient and health service provider.

Physiological measurements including oxygen saturation and forced expiratory volume in 1s (FEV1) are collected from the patient with Bluetooth-linked monitoring devices. Data is transmitted via a dedicated broadband link to a call center which in turn provides data access to the healthcare workers. At regular intervals, adjustable in frequency and in time of the day, the patient gets an automatic reminder for data entry from the phone. The patient is presented with a simple touch-interface to answer a set of questions about COPD symptoms and is subsequently asked to perform a spirometry test and pulseoximeter measurement. The results of the measurements are transmitted to the phone, and entered in a Bayesian network model to obtain the probability of exacerbation. In addition, the data is synchronized with the call center, which allows the responsible healthcare workers to examine the patient data; depending on the situation this may be a nurse specialized in lung diseases, GP or pulmonologist. If necessary, the patient can be advised to take action. Patients are provided with an action plan and an emergency supply of antibiotics and steroids which they are encouraged to start as soon as an exacerbation was recognized.

Published research suggests that such remote care has the capacity to reduce exacerbations and may improve the quality of life of the people using it in comparison to usual care. Emergency department visits are significantly reduced, as are hospital admissions in COPD patients, without clearly increasing morbidity or incurring excessive costs. Recent studies showed that patients with remote monitoring are much less likely to attend the emergency department than patients in the control group: OR 0.27 (95 % CI 0.11–0.66) (McLean et al. 2011). When patients with severe respiratory illness requiring long-term oxygen therapy were remotely monitored, hospital admissions decreased by 50 %, acute home exacerbations decreased by 55 % and hospitalization costs were reduced by 17 %, even after the costs of monitoring were included (Hebert et al. 2006). Monitoring of this kind can enable more sophisticated home care, detect deterioration prior to symptom development and minimize the need for complicated and cumbersome patient transportation to hospital or office appointments.

Management of Hypertension: The prevalence of hypertension (e.g., in European population, estimated approximately 20 %) is a serious concern in the healthcare service worldwide as it is one of the major risk factors for developing cardiovascular diseases such as CHF, stroke, coronary heart disease and renal failure, and it is one of the most frequent reasons for access to medical care. Inadequate blood pressure control and poor adherence to treatment remain among the major limitations in the management of hypertensive patients, particularly of those at high risk of cardiovascular events. No more than 30 % of treated hypertensive patients maintain a satisfactory blood pressure control. Compliance to therapy seems to be related to many factors including the high number of pills to be taken daily, the lack of motivation and poor patient's involvement in the management of this clinical condition, the absence of symptoms, the inability of the patient to understand the real long-term consequences of high blood pressure, and finally the psychological and personality traits of the patients (Patel and Taylor 2002). Home blood pressure monitoring is a well-established practice that has been shown to improve patient adherence to treatment regimens and to achieve target blood pressure levels. Many studies have been conducted on the effectiveness of telemedicine in managing hypertension. Another interesting feature about telemedicine is its ability to provide measurements of blood pressure without provoking the "white coat" effect. Accurate home blood pressure measurement improves blood pressure control, thereby reducing the risk of cardiovascular events. It may allow for more timely alterations in medical regimen and thus more rapid achievement of target blood pressure levels. It also can reduce the total cost of hypertension care compared with usual care (Parati et al. 2013).

Management of Diabetes and renal diseases: Diabetes is another chronic disease that is increasing alarmingly throughout the world and owing to its burden on patients and healthcare systems, it becomes necessary to identify patient management approaches that would ensure appropriate monitoring and treatment of diabetes while reducing the cost involved in the process. Evidence has shown that intensive diabetes treatment with frequent daily glucose monitoring of the blood may reduce the evolution of complications associated with the disease (Control and Group 1993) and a remote monitoring system could be an effective tool for that. A remote monitoring system for managing diabetic patients with better care relies on the integration of front end (patient unit) and backend webserver. The system is capable of reading glucose measurements from any glucose meter and sending them automatically via existing networks to the backend server (Fig. 1.5). This system helps the healthcare provider with data analysis: data visualization and decision support; furthermore allows them to send feedback and therapeutic advice to patients from anywhere using a browser enabled device. This telehealthcare system improves the management of glucose level of insulin dependent diabetic patients (Paré et al. 2010). However, an estimated 55.6 % of adults with diabetes do not meet the American Diabetes Association target for glycemic control (hemoglobin A1c (HbA1c) < 7.0 %). Factors that may contribute to suboptimal glycemic control include inadequate home glucose monitoring, non-compliance with medications or lifestyle changes, suboptimal patient education about the disease,

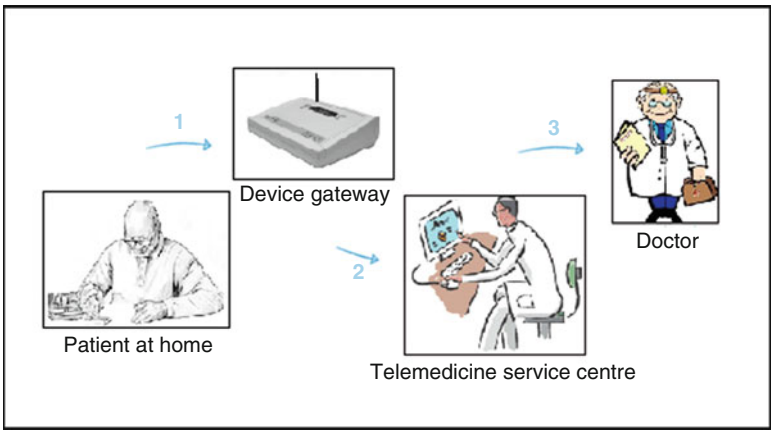


Fig. 1.5 Remote management system for diabetes

and limited access to providers for diabetes management. A significant decrease in HbA1c, an indicator of how well diabetes is being managed, was observed in several studies as a result of remote monitoring.

The increasing prevalence of diabetes has led to Diabetic Kidney Disease (DKD) becoming the leading cause of end stage of renal disease in many regions. The economic cost of DKD will grow to prohibitive amounts unless strategies to prevent its onset or progression are urgently implemented. Computer-assisted medical activity is increasing in several fields, with wide perspectives in nephrology and dialysis accounting for the peculiar characteristics of this population such as number, complexity, follow-up length and economic cost. Informatics procedures seem to be quite advisable in improving as clinical surveillance of the patients, as technical and managerial aspects of dialysis units. Home dialysis can improve the care and quality of life for patients with renal failure. Patients became comfortable with interactive healthcare delivery quickly, and preferred it as a means of receiving primary care as well as continuous physician monitoring during dialysis treatments. The technology provides access for families that might not be able to travel and it also allows them to monitor rural patients more frequently (Mitchell and Disney 1997).

1.5 Expectations of Clinicians from the Next-Generation Remote Healthcare Systems

Despite the existence of fragmented evidences of potential of remote healthcare in optimizing the care delivery model, it is in general well understood that the current form of remote healthcare systems need to make significant enhancements to become universally applicable. This leads to the development of the next-generation remote healthcare system that in effect should provide a platform

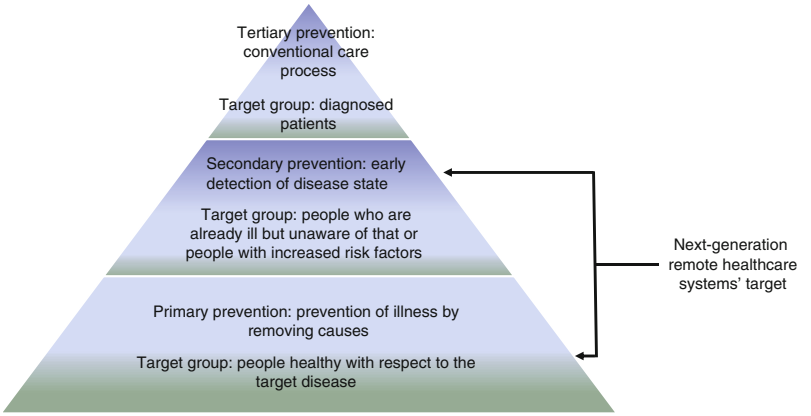


Fig. 1.6 The different levels of disease prevention and the most sought target levels of next-generation remote healthcare

enabling multiple functionalities ranging from developing integrated diagnostics and management model to knowledge harvesting. To reach this goal, a multidisciplinary approach amalgamating practical clinical knowledge about a disease, social requirements and technological advances is essential. Although the possibilities are endless, it is more important to understand how the clinicians expect to use such a system so as to avoid possible over-engineering—a typical case for any multidisciplinary design stemming out from the lack of understanding of inter-domain *modus operandi*.

1.5.1 Population Screening for Disease Prevention

Prevention has been projected as the best possible way to combat the escalation of disease burden in the society. As shown in Fig. 1.6, it can be categorized in three classes: primary, secondary and tertiary prevention and the highest level of impacts in terms of wellbeing and cost-effectiveness can be achieved at the primary and secondary prevention levels since it involves the majority of the population.

Primary prevention is more related to social awareness about the possible risk factors embedded in the lifestyle of a person. There are plethora of clinical studies linking the lifestyle (food habit, patterns of physical activities culture, environment etc.) factors with possible cause of disease. Although in the era of booming digital economy a significant portion of such information is available publicly through digital media, sometimes from the clinical experience it seems that the information is "too heavy" for a non-clinical person as well as is not presented in proper structurally integrated manner leaving room for false interpretation of the available information. Therefore, it is of utmost importance that the next-generation

healthcare system develops a strategy of making the information in a structured manner complying with the evidence-based practical clinical foundations.

The situation with secondary prevention is more intensive than the primary prevention. As the target group is either already ill or elevated risk-factors population, it needs to combine the strategy of dissemination of social awareness with technology-enabled monitoring. The main target here is to manage the population in clinical sense so that appropriate intervention could be initiated at proper time point before escalation of the disease. This is closer to the conventional telemonitoring system although needs to be enhanced significantly to acquire more descriptive heterogeneous physiological and lifestyle parameters in automated way and forming a clinical logic based on that information for ascertaining the overall health condition of a person. This calls for development of specialized sensors for multiple physiological parameter acquisition, advanced data processing and decision making techniques which could be embedded within the overall infrastructure of the next-generation remote healthcare.

The phase of tertiary prevention is more related to the conventional clinical care pathway where the diagnosed patients need to go through rigorous clinical follow-up process or even hospital admission. The main requirement here is to develop more efficient diagnostic modalities and patient information integration infrastructure from multiple heterogeneous sources.

Combining all the above mentioned features the next-generation remote healthcare could be an effective tool for reducing the disease burden in a society-wide manner—a much needed solution for alleviating the tremendous pressure the national health services are under.

1.5.2 Integrated Diagnostics Screening

A pre-requisite for effective prevention and management is to understand which patient in a high-risk group defined by conventional risk-factors requires immediate clinical attention. In the traditional reactive care model patients visit the clinicians only after the symptoms of disease start to manifest, therefore giving the clinicians possibly less time to intervene. An obvious result in most cases is hospitalization of even death. A major expectation from the next-generation remote healthcare system is to formulate the ability to identify those patients who may need clinical attention in short-term even before the symptoms are manifested partially/fully and thereby allowing the clinicians to move towards a proactive care pathway. This is more applicable and desired approach in the secondary prevention level although a cut-down version of it could be beneficial in the primary prevention level as well. The fundamental philosophy is shown in Fig. 1.7 for the case of cardiovascular disease monitoring as an example. The central to this approach is a decision-logic module that computes in an automated way the potential risk of an impending episode by information fusion of heterogeneous sensor, demographic data and evidence-based clinical diagnosis process. The main point here is that the

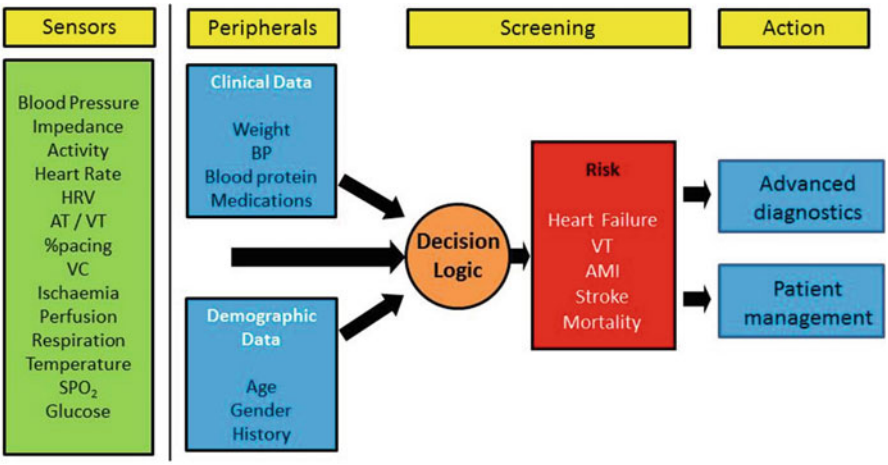


Fig. 1.7 The methodology for integrated diagnostic screening

decision-logic is not an automated diagnostic tool in the pure clinical sense, rather than that it only computes the "probability" of a possible episode based on available clinical knowledge. However, from the practical application point of view, it is more than sufficient that when the decision-logic module computes a high probability of an impending episode, it alarms the immediate caregiver warning him/her about the "potential" danger. In such a case the patient could be attended immediately and further rigorous mainstream diagnostic procedures (like radiography, echocardiography or high-resolution imaging etc.) could be initiated to ascertain the actual clinical condition of the patient. On the other hand, the patients not prone to develop impending episodes in short-term but belongs to the high-risk group, can also be monitored regularly using some quantitative metrics computed by the decision-logic module describing their regular health status in an automated way. This proactive care pathway is likely to give more reaction time in clinical sense which possibly will lead to reduction of hospitalization and eventual death.

1.5.3 Continuous Monitoring

Analysis of typical causes for hospitalization, in particular with patients suffering from long-term conditions, shows patient non-compliance and inadequate treatment are the major factors (Michalsen et al. 1998) in addition to the lack of proper patient awareness (Michalsen et al. 1998). Introduction of Heart Failure Specialist Nurse (HFSN) in the UK NHS system operating through structured telephone services and home visits has been reported to be significantly effective in mitigating these factors leading to reduced hospitalization and enhancing quality of life

(Pattenden et al. 2008). However, the steeply increasing prevalence rate of chronic diseases and the long-term conditions resulting from them, will raise the management issues (Pattenden et al. 2008) (e.g., caseload/nurse, and distance to be travelled by the nurse in a day for home visit—some of the nurses travel 40–70 miles a day) and cost and therefore such service may well become unsustainable in the future. Next-generation remote healthcare system should play an instrumental role in this respect. Automated patient monitoring based on objective evaluation criteria and semantically structured patient information preservation coupled with automated delivery of awareness program (both general and on-demand basis) anytime anywhere implies reduction of nurses' workload while reaching to the wider population with minimal time delay. Healthcare resources freed-up in this way (e.g., nurses' time) may be optimally allocated according to the actual need. On the other hand, the continuous monitoring through automated procedures ensures that the patient is in right health condition and actions could be taken immediately once some complication is detected.

1.5.4 Knowledge Harvesting

This is probably the most important scientific expectation from the next-generation remote healthcare system. It is a well-acknowledged fact that despite the fantastic achievements of clinical sciences, it is essentially limited, in several cases, by unavailability of appropriate data. Had such data be available, the probability of any clinical decision-errors could be reduced significantly. As an example, continuous trend calculation of physiological parameters in contextualized way has been deemed to be more effective in ascertaining a patient's actual clinical condition rather than trying to understand the same based on some discrete measurements—typical in the conventional healthcare system. Unfortunately, until now there was no infrastructure that could enable such continuous data acquisition over long time. The next-generation remote healthcare system opens up this possibility and thereby enabling an effective research environment for enhancing diagnostic quality through discovery of new biomarkers by exploring new types of data. A particularly important part is the patient-centric approach. Human body is a super complex organization and the inter-person variability in terms of physiological parameters could be significant. Such variability, in theory, always has finite potential of leading to decision mistakes. The next-generation remote healthcare infrastructure allows clinicians to study the characteristic trends of physiological parameters in a person-centric way which possibly will add significant knowledge in perfecting the diagnostic and treatment procedure. It's worth noting that the decision-logic described in Sect. 1.5.2 is implicitly related to this knowledge harvesting process. It is quite conceivable that the decision-logic will become more and more sophisticated and accurate as the generated knowledge is iteratively incorporated in its functionality making the system more and more reliable in the true clinical sense.

1.5.5 Provide Information, Don't Be an Artificial Doctor

While it is conceivable that the films like Star Wars allow us to think about super sophisticated systems in various fields where everything could be done by machines, the healthcare domain is unfortunately not something like that—at least not now. Healthcare plays with lives of human beings and nothing can be compromised in the quality of care delivery pathway. Doctors use their hard-earned experience to make a decision about the health condition of a patient. It is almost impossible to incorporate that "experience" in totality within a machine. Therefore there is finite probability of error in a machine's decision—however sophisticated the underlying algorithm is. Additionally, the sensors capturing the physiological signals have their inherent inaccuracies. More importantly, it may happen that the physiological signals captured may not be adequate enough to represent the disease phenomenon. Therefore it is more pragmatic to use this as a screening tool providing holistic picture of the patient and a potential complication pathway to the clinicians. The clinicians need to take the final decision, either based on the data provided or by taking it as a hint and applying sophisticated more directed diagnostic measures. In that sense the next-generation remote healthcare system should act as a virtual nurse rather than a virtual clinician.

The bottom-line is the clinicians expect the next-generation healthcare system should provide them wealth of information which can be used for developing extremely sophisticated diagnostic and management method while reducing the workload on all the stakeholders of a healthcare system.

1.6 Conclusion

By adapting to the population scale the potential of technology and intervening at the level of primary and secondary prevention, the next-generation remote healthcare could be an effective tool for reducing the disease burden in a society-wide manner and this may certainly help alleviating the tremendous pressure (and costs) the national health services are under. There is a need for the next-generation remote healthcare system to identify those individuals (in primary prevention) or patients (in secondary or tertiary care) who may need clinical attention in short-term even before the symptoms are manifested partially/fully and thereby allowing the clinicians to move towards a proactive care pathway. Accordingly, a revolution should occur and predictive/preventive instruments will come closer and closer to the events (weeks and days) which will be more and more foreseen in short or very-short-term (Puddu et al. 2012), as opposed to the classical terms (Klersy et al. 2009) whereby risks are evaluated at medium/long term (months and years ahead). Remote health system will provide the parameters to evaluate trends and short-term predictive indexes and will pave the road for this revolution in Medicine, based on decision-logic module that computes in an automated way the potential risk of an

impending episode by information fusion of heterogeneous sensor, demographic data and evidence-based clinical diagnosis process the CHIRON observational study is a good exemple of this phylosophy (see later in this book) and the results may well contribute to evaluate the feasability and potential usefulness of the new approach (Puddu et al. 2012). Clinicians will study the characteristic trends of physiological parameters in a person-centric way which possibly will add significant knowledge in perfecting the diagnostic and treatment. All these systems will however act as a virtual nurse rather than a virtual clinician.

References

European Union: European Union Action Plan 2012-2020 (2012) Innovative healthcare in the 21st century

Abraham WT, Adamson PB, Bourge RC et al (2011) Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. Lancet 377:658–666

Ahn AC, Tewari M, Poon C-S, Phillips RS (2006a) The limits of reductionism in medicine: could systems biology offer an alternative? PLoS Med 3:e208

Ahn AC, Tewari M, Poon C-S, Phillips RS (2006b) The clinical applications of a systems approach. PLoS Med 3:e209

Black AD, Car J, Pagliari C et al (2011) The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 8:e1000387.

Blobel B (2013) Knowledge representation and management enabling intelligent interoperability-principles and standards. In: Blobel B et al (Eds) Data and knowledge for medical decision support, p 3. http://www.deastore.com/book/data-and-knowledge-for-medical-b-blobel-ios-press/9781614992394.html

Bloland P, Simone P, Burkholder B et al (2012) The role of Public Health Institutions in global health system strengthening efforts: the US CDC's perspective. PLoS Med 9:e1001199

Bonfiglio S (2012) Fostering a continuum of care. Impact analysis of solutions for chronic disease prevention and management. Springer, Heidelberg, pp 35–41

Bradley EH, Herrin J, Wang Y et al (2006) Strategies for reducing the door-to-balloon time in acute myocardial infarction. N Engl J Med 355:2308–2320

Catanzariti D, Lunati M, Landolina M et al (2009) Monitoring intrathoracic impedance with an implantable defibrillator reduces hospitalizations in patients with heart failure. Pacing Clin Electrophysiol 32:363–370

Catwell L, Sheikh A (2009) Evaluating eHealth interventions: the need for continuous systemic evaluation. PLoS Med 6:e1000126

Catwell L, Sheikh A, Catwell L, Sheikh A (2009) Information technology (IT) system users must be allowed to decide on the future direction of major national IT initiatives. But the task of redistributing power equally amongst stakeholders will not be an easy one. Inform Prim Care 17:1

Chaudhry SI, Phillips CO, Stewart SS et al (2007) Telemonitoring for patients with chronic heart failure: a systematic review. J Card Fail 13:56–62

Chaudhry SI, Mattera JA, Curtis JP et al (2010) Telemonitoring in patients with heart failure. N Engl J Med 363:2301–2309

Chen C-C, Tseng C-H, Cheng S-H (2013) Continuity of care, medication adherence, and health care outcomes among patients with newly diagnosed type 2 diabetes: a longitudinal analysis. Med Care 51:231–237

Clark RA, Inglis SC, McAlister FA et al (2007) Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. BMJ 334:942

Clermont G, Auffray C, Moreau Y et al (2009) Bridging the gap between systems biology and medicine. Genome Med 1:88

Conraads VM, Tavazzi L, Santini M et al (2011) Sensitivity and positive predictive value of implantable intrathoracic impedance monitoring as a predictor of heart failure hospitalizations: the SENSE-HF trial. Eur Heart J 32:2266–2273

Control TD, Group CTR (1993) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 329:977–986

Davies AA, Borland RM, Blake C, West HE (2011) The dynamics of health and return migration. PLoS Med 8:e1001046

de la Torre-Díez I, González S, López-Coronado M (2013) EHR systems in the Spanish Public Health National System: the lack of interoperability between primary and specialty care. J Med Syst 37:1–7

Department of Health (2011) Whole System Demonstrator Programme: headline findings. www.dh.gov.uk (search for gateway ref 16972)

Diemberger I, Biffi M, Valzania C et al (2008) Telecardiology and remote monitoring of implanted electrical devices: the potential for fresh clinical care perspectives. J Gen Intern Med 23:73–77

Dy SM, Apostol C, Martinez KA, Aslakson RA (2013) Continuity, coordination, and transitions of care for patients with serious and advanced illness: a systematic review of interventions. J Palliat Med 16:436–445

Eapen ZJ, Reed SD, Curtis LH et al (2011) Do heart failure disease management programs make financial sense under a bundled payment system? Am Heart J 161:916–922

Eisner C, Sommer P, Piorkowski C et al (2006) A prospective multicenter comparison trial of home monitoring against regular follow-up in MADIT II patients: additional visits and cost impact. Comput Cardiol 2006:241–244

Fazio F, Simonetti G, Casolari P et al (2010) Criteri di appropriatezza clinica, tecnologica e strutturale nell'assistenza alle malattie del sistema cardiovascolare. In: Quaderni del Ministero della Salute, n°1, gennaio—febbraio 2010. Ministero della Salute, Roma. http://www.quadernidellasalute.it/quaderni-html/6-novembre-dicembre-2010.php#6

Fedele F, Scalvini S (2009) Modelli di telecardiologia attualmente in uso. G Ital Cardiol 10:31S–33S

Fedele F, Genuini I, De Lazzari C et al (2010) La telecardiologia 1° Edizione. Società Editrice Universo, Roma

Gabbrielli F (2013) Telemedicine R&D influencing incoming strategies and organization models. Telehealth Netw Hosp Serv New Methodol 250–264. http://www.igi-global.com/chapter/telemedicine-influencing-incoming-strategies-organization/74654

Gaddi A, Capello F, Manca M (2013) eHealth, care and quality of life. http://www.springer.com/medicine/book/978-88-470-5252-9

Gallagher N, MacFarlane A, Murphy AW et al (2013) Service users' and caregivers' perspectives on continuity of care in out-of-hours primary care. Qual Health Res 23:407–421

Geissbuhler A (2013) Lessons learned implementing a regional health information exchange in Geneva as a pilot for the Swiss national eHealth strategy. Int J Med Inform 82:E118–E124

Giovanni R, Gaddi A, Capello F (2013) Medical data, information economy and federative networks: the concepts underlying the comprehensive electronic clinical record framework. Nova Science, Hauppauge, NY

Granados D, Lefranc A, Reiter R et al (2005) Disability-adjusted life years: an instrument for defining public health priorities? Revue d'épidémiologie et de santé publique 53:111

Greenhalgh T, Russell J, Ashcroft RE, Parsons W (2011) Why national ehealth programs need dead philosophers: Wittgensteinian reflections on policymakers' reluctance to learn from history. Milbank Q 89:533–563

Hebert MA, Korabek B, Scott RE (2006) Moving research into practice: a decision framework for integrating home telehealth into chronic illness care. Int J Med Inform 75:786–794
Henderson C, Knapp M, Fernández J-L et al (2013) Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. BMJ 346:f2065
Hong J-S, Kang H-C (2013) Continuity of ambulatory care and health outcomes in adult patients with type 2 diabetes in Korea. Health Policy 109:158–165
Hsieh J, Yu K, Yang C-C (2009) The realization of ubiquitous 12-lead ECG diagnosis in emergency telemedicine. Telemed E Health 15:898–906
http://onlinelibrary.wiley.com/doi/10.1111/j.1756-5391.2010.01104.x/full
Inglis SC, Clark RA, Cleland JG et al (2010) Structured telephone support or telemonitoring programs for patients with chronic heart failure. Cochrane Database Syst Rev 8:CD007228
Jencks SF, Huff ED, Cuerdon T (2003) Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. JAMA 289:305–312
Kearns JW, Bowerman D, Kemmis K et al (2012) Group diabetes education administered through telemedicine: tools used and lessons learned. Telemed E Health 18:347–353
Khan M, Kozelhova M, Patel D (2006) Hospitalization due to atrial fibrillation in patients with cardiac resynchronization therapy. Heart Rhythm 3:P4–P104
Klersy C, De Silvestri A, Gabutti G et al (2009) A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 54:1683–1694
Koehler F, Winkler S, Schieber M et al (2011) Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure clinical perspective: the telemedical interventional monitoring in heart failure study. Circulation 123:1873–1880
Kuhn TS (1996) The structure of scientific revolutions. University of Chicago Press, Chicago
Lazarus A (2007) Remote, wireless, ambulatory monitoring of implantable pacemakers, cardioverter defibrillators, and cardiac resynchronization therapy systems: analysis of a world-wide database. Pacing Clin Electrophysiol 30:S2–S12
Marschollek M (2012) Decision support at home (DS@ HOME)-system architectures and requirements. BMC Med Inform Decis Mak 12:43
McBride A, Burey L, Megahed M et al (2013) The role of patient-held alert cards in promoting continuity of care for heart failure patients. Eur J Cardiovasc Nurs. doi: 10.1177/1474515113478531
McLean S, Nurmatov U, Liu J et al (2011) Telehealthcare for chronic obstructive pulmonary disease. Cochrane Database Syst Rev (7):CD007718
McMillan A, Trompeter J, Havrda D, Fox J (2013) Continuity of care between family practice physicians and hospitalist services. J Healthc Qual 35:41–49
McNamara RL, Herrin J, Bradley EH et al (2006) Hospital improvement in time to reperfusion in patients with acute myocardial infarction, 1999 to 2002. J Am Coll Cardiol 47:45–51
Michalsen A, Knig G, Thimme W (1998) Preventable causative factors leading to hospital admission with decompensated heart failure. Heart 80:437–441
Mitchell J, Disney A (1997) Clinical applications of renal telemedicine. J Telemed Telecare 3:158–162
Mortimer D, Segal L (2008) Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. Med Decis Making 28:66–89
Parati G, Omboni S, Compare A et al (2013) Blood pressure control and treatment adherence in hypertensive patients with metabolic syndrome: protocol of a randomized controlled study based on home blood pressure telemonitoring vs. conventional management and assessment of psychological determinants of adherence (TELEBPMET Study). Trials 14:22
Paré G, Moqadem K, Pineau G, St-Hilaire C (2010) Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 12:e21

Patel RP, Taylor SD (2002) Factors affecting medication adherence in hypertensive patients. Ann Pharmacother 36:40–45

Pattenden J, Coulton S, Spilsbury K et al (2008) The development and impact of the British Heart Foundation and Big Lottery Fund heart failure specialist nurse services in England: final report April 2008. The University of York, North Yorkshire

Poole JE, Johnson GW, Hellkamp AS et al (2008) Prognostic importance of defibrillator shocks in patients with heart failure. N Engl J Med 359:1009–1017

Prieto L, Sacristán JA et al (2003) Problems and solutions in calculating quality-adjusted life years (QALYs). Health Qual Life Outcomes 1:80

Puddu PE, Morgan JM, Torromeo C et al (2012) A clinical observational study in the CHIRON project: rationale and expected results. Impact analysis of solutions for chronic disease prevention and management. Springer, Heildelberg, pp 74–82

Riezebos RK (2011) ACP Journal Club. Telemonitoring did not reduce readmissions or mortality in patients recently hospitalized for heart failure. Ann Intern Med 154:JC3–JC8

Robertson A, Cresswell K, Takian A et al (2010) Implementation and adoption of nationwide electronic health records in secondary care in England: qualitative analysis of interim results from a prospective national evaluation. BMJ 341:c4564

Sable C (2001) Telecardiology: potential impact on acute care. Crit Care Med 29:N159–N165

Sack S, Wende CM, Nägele H et al (2011) Potential value of automated daily screening of cardiac resynchronization therapy defibrillator diagnostics for prediction of major cardiovascular events: results from Home-CARE (Home Monitoring in Cardiac Resynchronization Therapy) study. Eur J Heart Fail 13:1019–1027

Sackett DL (1969) Clinical epidemiology. Am J Epidemiol 89:125–128

Sackett DL (2002) Clinical epidemiology: what, who, and whither. J Clin Epidemiol 55:1161–1166

Sackett DL, Haynes RB, Tugwell P et al (1985) Clinical epidemiology: a basic science for clinical medicine. Little, Brown and Company, Boston

Seemungal TA, Donaldson GC, Paul EA et al (1998) Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 157:1418–1422

Sheikh A, Cornford T, Barber N et al (2011) Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from prospective national evaluation in "early adopter" hospitals. BMJ 343:888

Siddiqui M, Pollack CE (2013) Continuity in a changing health care environment. J Gen Intern Med 28:493–494

Small RS, Wickemeyer W, Germany R et al (2009) Changes in intrathoracic impedance are associated with subsequent risk of hospitalizations for acute decompensated heart failure: clinical utility of implanted device monitoring without a patient alert. J Cardiac Fail 15:475–481

Spinsante S, Antonicelli R, Mazzanti I, Gambi E (2012) Technological approaches to remote monitoring of elderly people in cardiology: a usability perspective. Int J Telemed Appl 2012:3

Steventon A, Bardsley M, Billings J et al (2012) Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ 344:e3874

Steventon A, Bardsley M, Billings J et al (2013) Effect of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial. Age Ageing 42:501–508

Swedberg K, Wolf A, Ekman I (2011) Telemonitoring in patients with heart failure. N Engl J Med 364:1078

Tompkins C, Orwat J (2009) A randomized trial of telemonitoring heart failure patients. J Healthc Manag 55:312–322

Topol EJ (2012) The creative destruction of medicine. Basic Books, New York

Varma N, Stambler B, Chun S (2005) Detection of atrial fibrillation by implanted devices with wireless data transmission capability. Pacing Clin Electrophysiol 28:S133–S136

Varma N, Epstein AE, Irimpen A et al (2010) Efficacy and safety of automatic remote monitoring for implantable cardioverter-defibrillator follow-up the Lumos-T Safely Reduces Routine Office Device Follow-Up (TRUST) trial. Circulation 122:325–332

Wade MJ, Desai AS, Spettell CM et al (2011) Telemonitoring with case management for seniors with heart failure. Am J Manag Care 17:71–79

Wang H-Q, Li J-S, Zhang Y-F et al (2013) Creating personalised clinical pathways by semantic interoperability with electronic health records. Artif Intell Med 58(2):81–89

Whellan DJ, Ousdigian KT, Al-Khatib SM et al (2010) Combined heart failure device diagnostics identify patients at higher risk of subsequent heart failure hospitalizations: results from PARTNERS HF (Program to Access and Review Trending Information and Evaluate Correlation to Symptoms in Patients With Heart Failure) study. J Am Coll Cardiol 55:1803–1810

White J et al (2013) Rethinking health IT standards. Expected savings will follow with the right framework for interoperability. Mod Healthc 43:26

Yu C-M, Wang L, Chau E et al (2005) Intrathoracic impedance monitoring in patients with heart failure correlation with fluid status and feasibility of early warning preceding hospitalization. Circulation 112:841–848

Chapter 2
System Overview of Next-Generation Remote Healthcare

Andrea Vitaletti and Stefano Puglia

Abstract Pervasive communication technologies and tiny wireless monitoring devices have provided a solid foundation for the idea of remote healthcare. This idea aims at making available over communication networks (e.g. the Internet) some of the fundamental healthcare services commonly available only in conventional medical facilities and clinical settings. Information and communication systems supporting ubiquitous and continuous remote healthcare provision have been effectively employed in a number of application scenarios. In this chapter we mainly focus on patient monitoring activities as they are general enough to demonstrate most of the technical and procedural challenges that need to be addressed in remote healthcare systems. The main contributions of the chapter are a detailed definition of necessary functional and non-functional requirements for such systems and the identification, design and discussion of a reference architecture conceived to properly meet these requirements.

2.1 Introduction

The availability of pervasive and cost effective communication technologies and the miniaturization of wireless monitoring devices have made it possible to conceive and develop the idea of remote healthcare.

For the purpose of this book, remote healthcare encompasses a wide range of technologies to make available over the Internet (or possibly a dedicated

A. Vitaletti (✉)
DIAG, "SAPIENZA" Università di Roma, via Ariosto 25, 00185 Rome, Italy

WLAB S.r.L., via Adolfo Ravà 124, 00142 Rome, Italy
e-mail: andrea.vitaletti@w-lab.it; vitaletti@dis.uniroma1.it

S. Puglia
WLAB S.r.L., via Adolfo Ravà 124, 00142 Rome, Italy
e-mail: stefano.puglia@w-lab.it

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_2,

communication network) some of the healthcare services, such as clinical care, diagnostics and patient monitoring commonly provided only in conventional medical facilities and clinical settings like public hospitals, private clinics and doctor's surgeries.

For the sake of simplicity, in this chapter we mainly focus on patient monitoring activities as they are general enough to demonstrate most of the technical and procedural challenges that need to be addressed in a remote healthcare system.

Remote healthcare systems have been effectively employed in a number of application scenarios, such as monitoring of heart failure (HF) patients (Martínez et al. 2006), data collection from glucose meters in the treatment of patients with diabetes Mellitus (Shea et al. 2006), monitoring and education of children with asthma (Chan et al. 2007) and recording movement activity of elderly people in their flats (Cerny 2010).

Interestingly, the same technologies that enable remote healthcare (i.e. healthcare everywhere) also support the continuum of care (i.e. healthcare any time), in which the personal healthcare information and clinical data of an individual are linked together to provide a coherent and high quality picture of his/her history as a patient.

The rest of the chapter is organized as follows: in Sect. 2.2 we briefly discuss the functional and non-functional requirements for remote healthcare, in Sect. 2.3 we present a high-level reference architecture for remote healthcare systems and in final Sect. 2.4 we give further details on the architecture's components.

2.2 Functional and Non-functional Requirements of Remote Healthcare Systems

In system engineering, functional requirements define the specific behaviours of a system in terms of a set of functions that must be implemented by the system or its components. A function is described as a precise set of inputs, behaviours and outputs. The definition of non-functional requirements is more controversial (Glinz 2007). We adopt here the following definition: a non-functional requirement specifies the constraints that have to be considered in the design and implementation of a functional requirement in order to achieve the expected quality.

A good set of requirements should be (IEEE Guide to Software Requirements Specifications 1984):

- *Unambiguous*. It uses a single unique term for each characteristic;
- *Complete*. It includes all significant requirements and responses to all inputs in all conditions and situations;
- *Verifiable*. Every requirement should be verifiable by some cost-effective checking process;
- *Consistent*. Individual requirements are not in conflict;

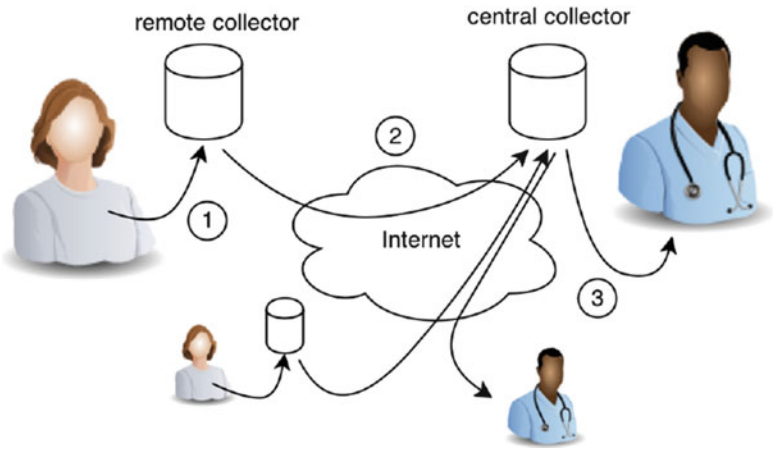


Fig. 2.1 A high-level overview of a remote healthcare system. Physiological and context data are collected in remote sites (1), transmitted to a central medical repository (2) and possibly accessed from everywhere by physicians (3)

- *Modifiable*. Any necessary change can be made easily while preserving completeness and consistency;
- *Traceable*. The origin of each requirement is clear and any modification can be easily tracked.

2.2.1 A Minimal Set of Functional Requirements

According to the definition of remote healthcare given in the introduction and illustrated in Fig. 2.1, a slightest set of functional requirements for a generic healthcare system can be defined as follows:

R1. A remote healthcare system must collect physiological data of a patient at his/her home;
R2. A remote healthcare system shall be able to collect physiological data of a patient also when he/she is not at home;
R3. The collection of physiological data as per [R1.] must be context-aware (i.e. any information is marked with a patient's "situation" such as location, carried out tasks, specific environment's conditions);
R3.1. A remote healthcare system must collect context data related to the physiological data;
R3.2. A remote healthcare system must annotate the acquired physiological and context data;

Table 2.1 Data rate, bandwidth and accuracy for some physiological and context sensors (Latré et al. 2011)

Application	Data rate	Bandwidth	Accuracy
Temperature	120 bps	0–1 Hz	8 bits
Blood saturation	16 bps	0–1 Hz	8 bits
Glucose monitoring	1600 bps	0–50 Hz	16 bits
EMG	320 kbps	0–10,000 Hz	16 bits
ECG (12 leads)	288 kbps	100–1000 Hz	12 bits
ECG (6 leads)	71 kbps	100–500 Hz	12 bits
Motion sensor	35 kbps	0–500 Hz	12 bits

R4. A remote healthcare system must deliver physiological and context data to a central medical repository;

R5. Healthcare professionals must be able to access their patients' data on the central medical repository from everywhere and at any time.

This is indeed a very minimal set of functional requirements, but we believe that too many details here would be beyond the scope of this chapter and, by definition of functional requirements, might not rely on general principles but on a very "application driven" basis. Instead, we deem it interesting to give some hints on the specific functions underlying the above mentioned requirements in order to better highlight the needed implementation steps. Focusing our attention on [R1.], the function to be implemented is the *collection of physiological data*. To improve the quality of this requirement, we should provide answers to questions like: What is the purpose of the monitoring activity? What kind of physiologic data are monitored? What are the relevant events to be monitored? How long is the monitoring activity expected to last?

Obviously, in order to avoid ambiguities and imprecisions produced by natural language, a more formal approach to describe functional requirements is advocated. What we mean by this is that a clearer descriptions of inputs, behaviours and outputs must be taken into account wherever and whenever possible. For instance, a typical ECG tracing of the cardiac cycle (heartbeat) consists of a P wave, a QRS complex, a T wave, and a U wave. Let us assume that the requirement [R1.] has been detailed in a set of sub-requirements, one of which [R1.n.] is: "the remote healthcare system must be able to extract the T waves[1] from the ECG tracing". In this case, the function to be implemented is the extraction of the T waves. The input is a 6 leads ECG tracing (e.g., 6 lists of 12 bits integers, see Table 2.1) and the output are the T waves (e.g., markers on the file in input).

We stress again that physiological data, such as ECG traces, must be enhanced with context data consisting of a subject's activity, psychological status and environmental conditions related to situations where physiological data is captured. As an example, the heart rate of a patient (i.e. physiological data) clearly depends on his/her ongoing activity (i.e., context data): a heart rate of 140 can be alarming when an individual is at rest, while it is fairly normal when the individual is running.

[1] The T wave represents the re-polarization of the ventricles.

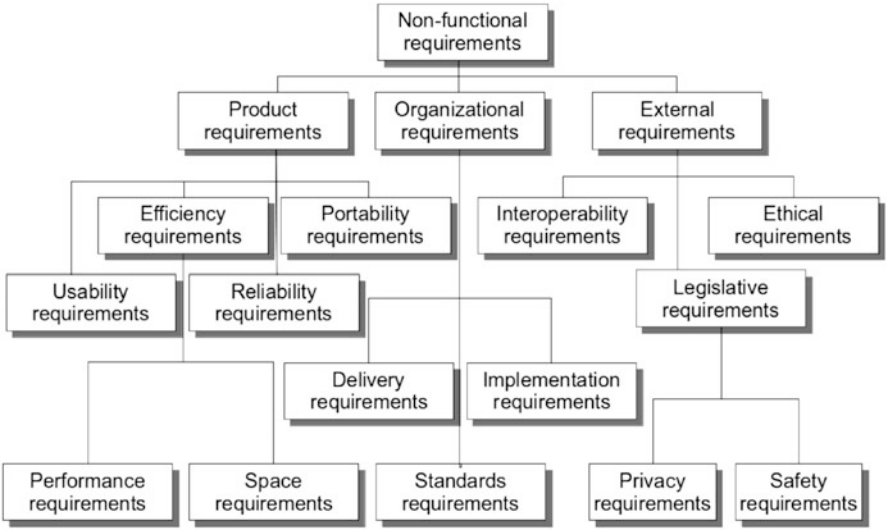


Fig. 2.2 A possible taxonomy of non-functional requirements

2.2.2 Non-functional Requirements

As observed before, the definition of non-functional requirements is more controversial. In Fig. 2.2 a possible taxonomy of non-functional requirements is provided. In the following, we briefly present some of these requirements and we discuss what the implications of such requirements are when applied to a remote healthcare system.

Usability A system is user-friendly if it is considered easy to use by its end-users. Observe that the usability of a system depends by definition on the expertise of the involved users. In the context of remote healthcare systems, the two main users, namely physicians and patients, are not expected to be particularly expert in the use of pervasive technologies. As a consequence, special care must be devoted to design a system's user experience specifically tailored to the needs of these two categories of users. In particular, from the patient's perspective, the sensors used to monitor the clinical and context parameters should be comfortable, ergonomic and "zero-configuration" (i.e. the user should not perform any action to make them work properly). In principle, sensors should not impose any constraint on the daily activity of the patients. From the doctor's perspective, on the other hand, the analysis of the collected patient data should be made simple and effective, offering both an intuitive and prompt access to a summary of relevant information and the possibility to further and deeper investigate the usually large amount of available underlying data.

Reliability Informally, a system or a component is reliable if it is able to perform the required functions under stated conditions for a specified period of time.

The reliability of a system is commonly evaluated by using a statistical approach, namely the probability that a system works as expected for a given time interval. This definition implicitly assumes that the customer usage, the operating environment conditions and what is meant by system "failure" have already been defined. Common metrics employed to evaluate the reliability of a system are: Mean-Time-To-Failure and Mean-Time-Between-Failures (Krasich 2009), Failure rate, BX life (the time by which X% of the units in a population will have failed).

In the context of remote healthcare systems, the reliability of physiological sensors is a critical factor. It is important to note that, while false positives (i.e., report of an issue when there are no issues to be reported) usually result in unexpected costs for the unnecessary actions, false negatives (i.e. no reported issues when there are actually some) can have serious consequences and impacts on the health of monitored patients and must therefore be avoided. Since most of the wireless sensors employed in a remote healthcare system are battery powered, when their energy levels fall below a certain threshold their behaviour can become unreliable.

Portability This requirement is mainly concerned with the system software components. A software is portable if it can run on heterogeneous environments (i.e. hardware platforms and/or operating systems). A remote healthcare system is a distributed system connecting heterogeneous environments, such as home sensor networks used to monitor patients' physiological and context parameters and physicians' information systems and devices. The supporting software should be able to run on any of the hardware and software platforms deployed in these environments, with possibly just minimal adaptation and customization required.

Interoperability This requirements is about the ability of a system to coexist and cooperate with other, possibly pre-existing, systems (e.g. legacy systems). Interoperability is usually achieved by the standardization of interfaces. Actually, all technologies and tools employed in remote healthcare systems need to interact with back-end systems already available and used in medical facilities. Similarly, information and data produced by remote healthcare systems should be also provisioned in a manageable format for software and tools physicians and medical personnel are familiar with.

Standards The wide adoption of global interoperability standards for e-health will facilitate the integration and interoperability of healthcare systems at a global scale and make it cost-effective. Historically, medical information systems have used proprietary formats only accepted by a single specific system and not inter-operable with the systems of other stakeholders. Electronic Health Records (EHR) (Baird et al. 2011; Eichelberg et al. 2005) aim at providing common digital formats and structures for integrating a variety of clinical information about a patient and support an easy exchange of such information by heterogeneous systems. The types of recorded information could include clinical observations, medical histories, treatments, medications, allergies, symptoms, etc.

Many standards organizations work in different areas of e-health. These include messaging (HL7), terminologies (SNOMED CT), clinical information and patient records (openEHR and HL7 CDA) and imaging (Digital Imaging and

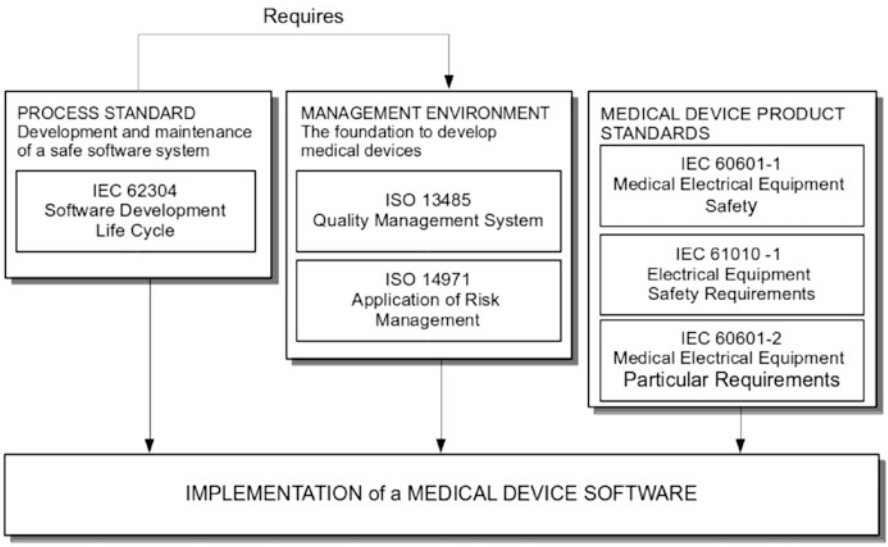


Fig. 2.3 Main medical standards for remote healthcare systems and their inter-connection

Communications in Medicine (DICOM)). In particular, the interoperability of EHRs is addressed in Health Level 7 (HL7) Clinical Document Architecture (CDA) (HL7 2000), CEN EN 13606 EHRcom (EN13606 2013) and openEHR (openEHR 2007). These standards aim to structure and mark-up the clinical content for the purpose of exchange (see Sect. 2.4.4 for further details) and their relationship is discussed in Peter et al. (2006), Khan et al. (2012b).

International Telecommunications Union (ITU)'s recommendations underlie much of the telecommunications infrastructure necessary to support remote healthcare services and standards like IEEE 802.15.6 (2012) and IEEE 802.15.4 (2013) are adopted on important components of remote healthcare systems, such as Body Sensor Networks (BSNs) and Wireless Sensor Networks (WSNs) to monitor physiological and context parameters of interest. Both BSNs and WSNs are based on Machine to Machine (M2M) communications. In this context, a dedicated ETSI technical committee is working on the development of standards for M2M communications (ETSI 2013).

ISO 14971 establishes the requirements for risk management to determine the safety of a medical device by the manufacturer during the product life cycle. Such activity is required by the ISO 13485 that includes the requirements for a comprehensive management system to design and manufacture medical devices. The standard IEC 62304 defines requirements for the whole life cycle of medical software and embedded software within medical devices. It applies to the development and maintenance of medical device software when software is itself a medical device or when software is an embedded or integral part of the final medical device. Compliance with IEC 62304 requires a quality management system (ISO 13485) and risk management (ISO 14971). The inter-connection of these medical standards for remote healthcare systems is shown in Fig. 2.3.

Safety A system is safe if it is protected against the consequences of failure, damage, error, accidents, harm or any other event which could be considered non-desirable. In the healthcare context, a system is safe if it does not create a hazard that could result in an injury for the users. In IEC/EN 62304 the manufacturer has to assign a safety class to the software system as a whole according to the following classification:

- *Safety Class A:* No injury or damage to health is possible;
- *Safety Class B:* Non serious injury is possible;
- *Safety Class C:* Death or serious injury is possible.

The effort to devise medical systems depends on these security classes and is clearly much higher (as are the associated costs) for class C devices and software than for class B or class A ones.

Privacy and Security Remote healthcare systems have the potentials to significantly increase the efficiency and effectiveness of healthcare delivery. These benefits, however, must be balanced with users' privacy and security concerns to make these systems socially acceptable (Hiller et al. 2011; Meingast et al. 2006). The availability of EHR on the Internet will expose such data to possible malicious attacks. Furthermore, the inherent distributed nature of remote healthcare systems will significantly increase the difficulty in implementing suitable and ad-hoc secure policies in every environment the system will be executed in. As an example, wireless communications at home should be protected against possible eavesdropping. In order to meet this requirement, the following questions should first be answered: who is the owner of the data? what data should be stored in the system? where? who can be granted access to such data? when? for which purpose? to whom should this information be disclosed without the patient's consent?

Performance In Table 2.1, there is a rough indication of the data rates necessary to transmit the data acquired by a number of physiological and context sensors as a function of the accuracy and of the bandwidth of the observed signal. Some other useful information can be found at PhysioNet that offers free web access to large collections of recorded physiological signals in its PhysioBank (2012) as well as related open-source software (PhysioToolkit) to analyse the traces.

It is worth noticing that the individual data rates are relatively small, but when combined together, they can pose a challenge to existing energy-efficient wireless technologies. Indeed, the data-rate of widely adopted wireless technologies such as IEEE 802.15.4 (2013) is normally up-to 250 kbps.

Most sensor nodes employed in remote healthcare systems are battery powered. As a consequence, a major problem for these devices is their lifetime; they must operate for a sufficiently long time to concurrently support long-time and reliable clinical observations without bothering users with unnecessary maintenance procedures (see usability) such as battery replace. A common technique to extend the battery life-time is to reduce the duty-cycling, namely the percent of time that a node is active, and thus consumes energy, as a fraction of the total time under consideration; a duty-cycle of 50%, in principle guarantees a doubled operation time. However, the use of such techniques decreases the responsiveness of the systems. When nodes are in non-active state, they cannot promptly react to specific

events. As a consequence, there is an increase in communication latency and the constraints usually associated with real-time systems can pose severe limits to the employment of duty-cycling. In summary, the trade-off between the latency of the system and its energy efficiency must be always carefully considered to select the most appropriate parameters and optimize the performance of the whole system.

2.3 High-Level Architecture Description of Remote Healthcare Systems

In order to meet the requirements introduced in the previous sections, we have conceived a reference architecture for remote healthcare systems. Figure 2.4 shows a first high-level description of this architecture by displaying the main actors involved together with the most relevant functional blocks which are further described in terms of more specific components. In the following, we adopt a top-down approach to better illustrate a proposed Service Oriented Architecture (SOA) underlying remote healthcare systems in order to promote alignment among other SOA clinical implementations and further interoperability among healthcare organizations (e.g. Healthcare Services Specification Project (HSSP) (2013)).

According to Fig. 2.4, patients' physiological data are gathered by a wearable body sensor network, equipped with suitable sensors (e.g. ECG, glucose meters etc.) to measure the physiological parameters of interest. The data produced by wearable body sensor networks, possibly enriched by additional context data (e.g. temperature, humidity, activity patterns) provided by a wireless sensor network, are eventually transmitted to a sink node called Healthcare Gateway (GW) (see Alemdar and Ersoy (2010) for a comprehensive survey on wireless sensor networks for healthcare). This gateway can locally store all patient's data in a Remote Repository before eventually transmitting them over a secured network (e.g. SSL/HTTPS Internet protocols) to a Global Repository. Physicians can access patient's data in the Global Repository with ad-hoc diagnostic Web-based tools and mobile applications (e.g. on smartphones and tablets) helping them properly carry out their medical practice in full respect of patients' privacy constraints and policies (e.g. see Shankar et al. (2012) for a privacy framework developed to design home-based computing systems for "aging in place" seniors). In other words, supported provision and increasingly sophisticated analysis of patients' physiological and context data can be performed securely from remote locations towards a ubiquitous and continuous diagnosis and treatment of diseases, injuries and other physical and mental impairments.

From a more technical perspective, the Global Repository is made up of a set of standard interfaces (API) sitting on top of a middleware system offering a number of services to properly collect, manage and provision patients' EHR (openEHR 2007). Please observe that in the context of the following sections, where we disclose more details on envisaged architecture's lower-level components, the Healthcare Gateway collecting patients' clinical and context data and the tools and applications developed for the benefits of physicians are meant to represent external applications

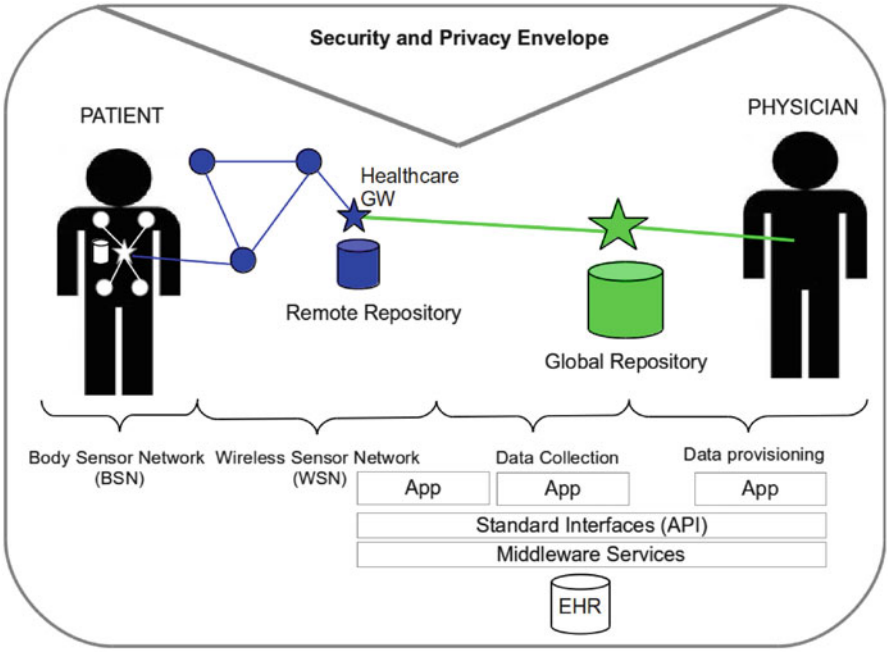


Fig. 2.4 High-level architectural description of the system

(with respect to the Global Repository) catering for data collection and data provision respectively. The architecture depicted in Fig. 2.4 is consistent with a system architecture for machine-to-machine communications in healthcare (Chen 2012; ETSI 2013).

2.4 Components of the System

In this section, the components sketched out in Fig. 2.4 and briefly introduced in the previous section are presented in greater details.

2.4.1 Context-Aware Patient Monitoring: Body Sensor Network

A BSN is a wireless network of wearable computing devices capable of gathering physiological and/or context data (Crosby et al. 2012; Khan et al. 2012a; Shaofeng Wang et al. 2011) and transmitting them to an on-body sink in order for this to store

and analyse such data. The sink is also able to transmit the collected data to the global repository by leveraging some middleware services, either by a direct connection (i.e., a mobile phone with a suitable data connection) or by a Wireless Sensor Network (see Sect. 2.4.2). The main difference between these two solutions is in terms of service duration and availability: while the former requires to recharge the phone battery almost every day, the latter is supposed to work for much larger periods of time. However, an on-body sink running on mobile phones (Nawka et al. 2011; Quero et al. 2007) greatly simplifies the overall architecture and will more conveniently support user mobility extending the availability of remote healthcare services also in outdoor environments.

Nowadays, BSNs are made of a limited number of nodes and in this context the star topology seems to be the most practical and efficient in terms of delay. Indeed, the IEEE 802.15.6 task group (TG6) (802.15.6 2012) recommends the 1-hop star and the 2-hop extended star (if necessary a node in the star behaves as a relay node to deliver data to the sink) as ideal network architecture candidates for BSNs. However, in the star topology, when Transmit Power Control (TPC) is employed, more distant nodes from the sink require more energy to transmit messages and can thus more quickly deplete their batteries. Furthermore, the proximity of antennas with the human body and the limited power of the employed wireless technologies result in a challenging environment in which, under particular circumstances, even nearby nodes are not able to communicate (Shah and Yarvis 2006). In Natarajan et al. (2007) the authors show that switching to a sufficiently dense multi-hop network, in which alternative routes can be exploited to deliver the messages, can be beneficial in terms of Packet Delivery Ratio, particularly in challenging outdoor settings where the nodes in a star need to significantly increase their power levels to successfully overcome the environmental impairments and reach the hub. Routing in a multi-hop BSN is an emerging area of research: there are protocols that try to reduce the issues related to radiation absorption and heating effects on the human body (Takahashi et al. 2008) or protocols that use clustering to reduce the number of direct transmissions to the sink (Watteyne et al. 2007). Anyway, the performance of routing algorithms in the case of multi-hop architectures or converge-cast algorithms in the context of stars are primarily dependent on the performance of the underlying MAC protocols. The IEEE 802.15.6 standard (802.15.6 2012) defines three physical layers: Narrowband (NB), Ultra wideband (UWB), and Human Body Communications (HBC) over which MAC protocols run. The standard defines three MAC access modes: Beacon Mode with super-frames, Non-Beacon Mode with super-frames and Non-Beacon Mode without super-frames. The details of these MACs are beyond the scope of this chapter and the interested reader can refer to 802.15.6 (2012), Tachtatzis et al. (2010), Kwak et al. (2010) while Latré et al. (2011), Gopalan and Park (2010) survey the research on MAC protocols for BSNs. We only mention here that the structure in frames defined in the IEEE standard allows the sink (i.e., the hub of the star) to schedule the allocation of slots to nodes in order to have contention-free periods where some QoS can be defined and guaranteed.

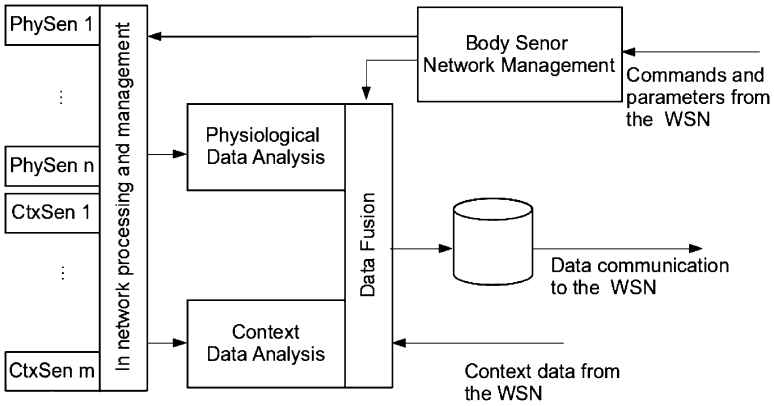
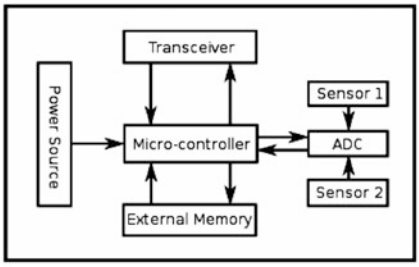


Fig. 2.5 Block diagram of a context-aware body sensor network

A block diagram describing the main components of a context-aware BSN and their mutual relationships is shown in Fig. 2.5. Physiological and context data are delivered to the BSN sink which analyses and fuses the received data to produce a context-aware picture of the physiological status of the observed patient. These enriched data are stored in a local repository and can be retrieved when necessary. Reliability and accuracy of context data can be improved by interacting with the WSN. Furthermore, the WSN can also provide new context data, such as the user's weight measured by a scale or the ambient temperature and humidity. The BSN Management Unit is in charge of managing all the activities connected with the acquisition and transmission of data in the BSN.

2.4.2 Extending the Scope of Monitoring: Wireless Sensor Network

Wireless Sensor Networks (WSNs) are wireless networks made of a number of autonomous sensor nodes capable to cooperate for monitoring a phenomena of interest and to report the observed data to a central node called sink (i.e., the healthcare gateway in Fig. 2.4). In some cases, WSNs also provide functions to control actuators deployed in the environment. In the reference architecture for healthcare systems depicted in Fig. 2.4, WSNs are used for three main purposes: (a) delivering the data acquired by the BSN to the healthcare gateway (i.e. the sink of the WSN), (b) getting contextual data that are difficult to gather in the BSN, such as the patient's weight and (c) controlling some environmental parameters by means of actuators to possibly increase the wellness of patients (e.g., the level of humidity in the environment).



Application	HTTP, CoAP
Transport	UDP, TCP
Routing (IP)	IETF RPL
Adaptation	IETF 6LowPAN
Medium Access	IEEE 802.15.4e
PHY	IEEE 802.15.4-2006

Fig. 2.6 The internal architecture of a sensor node (picture from Wikimedia Commons) and the proposed standards for WSNs protocol stack

Figure 2.6 shows a typical architecture for sensor nodes. It is important to note that all components of such architecture are resource constrained. A typical configuration could be: a microcontroller with 8 MHz of clock and 8K of RAM interfaced with a radio providing up to 250 kbps (see Table 1 in Healy et al. (2008) for a comprehensive comparison of sensor nodes hardware). Furthermore, sensor nodes are usually battery powered and the limited amount of energy available in nowadays batteries, which is not foreseen to be significantly improved in the near future (Paradiso and Starner 2005), imposes severe constraints to the design of solutions for WSNs. As an example, the power consumption of the TelosB, a state-of-the-art sensor node platform has been analysed in Prayati et al. (2010): the radio activities cost about 20 mA, while processing computations cost much less, from about 500 μA (1MHz clock) to about 2.4 mA (4MHz clock). Assuming the nodes powered by 2,000 mAh batteries and limiting our attention to radio activities only, a node would be able to continuously use the radio for only 100 h. A common technique to improve nodes lifetime is the duty-cycle, that is defined as the fraction of time in a given period the node is active; a duty-cycle of 10% can increase the node operation time by a factor 10. The price to pay for this improvement in operation time is latency. When nodes are not active, they cannot receive new incoming packets and thus the reception of such packets is deferred until the new active period of the next hop towards the destination. MAC protocols such as LPL (Polastre et al. 2004), S-MAC (Wei et al. 2002) and T-MAC (van Dam and Langendoen 2003) have been designed to exploit the duty-cycle principle. On top of such protocols, energy-efficient routing solutions have been proposed to reliably collect the data observed by the sensor nodes. Routing in WSN has been thoroughly investigated in recent years and there are a number of papers surveying the literature (Al-Karaki and Kamal 2004; Annie Uthra and Kasmir Raja 2012; Pantazis et al. 2013). In particular, some WSNs routing protocols propose a shift from address-centric routing to data-centric routing (Intanagonwiwat et al. 2003) which is claimed to be more

appropriate in the typical application scenarios for WSNs where users are more interested in collecting data rather than connecting with specific nodes.

Despite an intensive research activity during the last ten years, the effort to standardize the WSN protocol stack has only recently made some progress in view of the Internet of Things (IoT) paradigm. The IoT pursues a vision in which a globally interconnected continuum of uniquely identified devices, objects and things will provide an unprecedented quantity of data to measure, better understand and manage a plethora of phenomena. From this perspective, projects such as OpenWSN (2012) are working to provide an open-source implementation of a complete protocol stack based on the upcoming IoT standards. In particular, the physical layer is based on IEEE 802.15 (2013) which also standardizes a MAC layer. IEEE 802.15.4e is a working group proposing an amendment to the existing 802.15.4-2006 to better support industrial markets. The proposed solution introduces channel hopping to significantly increase robustness against external interference and multi-path fading. IETF 6LoWPAN is a working group developing encapsulation and header compression mechanisms to allow IPv6 packets to be efficiently delivered over 802.15.4 based networks. RPL (2011) is a routing protocol designed by the IETF ROLL group to operate on Low Power and Lossy Wireless Networks such as WSNs. In Fig. 2.4, the WSN is used to collect data in a multipoint (the node) to point (the sink) fashion. In this context, RPL builds a Destination Oriented Directed Acyclic Graph (DODAG) to route packets. A major benefit in implementing such protocol stack is that the transport protocols UDP and TCP can be used on top of RPL as well as HTTP and the Constrained Application Protocol COAP Internet-Draft (2013), a specialized Web transfer protocol that easily interacts with HTTP and is specifically designed to support a request/response interaction model and built-in discovery of services and resources in networks of constrained devices.

2.4.3 Data Collection and Provision: Applications and Middleware

Figure 2.7 shows an exploded view of the reference architecture proposed in Sect. 2.3, with a particular emphasis on the applications for data collection/provision and the middleware services deployed on top of the EHR maintained in the Global Repository (as depicted in Fig. 2.4).

The layered architectural design pictured in Fig. 2.7 is meant to loosely adhere to the Service Oriented Computing (SOC) paradigm based on the concept of "service" as a fundamental element for developing distributed service oriented architectures in a network based environment. As also reported in HSSP (2013), SOC provides an approach for business transformation by dividing complex environments into well defined, formally specified functions based on the activities that they perform (services). Each service has well defined responsibilities, authority and interfaces. These services then work together in collaboration to support the workflow of the

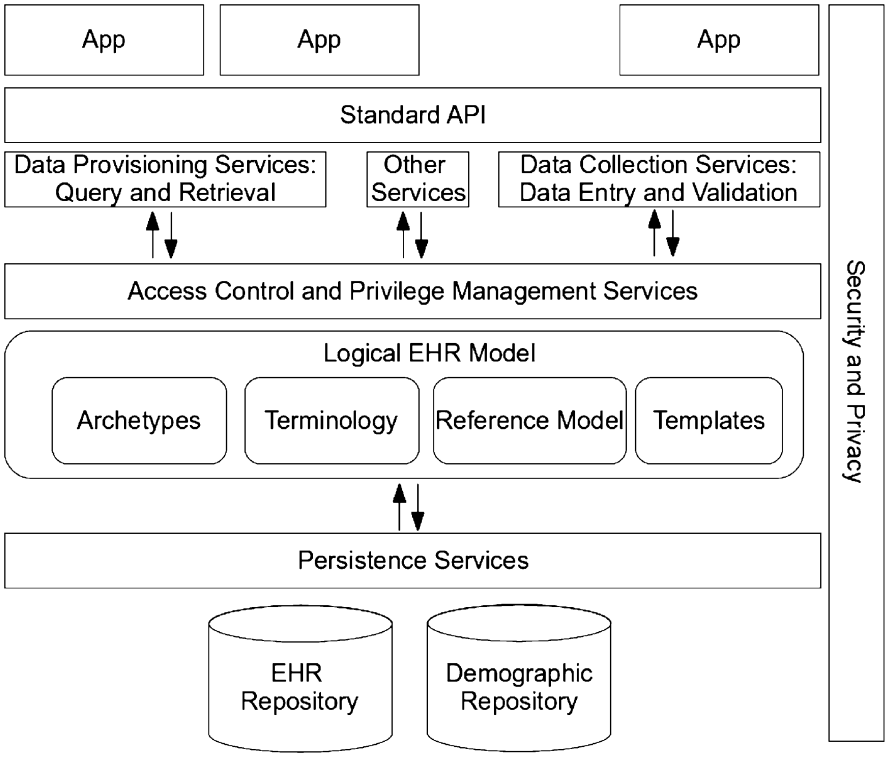


Fig. 2.7 Service oriented architecture for remote healthcare systems

business, all within the context of governance, oversighting and orchestration modules that manage their coordination and performance.

In the proposed reference architecture, applications are supposed to be Web-based, mobile and embedded software systems gathering and/or displaying patients' clinical and context data. According to this vision, both the wearable body sensor networks and the wireless sensor networks presented in previous Sects. 2.4.1 and 2.4.2 respectively, cooperate to gather patients' data. This data is conveyed to the middleware by means of an Healthcare Gateway interacting via standard API with available Data Collection Services concerned with the management of data entry and validation. Similarly, middleware's Data Provisioning Services are responsible for querying and retrieving patients' data upon specific requests made by physicians' applications (e.g. Chan et al. (2007) and Cerny (2010)). Data Collection and Data Provisioning Services may possibly interact with additional middleware services coping, for instance, with specific network communication protocols' rules (e.g. TCP/IP, HTTP, Web Services) and/or data formats (e.g. proprietary, XML, JSON). Moreover, to further support and achieve a higher degree of module independence, system robustness, scalability and general component distribution in a network environment,

different types of interaction paradigms and data access, such as asynchronous, possibly periodic, publish/subscribe messaging and synchronous on-demand polling policies, are envisaged.

Other traditional middleware services included in the reference architecture work on access control and privilege management. In addition to properly keeping track of users' identity and their possible authorisation set on applications' domains and other resources of interest, these services closely interact with architecture modules responsible for security and privacy issues. As already mentioned in Sect. 2.3, it is important that remote healthcare systems supporting patients and informing caregivers do not abrogate privacy (Shankar et al. 2012) when dealing and maintaining sensitive information such as those contained in the EHR.

Persistence services are included in the reference architecture's middleware to hide out physical details and organization of the system's data of interest, i.e. patients' electronic health records and demographic data. In line with the approach proposed in openEHR (2007), these services should include API such as the Virtual EHR API (vEHR). Such API mostly helps map the Logical EHR Model, addressed in the next section, with suitable key functions for CRUD (Create, Read, Update, Delete) operations to be executed in respect of performance target parameters.

2.4.4 Global Repository: Electronic Health Records

The first effort to standardize electronic healthcare records has been made by ENV13606. The *Knowledge* is a general statement, namely it applies to all members of a population, describing our understanding of the world, while *Information* is a statement on a specific member of the population. In ENV13606 the so called component oriented single model approach foresees that both knowledge and information are tightly coupled. This simplifies the implementation of practical solutions, but it makes extremely hard to follow the evolution of the knowledge in specific domains. As an example, the current knowledge about hypertension asserts that when the diastolic pressure is above 100 there might be consequences on the patient's health status. Thus, when a person provides the information that his/her pressure is above this threshold, the physician diagnoses a possible problem. Now, assume that further discoveries in the healthcare domain reduce the threshold to 95 instead of 100. If the system is designed to decouple knowledge from information, this new discovery has a minimal impact on the whole system. This is why the EN13606 (2013) has been designed as an evolution of ENV 13606 which actually provides a dual approach where knowledge and information are decoupled.

The EN 13606 is based on a Dual Model Architecture in which there is a clear separation between information and knowledge. Information is structured through a *Reference Model* that contains the basic entities to represent the information of the EHR. Knowledge is based on *Archetypes*, namely formal and computable definitions of a medical concepts expressed in the Archetype Definition Language (ADL) in the form of structured and constrained combinations of the entities of a Reference Model.

"The interaction of the Reference Model (to store data) and the Archetype Model (to semantically describe those data structures) provides an unseen capability of evolution to the information systems. Knowledge (archetypes) will change in the future, but data will remain untouched" (EN13606 2013). More in details, a Reference Model is formalized as an Object Oriented Model that defines the classes through which an EHR can be constructed. A reference Model contains: (a) a set of primitive types and data structures, (b) a set of classes that are the building blocks of EHRs; the six entities defined in EN13606 are: folder, composition, section, entry, cluster and element, (c) a set of auxiliary classes used to describe the context information attached to an EHR annotation, such as versioning information, (d) classes for describing demographic data and to communicate EHR fragments.

An archetype is made of three main sections: header, definition and ontology. The *header* contains meta-data such as an identifier and authoring information. In the *definition*, a structured and constrained combination of entities of a Reference Model is used to represent a particular clinical concept. The constraints used in the definition are of different types, such as constraints on the range of attributes of primitive types, on the existence and cardinality of attributes etc. The *ontology* section is where the entities identified in the definition section are described and bound to terminologies. Archetypes can be specialized by re-using previously defined archetypes and providing further constraint on information already expressed by such archetypes. The root of the specialization hierarchy is the Reference Model and more specialized archetypes are in the deeper levels of the hierarchy. Instances of specialized archetypes are also instances of parent, more general archetypes.

Archetypes define medical concepts that are combined in *Templates* to define use-case specific data-sets.

2.4.5 Security and Privacy

While remote healthcare systems will provide many benefits for advanced health care delivery, there are a number of security and privacy implications that must be explored and analysed in order to be technically reliable, ethically acceptable and meet fundamental social expectations. Some of the common security issues shared by traditional healthcare applications regarding EHRs are ownership of information, authenticity, authentication, non-repudiation, patient consent and authorization, integrity and confidentiality of data (Zhang and Liu 2010) and general governing policies. Although varying greatly from country to country about the addressed details, these issues are already controlled by many different regulations surrounding health care such as the ISO/TS 18308 standard (providing the definitions of security and privacy issue specifically for EHR) (ANSI 2003), the American Health Insurance Portability and Accountability Act of 1996 (HIPAA) (HIPAA 2011), the Health Information Technology for Economic and Clinical Health Act (HITECH 2011) and the European Advanced Informatics in Medicine/Secure Environment for Information Systems in MEDicine (AIM/SEISMED)

(Van Dorp and Dubbeldam 1994). However, security and privacy features and corresponding legal rules, procedures and laws need to be constantly updated and re-evaluated as new technology changes are introduced and overall supporting systems evolve over time.

In the context of the proposed vision of future remote healthcare systems, the integration of innovative WSNs and BSNs with more standard and conventional middleware and back-end systems poses new challenges for security and privacy related issues and requires a more careful review of mechanisms and approaches deemed suitable for peculiar and extremely resource-constrained components. Essentially, in order to be in accordance with the approach captured by the general architecture of Fig. 2.7 envisaging an overall security and privacy protection throughout the layers, special care must be devoted to properly combine traditional security and privacy preserving mechanisms in more controlled and dominated environments (i.e. LANs, WANs, the Web) with specific security issues in wireless healthcare applications using advanced innovative tools such as WSNs and BSNs.

An interesting EHR security reference model is proposed and discussed in Zhang and Liu (2010). By means of three core components and secure interactions among them to address and implement security and privacy requirements for patient care delivery in an EHR cloud-based system, it includes some core features that we find suitable for the proposed middleware services and back-end Global Repository. In particular, a first component is concerned with EHR secure collection, verification of authenticity, confidentiality, integrity and non-repudiation, compliance with regulations as well as integration of verified EHR data into a new composite EHR with a security certificate signed by the integrator. The other components are responsible for EHR secure storage and access management and EHR secure usage model respectively. Selection of encryption schemes and protocols, access control policies, cryptographic access control strategies and key management are all ensured and guaranteed as required by a standard secure enterprise architecture and therefore deemed satisfactory for the remote healthcare system's reference architecture proposed herein.

As per the previously discussed BSNs and WSNs, additional security and privacy vulnerabilities and threats need to be properly taken care of and some ad-hoc mechanisms already exist and are being devised that we consider and propose for inclusion in future remote healthcare systems. According to existing literature (Kumar and Lee 2012), typical security threats for wireless medical sensor networks involve malicious attacks and access to sensitive patient physiological and context data. Harmful examples of such attacks include monitoring and eavesdropping on patient vital signs, threats to information when in transit during a multi-hop route, routing threats to steal or modify data packets and forward the altered packets to a remote center that may cause false alarms or prevent true ones. Additional threats impact on patients' location, activity tracking and may trigger a Denial-of-Service (DoS) attack that diminishes or eliminates a network's capacity to perform its expected function. This could result as being even more disruptive in healthcare systems requiring always-on patient health monitoring (e.g. in-home, in-hospital) and possibly lead to very serious consequences such as the loss of a patient's life. Privacy

issues are more bound to be related to social impacts that an uncontrolled sharing of sensitive data may determine. Many actors are involved in a care giving process each having an important role in the context of a supporting information system: patients, physicians, insurance companies, health-coaches, patients' family. In Kumar and Lee (2012) authors report that in order to maintain privacy, patients should have the rights to determine which data should be collected, used or disclosed. Also which subject should be granted access to patient data should be carefully established. Any unauthorized collection or leakage of patient data could harm the patient. So, both identity threats and access threats need to be carefully considered to avoid eavesdropping and leakage of information potentially causing a misuse of that information that can result in patients being severely affected physically, psychologically, socially and financially. Security and privacy preserving mechanisms for traditional networks are usually inapplicable to resource-constrained sensor networks. As discussed in Sect. 2.4.2, computing power, memory, energy and execution times are the biggest issues that need to be tackled in order to ensure that these networks meet the security and privacy requirements. Strong cryptography, for instance, requires extensive computation and resource and therefore selecting appropriate cryptography are a challenging task (Kumar and Lee 2012). Furthermore, the selection of cryptography system depends on the computation and communication capability of the sensor nodes which may vary greatly depending on adopted technologies. Similarly, the related concept of key management can be impacted by the available sensor nodes. Some research suggests that key pre-distribution protocols based on symmetric key cryptography and self enforcing protocols using a public-key infrastructure can be equally suitable for wireless sensor networks and viable on resource constrained networks (see Kumar and Lee (2012) and refs within). Other features such as secure routing, resilience to node capture and secure localization may be harder to include given the technology state-of-art, whereas DoS countermeasures exist at different layers of WSN routing and are suitable enough for particular applications relying on WSNs.

In conclusion, a thorough assessment of security and privacy issues must be made for a real future remote healthcare system before it is released to the general public. While no particular problems exist in the proposed reference architecture to make middleware and back-end systems strongly secure and privacy-enabled, some open research challenges still impact on the BSN and WSN components. These are mostly due to an objective difficulty in adopting security and privacy preserving measures as-is on tiny devices with well-know constraints in terms of available computing and communication power, energy consumption and memory size.

2.5 Open Problems and Conclusions

In the present chapter we have presented a technical approach towards the definition, design and implementation of next-generation remote healthcare systems. In particular, we have identified some general functional and non-functional requirements that need to be properly met to ensure that

physiological data and environmental conditions where this data is captured can be dealt with over distributed communication networks (e.g. the Internet). Furthermore, we have designed a Service Oriented reference architecture serving the purpose of successfully addressing the aforementioned requirements and showing what specific functional blocks should be put in place to devise and implement actual future healthcare systems. The proposed and thoroughly discussed integration of many components very different from one another reveals a high complexity of such systems. A multi-faceted approach and complementary technical skills and scientific knowledge are envisaged to tackle all the challenges posed by required technologies as peculiar and specific as BSNs, WSNs, EHRs and a middleware software. Moreover, engineers and system developers should be especially concerned with security and privacy issues whose open research points and questions can be particularly hard on the tiny pervasive sensor components and networks (i.e. BSNs and WSNs). Next-generation remote healthcare systems constitute a fascinating and promising area in the context of future emerging e-health technologies. While continuous efforts are made and intensive activities are carried out to further refine the very concept of e-health and better point out all related initiatives (i.e. cultural, medical, sociological), we deem it advisable not to overlook a general system-based approach like the one suggested in this chapter when it comes to successfully engineer an effective supporting systems.

The majority of open problems that emerge in this overview chapter lies in the effective and efficient collection of physiological and context data in remote sites. The limited amount of energy available in current state-of-the-art batteries still dictates the design of solutions that try to strike a balance between energy efficiency and quality of data. Such solutions need to exploit advanced techniques developed both at hardware and software level and have to be designed to be compliant with the emerging standards. Most of the technologies employed in the other parts of the reference architecture can currently be considered commodities. However, the user-friendly principles that should drive the design of such systems, in which both patients and doctors are not necessarily ICT experts, always require an effort of simplification in the use of the available technologies. A major aspect of that stems out from these two vital points of energy consumption and ease of use is the number of sensors deployed in the system. There are several sensors already in the market that may capture different vital sign data. However, the use of lots of sensors is not practical both from the energy consumption and ease of use point of view. From practical experience, patients normally feel uncomfortable to have more than three patches on their body. Therefore, heterogeneous sensor integration and streamlining their communication is a vital aspect and maximal information from minimal number of sensors is also a key aspect. This is an open research question which needs to be addressed not only from the sensor manufacturing point of view, but also in terms of their communication to the gateway.

References

6lowpan Status Pages IPv6 over Low power WPAN (Active WG). http://tools.ietf.org/wg/6lowpan/

802.15.6-2012 (2012) IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks http://standards.ieee.org/findstds/standard/802.15.6-2012.html

Alemdar H, Ersoy C (2010) Wireless sensor networks for healthcare: A survey. Comput Network 54(15):2688–2710

Al-Karaki JN, Kamal AE (2004) Routing techniques in wireless sensor networks: a survey. IEEE Wireless Comm 11(6):6,28

Annie Uthra R, Kasmir Raja SV (2012) QoS routing in wireless sensor networks? a survey. ACM Comput Surv 45(1)

ANSI (2003) ISO/TS 18308 Health informatics-requirements for an electronic health record architecture, ISO 2003

Baird A, North F, Raghu TS (2011) Personal Health Records (PHR) and the future of the physician-patient relationship. In: Proceedings of the 2011 iConference (iConference '11). ACM, New York, NY, pp 281–288

Chan DS, Callahan CW, Hatch-Pigott VB, et al (2007) Internet-based home monitoring and education of children with asthma is comparable to ideal office-based care: results of a 1-year asthma in-home monitoring trial. Pediatrics 119(3):569–578

Chen K-C (2012) Machine-to-machine communications for healthcare. J Comput Sci Eng 6(2):119–126

Cerny M (2010) "Movement activity monitoring of elderly people," Second international conference on Computer Engineering and Applications (ICCEA), 2010, 2, pp 454,455, 19–21 March 2010

Crosby GV, Ghosh T, Murimi R, Chin CA (2012) Wireless body area networks for healthcare: a survey. Int J Ad Hoc Sensor Ubiquitous Computing (IJASUC) 3(3)

Eichelberg M, Aden, Jörg Riesmeier T, Dogac A, Laleci GB (2005) A survey and analysis of electronic healthcare record standards. ACM Comput Surv 37(4):277–315

ETSI (2013) Machine-to-Machine. http://www.etsi.org/technologies-clusters/technologies/m2m

Glinz M (2007) 15th IEEE international on non-functional requirements, requirements engineering conference, 2007. RE '07, pp 21–26

Gopalan SA, Park J-T (2010) 2010 International congress on energy-efficient MAC protocols for wireless body area networks: Survey, ultra modern telecommunications and control systems and workshops (ICUMT), pp 739,744, 18–20 Oct. 2010

Healthcare Services Specification Project, Health Level Seven, Object Management Group (2013) The Practical Guide for SOA in Health Care - A real-world approach to planning, designing, and deploying SOA. Version 1.0

Health Level 7 (HL7) (2000) Clinical Document Architecture (CDA). http://www.hl7.org/implement/standards/

Healy M, Newe T, Lewis E (2008) Wireless sensor node hardware: A review. IEEE Sensors 621–624

Hiller J, McMullen MS, Chumney WM, Baumer DL (2011) Privacy and security in the implementation of health information technology (electronic health records): U.S. and EU compared. Boston Univ J Sci Tech Law 17(1):1

IEEE 802.15 (2013) WPAN Task Group 4 (TG4). http://www.ieee802.org/15/pub/TG4.html

IEEE Guide to Software Requirements Specifications, IEEE Std 830–1984, 1984

Intanagonwiwat C, Govindan R, Estrin D, Heidemann J, Silva F (2003) Directed diffusion for wireless sensor networking. IEEE/ACM Trans Netw 11:1

Internet-Draft (2013) Constrained Application Protocol (CoAP). draft-ietf-core-coap-15 http://datatracker.ietf.org/doc/draft-ietf-core-coap/

Khan NA, Javaid N, Khan ZA, Jaffar M, Rafiq U, Bibi A (2012a) Ubiquitous healthcare in wireless body area networks, trust, security and privacy in computing and communications (TrustCom), 2012 11th international conference on IEEE, pp 1960,1967, 25–27 June 2012

Khan WA, Khattak AM, Lee S, Hussain M, Amin B, Latif K (2012b) Achieving interoperability among healthcare standards: building semantic mappings at models level. In: Proceedings of the 6th international conference on ubiquitous information management and communication (ICUIMC '12)

Krasich M (2009) How to estimate and use MTTF/MTBF would the real MTBF please stand up? Reliability and maintainability symposium, 2009. RAMS 2009, pp 353–359, 26–29 Jan. 2009

Kumar P, Lee H-J (2012) Security issues in healthcare applications using wireless medical sensor networks: a survey. Sensors (Basel) 12(1):55–91

Kwak K-S, Ullah S, Ullah N (2010) An overview of IEEE 802.15.6 standard. In: 3rd international symposium on applied sciences in biomedical and communication technologies (ISABEL), pp 1–6

Latré B, Braem B, Moerman I, Blondia C, Demeester P (2011) A survey on wireless body area networks. Wireless Network 17(1):1–18

Martínez A, Everss E, Rojo-Álvarez JL, Figal DP, García-Alberola A (2006) A systematic review of the literature on home monitoring for patients with heart failure. J Telemed Telecare 12:234–241

Meingast M, Roosta, T, Sastry S (2006) Security and privacy issues with health care information technology, engineering in medicine and biology society, 2006. EMBS '06. 28th Annual international conference of the IEEE, pp 5453, 5458

Natarajan A, Motani M, de Silva B, Yap K-K, Chua KC (2007) Investigating network architectures for body sensor networks. In: Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments (HealthNet '07)

Nawka N, Maguliri AK, Sharma D, Saluja P (2011) SESGARH: A scalable extensible smartphone based mobile gateway and application for remote health monitoring. In: 5th International Conference on internet multimedia systems architecture and application (IMSAA), 2011 IEEE, pp 1,6

Office for Civil Rights, United State Department of Health and Human Services Medical Privacy (2011) National Standards of Protect the Privacy of Personal-Health-Information. Available online: http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/index.html (accessed on 15 June 2011)

Office for Civil Rights, United State Department of Health and Human Services Medical Privacy (2011). National Standards of Protect the Privacy of Personal-Health-Information. Available online: http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/index.html (accessed on 15 June 2011)

openEHR (2007) S. Heard & T. Beale (eds.) "openEHR Architecture Overview". openEHR Foundation, https://github.com/openEHR/specifications/blob/master/publishing/architecture/overview.pdf?raw=true

OpenWSN (2012) https://openwsn.atlassian.net/wiki/display/OW/A+Standard+World

Pantazis NA, Nikolidakis S, Vergados DD (2013) Energy-efficient routing protocols in wireless sensor networks: A survey. IEEE Comm Surv Tutorials 15(2):551–591

Paradiso JA, Starner T (2005) Energy scavenging for mobile and wireless electronics. IEEE Pervasive Computing 4(1):18,27

Peter S, Thomas B, George H, Sam H, Heather L (2006) The relationship between CEN 13606, HL7, and OpenEHR. In: Health informatics conference (2006)

PhysioBank archive (2012) http://www.physionet.org/physiobank/

Polastre J, Hill J, Culler D (2004) Versatile low power media access for wireless sensor networks. In: Proceedings of the 2nd ACM international conference on Embedded networked sensor systems (SenSys '04), pp 95–107

Prayati A, Antonopoulos C, Stoyanova T, Koulamas C, Papadopoulos G (2010) A modeling approach on the TelosB WSN platform power consumption. J Syst Software 83(8):1355–1363

Quero JM, Tarrida CL, et al (2007) Health care applications based on mobile phone centric smart sensor network. In: 29th Annual International Conference of the IEEE Engineering in medicine and biology society, 2007. EMBS 2007, pp 6298–6301, 2007

RPL (2011) IPv6 Routing Protocol for Low power and Lossy Networks. Draft-ietf-roll-rpl-19 http://tools.ietf.org/html/draft-ietf-roll-rpl-19

Shah RC, Yarvis M (2006) Characteristics of on-body 802.15.4 networks. In 2nd IEEE workshop on Wireless Mesh Networks, 2006. WiMesh 2006, pp 138–139

Shankar K, Camp LJ, et al (2012) Aging, privacy, an home-based computing. IEEE Pervasive Computing 11(4):46–53

Shaofeng Wang, Lianying JI, Aiguang LI, Jiankang WU (2011) Body sensor networks for ubiquitous healthcare. J Control Theory Appl 9(1):3–9

Shea S, Weinstock R, Starren J, et al (2006) A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus. J Am Med Informat Assoc 13:40–51

Tachtatzis C, Di Franco, F, Tracey DC, Timmons NF, Morrison J (2010) An energy analysis of IEEE 802.15.6 scheduled access modes. In: IEEE GLOBECOM Workshops, 2010, pp 1270–1275

Takahashi D, Xiao Y, Hu F, Chen J, Sun Y (2008) Temperature-aware routing for telemedicine applications in embedded biomedical sensor networks. EURASIP J Wireless Comm Networking 2008(Article 26)

The CEN/ISO EN13606 standard (2013) http://www.en13606.org/

van Dam T, Langendoen K (2003) An adaptive energy-efficient MAC protocol for wireless sensor networks. In: Proceedings of the 1st ACM international conference on Embedded networked sensor systems (SenSys '03), pp 171–180

Van Dorp HD, Dubbeldam JF (1994) The AIM SEISMED guidelines for system development and design. Int J Biomed Comput 35(Suppl):179–186

Watteyne T, Augé-Blum I, Dohler M, Barthel D (2007) Anybody: A self-organization protocol for body area networks. In: Proceedings of the ICST 2nd international conference on Body area networks (BodyNets '07)

Wei Ye, Heidemann J, Estrin D (2002) An energy-efficient MAC protocol for wireless sensor networks, INFOCOM 2002. In: Twenty-First annual joint conference of the IEEE computer and communications societies, vol 3, pp 1567,1576

Zhang R, Liu L (2010) Security models and requirements for healthcare application clouds. Proceedings of IEEE 3rd international conference on cloud computing CLOUD 2010, vol 80(2), pp 268–275

Chapter 3
Vital Sign Sensing Technology

Andy Cranny, Andoni Beriain, Hector Solar, Gennaro Tartarisco, and Giovanni Pioggia

Abstract The four physiological measures of body temperature, pulse rate, respiration rate and blood pressure have for a long time been considered as vital signs in the diagnosis of a patient's health. It is also widely accepted that the routine measurement of other physiological or biological signals, possibly pathology specific, would help considerably in diagnosis and early stage treatment. Such measurements might include, for example, heart activity, brain activity, blood glucose level or mobility. Furthermore, the development of portable systems that can make a number of different health related measurements would prove beneficial in the monitoring of patients during treatment, recovery or rehabilitation. Technologies and instruments that can make these measurements have existed for some time, but factors such as their cost, lack of portability and in some instances, a requirement for expert knowledge, have restricted their wide scale use. Today, however, advances in information technology, communications and microfabrication techniques have made possible the realisation of truly portable systems for the measurement of a wide range of physiological signs at any medical intervention. This chapter describes the sensing technologies and systems currently being developed, or that are in use, for the measurement of a new, larger range of vital signs.

A. Cranny (✉)
School of Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK
e-mail: awc@ecs.soton.ac.uk

A. Beriain • H. Solar
Department of Electronics and Communication, Centre of Studies and Technical Research (CEIT) and School of Engineering of the University of Navarra (Tecnun), Madrid, Spain
e-mail: aberian@ceit.es; hsolar@ceit.es

G. Tartarisco • G. Pioggia
Institute of Clinical Physiology—CNR, Pisa, Italy
e-mail: gennaro.tartarisco@ifc.cnr.it; giovanni.pioggia@ifc.cnr.it

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_3,

3.1 Introduction

Modern discoveries in pharmacology and continuing improvements in medical healthcare have meant that we are living for longer. The World Health Statistics report for 2013 states that global life expectancy at birth for both sexes increased from 64 in 1991 to 70 in 2011 and furthermore, that global life expectancy at age 60 for both sexes increased from 18 in 1911 to 20 in 2011 (World Health Organization 2013). The same report also reveals that the annual growth rate for the world population over the decade 2001–2011 was on average 1.2 %, culminating in a total population in 2011 of over 6.9 billion people, of which 11 % were aged over 60. In 2000, only two of the countries within the Organisation for Economic Cooperation and Development (OECD, 34 member countries) had a total life expectancy of 80 years or more (OECD 2011). By 2009, 22 OECD countries had reached this milestone. These gains in life expectancy reflect large declines in mortality at all ages. Cardiovascular diseases (comprising mostly heart attack and stroke) remain the leading cause of death in OECD countries, though mortality rates have been cut by more than half since 1960.

This progressive rise in population and change in demographics is changing the needs and priorities of society, and brought about new problems that are being addressed in some instances by innovative solutions and the adoption of new paradigms. For example, concerns over the ever increasing healthcare costs of an ageing population, particularly those related to the treatment of chronic arthritis, cardiovascular diseases or neurodegenerative diseases, have driven research and development into telemedicine based systems that make possible cost efficient patient monitoring and rehabilitation within the home environment (Hung et al. 2004; van Halteren et al. 2004; Lorincz et al. 2004; Otto et al. 2006; Varshney 2007; Fletcher et al. 2010; Nangalia et al. 2010; Jovanov and Milenković 2011; Maharatna et al. 2012). Admission to a care home can be expensive, particularly long-term care, when compared with the provision of care directly at a patient's home and, generally, the patient prefers staying at home rather than being admitted to a care facility. In 1960, health spending across the OECD member countries accounted on average for under 4 % of gross domestic product (GDP), but by 2009 this had risen to 9.6 %, and in a dozen countries health spending accounted for over 10 % of GDP. For example, in 2009 the USA spent 17.4 % of their GDP on healthcare delivery, while this figure was around 11 % for major European countries (e.g. France, Germany) (OECD 2011).

Technological advances have been key in realising these changing scenarios of care transferral from hospital towards home. An important element is the ongoing digital revolution, which is inevitably contributing to increase the level of patient-caregiver-clinician connectivity. Today wireless computing infrastructures introduce an unprecedented opportunity to enable smart wearable systems (SWS) to deliver healthcare at home (Paradiso 2003; Bonato 2005; Sung et al. 2005; Pantelopoulos and Bourbakis 2010; Roggen et al. 2011; Poon et al. 2011; Patel et al. 2012; Chan et al. 2012). The aim of SWS is to match the living environment with the abilities of subjects affected by diseases or suffering from disabilities, in order to reduce as

much as possible the risks for those patients, to enhance their abilities, to support their independent living and to support their rehabilitation. SWSs in conjunction with the processing capabilities of smartphones or other personal communication devices can enable real-time analysis and feedback to patients, their carers and medical staff. This approach, also known as ecological intervention, can be provided to people during their everyday lives (i.e. in real time) and in natural settings (i.e. real world).

Intelligent, portable medical monitoring devices that are capable of processing data in real-time, are increasingly being developed. But this revolution in healthcare technology is not confined just to the hospital or home environment. Such systems are also employed in ordinary daily life, and are often used by athletes and healthy individuals because of their capability to give early feedback about personal health state (Ermes et al. 2008). Such systems are sometimes capable of triggering alarms when measured parameters reach levels that indicate the requirement for emergency medical intervention, and are often linked to primary care and ambulance switchboards (Lorincz et al. 2004; Lee et al. 2006). Mobile systems are also being developed that make possible the measurement and monitoring of vital physiological signs directly at the point of first response, whether this is at triage, at the scene of a traffic accident or during the critical stage of patient transportation to a medical facility (Lin et al. 2004; Gao et al. 2005; Suzuki et al. 2009).

SWS for monitoring health may include several devices or modules, both wearable and implantable, and can be classified as sensors and actuators, power supplies and wireless communication networks. Also included in this classification are decision support systems, algorithms for data processing, smart fabrics, processing units, software, multimedia devices and user interfaces. Another possible classification of such systems is based on the measured variables. These systems can be classified into, among for example, those regarding the measurement of vital signs, devices for the monitoring of skin temperature, heart rate, blood pressure, respiration rate, blood oxygen saturation, etc.

SWSs can be worn by an individual as an accessory, i.e. wristwatch, wristband, or as an electronic patch, chest belt, shirt, shoe, glove, or embedded in the user's outfit as part of their clothing. The data obtained by these kinds of sensors are directly sent by a wireless sensor network to a central processing hub, which may be worn on the body, or may be locally stored and downloaded to the hub at a convenient time (e.g. while recharging the system). This data is accessible to a medical centre, through the internet for example, in order to allow physicians to properly manage patients.

A large variety of laboratory prototypes, test beds and industrial products have already been produced. The role of SWS is to match the living environment with the physical and cognitive abilities and limitations of those suffering from disabilities or diseases, thereby enhancing performance and minimising the risk of illness, injury, and inconvenience. Presently, egocentric sensing systems for the unobtrusive monitoring of body-kinematics and physiological and behavioural signals are continuously improving. SWSs integrate smart sensors together with on-body signal conditioning and pre-processing, as well as the management of the energy consumption and wireless communication systems. Integrated wearable systems are able to measure

a variety of important medical signals including: arterial oxygen saturation; electrocardiographic signals (ECG); electroencephalographic signals (EEG); electromyographic signals (EMG); heart rate and respiratory values. Acquired information is correlated to obtain blood pressure, body temperature, hearth rate variability (HRV), end tidal CO_2 and thoracic impedance pneumographic values.

3.2 Vital Signs

Vital signs are measurements of the body's most basic functions. The four main vital signs routinely monitored by medical professionals and healthcare providers, with their normal ranges for the average healthy adult, are: body temperature (36.3–37.3 °C, average 36.9 °C); pulse rate (60–100 beats per minute); respiration rate (12–18 breaths per minute) and blood pressure (90/60 to 120/80 mmHg) (Simel 2011).

Normal vital signs change with age, sex, weight, exercise tolerance and overall health. Their measurement is therefore extremely useful in detecting, diagnosing or monitoring medical problems as well as being an essential part of a case presentation. The equipment needed to measure these vital signs is relatively simple, consisting of a thermometer, a sphygmomanometer and a watch, though a stethoscope may sometimes be required for assessing a patient with a very weak pulse. The portability and relatively low cost of these medical instruments means that the measurement of the traditional four vital signs is not constrained to a medical setting, but can be undertaken at home, at the site of a medical emergency, or elsewhere.

In recent times, many health professionals have proposed modernising the traditional group of four vital signs with the addition of more physiological measurements, including: activity (Greenwood et al. 2010; Coleman et al. 2012), blood glucose level (Rendell et al. 2013), emotional distress (Bultz and Carlson 2005), end-tidal CO_2 (Vardi et al. 2000), functional status (Bierman 2001), gait (Studenski et al. 2003), menstrual cycle (American College of Obstetricians and Gynecologists 2006), oxygen saturation (Tierney et al. 1997), pain (Lynch 2001; Davis and Walsh 2004; Mularski et al. 2006), pulse oximetry (Neff 1988; Mower et al. 1997; 1998), respiratory rate (Chen et al. 2009), sleep (Wilson 2005) and urinary continence (Joseph 2003). Some of these suggestions (for example, pain and emotional distress) could be considered as being subjective measurements of symptoms rather than objective measurements of health and can not easily be measured or quantified by a sensor.

Although various additional signs have been proposed, none have been officially adopted. In the past, the cost of the equipment necessary to perform these measurements has been one of the factors that prevented wide scale adoption of some of these measures as vital signs. Recent advances in IT, communication technologies and microfabrication, however, have meant that this is no longer the case. For example, mobile phone technology has brought about a new paradigm in mobile healthcare wherein a smartphone or PDA can now function as a data collection and analysis facility as well as providing digital and audio connectivity to other health or emergency services (through the internet, for example) (van Halteren et al. 2004;

Otto et al. 2006; Chen et al. 2005). Furthermore, advances in battery technology have resulted in new lightweight, smaller batteries with larger power capacity and longer operational lifetimes, which has been instrumental for the development of small, unobtrusive body worn sensors.

Other physiological signs, that traditionally were measured using complicated and bulky equipment, have been developed into smaller, portable systems, sometimes taking advantage of existing technology. For example, advances in microelectronics have resulted in the development of miniaturised, low-power, high performance amplification systems that are essential for the measurement of bioelectrical activity. When combined with the communication and processing capabilities of modern smartphones, these systems can provide point-of-care electrocardiography, electroencephalography and electromyography (Lee et al. 2006; Lin et al. 2004; Gao et al. 2005; Ghasemzadeh et al. 2010; Chi et al. 2010).

As well as improvements in healthcare offered by recent developments in portable, multi-parameter measurement systems, there is also a commercial incentive to invest in these systems, as is evident by the number of patents filed over recent years, a small representative sample of which are cited (Marx 1976; Steffen et al. 1981; Peters 1983; Pfohl 1991; Pross and Schlueter 1994; Money et al. 1999; Saltzstein et al. 1999; Groff and Mulvaney 2000; Jayaraman 2005; Fischell et al. 2007).

3.3 Constraints on System Design

The requirements of sensing technology in the context of remote healthcare present a challenge. The sensor node must provide continuous, accurate monitoring of physiological data while being unobtrusive to the patient. Continuous monitoring means high autonomy for the sensor node, and generally necessitates low power consumption. Continuous monitoring also requires sensor reliability so that the sensing capabilities can be ensured for a long period of time without data artefacts or malfunctioning.

The unobtrusive nature of the sensor requires low form in terms of size and weight. An unobtrusive sensor node is also related to autonomy as battery replacement or node recharging by the patient must be limited. In addition, the sensor node must be easily connectable to a central node, which involves a widespread communication protocol and enough communication range. These two last characteristics are in conflict with the low form requirements such that intelligent use of the available energy is compulsory.

Providing accurate results is also a primary requirement of the sensor node. A lack of accuracy for any sensor node, even one that fulfils the above characteristics, invalidates its applicability to healthcare.

Ideally any sensor or system should be designed with a requirement for minimum intervention by the patient or carer in terms of deployment (e.g. positioning sensors on the body, donning and doffing smart clothing), set-up and maintenance. However, for some types of measurement it is necessary that the sensor is in intimate contact with the patient's skin (e.g. bioelectrical signals and temperature),

which implies a certain level of involvement by the patient or carer and possibly some expertise. It is important that externally worn sensors (e.g. kinematic sensors) are responding to the patient's body movements rather than the movements of their clothing. Hence attachment methods must be secure, but not hinder the movement or interfere with normal activities of daily living (ADL).

The presentation of the sensor or measurement system is decisive in terms of wearability, usability and comfort. Therefore, in order to be a real solution, optimizing the system in these terms to the degree that the measuring principle allows is key. The main presentation forms for wearable sensors and patient monitoring systems are detailed in the following paragraphs.

Stand-alone sensor nodes This is the most used presentation form in commercial ehealth sensors due to the popularity of continuous glucose monitoring (CGM) systems, which are usually presented this way. In this form, the sensor and the communication module are presented in an independent and compact device that is carried by the person, usually next to the location where the sensing takes place. The main advantage of this presentation form is that the sensor does not depend on other external elements and can be easily carried in everyday use.

Wristwatch or wrist strap Similar to the stand alone sensor but in a wristwatch or wrist strap mode, this presentation is more comfortable and discrete than stand-alone solutions, but is not useful for monitoring many physiological signals due to the impossibility of measuring them at the wrist. Activity, galvanic skin response (GSR), blood oxygen level, pulse rate, and temperature are some of the signals that can be measured with devices using this presentation form.

Chest strap This presentation form is similar to the wrist strap in that it is comfortable and discrete, but is not appropriate for monitoring those physiological signals that do not manifest at the chest. It is especially common for heart rate monitoring sensors based on the electrocardiogram (ECG) principle and sensors to measure respiration rate. Kinematic sensors such as accelerometers and gyroscopes are also presented in the form of chest worn devices for the measurement of activity, posture and gait.

Head cap The head cap is a more application specific presentation form, generally associated only with the measurement of brain activity, which naturally requires the placement of electrodes in intimate contact with the skull.

Implantable devices The monitoring of parameters such as those based on electrocardiography (ECG or EKG) is sometimes done using implantable devices. The main disadvantage of this presentation form is that it is necessary to intervene surgically in order to insert the device in the body. Because of that, these devices are used only when the signal monitoring requires it or there are additional reasons for doing so, as in the case of pacemakers. Implantable microelectromechanical systems (MEMS) have been used for controlled drug release (Richards Grayson et al. 2004) and transcutaneous electronic devices are employed for electrical stimulation in patients with paralysis induced by central nervous system lesions

for the recovery of motor and sensory function (Rushton 1997; Hunter Peckham and Knutson 2005; Ragnarsson et al. 1998).

Smart textiles and gloves The sensors in this group are embedded into a glove or a shirt that is worn by the patient or individual. Special electronic inks and pastes, which are flexible and elastic, can be printed on to textiles to form electronic circuits in conjunction with miniaturised MEMS type sensors. Some of these printable pastes possess particular properties (e.g. high temperature coefficient of resistance) or are responsive to particular biological or chemical signals (e.g. water vapour) and therefore directly form the sensing element (e.g. temperature and sweat rate, in these two examples). Hence it is possible to integrate the whole sensor system in a shirt, leading to a very comfortable and aesthetic way of monitoring an individual. The method is especially useful in continuous remote monitoring, although everyday use requires a solution to ensure the cleanliness of the textile.

Patch This presentation form involves an adhesive patch that contains sensors as well as their associated processing circuitry, power sources and communication capabilities. This form is usually associated with the measurement of biophysical signals at the skin surface (e.g. electroencephalography, electrocardiography, electromyography, galvanic skin response) where intimate contact is required (Haahr et al. 2008). It is also used for secure location of kinematic sensors upon the body. Sensor deployment with this form, however, requires the use of non-allergenic materials.

3.4 Sensing Technology for Physiological Signals

There are a large number of different physiological signals that can currently be measured and we have chosen to classify them into categories that are loosely defined by the medium or domain in which the sensed signal manifests or operates, namely: biochemical, electrical, kinematic and physical. Of these, kinematic is a special case since these sensors should be classed as physical, but their use in so many disparate health related application areas merits their separate classification. In our classification system, we recognise the ambiguity in the definitions of biological and chemical sensors with respect to human monitoring. A chemical sensor can be defined as a system that measures changes in the concentrations or activities of the reactants or products in a chemical reaction. Now consider a human bodily function such as metabolism. Clearly the associated measurable parameters are chemical, but the process is a biological one. Adoption of the class identifier "biochemical" is therefore seen as a good compromise to this semantic problem.

Table 3.1 summarises physiological parameters that have been successfully telemonitored, defined as the use of information technology to monitor patients at a distance (Meystre 2005), and categorised according to our classification scheme. The following sections describe the sensing technologies available for the measurement of some of these signals.

Table 3.1 Physiological signs that have been successfully telemonitored. Adapted from Meystre (2005)

Physiological sign	Sensor class	Physiological sign	Sensor class
Basal metabolic rate	Physical	Intracranial pressure	Physical
Blood ethanol	Biochemical	Intrauterine pressure	Physical
Blood glucose	Biochemical	Intravesical pressure	Physical
Blood lactate	Biochemical	O_2 consumption	Physical
Blood pressure	Physical	Pacemaker parameters	Electrical
CO_2 production	Physical	Physical activity	Kinematic
Electrocardiograph (ECG)	Electrical	Posture	Kinematic
Electroencephalogram (EEG)	Electrical	Pulse oximetry	Physical
Electromyograph (EMG)	Electrical	Respiratory rate	Physical
Gait and cadence	Kinematic	Spirometry	Physical
Galvanic skin response (GSR)	Electrical	SpO_2	Physical
Heart rate	Physical	Temperature	Physical
Heart sounds	Physical		

3.4.1 Biochemical Signals

This class of physiological signal includes the measurement of bodily secretions such as sweat, which can indicate activity, fever or other diseases whose symptoms include raised metabolic rate. It also includes the measurement of specific chemicals or compounds that are only required with certain pathologies, e.g. the measurement of glucose for diabetes management.

MEMS biosensors In general, the quantities of products or reactants in a biochemical reaction are usually very small. To accurately measure them therefore requires sensing systems with high sensitivity and selectivity. This favours the use of MEMS where the effects of transduction on micro scale dimensions are greater than on their macro scale counterparts. For example, a common sensing technique in biochemical processes is to measure changes in the masses of products or reactants, and the measurement of a small mass change within a sensing structure due to (say) surface adsorption of a target analyte is far easier with a structure that has a low initial mass (effectively the background signal level) than with a structure of larger initial mass. A further benefit that MEMS structures offer is that being based on silicon facilitates the integration of processing circuitry and communications within the same device, making them ideally suited for remote, autonomous applications. It is therefore of no surprise that MEMS are increasingly being used as transduction structures for biochemical sensors (Bogue 2007).

In order to measure a biochemical parameter such as glucose concentration, the transducer needs to be coated with an active layer (called the receptor layer) that is responsive to, or has an affinity to, the measurand of interest. The function of the receptor layer is to interact with the analyte to be measured and transform that interaction into structural or electrical variations of a MEMS device (Smith 2005). The receptor layer may be biological, consisting of specific enzyme receptors.

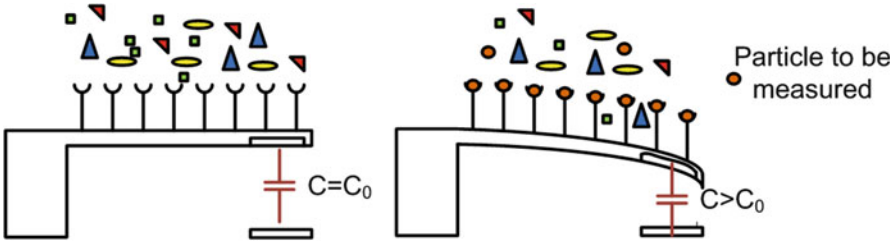


Fig. 3.1 Example of a MEMS biosensor built with a cantilever with affinity receptors

These facilitate biochemical reactions in the presence of a target substrate, which can be the analyte of interest. The enzymatic reaction may simply bind the substrate to the receptor layer or generate a proportional electrical signal in the transducer and therefore the amount of substrate element can be measured. Alternatively the receptor layer may be chemical, utilising affinity receptors which are based on specific chemical binding between the layer and the analyte to be detected.

The operation of a MEMS biosensor based on the mass change principal is illustrated in Fig. 3.1 In this case the superior part of a cantilever structure is coated with an affinity type receptor layer with the capability to bind to a certain particle. When this particle is not present around the cantilever, the bindings are empty. But as the concentration starts rising, the bindings start filling. The increasing mass of these stacked particles on the cantilever causes it to bend and reduce the distance between the metal plates located at its tip and the wafer substrate, leading to an increased capacitance value between both. Thus, the particle concentration in the cantilever's surroundings can be measured by measuring the change in capacitance. Alternatively, the change in strain caused by the cantilever deflection can be measured by a piezoresistive element located at the base of the cantilever, or if the cantilever is fabricated as a resonant structure, the change in mass can be detected as a proportional change in resonant frequency (temperature insensitive).

In addition to MEMS systems that operate by measuring changes in resistance or capacitance, other operating principles are being used in ehealth applications in recent years. For example, if the receptor layer was formed on the gate electrode of a field effect transistor, the accumulation of charge at the electrode as the target analyte binds to its surface will causes a change in the channel resistance between the source and drain of the transistor, which in turn results in a change in the measurable current.

It should be noted that the MEMS biosensors as described above function in accumulation mode only, meaning that they can only measure increases in the target species. This is acceptable for a single shot, disposable sensor. A method of removing adsorbed species must therefore be include in order to make the sensor reversible for continuous use and to be able to measure decreasing concentrations of the target species. Typically this may be achieved by including a heater element under the receptor layer.

Sweat rate The sweat rate can be defined as the flow of water vapour emitted by the skin. The interest in measuring sweat rate comes from the fact that a number of pathologies as well as physiological conditions may result in increased or decreased sweat rates. Therefore, being able to continuously measure this parameter may lead to important applications in remote healthcare and disease prevention (Salvo et al. 2010). However, diagnosis based solely on sweat rate should not be considered unless other environmental factors such as temperature and humidity are considered, since these affect sweat production. Similarly, the level of physical activity being undertaken at the time of a sweat rate measurement should also be considered as this too affects metabolism.

Several prototypes that take a different approach have also been presented (Salvo et al. 2010; Solar et al. 2013). In this case, the sweat rate is measured by means of a humidity sensor, which in general measures the sweat by variation in capacitance. The humidity sensor can be implemented as a smart textile sensor (Salvo et al. 2010) or even a commercial humidity sensor, with the advantage of not having to place the sensor in direct contact with the skin (Salvo et al. 2010; Solar et al. 2013).

Glucose Although there are many solutions for monitoring blood glucose concentration based on manually testing the level of glucose several times a day with finger prick tests, from the point of view of ehealth, continuous monitoring systems are more interesting and challenging. In addition, a patient's quality of life can be greatly improved by using an autonomous feedback system with a small insulin pump to regulate the insulin delivery based on the measured glucose levels. Such systems require the patient's glycemia to be measured accurately and continuously so that the insulin infusion rate and dosage can be correctly adjusted. Currently, commercial methods for continuous glucose monitoring are based on measuring the glucose concentration in the interstitial fluid with an external impedance spectroscopy device. From the ehealth perspective, interstitial fluid glucose concentration sensors are the most suitable as the measurements are done automatically and without any user interaction. In order to do this, a minimally invasive subcutaneous sensor, usually located in the hip, is used. The apparatus reaches the interstitial fluid using a needle or a small catheter and uses an enzyme that catalyzes the oxide reduction of glucose to trigger a reaction that ends with the generation of a small electrical current that is proportional to the glucose concentration. Devices presented in (Ichimori et al. 2006; http://www.medtronicdiabetes.com/products/paradigmrevelsystem; http://www.dexcom.com/dexcom-g4-platinum; http://www.abbottdiabetescare.co.uk/your-products/freestyle-navigator) use this principle and even include the support for transmitting the measured value to the hospital.

The problem with the systems mentioned above is that they are invasive and are therefore susceptible to the health risks associated with implanting foreign objects within the body. Advanced technologies for glucose monitoring therefore tend to focus on injection-free, transcutaneous measurements (Ferrante do Amaral and Wolf 2008; Girardin et al. 2009; English 2012). Of the various operating principles, the ones based on the specific Raman spectrum of the glucose molecules (http://www.c8medisensors.com) seem to be the most promising. Here the sensor

generates and sends a specific light beam to the skin and receives and processes the reflected light with a CCD sensor; once the data is processed the sensor identifies the amount of glucose. Because the measured value is transmitted through Bluetooth to an ordinary smartphone in the device described by (http://www.c8medisensors.com), it is possible to easily integrate the sensor into a remote healthcare system. The main disadvantage of these devices is that although they have been fully developed, currently they are only approved for research purposes.

3.4.2 Electrical Signals

The human body generates a wide range of electrical signals that vary with activity, environmental conditions and state of health. Most of these signals arise through the firing of neurons and associated electrical signal pathways, which respond to stimuli and control the body's physiological responses. These electrical signals can be detected at the surface of the body and the measurements are classed according to their functional source. Thus, the electrocardiogram (ECG or EKG) is a measure of the heart's electrical activity; the electroencephalogram (EEG) arises from the firing of neurons in the brain; and the electromyogram (EMG) results from muscular activity (Prance 2011). These measurements can be performed subcutaneously and, for clinical diagnosis, provide the most accurate, spatially resolved measurements. However, their invasive character means that they are unsuitable for extended monitoring, particularly in non-clinical and unsupervised settings such as the home (Hogrel 2005; Stegeman et al. 2000). In the context of long term or continuous health monitoring, it is necessary for measurement techniques to be non-invasive and straightforward for non-specialist personnel to undertake. This dictates, among other things, that the signals be acquired from the surface of the body, or off-body.

ECG Electrocardiography is a test that measures the electrical activity of the heart, producing a time varying waveform chart known as an electrocardiogram. The heart produces tiny, rhythmic electrical impulses which spread through the heart muscle to control the heart contraction. These impulses can be detected by small metal electrodes stuck on to the arms, legs and chest, and are recorded as a function of time by the ECG machine at each heartbeat. The electrodes on the different parts of the body detect electrical impulses coming from different directions within the heart. Because there are recognised, normal patterns for each electrode, various heart disorders can be diagnosed by analysing these patterns, including: abnormal heart rhythm, myocardial infarction and an enlarged heart.

This traditional method of performing ECG relies on the electrodes being held securely in intimate contact with the patient's skin (achieved with the use of conductive gels), and because of the locations of the electrodes, requires long leads to connect the electrodes to the amplification and signal recording system. Both these requirements tend to restrict or prevent patient movement, which is clearly undesirable of a portable, remote monitoring system. Non-contact, wireless

ECG sensors based on the principle of capacitive coupling are now being developed to overcome these problems (Peters 1983; Heimann et al. 2009). Furthermore, textile-based or knitted electrodes that are washable and fully integrated within clothing and wearable accessories are also being developed which may also overcome the requirements of accurate sensor placement (Paradiso et al. 2005a; Di Rienzo et al. 2010).

EEG Electroencephalography is the process of recording the brain's electrical activity using a number of electrodes placed in intimate contact with the scalp at particular locations. An EEG is mainly used to diagnose and manage epilepsy, though it can also be used to investigate other conditions that affect brain function, including: infections such as encephalitis, haemorrhages, head trauma, dementia and tumours. EEGs can also be used to diagnose and manage sleep disorders such as insomnia. In the context of mobile healthcare, EEG can also benefit from the advances in ECG electrode technology, particularly with regard to wireless connectivity (Casson et al. 2008).

EMG Electromyography is a test that measures the electrical activity of muscles and is often performed with another test that measures the conducting function of nerves. In some medical conditions the electrical activity of the muscles or nerves is not normal. Finding and describing these electrical properties in the muscle or nerve can help diagnose the patient's condition. EMG is measured using electrodes placed on the skin above the muscle being monitored. It is also therefore likely to benefit from the advances made in ECG and EEG electrode technologies. Scilingo et al. report, for example, on conductive fabric electrodes formed from stainless steel threads twisted around a standard viscose textile yarn used in EMG and ECG, and compare their performance against standard metal electrodes (Scilingo et al. 2005). Wearable EMG has also been demonstrated in gait analysis as part of a rehabilitation program (Freed et al. 2011) and in the measurement of infant's grasping behaviour over day long time scales in the home environment (Nishida et al. 2006).

Galvanic skin response In general, the available commercial products take advantage of the variability in skin conductance due to the effect of condensed sweat on the skin. This is referred to as the galvanic skin response (GSR) and it means that the skin itself is the sensor transducer. In this technique two electrodes are placed on the skin a fixed distance apart and the electrical impedance between them is measured. There is no accepted standard with respect to electrode placement sites, although the palm of the hand or two fingers are the most popular sites (Ming-Zher et al. 2010; http://bio-medical.com/products/sensors/gsr-sensors.html). From a long term monitoring perspective, however, these sites are highly susceptible to motion artefacts and may be uncomfortable or hindering during normal daily activities.

3.4.3 Kinematic Signals

Kinematic signals include acceleration, velocity, distance, rotation, rate of rotation, angle and time. These measurements are usually taken to determine how an entity moves (e.g. tracking the position of a limb segment in human motion studies) or to

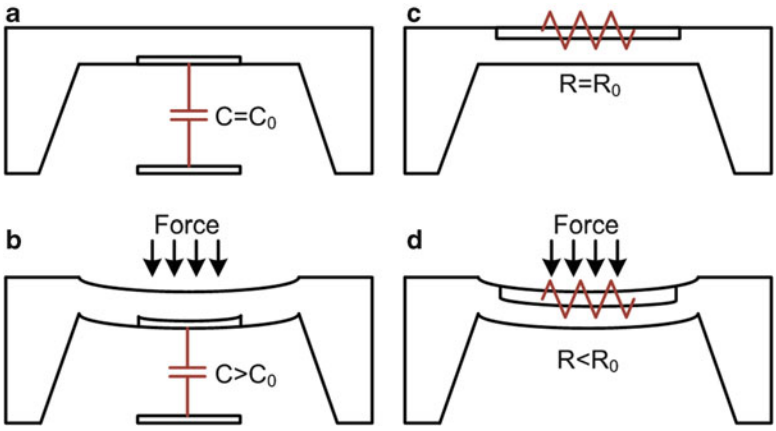


Fig. 3.2 Examples of MEMS pressure sensor operation principle

determine relative positional information (e.g. the angle of flexion of a limb joint). This class of signal is used in a wide range of health indicators including activity, balance, cadence, gait, posture, spasticity and tremor, some of which are subjective and rely on self-reporting. Many of these signals can be measured reliably as part of a mobile health system using kinematic sensors (accelerometers, gyroscopes and magnetometers) thereby removing subjectivity from symptomatic data.

Kinematic sensors are generally based on the transduction of a physical parameter to an electrical signal that is easier for the processing module to handle. In ehealth applications, where size and power consumption are key aspects, transducers based on MEMS are the most interesting (Bogue 2007). In these transducers, the electro-mechanical characteristics of the silicon are used to generate resistive or capacitive variations when the microstructure of the transducer is modified due to external forces such as acceleration, compression, pressure, temperature, etc.

Typical examples of MEMS capacitive and resistive transducers are shown in Fig. 3.2. Figure 3.2a, b show the effect of an external force on the MEMS structure, changing the distance between two metallic plates and therefore, the capacitance between them. Hence, the capacitance value is directly related to the force. By comparison, Fig. 3.2c, d show the change in the geometry of a piezoresistive material caused by the deformation of the MEMS structure due to the external force. The geometric variations in the material cause changes in its resistivity, which is directly related to the magnitude of the external force. These two basic operating principles are used with different topologies to optimize the sensitivity of the transducer toward different physical parameters and for different measurement ranges.

Accelerometers A MEMS accelerometer measures the effect of acceleration rather than acceleration directly, by making use of Newton's second law of motion ($F = ma$) which relates the acceleration (a) experienced by an object of mass (m) to the net force (F) acting on it. If the mass of the inertial element of a MEMS structure

is known, then the acceleration can be computed since the force can be measured by a MEMS structure as depicted in Fig. 3.2. Usually the mass does not need to be known since this can be calibrated out, and so the output signal from a calibrated MEMS accelerometer is directly proportional to the acceleration that it experiences.

A point to be aware of is that the accelerometer responds to the net force acting on it, which at all times includes a contribution from gravity (often referred to by the oxymoron "static acceleration"). Hence, when calculating kinematic signals such as joint angle, the output from the accelerometer must be adjusted to compensate for gravity before it can be processed (Lötters et al. 1998; Luinge and Veltink 2004). This is not necessarily a drawback, however. In the absence of any external acceleration, the accelerometer only responds to the force of gravity, which means that its orientation can be established when it is stationary (AN3182 Application Note 2010; Lee et al. 2010). But this also means that accelerometers are orientation sensitive and so care must be taken when re-positioning them on a patient during long term monitoring, unless provisions are made to counter this (Jiang et al. 2011). When correctly compensated, accelerometer data can be used in the home environment for the detection of posture, motion or activity, among others (Veltink et al. 1996; Foerster et al. 1999; Mathie et al. 2004).

Acceleration is not, however, the only parameter that can be measured with an accelerometer. Because acceleration is the first derivative of velocity and the second derivative of distance with respect to time, these two additional parameters can be determined by, respectively, a single or double integration of the acceleration data. To do this properly, however, requires prior knowledge of any integration constants, which usually implies required knowledge of boundary conditions. Care must also be taken when using these parameters in conjunction with the 'suvat' equations to establish, for example, movement trajectories or to reconstruct dynamic spatial co-ordinates, since these equations are only valid under the condition of constant acceleration. In these circumstances, computational algorithms need to be employed to predict the location of the sensor in space (Giansanti et al. 2003; Zhang et al. 2005; Yun and Bachmann 2006). The working range of an accelerometer is usually expressed in values of g, where g is the acceleration due to gravity. Consideration must therefore be given when selecting an appropriate accelerometer to operate within a particular measurement range; the omnipresent influence of gravity imposes a minimum working range of ± 1 g.

Gyroscopes When used in ehealth applications, gyroscopes are usually MEMS type devices because of their small size (unobtrusive), small mass and low power requirements. MEMS gyroscopes, or more correctly MEMS angular rate sensors, are used whenever rate of rotation (°/s) sensing is required without a fixed point of reference. This separates gyroscopes from any other means of measuring rotation, such as a tachometer or potentiometer. They are sometimes used in conjunction with accelerometers, with the combined pair being referred to as an *Inertial Measurement Unit* (IMU). In this scenario, the gyroscope is used not only to give a measure of angular rotation, but it also provides compensation to the accelerometer which demonstrates a dependence on orientation (Mayagoitia et al. 2002). The use of

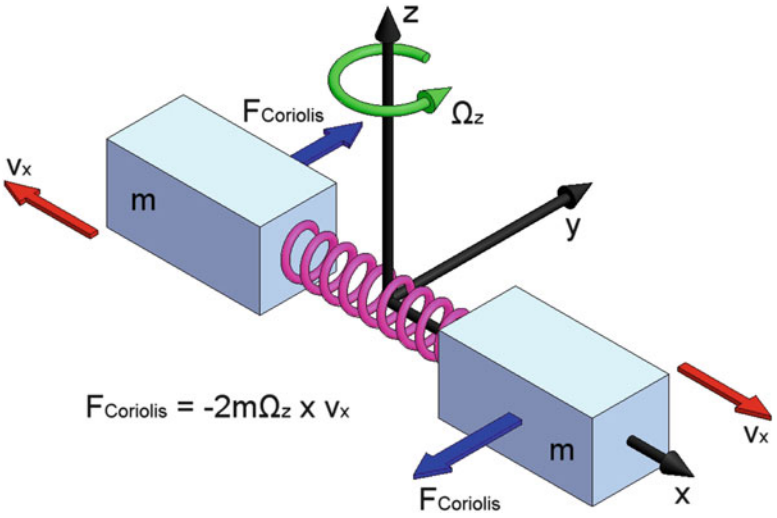


Fig. 3.3 Principle of operation for MEMS vibrating gyroscope. The x symbol in the equation for the Coriolis force is the cross product operator

gyroscopes and IMUs in ehealth is dominated by human motion studies (Tong and Granat 1999; Najafi et al. 2003; Luinge and Veltink 2005; Zhou et al. 2008; Hyde et al. 2008) and home based rehabilitation (Jovanov et al. 2005; Patel et al. 2010; Fulk and Sazonov 2011; Saywell et al. 2012).

The most common form of MEMS gyroscope is based on the tuning fork principal and contains a pair of identical masses (m) that are driven to oscillate with equal amplitude but in opposite directions, as shown in Fig. 3.3. If one of the masses is moving in the positive *x*-axis direction with velocity v_x and an angular rotation Ω_z is applied about the *z*-axis, then the mass will experience a force in the direction of the arrow shown as a consequence of the Coriolis force. Correspondingly, the other mass moving in the negative *x*-axis direction with the same velocity will also experience a Coriolis force of the same magnitude, but acting in the opposite direction. These two forces can be measured by other sensing mechanisms built into the MEMS structure (e.g. strain change measured with piezoresistor or deflection measured with capacitance, see Fig. 3.2).

Magnetometers A magnetometer is a device that responds to the strength and direction of a magnetic field and is therefore at all times influenced by the Earth's magnetic field. Although the strength, direction and inclination of the Earth's magnetic field varies over the surface of the Earth, it can be considered to be approximately constant within a geographical area the size of (say) a small town, and definitely constant within an area the size of a home. Because of this, the local magnetic field vector can be considered fixed in direction and magnitude and therefore a magnetometer's response in this location depends only on the orientation of the magnetometer with respect to the local field vector. This means that a magnetometer can be used as an orientation sensor.

There are many techniques used in magnetic field sensing, including: Hall Effect sensor, magneto-transistor, magneto-optical sensor, Lorentz force based MEMS sensor, search coil magnetic field sensor and SQUID magnetometer, amongst others. Of these, the MEMS-based magnetic field sensor offers a small-size solution for magnetic field sensing that does not involve the microfabrication of magnetic material and can therefore be produced at low cost. The Lorentz force based MEMS sensor relies on the mechanical motion of the MEMS structure due to the Lorentz force acting on a current-carrying conductor in the magnetic field. The mechanical motion of the micro-structure is sensed either electronically (see Fig. 3.2) or optically (e.g. by interferometry with laser or LED sources).

Although magnetometers have been used in biomechanical measurements (O'Donovan et al. 2007), they are not extensively used in ehealth applications. This is because they use the Earth's magnetic field as a reference which can be distorted by the presence of ferromagnetic materials (Kendell and Lemaire 2009). Thus they are not appropriate for telemonitoring of patients who require assisted technology for mobility (e.g. wheelchair with steel frame) or within the normal home environment where large ferrous structures are likely to be present (e.g. in the kitchen).

3.4.4 Physical Signals

In this context, a biophysical signal refers to any human signal that can be measured using a physical principal. This covers the traditional four "vital signs" of blood pressure, body temperature, pulse rate and respiration rate. Table 3.1 shows that there are a number of biophysical signs that have been telemonitored, but some of these (e.g. intercranial pressure) are only of use in very specific circumstances or are only applicable for use within a small range of pathologies. Sensor technologies that have been developed for some of the more commonly measured biophysical signals and that are appropriate for integration within ehealth systems are described in the following sections.

Blood pressure Currently, most blood pressure sensors are based on a digital or analogue sphygmomanometer, where a column of mercury is linked to a cuff around the arm and the height of the mercury column, in millimetres, indicates the blood pressure. The technology of these blood pressure sensors has advanced enough that users are notified by an alarm when the measurements are out of the normal range. They can be connected to the PC through a USB cable (http://www.medroninc.com) and they can even be connected via Wi-Fi to smartphones (http://www.withings.com/en/bloodpressuremonitor), making them easy to integrate in ehealth scenarios.

With the objective of improving these measurements and making them automatic, some research projects, such as (Asada 2009), are being developed based on the pulse wave velocity method, which allows blood pressure to be calculated by measuring the pulse at two points along a major artery. Alternatively, for people with severe vascular disease who have undergone surgery, implantable pressure sensors are available to control the pressure in a certain blood vessel or stent and

monitor the disease's evolution (Abraham et al. 2011). These sensors include a wireless communication protocol and can also be integrated in ehealth scenarios.

Body temperature The evaluation of body temperature is a routine diagnostic assessment method and an important sign of health and illness. However, body temperature is a complex physiological sign that is subject to many other variables such as age, diurnal variation, physical activity, seasonal changes, or the time of year. In addition, measuring body temperature will yield different results depending on where it is taken, although the most widespread and less invasive methods are rectal, oral, in the ear and under the armpit (axillary) (Sund-Levander et al. 2002; 2004). The temperature sensor nodes focus on non-invasive solutions and therefore it is actually the skin temperature that is measured. This can be a drawback in terms of accuracy, especially if none of the previous locations is chosen for temperature monitoring. However, the main advantages of temperature sensor nodes are the ability to monitor continuously and thus have data on temperature variability readily available. Valuable information can be obtained from those measurements if correlated with other sensors. For example, when combined with accelerometers, an ECG or the sweating rate, accurate activity and energy expenditure data can be obtained (Solar et al. 2013; Berntsen et al. 2010).

One of the main sensing techniques for temperature measurement implemented in commercial solutions is the use of thermocouples and thermistors (Pandian et al. 2008; Smith et al. 2010). The thermistor is a device whose resistance is directly correlated with temperature. The variation in resistance is converted into an electrical signal, typically using a Wheatstone bridge circuit, and further digitized with an analogue-to-digital converter (ADC). A thermocouple consists of two different materials, usually metal composites, which are connected at one end. If there is a temperature difference between the connected and non-connected ends, a thermoelectric voltage is produced whose magnitude is directly proportional to the temperature difference. In modern solutions, several thermocouples are connected in series in order to increase the voltage generated and to make the system more sensitive. The resulting system is defined as a thermopile. Modern thermopile sensors are small and fast, thanks to modern semiconductor technologies, so they are implemented in solutions like infrared thermometers which can be also found in wearable solutions for remote healthcare (Anliker et al. 2004; Johnstone et al. 2012).

In addition, fully integrated temperature sensors based on standard semiconductor technologies are also available. As their main objective is ultra low power consumption solutions, these sensors allow the implementation of tiny wireless sensor nodes, even with full passive RFID technologies, of very small sizes and high accuracies (Fernandez et al. 2012).

Heart rate The most common method for heart rate definition is based on the ECG, and there are many commercial devices that include these sensors and wireless communication protocols for integrating them in health systems, such as (http://www.corventis.com/us/nuvant.asp; http://www.preventice.com/products/bodyguardian; http://www.exmovere.com/pdf/Exmovere_Wearable_Sensor_Research.pdf; http://www.shimmer-

research.com/r-d). Pacemakers can also be considered implantable heart rate systems as they are performing continuous heart monitoring (http://www.medtronic.com/patients/bradycardia/device/index.htm). In addition, in the case where there is an oxygen saturation sensor, it is also possible to use the same device to simultaneously measure the heart rate using the same light refraction principle, as is done in (http://www.mir.abmedic.com/miroxi.htm; http://www.sensixa.com; http://www.lifewatch.com/wirelesshealthcare). Finally, most blood pressure sensor devices also have the ability to measure heart rate.

Oxygen saturation Also referred to as SpO_2 (saturation of peripheral oxygen), oxygen saturation is defined as the ratio of oxyhaemoglobin to the total concentration of haemoglobin present in the blood and is generally measured using a pulse oximeter. Blood oxygen saturation therefore measures the amount of oxygen carried in the blood. The normal reading for a healthy individual (non-smoker) should be between 96 and 99 %. It could be argued that these sensors are biochemical because they aim to detect a chemical species, but they have been classified here as a biophysical sensor since in general their measurement principal relies on a physical effect—optical absorption.

The most popular method for non-invasive estimation of SpO_2 is by means of photoplethysmography (Allen 2007). The method is based on measuring changes in the intensity of light reflected by the pulsatile arteriolar blood flow. An area of skin on the finger or ear, where there are superficial capillary blood vessels, is illuminated by monochromatic light emitted by LEDs in the range of 600–1,400 nm. This range of wavelengths includes those which are absorbed by haemoglobin and so depending on the amount of haemoglobin present in the bloodstream, the intensity of the reflected light coming from the tissue changes. By comparing the intensities of the transmitted and reflected light, the haemoglobin concentration, oxygen saturation and heart rate can be calculated. The devices presented in (http://www.lifewatch.com/wirelesshealthcare) are based on this principle and even include the Bluetooth protocol making their integration into remote monitoring systems much simpler. The main disadvantage of these sensors is that placing the transducer on the finger or the ear is not very comfortable or aesthetic. To overcome this problem, recent developments have focused on wearable devices (Asada et al. 2003; Izzetoglu et al. 2005).

Respiratory rate The respiration rate refers to the number of breaths a person takes per unit time. It is based on the measurements of the number of inspiration/expiration cycles per unit time. The aim of measuring a patient's respiratory rate is to determine whether the respirations are normal, abnormally fast (tachypnea), abnormally slow (bradypnea), or nonexistent (apnea).

The sensing techniques for respiration rate are varied. Some recent solutions use a smart textile with embedded piezoelectric or piezoresistive transducers. During breathing, the piezoelectric or piezoresistive sensor is subjected to stress due to the expansion of the thorax or the abdomen, thus providing information on the breath rate through a change in sensor resistance (piezoresistive sensor) or the variation of an electrical signal (piezoelectric sensor) (Lötters et al. 1998; Tura et al. 2003; Lymberis and Paradiso 2008; Paradiso et al. 2005b; Rovira et al. 2011; Taccini et al. 2008).

Other solutions make use of capacitive pressure sensors to detect the circumference expansion and contraction of the torso (Johnstone et al. 2012). Other solutions make use of electrodes (Mundt et al. 2005). Finally, respiratory inductance plethysmography (RIP) uses the magnetic fields generated by a weak alternating current flowing through a loop around the chest (Derchak 2011). An elastic belt with conductive wire is placed around the chest or abdomen. Breathing changes the shape of the magnetic field generated by the belt, which modifies the impedance of the wire that can be measured, for example, as a change in the frequency of the alternating current.

Spirometry By comparison to respiration, which is a measure of breathing rate, spirometry refers to the measurement of the volume of air that a patient inhales and exhales as a function of time. It essentially relies on flow rate sensing and its importance lies in the fact that spirometry tests assess the patient's condition in diseases like asthma or fibrosis.

The flow sensing mechanism is mainly based on one of two different concepts. The first is related to the measurement of differential pressure and the sensor is commonly referred to as a pneumotach. The measurement of the differential pressure takes place in a tube in which a pressure transducer is inserted that incorporates a resistive/capacitive electrical element. The air flow will provide a pressure drop across the transducer that is proportional to the flow rate. Depending on the flow direction, the differential pressure across the transducer is positive (expiration) or negative (inspiration). The pressure difference experienced by the electrical element within the pressure transducer is converted into an electrical and/or digital signal (Schlegelmilch and Kramme 2012).

The second method is an optical one that makes use of a turbine transducer built in a flow tube. Driven by the passing gas flow, the rotating elements interrupt or reflect the light from a light-emitting diode (LED). Photodiodes register the rotations, returning an electrical impulse frequency proportional to flow, while the total count is proportional to volume (Schlegelmilch and Kramme 2012).

Table 3.2 summarises examples of some commercial systems that employ the sensing technologies described here, indicating their presentation and communication standard. Some of these systems are described in more detail in the following sections.

3.5 Application Areas

An examination of the available scientific literature reveals that mobile healthcare technology is being used in application areas beyond traditional medicine. For example, the personal healthcare market is responding to the requirements of individuals who are becoming increasingly more health conscious and are providing a range of home based, self-diagnosis systems, ranging from blood pressure monitoring equipment to body mass index measurement systems. Sports science and leisure have also influenced the growth of remote healthcare technologies, driving the development of a plethora of activity and fitness monitoring devices (Armstrong 2007). Fundamental research into the biomechanics of human movement has also benefited from the miniaturisation of

Table 3.2 Features of commercially available physiological sensor solutions

Measurement	Technology	Presentation	Communication	Ref
Glucose	Interstitial fluid	Stand-alone	916.5/868.35	http://www.medtronicdiabetes.com/products/paradigmrevelsystem
Glucose	Interstitial fluid	Stand-alone	Wireless	http://www.dexcom.com/dexcom-g4-platinum
Glucose	Interstitial fluid	Stand-alone	Bluetooth	http://www.abbottdiabetescare.co.uk/your-products/freestyle-navigator
Glucose	Raman spectrum	Stand-alone	Bluetooth	http://www.c8medisensors.com
BP	Sphygmomanometer	Stand-alone	Wi-Fi	http://www.withings.com/en/bloodpressuremonitor
BP	Sphygmomanometer	Stand-alone	Bluetooth	http://www.lifewatch.com/wirelesshealthcare
BP	Sphygmomanometer	Stand-alone	Bluetooth	http://www.iem.de/stabil_o_graph_mobil2?_lang=1
BP	Pressure sensor	Implantable	Proprietary, LF	http://www.cardiomems.com/content.asp?display=aboutus&view=about%20us%20revised
Heart rate	ECG	Standalone	zLink	http://www.corventis.com/us/nuvant.asp
Heart rate	ECG	Chest strap	Bluetooth	http://www.preventice.com/products/bodyguardian
Heart rate	Light reflection	Wrist worn	GSM and ZigBee	http://www.exmovere.com/pdf/Exmovere_Wearable_Sensor_Research.pdf
Heart rate	ECG	Body worn	Bluetooth	http://www.shimmer-research.com/r-d
Heart rate	Light reflection	Finger	Bluetooth	http://www.mir.abmedic.com/miroxi.htm
Heart rate	Light reflection	Ear-mounted	Bluetooth	http://www.sensixa.com
Heart rate	Light reflection	Finger	Bluetooth	http://www.lifewatch.com/wirelesshealthcare
Heart rate	ECG	Stand-alone	Bluetooth	http://www.alivecor.com/
Heart rate	ECG	Chest strap	ISM 880/921	http://www.zephyranywhere.com/products/bioharness-3/
Heart rate	ECG	Wristwatch	2.4 GHz	http://www.suunto.com/en-GB/sports-watch-collections/Suunto-M-Series-Collection/

Heart rate	ECG	Implantable	Proprietary 401/406	http://www.medtronic.co.uk/index.htm
SpO_2	Light reflection	Finger	Bluetooth	http://www.mir.abmedic.com/miroxi.htm
SpO_2	Light reflection	Ear-mounted	Bluetooth	http://www.sensixa.com
SpO_2	Light reflection	Finger	Bluetooth	http://www.lifewatch.com/wirelesshealthcare
Respiration	Electrode	Stand-alone	Bluetooth	http://www.preventice.com/products/bodyguardian
Respiration	Piezoresistive	Smart textile	Bluetooth	Derchak (2011)
Respiration	Capacitive	Chest strap	ISM 880/921	http://www.zephyranywhere.com/products/bioharness-3/)
Respiration	Electrode	Stand-alone	Proprietary zLink	http://www.corventis.com/US/avivo.asp
Respiration	Plethysmography	Stand-alone	Bluetooth	http://www.thoughttechnology.com/proinf.htm; http://www.thoughttechnology.com/teleinfinit.htm; http://www.thoughttechnology.com/sensorsnew.htm
Temperature	Infrared	Chest strap	ISM 880/921	http://www.zephyranywhere.com/products/bioharness-3/
Temperature	Thermistor	Arm strap	2.4 GHz wireless	http://sensewear.bodymedia.com/
Temperature	Thermistor	Stand-alone	Bluetooth	http://www.thoughttechnology.com/proinf.htm; http://www.thoughttechnology.com/teleinfinit.htm; http://www.thoughttechnology.com/sensorsnew.htm
Spirometry	Turbine/optical	Stand-alone	Bluetooth	http://www.spirometry.com/ENG/Products/spirobank2.asp
Spirometry	Turbine/optical	Stand-alone	Bluetooth	http://sdidiagnostics.com/spirometers/astra300.php
Sweat rate	GSR	Body worn	Bluetooth	http://www.shimmer-research.com/r-d
Sweat rate	GSR	Chest strap	ISM 880/921	http://www.zephyranywhere.com/products/bioharness-3/
Sweat rate	GSR	Arm strap	2.4 GHz	http://sensewear.bodymedia.com/
Sweat rate	GSR	Stand-alone	Bluetooth	http://www.thoughttechnology.com/proinf.htm; http://www.thoughttechnology.com/teleinfinit.htm; http://www.thoughttechnology.com/sensorsnew.htm

wireless, inertial measurement systems (Mayagoitia et al. 2002; Tong and Granat 1999; Patel et al. 2010).

It is therefore appropriate to categorise current, available physiological measurement systems into application areas or themes. These can be loosely classified into the following categories: vital parameter monitoring, rehabilitation and home-based monitoring, sports science and leisure, personal healthcare and well being, and human activity monitoring. Many of the systems and devices described here, however, are not restricted to the application area under which they have been categorised.

3.5.1 Vital Parameter Monitoring

The Sotera ViSi Mobile Patient Safety System is a wireless platform for comprehensive vital signs monitoring through added body-worn sensors (Fig. 3.4). The system enables accurate, continuous monitoring of all core vital signs (heart rate, pulse rate, breath rate, ECG, SpO_2, blood pressure), as well as patient activity and posture (http://www.visimobile.com).

In 2010, Di Rienzo et al. (2010) worked within the MagIC project, aiming to measure vital signs via a textile-based unobtrusive system. They produced a vest incorporating a number of sensors for the monitoring of ECG, respiration rate, skin surface temperature and activity, with the objective of identifying particular cardiac conditions such as atrial and ventricular ectopic beats as well as atrial fibrillation. The system also included signal conditioning and pre-processing circuitry as well as Bluetooth communication for transmitting data to a local PC or PDA.

The Indian Defence Bioengineering and Electromedical Laboratory (DEBEL) of Bangalore realized a smart system, called Smart Vest (Pandian et al. 2008), aiming to monitor ECG, photoplethysmogram, heart rate, body temperature, blood pressure and galvanic skin response. A remote monitoring station receives the data acquired by the sensors in real-time, and reliability of the system seems to be successfully achieved. Chung et al. (Chung et al. 2008) developed the u-Healthcare system, that monitors ECG, blood pressure, SpO_2 and position (by accelerometers) in order to develop more patient-centred medical care systems. Sensors are placed into a chest belt and connected to a mobile phone via the IEEE 802.15.04 communication standard.

The AliveCor Heart Monitor is a clinical-quality, single-channel ECG recorder (Fig. 3.5) (http://www.alivecor.com/). It is cleared by the FDA for sale in the U.S. to licensed medical professionals and prescribed patients to record, display, store, and transfer single-channel ECG rhythms. The AliveCor Heart Monitor is a Class II medical device to be used with the iPhone model 4, 4S or 5. It wirelessly communicates with a dedicated application on the phone.

Another important tool was developed by Loew et al.: the BASUMA body worn system, developed to monitor several important physiological parameters in patients with chronic illness, including ECG, thoracic impedance, body temperature, respiration rate, lung function, breath condensate, blood pressure, pulse rate, and SpO_2 (Loew et al. 2007). All the sensors employed within this platform are equipped with the ZigBee compliant Philips AquisGrain platform.

Fig. 3.4 The Sotera wireless ViSi mobile device (http://www.visimobile.com)

Fig. 3.5 AliveCor heart monitor (http://www.alivecor.com/)

3.5.2 Rehabilitation and Home-Based Monitoring

In 2002, the Swiss Federal Institute of Technology (ETH) in Zürich, developed a wearable multi-parameter medical monitoring and alert system, AMON (Lukowicz et al. 2002), funded by the European FP5 IST program. The wrist-worn device comprises sensors to measure skin temperature, blood pressure, blood oxygen saturation (SpO_2) and includes a 1-lead-ECG. The device also contains a two-axis accelerometer as well as a GSM-based communication link, allowing data to be transmitted to a supervision centre for analysis and comparison. The aim of the AMON system is to continually monitor patients with severe cardiac or respiratory conditions in order to aid them to live independently. Risk alerts and other information can also be triggered and collected.

In 2005, the MIT Media Laboratory in Boston, reported on the development of LiveNet (Sung et al. 2005): a rehabilitation monitoring system based on a mobile wearable platform that comprises a Personal Digital Assistant (PDA), a software network and resource discovery application program interface (API), and a real-time machine that learns inference architecture. Developed to monitor health status, it consists of several sensors to measure the vital physiological signs of ECG, EMG and galvanic skin response, along with a 3-axis accelerometer to measure activity.

The communication between the applications is obtained by a three layer software architecture, which permits signal processing and development of real-time classifiers for wearable applications.

The measurement of several parameters, such as phonocardiography, ECG and body temperature was performed by a smart portable system developed in Taiwan (Chang and Tai 2005). Communication and reception of signals is performed via Bluetooth. Phonocardiography (PCG) is detected by a microphone, consisting of a capacitor, placed on a stethoscope, while ECG measurement is performed with a 3-lead system. A measuring board, including a CPU, a Bluetooth transmitter/receiver, an A/D module and a PDA, this latter controlling the system overall and having an external memory unit, is also present.

Vuorela et al. (Vuorela et al. 2010), of the University of Technology, Tampere, Finland, developed a portable physiological signal recorder in 2010 intended for continuous health monitoring. The system implemented measures bio-impedance (to detect respiration), ECG and user's activity and includes a bio-impedance block, an ECG block, an accelerometer, as well as a control and storage device, a radio communication interface, a rechargeable battery for long-term recording and a user-interface. The ECG system is operational with both custom-made textile electrodes and with conventional silver-silver chloride (Ag/AgCl) gel paste electrodes.

The Mermoth (Luprano 2006) project is a European Commission funded project aimed to develop smart fabrics and interactive textiles. The garment realised includes conductive and electrostrictive fabrics, as well as yarn and dry electrodes, and off-the-shelf accelerometers. These systems are able to measure ECG, respiratory inductance plethysmography, skin temperature and activity. A microcontroller, connected to a PDA, is used as an interface for the sensors, providing a link to a local PC for displaying and assessing data.

Smart pyjamas have been developed for infant monitoring during sleep as an aid in helping to prevent cot deaths. Using the European Space Agency's Technology Transfer Programme, the Mamagoose paediatric pyjamas exploit technology developed in two space applications: the analogue biomechanics recorder equipment and the respiratory inductive plethysmograph suit (http://www.esa.int/Our_Activities/Technology/TTP2/New_pyjamas_could_prevent_cot_deaths; http://www.esa.int/esaKIDSen/SEML8LXJD1E_Technology_1.html). The Mamagoose pyjamas have five special sensors positioned over the breast and stomach: three to monitor the infant's heart beat and two to monitor respiration. The special sensors are actually embeddd into the pyjama material and have no direct contact with the body, thus creating no discomfort to the baby.

3.5.3 Sports Science and Leisure

The Alive Heart and Activity Monitor is a wireless health monitoring system for screening, diagnosis and management of chronic diseases as well as for consumer health and fitness (http://www.alivetec.com/products.htm). Designed for use in the medical facility, home or gym, the monitor uses wireless Bluetooth and mobile

Fig. 3.6 Zephyr technology BioHarness (http://www.zephyranywhere.com/products/bioharness-3/)

phone networks to transmit ECG and accelerometer data to a mobile phone, computer, or central monitoring centre. The Alive Heart Monitor can be used for remote real time monitoring of exercise programs via the internet. Using this device, GPS and a mobile phone, the program supervisor or coach can remotely monitor the exact position, speed, ECG and heart rate in real time of an individual as they perform the exercises. For athletes and sports enthusiasts, this system provides real time feedback on their intensity and performance, allowing optimisation of the programs for maximum benefit and safety.

Zephyr technology produces the BioHarness chest strap that takes data from the body using smart fabric sensors (Fig. 3.6), allowing remote monitoring of human performance and condition in the real world (http://www.zephyranywhere.com/products/bioharness-3/). These sensors are comfortable, integratable and perform in the field under harsh environments and under extreme exercise. Zephyr interfaces to custom and industry standard radio interfaces—treating the radio and internet link as a transparent pipe—to get the information to where it is needed. Data is presented on PC or web applications making the data useful in the context of the situation and user. This may be to assist coaches in team training, medics involved in triage for injured people, or First Responder commanders concerned about the status of their crew-members. The device includes monitoring of heart rate, breathing rate, posture, activity level and peak acceleration.

3.5.4 Personal Healthcare and Well-being

Tools for personal healthcare and well-being are increasingly becoming available from commercial sectors. An example is the Basis B1 (http://www.mybasis.com/); a wrist-watch like health tracker and online personal dashboard designed to help people easily incorporate healthy habits into their daily routines. The Basis B1 band evaluates heart rate as a person moves through three natural states: sleeping, waking and “being active”. The Basis B1 band uses a patient-worn monitoring device and

transmits data to an online portal for analysis and feedback. Similarly, LifeSpan Fitness developed the MyBeat; a finger heart rate monitor in the form of a ring, providing the ability to check heart rate in real-time during exercise (http://www.lifespanfitness.com/products_heartratering.asp). It includes a stopwatch, a timer and a heart rate display able to count in the range 30–250 beats per minute.

HealthHub is a more complete and integrated personal health and lifestyle system (http://www.healthhub.com/info/what_is_HealthHub/Default.aspx). It is a small handheld telehealth monitor, designed for ease of use by patients and caregivers of all ages. In size, it is similar to a PDA and uses the principle of "store and forward". It enables measurement, entry and reading of information from a variety of sources. It can monitor vital signs parameters (weight, blood pressure, ECG), changes in symptoms, side effects, medication compliance, life style and quality of life, and then send encrypted information to the health centre.

The European Commission have funded several projects for the development of smart, wearable devices for health monitoring, through their various Framework funding programmes. Of these, Biotex developed textile based sensors to measure a particular set of physiological variables for personalised healthcare management (Coyle et al. 2010). These include sensors for sweat rate, ECG, respiration and SpO_2. By comparison, the MyHeart project focused on acquiring knowledge of an individual's actual health status and lifestyle as part of a study into cardio-vascular disease prevention. The system developed functional clothing with integrated textile sensors, allowing continuous monitoring of ECG vital signs (http://www.hitech-projects.com/euprojects/myheart/). Data collected from IMU sensors within the garments is also classified by activity, including resting, lying down, walking, running and going upstairs or downstairs.

3.5.5 Human Activity Monitoring

Objective and accurate information about the physical activities undertaken by an individual during normal daily life is of fundamental importance for research as well as for clinical practice. Furthermore, data collected provides useful feedback to the user, which facilitates education and self management. Within this category, systems for gait monitoring have proven popular and beneficial in human performance measurements worldwide. For example, Bamberg et al. developed and demonstrated a shoe containing a number of different kinematic and physical sensors to assess the characteristics of human gait, including gyroscopes, accelerometers and force-sensitive resistors, with their data being processed via data fusion algorithms (Bamberg et al. 2008).

Vivonoetics produce the Equivital™ EQ02 LifeMonitor: a body worn, multi-parameter platform for human performance data collection and analysis (http://vivonoetics.com/products/sensors/equivital/). The basic system provides data logging facilities and Bluetooth connectivity, as well as an analogue front end to interface with a collection of sensor modules including ECG, GSR, respiration, skin surface temperature, activity and SpO_2. It also includes an ingestible core body temperature

Fig. 3.7 The BodyMedia SenseWear armband (http://sensewear.bodymedia.com/)

module and a variety of bespoke software applications for data viewing and analysis on several different multimedia platforms.

The Intelligent Device for Energy Expenditure and Activity (IDEEA) (http://www.minisun.com/) is a device designed for physical activity and behaviour monitoring, gait analysis, energy expenditure estimation and posture detection. An external set of sensors consisting of five biaxial accelerometers are attached to the lower limbs and are wire-connected to a portable recorder worn at the waist. It uses a 32-bit microprocessor that enables real-time data acquisition and processing. The IDEEA has been used in monitoring physical activities of obese people in real life environments, and has been validated in a study of ambulatory measurement for gait analysis, and energy expenditure estimation of physical activity.

The SenseWear system (http://sensewear.bodymedia.com/) is a wearable armband developed by BodyMedia to enable automatic monitoring of energy consumption, activity and sleep efficiency, providing the caregiver with a very useful tool to assess and educate patients to change their behaviour (Fig. 3.7). It employs a multisensory array composed of a 2-axis accelerometer, a near-body ambient temperature sensor, a skin temperature sensor, a heat flux sensor and a sensor for galvanic skin response. Bespoke software uses sensor data to calculate total energy expenditure, active energy expenditure, step counter, physical activity levels and duration, sleep duration and efficiency, lying down time.

StayHealthy has developed two motion monitor products, the CT1 Calorie Tracker and the RT3 (http://www.stayhealthy.com/en_us/main/). Both products can be worn with the placement of a clip at the waist. CT1 is a FDA cleared Class II medical device for accurate energy expenditure estimation. RT3 is an activity monitor that uses a piezoelectric 3-axis accelerometer to provide metabolic

equivalent of tasks (METs) for clinical and research applications. RT3 also replaces the previous version Tritrac-R3D, which has been widely used in a number of studies and research applications. A validation of RT3 for the assessment of physical activity (PA) reported that RT3 was a reliable assessor of PA for boys and men. RT3 has been used in recording temporal patterns of activity in chronic obstructive pulmonary disease (COPD) patients.

The AMP 331 produced by Dynastream Innovations is an activity monitor positioned on the back of the ankle. With the proprietary "SpeedMax" technology, AMP 331 uses accelerometers to measure the forward and vertical accelerations to determine the position of the foot in space. Major gait parameters, such as stride length, speed and travelled distance during walking or running can be calculated. The recorded data in AMP 331 can be downloaded to PCs via a 916 MHz wireless radio receiver. The company showed that the accuracy in distance computation is about 97 % rising to approximately 99 % after proper calibration. A study was conducted to validate the AMP 331 in assessing energy expenditure (EE) (Darter et al. 2006). This study recruited 41 subjects whose 12-hour daily activities in a field environment were recorded. The EE estimate from the AMP 331 and diary record were compared and produced a Pearson correlation coefficient of 0.651. The AMP 331 was reported to better estimate EE than other wearable sensors during walking using the manufacturer's estimation equation. The accuracy of the AMP 331 to detect atypical gait was also studied. The AMP 331 performed better than other sensors (compared with data obtained from video recordings) in detecting structured walking and stair ascent/descent.

The GT1M by ActiGraph uses a single axis accelerometer to measure acceleration at 30 Hz sampling rate and with 12-bit resolution over the range 0.05–2.5 g (http://www.actigraphcorp.com/support/devices/gt1m/). The sampled signals are bandpass filtered between 0.25 and 2.5 Hz to remove high frequency noise and low frequency artefacts. The GT1M can be worn at the waist to measure activity counts, step counts, activity level and energy expenditure. It can also be worn on the wrist for sleep monitoring. The data can be downloaded to the PC software "ActiLife" via USB connection. GT1M has been used in evaluating physical activity levels in children and adolescents. The ActiGraph series was the most studied activity monitor, and many studies have validated its reliability and performance.

The Activ8 physical activity monitor is a small (34 × 30 × 10 mm) device to track physical activity during the entire day by using 3D accelerometer technology (http://www.vitamove.nl/products/activ8/). The measured physical activity levels are converted into accurate kilocalorie equivalent values that reflect physical energy expenditure. Activ8 can also detect major activity classes like lying, sitting, standing, walking, cycling and running, providing useful information about what the user has been doing. By monitoring physical energy expenditure, Activ8 provides an ability to monitor and analyse user's lifestyle. The device does not need to be in intimate contact with the user and can simply be carried in a pocket. Therefore it is less likely to restrict or hinder normal movement or activities. The Activ8 physical activity monitor is delivered with easy to use PC software to start/stop and to display a recording over time. Furthermore it includes a reporting service, providing an overview of the entire recording period.

3.5.6 Other

ProeTEX is a European Commission funded project aimed at developing smart garments for emergency-disaster personnel (Curone et al. 2010). Portable devices and sensors are included in such garments in order to continuously monitor, in real-time, the risks endangering lives of rescuers. The system realized within this project allows detection of vital parameters such as heart rate, breathing rate, body temperature, SpO_2, position, activity and posture. It also has the capabilities to measure environmental variables including the presence of toxic gases (CO_2), external temperature and heat flux through the garment. An electronic box collects data acquired by the sensors and sends them to a local operation system via wireless. Preliminary results show a good efficiency of this approach in monitoring of such parameters.

3.6 Challenges Ahead

The interest in wireless sensor development for remote healthcare and the different available solutions have been presented in this chapter. In order to succeed and gain popularity, however, these wireless sensors need to overcome several technological challenges. They can be categorised as follows.

Quality and reliability challenges The sensor is expected to last for a long time in order to avoid uncomfortable sensor replacement or re-calibration. Any device has its own lifetime but biomedical applications can in some cases require proper running of sensors for several years of continuous use. Because of this the sensor must be extremely robust. The reliable function of the sensors must also be accompanied by the robustness of the sensor network. Therefore the network needs to be tested in coexistence with other wireless networks that can interfere with the communications. The coexistence with other networks is especially critical in hospitals where many types of medical apparatus use wireless communication and the interferences between them can literally be fatal. Hence the sensor network must also be robust and fault tolerant by means, for example, of distributed operation architectures in which the sensors cooperate with the surrounding devices. This, however, requires high computational capabilities to power and size constrained devices. In addition, since the biomedical biosensors are going to be used to diagnose and treat diseases, they need to comply with strict regulatory requirements. It is therefore mandatory to demonstrate that the information obtained by them is sufficiently reliable and that their use will not only be harmless, but also of demonstrated value to patients. This is particularly relevant when the signal measured by the sensor is considered to be a critical or vital sign, whereupon solid clinical trials are required with the consequent increase in costs, time and the potential rejection of the sensor as a beneficial and compliant medical device.

Communication and security challenges The communication between the sensor and the network should be done with a standardised communication protocol in order to use a single network access point with different wearable sensors and to reduce the implementation cost of the system. These communication standards should transmit data securely and privately since the information obtained by the sensor is very intimate and should only be accessible to authorised users. However, most available communication standards do not usually provide enough security, so extra encryption mechanisms may be required. These extra security mechanisms may be in conflict with the limited computation capabilities of the sensor, so that new low-computational encryption solutions are required in which privacy is correctly ensured.

Biocompatibility and wearability challenges Wearable sensors need to be comfortable and biocompatible. This is especially true for implantable solutions where sensors must be miniaturised and built from biocompatible materials to ensure that the body does not react negatively to them. The miniaturisation of sensors is a great challenge and there is continuing research being undertaken to address the issues and provide new solutions. For example, the size of the device could be reduced if the transducer and the necessary signal conditioning and processing electronics are combined into a single integrated circuit by using CMOS and MEMS compatible fabrication processes. Recent advances in microfluidics (Lewis et al. 2013) and "lab-on-chip" (LOC) technologies may also present new form factors for wearable solutions, as well as offering new biochemical signal sensing opportunities. In particular, LOC systems incorporating biological or biomedical MEMS devices are poised to have a significant impact on ehealth. These devices also referred to as point-of-care (POC) systems, represent a significant opportunity in various patient-centric settings, including at home, at the doctor's office, in ambulances, in emergency rooms, etc. (Watkins et al. 2011; Wolf and Herzog 2013).

Power consumption challenges In an ideal scenario the communication between the sensor and the network access point that receives the information should be reliable over large distances, particularly with high sensor autonomies. However, miniaturised sensor devices are powered by batteries with limited and (generally) low power, which makes achieving long communication distances difficult. In addition, even when considering the shorter communication ranges of body area networks (BAN), biomedical sensors should be designed for high autonomy. For example, in the case of implanted sensors, the sensor must remain in operation for several years once it is implanted. Wearable sensors may not have such demanding autonomies by comparison, but the acceptance of sensors is negatively impacted if the autonomy is not sufficient for purpose.

Because of these requirements, much research effort is being directed toward reducing the power consumption of sensors in order to increase their autonomy. Interesting research is being conducted in the field of energy harvesting (Beeby et al. 2006; Cook-Chennault et al. 2008; Leonov 2011). In this technology the energy is captured from the environment and converted into electrical energy. This technology may extend the autonomy of a battery powered sensor node or even

Table 3.3 Estimated power output levels per harvesting source (Fiorini et al. 2008)

Source	Harvested power
Photovoltaic	10 $\mu W/cm^2$ to 10 mW/cm^2
Vibration/motion	4–100 $\mu W/cm^2$
Thermal energy	25 $\mu W/cm^2$ to 10 mW/cm^2
RF (GSM Standard)	0.1 $\mu W/cm^2$

replace the battery itself. The energy can be captured from a number of sources: vibration harvesting, thermal harvesting, photovoltaic harvesting, RF harvesting, etc. Of course, energy harvesting mechanisms require efficient and smart power management in order to ensure that the greatest amount of energy is available to the sensor node and to maximize the sensor node operation from a limited energy source. Table 3.3 shows various energy sources along with an estimate of the power that could be harvested (Fiorini et al. 2008).

Techniques such as the ones presented in Table 3.3 allow increases to the communication range or the sensor autonomy without increasing the battery size. Although being a promising area of research, energy harvesting presents its own challenges. The energy harvested by these techniques may be low and highly sporadic and therefore smart management of the available energy is required for its efficient use. This requires more computational power from the sensors. Furthermore, the increment in size for sensors including power harvesting modules may be unacceptable for many applications.

In the future, virtual reality, immersive environments, haptic feedback and nanotechnology promise new ways to improve the capabilities of ehealth. Nevertheless, the technology for telemonitoring is currently far from mature and there are still technological issues to be addressed as identified above. There will also be challenges for adoption of such systems because individuals may see constant physiological surveillance as intrusive. There are also potential cultural problems to be tackled in relation to the deployment of such technology in health care organisations, as they produce requirements for new ways of working. However, despite all the potential hurdles, it is likely to be only a matter of time before smart systems continuously monitor every patient from the moment they are admitted to the point of discharge from hospital (and possibly beyond) (Nangalia et al. 2010).

References

Abbott. Abbott diabetes care—FreeStyle navigator. http://www.abbottdiabetescare.co.uk/your-products/freestyle-navigator. Accessed 25 March 2013

Abraham WT, Adamson PB, Bourge RC et al (2011) Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. Lancet 377(9766):658–666

Activ8. http://www.vitamove.nl/products/activ8/. Accessed 25 March 2013

Alive Heart. http://www.alivetec.com/products.htm. Accessed 23 March 2013

AliveCor Heart Monitor. http://www.alivecor.com/. Accessed 5 July 2013

Allen J (2007) Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 28:R1–R39

American College of Obstetricians and Gynecologists (2006) Menstruation in girls and adolescents: using the menstrual cycle as a vital sign. ACOG Committee Opinion No. 349. Obstet Gynecol 108(5):1323–1328
AN3182 Application Note (2010) Tilt measurement using a low-g 3-axis accelerometer. Rev 1. STMicroelectronics. http://www.st.com/web/en/resource/technical/document/application_note/CD00268887.pdf. Accessed 4 July 2013
Anliker U, Ward JA, Lukowicz P et al (2004) Amon: a wearable multiparameter medical monitoring and alert system. IEEE Trans Inf Technol Biomed 8(4):415–427
Armstrong S (2007) Wireless connectivity for health and sports monitoring: a review. Brit J Sports Med 41:285–289
Asada H (2009) Wearable blood pressure sensor offers 24/7 continuous monitoring. MIT Tech Talk 53(21):4
Asada HH, Shaltis P, Reisner A et al (2003) Mobile monitoring with wearable photoplethysmographic biosensors. IEEE Eng Med Biomag 22(3):28–40
Astra 300 spirometer. http://sdidiagnostics.com/spirometers/astra300.php. Accessed 6 July 2013
Avivo mobile patient management system. http://www.corventis.com/US/avivo.asp. Accessed 6 July 2013
Bamberg SJM, Benhasat AY, Scarborough DM et al (2008) Gait analysis using a shoe-integrated wireless sensor system. IEEE Trans Inf Technol Biomed 12(4):413–423
Basis B1. http://www.mybasis.com/. Accessed 23 March 2013
Beeby SP, Tudor MJ, White NM (2006) Energy harvesting vibration sources for microsystems applications. Meas Sci Technol 17:R175–R195
Berntsen S, Hageberg R, Aandstad A et al (2010) Validity of physical activity monitors in adults participating in free-living activities. Br J Sports Med 44(9):657–664
Bierman AS (2001) Functional status: the sixth vital sign. J Gen Intern Med 16(11):785–786
BioHarness. http://www.zephyranywhere.com/products/bioharness-3/. Accessed 1 July 2012
Bio-Medical Instruments, Inc. Galvanic skin response sensors. http://bio-medical.com/products/sensors/gsr-sensors.html. Accessed 28 March 2013
Bogue R (2007) MEMS sensors: past, present and future. Sens Rev 27(1):7–13
Bonato P (2005) Advances in wearable technology and applications in physical medicine and rehabilitation. J Neuroeng Rehabil 2:2–5
Bultz BD, Carlson LE (2005) Emotional distress: the sixth vital sign in cancer care. J Clin Oncol 23(26):6440–6441
C8 MediSensors Optical Glucose Monitor System. http://www.c8medisensors.com. Accessed 25 March 2013
Cardiomems. http://www.cardiomems.com/content.asp?display=aboutus&view=about%20us%20revised. Accessed 6 July 2013
Casson AJ, Smith S, Duncan JS et al (2008) Wearable EEG: what is it, why is it needed and what does it entail? In: Proceedings of the 30th annual international IEEE EMBS conference, Vancouver, BC, 20–24Aug 2008
Chan M, Estève D, Fourniols JY et al (2012) Smart wearable systems: current status and future challenges. Artif Intell Med 56(3):137–156
Chang JR, Tai CC (2005) A new wireless-type physiological signal measuring system using a PDA and the Bluetooth technology. Biomed Eng Appl Basis Comm 17:229–235
Chen W, Wei D, Zhu X et al (2005) A mobile phone-based wearable vital signs monitoring system. In: Proceedings of the fifth international conference on computer and information technology, Shanghai, China, 21–23 Sept 2005
Chen L, Reisner AT, Gribok A (2009) Can we improve the clinical utility of respiratory rate as a monitored vital sign? Shock 31(6):575–581
Chi YM, Jung TP, Cauwenberghs G (2010) Dry-contact and noncontact biopotential electrodes: methodological review. IEEE Rev Biomed Eng 3:106–119
Chung WY, Lee SC, Toh SH (2008) WSN based mobile u-healthcare system with ECG, blood pressure measurement function. In: Proceedings of the 30th annual international conference of the IEEE EMBS, Vancouver, BC, 20–24 Aug 2008

Coleman KJ, Ngor E, Reynolds K et al (2012) Initial validation of an exercise "vital sign" in electronic medical records. Med Sci Sports Exerc 44(11):2071–2076
Cook-Chennault KA, Thambi N, Sastry AM (2008) Powering MEMS portable devices—a review of non-regenerative and regenerative power supply systems with special emphasis on piezoelectric energy harvesting systems. Smart Mater Struct 17(4):1–33
Corrventis. NUVANT Mobile Cardiac Telemetry (MCT) System. http://www.corventis.com/us/nuvant.asp. Accessed 25 Mar 2013
Coyle S, Lau KT, Moyna N et al (2010) BIOTEX—biosensing textiles for personalised healthcare management. IEEE Trans Inf Technol Biomed 14(2):364–370
Curone D, Secco EL, Tognetti A et al (2010) Smart garments for emergency operators: the Proetex project. IEEE Trans Inf Technol Biomed 14(3):694–701
Darter BJ, Janz KF, Puthoff ML et al (2006) Reliability and accuracy of the AMP 331 for activity monitoring and energy expenditure prediction in young adults. J Phys Act Health 3(3):277–291
Davis MP, Walsh D (2004) Cancer pain: how to measure the fifth vital sign. Cleve Clin J Med 71 (8):625–632
Derchak PA (2011) The LifeShirt preclinical system is a significant advance towards answering the call of the 3Rs. National Centre for the Replacement, Refinement, and Reduction of Animals in Research:1–9
Dexcom. Dexcom G4 Platinum. http://www.dexcom.com/dexcom-g4-platinum. Accessed 25 March 2013
Di Rienzo M, Meriggi P, Rizzo F et al (2010) Textile technology for the vital signs monitoring in telemedicine and extreme environments. IEEE Trans Inf Technol Biomed 14(3):711–717
English LB (2012) In pursuit of an ideal—a perspective on non-invasive continuous glucose monitoring. Eur Endocrinol 8(1):18–21
Ermes M, Parkka J, Mantyjarvi J et al (2008) Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. IEEE Trans Inf Technol Biomed 12(1): 20–26
Exmovere. The Empath Watch. http://www.exmovere.com/pdf/Exmovere_Wearable_Sensor_Research.pdf. Accessed 25 March 2013
Fernandez E, Solar H, de No J et al (2012) A 150nW CMOS novel temperature sensor for remote patient monitoring based on an auto-resonant active inductor architecture. In: Proceedings of the IEEE international symposium on circuits and systems (ISCAS 2012), Seoul, Korea, 20–23 May 2012
Ferrante do Amaral CE, Wolf B (2008) Current development in non-invasive glucose monitoring. Med Eng Phys 30(5):541–549
Fiorini P, Doms I, Van Hoof C et al (2008) Micropower energy scavenging. In: Proceedings of the IEEE European solid-state circuits conference (ESSCIRC 2008), Edinburgh, Scotland, 15–19 Sept 2008
Fischell DR, Fischell TA, Harwood J et al (2007) Implantable device for vital signs monitoring. US Patent 2007/0016089 A1, 18 Jan 2007
Fletcher RR, Dobson K, Goodwin MS et al (2010) iCalm: wearable sensor and network architecture for wirelessly communicating and logging autonomous activity. IEEE Trans Inf Technol Biomed 14(2):215–223
Foerster F, Smeja M, Fahrenberg J (1999) Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. Comput Human Behav 15:571–583
Freed A, Chan ADC, Lemaire ED et al (2011) Wearable EMG analysis for rehabilitation (WEAR)—surface electromyography in clinical gait analysis. In: Proceedings of IEEE international workshop on medical measurements and applications (MeMeA), Bary, Italy, 20–31 May 2011
Fulk GD, Sazonov E (2011) Using sensors to measure activity in people with stroke. Top Stroke Rehabil 18(6):746–757
Gao T, Greenspan D, Welsh M et al (2005) Vital signs monitoring and patient tracking over a wireless network. In Proceedings of the 27th annual international conference of the IEEE EMBS, Shanghai, 1–4 Sept 2005

Ghasemzadeh H, Jafari R, Prabhakaran B (2010) A body sensor network with electromyogram and inertial sensors: multimodal interpretation of muscular activities. IEEE Trans Inf Technol Biomed 14(2):198–206
Giansanti D, Macellari V, Maccioni G et al (2003) Is it feasible to reconstruct body segment 3-D position and orientation using accelerometric data? IEEE Trans Biomed Eng 50(4):476–483
Girardin CM, Huot C, Gonthier M et al (2009) Continuous glucose monitoring: a review of biochemical perspectives and clinical use in type 1 diabetes. Clin Biochem 42(3):136–142
Greenwood JL, Joy EA, Stanford JB (2010) The physical activity vital sign: a primary care tool to guide counselling for obesity. J Phys Act Health 7(5):571–576
Groff CP, Mulvaney PL (2000) Wearable vital sign monitoring system. US Patent 6,102,856, 15 Aug 2000
GT1M. http://www.actigraphcorp.com/support/devices/gt1m/. Accessed 24 March 2013
Haahr RG, Duun S, Thomsen EV (2008) A wearable "electronic patch" for wireless continuous monitoring of chronically diseased patients. In: Proceedings of the fifth international workshop on wearable and implantable body sensor networks, Hong Kong, China, 1–3 June 2008
HealthHub. http://www.healthhub.com/info/what_is_HealthHub/Default.aspx. Accessed 5 July 2013
Heimann K, Steffen M, Bernstein N et al (2009) Noncontact monitoring of heart and lung activity using magnetic induction measurement in a neonatal animal model. Biomed Tech 54(6): 337–345
Hogrel J-Y (2005) Clinical applications of surface electromyography in neuromuscular disorders. Neurophysiol Clin 35:59–71
Hung K, Zhang YT, Tai B (2004) Wearable medical devices for tele-home healthcare. In: Proceedings of the 26th annual international conference of the IEEE EMBS, San Francisco, CA, 1–5 Sept 2004
Hunter Peckham P, Knutson JK (2005) Functional electrical stimulation for neuromuscular applications. Annu Rev Biomed Eng 7:327–360
Hyde RA, Ketteringham LP, Neild SA et al (2008) Estimation of upper-limb orientation based on accelerometer and gyroscope measurements. IEEE Trans Biomed Eng 55(2):746–754
Ichimori S, Nishida K, Shimoda S et al (2006) Development of a highly responsive needle-type glucose sensor using polyimide for a wearable artificial endocrine pancreas. J Artif Organs 9 (2):105–113
IDEEA. http://www.minisun.com/. Accessed 24 March 2013
Izzetoglu M, Izzetoglu K, Bunce S et al (2005) Functional near-infrared neuroimaging. IEEE Trans Neural Syst Rehabil Eng 13(2):153–159
Jayaraman S (2005) Fabric-based sensor for monitoring vital signs. US Patent 6,970,731, 29 Nov 2005
Jiang M, Shang H, Wang Z et al (2011) A method to deal with installation errors of wearable accelerometers for human activity recognition. Physiol Meas 32:347–358
Johnstone JA, Ford PA, Hughes G et al (2012) Bioharness multivariable monitoring device. Part I: validity. J Sports Sci Med 11:409–417
Joseph A (2003) Continence: the sixth vital sign? Let's not ignore urinary incontinence. Am J Nurs 103(7):11
Jovanov E, Milenković A (2011) Body area networks for ubiquitous healthcare applications: opportunities and challenges. J Med Syst 35:1245–1254
Jovanov E, Milenković A, Otto C et al (2005) A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. J Neuroeng Rehabil 2:6–15
Kendell C, Lemaire ED (2009) Effect of mobility devices on orientation sensors that contain magnetometers. J Rehabil R D 46(7):957–962
Lee D-S, Lee Y-D, Chung W-Y et al (2006) Vital sign monitoring system with life emergency event detection using wireless sensor network. In: Proceedings of the IEEE fifth sensors conference, Daegu, Korea, 22–25 Oct 2006
Lee GX, Low KS, Taher T (2010) Unrestrained measurement of arm motion based on a wearable wireless sensor network. IEEE Trans Instrum Meas 59(5):1309–1317

Leonov V (2011) Energy harvesting for self-powered wearable devices. In: Bonfiglio A, De Rossi D (eds) Wearable monitoring systems. Springer, New York, pp 27–49
Lewis AP, Cranny A, Harris NR et al (2013) Review on the development of truly portable and in-situ capillary electrophoresis systems. Meas Sci Technol. doi:10.1088/0957-0233/24/4/042001
LifeMonitor. http://vivonoetics.com/products/sensors/equivital/. Accessed 5 July 2013
Lifewatch. PMP4 Wireless Healthcare System. http://www.lifewatch.com/wirelesshealthcare. Accessed 25 March 2013
Lin Y-H, Jan I-C, Ko PC-I et al (2004) A wireless PDA-based physiological monitoring system for patient transport. IEEE Trans Inf Technol Biomed 8(4):439–447
Loew N, Winzer KJ, Becher G et al (2007) Medical sensors of the BASUMA body sensor network. In: Proceedings of the fourth international workshop on wearable and implantable body sensor networks, Aachen, Germany, 26–28 March 2007
Lorincz K, Malan DJ, Fulford-Jones TRF et al (2004) Sensor networks for emergency response: challenges and opportunities. Pervasive Comput 3(4):16–23
Lötters JC, Schipper J, Veltink PH et al (1998) Procedure for in-use calibration of triaxial accelerometers in medical applications. Sens Actuators A 68:221–228
Luinge HJ, Veltink PH (2004) Inclination measurement of human movement using a 3-D accelerometer with autocalibration. IEEE Trans Neural Syst Rehabil Eng 12(1):112–121
Luinge HJ, Veltink PH (2005) Measuring orientation of human body segments using miniature gyroscopes and accelerometers. Med Biol Eng Comput 43:273–282
Lukowicz P, Anliker U, Ward J et al (2002) AMON: a wearable medical computer for high risk patients. In: Proceedings of the IEEE sixth international symposium on wearable computers, Seattle, WA, 7–10 Oct 2002
Luprano J (2006) European projects on smart fabrics, interactive textiles: sharing opportunities and challenges. In: Proceedings of workshop on wearable technology and intelligent textiles, Helsinki, Finland, 22 Nov 2006
Lymberis A, Paradiso R (2008) Smart fabrics and interactive textile enabling wearable personal applications: R&D state of the art and future challenges. In: Proceedings of the IEEE international conference on engineering in medicine and biology society (EMBS 2008), Vancouver, BC, 20–24 Aug 2008
Lynch M (2001) Pain as the fifth vital sign. J Intraven Nurs 24(2):85–94
Maharatna K, Mazomenos EB, Morgan J et al (2012) Towards the development of next-generation remote healthcare system: Some practical considerations. In: Proceedings of the IEEE international symposium on circuits and systems (ISCAS), Seoul, Korea
Mamagoose paediatric pyjamas. http://www.esa.int/esaKIDSen/SEML8LXJD1E_Technology_1.html. Accessed 5 July 2013
Mamagoose paediatric pyjamas. http://www.esa.int/Our_Activities/Technology/TTP2/New_pyjamas_could_prevent_cot_deaths. Accessed 5 July 2013
Marx AJ (1976) Patient vital-signs automated measuring apparatus. US Patent 3,996,928, 14 Dec 1976
Mathie MJ, Coster ACF, Lovell NH et al (2004) Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiol Meas 25:R1–R20
Mayagoitia RE, Nene AV, Veltink PH (2002) Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. J Biomech 35:537–542
Medical International Research MirOXI digital multimeter. http://www.mir.abmedic.com/miroxi.htm. Accessed 25 March 2013
Medron Inc. Medron automatic blood pressure monitor. http://www.medroninc.com. Accessed 25 March 2013
Medtronic. About Pacemakers. http://www.medtronic.com/patients/bradycardia/device/index.htm. Accessed 25 March 2013
Medtronic. http://www.medtronic.co.uk/index.htm. Accessed 6 July 2013

Medtronic. MiniMed Paradigm Real-Time Revel System. http://www.medtronicdiabetes.com/products/paradigmrevelsystem. Accessed 25 March 2013
Meystre S (2005) The current state of telemonitoring: a comment on the literature. Telemed J E Health 11(1):63–69
Ming-Zher P, Swenson NC, Picard RW (2010) A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. IEEE Trans Biomed Eng 57(5):1243–1252
Money EW, Caldwell R, Sciarra M (1999) Vital sign remote monitoring device. US Patent 5,919,141, 6 July 1999
Mower WR, Sachs C, Nicklin EL et al (1997) Pulse oximetry as a fifth pediatric vital sign. Pediatrics 99:681–686
Mower WR, Myers G, Nicklin EL et al (1998) Pulse oximetry as a fifth vital sign in emergency geriatric assessment. Acad Emerg Med 5(9):858–865
Mularski RA, White-Chu F, Overbay D et al (2006) Measuring pain as the 5th vital sign does not improve quality of pain management. J Gen Intern Med 21(6):607–612
Mundt CW, Montgomery KN, Udoh UE et al (2005) A multiparameter wearable physiologic monitoring system for space and terrestrial applications. IEEE Trans Inf Technol Biomed 9(3):382–391
MyBeat. http://www.lifespanfitness.com/products_heartratering.asp. Accessed 23 March 2013
MyHeart. http://www.hitech-projects.com/euprojects/myheart/. Accessed 23 March 2013
Najafi B, Aminian K, Paraschiv-Ionescu A et al (2003) Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. IEEE Trans Biomed Eng 50(6):711–723
Nangalia V, Prytherch DR, Smith GB (2010) Health technology assessment review: remote monitoring of vital signs—current status and future challenges. Crit Care 14:233–240
Neff TA (1988) Routine oximetry: a fifth vital sign? Chest 94(2):227
Nishida Y, Kawakami G, Mizoguchi H (2006) Everyday grasping behavior measurement with wearable electromyography. In Proceedings of the fifth IEEE conference on sensors, Daegu, Korea, 22–25 Oct 2006
O'Donovan KJ, Kamnik R, O'Keeffe DT et al (2007) An inertial and magnetic sensor based technique for joint angle measurement. J Biomech 40:2604–2611
OECD (2011) Health at a glance 2011: OECD indicators. OECD Publishing. http://dx.doi.org/10.1787/health_glance-2011-en. Accessed 1 July 2013
Otto C, Milenković A, Sanders C et al (2006) System architecture of a wireless body area sensor network for ubiquitous health monitoring. J Mob Multimed 1(4):307–326
Pandian PS, Mohanavelu K, Safeer KP et al (2008) Smart Vest: wearable multi-parameter remote physiological monitoring system. Med Eng Phys 30(4):466–477
Pantelopoulos A, Bourbakis NG (2010) A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans Syst Man Cybern C 40(1):1–12
Paradiso R (2003) Wearable health care system for vital signs. In: Proceedings of the fourth annual IEEE conference on information technology applications in biomedicine, Birmingham, UK, 24–26 April 2003
Paradiso R, Loriga G, Taccini N et al (2005a) WEALTHY—a wearable healthcare system: new frontier on e-textile. J Telecomm InfTechnol 4:105–113
Paradiso R, Loriga G, Taccini N (2005b) A wearable health care system based on knitted integrated sensors. IEEE Trans Inf Technol Biomed 9(3):337–344
Patel S, Hughes R, Hester T et al (2010) A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology. Proc IEEE 98(3):450–461
Patel S, Park H, Bonato P et al (2012) A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 9:21–32
Peters JL (1983) Vital signs monitoring apparatus. US Patent 4,383,534, 17 May 1983
Pfohl RL (1991) Vital signs monitoring and communication system. US Patent 4,981,139, 1 Jan 1991
Poon CCY, Liu Q, Gao H et al (2011) Wearable intelligent systems for E-health. J Comput Sci Eng 5(3):246–256

Prance H (2011) Sensor developments for electrophysiological monitoring in healthcare. In: Gargiulo GD, McEwan A (eds) Applied biomedical engineering, InTech, Croatia, pp 265–286
Preventice. BodyGuardian Remote Monitoring System. http://www.preventice.com/products/bodyguardian. Accessed 25 March 2013
Procomp Infiniti system sensors. http://www.thoughttechnology.com/sensorsnew.htm. Accessed 6 July 2013
Procomp Infiniti system wireless adapater. http://www.thoughttechnology.com/teleinfinit.htm. Accessed 6 July 2013
ProComp Infiniti system. http://www.thoughttechnology.com/proinf.htm. Accessed 6 July 2013
Pross G, Schlueter M (1994) Method and system for monitoring vital signs. US Patent 5,343,869, 6 Sept 1994
Ragnarsson KT, Pollack S, O'Daniel W Jr et al (1998) Clinical evaluation of computerized functional electrical stimulation after spinal cord injury: a multicenter pilot study. Arch Phys Med Rehabil 69(9):672–677
Rendell M, Saiprasad S, Trepp-Carrasco A et al (2013) The future of inpatient diabetes management: glucose as the sixth vital sign. Expert Rev Endocrinol Metab 8(2):195–205
Richards Grayson AC, Scheidt Shawgo R, Li Y, Cima MJ (2004) Electronics MEMS for triggered delivery. Advanced Drug Delivery Reviews 56(2):173–184
Roggen D, Magnenat S, Waibel M et al (2011) Wearable computing. IEEE Robot Autom Mag 18:83–95
Rovira C, Coyle S, Corcoran B et al (2011) Integration of textile-based sensors and Shimmer for breathing rate and volume measurement. In: Proceedings of the IEEE fifth international conference on pervasive computing technologies for healthcare, Dublin, Ireland, 23–26 May 2011
Rushton DN (1997) Functional electrical stimulation. Physiol Meas 18:241–275
Saltzstein WE, Sabri M, Dobaj AP et al (1999) Medical patient vital signs monitoring apparatus. US Patent 5,931,791, 3 Aug 1999
Salvo P, Di Francesco F, Costanzo D et al (2010) A wearable sensor for measuring sweat rate. IEEE Sens 10(10):1557–1558
Saywell N, Vandal AC, Brown P et al (2012) Telerehabilitation to improve outcomes for people with stroke: study protocol for a randomised controlled trial. Trials 13:233–243
Schlegelmilch RM, Kramme R (2012) Pulmonary function testing. In: Kramme R, Hoffmann KP, Pozos RS (eds) Handbook of medical technology. Springer, Berlin
Scilingo EP, Gemignani A, Paradiso R et al (2005) Performance evaluation of sensing fabrics for monitoring physiological and biomechanical variables. IEEE Trans Inf Technol Biomed 9(3):345–352
SenseWear. http://sensewear.bodymedia.com/. Accessed 1 July 2013
Sensixa e-AR with SpO_2. http://www.sensixa.com. Accessed 25 March 2013
Shimmer Wearable Wireless Sensor Platform. http://www.shimmer-research.com/r-d. Accessed 25 March 2013
Simel DL (2011) Approach to the patient: history and physical examination. In: Goldman L, Schafer AI (eds) Goldman's cecil medicine, 24th edn. Elsevier Saunders, Philadelphia
Smith JP (2005) Medical and biological sensors: a technical and commercial review. Sens Rev 25(4):241–245
Smith AD, Crabtree DR, Bilzon JL et al (2010) The validity of wireless iButtons and thermistors for human skin temperature measurement. Physiol Meas 31(1):95–114
Solar H, Fernández E, Tartarisco G et al (2013) A non invasive, wearable sensor platform for multi-parametric remote monitoring in CHF patients. Health Technol 3(2):99–109
Sotera ViSi Mobile. http://www.visimobile.com. Accessed 5 July 2013
Spirobank II. http://www.spirometry.com/ENG/Products/spirobank2.asp. Accessed 6 July 2013
Stabil-O-Graph. http://www.iem.de/stabil_o_graph_mobil2?_lang=1. Accessed 6 July 2013
StayHealthy. http://www.stayhealthy.com/en_us/main/. Accessed 24 March 2013
Steffen DA, Sturm RE, Rinard GA (1981) Vital signs monitoring system. US Patent 4,270,547, 2 June 1981

Stegeman DF, Blok JH, Hermens HJ et al (2000) Surface EMG models: properties and applications. J Electromyogr Kinesiol 10:313–326
Studenski S, Perera S, Wallace D et al (2003) Physical performance measures in the clinical setting. J Am Geriatr Soc 51(3):314–322
Sund-Levander M, Forsberg C, Wahren LK (2002) Normal oral, rectal, tympanic and axillary body temperature in adult men and women: a systematic literature review. Scand J Caring Sci 16(2):122–128
Sund-Levander M, Grodzinsky E, Loyd D et al (2004) Errors in body temperature assessment related to individual variation, measuring technique and equipment. Int J Nurs Pract 10(5): 216–223
Sung M, Marci C, Pentland A (2005) Wearable feedback systems for rehabilitation. J Neuroeng Rehabil 2:17–28
Suunto heart rate monitors. http://www.suunto.com/en-GB/sports-watch-collections/Suunto-M-Series-Collection/. Accessed 6 July 2013
Suzuki S, Matsui T, Kawahara H et al (2009) A non-contact vital sign monitoring system for ambulances using dual-frequency microwave radars. Med Biol Eng Comput 47:101–105
Taccini N, Loriga G, Pacelli M et al (2008) Wearable monitoring system for chronic cardio-respiratory diseases. In: Proceedings of the IEEE international conference on engineering in medicine and biology society (EMBS 2008), Vancouver, BC, 20–24 Aug 2008
Tierney LM, Whooley MA, Saint S (1997) Oxygen saturation: a fifth vital sign? West J Med 166 (4):285–286
Tong K, Granat MH (1999) A practical gait analysis system using gyroscopes. Med Eng Phys 21:87–94
Tura A, Badanai M, Longo D et al (2003) A medical wearable device with wireless Bluetooth-based data transmission. Meas Sci Rev 3:1–4
van Halteren A, Bults R, Wac K et al (2004) Mobile patient monitoring: the MobiHealth system. J Inf Technol Healthcare 2(5):365–373
Vardi A, Levin I, Paret G et al (2000) The sixth vital sign: end-tidal CO_2 in pediatric trauma patients during transport. Harefuah 139(3–4):85–87
Varshney U (2007) Pervasive healthcare and wireless health monitoring. Mob Netw Appl 12(2–3): 113–127
Veltink PH, Bussmann HBJ, de Vries W et al (1996) Detection of static and dynamic activities using uniaxial accelerometers. IEEE Trans Rehabil Eng 4(4):375–385
Vuorela T, Seppä VP, Vanhala J (2010) Design and implementation of a portable long-term physiological signal recorder. IEEE Trans Inf Technol Biomed 14(3):718–725
Watkins N, Irimia D, Toner M et al (2011) On a chip. IEEE Pulse 2(6):19–27
Wilson JF (2005) Is sleep the new vital sign? Ann Intern Med 142(10):877–880
Withings. Withings blood pressure monitor. http://www.withings.com/en/bloodpressuremonitor. Accessed March 2013
Wolf B, Herzog K (2013) Electronics for better healthcare. Eur J Prev Cardiol 20(2):2–7
World Health Organization (2013) World health statistics. WHO, Geneva. ISBN 978-92-4-156458-8
Yun X, Bachmann ER (2006) Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking. IEEE Trans Robot 22(6):1216–1227
Zhang Y, Hu H, Zhou H (2005) Study on adaptive Kalman filtering algorithms in human movement tracking. In: Proceedings of the IEEE international conference on information acquisition, Hong Kong and Macau, China, 27 June–3 July 2005
Zhou H, Stone T, Hu H et al (2008) Use of multiple wearable inertial sensors in upper limb motion tracking. Med Eng Phys 30:123–133

Chapter 4
Signal Processing Architecture Implementation Methodologies for Next-Generation Remote Healthcare Systems

Amit Acharyya

Abstract Remote healthcare, an emerging application with limited resources, requires sophisticated signal processing algorithms and their efficient architectural implementation methodologies. Direct mapping of traditional signal processing algorithms to hardware may not be suitable for such resource constrained applications. It is therefore necessary to explore the signal processing algorithms used for this applications and their corresponding low complexity and low power consumption architectural implementation using an algorithm-architecture holistic optimization approach. This chapter identifies and does the review on the traditional computationally intensive signal processing algorithms which are of significant use for healthcare related applications, focuses on the corresponding low-complexity architecture design and formulates the model of computing the overall hardware complexity of these architectures.

4.1 Role of Signal Processing in Remote Healthcare

Emerging resource constrained applications including remote healthcare require sophisticated signal processing algorithms and corresponding architectural implementation with limited area and low power consumption. Sensors used for such application capture the signals and transmit to some central node on continuous basis. Most of the processing are performed within the central node and it is commonly assumed that this central node has continuous power supply. However since the sensor nodes are mostly run by battery back-up and their radio front end consumes significantly high power, these nodes run out of power soon in such conventional approach. To overcome this problem, recently on-body-sensor

A. Acharyya (✉)
Indian Institute of Technology, Hyderabad-502205, India
e-mail: amit_acharyya@iith.ac.in

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_4,

resource constrained signal processing algorithms and architectures have been envisaged which would eliminate the need of such continuous data transmission. In general the physiological signals captured by the on-body sensors in the remote health monitoring application are corrupted with noise and other signals commonly known as artifacts. Therefore, in this chapter, we explore primarily those signal processing algorithms which can help denoising and signal separation. Wavelet Transform (WT), Kalman filtering and Short Time Fourier Transform (STFT) are widely used for denoising purpose (Addison 2004; Karvounis et al. 2007; Li et al. 1995; Pan et al. 1999; Senhadji et al. 1995). Approaches like Blind Source Separation (BSS), in particular Independent Component Analysis (ICA), have been used for accurately separating out the target signal from the mixture of several signals and in its feature extraction (Chien et al. 2006; He et al. 2006). On the other hand such applications, due to the scarcity of the available resources in terms of power and area, demands low-power low-complexity architectural designs of the available signal processing algorithms. Therefore co-optimization of signal processing algorithms and architectures plays a major role in overall energy minimisation for remote healthcare related applications. Thus in this chapter, we will also focus on the low-complexity architecture design of the above mentioned denoising and signal separation algorithms and also provide an insight into the computation of the complexity of the architectures.

Fast Fourier Transform is very well known technique used for signal denoising. However since the real-life biomedical signals, in general, are non-stationary in nature, this approach is deemed to undermine the achievable quality besides being computationally expensive. A better way is to apply the Discrete Wavelet Transform (DWT) by exploiting its time-frequency localisation property to capture the essence of non-stationary nature of the real-life signals. However DWT is computationally intensive and thus its direct mapping in hardware is not suitable for such resource constrained applications. Therefore without sacrificing its algorithmic accuracy designing its corresponding low-complexity architecture is a potential research challenge which would be covered later in this chapter.

Separation of a target signal from a mixture of several signals can be categorized under the classical BSS problem. Typical BSS algorithms like ICA require complicated mathematical operations in terms of divisions, multiplications, square root evaluations resulting in significant power consumption and silicon area. Therefore although effective for traditional off-line workstation based processing these algorithms pose big problem from energy consumption point of view for resource constrained systems like the personalized remote health monitoring system. Therefore design of an efficient but low-complexity BSS scheme and corresponding maximally power efficient architecture for the same are of utmost significance. In this chapter, as a part of well known ICA algorithms, the concept of FastICA will be explored along with its low-complexity architecture formulation which will be shown to reduce the hardware complexity of the conventional FastICA significantly.

4.2 Fundamentals of Biomedical Signal Processing

This section begins with a review on the evolution of the concept of WT from the Fourier Transform and explains its applicability in healthcare. WT is not a new technique, it is more than two decades ago when J. Morlet and A. Grossman proposed the concept of wavelet analysis to reach automatically the best trade-off between time-frequency resolution and developed the mathematical foundation of wavelets. Later this proposition has been considered as the extension of the ideas given by Haar in 1910 and Gabor in 1946. As any discovery in science, wavelets resulted from numerous contributions, they are based on the concepts that already existed before this proposed idea Truchetet and Laligant (2004). But at that stage wavelets were still very much in the realms of pure mathematics and concentrated more on the theory rather than the application (Bentley and McDonnell 1994). Only a handful of scientific papers used to be published each year and that was also mainly by the mathematics community (Addison 2004). In Strang (1989) gives the basic introduction of wavelets and dilation equation. But Mallat, Daubechies and Meyer changed this by defining the connection between wavelets and digital signal processing. In Mallat (1999) multi-resolution signal decomposition using WT technique is proposed. In Rioul and Vetterli (1991) the relationship between the WT and its applicability in signal processing discussed in detail. The design aspects of wavelets and its theoretical detail have been combined and presented nicely in Vetterli and Herley (1992). Design and development of fast algorithm for both discrete as well as continuous time WT has been depicted in Rioul and Duhamel (1992). In Rioul (1993) discrete time multi-resolution theory of WT is proposed which came out to be pioneering work in the domain of wavelet architecture design in digital signal processing research.

Approximation using superposition of functions has existed since the early 1800s, when Fourier discovered that he could superpose sines and cosines to represent other functions. Let us consider a signal, $x(t) = \cos(2\pi 10t) + \cos(2\pi 25t) + \cos(2\pi 50t) + \cos(2\pi 100t)$. This is an example of a stationary signal because it has frequencies of 10, 25, 50 and 100 Hz. For this stationary signal, Fourier transform provides the necessary frequency information. But, on the contrary, now let us consider another signal having four different frequency components at four different time intervals; 100 Hz sinusoid within 0–300 ms interval, 50 Hz sinusoid in 300–600 ms, 25 Hz within 600–800 ms interval and 10 Hz sinusoid within 800–1, 000 ms interval. This signal is a representative of non-stationary signal. In this case Fourier transform fails to localize frequency with its respective time of occurrence. Here comes the advantage of wavelet analysis where the fundamental idea is to analyze the data at different scales and resolutions. If a signal is looked with a large "window", its gross features would come out, whereas with a small "window", the small features can be noticed. The wavelet analysis procedure is to adopt a wavelet prototype function, called *analyzing wavelet* or *mother wavelet*. Temporal analysis is performed with a contracted, high frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low frequency version of the same wavelet. Because the

original signal or function can be represented in terms of a wavelet expansion i.e. using coefficients in a linear combination of the wavelet functions and data operations can be performed using just the corresponding wavelet coefficients (Graps 1995).

Considering $s(t)$ be the time-domain description of the incoming signal, $\psi(t)$ be the mother wavelet, a and b be the scale and time-shift variable respectively, WT can be given as (Bentley and McDonnell 1994; Rioul 1993):

$$WT(a,b) = a^{-1/2} \int s(t)\psi^* \left(\frac{t-b}{a} \right) \mathrm{d}t \tag{4.1}$$

where " $*$ " denotes the "complex conjugate" notation. From (4.1), it can be seen that the WT performs decomposition of signal $s(t)$ into a weighted set of basis functions which are scaled version of the mother wavelet $\psi(t)$ (Bentley and McDonnell 1994). Decomposition into these basis functions can be seen as time varying spectral analysis in which scale a plays a role of local frequency: with the increase of a, wavelets are stretched and analyze low frequencies and so on (Rioul 1993). In Cohen and Kovacevic (1996) a clear and comprehensive mathematical background of wavelets is provided. A comparative study of classical Fourier analysis and comparatively new wavelet analysis has been presented in Walker (1997). The necessary theoretical detail of WT has been well depicted in book (Daubechis 1992) and the applications of WT in different signal processing tasks has been well explained in Chui (1992). Typically DWT and Inverse DWT (IDWT) are computed using Analysis and Synthesis filter banks respectively as shown in Figs. 4.1 and 4.2 for four levels of resolution. In Analysis Bank the original discrete time sequence is decomposed into coarser half resolution approximation level with a half-band low-pass and high-pass filters. In dyadic space, this analysis filtering scheme at each stage can be depicted as (Chen et al. 1996; Parhi and Nishitani 1993):

$$w_k^j = \sum_n x_n h_{2k-n}, \qquad n,k \in Z \tag{4.2}$$

$$r_k^j = \sum_n x_n g_{2k-n}, \qquad n,k \in Z \tag{4.3}$$

where, x_n is input discrete time sequence to the filter, h_n and g_n are high-pass and low-pass filter impulse response respectively; and w_k^j and r_k^j are their corresponding outputs (wavelet and residue) for jth stage. The relationship between these two filters is given as (Chen et al. 1996):

$$h_{L-1-n} = (-1)^n g_n \tag{4.4}$$

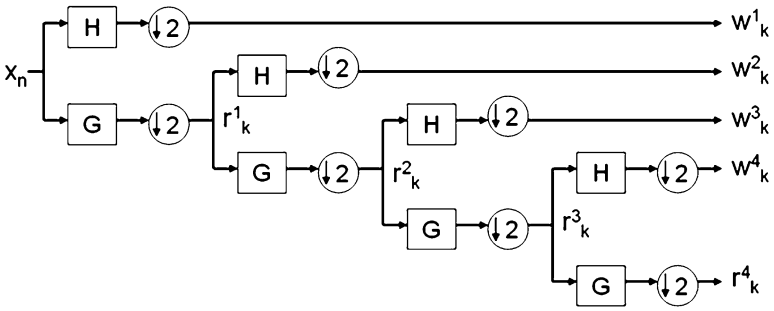


Fig. 4.1 Four-resolution level Analysis Bank for DWT

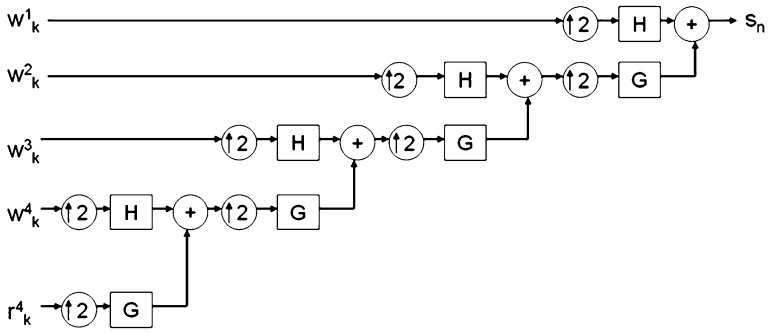


Fig. 4.2 Four-resolution level Synthesis Bank for IDWT

where, L is filter length and $(-1)^n$ is the modulation parameter which transforms low-pass filter into high-pass one. Similarly, IDWT at each resolution level of synthesis bank can be described as (Chen et al. 1996; Parhi and Nishitani 1993):

$$s_n = \sum_k [w_k^j h_{n-2k} + r_k^j g_{n-2k}], \quad n, k \in Z \tag{4.5}$$

where s_n is the reconstructed data output from the Synthesis Bank.

WT technique differs from the traditional Fourier transform by the way in which it localizes the information in the time-frequency plane; in particular, it is capable of trading one type of resolution for the other, which makes it suitable for analysing non-stationary signals. Most biomedical signals are examples of non-stationary signals and they typically have highly complex time-frequency characteristics. Frequently, they consist of brief, high-frequency components closely spaced in time, accompanied by long-lasting, low-frequency components closely spaced in frequency. Any appropriate analysis method for dealing with them should therefore exhibit good frequency resolution along with fine time resolution—the first to localize the low-frequency components, and the second to resolve high-frequency components. One application, where these properties have been found to be

relevant is healthcare. Due to the wide variety of signals and problems encountered in medicine and biology, the spectrum of applications of WT has been extremely large. It ranges from the analysis of the more traditional physiological signals such as Electrocardiogram (ECG) to the recent imaging modalities including Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) (Aldroubi and Unser 1996; Unser and Aldroubi 1996).

In practice, it is often possible to treat non-stationary signals as stationary ones by dividing them into blocks of short, pseudo-stationary segments—that is, segments whose statistics remain essentially unchanged for their duration. This approach appears to be valuable in many applications and popularly known as Short Time Fourier Transform (STFT). However, it also has one potential disadvantage: if the time-domain analysis window is made too short, frequency resolution will suffer leading to scenarios where frequency components belonging to different time windows start overlapping with each other. Lengthening of the window size, on the other hand, may invalidate the assumption of stationarity within the window. So for many practical signals such as physiological signals, STFT may not be a very suitable candidate for such time-frequency analysis. An alternate way to analyze non-stationary biomedical signals is to expand them onto basis functions created by expanding, contracting, and shifting the mother wavelet, specifically selected for the signal under consideration (Akay 1997; Addison 2006).

Figure 4.3 illustrates the shortcomings of STFT analysis in detecting signal features of short duration. This figure contains a scalogram[1] and a spectrogram[2] corresponding to the rhythmic ECG signal shown at the top of the figure. The time domain signal is derived from a pig heart which has been shocked several times. The spectrogram is generated from an STFT employing a 3.4s Hanning window typical for the analysis of this type of signal. The smearing and hence loss of local information across the spectrogram over these time scales is evident in the plot. The main difficulty in dealing with biomedical objects is the extreme variability of the signals and the necessity to operate on a case by case basis. Often, one does not know a priori what is the pertinent information and/or at which scale it is located. For example, it is frequently the deviation of some signal feature from the normal that is most relevant information for diagnosis. As a result, the problems tend to be less well defined than those in engineering and the emphasis is more on designing robust methods that work in most circumstances, rather than procedures that are optimal under very specific assumptions. Another important aspect of biomedical signals is that the information of interest is often a combination of features that are well localized temporally or spatially (e.g. spikes and transients in Electroencephalograph (EEG) signals and micro-calcifications in mammograms) and others that are more diffuse (e.g. small oscillations, bursts and texture). This requires the

[1] Distribution of energy i.e. square modulus of the wavelet coefficients of the signal in time-scale plane (Rioul and Vetterli 1991).

[2] Distribution of energy i.e. square modulus of the Fourier coefficients in time-frequency plane (Rioul and Vetterli 1991).

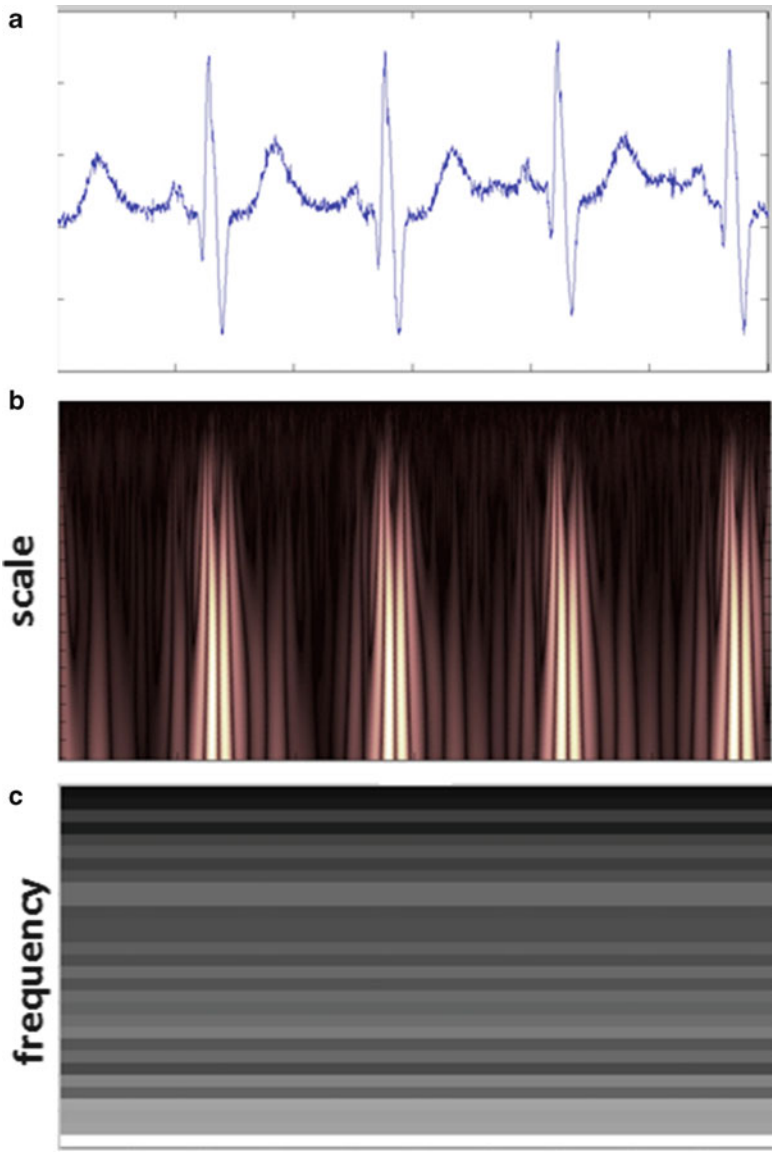


Fig. 4.3 Wavelet scalogram versus STFT spectrogram for rhythmic signal. (**a**) original time-domain rhythmic ECG signal (X-axis: time, Y-axis: amplitude). (**b**) Morlet based scalogram corresponding to (**a**) (X-axis: time, Y-axis: Scale as defined by a in (4.1)). (**c**) Spectrogram corresponding to (**a**) generated using STFT with a 3. 4s Hanning window (X-axis: time, Y-axis: frequency)

analysis methods sufficiently versatile in terms of their time-frequency localization (Unser and Aldroubi 1996). WT can provide both good time resolution at high frequencies and good frequency resolution at low frequencies. This excellent combination of time-frequency resolution makes wavelets potentially invaluable in numerous applications including healthcare. To name a few, healthcare application includes the early discovery of the precursors of heart disease, studies of fetal breathing, fetal electrocardiogram detection, heart rate variability, extraction of speech from background noise in digital hearing aids, the detection of breast cancer, medical image compression and many more (Akay 1997).

It has been pointed out in last section that in remote healthcare, signal separation is one challenging task. In a large number of cases, the signal received by a sensor is actually the mixture of several signal sources. For example, the signal received by an ECG sensor is a superposition of ECG, noise and artefacts. Generally, sources as well as mixtures are unknown and these are estimated without any prior knowledge of the sources—this problem is called "Blind Source Separation" (BSS) (Jutten and Herault 1991). The lack of prior knowledge about the mixture is compensated by a statistically strong but often physically plausible assumption of independence between the source signals (Cardoso 1998). ICA is a concept which is employed for solving this BSS problem (Roberts and Everson 2001). ICA has very well been formulated in Cichocki and Amari (2002) and Stone (2004) and also its various applications are discussed. In Cichocki (1999) the importance of ICA in brain science has been discussed in detail. Thrust has been given there on developing novel and advanced ICA algorithm to analyze multi-sensory biomedical signals, especially electroencephalographic, magnetoencephalographic (MEG) and electromyographic (EMG). Novel and innovative techniques in BSS are required in Brain Science since brain sources are unidentified, extremely weak, non-stationary, distorted by noise and other interferences. In Pourazad et al. (2005) and Chien et al. (2006) the application of ICA in healthcare has been described considering one important task in biomedical signal processing—separating heart sound from lung sound. Acoustical analysis of lung sounds provides important and helpful information in the diagnosis and monitoring of lung diseases. But, for the cardiologists, the main interest is in heart sound and lung sound is considered as noise. So, both of these signals are of interest for different set of medical practitioners. Lung and heart sounds are considered as independent source signals. So, separating the mixed signals recorded on the skin is of utmost importance. In Pourazad et al. (2005), the applicability and feasibility of ICA for such signal processing problem has been studied and in Chien et al. (2006), ICA has been applied successfully for separating these two signals. According to Calhoun and Adali (2006), ICA has recently demonstrated considerable promise in characterizing functional Magnetic Resonance Imaging (fMRI) data. ICA has been successfully used in a number of exciting fMRI applications, including the identification of various signal types such as task and transiently task-related and physiology related signals in the spatial and temporal domain (Calhoun and Adali 2006). ICA especially FastICA can be applied to the feature extraction of pulse wave of human being (Aoyagi et al. 2002). Pulse wave is the rhythmic change

of the inflow of blood to the different parts of the body in the form of waves as measured from the outside of the body. This wave is modified by the physiological conditions such as the heart beat, the circulation of the blood, and changes in the state of the minor artery system, which leads to the distortion of the shape of the waves (Aoyagi et al. 2002). FastICA can be successfully applied to remove out these distortions and get back the original pulse wave. ICA is also frequently applied in Electrocardiography (Lo et al. 2006). Routinely recorded ECG are often corrupted by different types of artefacts and many methods have been made to enhance their quality by reducing the noise and artefacts (He et al. 2006). ICA has shown tremendous potential for artefacts removal and has been successfully used in He et al. (2006), Barros et al. (1998). ICA can also be applied to separate spinal cord motor signals (Tie and Sahin 2005). ICA has also got potential applications in detecting fetal arrhythmias from fetal Magnetocardiographs (fMCG) obtained by ultra-sonographic imaging (Comani et al. 2004). Besides, other applications of ICA algorithm has been well depicted in Hyvärinen and Oja (2000) and part IV of the book (Hyvärinen et al. 2001) with necessary minute detail including feature extraction, brain imaging, telecommunication and multi-user detection.

Application and quality assessment of different available ICA algorithms have been done in Mutihac and Mutihac (20007) on EEG recordings to separate the independent source signals as well as to remove the possible artifacts in those recordings. A comparative study of different ICA algorithms—JADE, BS Infomax, FastICA and EGLD ML—has been performed in terms of absolute activation correlations, variance in the reconstructed signals, absolute correlation between source signals and corresponding ICA estimated activations, Signal to Interference Ratio (SIR), Cross Talk Error (CTE) and CPU processing time (Mutihac and Mutihac 20007). Depending upon all these figures of merit, the performance of these algorithms has been classified and it has been found that FastICA is the superior of them all (Mutihac and Mutihac 20007) and it outperforms most of the existing ICA algorithms in terms of higher convergence speed (Oja and Yuan 2006). Therefore we will limit our discussion here only in FastICA. Denoting independent sources by S, mixed signal (X) can be defined as Shyu et al. (2008):

$$X = AS \tag{4.6}$$

where, n = number of independent sources, $X = \{x_i\}$, $S = \{s_i\}$, $i \in (1, n)$; A is a full-rank $n \times n$ mixing matrix; $s_i = \{s_{i,j}\}$, $x_i = \{x_{i,j}\}$ where $j \in (1, m)$ and m is equal to the frame-length. FastICA comprises of two steps—Preprocessing and FastICA Iteration. Preprocessing can further be divided into to steps—Centering and Whitening. In Centering, mean vector of $x_{i,j}$ is computed and then subtracted from $x_{i,j}$ to make the mixed signal a zero-mean matrix. It can be represented as (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$x_{i,j} = x_{i,j} - \mathscr{E}[x_{i,j}] \tag{4.7}$$

where $\forall i \in (1, n)$, $j \in (1, m)$. The next preprocessing step is whitening which linearly transforms vector X into a new vector Z which is white, i.e., its components are uncorrelated and their variances are equal to unity. One way to do this is Eigen Value Decomposition (EVD) which can be defined as (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$\mathscr{E}[XX^T] = C_X = EDE^T \tag{4.8}$$

where C_X is the covariance matrix of X; $E = \{e_i\}$ is the orthogonal matrix of eigenvectors e_i of C_X and $D = diag(\{d_i\})$ is the diagonal matrix of its eigenvalues $\forall i \in (1, n)$. The whitening process can be described by Roberts and Everson (2001), Hyvärinen and Oja (2000):

$$Z = PX \tag{4.9}$$

where $Z = \{z_i\}$ and P is the whitening matrix and can be defined as (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$P = D^{-1/2} \times E^T \tag{4.10}$$

The utility of whitening resides in the reduction of number of parameters to be estimated (Hyvärinen and Oja 2000; Roberts and Everson 2001). The FastICA algorithm produces the estimated output vector ($\hat{S}$) from the whitened mixed signal vector (Z) by estimating the unmixing matrix B of dimension $n \times n$ which can be defined as (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$\hat{S} = B^T Z \tag{4.11}$$

The k^{th} column of B represents the weight vector w_k associated with k^{th} estimated IC where $k \in (1, n)$. FastICA algorithm for estimating n ICs can be given as follows (Hyvärinen and Oja 2000; Roberts and Everson 2001):

(i) Choose n, the number of ICs to estimate and set counter $k \leftarrow 1$.
(ii) Choose an initial value of unit norm of $\mathbf{w_k}$.
(iii) FastICA introduces the optimized contrast function (g) within the basic iterative equation as follows:

$$\mathbf{w_k} \leftarrow \mathscr{E}\{Z\dot{g}(\mathbf{w_k}^T Z)\} - \mathscr{E}\{\ddot{g}(\mathbf{w_k}^T Z)\}\mathbf{w_k} \tag{4.12}$$

where $\dot{g}$ and $\ddot{g}$ are the first and second derivative of contrast function g respectively. Using Kurtosis-based contrast function in (4.12) we get (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$\mathbf{w_k} \leftarrow \mathscr{E}\{Z(\mathbf{w_{k-1}}^T Z)^3\} - 3\mathbf{w_{k-1}} \tag{4.13}$$

(iv) To prevent different weight vectors w_k from converging to the same maxima, the next step is to orthogonalize the vectors after every iteration. Considering the *deflationary orthogonalization* and following Gram–Schmidt orthogonalization, this step can be defined as (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$\mathbf{w_k} \leftarrow \mathbf{w_k} - \sum_{j=1}^{k-1} (\mathbf{w_k}^T \mathbf{w_j}) \mathbf{w_j} \tag{4.14}$$

(v) Normalize w_k as follows (Hyvärinen and Oja 2000; Roberts and Everson 2001):

$$\mathbf{w_k} \leftarrow \mathbf{w_k} / \|\mathbf{w_k}\| \tag{4.15}$$

(vi) If $\mathbf{w_k}$ has not converged, go back to step (iii), if converged go to next step.
(vii) Set $k \leftarrow k + 1$. If $k \leq n$, go back to step (ii).

Now using computed B in (4.11) estimated output $\hat{S}$ is obtained. The detailed procedure and corresponding block diagrams can be found in Shyu et al. (2008) and Hyvärinen (1999). Now in the next section we will focus on designing the low-complexity architectures of these algorithms so that these can be used in remote healthcare.

4.3 Constraints for Applying Traditional Signal Processing Techniques in Remote Healthcare

We have already seen in the last section that the DWT is of great research interest in remote healthcare application. Over last two decades numerous architectures have been proposed in literatures for implementing Forward and Inverse DWT (IDWT). First such architecture was designed by Knowels (1990). Parhi and Nishitani have proposed folded and digit-serial architectures for One Dimensional DWT (1-D DWT) (Parhi and Nishitani 1993). Grzeszczak, Yeap and Pachanathan proposed the DWT architecture with less number of multipliers than the existing methods (Grzeszczak et al. 1994). Fridman and Manolakos came up with the first fastest parallel architecture of 1-D DWT (Fridman and Manolakos 1994). Then onwards various implementation schemes have been proposed (Alam et al. 2003; Cao et al. 2006; Chen et al. 1996; Denk and Parhi 1997; Grzeszczak et al. 1996; Jou et al. 2001; Martina and Masera 2006, 2007; Oweiss et al. 2007; Vishwanath et al. 1995; Yu et al. 1997) and improvement of architectures has been made over the period of time. In general these implementations have been influenced by certain characteristics such as computational time and complexity, regularity, smoothness and linear phase characteristics, transform efficiency as well as perfect reconstruction of the decomposed signals (Oweiss et al. 2007). But, most

of the reported hardware approaches, though focuses on improving computational speed, does not adequately address severe power and area constraints issues (Oweiss et al. 2007); which often are the two most important metrics in today's high performance signal processing systems. As for example sophisticated applications such as remote healthcare system require light-weight signal processing with particular emphasis on low power and low area due to scarcity of resources. So, design of optimal Application Specific Integrated Circuit (ASIC) tailored to such applications consisting of DWT/IDWT demands power and area efficient implementation of DWT algorithms.

The main power consuming operation in DWT/IDWT computation is the filtering which requires significant number of multiplications (Alam et al. 2003). Distributed Arithmetic (DA) can be adopted for eliminating the requirement of multiplications (Alam et al. 2003) which may lead to reduction of power consumption. However in the conventional DA based approach one needs to store all the possible combinations of filter coefficients in a memory which increases exponentially in size with the frame-length (Alam et al. 2003). Thus for longer frame-length the advantage of using DA may eventually be lost owing to the significant increase in memory size. Here a recent methodology is mentioned for memory reduction in DA based design of DWT/IDWT architecture by exploiting its inherent algorithmic symmetry resulting in data repeatability. Subsequently a 16-bit fixed-point DWT/IDWT architecture is developed for frame-length 16, which requires less silicon area and power consumption compared to some of the available DWT/IDWT architectures. This architecture will be independent of any specific mother wavelet and thus can be reconfigured for any type of wavelet function. DA was proposed more than two decades ago and has since used widely in VLSI implementation of Digital Signal Processing (DSP) architectures (Alam et al. 2003; Bellaouar and Elmasry 1995; Burleson 1990; Cao et al. 2006; Taylor 1986). Although DA based methods are advantageous, but in the context of convolution-based filtering such as DWT/IDWT (as an example, 9/7—5/3 wavelet vlsi architecture (Martina and Masera 2006, 2007)), its memory requirement has exponential relationship with the incoming data-frame length (Alam et al. 2003). Here we will demonstrate that it is possible to restrict this exponential growth in memory requirement through pre-computing carefully various combinations of filter coefficients. DA is a bit serial computational operation which is used to compute the inner product of a pair of vectors without using multipliers. The basic DA equation is given as follows (Bellaouar and Elmasry 1995; Burleson 1990; Taylor 1986):

$$x_n = -x_{n,l}.2^l + \sum_{b=0}^{l-1} x_{n,b}.2^b \tag{4.16}$$

where $l =$ (total number of bits/sample). In dyadic space, a convolution based wavelet filter can be represented as:

$$w_a = \sum_{n=0}^{3} x_n h_{2a-n}, \tag{4.17}$$

Table 4.1 Conventional and reduced memory requirement to store combinations of filter coefficients in DA for data-frame length = 4 and 4 bits/sample case

Address	Data	Address	Data	Address	Data
0000	0	0001	h_2	0010	h_1
0011	$h_2 + h_1$	0100	h_0	0101	$h_0 + h_2$
0110	$h_0 + h_1$	0111	$h_0 + h_1 + h_2$	1000	0
1001	h_2	1010	h_1	1011	$h_2 + h_1$
1100	h_0	1101	$h_0 + h_2$	1110	$h_0 + h_1$
1111	$h_0 + h_1 + h_2$	—	—	—	—
00	h_0	01	h_2	10	h_1
11	$h_1 + h_2$	—	—	—	—

where x_n and h_n are input samples and filter coefficients respectively. Considering frame-length = 4 and wordlength = 4, (as an example) and using (4.16) in (4.17) and assuming $a = 1$,

$$w_1 = -[x_{03}h_0 + x_{13}h_1 + x_{23}h_2]2^3 + \ldots\ldots + [x_{00}h_0 + x_{10}h_1 + x_{20}h_2]2^0 \quad (4.18)$$

where $x_{i\,j}$ = ith sample jth bit of the input data. The possible combinations of filter coefficients obtained from (4.18) are shown in the first 6 rows of Table 4.1 which occupies 16 memory locations. However, it can be observed from Table 4.1 that there exists redundant (such as "0") and repetitive filter coefficients (such as "h_0", "h_1", "h_2", "$h_1 + h_2$", "$h_0 + h_1$", "$h_0 + h_2$", "$h_0 + h_1 + h_2$") occupying more than a single memory location. Thus if only the unique combinations of the filter coefficients are stored in the memory the other filter coefficients can be obtained on-the-fly using simple addition operation. In this particular example, the proposed methodology leads to only four memory locations, as shown in the last two rows of Table 4.1, rather than 16 locations in the conventional approach.

Considering wavelet computation in dyadic space for frame-length (p) = number of filter coefficients; and assuming ith level of resolution consists of j number of filter coefficients, the following relationships hold: $\forall i \in [1, \log_2 p]$, $j \in [1, p/2^i]$, where i and j are integers. Now if the jth coefficient consists of s number of data samples, considering the causality of the system, s can be represented as:

$$s = 1 + (j-1)2^i \quad (4.19)$$

The maximum number of samples (s_{max}) present at the ith level of resolution can be obtained by substituting $j = p/2^i$ in (4.19) which leads to :

$$s_{max} = 1 + p - 2^i \quad (4.20)$$

It is to be noted from (4.20) that ($s_{max} - 1$) is an even number. In this methodology, ($s_{max} - 1$) number of samples is divided into k sub-frames. Each of the sub-frames

consists of a pair of data samples x_{n-1} and x_n. As shown later in this section, this approach will lead to reduction in memory. k can be represented as:

$$k = (p/2) - 2^{(i-1)} \tag{4.21}$$

At each time instant the binary value of m^{th} bit of x_{n-1} and x_n are checked and depending on their combination the appropriate linear combination of the filter coefficients is fetched from the memory (as can be obtained by expanding (4.17)). However, one of these combinations is "00" which means no filter coefficient needs to be multiplied with the input data. As an example in (4.18) if "$x_{1,3}x_{0,3}$" = "00" then multiplication of these input bits with h_1 and h_2 will always yield 0. Thus one needs to store the combinations of filter coefficients corresponding to only three bit-combinations of the sample pair (namely "01", "10" and "11"). Arranging $(s_{max} - 1)$ samples from s_{max} samples at the i^{th} resolution level, we are left with only one sample. Since m^{th} bit of this sample can assume either "0" or "1", only one combination of filter coefficients needs to be stored in the memory corresponding to this sample. Thus, the total memory requirement (M_i) in this methodology for ith level of resolution can be given by:

$$M_i = k \times 3 + 1 = [(p/2) - 2^{(i-1)}] \times 3 + 1 \tag{4.22}$$

The memory requirement for the analysis bank can be computed by summing M_i for all i. But, at the last resolution level of the analysis bank, all is left is one residue signal along with the wavelet coefficient. This residue signal consists of one sample for which, as discussed above, two combinations of filter coefficients are possible in which one is "0". Therefore, the combination of filter coefficients corresponding to this residue signal occupies one memory block only. Thus, according to this methodology, the Total Memory Requirement (*TMR*) can be represented as:

$$TMR = 1 + \sum_{i=1}^{\log_2 p} M_i \tag{4.23}$$

From (4.23) it is evident that *TMR* grows nearly linearly with the increase of frame-length. Since the filter coefficients used in this methodology are generic, these can be reconfigured to generate any type of mother wavelet and thus can be used for any kind of WT architecture design.

The block diagram of the DWT/IDWT architecture is shown in Fig. 4.4. At the block level the architecture is similar to a standard DA based architecture. However, the main novelty of the architecture lies in the formulation of a new addressing scheme and the corresponding address generation unit design for the reduced memory unit (shown as Generic Memory in Fig. 4.4). In a typical DA based approach the bits of the input samples are checked and addresses are generated to fetch data from the Memory. These data are then used in the Analysis Bank to compute DWT by applying (4.18) in (4.2) and (4.3). Similarly, IDWT is computed

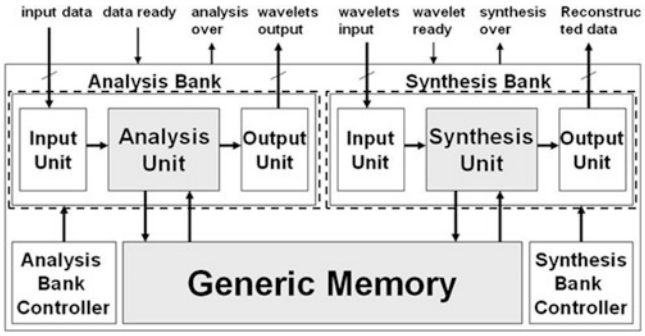


Fig. 4.4 VLSI architecture of DWT and IDWT

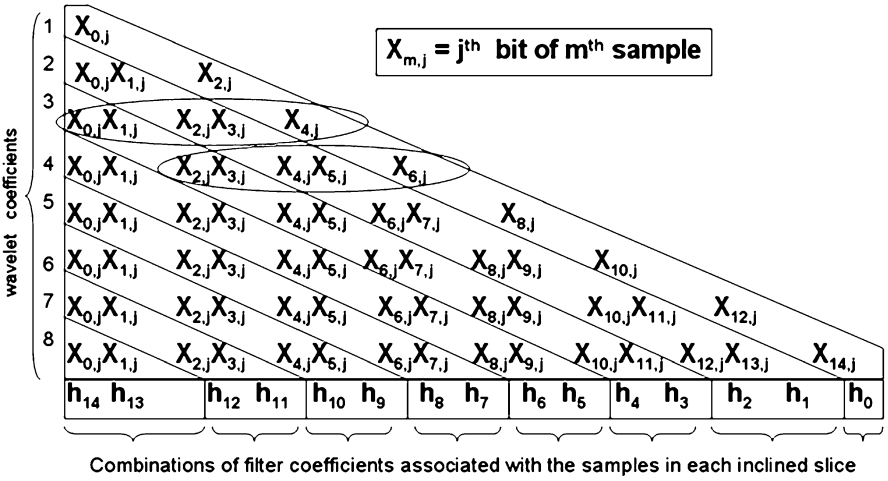


Fig. 4.5 Example of address generation logic for 1st resolution-level wavelet coefficients with frame-length 16

in the Synthesis Bank by applying (4.18) in (4.5). The memory unit of this architecture only the non-repetitive combinations of filter coefficients are stored from the set of data. The strategy for address generation is explained here with an example of frame-length 16, for the 1^{st} resolution level of the analysis bank.

Figure 4.5 shows the incoming data-samples and the associated filter coefficients required for computing each wavelet coefficient in this case. It can be noted from Fig. 4.5 that the filter coefficients are the same for different samples present in the same inclined slices. For example, as shown in the encircled regions in Fig. 4.5, the filter coefficients associated with the samples for computing the 3^{rd} wavelet coefficient are the same for the last five samples of the 4^{th} wavelet coefficient computation. Mathematically, the filter coefficients associated with the q^{th} data sample of r^{th} and $(q + 2)^{th}$ data sample of $(r + 1)^{th}$ wavelet coefficients are the same (symmetry

Table 4.2 Example of address generation logic samples x_2 and x_3 of 3rd wavelet coefficient in 1st resolution level

$x_{2,j}$	$x_{3,j}$	Fetch from Memory
0	0	No fetch
0	1	h_1
1	0	h_2
1	1	$h_1 + h_2$

property). This symmetry can be represented in generalized form for *i*th level of resolution as follows:

$$x_m^j = x_{m+2^i}^{j+1} \tag{4.24}$$

where, m = sample number and j = wavelet coefficient. While designing the address generation logic, this symmetry has been exploited at different levels of resolution. Table 4.2 shows how the appropriate filter coefficients are fetched from the memory for different combinations of j^{th} bit of samples x_2 and x_3 for the 3rd wavelet coefficient computation.

The Analysis and Synthesis banks are designed following the equations given in Parhi and Nishitani (1993) and Chen et al. (1996), which represent wavelet analysis and synthesis in dyadic space. DA based filter design technique has been applied to realize the FIR filters of these blocks which replaces multiplication operations by additions only and thereby making the entire architecture multiplier-less. In addition; there are two control units responsible for controlling the analysis and synthesis banks. As shown in Fig. 4.4, two active-low single bit signals—"data ready" and "wavelet ready" enables the Analysis Bank and Synthesis Bank respectively. These signals may be enabled by the user or the previous processing unit in case of a complete system. When the output of these banks are ready, the respective controllers produce two active-low single bit signals—"analysis over" and "synthesis over" which indicates the completion of the process. However, reducing memory at the expense of adder arises one important issue. The hardware savings obtained due to reduction in memory size, can be negated if the total number of adders used in the design is more than a certain limit. It will be discussed in the next section where we will deal with the computation of hardware complexity.

Now let us turn our attention into the architecture formulation for FastICA algorithm. Most of the reported articles concentrated on off-line computation of FastICA. But, the problem in that case, such methods can not be applied in real time applications including remote healthcare. So architectural studies would be ideal for the hardware implementation of this algorithm (Du et al. 2007) so that it can be applied in such real time signal processing applications. In Du et al. (2007), it has been identified that VLSI implementation of ICA algorithm requires extremely efficient hardware designs and sufficient IC resources and it has been anticipated that due to the rapid growth in VLSI industries, significant solutions of complex ICA algorithms will be possible. However there have been very few studies about the real-time implementation of ICA in hardware and so there was no significant

reported research articles on VLSI implementation of FastICA. One reason behind it may be FastICA in its present form requires complex mathematical operations including division, square root evaluation and multiplication and thus demands large silicon area and power consumption and thus is not suitable for resource constrained applications like mobile healthcare because the body-worn sensors are typically battery powered, unobtrusive, tiny and are expected to operate for long time. Recently, the first VLSI implementation of FastICA based on floating-point arithmetic has been reported in Shyu et al. (2008). This is direct algorithm into architecture mapping and such implementation contributes significantly to the silicon area and consume significant power. Thus fixed-point arithmetic may be an optimal choice for low complexity hardware design of FastICA although a compromise with accuracy may be necessary. Although the FastICA algorithm has received tremendous attention over the last decade, still potential gaps are there in algorithm and architecture holistic optimization approach.

we will perform our initial research study on designing a low complexity fixed-point arithmetic based VLSI architecture for FastICA algorithm and show that it will consume low power and occupy small silicon area. We have already discussed in the last section that FastICA algorithm consists of two main blocks—Preprocessing and FastICA (Hyvärinen 1999). The preprocessing has two units—Centering and Whitening and their operation is to make the incoming data a zero mean variable and decorrelate these zero mean signals. The FastICA block estimates the final output vector. Centering requires add and accumulation unit whereas whitening needs computation of eigen values and the corresponding eigen vectors of the covariance matrix formed using the mixed sensory data set (Hyvärinen 1999). Coordinate Rotation Digital Computer (CORDIC) is the widely used technique for implementing Eigen Value Decomposition (EVD) in hardware (Bravo et al. 2006, 2008; Brent et al. 1985; Cavallaro and Luk 1988; Götze and Hekstra 1995; Götze et al. 1993; Hsiao and Delosme 1996). The main FastICA Iterative steps consists of costly arithmetic operations involving division, square root evaluation and multiplications. The direct implementation of these operations increases the hardware complexity and contribute to high power consumption. Recently an architecture for low-complexity implementation of 2D FastICA was proposed in Shyu et al. (2008) and Acharyya et al. (2009b). However such architectural modifications are obtained using algebraic method which can not be extended in higher dimensions where $n > 2$. Therefore the FastICA architectures discussed in Shyu et al. (2008) and Acharyya et al. (2009b) can not be generalized for nD cases where $n \geq 2$. Now, as recently shown in Acharyya et al. (2010) and Acharyya et al. (2011), we aim to make an attempt to merge the whitening stage with the FastICA Iterative step so that the same hardware unit can be reused for both of these stages. Since, CORDIC is popularly known technique for implementing EVD, the research goal in this chapter is to explore how, if at all possible, the concept of co-ordinate rotation can be applied in this context for implementing the FastICA Iterative step in a generalized sense. CORDIC is an efficient implementation technique for vector rotation and arctangent computation and since it can be realized using simple shift and add operations, it is effective in

terms of low hardware complexity (Hu 1992; Volder 1959; Walther 1971). CORDIC, in general, can be operated in two modes—rotation and vectoring (Hu 1992). In rotation mode, given the angle of rotation and the initial vector, final vector is computed and in the vectoring mode, the angle between the initial vector and the principal coordinate axis is computed. Considering the rotation in clockwise sense, the basic CORDIC expressions can be expressed as (Hu 1992; Volder 1959; Walther 1971):

$$\begin{bmatrix} x_f \\ y_f \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{4.25}$$

where x_0, y_0 and x_f, y_f are the initial and final components of the vector and the angle of rotation is θ. In Rotation mode, angle θ is approximated by a sequence of micro-angles and after a finite number of iterations x_f and y_f are generated. In vectoring mode, where angle θ is unknown, y_f is forced to zero after finite number of iterations. In the case under consideration for 2*D* FastICA (i.e. $n = 2$), expanded form of (4.13) can be written as:

$$\begin{bmatrix} {w_{1,1}}^{(p+1)} \\ {w_{1,2}}^{(p+1)} \end{bmatrix} = \begin{bmatrix} \mathscr{E}\left[z_{1,j}\left\{z_{1,j}{\underline{w}_{1,1}}^{(p)} + z_{2,j}{\underline{w}_{1,2}}^{(p)}\right\}^3\right] \\ \mathscr{E}\left[z_{2,j}\left\{z_{1,j}{\underline{w}_{1,1}}^{(p)} + z_{2,j}{\underline{w}_{1,2}}^{(p)}\right\}^3\right] \end{bmatrix} - 3\begin{bmatrix} {\underline{w}_{1,1}}^{(p)} \\ {\underline{w}_{1,2}}^{(p)} \end{bmatrix} \tag{4.26}$$

where p denotes the number of iteration stage, $z_{i,\,j}$ represents the ith whitened data containing j number of samples where $i = \{1, 2\}$ and $j \in (1, m)$ where m denotes the frame-length, ${w_{1,\,q}}^{(p+1)}$ is the 1st column of the unmixing matrix after pth iteration where $q = \{1, 2\}$ and ${\underline{w}_{1,q}}^{(p)}$ indicates the *normalized value* of ${w_{1,\,q}}^{(p)}$ used in pth iteration as given by (4.15). Without any loss of generality, ${\underline{w}_{1,1}}^{(p)}$ and ${\underline{w}_{1,2}}^{(p)}$ can be considered as the components of a unit norm vector ${\underline{\mathbf{w}}_1}^{(p)}$. Now transforming from Cartesian to Polar co-ordinate system, ${\underline{\mathbf{w}}_1}^{(p)}$ can be re-written as:

$${\underline{\mathbf{w}}_1}^{(p)} = \left[{\underline{w}_{1,1}}^{(p)}\ {\underline{w}_{1,2}}^{(p)}\right]^T = \left[\cos\theta_p\ \ \sin\theta_p\right]^T \tag{4.27}$$

where polar angle $\theta_p = \tan^{-1}\left({\underline{w}_{1,2}}^{(p)}/{\underline{w}_{1,1}}^{(p)}\right)$ at pth iteration stage. Since the vector ${\underline{\mathbf{w}}_1}^{(p)}$ is of unit-norm, the magnitude of the radial vector is one and thus does not appear in (4.27). From (4.25) it can be written as:

$$x_f = x_0\cos\theta + y_0\sin\theta \tag{4.28}$$

Using (4.27) in (4.26) and considering the bracketed portion of (4.26) on which cubing has to be performed, a similarity can be found with (4.28) where $x_0 = z_{1,\,j}$, $y_0 = z_{2,\,j}$ and $\theta = \theta_p$ at the pth FastICA iteration stage.

Instead of using the complete expression shown in (4.25), one concise notation can be introduced as follows for brevity: $x_f = \mathscr{R}\ o\ t_x(x_0, y_0, \theta)$ and $y_f = \mathscr{R}\ ot_y$

(x_0, y_0, θ) where $\mathscr{R}\,ot_{x/y}(\cdot)$ denotes the x/y component of the rotation mode CORDIC outputs. Using this notation, (4.26) can be modified as:

$$\begin{bmatrix} w_{1,1}{}^{(p+1)} \\ w_{1,2}{}^{(p+1)} \end{bmatrix} = \begin{bmatrix} \mathscr{E}[z_{1,j}\{\mathscr{R}ot_x(z_{1,j}, z_{2,j}, \theta_p)\}^3] \\ \mathscr{E}[z_{2,j}\{\mathscr{R}ot_x(z_{1,j}, z_{2,j}, \theta_p)\}^3] \end{bmatrix} - 3\begin{bmatrix} \underline{w}_{1,1}{}^{(p)} \\ \underline{w}_{1,2}{}^{(p)} \end{bmatrix} \tag{4.29}$$

From (4.29) it can be seen that rotation mode CORDIC can be used in the main FastICA Iterative stage. However explicit information of θ_p is not available and has to be derived from $\underline{w}_{1,1}{}^{(p)}$ and $\underline{w}_{1,2}{}^{(p)}$ instead. From the line following (4.27) it has already been seen that θ_p is the *arctangent* of two components of the vector $\underline{\mathbf{w}}_1{}^{(p)}$ which is exactly the same as the CORDIC operating in Vectoring mode as given in (4.25). Again, introducing another concise notation for the sake of simplicity, θ_p can be denoted as $\mathscr{V}ec_\theta(\underline{w}_{1,1}{}^{(p)}, \underline{w}_{1,2}{}^{(p)})$ where the x and y-inputs of the vectoring mode CORDIC is denoted by $\underline{w}_{1,1}{}^{(p)}$ and $\underline{w}_{1,2}{}^{(p)}$ respectively. Using this notation, (4.29) can be modified as:

$$\begin{bmatrix} w_{1,1}{}^{(p+1)} \\ w_{1,2}{}^{(p+1)} \end{bmatrix} = \begin{bmatrix} \mathscr{E}[z_{1,j}\{\mathscr{R}ot_x(z_{1,j}, z_{2,j}, \mathscr{V}ec_\theta(\underline{w}_{1,1}{}^{(p)}, \underline{w}_{1,2}{}^{(p)}))\}^3] \\ \mathscr{E}[z_{2,j}\{\mathscr{R}ot_x(z_{1,j}, z_{2,j}, \mathscr{V}ec_\theta(\underline{w}_{1,1}{}^{(p)}, \underline{w}_{1,2}{}^{(p)}))\}^3] \end{bmatrix} - 3\begin{bmatrix} \underline{w}_{1,1}{}^{(p)} \\ \underline{w}_{1,2}{}^{(p)} \end{bmatrix} \tag{4.30}$$

Thus (4.30) represents the mapping of the conventional 2*D* FastICA Iterative stage in terms of the rotation and vectoring mode CORDIC.

Since the basic FastICA Iteration is not a norm-preserving operation, normalization of the newly obtained vector is necessary as pointed out in (4.15). This step can also be accomplished using the same CORDIC in the following way. Denoting the normalized components of the vector obtained after pth iteration using (4.26) by $\underline{w}_{1,1}{}^{(p+1)}$ and $\underline{w}_{1,2}{}^{(p+1)}$ the following equation holds:

$$\underline{w}_{1,i}{}^{(p+1)} = \frac{w_{1,i}{}^{(p+1)}}{\sqrt{|\mathbf{w}_1{}^{(p+1)}|^2}} \tag{4.31}$$

where $i = \{1, 2\}$. Using Cartesian to Polar Co-ordinate transformation following (4.27), these two components can be represented as:

$$\underline{\mathbf{w}}_1{}^{(p+1)} = [\underline{w}_{1,1}{}^{(p+1)}\ \underline{w}_{1,2}{}^{(p+1)}]^T = [\cos\theta_{(p+1)}\ \ \sin\theta_{(p+1)}]^T \tag{4.32}$$

Now, considering $x_0 = 0$ and $y_0 = 1$ in (4.25), it can be found that $x_f = \sin\theta$ and $y_f = \cos\theta$. Thus a similarity can be established between this (y_f, x_f) and (4.32) and following the same notations used in (4.29) and (4.30), (4.32) can be written as:

$$\begin{bmatrix} \underline{w}_{1,1}{}^{(p+1)} \\ \underline{w}_{1,2}{}^{(p+1)} \end{bmatrix} = \begin{bmatrix} \mathscr{R}ot_y(0, 1, \mathscr{V}ec_\theta(w_{1,1}{}^{(p+1)}, w_{1,2}{}^{(p+1)})) \\ \mathscr{R}ot_x(0, 1, \mathscr{V}ec_\theta(w_{1,1}{}^{(p+1)}, w_{1,2}{}^{(p+1)})) \end{bmatrix} \tag{4.33}$$

It is to be observed from (4.33) that $\underline{w}_{1,1}{}^{(p+1)}$ and $\underline{w}_{1,2}{}^{(p+1)}$ are the y and x output of the Rotation mode CORDIC respectively. After finite number of FastICA Iterations once the normalized vector converges, using (4.11), one independent component is estimated. Here it will be shown that unlike the conventional FastICA algorithm, complete array of multiplication operations can be removed by reusing CORDIC. Denoting the converged normalized vector by $\underline{\mathbf{w}}_1{}^c = [\underline{w}_{1,1}{}^c \ \underline{w}_{1,2}{}^c]^T$ where superfix "c" stands for *convergence*, denoting the estimated component by $\hat{s}_1 = \{\hat{s}_{1,j}\}$ where $j \in (1, m)$ and m is the frame-length and using the same set of arguments used to derive (4.29) and (4.30), (4.11) can explicitly be written as:

$$\begin{aligned} \hat{s}_{1,j} &= z_{1,j}\underline{w}_{1,1}{}^c + z_{2,j}\underline{w}_{1,2}{}^c \\ &= \mathscr{R}ot_x(z_{1,j}, z_{2,j}, \mathscr{V}ec_\theta(\underline{w}^c_{1,1}, \underline{w}^c_{1,2})) \end{aligned} \tag{4.34}$$

Relinquishing "c" from (4.34), $\hat{s}_{1,j}$ can be rewritten as:

$$\hat{s}_{1,j} = \mathscr{R}ot_x(z_{1,j}, z_{2,j}, \mathscr{V}ec_\theta(w_{1,1}, w_{1,2})) \tag{4.35}$$

where $w_{1,\,1}$ and $w_{1,\,2}$ are the components of the unnormalized vector obtained after FastICA iteration prior to convergence checking. Figure 4.6 presents the architecture of the co-ordinate rotation based 2*D* FastICA algorithm proposed in the last section. Since it has already been mentioned in the first section that CORDIC is widely used for EVD which is the fundamental operation in the preprocessing step of FastICA and since the proposed algorithm in the last section has already shown how CORDIC can also be used for implementing the main FastICA step, the controller design of the overall architecture becomes an important task to ensure proper data enters the CORDIC unit at the right time. As outlined in the last section, the controller operation can also be divided into three modes of operations—(*i*) *iteration*, (*ii*) *normalization* and (*iii*) *estimation*. These three modes can be used to generate *select* signals for the multiplexers and demultiplexer at the inputs and outputs of the CORDIC shown in Fig. 4.6d.

When *iteration* mode is active, the x and y inputs of the Vectoring mode CORDIC are $\underline{w}_{1,1}{}^{(p)}$ and $\underline{w}_{1,2}{}^{(p)}$ respectively as also shown in (4.30) where p denotes the iteration number (see Fig. 4.6a). The accumulated angle θ_p is obtained at the output of the Vectoring mode CORDIC which can be connected to the θ-input of the Rotation mode CORDIC as shown in Fig. 4.6d. The x and y inputs of the Rotation mode CORDIC under this active *iteration* mode become $z_{1,\,j}$ and $z_{2,\,j}$ respectively as also shown in (4.30). These data can be accessed serially from the respective memories. The x-output of the Rotation mode CORDIC becomes $\mathscr{R}$ o $t_x(z_{1,\,j}, z_{2,\,j}, \theta_p)$. Corresponding y-output is not used and should be ignored (see Fig. 4.6a).

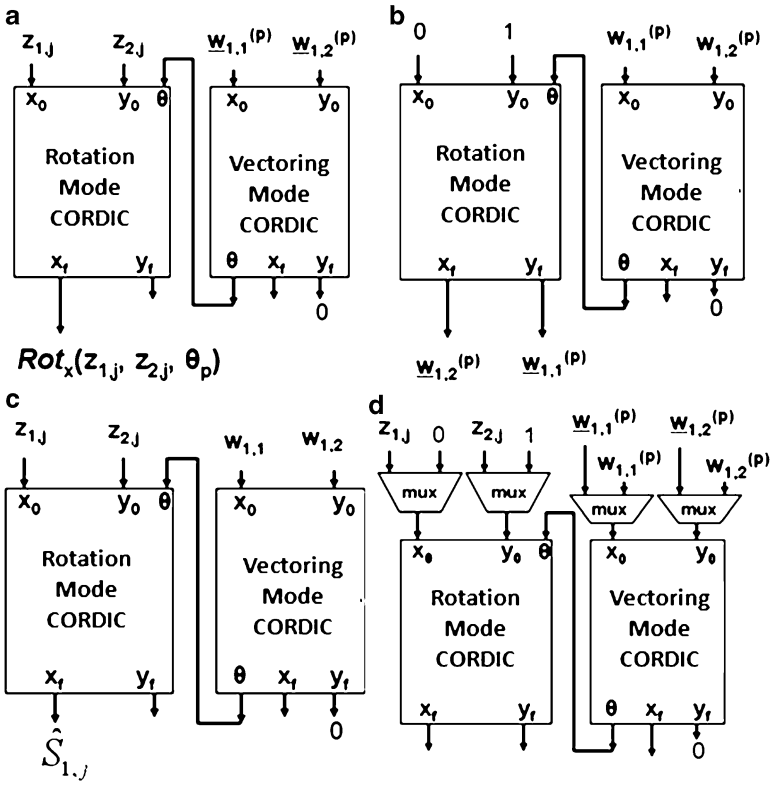


Fig. 4.6 CORDIC based 2*D* FastICA. (**a**) Iteration, (**b**) Normalization, (**c**) Estimation stage and (**d**) Multiplexed architecture

When *normalization* mode is active, the unnormalized vector components $w_{1,1}{}^{(p)}$ and $w_{1,2}{}^{(p)}$ are to be considered as the x and y inputs of the Vectoring mode CORDIC respectively as shown in Fig. 4.6b. The accumulated angle θ_p is computed and fed to the input of the Rotation mode CORDIC where 0 and 1 are considered as the inputs to x and y inputs respectively as shown in (4.33) (see Fig. 4.6b). As mentioned in the line following (4.33) care has to be taken while connecting the outputs of Rotation mode CORDIC with the next computational block in the *normalization* mode. The x and y outputs of the Rotation mode CORDIC become $\underline{w}_{1,2}{}^{(p)}$ and $\underline{w}_{1,1}{}^{(p)}$ respectively (see Fig. 4.6b).

When *estimation* mode is active, Vectoring mode CORDIC needs not to be used because the accumulated angle information is already computed during the previous normalization step and can be reused as the θ input of the Rotation mode CORDIC. This is the reason why "c" is relinquished from (4.34) to derive (4.35). Similar to *iteration* mode, in this *estimation* mode also $z_{1,j}$ and $z_{2,j}$ are to be considered as the x and y inputs of the Rotation mode CORDIC and here also the y-output has to be ignored (see Fig. 4.6c).

Since iteration, normalization and estimation—all stages are executed sequentially, same CORDIC unit can be reused for implementing these stages only at the expense of multiplexers at the inputs of the rotation and vectoring mode CORDIC. Combining these three stages, it is possible to model the multiplexers based single CORDIC unit for 2*D* FastICA implementation as shown in Fig. 4.6d. It can be noted from Fig. 4.6d that one 2 : 1 multiplexer is used at each of the x and y inputs of the Rotation and Vectoring mode CORDIC. Use of the demultiplexer as shown in Fig. 4.6d may not be mandatory and can be replaced by direct hardware wiring. Now, following Acharyya et al. (2011), we generalize this 2*D* FastICA algorithm in *nD* vector space using co-ordinate rotation based approach. Since the 2*D* algorithm mentioned above is based on the topological rather than the algebraical pattern exploitation of the basic FastICA algorithm, it is indeed possible to extend the 2*D* concept in *nD* vector space and in that process increase of the controller complexity will be inevitable. It is to be noted from Fig. 4.6d that the y-output of the Vectoring mode CORDIC is 0 (approximately) and x-output provides the magnitude of the radial vector in 2*D* $x - y$ plane. This x-output, in case of 3*D*, can be used as the x-input of the Vectoring mode CORDIC along with the third component of the vector of the unmixing matrix as the y-input. In this way, this concept of 2*D* co-ordinate rotation based FastICA can be extended from $x - y$ plane to $x - y - z$ plane and higher dimensions (Acharyya et al. 2011). Thus it is possible to propose a generalized theorem which may provide the theoretical basis of *n*D FastICA Iteration stage realization with the recursive use of its immediate lower dimension and ultimately boils down to the 2D CORDIC *Rotation* and *Vectoring* mode (Acharyya et al. 2011). Detailed proofs of the following theorems can be found in Acharyya et al. (2011).

Theorem 1 (CORDIC based Recursive Formulation of the Iteration Stage of *n*D FastICA Algorithm:)

$$\begin{bmatrix} w_{1,1}{}^{(p+1)} \\ w_{1,2}{}^{(p+1)} \\ w_{1,3}{}^{(p+1)} \\ . \\ . \\ . \\ w_{1,n}{}^{(p+1)} \end{bmatrix} = \begin{bmatrix} \mathscr{E}\left[z_{1,j}\{\mathscr{R}_{x,j}^{nD}\}3\right] \\ \mathscr{E}\left[z_{2,j}\{\mathscr{R}_{x,j}^{nD}\}3\right] \\ \mathscr{E}\left[z_{3,j}\{\mathscr{R}_{x,j}^{nD}\}3\right] \\ . \\ . \\ . \\ \mathscr{E}\left[z_{n,j}\{\mathscr{R}_{x,j}^{nD}\}3\right] \end{bmatrix} - 3 \begin{bmatrix} \underline{w}_{1,1}{}^{(p)} \\ \underline{w}_{1,2}{}^{(p)} \\ \underline{w}_{1,3}{}^{(p)} \\ . \\ . \\ . \\ \underline{w}_{1,n}{}^{(p)} \end{bmatrix} \tag{4.36}$$

where, considering $\mathscr{R}_{x,j}^{2D} = \mathscr{R}ot_x(z_{1,j}, z_{2,j}, \mathscr{V}ec_\theta(\underline{w}_{1,1}{}^{(p)}, \underline{w}_{1,2}{}^{(p)}))$ *as the basic Rotation mode and* $\mathscr{V}_\theta^{2D} = \mathscr{V}ec_\theta(\underline{w}_{1,1}{}^{(p)}, \underline{w}_{1,2}{}^{(p)})$; $\mathscr{V}_x^{2D} = \mathscr{V}ec_x(\underline{w}_{1,1}{}^{(p)}, \underline{w}_{1,2}{}^{(p)})$ *as the basic Vectoring mode of CORDIC operation, for* $n \geq 3$*, Iteration stage of the nD FastICA algorithm can be expressed in recursive way as follows:*

(i)

$$\mathscr{R}_{x,j}^{nD} = \mathscr{R}ot_x(z_{n,j}, \mathscr{R}_{x,j}^{(n-1)D}, \mathscr{V}ec_\theta(\underline{w}_{1,n}^{(p)}, \mathscr{V}_x^{(n-1)D})) \tag{4.37}$$

(ii)

$$\mathcal{V}_x^{nD} = \mathcal{V}ec_x(\underline{w}_{1,n}^{(p)}, \mathcal{V}_x^{(n-1)D}) \tag{4.38}$$

(iii)

$$\mathcal{V}_\theta^{nD} = \mathcal{V}ec_\theta(\underline{w}_{1,n}^{(p)}, \mathcal{V}_x^{(n-1)D}) \tag{4.39}$$

Theorem 2 (CORDIC based Recursive Formulation of the Normalization Stage of *n*D FastICA Algorithm:)

(a). *When* $n = 2$, $\underline{w}_{1,1}$ *and* $\underline{w}_{1,2}$ *can be represented as:*

$$\underline{w}_{1,1} = \mathcal{R}ot_y(0, 1, \mathcal{V}_\theta^{2D}) \tag{4.40}$$

$$\underline{w}_{1,2} = \mathcal{R}ot_x(0, 1, \mathcal{V}_\theta^{2D}) \tag{4.41}$$

(b). *When* $n = 3$, $\underline{w}_{1,1}$, $\underline{w}_{1,2}$ *and* $\underline{w}_{1,3}$ *can be represented as:*

$$\underline{w}_{1,1} = \mathcal{R}ot_y(0, \mathcal{R}_x^{3D}, \mathcal{V}_\theta^{2D}) \tag{4.42}$$

$$\underline{w}_{1,2} = \mathcal{R}ot_x(0, \mathcal{R}_x^{3D}, \mathcal{V}_\theta^{2D}) \tag{4.43}$$

$$\underline{w}_{1,3} = \mathcal{R}ot_y(0, 1, \mathcal{V}_\theta^{3D}) \tag{4.44}$$

(c). *When* $n \geq 4$,

(i). *for* $i = (n-1), (n-2), \ldots, 3$, $\underline{w}_{1,i}$ *can be represented as:*

$$\underline{w}_{1,i} = \mathcal{R}_y^{iD} = \mathcal{R}ot_y(0, \mathcal{R}_x^{(i+1)D}, \mathcal{V}_\theta^{iD}) \tag{4.45}$$

and

$$\underline{w}_{1,n} = \mathcal{R}_y^{nD} = \mathcal{R}ot_y(0, 1, \mathcal{V}_\theta^{nD}) \tag{4.46}$$

(ii). $\underline{w}_{1,1}$ *and* $\underline{w}_{1,2}$ *can be represented as:*

$$\underline{w}_{1,1} = \mathcal{R}ot_y(0, \mathcal{R}_x^{3D}, \mathcal{V}_\theta^{2D}) = \mathcal{R}_y^{2D} \tag{4.47}$$

$$\underline{w}_{1,2} = \mathcal{R}ot_x(0, \mathcal{R}_x^{3D}, \mathcal{V}_\theta^{2D}) = \mathcal{R}_x^{2D} \tag{4.48}$$

Theorem 3 (CORDIC based Recursive Formulation of the Estimation Stage of *n*D FastICA Algorithm:)

$$\hat{s}_{1,j} = \mathscr{R}_{x,j}^{nD} \tag{4.49}$$

where, considering $\mathscr{R}_{x,j}^{2D} = \mathscr{R}ot_x(z_{1,j}, z_{2,j}, \mathscr{V}ec_\theta(w_{1,1}, w_{1,2}))$ *as the basic Rotation mode and* $\mathscr{V}_\theta^{2D} = \mathscr{V}ec_\theta(w_{1,1}, w_{1,2})$; $\mathscr{V}_x^{2D} = \mathscr{V}ec_x(w_{1,1}, w_{1,2})$ *as the basic Vectoring mode of CORDIC operation, for* $n \geq 3$, *Estimation stage of the nD FastICA algorithm can be expressed in recursive way as follows:*

(i)

$$\mathscr{R}_{x,j}^{nD} = \mathscr{R}ot_x(z_{n,j}, \mathscr{R}_{x,j}^{(n-1)D}, \mathscr{V}ec_\theta(w_{1,n}, \mathscr{V}_x^{(n-1)D})) \tag{4.50}$$

(ii)

$$\mathscr{V}_x^{nD} = \mathscr{V}ec_x(w_{1,n}, \mathscr{V}_x^{(n-1)D}) \tag{4.51}$$

(iii)

$$\mathscr{V}_\theta^{nD} = \mathscr{V}ec_\theta(w_{1,n}, \mathscr{V}_x^{(n-1)D}) \tag{4.52}$$

Now in the next section we will do the hardware complexity analysis of the forementioned architectures.

4.4 Arithmetic Complexity and Trade-off Analysis

In this section we will do the arithmetic complexity and the tradeoff analysis especially for the biomedical signal processing architectures described before. However the same approach can be used to analyze the complexity of any architecture pertaining to any signal processing algorithm. The memory reduction methodology mentioned in the last section to implement low-complexity DWT and IDWT, there is a possibility that the area saving due to reduction of memory size may be outweighed by the requirement of extra adders. This section explicitly addresses this issue in terms of total number of transistor saving. To do that the total memory requirement using conventional DA is derived first. For simplicity, here only the memory requirement for the Analysis bank is considered. The memory requirement for the synthesis bank can be derived using the same approach. Following the same notations adopted in Acharyya et al. (2009a), the number of memory block (N_{ij}) required for computing jth wavelet coefficient at the ith resolution level using conventional DA can be given by:

$$N_{ij} = 2^s = 2^{1+(j-1)2^i} \tag{4.53}$$

Since each wavelet coefficient is a convolution sum of the input samples and filter coefficients, computation of j number of wavelet coefficients at the ith level of resolution requires j times application of DA. It means that (4.53) has to be iterated j times where j varies from 1 to $p/2^i$ for each i. Thus the total memory required for ith level of resolution can be expressed as:

$$N_i = \sum_{j=1}^{p/2^i} N_{ij} = 2 \times (2^p - 1)/((4)^{2^{(i-1)}} - 1) \tag{4.54}$$

The memory requirement for the analysis bank can be computed by summing N_i for all i. But, since the last level of resolution consists of one wavelet coefficient and one residue signal which requires two more memory blocks (Acharyya et al. 2009a), the total Conventional Memory Requirement (CMR) for the complete analysis bank can be expressed as:

$$CMR = 2 + \sum_{i=1}^{\log_2 p} N_i \tag{4.55}$$

Compared to CMR the total memory requirement in this proposed methodology is given by (4.23). The total hardware cost required in this approach is the summation of TMR and the number of extra adders required for generating the appropriate filter coefficients on-the-fly. To find out the number of required adders we define a parameter "Adder Penalty" (AP). Unlike TMR, AP is dependent on the number of coefficients present per level of resolution. In the proposed methodology, for computing j^{th} wavelet coefficient at ith resolution level $(s-1)/2$ number of W-bits adders is needed where s is given by (4.19). Denoting the adder requirement for ith stage as P_i and expressing s in terms of j we get:

$$\begin{aligned} P_i &= \sum_{j=1}^{p/2^i} (j-1)2^{(i-1)} \times W \\ &= (p/2^2) \times [(p/2^i) - 1] \times W \end{aligned} \tag{4.56}$$

Thus the total adder requirement for the analysis bank can be given by

$$AP = \sum_{i=1}^{\log_2 (p)-1} P_i \tag{4.57}$$

In (4.57), the upper limit of the summation is set to $\log_2(p) - 1$ due to the fact that the last resolution level does not need any addition because of the presence of only one sample in the corresponding wavelet coefficient. Considering a single-bit memory cell consists of t number of transistors and one single-bit adders consists of Kt number of transistors, Transistor Savings (TS) can be given as:

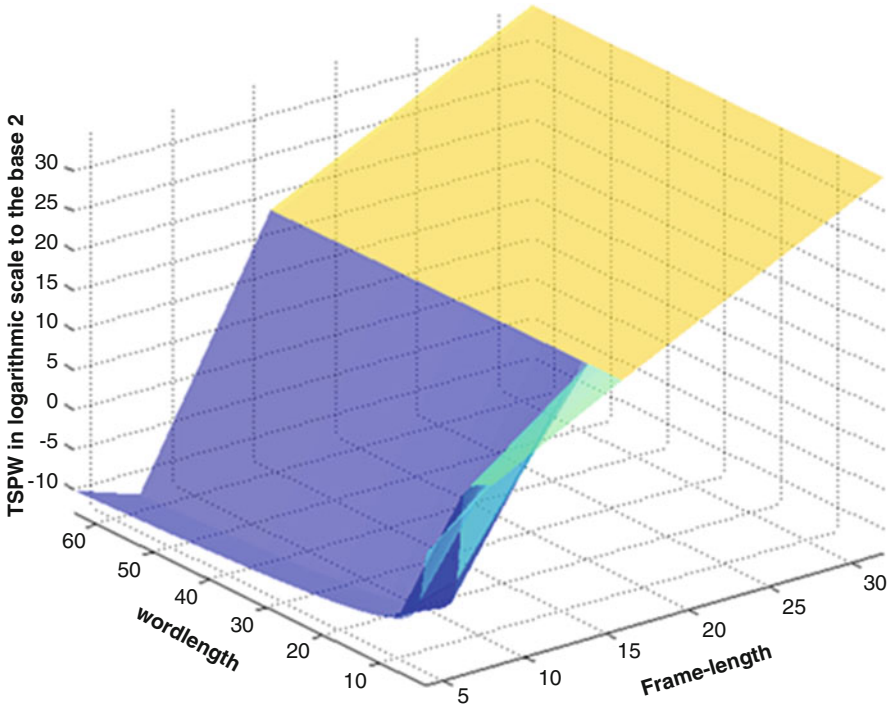


Fig. 4.7 Variation of TSPW with frame-length and word-length

$$TS = [CMR \times W - (TMR \times W + AP \times K \times W)] \times t \tag{4.58}$$

More specifically, from (4.58), a new metric called Transistor Savings Per Word-length (*TSPW*) can be defined as:

$$TSPW = [CMR - (TMR + AP \times K)] \times t \tag{4.59}$$

To obtain an effective savings in hardware in terms of total transistor count in the proposed methodology $TSPW$ should to be positive.

Considering a single bit SRAM cell requires 6 transistors and one single bit adder requires 18 transistors (Bellaouar and Elmasry 1995), the variation of TSPW with four different discrete values of frame-length (= 4, 8, 16 and 32) and word-length is plotted in Fig. 4.7. It can be observed that for smaller frames (< 8), $TSPW$ is negative meaning that for frame-length < 8, the proposed methodology does not achieve hardware savings over conventional DA based method. However for frame-length $= 8$, TSPW is positive for word-length up to 7, but above that *TSPW* becomes negative. This means that over word-length 7, the rate of quadratic growth in AP dominates over the rate of exponential growth of *CMR*. But for longer frame-length (> 8), *TSPW* increases for a fixed word-length. This means for longer

frames, the value of *CMR* dominates over the summation of *AP* and *TMR* resulting in hardware savings with the proposed methodology. It is to be noted that for longer frames also, keeping the frame-length fixed, *TSPW* starts falling gradually with the increase in word-length. But, this falling rate is much less compared to the rate of increase in *CMR* for longer frames. Following things are also to be observed from this figure. For the word-length < 8, TSPW improves approximately by 2^8 from the frame-length 8 to 16. When word-length $= 8$, as mentioned above, TSPW becomes negative for frame-length $= 8$ for the proposed algorithm. However, when frame-length $= 16$, TSPW is positive throughout and its improvement over the frame-length 8 for the fixed word-length $= 8$ is approximately by 2^{22} which increases with the higher word-length by an approximate amount of 2^{30}. Such improvement in TSPW from 2^8 to 2^{22} to 2^{30} can also be observed as the sharp discontinuity in the Fig. 4.7. This overall analysis demonstrates that such methodology is suitable for longer frames.

Now we turn our attention into the hardware complexity analysis of the FastICA architecture discussed in the last section. Throughout the analysis we keep a generalized view of frame-length m and word-length b. Since FastICA is an iterative procedure, we consider only one single iteration because the same hardware resources can be reused for the next iterations. Now considering the same assumption used in Acharyya et al. (2009b)—no resource sharing or parallel computing capability in one iteration, each arithmetic operation can be translated in terms of one arithmetic unit.

To provide a comparison on a uniform platform we consider only Ripple Carry Adder (RCA), Conventional Array Multiplier (CAM), Non-restoring Iterative Cellular Square Rooter (SQRT) and Non-restoring Array Divider (NAD) as the means of implementing the arithmetic operations. Considering a b-bit RCA requires b Full Adders (FA) (in a simplified view) (Acharyya et al. 2009b), $b \times b$ CAM requires $b(b-2)$ FA plus b Half Adders (HA) and b^2 AND gates (Acharyya et al. 2009b). Similarly one $b \times b$ NAD consists of $0.5 \times b(3b-1)$ FA and $0.5 \times b(3b-1)$ XOR gates (Acharyya et al. 2009b) and one b-bit SQRT needs $0.125 \times (b+6)b$ FA and XOR gates (Majithia 1973). In addition, considering one FA cell requires 24 transistors, one HA cell and one two input XOR gates consists of 12 transistors and a two input AND gates consists of 6 transistors (Acharyya et al. 2009b), we can calculate $TC_A = 24b$, $TC_M = 6b(5b-6)$, $TC_D = 18b$ (3b-1)and $TC_{SQ} = 18(b/2+1)(b/2+3)$, where TC_* are the transistor counts for RCA, CAM, NAD and SQRT respectively. One basic single bit $2:1$ Transmission Gate Multiplexer comprises of 4 transistors and thus b-bit $2:1$ multiplexer array has got $TC_{mux}^{2:1} = 4b$ (Weste and Harris 2005). Any larger multiplexer can be realised using cascaded $2:1$ multiplexer (Weste and Harris 2005) and thus total transistor count of $n:1$ multiplexer ($TC_{mux}^{n:1}$) can be expressed in terms of $TC_{mux}^{2:1}$ as: $TC_{mux}^{n:1} = (n-1) * TC_{mux}^{2:1}$. Using all these results and Considering the 2*D* FastICA case first, in Iteration phase, comparing (4.26) with (4.29), it can be found that 2 multiplications and 1 addition operation is saved per sample per frame in each iteration. Therefore, for the whole frame m, total saving is—$2m$ multiplications and m additions. In Normalization phase, comparing (4.31) and

(4.33), it can be observed that 2 multiplications, 1 addition, 1 square-rooting and 2 divisions operations are saved per iteration. In the output estimation phase, similar to the iteration step, $2m$ number of multiplications and m number of addition operations are saved. Considering all these operations in account, it can be found that overall saving per iteration is—$(4m + 2)$ multiplications, $(2m + 1)$ additions, 1 square-rooting and 2 divisions. Using all these values of TC_* in the arithmetic operations savings as derived in the last paragraph, overall Transistor Saving (TS) can be computed as:

$$TS_{2D} = (4m + 2)TC_M + (2m + 1)TC_A + TC_{SQ} + 2TC_D \tag{4.60}$$

The hardware saving obtained for CORDIC based 2*D* FastICA at (4.60), is mitigated somewhat by the insertion of the multiplexer arrays as shown in Fig. 4.6 and mentioned in Acharyya et al. (2010) and Acharyya et al. (2011). Therefore, using Fig. 4.6d and overall penalty for CORDIC based 2*D* architecture (P_{2D}) can be given as (Acharyya et al. 2011):

$$P_{2D} = 4 * TC_{mux}^{2:1} \tag{4.61}$$

Effective hardware saving for the proposed CORDIC based 2*D* architecture (ES_{2D}), if there is any, can be given by:

$$ES_{2D} = TS_{2D} - P_{2D} \tag{4.62}$$

where TS_{2D} and P_{2D} are obtained from (4.60) and (4.61) respectively. Using the previous assumption on arithmetic complexity in (4.62), ES_{2D} can be represented in terms of total number of transistors saved. Then effective TS obtained above for 2*D* case can be normalized with respect to b and a metric—Transistor Saving Per Word-length ($TSPW$) can be formulated using the approach presented in Acharyya et al. (2009a). Being the function of m and b, Fig. 4.8 shows the variation of $TSPW$ for the Proposed 2*D* Architecture over Shyu et al. (2008) and Acharyya et al. (2009b) with respect to these parameters considering $64 \leq m \leq 4096$ and $4 \leq b \leq 32$. It can be observed from Fig. 4.8 that $TSPW$ for the Proposed Algorithm is significantly higher than that presented in Acharyya et al. (2009b) for higher m and b. Following the same procedure adopted to analyze the performance of the 2*D* FastICA architectures above, the performance analysis of the CORDIC based *nD* FastICA architecture in generalized fashion can also be done. In the Iteration phase of the *n D* FastICA architecture, as obtained from *Theorem-1* in last section, it can be found that n multiplications and $(n - 1)$ addition operations are saved per sample per frame in each iteration. Therefore, for the whole frame m, total saving is—$n\,m$ multiplications and $(n - 1)m$ additions. In *nD* Normalization phase, from *Theorem-2* in last section, it can be observed that n multiplications, $(n - 1)$ addition, 1 square-rooting and n divisions operations are saved per iteration. In the output estimation phase, similar to the iteration step, nm number of

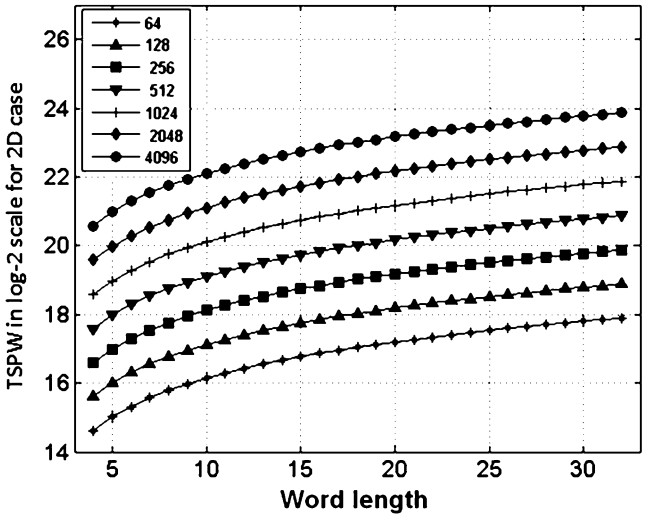


Fig. 4.8 Variation of transistor saving per word-length ($TSPW$) of the proposed algorithm for $2D$ with word-length and frame-length

multiplications and $(n-1)m$ number of addition operations are saved. Considering all these operations into account, it can be found that overall saving per iteration is—$(2m+1)n$ multiplications, $(2m+1)(n-1)$ additions, 1 square-rooting and n divisions. Following the approach mentioned before to derive (4.60), TS for nD case can be given as follows:

$$TS_{nD} = n(2m+1)TC_M + (n-1)(2m+1)TC_A + TC_{SQ} + nTC_D \tag{4.63}$$

The hardware saving obtained for CORDIC based nD FastICA at (4.63), is reduced somewhat by the insertion of the multiplexer arrays as mentioned in Acharyya et al. (2011). Therefore, overall penalty for CORDIC based nD architecture (P_{nD}) can be given as (Acharyya et al. 2011):

$$\begin{aligned} P_{nD} &= 4*TC_{mux}^{2:1} + 5*TC_{mux}^{(n-1):1} + TC_{mux}^{2:1} + 2*(k+1)*TC_{mux}^{(n-1):1} \\ &= 5*TC_{mux}^{2:1} + 5*(n-2)*TC_{mux}^{2:1} + 2*(k+1)*(n-2)*TC_{mux}^{2:1} \\ &= (5*(n-1) + 2*(k+1)*(n-2))*TC_{mux}^{2:1} \end{aligned} \tag{4.64}$$

where, $(k+1)$ denotes total number of CORDIC micro-rotations (Acharyya et al. 2011). In (4.64), it has been assumed that TC of a multiplexer is same as that of a demultiplexer for brevity.

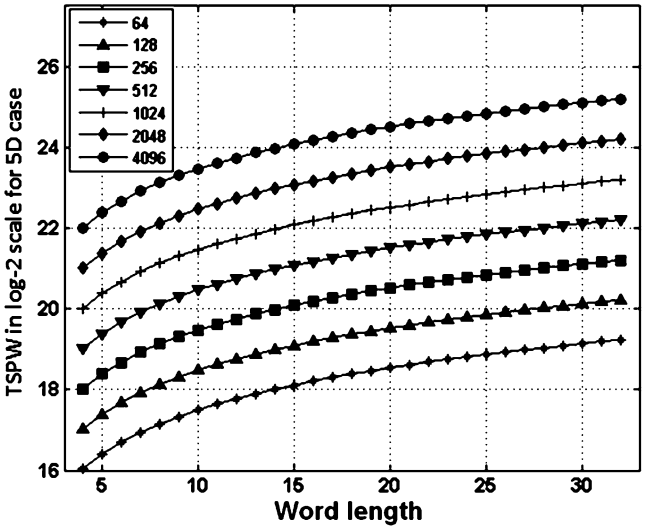


Fig. 4.9 Variation of transistor saving per word-length ($TSPW$) of the proposed algorithm for $5D$ with word-length and frame-length

Following the same procedure adopted for the proposed $2D$ architectures above, effective hardware saving for the proposed CORDIC based $n\,D$ architecture (ES_{nD}), if there is any, can be given by:

$$ES_{nD} = TS_{nD} - P_{nD} \tag{4.65}$$

where TS_{nD} and P_{nD} are obtained from (4.63) and (4.64) respectively. Using previous assumptions in (4.65), ES_{nD} can be represented in terms of total number of transistors saved. Then effective TS obtained above for nD case can be normalized with respect to b and a metric—Transistor Saving Per Word-length ($TSPW$) can be formulated using the approach presented in Acharyya et al. (2009a). Considering, $n = 5$ and $n = 6$ in (4.65), $TSPW$ for the proposed $5D$ and $6D$ cases can be obtained. Being the function of m and b, Figs. 4.9 and 4.10 show the variation of $TSPW$ for the Proposed $5D$ and $6D$ Architectures over the conventional FastICA architectures (direct mapping of $5D$ and $6D$ FastICA algorithms into architectures) with respect to these parameters considering $64 \leq m \leq 4096$ and $4 \leq b \leq 32$. It can be observed from these figures that, like Fig. 4.8, here also $TSPW$ for the Proposed Algorithm is significantly higher than the corresponding conventional architectures for higher m and b. Figure 4.11 shows the comparative $TSPW$ variation analysis for the proposed $2D$ to $6D$ FastICA cases with respect to different frame-length ($64 \leq m \leq 4096$) and word-length ($4 \leq b \leq 32$). It can be observed from Fig. 4.11 that $TSPW$ for the Proposed Algorithm increases even with the increase in dimension n as well with the increase of frame-length m and word-length b.

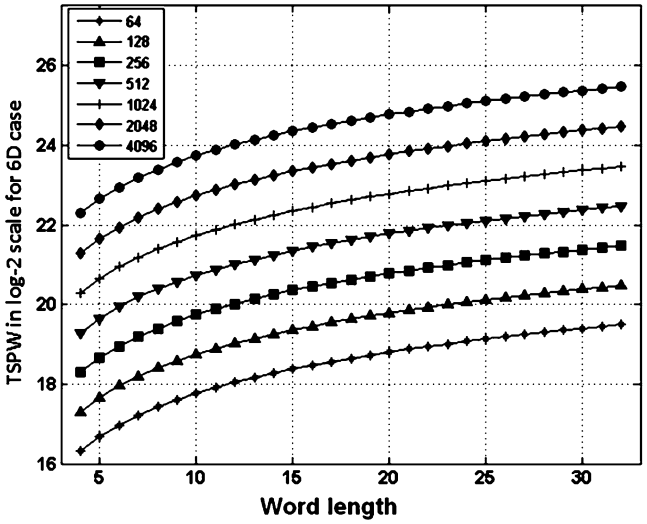


Fig. 4.10 Variation of transistor saving per word-length ($T\ S\ P\ W$) of the proposed algorithm for $6D$ with word-length and frame-length

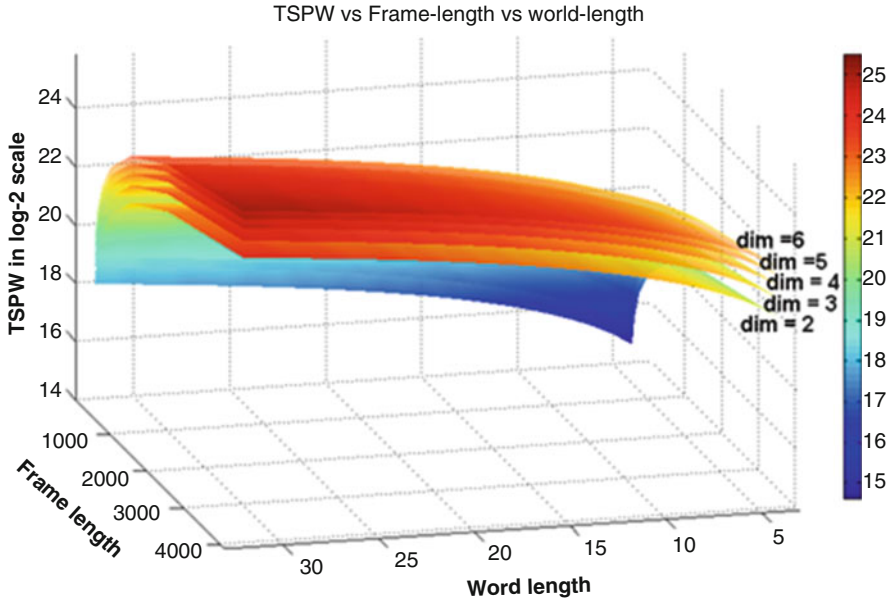


Fig. 4.11 Comparative variation of transistor saving per word-length ($T\ S\ P\ W$) of the proposed algorithm with frame-length and word-length for $2D$ to $6D$ (denoted by dim = 2 to dim = 6)

4.5 Possible Research Questions

Now before we conclude this chapter let us explore how the VLSI architectures implementation methodologies discussed above can be applied in some of the emerging areas of remote healthcare applications. One such area is cardio-vascular disease (CVD) detection, diagnosis and prognosis in a remote healthcare environment where the ECG signals are sensed, processed on a chip and transmitted to the central server whenever any abnormality is detected. Here the faithful reconstruction of standard 12-lead ECG from three or four leads is one potential research challenge. On-chip ECG signal compression would be another important area of research in such environment. Apart from this, real-time analysis of the EEG signals followed by online diagnosis in a cyber-physical system frame-work would help immensely the autistic children to join the main-stream in less time than that of the conventional way of treating them using behavioral therapy. ICA and WT would find significant applications here to separate EEG signals captured by different leads from each other followed by the removal of noise and artifacts including eye-blinks and muscular artifacts. Properties of ICA can be adopted here for the separation of EEG signals and WT would be useful to remove the artifacts and denoise it. However ICA used for solving the well-determined system may not be applicable in such real-time environment because it is very difficult to get number of mixed sensory data same as the number of sources. Therefore ICA for the under-determined system can be investigated and the corresponding low-complexity architectures can be formulated following the methodologies discussed above. One more rapidly emerging area is proteomic and its applications in CVD and Cancer. The state-of-the approaches are based on the laboratory based bio-chemical techniques which are time-consuming. Given the rate of increase in the number of patients suffering from these two diseases, technology can be used to expedite the procedure of analysis of the protein data obtained from the Mass Spectrometer envisaging a tiny chip packed with several signal processing architectures and capable of separating the protein profiles, classifying the patients in different diseased classes and suggesting the probable medicines and diagnostics.

4.6 Conclusion

This chapter identified denoising and signal separation as the most important need of remote healthcare application and dealt with WT and ICA algorithms and their corresponding low-complexity architecture design. It is to be noted that the same algorithms also can be utilised for extracting several features of diagnostic values from the signals sensed through the on-body sensors and also used for classifying the signals under different diseased classes. WT can be used for signal compression too by exploiting the sparsity of the signals. Apart from the symmetric FastICA algorithms and architectures described here, there are also more realistic ICA for

underdetermined-system and single-channel ICA where number of sensors are less than the number of sources under consideration. There are potential research challenges involved in proposing the low-complexity architectures for such algorithms. In this chapter one mathematical model for computing the arithmetic complexity has also been described which can again be used for any signal processing architectures.

References

Acharyya A, Maharatna K, Al-Hashimi BM (2010) Co-ordinate rotation based low complexity 2D FastICA algorithm and architecture. In: IEEE International Conference on Green Circuits and Systems, pp 60–64, 21–23 June, 2010, Shanghai, China

Acharyya A, Maharatna K, Al-Hashimi BM, Gunn SR (2009a) Memory reduction methodology for distributed-arithmetic-based DWT/IDWT exploiting data symmetry. IEEE Trans Circ Syst II Express Briefs 56(4):285–289

Acharyya A, Maharatna K, Al-Hashimi BM (2009b) Hardware reduction methodology for 2-dimensional kurtotic fast ICA based on algorithmic analysis and architectural symmetry. In: IEEE workshop on signal processing systems, October, 2009

Acharyya A, Maharatna K, Al-Hashimi BM, Reeve J (2011) Co-ordinate rotation based low complexity N-D FastICA algorithm and architecture. IEEE Trans Signal Process 59 (8):3997–4011

Addison P (2004) The little wave with the big future. Phys. world 17:35–39

Addison PS (2006) Wavelet transforms and the ECG: a review. Physiol Meas 26

Akay M (1997) Wavelet applications in medicine. IEEE Spectrum 50–56

Alam M, Rahman CA, Badawy W, Jullien G (2003) Efficient distributed arithmetic based DWT architecture for multimedia applications. In: IEEE international workshop on system-on-chip for real-time applications, pp 333–336

Aldroubi A, Unser MA (1996) Wavelets in medicine and biology. CRC Press, Boca Raton, FL. http://www.crcpress.com/product/isbn/9780849394836

Aoyagi T, Tokutaka H, Fujimura F, Maniwa T (2002) Application of FastICA to pulse wave. In: Proceedings of the 9th international conference on neural information processing (ICONIP), vol 2, pp 769–772

Barros AK, Mansour A, Ohnishi N (1998) Removing artifacts from electrocardiographic signals using independent component analysis. Neurocomputing 22:173–186

Bellaouar A, Elmasry MI (1995) Low-power digital VLSI design: circuits and systems. Springer. http://www.springer.com/engineering/circuits+%26+systems/book/978-0-7923-9587-4

Bentley PM, McDonnell JTE (1994) Wavelet transforms: an introduction. Electron Comm Eng J 175–186

Bravo I, Jim*é*nez P, Mazo M, L*á*zaro JL, Gardel A (2006) Implementation in FPGAs of Jacobi method to Solve the eigenvalue and eigenvector problem. In: International conf. field prog. logic and applications

Bravo I, et al (2008) Novel HW architecture based on FPGAs oriented to solve the eigen problem. IEEE Trans VLSI Syst 16(12)

Brent RP, Luk FT, Loan CV (1985) Computation of the singular value decomposition using mesh connected processors. J VLSI Comput Syst 1(3):242–270

Burleson WP (1990) Polynomial evaluation in VLSI using distributed arithmetic. IEEE Trans Circ Syst 37(10):1299–1304

Calhoun VD, Adali T (2006) Unmixing fMRI with independent component analysis. IEEE Eng Med Biol Mag 79–90

Cao X, Xie Q, Peng C, Wang Q, Yu D (2006) An efficient VLSI implementation of distributed architecture for DWT. In: IEEE 8th workshop multimedia signal processing, pp 364–367
Cardoso JF (1998) Blind signal separation: statistical principles. Proc IEEE 86(10):2009–2025
Cavallaro JR, Luk FT (1988) CORDIC arithmetic for an SVD processor. J Parallel Distr Comput 5:271–290
Chen X, Zhou T, Zhang Q, Li W, Min H (1996) A VLSI architecture for discrete wavelet transform. IEEE Int Conf Image Process 1:1003–1006
Chien JC, Huang MC, Lin YD, Chong F (2006) A study of heart sound and lung sound separation by independent component analysis technique. In: 28th Annual intl conference of the IEEE engg in medicine and biology society, pp 5708–5711
Chui CK (1992) An introduction to wavelets. Academic, San Diego, CA
Cichocki A (1999) What is intelligent blind signal processing and why it is important in brain science? RIKEN Brain Science Institute News, Japan, no. 4, June, 1999
Cichocki A, Amari S (2002) Adaptive blind signal and image processing: learning algorithms and applications. Wiley. http://as.wiley.com/WileyCDA/WileyTitle/productCd-0471607916.html
Cohen A, Kovacevic J (1996) Wavelets: The mathematical background. Proc IEEE 84(4):514–522
Comani S, Liberati M, Mantini D, Gabriele E, Brisinda D, Luzio SD, Fenici R, Luca G (2004) Characterization of fetal arrhythmias by means of fetal magnetocardiography in three cases of difficult ultrasonographic imaging. Pacing Clin Physiol (PACE) 27(12):1647–1655
Daubechis I (1992) Ten lectures on wavelets. SIAM. http://www.cambridge.org/us/academic/subjects/mathematics/abstract-analysis/ten-lectures-wavelets
Denk TC, Parhi KK (1997) VLSI architectures for lattice structure based orthonormal discrete wavelet transforms. IEEE Trans Circ Syst II Analog Digital Signal Process 44(2):129–132
Du H, Qi H, Wang X (2007) Comparative study of VLSI solutions to independent component analysis. IEEE Trans Ind Electron 54(1)
Fridman J, Manolakos ES (1994) Distributed memory and control VLSI architectures for the 1-D discrete wavelet transform. IEEE Workshop VLSI Signal Process VII:388–397
Götze J, Hekstra GJ (1995) An algorithm and architecture based on orthonormal μ-rotations for computing the symmetric EVD. Integrat VLSI J 20:21–39
Götze J, Paul S, Sauer M (1993) An efficient Jacobi-like algorithm for parallel eigenvalue computation. IEEE Trans Comput 42(9)
Graps A (1995) An introduction to wavelets. IEEE Comput Sci Eng 2(2):50–61
Grzeszczak A, Yeap TH, Panchanathan S (1994) VLSI architecture for discrete wavelet transform. Can Conf Electr Comput Eng 2:461–464
Grzeszczak A, Mandal MK, Panchanathan S, Yeap T (1996) VLSI implementation of discrete wavelet transform. IEEE Trans Very Large Scale Integrat (VLSI) Syst 4(4):421–433
He T, Clifford G, Tarassenko L (2006) Application of independent component analysis in removing artefacts from the electrocardiogram. Neural Comput Appl 15:105–116
Hsiao SF, Delosme JM (1996) Parallel singular value decomposition of complex matrices using multidimensional CORDIC algorithms. IEEE Trans Signal Process 44(3)
Hu YH (1992) CORDIC based VLSI architecture for digital signal processing. IEEE Signal Process Mag 16–35
Hyvärinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Network 10(3)
Hyvärinen A, Oja E (2000) Independent component analysis: Algorithms and applications. Neural Networks 13(4–5):411–430
Hyvärinen A, Karhunen J, Oja E (2001) Independent component analysis. Wiley, New York
Jou JM, Shiau YH, Liu CC (2001) Efficient VLSI architectures for the biorthogonal wavelet transform by filter bank and lifting scheme. IEEE Int Symp Circ Syst 529–532
Jutten C, Herault J (1991) Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. J Signal Process 24:1–10. Elsevier Science Publishers

Karvounis EC, Tsipouras MG, Fotiadis DI, Naka KK (2007) An automated methodology for fetal heart rate extraction from the abdominal electrocardiogram. IEEE Trans Inform Tech Biomed 11:628–638

Knowels G (1990) VLSI architecture for the discrete wavelet transform. Electron Lett 26 (15):1184–1185

Li C, Zheng C, Tai C (1995) Detection of ECG characteristic points using wavelet transforms. IEEE Trans Biomed Eng 42:21–28

Lo B, Deligianni F, Yang GZ (2006) Source recovery for body sensor network. In: IEEE international workshop on wearable and implantable body sensor networks, April, 2006

Majithia JC (1973) Pipeline array for square-root extraction. Electron Lett 9(1):4–5

Mallat SG (1999) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11(7)

Martina M, Masera G (2006) Low-complexity, efficient 9/7 wavelet filters VLSI implementation. IEEE Trans Circ Syst II Express Briefs 53(11):1289–1293

Martina M, Masera G (2007) Multiplierless, folded 9/7 – 5/3 wavelet VLSI architecture. IEEE Trans Circ Syst II Express Briefs 54(9):770–774

Mutihac R, Mutihac RC (20007) A comparative study of independent component analysis algorithms for electroencephalography. Rom Rep Phys 59(3):827–853

Oja E, Yuan Z (2006) The FastICA algorithm revisited: convergence analysis. IEEE Trans Neural Network 17(6)

Oweiss KG, Mason A, Suhail Y, Kamboh AM, Thomson KE (2007) A scalable wavelet transform VLSI architecture for real-time signal processing in high-density intra-cortical implants. IEEE Trans Circ Syst I Regular Papers 54(6)):1266–1278

Pan Q, Zhang L, Dai G, Zhang H (1999) Two denoising methods by wavelet transform. IEEE Trans Signal Process 47:3401–3406

Parhi KK, Nishitani T (1993) VLSI architectures for discrete wavelet transforms. IEEE Trans Very Large Scale Integrat (VLSI) Syst 1(2):191–202

Pourazad MT, Moussavi Z, Farahmand F, Ward RK (2005) Heart sounds separation from lung sounds using independent component analysis. In: IEEE Annual conference on engineering in medicine and biology, pp 2736–2739

Rioul O, Vetterli M (1991) Wavelets and signal processing. IEEE Signal Process Mag 14–38

Rioul O, Duhamel P (1992) Fast algorithms for discrete and continuous wavelet transforms. IEEE Trans Inform. Theory 38(2):569–586

Rioul O (1993) A discrete-time multiresolution theory. IEEE Trans Signal Process 41(8):2591–2606

Roberts S, Everson R (2001) Independent component analysis: Principles and practice. Cambridge University Press, Cambridge

Senhadji L, Carrault G, Bellanger JJ, Passariello G (1995) Comparing wavelet transforms for recognizing cardiac patterns, IEEE Eng Med Biol Mag 14:167–173

Shyu KK, Lee MH, Wu YT, Lee PL (2008) Implementation of pipelined FastICA on FPGA for real-time blind source separation. IEEE Trans Neural Network 19(6):958–970

Stone JV (2004) Independent component analysis: A tutorial introduction. MIT Press

Strang G (1989) Wavelets and dilation equations: a brief introduction. SIAM Rev Soc Ind Appl Math 31(4):614–6127

Taylor FJ (1986) An analysis of the distributed arithmetic digital filter. IEEE Trans Acoust Speech Signal Process 34(5):1165–1170

Tie Y, Sahin M (2005) Separation of spinal cord motor signals using the FastICA method. J Neural Eng 2:90–96

Truchetet F, Laligant O (2004) Wavelets in industrial applications: a review. In: Wavelet Application in Industrial Processing II, Proceedings of SPIE, vol 5607, pp 1–14, November, 2004

Unser M, Aldroubi A (1996) A review of wavelets in biomedical applications. Proc IEEE 84 (4):626–638

Vetterli M, Herley C (1992) Wavelets and filter banks: theory and design. IEEE Trans Signal Process 40(9):2207–2232

Vishwanath M, Owens RM, Irwin MJ (1995) VLSI architectures for the discrete wavelet transform. IEEE Trans Circ Syst II Analog Digital Signal Process 42(5):305–316
Volder JE (1959) The CORDIC trigonometric computing technique. IRE Trans Electron Comput EC-8:330–334
Walker JS (1997) Fourier analysis and wavelet analysis. Notice AMS 44(6):658–670
Walther JS (1971) A unified algorithm for elementary functions. In: Spring joint computer conference, pp 379–385
Weste NHE, Harris D (2005) CMOS VLSI design: a circuits and systems perspective, chap-1, 3rd international edition. Pearson-Addison Wesley. http://www.mypearsonstore.com/bookstore/cmos-vlsi-design-a-circuits-and-systems-perspective-9780321547743?xid=PSED
Yu C, Hsieh CA, Chen SJ (1997) Design and implementation of a highly efficient VLSI architecture for discrete wavelet transform. In: IEEE Custom Integrated Circuit Conference, pp 237–240

Chapter 5
Machine Learning Techniques for Remote Healthcare

Saptarshi Das and Koushik Maharatna

Abstract In this chapter, popular machine learning techniques are discussed in the context of remote healthcare. In this domain the main challenges are low computational complexity and hardware implementation, and not just conventional way of mathematical analysis of machine learning algorithms. Statistical view-point of different machine learning techniques, standard parametric and nonparametric algorithms for classification and clustering are briefly discussed. A practical 12-lead Electrocardiogram (ECG) signal based myocardial scar classification example has also been shown as a representative example. Complexity of few classification algorithms, online implementation issues for statistical feature extraction and some open research problems have also been introduced briefly.

5.1 Role of Machine Learning in Remote Healthcare

Machine learning or pattern recognition mainly deals with automatic selection of patterns into few classes using some statistical measures of the patterns known as features. The application of machine learning in biomedical signals is mostly to detect healthy and unhealthy electrophysiological condition in an automated way and have already become a celebrated research topic (Kononenko 2001; Magoulas and Prentza 2001). The electrical signals generated from human body like ECG, EEG, EMG etc. often carry some signature of the physical condition of a patient. The role of machine learning is to generate meaningful and informative features from these recordings. In most of the biomedical applications, the classification is done based on two classes denoting the healthy and unhealthy physical condition, though classification with more than two classes may also be required representing the degree of severity of disease or discrimination between different diseases.

S. Das (✉) • K. Maharatna
School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
e-mail: sd2a11@ecs.soton.ac.uk; km3@ecs.soton.ac.uk

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_5,

The task of the classifier here is to judge the condition of a patient from a-few number of variables, commonly derived from popular signal or image processing algorithms. For remote monitoring there might be two possibilities for classification. In the first category, the biological signal could be sensed and after that using some low complexity operations corresponding to meaningful and discriminative feature extraction, followed by classifier training. The on-body classifier will then give the indication of normal or abnormal health condition. In the second category, the biological signals may be acquired by the on-body sensor and then transmitted to a computer for analysis and decision making. Few machine learning algorithms, due to having their high complexity in the feature extraction and decision making process needs high computational power. These implementation issues are the most critical decisions in applying most of the established machine learning algorithms for remote healthcare since the battery running the complex algorithms will be exhausted quickly and hence cannot be implemented on body along with the sensor. Among many other techniques, the two most popular paradigms in machine learning are known as classification and clustering. In classification, the algorithm or classifiers is first trained using few set of statistical features, called the training data-set and known class information. Now, the task of the classifier is to assign patterns from respective features to one of these classes using a new feature vector called test data-set. In contrast, the clustering algorithms themselves try to find out the class information from the feature vectors.

Historically, the theoretical basis of classification was devised from the Bayes decision theory which finds the optimum separation boundary between different classes. Theoretically, the Bayesian approach guarantees optimal classification if the probability distributions functions (pdfs) of different classes are known a priori. With the consideration of Gaussian pdf, the knowledge of accurate mean and covariance matrices of different classes need to be known, which is quite impractical in most cases. The other implementation hurdle for optimal Bayesian classification approach is that it assumes that the data has been generated from Gaussian distribution which is often violated in practice and also more when we have limited number of data-points to train the classifier. In spite of being strong theoretical basis, the optimal Bayesian classifier has these practical limitations to deal with, particularly in biomedical signal processing and machine learning related applications. This motivates the use of several parametric and non-parametric classification techniques like discriminant analysis to the use of neural networks and several others heuristic technique based classifier and clustering approaches (Das et al. 2009).

Classification is generally done using two different strategies i.e. unsupervised and supervised learning. In unsupervised learning, the classifier tries to find out patterns within the data itself even without having any prior knowledge of the class labels of training-data and sometimes even without the knowledge about the number of underlying classes. In supervised learning, the classifier is first trained using a training data-set and then it is tested on a different test data-set. The classifier's parameters are adapted during the training phase by minimization of the classification error. One of the often encountered problems in designing

classification algorithms is known as over-fitting of the algorithm which gives very low classification error for the training-set but fails to generalize similar unknown test-set. If in the scatter diagram between two different features, the data points can be separated to their respective classes using a straight line the data-set is known as linearly separable and the line above and below which the classes are defined is known as the decision boundary. If the data-points become mixed and can only be separated using nonlinear kernels, then it is known as nonlinearly separable. The goal of classification in biomedical application should be directed towards finding such features which makes the problem linearly separable since introducing complicated nonlinearity are prone to give over-fitting. This is essential due the simplicity in implementation in low power circuits which may be mounted on the patient's body. This is an active and growing research topic now days in the field of wearable pervasive computing (Yang 2006).

A good classifier should be able to faithfully classify any unknown data that it has never seen. As the complexity of the classifier increases, the training error reduces but the validation error increases. In most of the practical biomedical applications, it is difficult to get a separate validation data-set to test classifier's performance or we want to use the whole information in the data-set to train the classifier rather than keeping a part of the data-set aside for validation. In such cases the data-set is divided in several (say N) segments or folds and the classifier is trained without the Nth segment into account and then tested on that segment. Then we rotate the fold and pick a different fold and test the classifier again by taking rest of the data-sets for training. The average classification accuracy is then computed over all the folds and the technique is known as N-fold cross validation of the classification algorithm. The original data-set can also be randomized before running the classifier in order to ensure that the data-points in each fold are different for each run which helps to generalize the classifier's performance. It is a common practice to take an average accuracy of multiple repetitions of the N-fold cross validation with randomized data-segments each time, as a preventive measure for over-fitting. As an extreme case, when N equals the number of data-points that we have for training is known as the Leave-One-Out Cross-Validation (LOOCV) which is in fact computationally heavy for large data-points but becomes necessary when we have less number of data-points in each fold of the N-fold cross validation (Rogers and Girolami 2011).

5.2 Constraints for Applying Traditional Machine Learning Techniques in Remote Healthcare

Machine learning techniques for Magnetic Resonance Imaging (MRI), Diffusion Tensor Imaging (DTI) analysis and other image analysis need more computational power. The main thrust for implementable low complexity algorithms mostly deal with the Electrocardiography (ECG), Electroencephalography (EEG), Magnetoencephalography (MEG) and Electromygraphy (EMG) processing related machine

learning techniques (Sörnmo and Laguna 2005). Historically the naïve Bayesian classifier, back-propagation neural networks, and symbolic learning were used in several biomedical applications, but recently improved algorithms are being encouraged that can handle missing and noisy data, also give insight and transparency of the diagnostic knowledge, and performs well even with small set of data points (Kononenko 2001). Machine learning pervaded several interesting applications in biomedical engineering for example non-invasive estimation of blood glucose level and blood pressure (Monte-Moreno 2011), kernel based classification approaches of epilepsy diagnosis from EEG (Lima and Coelho 2011), classifier ensemble to tackle missing features (Nanni et al. 2012), ECG analysis from statistical, geometric and nonlinear features (Jovic and Bogunovic 2011; Maglaveras et al. 1998), application of Bayesian network in bi-level multi-classifier (Sierra et al. 2001) and image interpretation (Velikova et al. 2013), machine learning approaches for Dementia scoring (Mani et al. 1999) etc. Regarding histo-pathological image analyses, several biomedical applications of machine learning techniques like object level, spatially related and multi-scale feature extraction, graph based modeling, feature selection, dimensionality reduction and manifold-learning are reviewed in (Gurcan et al. 2009). Application of various supervised and unsupervised learning techniques in EEG, MEG and MRI data analysis has been reviewed in (Bießmann et al. 2011). Application of state of the art clustering algorithms in biomedical applications, especially MRI analysis using several distance and similarity measures in hard clustering, soft clustering, kernel based, evolutionary computational based and neural network based clustering are categorized in (Xu and Wunsch 2010). In (Sajda et al. 2009) application of multivariate signal processing and machine learning techniques have been reviewed for neuroimaging in human perceptual decision making using functional MRI (fMRI) and EEG data.

The application of traditional machine learning techniques in many biomedical signals and images are not so straight forward. Most of the biomedical signal due to their non-stationary nature, behave in a very complex way. Therefore, standard statistical feature extraction techniques may not be valid in many realistic scenarios. Signal processing techniques considering the non-stationarity of the signal like use of Wavelet and Hilbert transform involves several heuristics like choice of basis function, window size etc. which are generally problem dependent. The main constraints in applying traditional probabilistic modeling technique for biomedical signals are the unavailability of robust techniques to estimate the exact pdfs of the classes. In recent years, many biomedical signal classification strategies are developed for condition monitoring of patients. For example, in stroke rehabilitation the EMG signals and neurophysiological condition monitoring the EEG signals may be acquired in a pervasive way and then transferred to a personal computer for processing. For ECGs the three steps i.e. signal processing, feature extraction and classification is quite easier to be carried out on the body worn sensor itself, since ECG signals are periodic and less complex than the random signals like EEG and EMG. While dealing with random signals like EEG and EMG, one may wish to extract its statistical properties as features (Herman et al. 2008;

Lucas et al. 2008), whereas morphology based feature extraction may be preferable for periodic signals like ECG (Mazomenos et al. 2012, 2013). Although there are many feature selection algorithms available in the literature, most of these techniques include heuristic optimization technique which are difficult to implement in low-complexity hardware. The big challenge in the field of biomedical signal classification is to extract application-specific meaningful and reliable features having minimum correlation amongst them. The curse of dimensionality is another big constraint since with arbitrarily high number of features two classes may be separable, but an engineering approach would be looking as much information as one can extract from those features using Principal Component Analysis (PCA) (Subasi and Ismail Gursoy 2010).

5.3 Overview of Feature Extraction Principles

5.3.1 Role of Advanced Signal Processing Techniques in Feature Extraction

For most of random signals occurring in biomedical applications, few statistical features are commonly used to build a classifier which can automatically predict about the patient's health condition by only looking at those set of features. Basic signal processing tools like frequency domain transforms and filters are employed to clean the biological signals from artifacts and noise. Few advanced signal processing tools like joint time-frequency domain exploration, single channel and multi-channel signal processing techniques are also very popular for feature extractions which are outlined briefly in this subsection.

The basic four features may be considered from the first order statistics of a random biological signal like EEG or MEG. The mean is given by the first moment of a random variable X i.e. $\mu = \overline{X} = E[X]$. The higher ($k^{th}$) central moments ($m_k$) are often used as features which comes from the first order statistics of the signal as follows.

$$m_k = E\left[\left(X - \overline{X}\right)^k\right] \tag{5.1}$$

The second central moment is known as the variance ($m_2 = \sigma^2$). In (5.1), simply putting $k = 3$, 4 and normalizing by k^{th} power of standard deviation gives the third and fourth standardized moments known as skewness (γ) and kurtosis (β). Among second order moments of biomedical images, the popular measures are angular second moment (ASM), contrast, correlation, homogeneity, entropy etc. Using first and second order statistical features for brain MRI images, a classification has been attempted in (Aggarwal and Agrawal 2012). Considering the issue of low computational complexity for feature extraction in random biomedical signals like EEG and EMG, several statistical feature extraction methods from time-series data are given in (Kugiumtzis and Tsimpiris 2010).

Considering common frequency domain transforms for biomedical signal analysis, the features can be both in time and frequency domain, as well as in joint time-frequency domain. Also, the features may capture the characteristics of one electrode or multiple electrodes together using popular univariate and multivariate signal processing techniques. Due to the uncertainty principle in time-frequency localization there is no single universal method which is capable of capturing both domain information as a part of the feature extraction process. The time-frequency inequality is given by (5.2).

$$\sigma_t^2 \sigma_f^2 \geq \frac{1}{16\pi^2} \tag{5.2}$$

Here, $\{\sigma_t^2, \sigma_f^2\}$ are the time and frequency spans which can be represented in terms of mean time (t_0), mean frequency (f_0) and total energy (E_x) of a biomedical signal $x(t)$ as given in (Blinowska and Zygierewicz 2012).

$$\begin{aligned}
\sigma_t^2 &= \frac{1}{E_x}\int_{-\infty}^{\infty}(t-t_0)|x(t)|^2 dt, \quad t_0 = \frac{1}{E_x}\int_{-\infty}^{\infty} t|x(t)|^2 dt, \\
\sigma_f^2 &= \frac{1}{E_x}\int_{-\infty}^{\infty}(f-f_0)|X(f)|^2 df, \quad f_0 = \frac{1}{E_x}\int_{-\infty}^{\infty} f|X(f)|^2 df \\
E_x &= \int_{-\infty}^{\infty}|x(t)|^2 dt = \int_{-\infty}^{\infty}|X(f)|^2 df
\end{aligned} \tag{5.3}$$

Different time-frequency transform techniques like Short-Time Fourier Transform (STFT), Hilbert transform, Continuous Wavelet Transform (CWT) etc. employs various time-frequency resolution but in each of the cases the inequality (5.2) holds. Before the feature extraction, biological signals are generally passed through the digital filters to retain only the useful spectral components e.g. to isolate a specific EEG band like delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz) or gamma (30 Hz and above) (Sörnmo and Laguna 2005). In order to do that the Finite Impulse Response (FIR) or Moving Average (MA) filters and Infinite Impulse Response (IIR) or Auto-Regressive (AR) and more general Auto-Regressive Moving Average (ARMA) filters are employed with specifications of filtering in frequency and time domains respectively. The frequency domain FIR/IIR filtering includes specifying the pass and stop band cut-off frequencies and associated ripples with a provision of choosing different windows and design paradigms. Other filtering techniques like matched filters maximize the output Signal to Noise Ratio (SNR) with a precise knowledge of the signal template and noise covariance. Other stochastic filtering paradigms like Wiener filtering and Least Mean Square (LMS) have been successfully used for eye blinking artifact removal from EEG signals by considering the Electro-oculogram (EOG) as the disturbance and also in noise reduction in evoked potentials (Sörnmo and Laguna 2005). For artifact removal the use of Independent Component Analysis (ICA) and

Empirical Mode Decomposition (EMD) and their hybridization is becoming increasingly popular (Lindsen and Bhattacharya 2010; Mijovic et al. 2010). The concept of state in biomedical applications using Hidden Markov Model (HMM) and Kalman Filter are detailed in (Blinowska and Zygierewicz 2012).

Advanced signal processing techniques with the consideration of stationarity of the signal (like entropy, autocorrelation, power spectral density estimates based on Fourier transform or by different windowing techniques, parametric models like AR and ARMA) have already been applied on biological signals in (Blinowska and Zygierewicz 2012; Cerutti and Marchesi 2011). There are also similar techniques available with a nonstationary consideration of the biomedical signal e.g. the use of Hilbert transform, time-frequency energy distributions (Wigner-Ville distributions, Cohen class, Choi-Williams distribution etc.), STFT based spectrogram, CWT based scalogram etc. The above mentioned spectrogram or scalogram considers a-priori trade-off for the time-frequency resolution, imposed by (5.2). For understanding most of the time-frequency domain representations, it is hard to locate which aspect is due to the signal and which one due to the particular choice of that specific method. In order to overcome this problem, specialized techniques like Matching Pursuit (MP) has been developed which automatically adjusts the local signal properties by matching residue of the decomposed signal. In a similar technique, known as EMD decomposes a nonstationary biomedical signal into several intrinsic mode functions (IMFs) by obtaining the upper and lower envelopes of the signal and checking the residues which has been successfully applied in artifact separation in EEG signals (Mijovic et al. 2010). Nonlinear methods in signal processing are also becoming increasingly popular for patient's health condition monitoring e.g. the use of phase space diagrams, Lyapunov exponent for detecting chaos, correlation dimension, fractal dimension, detrended fluctuation analysis (DFA), recurrence plots, approximate entropy etc. Care should be taken while using nonlinear methods in choosing the bins for calculating probabilities, time delay embedding in phase space reconstruction (PSR) and also the level of noise present in the signal as detailed in (Blinowska and Zygierewicz 2012; Cerutti and Marchesi 2011).

In most biomedical signal processing applications, we have signals for more than one channel like standard 12-lead ECG or 128-electrode EEG systems. The cross-electrode statistical measures are being very useful in capturing any diseased condition using the concepts of cross-correlation, cross-power spectrum, coherence etc. All these tools are based on bivariate statistical analyses to understand a multivariate problem. For cases where two or more channels are driven by the same source, such existence of correlation does not necessarily mean causation. Therefore, improved measures like partial cross-spectrum and partial coherence and more generally speaking, multiple cross-coherence is introduced to find out the common information in different channels. Similar to the AR model of single channel data, the technique can also be generalized to multivariate autoregressive (MVAR) type parametric models and can be estimated both in time or frequency domains. There are also measures of directedness like Granger causality index (GCI), directed transfer function (DTF), partial directed coherence (PDC) and

multivariable nonlinear measures like entropy, mutual information and phase synchronization etc. The two most popular multivariate signal decomposition algorithms include independent component analysis (ICA) and principal component analysis (PCA), which along with other mentioned signal processing tools are widely used as a part of the feature extraction process before applying any sort of supervised or unsupervised learning technique on biomedical signals.

5.3.2 *Preprocessing of Features and Feature Ranking*

Often the features (x) are normalized or standardized so as not to bias the training of the classifier with features having high value. The normalization is done to scale the feature vector so that they are scaled within the maximum ($x_{\max}$) and minimum ($x_{\min}$) value, whereas the standardization is done by removing the mean (μ_x) and normalizing it by the standard deviation (σ_x).

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, x_{std} = \frac{x - \mu_x}{\sigma_x} \tag{5.4}$$

Nonlinear sigmoid function type kernels may be used for data standardization if the data-points are not evenly distributed around the mean which is known as the softmax scaling (Theodoridis and Koutroumbas 2010) with a user defined parameter r.

$$x_{\text{softmax}} = 1 \Big/ \left(1 + \exp\left(-\left(\frac{x - \mu_x}{r\sigma_x}\right)\right)\right) \tag{5.5}$$

Outlier rejection is an important step in pre-processing of the feature matrix. Generally data-points lying beyond the range of 2–3 standard deviations are considered as outlier and are rejected.

The Fisher's Discriminant Ratio (FDR) is a common measure to explore the discriminating power of a feature to separate two classes. The FDR can be represented using the mean and variance of the individual classes i.e. $\{\mu_1,\mu_2\}$ and $\{\sigma_1^2,\sigma_2^2\}$ as (5.6).

$$FDR = \frac{(\mu_1 - \mu_2)^2}{(\sigma_1^2 + \sigma_2^2)} \tag{5.6}$$

Clearly, higher ranking based on FDR will be assigned to those features which have higher difference in the mean values and small standard deviation implying compact distantly located clusters. There are also few measures of feature ranking which are defined pairwise (for two selected features) e.g. divergence, Bhattacharya distance etc. A more generalized feature ranking technique are the following three measures based on scatter matrices (Theodoridis and Koutroumbas 2010). Scatter

matrix is considered as a generalization of FDR and quantifies how the features are scattered in the feature space.

$$J_1 = \frac{trace\{S_m\}}{trace\{S_m\}}, J_2 = \frac{|S_m|}{|S_m|}, J_3 = trace\{S_w^{-1} S_b\} \tag{5.7}$$

Any one measure among the three in (5.7) may be used to investigate which feature has a high 'between-class distance' and low 'within-class variance'. In (5.7) $\{S_m,S_w,S_b\}$ represents the *mixture* scatter matrix, *within-class* scatter matrix and *between-class* scatter matrix and can be mathematically represented as (5.8).

$$S_w = \sum_{i=1}^{c} P_i S_i, \quad S_b = \sum_{i=1}^{c} P_i (m_i - m_0)(m_i - m_0)^T, S_m = S_w + S_b \tag{5.8}$$

5.3.3 Dimensionality Reduction of the Feature Space Using PCA

Often with high dimensional feature vectors, one may wish to represent the same amount of information contained in the data with less number of variables or produce more meaningful features from the original feature vectors. The latent information in the data-set is often done using principal component analysis and is considered as an efficient tool for dimensionality reduction (Rogers and Girolami 2011; Semmlow 2004). Let us consider the old and new feature vectors are denoted by x_j and y_i respectively and the weights as w_{ij}.

$$y_i = \sum_{j=1}^{M} w_{ij} x_j \quad \forall i = 1, \cdots, N \qquad \Rightarrow Y = WX \tag{5.9}$$

The PCA is based on applying a linear transformation on the original feature vectors which would cause rotation and scaling of the vectors (Semmlow 2004). The PCA approach finds the transformation matrix W that will remove correlation amongst the features or decorrelate them while also reducing the number of variables without losing the information contained in the data. In PCA, the feature vectors are rotated until the covariance or correlation goes to 0. The principal components and the eigenvalues of the covariance matrix of the feature matrix are commonly determined by using singular value decomposition (SVD). In PCA, the feature vector X is decomposed into the diagonal matrix D containing the square root of the eigenvalues and the U represent the principle component matrix.

$$X = U^* D^{1/2} U' \tag{5.10}$$

The eigenvalues of the covariance matrix describe how much variance is explained by the principal components and the eigenvalues are sorted in decreasing order of magnitude. This information is helpful in getting the insight which principal component is significant and which is not. Principal components,

corresponding to eigenvalues almost closer to 0 has less significance and can be eliminated to reduce the size of the feature vectors. The total variance of the feature vectors (with zero mean) is equal to the sum of the eigenvalues. It is a common practice that to represent a fraction of the total variance (say 90 %), the insignificant eigenvalues and the associated principal components are neglected that falls below that fraction with respect to the sum of all eigenvalues. The scree plot i.e. the plot of eigenvalues versus the principal component numbers (in decreasing order of importance) are also used to quantify the accuracy of the dimensionality reduction process in the new hybrid feature space.

Apart from the linear method of dimensionality reduction, the well-known kernel PCA can be used when the data lie in a nonlinear manifold. In this method the data is first transformed into a higher-dimensional space using a nonlinear mapping and then PCA is applied on the transformed data. Usually, a polynomial and exponential kernel function are used in this method (Theodoridis and Koutroumbas 2010).

5.3.4 *Feature Selection*

Feature selection deals with finding a subset from a large pool of available features to get an acceptable level of classification accuracy. Selecting highly informative features involve retaining only those features having a large between-class distance and a small within-class variance. After data normalization and removing the outliers, the feature selection technique is applied. The feature selection techniques help in order to avoid the problem of over-fitting and mimicking small inconsistent patterns by the classifier. It also heavily depends on the number of total data-points available to train the classifier.

There are mainly two steps known as scalar feature selection and feature vector selection (Theodoridis and Koutroumbas 2010). The scalar feature selection is done using any one of the discussed feature ranking criteria like FDR or scatter matrices. The feature ranking does not take the correlation among the feature into consideration which is addressed in scalar feature selection. In the first step, all the features are ranked in decreasing order of their priority using some ranking or criteria say C. Considering i_1 as the index of the best feature, the correlation co-efficient of the first top ranked feature with all the other features are computed which are then used to calculate the next-best index i_2 for the second most important feature x_{i_2} as (5.11).

$$i_2 = \max_j \left\{ a_1 C_j - a_2 \left| \rho_{i_1,j} \right| \right\}, \quad j \neq i_1 \tag{5.11}$$

This index includes both the criterion C for j^{th} feature and the correlation-coefficient $\rho_{i_1,j}$ between the best feature and rest of the features leaving the best one ($j \neq i_1$) after weighting them by two user-defined constants $\{a_1,a_2\}$. Rest of the features are then ranked using the following relation (5.12).

$$i_k = \max_j \left\{ a_1 C_j - \frac{a_2}{k-1} \sum_{r=1}^{k-1} \left| \rho_{i_r,j} \right| \right\}, \quad j \neq i_r, \quad r = 1, 2, \cdots, k-1 \tag{5.12}$$

for $k = 3, 4, \cdots, m$. By this technique the average correlation with the previously correlated features are considered.

In feature vector selection, different combinations of features are considered which have survived from the previous step, to obtain the best combination. By selecting all possible combinations of the scalar feature selection, the classifier is trained and its performance is evaluated each time by say leaving one feature. In exhaustive search method all possible combinations are considered which are computationally intensive. Whereas several sub-optimal feature vector selection techniques are available like sequential forward and backward selection etc.

5.4 Overview of Different Classification Techniques

5.4.1 Optimal Bayesian Classifier

Theoretically the classical Bayesian classifier is the most optimal one over any other techniques, provided the pdfs of the classes are Gaussians (Duda et al. 2012; Theodoridis and Koutroumbas 2010). Given l dimensional feature vector $x = [x(1) \quad x(2) \quad \cdots \quad x(l)]^T \in \Re^l$ and a set of classes ω_i, $i = 1, 2, \cdots, c$, the task is to classify a pattern to one out of c classes. According to the Bayes theory

$$P(\omega_i|x)p(x) = p(x|\omega_i)P(\omega_i) \tag{5.13}$$

where, $p(x) = \sum_{i=1}^{c} p(x|\omega_i)P(\omega_i)$,

$P(\omega_i)$ is the a priori probability of class ω_i,

$P(\omega_i|x)$ is the a posteriori probability of class ω_i given the value of x,

$p(x)$ is the probability density function of x,

$p(x|\omega_i)$ is the class conditional probability of x given ω_i or likelihood of ω_i with respect to x.

According to the Bayesian classification technique, x is assigned to the class ω_i if

$$P(\omega_i|x) > P\left(\omega_j\middle|x\right), \quad \forall j \neq i \tag{5.14}$$

If the probability of occurrence is considered to be same for all classes, then (5.14) can be multiplied by the $p(x)$ in the two side of the inequality which gives similar expressions as in (5.13). Under this assumption, (5.14) can be rewritten as

$$p(x|\omega_i)P(\omega_i) > p\big(x|\omega_j\big)P\big(\omega_j\big), \quad \forall j \neq i \tag{5.15}$$

The Bayesian classifier theoretically does best possible (optimal) classification since it minimizes the probability of error, if the parameters characterizing the pdf are accurately known. As can be understood from (5.15) that the classification result depends on both the conditional probability of x given ω_i as well as the a priori probability of class ω_i, while considering the pdf of x i.e. $p(x)$ as constant. According to the central limit theorem the pdf of the sum of few statistically independent random variables tends to the Gaussian as the number of summands tends to infinity. Practically with a large of summands this criterion is generally met which attracted the attention of pattern recognition community to use Gaussian pdf in various applications as well as due to its mathematical tractability.

The multidimensional Gaussian pdf is represented by

$$p(x) = \frac{1}{(2\pi)^{l/2}|S|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T S^{-1}(x-\mu)\right) \tag{5.16}$$

where, mean vector $\mu = E[x]$ with E being the expectation operator,

covariance matrix $S = E[(x-\mu)(x-\mu)^T]$ and its determinant is $|S|$.

For the one-dimensional case (5.16) reduces to (5.17) with corresponding change of covariance as simply the variance σ^2 of the random variable x, which is often known as normal pdf denoted as $\mathcal{N}(\mu, \sigma^2)$

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \Re \tag{5.17}$$

For a two-dimensional vector with Gaussian pdf $\mathcal{N}(\mu, S)$ and covariance matrix $S = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$, the vectors are uncorrelated if $\sigma_{12} = 0$. The diagonal elements denote the individual variances along the X or Y direction in a scatter diagram of the data points. For $\sigma_1^2 = \sigma_2^2$, $\sigma_{12} = 0$, the data points will look like a circle (spherical for multivariable data). The scatter diagram takes an elliptic shape with the major axis lying along the greater value among $\{\sigma_1^2, \sigma_2^2\}$. Correlated multivariable data is always ellipsoidal in the scatter diagram even if the variances be the same. The slope of the major axis with the X-axis becomes positive for positively correlated two-dimensional data i.e. $\sigma_{12} > 0$ and the slope is negative for negatively correlated data denoting $\sigma_{12} < 0$.

In order to make Bayesian classifiers, uniform priors and Gaussian class conditional densities are generally assumed. Generally, to fit a D-dimensional Gaussian involves finding $(D + D + (D(D-1)/2))$ number of parameters i.e. D for the mean and $(D + (D(D-1)/2))$ for the covariance matrix. In naïve Bayes assumption, the D-dimensional multivariate class conditional densities are approximated with a product of D-univariate distributions (Rogers and Girolami 2011). Each univariate Gaussian class-conditional densities involve two parameters $\{\mu, \sigma^2\}$ i.e. 2D parameters in total for D-dimensional Gaussian. In naïve Bayes classifier,

the complexity or the number of parameters to be determined is reduced at the cost of reduced model flexibility. This essentially means the class conditionals are aligned only along the axes and hence cannot model any within class dependencies.

5.4.2 Minimum Distance Classifier

The minimum distance classifier is a slightly relaxed case of the optimal Bayesian classifier (Duda et al. 2012; Theodoridis and Koutroumbas 2010). The Bayesian classifier can be simplified under certain assumptions as follows in order to design the minimum distance classifier.

1. The classes are equi-probable
2. The data in all classes follow a Gaussian distribution with the pdf $\mathcal{N}(\mu, S)$
3. The covariance matrix is same for all classes
4. The covariance matrix is equal and all the diagonal elements of it are equal which reduces to $S = \sigma^2 I$, with I being the identity matrix of the same order as the dimension of the feature space.

Any unknown x can be assigned to class ω_i based different distance measures such as the Euclidean distance (d_E), Mahalanobis distance (d_M) etc. The Euclidean distance classifier assigns a pattern to a class whose mean is the closest to it with respect to the Euclidean norm or distance ($\|\cdot\|$).

$$\|x - \mu_i\| \equiv \sqrt{\left[(x - \mu_i)^T (x - \mu_i)\right]} < \|x - \mu_j\|, \quad \forall j \neq i \tag{5.18}$$

This minimum Euclidean distance classifier can be generalized or improved for multivariable data-set with different variances in each dimension with some correlation among them i.e. by removing the fourth assumption as the covariance matrix being diagonal with equal elements. Considering a common covariance matrix for the two classes defining the shape of the multivariable Gaussian pdf, any unknown x can be assigned to class ω_i based on the minimum Mahalanobis distance with the criterion

$$\sqrt{\left[(x - \mu_i)^T S^{-1} (x - \mu_i)\right]} < \sqrt{\left[(x - \mu_j)^T S^{-1} (x - \mu_j)\right]} \quad \forall j \neq i \tag{5.19}$$

In both the minimum Euclidean or Mahalanobis distance classifier design, the statistical properties of the classes i.e. mean and common covariance matrix are to be estimated using the well-known "maximum likelihood" parameter estimation for Gaussian pdfs. In many practical problems the parameters of the pdfs are not known. Assuming that the data has been generated from a Gaussian distribution, the task is to estimate mean and covariance characterizing the Gaussian pdf. Maximum likelihood (ML) is a traditional method of estimating unknown parameters of a Gaussian from a significantly large amount of experimental data.

Assuming a randomly generated l-dimensional data recorded for N samples i.e. $x_i \in \Re^l, i = 1, 2, \cdots, N$, the maximum likelihood estimate of the unknown mean and covariance of the Gaussian can be estimated as

$$\mu_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{5.20}$$

$$S_{ML} = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu_{ML})(x_i - \mu_{ML})^T \tag{5.21}$$

It is obvious that the ML estimate of the mean and covariance approaches their true value as $N \to \infty$ implying a dataset with large number of recordings. In fact, for practical implementation in biomedical applications, correct estimation of the covariance matrix is a very significant step and care should be taken when limited number of data points is available. With increase in less meaningful and discriminative features the dimension of the covariance matrix increases and the parametric variance of the ML estimates become high and becomes highly unreliable.

In many practical cases, it may so happen that the above assumption of the data generation process from a Gaussian pdf is not true. The underlying true pdf behind the data generation may be multimodal or many-peaked pdf. If the data cannot be efficiently fitted with a single Guassian pdf with constant mean and covariance, it is a common practice to represent the arbitrary pdf as the sum of several Gaussians with different mean and covariance $\mathcal{N}(\mu_j, S_j), j = 1, 2, \cdots, J$. This technique is known as *mixture model* which says any arbitrary pdf can be modeled as a linear combination of several (say J number of pdfs) with the value of J being sufficiently high.

$$p(x) = \sum_{j=1}^{J} P_j p(x|j) \tag{5.22}$$

where, $\sum_{j=1}^{J} P_j = 1, \quad \int p(x|j)dx = 1$

Similar to the ML estimation of the pdf parameters of a single Gaussian, for mixture models the *expectation-maximization* (EM) algorithm iteratively finds out the statistical properties of the J pdfs from the N data-points, $x_i \in \Re^l, i = 1, 2, \cdots, N$. Therefore the task is to estimate the a priori probabilities P_j, $j = 1, 2, \cdots, J$ and parameters associated with each of the terms $p(x|j)$, $j = 1, 2, \cdots, J$. Assuming each of the summand pdfs to be Gaussian with covariance $\sigma^2 I$ and

$$p(x|j) = \frac{1}{(2\pi)^{l/2}\sigma_j^l} \exp\left(-\frac{1}{2\sigma_j^2}(x-\mu)^T(x-\mu)\right), \quad j = 1, 2, \cdots, J \tag{5.23}$$

The unknown parameters which need to be estimated are the means m_j, $j = 1, 2, \cdots, J$ (lJ number of parameters) and covariances $S_j = \sigma_j^2 I$, $j = 1, 2, \cdots, J$

(J number of parameters). The expectation-maximization algorithm iteratively estimates these unknown parameters i.e. a priori-probability, mean and covariances of the summand pdfs starting from an initial guess. The technique is susceptible to the choice of initial guess of the unknowns and the choice of J. The popular method of choosing the number of summand pdfs is information-based criteria if the number of dense regions or clusters in the data vector can be identified. After obtaining the pdfs of the classes using the mixture models, the classical Bayesian classifier can be built as given by (5.15).

5.4.3 *Discriminant Analysis Based Classifiers*

The linear discriminant analysis (LDA) classification algorithm is capable of separating two different classes using a decision boundary that is linear. The decision boundary may be a straight line for considering two features, a plane for three features or a hyper-plane for the consideration of four or more features. The output predicted by the linear discriminator is given by (5.24).

$$y = \sum_{i=1}^{M} x_i w_i + b \tag{5.24}$$

where, y is the predicted class, M is the number of distinct features, x_i is the i^{th} feature. Classifier tunable parameters are given by weights (w_i) and bias or off-set (b). If the classifiers output is greater than 0.5, it signifies that the input data-point belongs to "class-A" (say). On the other hand if the classifier's output is less than or equal to 0.5, then the data point belongs to say "class-B". Often the bias in (5.24) is lumped as the last weight vector and by putting the constant 1 to the last input data set, for simplifying the calculation. Therefore the input feature vector becomes $x_i = [x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_M \quad 1]$ and the output of the classifier can be written as

$$y = \sum_{i=1}^{M+1} x_i w_i = Xw \tag{5.25}$$

where, w is the combined weight-bias vector and X is the matrix of features with last column being unity.

This classifier can be trained by minimizing the squared error of the predicted class and actual class which is commonly known as least square approach.

$$\varepsilon^2(w) = \sum_{i=1}^{N+1} \left(d_i - x_i^T w\right)^2 \tag{5.26}$$

Here, d is the training-set containing the correct class information for the data-set. For the minimization of (5.26), it is differentiated with respect to w and set to 0,

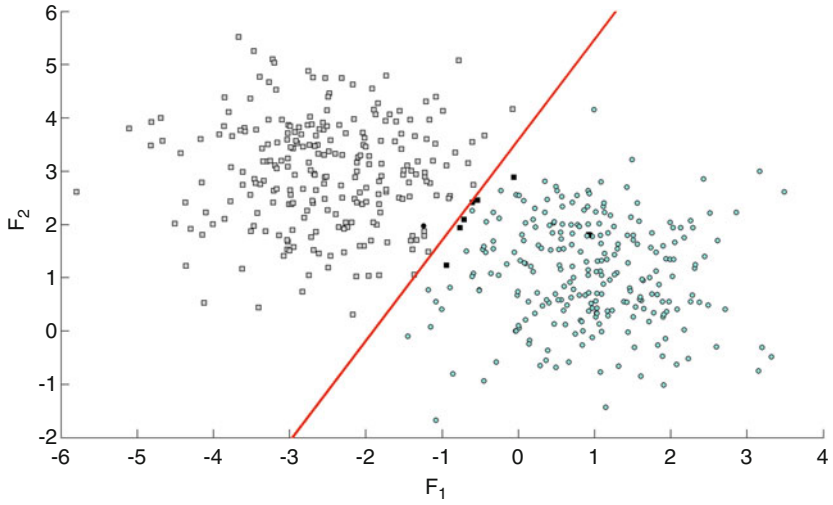


Fig. 5.1 Typical example of classification using LDA

which gives the classifier weights w in the form of pseudo-inverse of the inputs, multiplied by the class information vector.

$$w = \left(X^T X\right)^{-1} X^T d \tag{5.27}$$

In general the least square method does not give good generalization of the classification performance if small training set is used to determine the weights. In Fig. 5.1, one representative example of LDA classifier has been shown which separates two classes by drawing a decision boundary between them in 2D feature space. The black filled circle and boxes denote the misclassified data points for the two respective classes.

The classifier's performance can be evaluated as using confusion matrix and several measures sensitivity, specificity, accuracy, positive predictive value etc. The confusion matrix is often referred for multi-class problems. The principal diagonal of the confusion matrix indicates the number of correct classification and other off-diagonal elements of it represent various combinations of misclassification among the true and predicted classes. The confusion matrix is clinically highly significant in order to judge the performance of the classifier. It is due to the fact that incorrect classification of unhealthy persons as healthy may have more severe consequence than incorrect classification of healthy persons as unhealthy. For classification problems with two classes, the standard measures in biomedical applications are sensitivity and specificity. Often, the unusual class or abnormal condition is termed as positive (P) and normal class as negative (N). Correct detection or in other term classification of abnormal conditions is known as true positive (TP). Similarly correct classification of normal condition is true

negative (*TN*). Two type of incorrect classification may occur like classifying abnormal as normal (false negative) or *FN* and classifying normal as abnormal (false positive) or *FP*. The confusion matrix for two class problem is given by (5.28).

$$\text{Confusion matrix} = \overset{\text{Predicted class}}{\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}} \left.\vphantom{\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}}\right\} \text{True class} \tag{5.28}$$

The sensitivity and specificity are also known as true positive rate (*TPR*) and true negative rate (*TNR*) which are given by (5.29).

$$\begin{aligned} \text{Sensitivity or TPR} &= \frac{\text{TP}}{\text{TP}+\text{FN}} \times 100\% = \frac{\text{TP}}{\text{P}} \times 100\% \\ \text{Specificity or TNR} &= \frac{\text{TN}}{\text{TN}+\text{FP}} \times 100\% = \frac{\text{TN}}{\text{N}} \times 100\% \end{aligned} \tag{5.29}$$

Since the false negative is much more serious issue in a clinical sense, hence a classifier should show high sensitivity to unknown data-set before it can be used for detection of abnormalities in practice. The overall classification accuracy (*ACC*) can also be derived easily from the confusion matrix (5.28). There is another index to evaluate classifier's performance known as positive predictive value (*PPV*) which is equivalent to precision. The *ACC* and *PPV* are given by (5.30) in terms of different elements of the confusion matrix (5.28).

$$\begin{aligned} \text{Accuracy or ACC} &= \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \times 100\% = \frac{\text{TP}+\text{TN}}{\text{P}+\text{N}} \times 100\% \\ \text{PPV} &= \frac{\text{TP}}{\text{TP}+\text{FP}} \times 100\% = \frac{\text{TP}}{\text{TP}+(N-TN)} \times 100\% \end{aligned} \tag{5.30}$$

Different threshold values instead of 0.5 could have been set to check the corresponding sensitivity and specificity. A good classifier is judged by the trade-off between these two performance measures and then only a reliable threshold can be correctly determined (Semmlow 2004). This trade-off curve between sensitivity and (100% − specificity) is known as the receiver operator characteristic (ROC) curve. An improvement in classifier's performance is judged by the ROC curve by measuring how close the curve is to the left top corner as shown in Fig. 5.2. Commonly, the two axes of the ROC curve are represented as percentage or a fraction less than 1 (<100 %). For an ideal classifier the area under the ROC curve should be 1 and for a bad classifier predicting the classes randomly will have a ROC area as 0.5. Higher the area of ROC curve, better is the classifier's performance for varying threshold.

The LDA classifier gives good result if the data is linearly separable. For many practical biomedical applications, more complex decision boundaries may be required. The easiest way to address this problem is to use higher dimensional kernels.

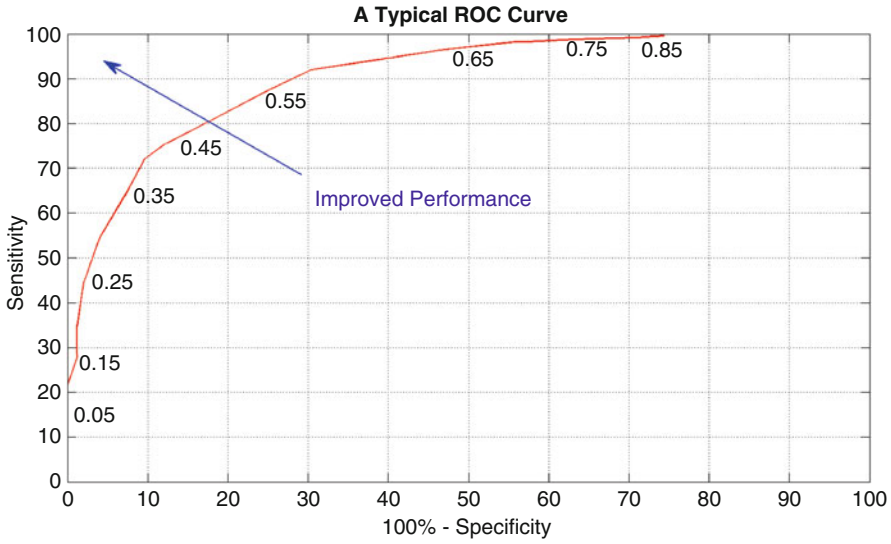


Fig. 5.2 A typical example of ROC curve

Any input feature vector may be transformed to high dimensional feature space using a polynomial kernel. The Cover's theorem says that it is always possible to obtain a linear boundary or hyper-plane if the order of the kernel is gradually increased when the data-set becomes linearly separable (Semmlow 2004). The main problem of this approach is the classifier would not generalize unless the training-set contains large number of data-points and also huge increase of the computational complexity. The curse of dimensionality is the main obstacle for using a generalized classifier from limited number of experimental data in most of the biomedical signal classification related tasks. To visualize the effect of higher order kernels, let us consider a 2D feature vector $\{x_1,x_2\}$. A quadratic kernel would map the feature vectors to form a new feature vector as $\{x_1, x_2, x_1^2, x_2^2, x_1x_2\}$. Clearly, the new feature vector consists of the original feature vectors, their squared values and their cross-products. Similarly, for a 3D feature vector $\{x_1,x_2,x_3\}$, use of a quadratic kernel will produce a new feature vector as $\{x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, x_2x_3\}$. Using these higher dimensional features for similar least square based discriminant analysis would produce the well-known quadratic discriminant analysis (QDA) classifier. Apart from polynomial kernel, other popular kernels are Gaussian, various trigonometric functions combined with polynomials etc.

5.4.4 Support Vector Machine (SVM)

The decision boundaries obtained by the least square method for the discriminant analysis are not the best or optimal, since it may vary with change in the size of the training data-set. The least-square method for training LDA/QDA classifiers gives

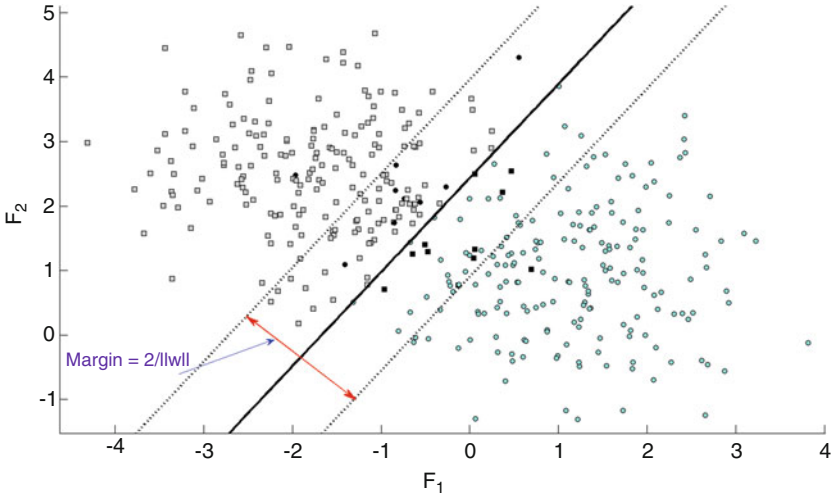


Fig. 5.3 Typical example of classification using SVM

too much emphasis on the data-points of the training set that are not critically lying between the two classes. For example in Fig. 5.3, an approach could be to maximize the distance between the critical points those are close to the other class. The points closest to the class separation boundary are called support vectors and the classifier that maximizes the distance between these critical support vectors is known as SVM classifier. Since SVM classifier maximizes the margin between the classes, it is also known as maximum margin classifier. If the data is linearly separable, the SVM classifier tries to find the optimum hyper-plane which maximizes the margin (M) using an optimization routine, subjected to a constraint that all the data points are lying on the appropriate side of the decision boundary.

To simplify the mathematical representation it is considered that the classes are assumed to have a class label of $y = \pm 1$ which implies the decision boundary to be represented as $y = 0$. If the input patterns, weight vector and the bias are represented as x_i, w, b respectively, the equation of the decision boundary can be written as (5.31).

$$y = x_i w + b = 0 \tag{5.31}$$

The equation of the lines going through the support vectors can also be written in as same way as (5.32).

$$\begin{aligned} x_i w + b &\geq 1 \quad \text{when } y = +1 \\ x_i w + b &\leq 1 \quad \text{when } y = -1 \end{aligned} \tag{5.32}$$

The above two equation can be combined as follows.

$$y_i(x_i w + b) \geq 1 \tag{5.33}$$

Equation (5.33) signifies that if the parameters $\{w,b\}$ can be chosen by a suitable optimization framework, the two classes will lie on the appropriate side of the support vectors. From (5.31), it can derived that the distance of the hyper-plane from the origin is $-\ b/\|w\|$. Here, $\|w\|$ denotes the norm of w given by .

$$\|w\| = \sqrt{w_1^2 + w_2^2 + \cdots + w_n^2} = w^T w \tag{5.34}$$

Also, the distance of the hyper-planes in (5.32) to the origin is $(\pm 1\ -\ b)/\|w\|$. This is can be used to find out the distance between two lines defined by (5.32) by simply subtracting the two distances as follows.

$$M = \frac{(1-b)}{\|w\|} - \frac{(-1-b)}{\|w\|} = \frac{2}{\|w\|} \tag{5.35}$$

Therefore, the maximum margin (M) can be obtained by minimizing $\|w\|$. The minimization is done using quadratic programming (QP) subjected to the constraint given by (5.33) which ensures that the boundaries are on the right side. The linear SVM is effective if the data is linearly separable. Otherwise, the data can be transformed in higher dimensional spaces using nonlinear kernels. Optimization process of the SVM classifier with high nonlinear kernels is computationally intensive and is prone to give over-fitting of the training data-set. In most real world applications, SVM is augmented with polynomial kernels due to its mathematical tractability, although complex kernels like Radial Basis Function (RBF), sigmoid, b-spline etc. are also possible (Gunn 1998).

As expected with higher order kernels the decision boundaries become complicated as shown in Fig. 5.4. The complex decision boundaries may perform well on the training data-set, but fail to generalize in most cases with increasing order of complexity for the kernel. Here, only an increase in the order of the polynomial kernel is shown from second order to fourth order, as an representative case which can easily be extended for even more complex kernels. In this situation, the classifier becomes more prone to capture the small inconsistent patterns among different data-sets of the underlying process. Therefore, multiple-run and multiple-fold cross validation must be done to generalize the classifier's discriminative capability.

5.4.5 k-*Nearest Neighbor (*k-*NN) Classifier*

In k-NN classification, when a new data point in the feature space, we first try to find k-nearest neighbors which are closest to the data point using some distance matrices like Euclidean or Mahalanobis. The new data point is assigned to the class in which

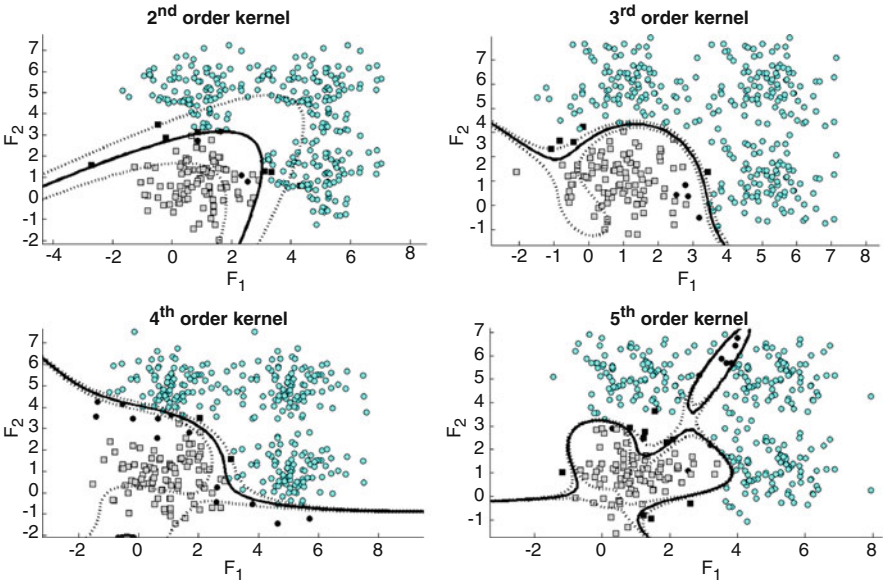


Fig. 5.4 Example of SVM classification decision boundaries with different order of the polynomial kernel

the majority of the *k*-nearest neighbors belong. One limitation of *k*-NN is when k is a multiple of the number of classes; there arise an issue of tie. In order to avoid this situation, for binary class classification for example, *k* should be always chosen as odd numbers (Rogers and Girolami 2011). The *k*-NN classifier tends towards the optimal Bayesian classifier as *k* tends to infinity. For very small value of *k*, the *k*-NN classification has less robustness against noisy data. Whereas, for very high value of *k*, the decision boundary becomes too smooth which may bias the misclassification rate. Also, the biggest implementation problem of *k*-NN in remote healthcare is its high computational complexity. The search for nearest neighbors repeats every time whenever a new data point comes in for getting classified to one of the classes. The search becomes even more complex for high dimensional feature space (Duda et al. 2012; Theodoridis and Koutroumbas 2010).

5.5 Least Square, Its Regularization to Maximum Likelihood and Bayesian Inference: Few Issues

The Least square estimation (LSE) technique derived to design LDA/QDA classifier can be obtained using an alternative approach. Let us define that x_n being the measurements and w being the classifier weights such that it minimizes the error

between the observations t_n and estimation, where the loss function is given by (5.36) and its derivative with respect to w gives the optimum weights of the classifier as shown in (5.37).

$$\begin{aligned}\mathcal{L} &= \frac{1}{N}\sum_{i=1}^{N}\left(t_n - w^T x_n\right) = \frac{1}{N}(t - Xw)^T(t - Xw)\\ &= \frac{1}{N}\left(t^T - X^T w^T\right)(t - Xw)\\ &= \frac{1}{N}\left(t^T t - t^T Xw - w^T X^T t + w^T X^T Xw\right)\\ &= \frac{1}{N}t^T t + \frac{1}{N}w^T X^T Xw - \frac{2}{N}w^T X^T t\end{aligned} \tag{5.36}$$

$$\begin{aligned}\frac{d\mathcal{L}}{dt} &= \frac{2}{N}X^T Xw - \frac{2}{N}X^T t = 0\\ &\Rightarrow X^T Xw = X^T t\\ &\Rightarrow \hat{w} = \left(X^T X\right)^{-1}X^T t\end{aligned} \tag{5.37}$$

The problem of over-fitting is inherent in the least square based training of the classifier as discussed earlier due to too much emphasis on the data-points which may contain outliers or inconsistent patterns. The Regularized Least Square (RLS) technique is capable of handling this problem by introducing a trade-off between the penalty for not fitting the data well and penalty for choosing to complex model of the classifier, with the help of an external parameter λ. Here, rather than penalizing the squared loss function, an extra term is added to penalize over-complexity as given in (5.38).

$$\begin{aligned}\mathcal{L}' &= \mathcal{L} + \lambda w^T w\\ &= \frac{1}{N}t^T t + \frac{1}{N}w^T X^T Xw - \frac{2}{N}w^T X^T t + \lambda w^T w\end{aligned} \tag{5.38}$$

The optimum value of w can now be found in a similar way as done above by setting the derivative of the new loss function to 0.

$$\begin{aligned}\frac{d\mathcal{L}'}{dt} &= \frac{2}{N}X^T Xw - \frac{2}{N}X^T t + 2\lambda w = 0\\ &\Rightarrow \left(X^T X + N\lambda I\right)w = X^T t\\ &\Rightarrow \hat{w} = \left(X^T X + N\lambda I\right)^{-1}X^T t\end{aligned} \tag{5.39}$$

The parameter λ can be chosen externally to control over-fitting and generalization capability. Making λ too small makes the model too complex and too high value of λ fails to capture underlying trends in the data-set (Rogers and Girolami 2011).

The weights of the classifier also changes with the number of data points in (5.39) which was not the case in the classical version of least square (5.37).

On contrary, the maximum likelihood approach treats the classification error as noise or random variable ε_n having some known pdf. In this case, the model would take the form (5.40).

$$t_n = w^T x_n + \varepsilon_n, \quad \varepsilon_n \sim \mathcal{N}\left(0, \sigma^2\right) \tag{5.40}$$

In (5.40), there are one deterministic ($w^T x_n$) and one random (ε_n) component which makes the prediction (t_n) also random. Since it is considered that the errors have Gaussian distribution and hence the predictions t_n have also the same variance σ^2 with its mean shifted from 0 to $w^T x_n$. i.e.

$$p\left(t_n | x_n, w, \sigma^2\right) = \mathcal{N}\left(w^T x_n, \sigma^2\right) \tag{5.41}$$

The quantity $p(t)$ cannot be termed as probability since the area under the curve is not equal to unity. For a particular value t_n the height of the curve gives the measure of how likely that particular observation may appear t_n given the data point x_n and thus is known as the likelihood function. The concept of finding parameters w and σ^2, such that it maximizes the likelihood function is known as the Maximum Likelihood (ML) technique. In fact, for a given problem we need to maximize the likelihood of all the data-points which can be assumed to be independent given a value for w. This reduces the problem of finding joint density of all data-points to the product of individual densities of all data-points. The likelihood then can be represented as

$$L = \prod_{n=1}^{N} p\left(t_n | x_n, w, \sigma^2\right) = \prod_{n=1}^{N} \mathcal{N}\left(w^T x_n, \sigma^2\right) \tag{5.42}$$

For analytical tractability, the natural logarithm of the likelihood function is minimized instead of the likelihood itself, in order to find out the optimum parameters of w and σ^2 by setting its partial derivatives with respect to the parameters to 0. Incidentally, for Gaussian distribution assumption of the error the optimum condition of the weights comes as (5.37) and the estimate of the variance comes as (5.21), which is the same that can be derived using the least square approach. The idea is discussed here because in many real world applications, the Gaussian pdf assumption of the error does not hold which will give different conditions for optimum weight selection and can be derived analytically if the expression for the pdf can be obtained. Also, in order to ensure that the maximum of the likelihood is obtained, its second derivative i.e. the Hessian matrix needs to be shown as negative definite. Just similar to the least square approach, the maximum likelihood technique also favors complex models and can be generalized by looking at the bias-variance trade-off of the model (Rogers and Girolami 2011).

In Bayesian approach of machine learning, where $p(r)$ is the belief in the light of new observation y_N, for the sake of analytical tractability, the likelihood ($P(y_n|r)$)

and prior ($p(r)$) conjugates pair are first searched for, where the posterior ($p(r|y_n)$) takes the same form as the prior e.g. Gaussian-Gaussian, Beta-Binomial, Gamma-Gaussian, Dirichlet-Multinomial etc. This helps to get rid of the problem of computing the denominator or the marginal likelihood ($P(y_n)$) in the Bayes rule (5.43) to obtain analytical expression of the posterior.

$$p(r|y_n) = \frac{P(y_n|r)p(r)}{P(y_n)} \tag{5.43}$$

The marginal likelihood or the denominator which is often challenging to derive acts as a normalizing constant to ensure the posterior density is properly scaled and can be computed by integrating r out of the joint density or likelihood.

$$P(y_N) = \int_{r=0}^{r=1} p(y_N, r) = \int_{r=0}^{r=1} P(y_N, r)p(r) \tag{5.44}$$

For Gaussian density consideration of the prior as $\mathcal{N}(Xw, \sigma^2 I_N)$ and likelihood as $\mathcal{N}(\mu_0, S_0)$, the maximum a posteriori (MAP) estimate of the Gaussian posterior becomes $\mathcal{N}(\mu_w, S_w)$ with the covariance and mean of the posterior density as

$$S_w = \left(\frac{1}{\sigma^2}X^T X + S_0^{-1}\right)^{-1}, \mu_w = S_w\left(\frac{1}{\sigma^2}X^T t + S_0^{-1}\mu_0\right) \tag{5.45}$$

After obtaining the posterior density, an expectation with respect to this density will give us the class information. If the posterior density is not of any standard form, although the numerator from Bayes' rule (5.43) can be obtained by multiplying the prior (commonly considered as Gaussian) and the likelihood, getting the expression for marginal likelihood (5.44) is difficult to get analytically by performing the integration. Therefore, it is a common approach to maximize the numerator in (5.43) instead of the posterior which is known as the MAP estimate of the classifier weights w. The Newton–Raphson method is often used to maximize the logarithm of the numerator which is proportional to the posterior density.

In machine learning community, tricky posterior densities are often approximated by Gaussian density which works reliably around the mode and the performance decreases away from the mode. This is known as the Laplace approximation and can be applied easily to approximate any density for which we can find out the mode. Also, in order to calculate expectation the true posterior often becomes analytically intractable. The Metropolis-Hastings (MH) algorithm is a well-known technique to directly sample from the posterior without obtaining any analytical expression of the posterior. The three techniques i.e. MAP while maximizing the numerator of Bayes rule, approximating any posterior density with Gaussian and directly sample from the posterior using MH algorithm to compute expectation, together made the foundation of Bayesian inference in machine learning (Rogers and Girolami 2011).

5.6 A Practical Case Study: Scar Classification from ECG Signal

There are several literatures dealing with ECG signal classification (Gacek 2012) but recent investigations has reported that the myocardial scar can even be characterized from the ECG signal, using the well-known Selvester scoring (Strauss et al. 2008). Here, a practical example of scar classification has been shown from 12-lead ECG signals. An ECG database with 108 scar patients and 46 healthy subjects, provided by the Cardiology Department of the University Hospital Southampton NHS Trust, has been used to build various classifiers. The percentage of scar was confirmed from Cardiac Magnetic Resonance Imaging (CMR) and >0 % scar has been considered as the abnormal class and with no scar from CMR as the normal class. The classifiers are trained with ECG data-sets of 10 s with a sampling frequency of 500 Hz. The focus of the present study is to look at the cross-lead information in time and frequency domain and built the classifier using physically meaningful features in amplitude and phase domain of power spectrum and auto/cross-correlation domains. Since, myocardial scar is manifested as a conduction delay in the ECG signal propagation path, considering the same source in the heart, the phenomena is expected to be maximally captured in few of the leads while the ECG signal is passed through the scar tissue while other leads may give completely healthy ECG signal. Therefore, the cross-lead features are explored here for myocardial scar classification from ECG signal which may serve as a preliminary screening before sending the patient for CMR.

5.6.1 Feature Extraction, Ranking and Selection for 12-Lead ECG

For both healthy and myocardial scar subjects the following four cross-lead domains have been explored for feature extraction:

1. Cross power spectrum domain
2. Frequency weighted cross-power spectrum domain
3. Coherence domain
4. Cross-correlation domain

As a pre-processing of the ECG signal before feature extraction, we applied mean removal and band-pass filtering in a frequency range of [1,30] Hz. In Fig. 5.5, we have shown a representative example of two ECG signals $x(t)$ and $y(t)$ respectively, for lead-I and lead-II. Considering the corresponding Fourier transforms as $X(f)$ and $Y(f)$ respectively, the cross power spectrum can be calculated as expectation of the product of complex frequency spectrum of the former and the conjugate of the latter (Nolte et al. 2004).

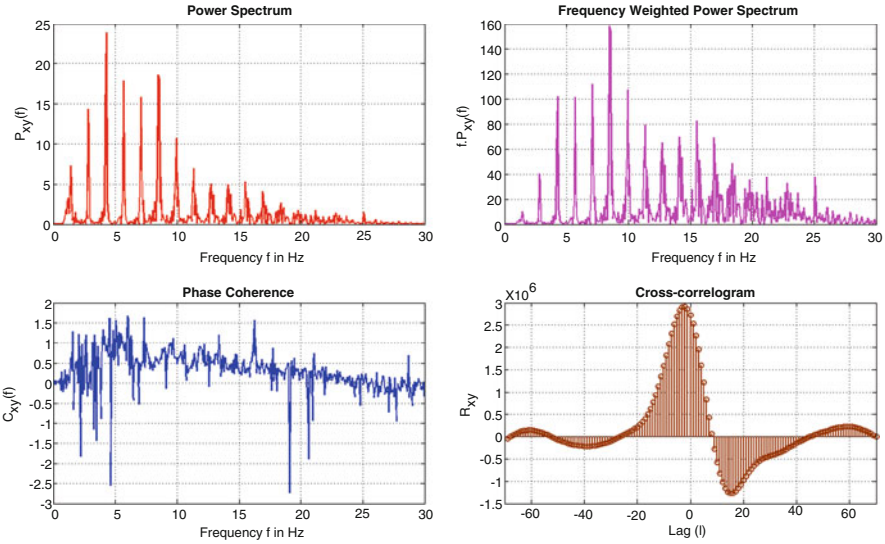


Fig. 5.5 Four cross-lead domains for feature extraction

$$P_{xy}(f) = E[X(f)Y^*(f)] \tag{5.46}$$

The coherence of the two signals is given by the cross-spectrum normalized by the respective power spectral densities (PSDs) $P_{xx}(f)$ and $P_{yy}(f)$. It is well known that the magnitude of the coherence in (5.47) is always one due to the normalization but its phase is informative and expected to capture the conduction delay in scar tissue within the heart.

$$C_{xy}(f) = \frac{P_{xy}(f)}{\sqrt{P_{xx}(f)P_{yy}(f)}} \tag{5.47}$$

The autocorrelation function is given in terms of the time shift parameter or lag l

$$R_{xy}(l) = E[x(n)y(n+l)] \tag{5.48}$$

where, E is the expectation operator.

Four set of features have been generated from each combination of two ECG leads as follows. Cross power can be obtained by integrating the cross-power spectral density within the frequency range of interest i.e. 30 Hz (Sörnmo and Laguna 2005).

$$\text{Cross-power}_{xy} = \underset{\varepsilon \to 0}{Lim} \int_{\varepsilon}^{30} P_{xy}(f)df \tag{5.49}$$

The integral in the above equation has been evaluated using the well-known trapezoidal rule of numerical integration. For any two given signal the area under

the drooping power spectrum may be closer. Still there may exist high peaks in the higher frequency regions which motivated us to introduce another index to capture these particular frequency domain characteristics. The frequency weighted cross spectrum is introduced as follows:

$$\text{Freq weighted cross-power}_{xy} = \underset{\varepsilon \to 0}{Lim} \int_{\varepsilon}^{30} f \cdot P_{xy}(f) df \tag{5.50}$$

The frequency averaged phase coherence or mean phase coherence (Mormann et al. 2000) is given by (5.51).

$$\text{Frequency averaged coherence}_{xy} = \frac{1}{N} \sum_{i=1}^{N} C_{xy}(i), \quad N \cdot F_s = 30 \text{ Hz} \tag{5.51}$$

where, $\{N,F_s\}$ are the number of data points within the frequency spectrum of 30 Hz and sampling frequency respectively.

The centroid of the cross-correlogram (Chandaka et al. 2009) is given by (5.52).

$$\text{Centroid-Correlogram}_{xy} = \frac{\sum_{m=-M}^{M} mR_{xy}(m)}{\sum_{m=-M}^{M} R_{xy}(m)} \tag{5.52}$$

For a good estimation of the auto-correlation function the maximum lag M should be small enough compared to the whole data-set and has been taken as 200 samples in the present study.

Since, the cross-lead information are the same for *x vs. y* and *y vs. x*. Thus we have unique 78 combinations of each feature by combining 12 lead ECG signals in groups of two. Each 12-lead cross-electrode information can be represented as a symmetric matrix where the diagonal terms are simply corresponding to the autocorrelation function and (auto) power spectral densities for each lead. The diagonal terms for the frequency averaged coherence domain are 0. The above mentioned features were generated for each pairs from the 12 lead ECG data i.e. I, II, III, aVR, aVL, aVF, V1–V6. Mean, variance, skewness and kurtosis have been calculated from all the 78 unique combinations formed by 12 lead ECG signals.

Using the cross-lead informations we now calculate four basic first order statistical measures (mean, variance, skewness and kurtosis) along four domains i.e. correlation, power spectrum, frequency weighted power spectrum and coherence domains. For exploring the discriminative nature of the obtained features to separate the scar and non-scar subjects, Fisher's discriminative ratio (FDR) has been used to rank them according to their decreasing order of importance, as shown

Table 5.1 Feature ranking based on FDR

Feature ranking	FDR	Feature description
F_1	0.064055512	Kurtosis of the frequency averaged coherence
F_2	0.060913144	Variance of the frequency weighted cross-power
F_3	0.037831304	Skewness of the cross-power
F_4	0.037293165	Variance of the cross-power
F_5	0.03096473	Mean of the cross-power
F_6	0.029735205	Skewness of the frequency weighted cross-power
F_7	0.01983636	Skewness of the centroid of cross-correlogram
F_8	0.019668996	Variance of the frequency averaged coherence
F_9	0.018362378	Mean of the frequency weighted cross-power
F_{10}	0.017618007	Mean of the centroid of cross-correlogram
F_{11}	0.008805947	Mean of the frequency averaged coherence
F_{12}	0.005730008	Variance of the centroid of cross-correlogram
F_{13}	0.0044085	Kurtosis of the cross-power
F_{14}	0.004324487	Kurtosis of the centroid of cross-correlogram
F_{15}	0.00305714	Skewness of the frequency averaged coherence
F_{16}	0.001385125	Kurtosis of the frequency weighted cross-power

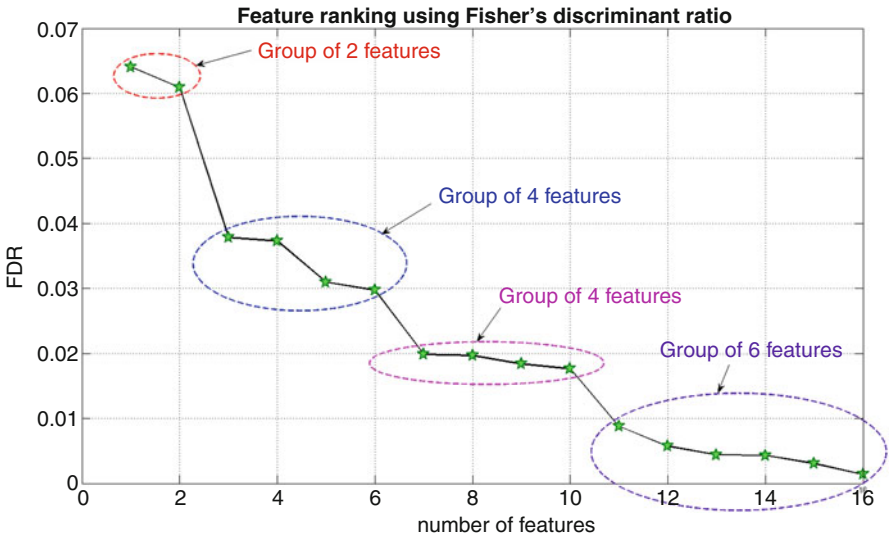


Fig. 5.6 Ranking of 16 features using Fisher's discriminant ratio

in Table 5.1. It is observed that the coherence and cross-power spectrum domain features are ranked higher than the correlation domain features.

The above mentioned 16 features are now plotted in Fig. 5.6 with decreasing order of importance using the FDR criterion (5.6). From Fig. 5.6, it is evident that there exist four groups, i.e. first group with top 2 features and next groups with four,

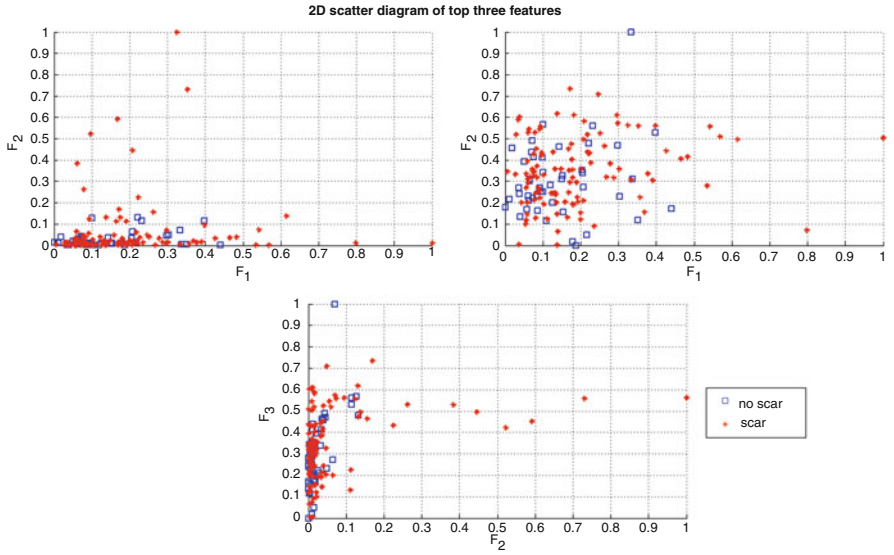


Fig. 5.7 2D scatter diagram of top three features for the two classes

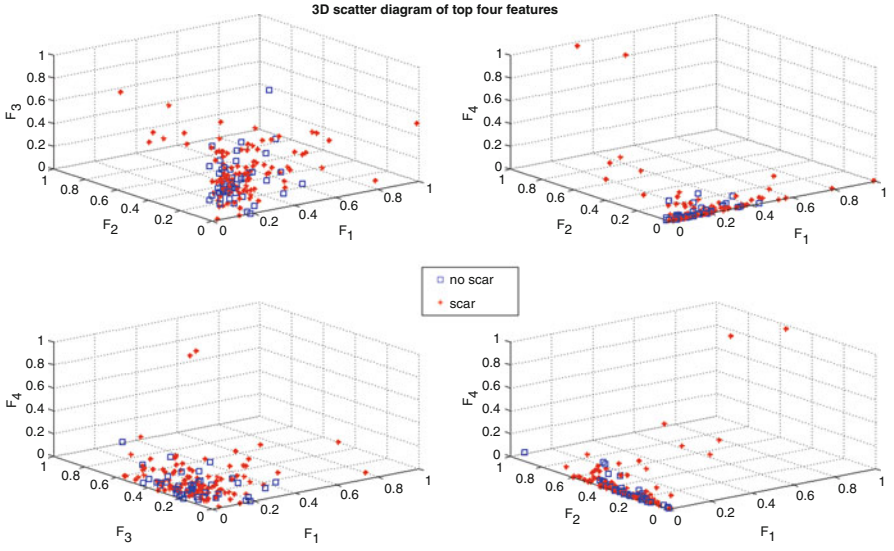


Fig. 5.8 3D scatter diagram of top three features for the two classes

four and six features respectively. Classification accuracies are compared in the following section with these ranked features based on FDR. Also, the scatter diagram of the most discriminative three and four features in Table 5.1 are shown in Figs. 5.7 and 5.8 respectively. From the scatter diagrams it is evident that the two

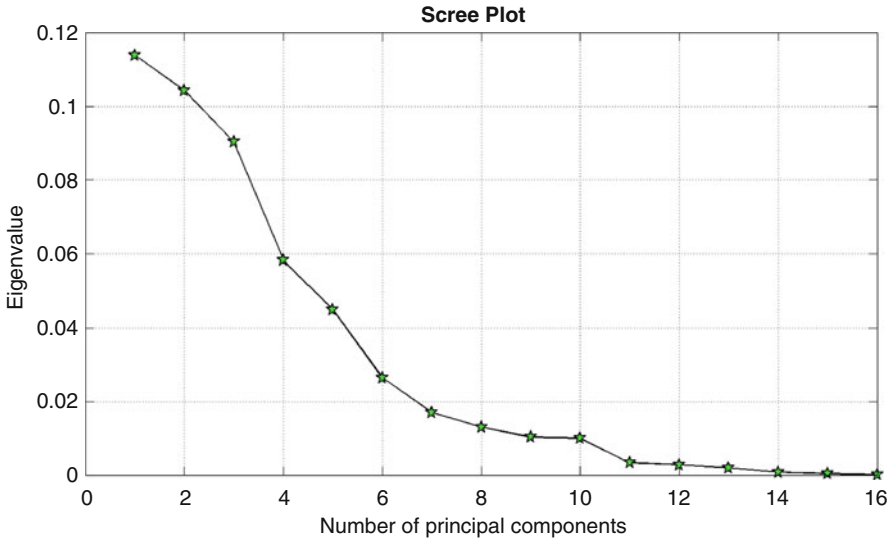


Fig. 5.9 Eigenvalues of the PCs in decreasing order

Table 5.2 Number of principal components to considered

Variance explained (%)	Principal components to be considered
100	All 16 PCs
>99	PCs 1–12
>98	PCs 1–10
>95	PCs 1–9
>90	PCs 1–7
>80	PCs 1–5

classes overlapped and not geometrically isolated by simple boundaries in the feature space. This motivates us to use higher order kernels in the classifiers for the particular binary class classification problem. Also, from the above listed 16 features, we applied the principal component analysis for dimensionality reduction of the feature set. The scree plot, shown in Fig. 5.9 indicates that first 5 or 7 hybrid features obtained using PCA are highly informative and rest of the hybrid features may not at all be considered without significantly loosing much information contained in the original feature vectors. In Table 5.2, the total variance explained by top PCA based hybrid features are shown. It is evident that 98 % information in the original feature set is contained within the top 10 principal components. The main advantage of taking the principal components as features is that they are no longer correlated as was the case for the original feature vector.

5.6.2 *Parametric and Non-parametric Classification with Ranked Features and the Principal Components*

The classification of subjects with scar is attempted with linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) using the feature groups obtained from fisher ratio analysis (5.6). From the plot of Fisher ratios for all the 16 features under consideration in descending order of importance in Fig. 5.6, we found that there exist four groups i.e. F_1–F_2, F_3–F_6, F_7–F_{10}, F_{11}–F_{16}. We trained the LDA and QDA classifiers with different feature sets of decreasing importance i.e. top 2, 6, 10 and all 16 features. Since the LDA and QDA classifiers' weights can be easily determined by simple matrix inversion, it is quite fast compared to the other non-parametric classification techniques like SVM and *k*-NN. Due to its low computational complexity, we ran the experiments to simulate tenfold cross validation with 100 runs of randomized data set. Similar analysis has also been carried out with the top 5, 7, 9 and 10 hybrid PCA based features, as ranked in Fig. 5.9 and Table 5.2. The average accuracies for the LDA and QDA classifiers are given in Table 5.3 with top FDR based and PCA based features. It can also be observed from Table 5.3 that as the number of features increases the average accuracy falls gradually. Therefore, to get a good accuracy and relatively low computational complexity for remote healthcare applications, the best ranked features may be reliably used for classification. It is also important to note that except the top 2 features, increasing the number of features has worsened the average classification accuracy for QDA compared to LDA. Therefore, LDA can be considered as a more robust classifier than QDA subjected to equal increase in number of relatively less significant features. This particular observation signifies that though the class separation boundaries may be better obtained using higher dimensional kernels, in many cases it complicates the whole nonlinear mapping of features space which may not be consistent for multiple trials of classifier design.

For support vector machine classifier, we used only top 2 and top 3 features obtained from FDR ranking and the top 2–3 principal components. Since the SVM includes an inherent quadratic programming optimization routine to find out the support vectors, it is quite slow compared to the discriminant analysis based classifiers. Therefore, we here opted for tenfold cross-validation with 10 runs of randomized data-sets. For SVM we have only chosen increasing the order of the polynomial kernel among many other options like use of radial basis function (RBF), b-spline etc. The fall in the average accuracy of 10 runs of tenfold cross-validation with higher dimensional kernels in SVM is more prominent with the PCA based hybrid features in Table 5.4. This suggests that the maximum separation boundary can be reliably found with less complex kernels and this is also beneficial from hardware implementation point of view.

Since the problem under consideration is a binary class classification problem, the choice of *k* in the *k*-NN algorithm should be an odd number and the optimum setting may be achieved by varying the value of *k*. Table 5.5 shows the average accuracy of the top two FDR based features and top two principal components while increasing the value of *k* from 3 to 25. As discussed before, high value of

Table 5.3 Performance of LDA and QDA classifier with different set of features with FDR based feature ranking and PCA based dimension reduction

	FDR based		PCA based	
Classifier	Feature set	Accuracy	Feature set	Accuracy
LDA	Top 2	71.42857	Top 5 PCs	70.75714
	Top 6	70.82857	Top 7 PCs	70.60714
	Top 10	69.26429	Top 9 PCs	70.35714
	All 16	65.85714	Top 10 PCs	70.47143
QDA	Top 2	71.42143	Top 5 PCs	64.67143
	Top 6	63.00714	Top 7 PCs	63.41429
	Top 10	54.5	Top 9 PCs	60.50714
	All 16	50.52143	Top 10 PCs	60.25714

Table 5.4 Performance of SVM classifier with different set of FDR/PCA based features and increasing the kernel order/complexity

Feature set	Kernel	Accuracy with FDR based features	Accuracy with PCA based features
Top 2 features or principal components	Linear	71.42857	70.85714
	Polynomial 2	71	69.92857
	Polynomial 3	67.28571	70.35714
	Polynomial 4	67.07143	67.71429
	Polynomial 5	68.35714	66.78571
Top 3 features or principal components	Linear	71.28571	71.07143
	Polynomial 2	66.71429	69.57143
	Polynomial 3	65.71429	65.92857
	Polynomial 4	63.42857	60.92857
	Polynomial 5	63.35714	62.21429

Table 5.5 Performance of k-NN classifier with top two features and varying k

k	Accuracy with top two FDR based features	Accuracy with top two PCA based features
3	59.92857	63.07143
5	63.78571	64.85714
7	65.85714	66.64286
9	68.64286	65.85714
11	69.28571	66.92857
13	68.64286	67.78571
15	69.64286	69.28571
17	70.64286	69.14286
19	71.42857	69.92857
21	71.14286	70.57143
23	71.14286	71.21429
25	71.28571	71.42857

k smoothens the decision boundary and becomes robust against noisy data. From the above case study, it is observed that the maximum accuracy of 71.4 % is achieved by the LDA classifier with the top two FDR based features.

5.7 Overview of Clustering Techniques

Clustering falls under the category of unsupervised learning where the class labels are considered to be unknown for the given input patterns. The task is to find out clearly identifiable groups or clusters in the feature space, considering that similar patterns falls in one cluster and less similar clusters falls in other different clusters. Any pattern may be considered to fall in one cluster or it may belong to many clusters simultaneously which brings the concept of hard clustering and soft clustering (Theodoridis and Koutroumbas 2010). In biomedical applications, the most popular algorithms are based on a cost function optimization method e.g. application of k-means, fuzzy c-means algorithm in EEG (Günes et al. 2010; Wahlberg and Lantz 2000). Most of the advanced clustering techniques have been applied in biomedical image processing (Gonzalez and Romero 2010) which requires lots of computational resources compared to simpler physiological signals like EEG, EMG etc.

5.7.1 Hard Clustering

In hard clustering e.g. the k-means clustering algorithm, every data point in the data-set X is considered to lie strictly in any one of the clusters C_j which are again represented by their mean values θ_j (cluster representative). This algorithm assumes the number of underlying clusters in the data-set (m) is known and the algorithm tries to move θ_j into the regions where data-points are densely located. The k-means algorithm starts using initial guess values for the cluster representation and the solution is refined in an iterative way by minimizing a cost function. The initialization should not be too close for different clusters since the distance of the data points from θ_j in the early stage of the algorithm would be more or less same which may produce wrong estimation of the cluster representation. The diversity in the initial guess for initial guesses of θ_j helps to get a good solution but too far initialization (generally more than few standard deviations of any feature) also fails to reach the true center of the cluster. The cost function in k-means clustering algorithm can be represented as

$$J(\theta, U) = \sum_{i=1}^{N} \sum_{j=1}^{m} u_{ij} \left\| x_i - \theta_j \right\|^2 \tag{5.53}$$

Here, $\theta = \begin{bmatrix} \theta_1^T & \cdots & \theta_m^T \end{bmatrix}^T$, $\| \cdot \|$ represents Euclidean distance and $u_{ij} = 1$ if x_i lies closest to θ_j, otherwise $u_{ij} = 0$. The k-means clustering finds the minimum of J, when all θ_j lies almost at the center of each cluster. In this algorithm, the number of

cluster representatives (m) should be estimated correctly. If the number of representative is larger than the true number of clusters, at least one physical cluster will be identified as more, whereas with less number of representatives two or more number of clusters may be identified by one representative. Generally, the k-means algorithm is run several times with random initial guess and by increasing the value of m. In the J_m vs. m plot, the location at which a significantly sharp *knee* is observed, it can be considered as the optimal number of clusters underlying the data. If no such knee is observed with increasing m, it can be concluded that there is no compact cluster in the data-set. Among other hard clustering techniques the Generalized Mixture Decomposition Algorithmic Scheme (GMDAS) is also very popular (Theodoridis and Koutroumbas 2010). This algorithm iteratively finds out the mixing probabilities, and the mean, covariances of the mixture of Gaussian model as shown in Sect. 5.4.2. The GMDAS algorithm is much computationally intensive than the k-means algorithm.

The k-means clustering is mostly done with Euclidean and Mahalanobis distance measures as already introduced in Sect. 5.4.2. The Euclidean distance based clustering gives optimum performance for spherical shaped clusters in the features space whereas the Mahalanobis distance based clustering can identify ellipsoidal clusters. In literature for example in (Mao and Jain 1996), there is evidence that the Mahalanobis distance based clustering often identifies closely located clusters as one big cluster and remotely located outliers as another cluster as this particular technique gives too much importance to the variance at the time of formation of the clusters. Thus it is recommended in (Mao and Jain 1996), to start with k-means clustering with Euclidean distance (d_E) and then increasing the effect of Mahalanobis distance (d_M) to make a balance. The distance measure (d) in such cases can be represented as a weighted sum of d_E and d_M, where the weight (λ) may be varied from 1 to 0, over the iterations due to unavailability of large number of samples, required to estimate the covariance matrix.

$$d = \lambda d_E + (1 - \lambda) d_M \tag{5.54}$$

5.7.2 Soft Clustering

The most popular soft clustering technique i.e. fuzzy c-means algorithm (FCM) assumes that any data point can simultaneously belong to more than one cluster up to a certain degree. The similar variable u_{ij} like in k-means algorithm, signifies here the degree of membership of data-point x_i in cluster C_j. The cost function in FCM is given by

$$J(\theta, U) = \sum_{i=1}^{N} \sum_{j=1}^{m} u_{ij}^{q} \left\| x_i - \theta_j \right\|^2 \tag{5.55}$$

Here, $\theta = \begin{bmatrix} \theta_1^T & \cdots & \theta_m^T \end{bmatrix}^T$ represent the cluster representative. The optimization is carried out subjected to the constraint $u_{ij} \in [0, 1], \sum_{j=1}^{m} u_{ij} = 1$. Typical values

of the fuzzifier lies in the interval $q \in [1.5,3]$. Rest of the part of the algorithm like initialization and finding dense clusters in the feature space is similar to that in k-means clustering. The FCM algorithm returns the cluster representatives θ_j and the degree of membership of all data-points x_i in the cluster m. The fuzzy c-means clustering algorithm is somewhat robust against the outliers compared to the k-means clustering algorithm. This is due to the fact that k-means contributes 100 % in the estimation of cluster representatives due to the hard boundary between the clusters, whereas the fuzzy c-means allows less contamination of the cluster representation estimates by the outliers due to having a concept of fuzzy membership or degree of association. The FCM has found several applications in the field of biomedical image segmentation (Ji et al. 2012), which is computationally more intensive than applications in bioelectrical signals.

There is another popular soft clustering technique known as possibilistic c-means (PCM) algorithm. This algorithm is quite similar to the FCM except the first constraint being relaxed, i.e. summation of all u_{ij} need not to be equal to one in this case though they may lie in the interval $u_{ij} \in [0,1]$. Hence, the u_{ij}'s for each x_i is no longer inter-related and do not therefore signify the degree of membership of x_i in cluster C_j. Rather, here u_{ij} represent the degree of compatibility between x_i and C_j and is independent of that between x_i and remaining clusters. The PCM algorithm is also less sensitive to the prior knowledge of underlying clusters and an overestimated value may be assigned for the number of clusters. Similar to FCM, the PCM returns the cluster representatives and degree of compatibility the data points.

Apart from the assumption of compact clusters, other shape of clusters can be easily identified with much complex algorithms like the spectral clustering etc. Other philosophies like representing the clustering as a nesting are mainly done by two ways viz. agglomerative and divisive hierarchical clustering. For hierarchical clustering the proximity dendogram is a most common tool of visualization (Theodoridis and Koutroumbas 2010).

5.8 Arithmetic Complexity Analysis and Critical Evaluation of Trade-offs Between Complexity and Accuracy

5.8.1 Complexity Analysis of Few Popular Classifiers

Hardware complexity of popular classification algorithms have been introduced in (Chen et al. 2013). Table 5.6 shows that the hardware complexity increases with the number of support vectors (M) and the number of training sample (S) for SVM and k-NN classifiers. Whereas, the complexity of the LDA and QDA classification algorithm is relatively low as it only depends on the dimension of the feature vector (N) and is therefore recommended for low power integrated circuit design. In terms

Table 5.6 Arithmetic operations involved in popular classification techniques

Classifier	Addition (+)	Multiplication (x)	Square ($(.)^2$)	Square-root ($\sqrt{}$)
LDA	N	N	0	0
QDA	$N^2 + N$	$N^2 + N$	N	0
SVM with linear kernel	$(N + 1)M - 1$	$(N + 2)M$	0	0
SVM with quadratic kernel	$(N + 2)M - 2$	$(N + 2)M$	M	0
k-NN	$2S(N + 1) - 6$	0	SN	S

of lowest computational complexity, the LDA is the best classifier since its complexity increases linearly with an increase in the feature vector dimension. Considering M and S comparable to N, it can be concluded that for the other classifiers', an almost quadratic increase in the complexity is obtained. Since as an arithmetic block, the hardware multiplier consumes more energy than adder, therefore the LDA will consumes less energy as compared to the others.

It has been shown in (Chen et al. 2013) that for two input NAND gate complexity and b-bit word-length implementation, the number of transistor requirement for basic arithmetic operations in Table 5.6 can be deduced as $T_+ = 24b$, $T_\times = T_{(\cdot)^2} = 30b^2 - 36b$, and $T_{\sqrt{\cdot}} = 18\left(\left(\frac{b}{2}+1\right)\right)\left(\left(\frac{b}{2}+3\right)\right)$, where T_* denote the transistor count for arithmetic operation (*). It is well known that a two input NAND gate requires four transistors, therefore the equivalent gate number can be derived from the above transistor numbers as $G_+ = 6b$, $G_\times = G_{(\cdot)^2} = \frac{15}{2}b^2 - 9b$, and $G_{\sqrt{\cdot}} = \frac{9}{2}\left(\left(\frac{b}{2}+1\right)\right)\left(\left(\frac{b}{2}+3\right)\right)$.

5.8.2 *Complexity Reduction of the Feature Extraction Process*

For statistical features involving higher moments of a random biomedical signals like mean (μ), variance (σ^2), skewness (γ) and kurtosis (β) the algorithms can be suitably modified to compute the parameters in an online fashion. This alleviates the requirement of the storing all the samples every time to compute the statistical features. For example online computation of two representative statistical features mean and variance are shown below. Considering a time series data $X = \{x_i\}, i \in \mathbb{N}_+, i \in [1, n]$, the mean is given as its expected value as follows.

$$\mu_n = E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{5.56}$$

Now the goal is to obtain the updated mean (μ_{n+1}) considering the next sample (x_{n+1}) from the old estimate μ_n as follows.

$$\begin{aligned}\mu_{n+1} &= \frac{1}{n+1}\sum_{i=1}^{n+1} x_i = \frac{n}{n+1}\cdot\frac{x_1+x_2+\cdots+x_{n+1}}{n} \\ &= \frac{n}{n+1}\left(\frac{x_1+x_2+\cdots+x_n}{n}+\frac{x_{n+1}}{n}\right) = \frac{n}{n+1}\left(\mu_n+\frac{x_{n+1}}{n}\right) \\ &\Rightarrow \mu_{n+1} = \frac{1}{n+1}(n\mu_n + x_{n+1})\end{aligned} \tag{5.57}$$

From (5.57), it is clear that the updated mean can be obtained without the requirement of storing the whole data-set in a buffer to extract its statistical property. The old to new estimate of mean is only related by simple algebraic operations which are suitable in hardware circuit implementation and a counter (n).

Similar treatments can be done for online calculation of variance as well where the new variance (σ_{n+1}^2) can be represented as a function of old variance (σ_n^2), old mean (μ_n), and the new mean (μ_{n+1}). This online implementation scheme is far less complex from the perspective of hardware design as the buffer requirement for storing the time series is alleviated. The variance of the time series is given by (5.58).

$$\begin{aligned}\sigma^2 &= E[X-\mu]^2 = E\left[X^2\right] - 2E[X]\mu + \mu^2 \\ &= E\left[X^2\right] - \mu^2 \\ &\Rightarrow \sigma_n^2 = \frac{1}{n}\left(x_1^2+x_2^2+\cdots+x_n^2\right) - \mu_n^2\end{aligned} \tag{5.58}$$

Therefore, the updated variance considering the next sample becomes

$$\begin{aligned}\sigma_{n+1}^2 &= \frac{1}{n+1}\left(x_1^2+x_2^2+\cdots x_n^2+x_{n+1}^2\right) - \mu_{n+1}^2 \\ &= \frac{n}{n+1}\cdot\frac{1}{n}\cdot\left(x_1^2+x_2^2+\cdots x_n^2+x_{n+1}^2\right) - \mu_{n+1}^2 \\ &= \frac{n}{n+1}\left(\frac{x_1^2+x_2^2+\cdots x_n^2}{n}+\frac{x_{n+1}^2}{n}\right) - \mu_{n+1}^2 \\ &= \frac{n}{n+1}\left(\frac{x_1^2+x_2^2+\cdots x_n^2}{n} - \mu_n^2 + \mu_n^2\right) + \frac{x_{n+1}^2}{n+1} - \mu_{n+1}^2 \\ &\Rightarrow \sigma_{n+1}^2 = \frac{n}{n+1}\left(\sigma_n^2+\mu_n^2\right) + \frac{x_{n+1}^2}{n+1} - \mu_{n+1}^2 \\ &= \frac{1}{n+1}\left[n\left(\sigma_n^2+\mu_n^2\right)x_{n+1}^2\right] - \mu_{n+1}^2\end{aligned} \tag{5.59}$$

From (5.59), it is evident that the updated variance can easily be computed using the old estimates σ_n^2 and μ_n^2, new mean μ_{n+1}^2 given in (5.57), the new sample x_{n+1} and the sample counter n.

Among higher order moments the skewness (γ) and kurtosis (β) are often used for characterizing non-Gaussian distributions of data-set and have been widely used in the statistical pattern recognition community as features for data classification. Several distributions like student's t-distribution, Rayleigh distribution, Laplace distribution, exponential distribution, Poisson distribution, logistic distribution are known as super-Gaussian whereas the Bernoulli distribution is known sub-Gaussian based on the fact whether they are leptokurtic and platykurtic.

The skewness (γ) of the time series is given by

$$\gamma = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{E[X^3] - 3E[X^2]\mu + 3E[X]\mu^2 - \mu^3}{\sigma^3}$$
$$= \frac{E[X^3] - 3\mu(E[X^2] - \mu^2) - \mu^3}{\sigma^3} = \frac{E[X^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3}$$

$$\Rightarrow \gamma_n = \frac{\frac{1}{n}\left(x_1^3 + x_2^3 + \cdots + x_n^3\right) - 3\mu_n\sigma_n^2 - \mu_n^3}{\sigma_n^3} \tag{5.60}$$

Therefore, the updated skewness considering the next sample becomes

$$\gamma_{n+1} = \frac{\frac{1}{n+1}\left(x_1^3 + x_2^3 + \cdots + x_n^3 + x_{n+1}^3\right) - 3\mu_{n+1}\sigma_{n+1}^2 - \mu_{n+1}^3}{\sigma_{n+1}^3}$$
$$= \frac{\frac{n}{n+1}\cdot\frac{\left(x_1^3 + x_2^3 + \cdots + x_{n+1}^3\right)}{n} - 3\mu_{n+1}\sigma_{n+1}^2 - \mu_{n+1}^3}{\sigma_n^3}\cdot\frac{\sigma_n^3}{\sigma_{n+1}^3}$$
$$= \left(\frac{n}{n+1}\cdot\frac{\left(x_1^3 + x_2^3 + \cdots + x_n^3\right)/n}{\sigma_n^3} + \frac{x_{n+1}^3}{(n+1)\sigma_n^3} - \frac{3\mu_{n+1}\sigma_{n+1}^2}{\sigma_n^3} - \frac{\mu_{n+1}^3}{\sigma_n^3}\right)\cdot\frac{\sigma_n^3}{\sigma_{n+1}^3}$$
$$= \left(\begin{array}{l}\frac{n}{n+1}\cdot\dfrac{\frac{\left(x_1^3 + x_2^3 + \cdots + x_n^3\right)}{n} - 3\mu_n\sigma_n^2 - \mu_n^3 + 3\mu_n\sigma_n^2 + \mu_n^3}{\sigma_n^3} \\ \qquad + \dfrac{x_{n+1}^3}{(n+1)\sigma_n^3} - \dfrac{3\mu_{n+1}\sigma_{n+1}^2}{\sigma_n^3} - \dfrac{\mu_{n+1}^3}{\sigma_n^3}\end{array}\right)\cdot\frac{\sigma_n^3}{\sigma_{n+1}^3}$$
$$= \left(\frac{n}{n+1}\cdot\frac{\gamma_n + 3\mu_n\sigma_n^2 + \mu_n^3}{\sigma_n^3} + \frac{x_{n+1}^3}{(n+1)\sigma_n^3} - \frac{3\mu_{n+1}\sigma_{n+1}^2}{\sigma_n^3} - \frac{\mu_{n+1}^3}{\sigma_n^3}\right)\cdot\frac{\sigma_n^3}{\sigma_{n+1}^3}$$
$$= \frac{n}{(n+1)\sigma_{n+1}^3}\left(\gamma_n\sigma_n^3 + 3\mu_n\sigma_n^2 + \mu_n^3 + x_{n+1}^3\right) - \frac{3\mu_{n+1}\sigma_{n+1}^2}{\sigma_{n+1}^3} - \frac{\mu_{n+1}^3}{\sigma_{n+1}^3}$$

$$\begin{aligned}\Rightarrow \gamma_{n+1} &= \frac{n}{(n+1)\sigma_{n+1}^3}\left(\gamma_n\sigma_n^3 + 3\mu_n\sigma_n^2 + \mu_n^3 + x_{n+1}^3\right) - \frac{3\mu_{n+1}}{\sigma_{n+1}} - \frac{\mu_{n+1}^3}{\sigma_{n+1}^3} \\ &= \frac{n}{(n+1)\sigma_{n+1}^3}\left(\gamma_n\sigma_n^3 + 3\mu_n\sigma_n^2 + \mu_n^3 + x_{n+1}^3\right) - \frac{\mu_{n+1}}{\sigma_{n+1}^3}\left(3\sigma_{n+1}^2 + \mu_{n+1}^2\right)\end{aligned} \tag{5.61}$$

The above formula indicates the online skewness calculation involves few set of algebraic operations with the already evaluated lower moments in the present and the past time instant and the old skewness.

The kurtosis (β) of the time series is given by

$$\begin{aligned}\beta &= E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] \\ &= \frac{E\left[X^4\right] - 4E\left[X^3\right]\mu + 6E\left[X^2\right]\mu^2 - 4E[X]\mu^3 + \mu^4}{\sigma^4} \\ &= \frac{E\left[X^4\right] - 4E\left[X^3\right]\mu + 6E\left[X^2\right]\mu^2 - 3\mu^4}{\sigma^4} \\ &= \frac{E\left[X^4\right]}{\sigma^4} - \frac{4\mu E\left[X^3\right]}{\sigma^4} + \frac{6\mu^2 E\left[X^2\right]}{\sigma^4} - \frac{3\mu^4}{\sigma^4} \\ &= \frac{E\left[X^4\right]}{\sigma^4} - \frac{4\mu}{\sigma}\left(\frac{E\left[X^3\right] - 3\mu\sigma^2 - \mu^3 + 3\mu\sigma^2 + \mu^3}{\sigma^3}\right) \\ &\quad + \frac{6\mu^2}{\sigma^4}\left(E\left[X^2\right] - \mu^2 + \mu^2\right) - \frac{3\mu^4}{\sigma^4} \\ &= \frac{E\left[X^4\right]}{\sigma^4} - \frac{4\mu}{\sigma}\left(\frac{\gamma + 3\mu\sigma^2 + \mu^3}{\sigma^3}\right) + \frac{6\mu^2}{\sigma^4}\left(\sigma^2 + \mu^2\right) - \frac{3\mu^4}{\sigma^4} \\ &= \frac{E\left[X^4\right]}{\sigma^4} - \frac{4\mu\gamma}{\sigma^4} - \frac{12\mu^2\sigma^2}{\sigma^4} - \frac{4\mu^4}{\sigma^4} + \frac{6\mu^2\sigma^2}{\sigma^4} + \frac{6\mu^4}{\sigma^4} - \frac{3\mu^4}{\sigma^4} \\ &= \frac{E\left[X^4\right] - 4\mu\gamma - 6\mu^2\sigma^2 - \mu^4}{\sigma^4} \\ \Rightarrow \beta_n &= \frac{\dfrac{1}{n}\left(x_1^4 + x_2^4 + \cdots + x_n^4\right) - 4\mu_n\gamma_n - 6\mu_n^2\sigma_n^2 - \mu_n^4}{\sigma_n^4}\end{aligned} \tag{5.62}$$

Therefore, the updated skewness considering the next sample becomes

$$\beta_{n+1} = \frac{\frac{1}{n+1}\left(x_1^4 + x_2^4 + \cdots + x_n^4 + x_{n+1}^4\right) - 4\mu_{n+1}\gamma_{n+1} - 6\mu_{n+1}^2\sigma_{n+1}^2 - \mu_{n+1}^4}{\sigma_{n+1}^4}$$

$$= \frac{n}{(n+1)\sigma_{n+1}^4}\left(\frac{x_1^4 + x_2^4 + \cdots + x_n^4}{n} + \frac{x_{n+1}^4}{n}\right) - \frac{4\mu_{n+1}\gamma_{n+1}}{\sigma_{n+1}^4} - \frac{6\mu_{n+1}^2\sigma_{n+1}^2}{\sigma_{n+1}^4} - \frac{\mu_{n+1}^4}{\sigma_{n+1}^4}$$

$$= \frac{n\sigma_n^4}{(n+1)\sigma_{n+1}^4}\left(\begin{array}{l}\frac{x_1^4 + x_2^4 + \cdots + x_n^4}{n} - 4\mu_n\gamma_n - 6\mu_n^2\sigma_n^2 \\ \frac{-\mu_n^4 + 4\mu_n\gamma_n + 6\mu_n^2\sigma_n^2 + \mu_n^4}{\sigma_n^4}\end{array}\right) + \frac{x_{n+1}^4}{(n+1)\sigma_{n+1}^4} - \frac{4\mu_{n+1}\gamma_{n+1}}{\sigma_{n+1}^4} - \frac{6\mu_{n+1}^2\sigma_{n+1}^2}{\sigma_{n+1}^4} - \frac{\mu_{n+1}^4}{\sigma_{n+1}^4}$$

$$= \frac{n\sigma_n^4}{(n+1)\sigma_{n+1}^4}\left(\beta_n + \frac{4\mu_n\gamma_n}{\sigma_n^4} + \frac{6\mu_n^2\sigma_n^2}{\sigma_n^4} + \frac{\mu_n^4}{\sigma_n^4}\right) + \frac{x_{n+1}^4}{(n+1)\sigma_{n+1}^4} - \frac{4\mu_{n+1}\gamma_{n+1}}{\sigma_{n+1}^4} - \frac{6\mu_{n+1}^2\sigma_{n+1}^2}{\sigma_{n+1}^4} - \frac{\mu_{n+1}^4}{\sigma_{n+1}^4}$$

$$\Rightarrow \beta_{n+1} = \frac{n}{(n+1)\sigma_{n+1}^4}\left(\beta_n\sigma_n^4 + 4\mu_n\gamma_n + 6\mu_n^2\sigma_n^2 + \mu_n^4\right) + \frac{x_{n+1}^4}{(n+1)\sigma_{n+1}^4} - \frac{1}{\sigma_{n+1}^4}\left(4\mu_{n+1}\gamma_{n+1} + 6\mu_{n+1}^2\sigma_{n+1}^2 + \mu_{n+1}^4\right)$$

$$= \frac{1}{\sigma_{n+1}^4}\left[\begin{array}{l}\left(\frac{1}{(n+1)}\left(n\beta_n\sigma_n^4 + 4n\mu_n\gamma_n + 6n\mu_n^2\sigma_n^2 + n\mu_n^4 + x_{n+1}^4\right)\right) \\ -\left(4\mu_{n+1}\gamma_{n+1} + 6\mu_{n+1}^2\sigma_{n+1}^2 + \mu_{n+1}^4\right)\end{array}\right] \tag{5.63}$$

5.8.3 *Open Research Problems in Low Complexity Machine Learning Algorithms*

Application of machine learning techniques in a resource constrained environment for biomedical applications comes from the basic implementation issues of wireless sensor network for health applications (Yang 2006). Advancement of low power computation on the sensor nodes is being an active research topic though most of

the effort has been focused on the communication side rather than implementation of an automatic decision making processes on the sensors using machine learning techniques which can be considered as the first step for biomedical signals sensing and diagnosing patient's health condition. In the last few decades several efforts has been made to achieve high speed low complexity hardware design for basic signal processing tools involving different FIR/IIR filtering and transform techniques (Maharatna et al. 2004; Maharatna et al. 2001; Maharatna and Banerjee 2001; Parhi 2007). But most of the advanced signal processing algorithms e.g. multichannel statistical feature extraction as discussed in Sect. 5.3.1, time-frequency representation like spectrogram or scalogram, signal decomposition methods like ICA, PCA, EMD etc. are still an open problem. Also probabilistic techniques like HMM, Bayesian inferencing are mostly confined as a theoretical technique which can be solved using a computer very easily but equivalent low power integrated circuit design techniques are yet to come for application in the field of biomedical informatics. It is quite easy for theoreticians and computer scientists to calculate prior, conditional, joint probabilities and formulate recursive algorithms based on the concept of likelihood function, posterior density etc. But for its applicability in low power application for remote healthcare the engineering part involves efficient way of dividing the features in several bins in order to extract essential properties of the feature histograms and sometime in several dimensions. Also iterative techniques in machine learning like expectation maximization, hard and soft clustering, use of computational intelligence techniques like neural networks, fuzzy logic and rough set theory, gradient based optimization, multiple agent based evolutionary and swarm optimization algorithms which made the foundation of biomedical signal and image analysis and machine learning needs intense research focus in order to spread the algorithms in a mobile and resource constrained environment. We believe the brief overview of different machine learning tools, presented in this chapter will pave the way of future hardware designers to precisely focus on the real electronic circuit implementation issues for cutting edge techniques for complex, iterative, recursive, multi-dimensional algorithms involving optimization routines rather than simple algebraic operations. Also, significant advancement of medical image processing and automated analyses where most of the machine learning tools are now a days used offline, can be implemented in future in resource constrained environment for a better remote healthcare system for online patients' condition monitoring.

5.9 Conclusion

Various machine learning techniques mainly in the field of biomedical pattern classification and clustering and associated feature extraction principles are presented in a brief way in this chapter. Practical issues of implementing established machine learning techniques are highlighted from the perspective of low-power hardware design in biomedical pattern recognition related applications. Overcoming hardware implementation issues of established machine learning

algorithms with a specific focus of biomedical-signal and image analysis need intense research focus in future and are still open in the contemporary research community.

References

Aggarwal N, Agrawal R (2012) First and second order statistics features for classification of magnetic resonance brain images. J Signal Inf Process 3:146–153

Bießmann F, Plis S, Meinecke FC et al (2011) Analysis of multimodal neuroimaging data. IEEE Rev Biomed Eng 4:26–58

Blinowska KJ, Zygierewicz J (2012) Practical biomedical signal analysis using MATLAB. CRC Press, Boca Raton

Cerutti S, Marchesi C (2011) Advanced methods of biomedical signal processing. Wiley-IEEE Press, Piscataway

Chandaka S, Chatterjee A, Munshi S (2009) Cross-correlation aided support vector machine classifier for classification of EEG signals. Expert Syst Appl 36:1329–1336

Chen T, Mazomenos EB, Maharatna K et al (2013) Design of a low-power on-body ECG classifier for remote cardiovascular monitoring systems. IEEE J Emerg Sel Top Circuits Syst 3:75–85

Das S, Abraham A, Konar A (2009) Metaheuristic clustering. Springer, Germany

Duda RO, Hart PE, Stork DG (2012) Pattern classification. Wiley-Interscience, New York

Gacek A (2012) ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence. Springer, Germany

Gonzalez FA, Romero E (2010) Biomedical image analysis and machine learning technologies: applications and techniques. Medical Information Science Reference, Hershey

Günes S, Polat K, Yosunkaya S (2010) Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. Expert Syst Appl 37:7922–7928

Gunn SR (1998) Support vector machines for classification and regression. Technical Report, School of Electronics and Computer Science, University of Southampton

Gurcan MN, Boucheron LE, Can A et al (2009) Histopathological image analysis: a review. IEEE Rev Biomed Eng 2:147–171

Herman P, Prasad G, McGinnity TM, Coyle D (2008) Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. IEEE Trans Neural Syst Rehabil Eng 16:317–326

Ji Z, Sun Q, Xia Y et al (2012) Generalized rough fuzzy c-means algorithm for brain MR image segmentation. Comput Methods Programs Biomed 108:644–655

Jovic A, Bogunovic N (2011) Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. Artif Intell Med 51:175–186

Kononenko I (2001) Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med 23:89–109

Kugiumtzis D, Tsimpiris A (2010) Measures of analysis of time series (MATS): a MATLAB toolkit for computation of multiple measures on time series data bases. Journal of Statistical Software 33(5):1–30

Lima CA, Coelho AL (2011) Kernel machines for epilepsy diagnosis via EEG signal classification: a comparative study. Artif Intell Med 53:83–95

Lindsen JP, Bhattacharya J (2010) Correction of blink artifacts using independent component analysis and empirical mode decomposition. Psychophysiology 47:955–960

Lucas M-F, Gaufriau A, Pascual S et al (2008) Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. Biomed Signal Process Control 3:169–174

Maglaveras N, Stamkopoulos T, Diamantaras K et al (1998) ECG pattern recognition and classification using non-linear transformations and neural networks: a review. Int J Med Inform 52:191–208

Magoulas GD, Prentza A (2001) Machine learning in medical applications. Machine learning and its applications, Lecture Notes in Computer Science 2049:300–307

Maharatna K, Banerjee S (2001) A VLSI array architecture for Hough transform. Pattern Recognit 34:1503–1512

Maharatna K, Dhar AS, Banerjee S (2001) A VLSI array architecture for realization of DFT, DHT, DCT and DST. Signal Processing 81:1813–1822

Maharatna K, Grass E, Jagdhold U (2004) A 64-point Fourier transform chip for high-speed wireless LAN application using OFDM. IEEE J Solid-State Circuits 39:484–493

Mani S, Shankle WR, Dick MB, Pazzani MJ (1999) Two-stage machine learning model for guideline development. Artif Intell Med 16:51–71

Mao J, Jain AK (1996) A self-organizing network for hyperellipsoidal clustering (HEC). IEEE Trans Neural Netw 7:16–29

Mazomenos E, Chen T, Acharyya A et al (2012) A time-domain morphology and gradient based algorithm for ECG feature extraction. In: 2012 I.E. international conference on Industrial technology (ICIT), Athens, pp 117–122

Mazomenos EB, Biswas D, Acharyya A et al (2013) A low-complexity ECG feature extraction algorithm for mobile healthcare applications. IEEE J Biomed Health Inform 17:459–469

Mijovic B, De Vos M, Gligorijevic I et al (2010) Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. IEEE Trans Biomed Eng 57:2188–2196

Monte-Moreno E (2011) Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques. Artif Intell Med 53:127–138

Mormann F, Lehnertz K, David P, Elger CE (2000) Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. Physica D 144:358–369

Nanni L, Lumini A, Brahnam S (2012) A classifier ensemble approach for the missing feature problem. Artif Intell Med 55:37–50

Nolte G, Bai O, Wheaton L et al (2004) Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin Neurophysiol 115:2292–2307

Parhi KK (2007) VLSI digital signal processing systems: design and implementation. Wiley, New York

Rogers S, Girolami M (2011) A first course in machine learning. Chapman & Hall, London

Sajda P, Philiastides MG, Parra LC (2009) Single-trial analysis of neuroimaging data: inferring neural networks underlying perceptual decision-making in the human brain. IEEE Rev Biomed Eng 2:97–109

Semmlow JL (2004) Biosignal and biomedical image processing: MATLAB-based applications. CRC press, Boca Raton

Sierra B, Serrano N, Larrañaga P et al (2001) Using Bayesian networks in the construction of a bi-level multi-classifier. A case study using intensive care unit patients data. Artif Intell Med 22:233–248

Sörnmo L, Laguna P (2005) Bioelectrical signal processing in cardiac and neurological applications. Academic, Burlington

Strauss DG, Selvester RH, Lima JA et al (2008) ECG quantification of myocardial scar in cardiomyopathy patients with or without conduction defects: correlation with cardiac magnetic resonance and arrhythmogenesis. Circ Arrhythm Electrophysiol 1:327–336

Subasi A, Ismail Gursoy M (2010) EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst Appl 37:8659–8666

Theodoridis S, Koutroumbas K (2010) Introduction to pattern recognition: a matlab approach. Academic, Amsterdam

Velikova M, Lucas PJ, Samulski M, Karssemeijer N (2013) On the interplay of machine learning and background knowledge in image interpretation by Bayesian networks. Artif Intell Med 57:73–86

Wahlberg P, Lantz G (2000) Methods for robust clustering of epileptic EEG spikes. IEEE Trans Biomed Eng 47:857–868
Xu R, Wunsch DC (2010) Clustering algorithms in biomedical research: a review. IEEE Rev Biomed Eng 3:120–154
Yang G-Z (2006) Body sensor networks. Springer, Berlin

Chapter 6
Patient Health Record (PHR) System

Artur Krukowski, Carlos Cavero Barca, Emmanouela Vogiatzaki, and Juan Mario Rodríguez

Abstract This chapter focusses on the introduction of the Electronic (EHR) and Personal Health Records (PHR) as new technological approaches aimed at standardising electronic management of medical information between the patient and its physicians, as well as among medical organisations collaborating in providing integrated medical care services. We start with the introduction of the roles and aims behind EHR, followed by the description of the applicable legal and standardisations frameworks and European activities in this area, finishing with the introduction to most common commercial as well as open-source implementations of such systems. As a conclusion we also signal ethical and privacy concerns that are relevant to using and exchanging electronic health information.

6.1 Role and Structure of PHR

The Electronic Health Record (EHR) of a patient can be defined as digitally stored health care information about individual's lifetime with the purpose of supporting continuity of care, education and research, and ensuring confidentiality at all times. A patient's healthcare information may be spread out over a number of different institutes that do not interoperate. In order to provide continuity of care, clinicians should be able to capture the complete clinical history of the patient.

A. Krukowski (✉)
Intracom Telecom, Athens, Greece
e-mail: krukowa@intracom.com

C.C. Barca • J.M. Rodríguez
ATOS Research and Innovation (ARI), ATOS, Madrid, Spain
e-mail: carlos.cavero@atos.net; juan.rodriguezp@atos.net

E. Vogiatzaki
RFSAT Ltd., Sheffield, UK
e-mail: emmanouela@rfsat.com

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_6,

The Personal Health Record (PHR) is the electronic part of the health-related information of a person (such as diagnoses, medications, allergies, lab test results, immunization records but also administrative tasks such as appointment or prescription renewals) that can be extracted from multiple sources, but always under the control of the consumer, patient or informal caregiver. This is the relevant difference between the PHR and the EHR or electronic medical record, which is maintained by the healthcare providers and payers (Heubusch 2008).

The EHR standardisation aims to ensure that patient records are used to support shared care among clinicians with different specialisations, while enabling the mobility within and among countries for people who give and receive healthcare.

From the viewpoint of standardization, the single most important characteristic of the EHR is the ability to share EHR information between different authorized users. In technical terms, this requires interoperability of information in the EHR and interoperability of EHR systems that exchange and share this information. We distinguish two major levels of interoperability (info sharing) of information (ISO/TC 215 2003):

- Functional interoperability: it is the ability of two or more systems to exchange information (so that it is human readable by the receiver); and
- Semantic interoperability: it is the ability for information shared by systems to be understood at the level of formally defined domain concepts (so that information is digitised by the receiving system). Semantic interoperability is not an all-or-nothing concept. The degree of semantic interoperability depends on the level of agreement on terminology and on the content of archetypes and templates used by the sender and receiver of information.

One of the key requirements for interoperability of the EHR is to break the nexus between the EHR and the EHR system (i.e. the EHR should conform to an information model independent of both the physical database schema used for local storage and the applications, which create, maintain, and retrieve EHR). This EHR information model should be independent of any particular implementation technology (i.e. it should be a logical information model). Technology independence is essential to make EHR 'future proof' to enable a lifetime EHR possibility.

In order to achieve semantic interoperability of EHR information, there are three prerequisites, with the first ones being required for functional interoperability (HL7 2004):

1. A standardized EHR reference model, i.e. the EHR information architecture, between the sender (or sharer) and receiver of the information;
2. Standardized service interface models to provide interoperability between the EHR service and other services such as demographics, terminology, access control and security services in a comprehensive clinical information system;
3. A standardized set of domain-specific concept models, i.e. archetypes and templates for clinical, demographic, and other domain-specific concepts; and

Standardized terminologies, which underpin the archetypes. This does not mean that there is need to have a single standardized terminology for each health domain but rather, terminologies used should be associated with controlled vocabularies.

6.2 EHR Standardization Bodies

One of the main factors hindering the widespread adoption of integrated PHRs is the lack of technical standards for interoperability, which is the ability of systems to exchange information using the same mechanisms. "The immaturity and slow diffusion of standards for interoperability and data portability are key barriers to the integration and exchange of structured data among PHRs and the range of relevant entities that provide and finance health care" (Detmer et al. 2008). The success and final adoption of the PHR systems depends on the capability of interacting with Electronic Health Records (EHRs) and other sources of personal health data (National Committee on Vital and Health Statistics 2006). Currently after years of pursuing the interoperability, the EHR standards continue lacking of the public adoption and immaturity mentioned.

A number of standardization efforts are progressing to provide the interoperability of EHRs such as CEN/TC 251 EHRcom (http://www.centc251.org), openEHR (http://www.openehr.org) and HL7 (http://xml.coverpages.org/CDA-20040830v3.pdf). These standards aim to structure and mark-up the clinical content for the purpose of exchange. A complementary initiative addressing the issue of how to exchange EHR complying with different content standards is the Integrating the Healthcare Enterprise (IHE) (http://www.ihe.net) Cross-Enterprise Document Sharing (XDS) integration profile (http://www.ihe.net/Technical_Framework).

6.2.1 ISO/TC215

CEN/TC 251 is the technical committee on Health Informatics of the European Committee for Standardization. Its mission is to achieve compatibility and interoperability between independent health systems and to enable modularity by means of standardization. This includes requirements on health information structure to support clinical and administrative procedures, technical methods to support interoperable systems as well as requirements regarding safety, security and quality (Eichelberg et al. 2005a). The CEN pre-standard ENV 13606:2000 "Electronic Healthcare Record Communication" is a message-based standard for the exchange of EHR content. This standard defines an EHR information model, called the "extended architecture" since it is an extension of the earlier pre-standard ENV 12265. It also defines a list of machine-readable domain terms that can be used to structure EHR content, a method of specifying "distribution rules", that is, rules under which certain EHR content may be shared with other systems and, finally, request and response messages that allow systems to exchange subsets of an EHR. ENV 13606 does not attempt to specify a complete EHR system; instead, it focuses on the interfaces relevant for a communication between EHR systems.

ENV 13606 was intended to be the first fully implementable EHR standard, and subsets of it were implemented in a number of EHR projects in the UK, Denmark, the Netherlands, Sweden, and Norway. However, none of these projects used the complete ENV 13606 specification; moreover, the implementation experience

showed a number of weaknesses in the standard that limited its usefulness and market uptake: the single-level modelling approach made the information model extremely complex, with lot of optionality and a level of abstraction that made quite difficult to comprehend and implement the model.

In 2001, CEN/TC 251 decided to revise ENV 13606 into a full European Standard, taking into account the existing implementation experience and to adopt the openEHR archetype methodology. ENV 13606 is a standard that is now gradually being approved, and consists of five parts:

- Reference Model: it defines the hierarchy of generic building blocks of the EHR through a set of classes. It represents the stable characteristics of the EHR entries, how they are aggregated, and the context information required to meet ethical, legal and provenance requirements;
- Archetype Interchange Specification: each archetype defines legal combinations of the building block classes defined in the Reference Model for particular clinical domains, organizations. The archetype model is syntactically equivalent to those of the Good Electronic Health Record project, and the openEHR standard;
- Reference Archetypes and Term Lists: it includes the vocabularies for attributes, and archetypes to represent HL7 specialized Acts and openEHR specialised ENTRYs;
- Security Features: it defines an interoperable specification for EHR disclosure consent, and an interoperable disclosure log;
- Exchange Models: this part is still under discussion.

ENV 13606 standard was not intended to specify the internal architecture or database design of EHR systems or components. Nor is it intended to prescribe the kinds of clinical applications that might request or contribute EHR data in particular settings, domains or specialties. For this reason, the information model proposed there was called the *EHR Extract*, and might be used to define a message, an XML document or schema, or an object interface. The information model in this European Standard is an ISO RM-ODP Information Viewpoint of the EHR Extract. This European Standard considers the EHR to be the persistent longitudinal and potentially multi-enterprise or multi-national record of health and care provision relating to a single subject of care (the patient), created and stored in one or more physical systems in order to inform the subject's future health care and to provide a medico-legal record of care that has been provided.

6.2.2 ISO/EN EN13606

ISO/EN EN13606 is a norm designed to achieve semantic interoperability in EHR-related data communication among different Health Information Systems (HIS) (CEN/TC251–ISO/TC215 I. E 2010a). Its main goal is to define a stable and reliable information structure in order to communicate EHR parts of the same

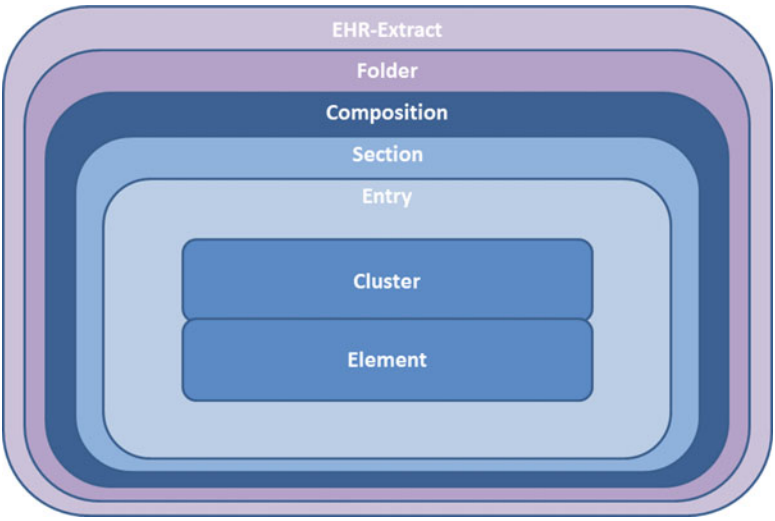


Fig. 6.1 Components of the ISO/EN 13606 (EN13606 Association 2009)

patient. The first version of the 13606 four-part pre-standard was published in 1999–2000 but attempts to implement this pre-standard in software proved to be difficult and those implementations which were undertaken suffered from the "HL7 v2 problem" of too much optionality. In 2002 CEN made a decision to revise the 13606 pre-standard and upgrade it to a full normative European standard (EN 13606, also called EHRcom). The ISO/EN13606 standard were completed and ratified after Part 5 by ISO and CEN in February 2010. ISO/EN 13606 architecture provides a framework to communicate EHR data using the dual model approach (reference model and archetypes) to provide the semantic interoperability. The ISO/EN 13606 consists of five parts:

1. **Part 1**: CEN 2007: ISO 2008: The Reference Model, the generic common information model. The global characteristics of health record components.
2. **Part 2**: CEN 2007: ISO 2008: Archetype Interchange Specification, information model of the metadata to represent the domain-specific characteristics of electronic health record entries. This chapter defines how to share archetypes, and not how to exchange them within particular systems.
3. **Part 3**: CEN 2008: ISO 2009: Reference Archetypes and Term Lists, establishing the normative terminologies and controlled vocabularies.
4. **Part 4**: CEN 2007: ISO 2009: Security Features, covering security mechanisms and methodology.
5. **Part 5**: CEN/ISO 2010: Exchange Models, interface designed to request specific extracts, archetypes or audit log.

The relevant components of the generic reference model are (Fig. 6.1): an "EHR_Extract" is the root node of the EHR and contains "Compositions" which

can be organised using "Folders" (in the same way as Microsoft Windows explorer folders). The "Entry" can be an observation, medication order, diagnoses and can be organized within "Sections". The leaf nodes (data) are "Elements" which are included in the "Entry" and optionally organized within "Clusters".

The ISO/EN 13606 is a subset of the full openEHR specification (Schloeffel et al. 2006). Within the shared classes the main difference between with the openEHR reference model is that "Entries" are broken down in the corresponding kind of information stored (Muñoz et al. 2011).

6.2.3 GEHR/OpenEHR

The GEHR/openEHR initiative was started in 1992 as an EU research project, called "Good European Health Record", in the third FP. The initiative was later continued under the name "Good Electronic Health Record" with strong participation from Australia. Currently it is maintained by the openEHR Foundation, a non-profit organization defining itself as "*an international*, *on-line community whose aim is to promote and facilitate progress towards EHRs of high quality*, *to support the needs of patients and clinicians everywhere*".

The openEHR is a foundation that supports the development of an open and semantic-connected platform for eHealth systems (OpenEHR Foundation 2012). It is based on 15 years research, focused engineering design and real-world implementation experience, rather than being created as a formal consensus standard. However, over the last years it has had a significant influence over the development of EHR standards by the three main international eHealth standards organisations: CEN, HL7 and ISO.

The information model covers the EHR architecture and describes classes such as Folder, Composition, Section and Entry, and of the basic data structure and types. The Care Entry class "define the semantics of all the 'hard' information in the record" (Beale et al. 2007), the Admin Entry represents information recorded during administrative issues. Figure 6.2 shows graphically the ontology leading to the Entry model.

The most noteworthy concept introduced by GEHR/openEHR is the "archetype" concept. This approach uses a two-level methodology to model the EHR structure. In the first level, a generic reference model that is specific to the healthcare domain but still very general is developed. This model typically contains only a few classes (e.g. role, act, entity, participation) and must be stable over time. In the second level, healthcare and application specific concepts such as blood pressure, lab results etc. are modelled as archetypes, that is, constraint rules that specialize the generic data structures that can be implemented using the reference model. As an example, a constraint may restrict a generic "Observation" class to, e.g., "Blood Pressure" archetype.

An archetype definition consists of three parts: descriptive data, constraint rules and ontological definitions. The descriptive data contains a unique identifier for the

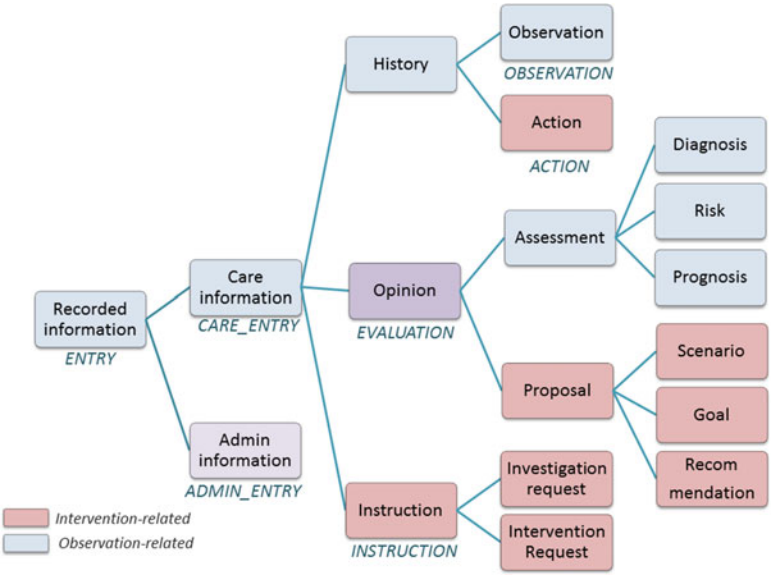


Fig. 6.2 The OpenEHR ontology of recorded information (Beale and Heard 2008)

archetype, a machine-readable code describing the clinical concept modelled by the archetype and various metadata such as author, version, and purpose. It also states whether an archetype is a specialization of another archetype. The constraint rules are the core of the archetype and define restrictions on the valid structure, cardinality and content of EHR record component instances complying with the archetype. The ontological part defines the controlled vocabulary (that is, machine-readable codes) that may be used in specific places in instances of the archetype. It may contain language translations of code meanings and bindings from the local code values used within the archetype to external vocabularies such as SNOMED or LOINC. It may also define additional constraints on the relationship between coded entries in the archetype based on the code value.

As mentioned above the "Care Entry" concept covers the most common and medically relevant information. In the openEHR four types of entries can be distinguished: observations, evaluations, instructions, and actions. The "observation" and "action" classes represent statements about the past events of the individual subject of record. The "evaluation" classes represent current assessment by the attending health professional, including "diagnosis" and "prognosis", as well as the representation of the imagined future, like "goals" and "scenarios". "Instructions" represents future events that should take place as prescribed by the health professional.

The openEHR framework includes a reference information model, the Archetype Definition Language (ADL) for expressing archetypes, an archetype library, implementation technology specifications (XML schemas, IDL specifications etc.) and a collection of open source implementations of the openEHR specifications.

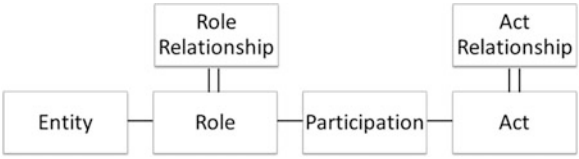


Fig. 6.3 RIM structure

6.2.4 The HL7 Family of Standards

Hl7 is an international organization founded in 1987 and supported by ANSI with the goal of develops global standards related with eHealth. This organization has already defined a set of standards for clinical information interchange, whose name is HL7 standards. Among them HL7 CDA (Clinical Document Architecture) defines the Architecture of electronic documents used within Health domain and it is HL7's current main strategy for EHR interoperability (hl7 2007). Besides, HL7 supports RIM (Reference Information Model), a model of healthcare information as viewed within the scope of HL7 standards, which provides a static view of information needs along with use case models, interaction models, data type models, terminology models, and other types of models to provide a complete view of the requirements and design of HL7 standards, thus giving a valid starting point for any HL7-compliant Architecture Design (Ferranti et al. 2006).

A set of meta-classes can be identified in the RIM (Dolin et al. 2006), as observed in Fig. 6.3:

- **Act**: actions in the healthcare management
- **Participation**: context for an act: Who? For whom? Where?
- **Entity**: physical things, subjects or targets taking part in healthcare act.
- **Role**: establishes roles that entities play in its participation in healthcare acts.
- **Act Relationship**: represents a relationship between two acts.
- **Role Link**: represents a dependency between roles.

Finally, the hl7 v3 Template is "an expression of a set of constraints on the RIM which is used to apply additional constraints to a portion of an instance of data which is expressed in terms of some other Static Model. Templates are used to further define and refine these existing models within a narrower and focused scope" (HL7 2007).

6.2.5 Relationship Between Standards

The openEHR supports the creation, storage, maintenance, and querying of complete EHRs. ISO/EN 13606 is a subset of the full openEHR implementation and it is

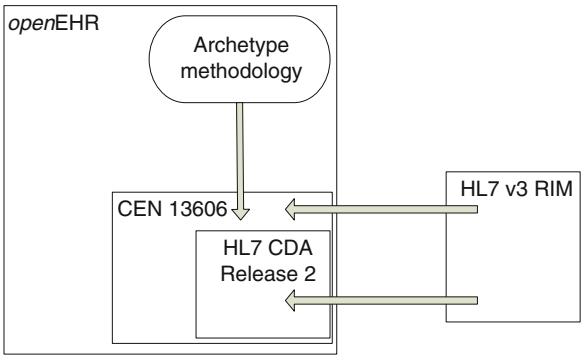


Fig. 6.4 Relationships between standards (openEHR, EN13606 and hL7 CDA)

an appropriate standard for exchange of EHR extracts. At the same time, ISO/EN 13606 offers a partial alignment with HL7 Clinical Document Architecture (CDA) (Fig. 6.4).

Release 2.0, which complies with hl7 RIM. e.g., hl7 CDA is based on the hl7 RIM, but it was designed to represent patient summaries, not thinking on providing decision support capabilities. All the aforementioned reference models are archetype-based. hl7v2.x messaging is an appropriate standard, at least in short/medium term, for transmission of information from clinical information systems to EHR systems.

6.3 Data Structuring Algorithms in PHR

According to European Committee for Standardisation, the EHR is "*the persistent longitudinal and potentially multi-enterprise or multinational record of health and care provision relating to a single subject of care* (*the patient*), *created and stored in one or more physical systems in order to inform the subject's future health care and to provide a medico-legal record of care that has been provided*" (CEN/TC251–ISO/TC215 I. E 2010b).

Patient data is managed, in current Hospital Information Systems (HIS), as a digital and well-structured record, which contains all individual-related health data, such as, demographic, diseases, allergies, history and activity during illness periods, etc.

The semantic interoperability of health data can be achieved only with the standardization of EHR. During the last 10 years, the international organizations have been driving great efforts to define the architecture for exchanging properly the health information between EHR coming from different systems (Muñoz et al. 2011).

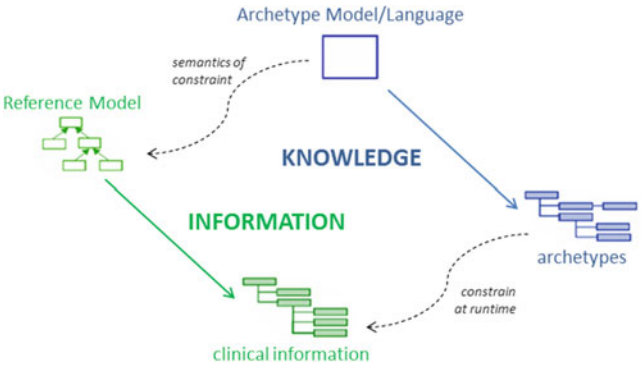


Fig. 6.5 Dual model approach

6.3.1 The Dual Model Approach

The huge amount of clinical concepts and volatility of the information are the main drawbacks to deploy long-term EHR systems due to the single model approach. The ad-hoc solutions are implemented by technical stuff gathering the user requirements from the clinicians providing tools probably perfect at design time but a predictable out-of-date system requiring new releases in the near future if additional information is needed. The "dual model approach", also called the two-level modelling (Rector et al. 1993; Johnson 1996; Beale 2002) separates out the clinical knowledge (volatile) and the reference model (static). The medical concepts are modelled using archetypes based on a stable reference model (Goossen et al. 2010). This approach is attractive for the stake-holders due to its stability, the EHR products are installed once and additional clinical functionalities could be extended using archetypes, on the contrary the ever changing clinical practice jeopardises the return on EHR investment in ad-hoc systems.

The dual model approach defines a generic information model, the reference model with domain-invariant classes to be instantiated as well as specific clinical models, which support semantic interoperability, which are called archetypes, containing specific clinical information designed using a common language as it is shown in Fig. 6.5. This approach has been supported by many working groups within different international initiatives.

6.3.2 Detailed Clinical Models (DCM)

The Detailed Clinical Models are actually "a new way to structure medical information. It combines expert knowledge, data specification and terminology and enables various technical applications". They specify the information models and the way the

data is exchanged: user interfaces, data management, decision support systems and so on. It could be considered equivalent to Archetypes, Templates or Clinical Statements. The Clinical Information Modelling Initiative (CIMI) is working to establish those common models from different standards, initially modelling these DCMs in both ADL (archetype definition language) and UML (www.uml.org).

6.4 Standardisation of User interfaces to PHR

6.4.1 Continuity of Care Record (CCR)

The American Society for Testing and Materials (ASTM International) Continuity of Care Record (CCR) is a clinical framework that was first developed by health care practitioners to meet the information exchange needs of primary care providers (Kibbe et al. 2004). ASTM International (2005) defines the CCR as a "summary of the patient's health status (e.g., problems, medications, allergies) and basic information about insurance, advance directives, care documentation, and care plan recommendations" (ASTM WK4363 2004).

6.4.2 CCD (Continuity of Care Document)

A CCD is a joint effort between HL7 International and ASTM approved as an ANSI standard in 2007 in order to use HL7 CDA for sharing the CCR (Continuity of Care Record) patient summary. It represents a complete implementation of CCR, combining the best of HL7 technologies with the richness of CCRs clinical data representation, and does not disrupt the existing data flows in payer, provider of pharmacy organizations. The Continuity of Care Document (CCD) establishes a detailed set of constraints and templates, covering the main sections of the summary record to be represented as CDA elements (HL7/ASTM 2007; European Commission—DG INFSO/Q-REC 2007).

6.4.3 Integrating the Healthcare Enterprise (IHE)

International is a global association of healthcare IT vendors, user organisations, clinical professional societies, and advocacy groups that promotes interoperability through the co-ordinated use of established standards such as Digital Imaging and Communications in Medicine (DICOM) and HL7. IHE, more than any other single organisation, has paved the way for practical healthcare interoperability (Eichelberg et al. 2005b; Integrating the Healthcare Enterprise (IHE) 2006).

6.5 Terminologies

Coded elements are used in the healthcare environment to precisely define the clinical concepts language-independently. The use of medical terminology is one of the bases to provide semantic interoperability.

- **Systematised Nomenclature of Medicine Clinical Terms** (**SNOMED-CT**) (2009) consists of controlled medical vocabularies (CMVs),—accumulated medical concepts updated in a rigorous fashion. SNOMED-CT has been gaining momentum as the primary coding method for clinical concepts. SNOMED has become the presumptive source of clinical codes and concepts within its member countries.
- **Logical Observation Identifiers Names and Codes** (**LOINC**) (1994) is a database and universal standard for identifying medical laboratory observations. It was developed and is maintained by the Regenstrief Institute, a US non-profit medical research organization, in 1994. LOINC was created in response to the demand for an electronic database for clinical care and management and is publicly available at no cost.
- **International Classification of Diseases** (**ICD**) (2013) is the standard diagnostic tool for epidemiology, health management and clinical purposes. This includes the analysis of the general health situation of population groups. It is used to monitor the incidence and prevalence of diseases and other health problems. It is used to classify diseases and other health problems recorded on many types of health and vital records including death certificates and health records. In addition to enabling the storage and retrieval of diagnostic information for clinical, epidemiological and quality purposes, these records also provide the basis for the compilation of national mortality and morbidity statistics by WHO Member States. It is used for reimbursement and resource allocation decision-making by countries.

The crucial issue regarding the deployment of an EHR product in the medical environment is the ease of mapping to existing local data stores, as well as to national specifications (i.e. as an interface specification, for instance from hl7 v3 to archetypes). In order to provide the semantic interoperability the connection with external health terminology bindings is mandatory to be language independent (McCay et al. 2008). The term binding is supported for manual or semi-automatic creation between archetypes and the concepts in terminology systems (Sundvall et al. 2008; Qamar and Rector 2007) but the main drawbacks still reside in,

- The big amount of clinical concepts and their relationships, which makes necessary to filter out the terms using a powerful and intelligent tool.
- Mapping and translating concepts across vocabularies.

6.6 European R&D Projects Related to EHR Standardization

According to an eHealth ERA report related to eHealth priorities and strategies in European countries (European Commission 2007), the achievement of a European EHR is not yet an overarching goal, but collaboration on developing individual countries' EHR or even basic patient summaries is a first step.

Electronic Health Record is a rather fuzzy term, which has various definitions. A long-term objective of most European countries is a system of regional or nationwide summaries, or sometimes even full (occasionally life-long) document-based or deeply structured records for each citizen. Such a summary or record may be viewed by any of the following:

- Either all the necessary persons concerned;
- Only by those who need access in order to ensure good quality and safe health services; or
- Only by those who have been directly authorised by the patient.

The eventual development of EHR is evident in 25 out of the 32 countries reviewed during the preparation of the above-mentioned report. Six countries report that they currently have widespread local EHR in hospitals and other health provider organisations, which, however, are not yet fully interconnected. Three countries have a national EHR, although they are yet restricted in scope. Luxembourg, for example, maintains radiology records for its citizens; and in Sweden, citizens have a medication record. Germany, Sweden and Turkey are currently developing the structures of a patient summary or minimal data set. Consistent with its regions-based healthcare system, Spain is developing this work on a regional level. Only one country has a fully implemented EHR system of a countrywide scope—the Czech Republic. The Danish MedCom infrastructure supporting the electronic exchange of various healthcare related messages between healthcare and other service providers is expanded towards a countrywide EHR system as well.

Interoperability seems not to be as high on most countries' agendas as one might expect, given that it is one of the key issues in the EU eHealth Action Plan, it is a core element of current discussions among European Member States, and is also vividly present in international discussions. Only about one-third of the countries' fact sheets mention interoperability explicitly. With the exception of Italy, Romania and Spain, which have made technical and semantic interoperability priority issues, interoperability is seen as a challenge that needs to be addressed as part of a larger initiative. In Denmark, for example, MedCom has already developed a platform for technical standards and interoperability for eMessages—the Danish Health Data Network—, and SNOMED CT (Systematised Nomenclature of Medicine Clinical Terms) is currently being translated to provide semantic interoperability. Figure 6.6 summarizes the status in several European countries regarding existence of EHR.

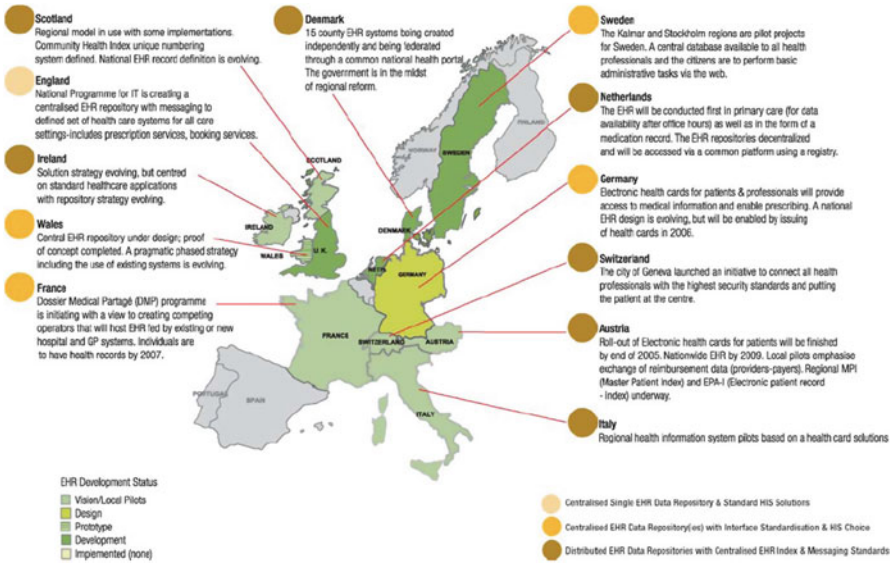


Fig. 6.6 Overview of status of implementation and uptake of hospital information systems and electronic health records throughout Europe (Kongstedt et al.)

6.6.1 European R&D Projects Related to EHR Standardization

The objectives of the EHR-IMPLEMENT project (http://www.ehr-implement.eu) (EU funded project 2007–2010) are to collect, analyse and compare national initiatives of broad scale EHR implementations among European countries focusing on socio-organisational issues and to provide best practice, policy and strategic recommendations to facilitate EHR implementation throughout Europe. The project aimed to:

- Analyse selected national policies, strategies and initiatives for broad scale EHR implementation taking into account cultural and organizational diversities of health systems in six European Member States (Belgium, Denmark, France, Ireland, Slovenia and United Kingdom);
- Carry out a survey of national policy and action plans for broad scale EHR implementation across European Member States;
- Identify the best practices towards broad scale EHR implementations in European countries;
- Raise the awareness of decision and policy makers regarding socio-cultural and organizational issues of broad scale EHR implementation; and
- Support the creation of a multidisciplinary community of scientific experts, technical personnel and National Health System representatives to promote information sharing and mutual learning

6.7 PHR/EHR Implementations

6.7.1 Commercial Implementations

Since the beginning major companies have invested on developing proprietary though freely accessible, own brands of Personal and Electronic Health systems, most well-known ones are listed below.

6.7.1.1 Google Health (http://www.google.com/health)

Google Health was a personal health information centralization service, introduced by Google in 2008 and cancelled in 2011. Google Health was under development from mid-2006. In 2008, the service underwent a 2-month pilot test with 1,600 patients of The Cleveland Clinic. Starting on May 20, 2008, Google Health was released to the public as a service in beta test stage. On 15 of September, 2010 Google updated Google Health with a new look and feel. On 24 of June, 2011 Google announced it was retiring Google Health on 1st of January 2012. The reason Google gave for abandoning the project was a lack of widespread adoption.

The service allowed Google users to volunteer their health records—either manually or by logging into their accounts at partnered health services providers—into the Google Health system, thereby merging potentially separate health records into one centralized Google Health profile.

Volunteered information could include "health conditions, medications, allergies, and lab results". Once entered, Google Health used the information to provide the user with a merged health record, information on conditions, and possible interactions between drugs, conditions, and allergies. Google Health's API was based on a subset of the Continuity of Care Record.

Google Health was an opt-in service, meaning it could only access medical information volunteered by individuals. It did not retrieve any part of a person's medical records without his or her explicit consent and action. However, it did encourage users to set up profiles for other individuals.

6.7.1.2 Microsoft Health Vault (http://www.healthvault.com/)

Microsoft HealthVault is a WEB-based platform from Microsoft to store and maintain health and fitness information. It started in October 2007 in the US. As of 2013, the website addresses both individuals and healthcare professionals in the UK and Germany and the list of national deployments constantly grows.

A HealthVault record stores an individual's health information. Access to a record is through a HealthVault account, which may be authorized to access records for multiple individuals, so that a mother may manage records for each of her children or a son may have access to his father's record to help the father deal with

medical issues. Authorization of the account can be through Windows Live ID, Facebook or a limited set of OpenID providers.

An individual interacts with their HealthVault record through the HealthVault site, or, more typically, through an application that talks to the HealthVault platform. When an individual first uses a HealthVault application, they are asked to authorize the application to access a specific set of data types, and those data types are the only ones the application can use. An individual can also share a part (some data types) or the whole of their health record with another interested individual such as a doctor, a spouse, a parent, etc.

HealthVault Connection Centre allows health and fitness data to be transferred from devices (such as heart rate watches, blood pressure monitors, Withings Wi-Fi body scales etc.) into an individual's HealthVault record. User can find and download drivers for medical devices. A dedicated Device Driver Development Package (http://www.microsoft.com/en-us/download/details.aspx?id=26801) allows also device manufacturers to develop the software support for their devices such that they can communicate with the Health Vault.

HealthVault supports storage of DICOM (http://dicom.nema.org/) based medical imaging. Consumers can upload and download medical imaging DVD through HealthVault connection centre. Third parties can also upload and download medical imaging to/from HealthVault. In addition, there has been plethora of HealthVault medical imaging viewers released by the third party to connect to HealthVault even on mobile phones.

HealthVault supports a number of exchange formats including industry standards such as the Continuity of Care Document (CCD) (http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6), Continuity of Care Record (CCR) and Clinical Document Architecture (CDA) (http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7). Support for industry standards makes it possible to integrate with diverse personal health record solutions.

A list of WEB applications from third-party providers is available at the HealthVault website. Health service providers can develop their own support for MS Health Vaults via an HealthVault .NET Software Development Kit (http://msdn.microsoft.com/en-us/healthvault/jj126730). Examples include:

- InstantPHR released by Get Real Health
- HealthUnity PHR Gateway offered by HealthUnity
- PassportMD developed by PassportMD
- ActivePHR released by ActiveHealth

6.7.1.3 World Medical Card

World Medical Card is a product and registered trademark belonging to World Medical Centre, a Norwegian company headquartered in Bergen, Norway. The company's business is Health information technology, more specifically a supplier of Personal health records.

The World Medical Centre was established in 1998 for creating a system for improving the safety of people in situations requiring immediate medical treatment by a doctor, not familiar with the person's medical history. The international, personal medical card (World Medical Card) system was developed in cooperation with specialists in acute medicine and the University of Bergen, to allow individuals to carry essential medical information with them at all times and everywhere in the world.

Early versions included smart card and data matrix versions, but were abandoned as they required specific infrastructure to be installed at the facility receiving the patient in order to be useful. The latest generation of cards include the information in printed letters, but is only accessible by physically cutting open the card. This also reveals an emergency code that allows the medical professional access to a read-only web profile describing the cardholder.

Later the product portfolio was extended to include a WEB-based Personal health record, allowing the user to manage personal information while also serving as the source of information for producing the card, and a mobile application. Currently it is offered today in a form of three main elements:

1. Online ("onWeb") (http://www.wmc-card.com/index.php?mapping=6) health profile for adding and editing personal health data
2. Multilingual ("onMobile") (http://www.wmc-card.com/index.php?mapping=5) WAP phone application for accessing same data
3. Sealed ("onCard") (http://www.wmc-card.com/index.php?mapping=4) physical card containing compact holder's health data

The company has been a pioneer in promoting the use of ICD-10 (http://www.who.int/classifications/icd/en/) and ATC codes (http://www.who.int/classifications/atcddd/en/) as defined by World Health Organization for processing all medical data.

6.7.2 Implementations from European Research Projects

Many EHR platforms have been developed through European activities, funded by the European Commission through programs such as ICT, ICT-PSP, AAL, ARTEMIS and other ones. The list of most commonly known is provided below.

6.7.2.1 LinkCare Platform (http://www.linkcare-eu.org)

It has been developed by the LinkCare Alliance (www.linkcarealliance.org) as part of the eTEN Project. It aimed to deliver proven information systems for chronic care, linking hospital care, primary care and home care. This concept fills a critical unmet need in current healthcare systems that are challenged by the need of different actors to cooperate in specific scenarios, notably those corresponding to the management of chronic patients.

From the end-user standpoint, LinkCare provides a single vendor integrated solution, a computer supported cooperative work environment, targeting core business areas where costs are greatest and where effective, timely and accurate communication between numerous institutions and actors is critical.

In practical terms, this means that healthcare providers can find in LinkCare technology a supporting tool for those services targeting long-term care. LinkCare facilitates professionals' tasks related to chronic case management, clinical documentation, patient tracking, data analysis, customer relationship management, patient education, professional communication as well as performance evaluation.

The market for information services in chronic care management is truly emerging and LinkCare aims at consolidating a position on it, not only based on the ICT platform but also on the accumulated experience on new models of delivery of care services acquired through the project and through other experiences. This is packed as accompanying consultancy services and/or system integration services to those stakeholders interested in adopting LinkCare in their work practices.

The LinkCare services portfolio includes:

- Electronic case management module and embedded EPR Interface: Capability to support integration and communication with the existing Customer's systems using industry standard protocol (namely, HL7 and XML)
- CRM module—"call-centre" support tools: Core to the LinkCare services is the existence of a single point of access for customers and networked professionals from where the different actions can be decided, ordered and transferred or executed. This requires a call centre supporting advanced CRM (customer relationship management) features. Furthermore, the specificity of the targeted health services means some extra capabilities such as the link to health information resources (corporate HIS, departmental solutions...)
- Professional's mobile support tools: Most of the services that LinkCare will support are based on the mobility of the professionals providing the service. The tools incorporated into LinkCare should fully support these new work practices allowing the professional to minimise the need to contact the institution. Examples of services should include:
  - Enhanced off-line communication features: possibility of sending messages to pagers, SMS messages/emails/voices messages...
  - Enhanced online communication features: MMS, video-clips, video-conference
  - Automatic tracking of pending tasks with professional agenda update. Warning systems to help in the avoidances of delays
- Patient's mobile support tools: Similar to the previous point, the potential for deploying healthcare services to patients will depend on the level of monitoring capabilities that LinkCare tools could offer. Dealing with more severe patients directly links to the availability of more continuous, long-term monitoring, and the possibility of summarizing data, such as:
  - Services in the area of sleep disorders in COPD patients
  - Distant on-line supervision of rehab sessions.

Both examples illustrate the need for:

- Extended periods of data collection (1–8 h)
- Significant amount of data transmitted simultaneously
- Need for tools for summarizing the information

- Computer Supported Cooperative Work module including workflow: Support for modelling the clinical processes at customer's site, providing the necessary communication and interaction tools among the involved actors
- On-line education and reference access to current content, docs and research: Emerging models of care provision maximize the importance of patients and care takers as key partners in the management of health conditions. LinkCare should be a comprehensive and trusted repository/broker for information content that could be passively provided (upon user's will, pull paradigm) or actively suggested (push paradigm, in alignment with the program where the patient is currently treated)
- Performance monitoring and evaluation module ("control panel/dashboard" & reporting tools): a set of decision management tools providing essential information about key indicators of the clinical and business processes, to allow timely intervention for corrective interventions and analytical support to improvement actions

6.7.2.2 intLIFE60 PHR

The intLIFE PHR (intLIFE) is a platform developed internally by Intracom S. A. Telecom Solutions, enhanced and geared to diverse health application through a number of FP7 funded research projects, such as ICT-PSP-NEXES, AAL-PAMAP, FP7-StrokeBack, FP7-ARMOR and others. The intLIFE platform consists of the following subsystems (shown in Fig. 6.7):

- Electronic Health Record Subsystem
- Vital Signs Monitoring Subsystem
- Personal Health Record Subsystem
- intLIFE Management Subsystem

The list features and functionality of the components currently available includes:

- **EHR Subsystem**: the intLIFE EHR application enables clinicians and paramedical personnel to edit/review health related information of the monitored subjects. An Overview tab will provide the user with a quick and printable outline of selected information (e.g. diagnoses, medications, surgeries, specific measurements, etc.). The following information fields are available:
  - Patient's General Health Profile; family health history, habits & social history (e.g. smoking, alcohol consumption), allergies, vaccinations
  - Visits; organ system findings, manual entry of symptoms & measurements

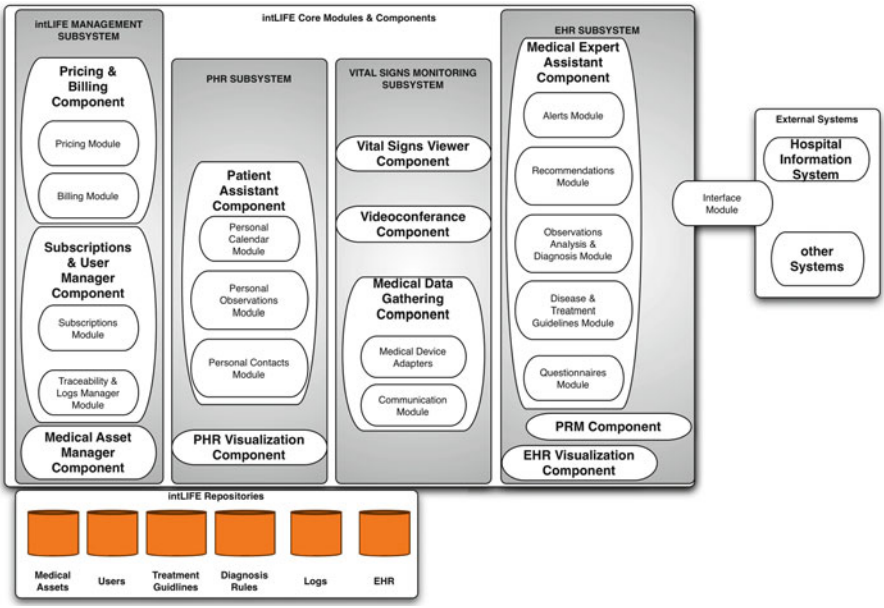


Fig. 6.7 intLIFE core modules & components: electronic health record subsystem

- Medical Tests; test orders, manual entry of test results, test results overview and graphic representation
- Diagnosis Management; insert new diagnoses, using the ICD-10 nomenclature, search for past diagnoses
- Treatment Management; surgeries, medication

- **EHR Visualization Component** is responsible for presenting the content of the Electronic Health Record of a patient. It stores data to and retrieves data from the Electronic Health Record database of the intLIFE platform. The EHR Visualization component comprises one of the major means of interaction of the Medical Expert User with the intLIFE platform. In order to plan the patient's treatment, a Medical Expert User needs to have current data and information concerning the patient's medical history. Based on data from the patient's EHR the Medical Expert User can decide how to proceed, decide whether intervention is required and determine a success of patient's therapy.
- **Medical Expert Assistant Component** is part of the intLIFE EHR and includes a set of tools that are supportive to the utilization of the intLIFE system by the Medical Expert (Questionnaire Module, Alerts Module, Recommendations Module, Observation Analysis & Diagnosis Module, Disease & Treatment Guidelines Module and Reports Generation Module). It is designed in a modular way, so that new tools can be added in the future. Currently, only the Questionnaire Module is available.

- **Vital Signs Monitoring Subsystem**
  - Medical Data Gathering Component; is responsible for gathering measurements from the peripheral medical devices. It is designed in a modular way, so that new devices can be added in the future. It incorporates a Medical Device Adapter Module for each device to communicate with, and a Comms Module to transfer measurements to intLIFE server.
  - Medical Device Adapters Module; They implement the interfaces between the medical devices and the intLIFE platform, translating the vendor- or even device- specific message structures to a common structure, in order for the measurements to be seamlessly integrated to EHR.
  - Communication Module; It securely transfers the measurements collected from the peripheral medical devices. It can use different protocols (HTTP, FTP, etc.) and different encryption algorithms. It provides graphical user interface for configuring the communication parameters.
  - Vital Signs Viewer Component; The Vital Signs Viewer Component provides the Healthcare Professional User with an effective graphical user interface through which he/she may retrieve from the EHR database and visualize measurements from peripheral Medical Devices.
- **Personal Health Record application**: provides an IP-TV interface to the intLIFE EHR. Automatic login, reveals only those EHR tabs that are relevant to the user (e.g. exclude visits, etc.). Moreover, in addition to the typical EHR the PHR application supports personalized messages, personal reminders, personal rehabilitation plans, questionnaires, personal trainer presenting educational material, and videoconference module.
- **intLIFE Management Subsystem**: the Administrator Web Interface enables the System administrator to have access to a set of administrative tools:
  - Users Administration; this process activates the necessary web user interface controls that enable the administrator to manage intLIFE users. The administrator is able to add (register), update and deactivate or reactivate intLIFE users.
  - Equipment Administration; This process activates the necessary web interface controls that enable the administrator to manage intLIFE equipment, i.e. it is a device manager that associates medical devices, STBs and other terminal equipment to physical or logical entities (patients/healthcare professionals or network nodes, respectively).

6.7.3 FREE and Open Source Implementations (http://www.goomedic.com/open-source-emr-list#)

The popularity of EHR/PHR systems have given raise to development of Open-Source platform too. Their capabilities and features can closely compete with

commercial and proprietary implementation, except when it comes to interoperability and flexibility in developing of add-on services and applications. The list of most attractive Open Source EHR/PHR solutions suitable for serving as a base for the development of the project solution include:

- **INDIVO Health** (http://indivohealth.org/developer-community) original personally controlled health record (PCHR) system. A PCHR enables an individual to own and manage a complete, secure, digital copy of her health and wellness information. INDIVO integrates health information across sites of care and over time. INDIVO is free and open-source, uses open, unencumbered standards, and is actively deployed in diverse settings, in particular our own Children's Hospital Boston and the Dossia Consortium.
- **TOLVEN Patient/Clinician HR** (http://www.tolven.org) focusing on delivering:
  - Personal Health Record (ePHR) that enables patients to record and selectively share healthcare information about themselves and their loved ones in a secure manner.
  - Clinician Health Record (eCHR) enables healthcare actors to securely access healthcare information collated from any number of trusted sources relating to individual patients in a structured and easy accessible way.
  - Healthcare Informatics Platform enables all healthcare data to be stored and accessed via ePHR and eCHR solutions. It uses industry standard technologies and data models.
  - Health Analytics solution that enables all data stored in the TOLVEN Platform to be extracted or analyzed for statistical purposes
- **HealtheMe** from KRM Associates Inc (http://www.krminc.com), an open source PHR system developed as part of a Medicaid eHealth transformation initiative for use in West Virginia, known also as HealtheMountaineer.
- **OpenEMR** (http://www.open-emr.org) is a Free and Open Source electronic health records and medical practice management application that can run on Windows, Linux, Mac OS X, and many other platforms. OpenEMR is ONC Complete Ambulatory EHR certified and is one of the most popular open source electronic medical records in use today. OpenEMR is supported by a strong community of volunteers and professionals. The OpenEMR community maintains OpenEMR's as a free, open source software solution for medical practices.
- **OpenMRS** (http://www.open-emr.org) is both software and a community. OpenMRS is a Java-based, web-based electronic medical record. It started from a simple data model, wrapped into an API, and then built a web-based application that uses the API. The OpenMRS API works like a "black box," hiding the complexities of the data model beneath it and ensuring that applications and modules using the API work with a similar set of business rules for managing the electronic medical record system data.

At the heart of OpenMRS is a concept dictionary. It defines all of the unique concepts used throughout the system. Using combinations of questions and answers, observations (observable data) can be defined as well as forms that gather multiple observations within a single encounter.

OpenMRS is constructed to support modules. Using modules, implementations are able to modify the behaviour of the system to meet their local needs without everyone having to agree on a single approach. Modules have full access to the system, so they can add tables in the database, alter behaviour of the API, and/or add or change web pages in the web application as needed to meet their needs.

6.8 Adoption Problems

Despite the need for centralizing patient information, the adoption of PHR has been very slow. A study (Kim and Johnson 2002) made to assess the functionality and utility of online PHRs, identified 19 websites offering different versions of PHRs. Centralized PHRs should help patients relate accurate history during clinical encounters, check for drug interactions, eliminate unnecessary duplication of laboratory tests and diagnostic studies, and serve as an information hub for patients' health management. An analysis of web-based PHR systems has revealed that most websites did provide access to personal medical information. However, each system demonstrated limited capacity in a different way.

From the 19 sites examined, four were applicable only to certain diseases; another four had recurrent technical problems or connections to a specific hospital's information system. The remaining 11 sites did not provide patients with sufficient guidance as to how they should enter personal data. Some of the sites allowed patients to select medical conditions from categorized lists, which did not cover the patients' complete health condition while others allowed free text entry. To formulate medication history, sites that required patients to choose medication from lists requested them to enter a wide range of descriptive information for each medication such as prescribed dose, administration frequency, start date, name of pharmacy that issued the medication and name of provider that prescribed the medication. With respect to laboratory tests, only two allowed patients to import results from outside sources. From these two sites, only one was functional. Not every site allowed patients to enter insurance coverage information. Majority of the sites required patients to enter date and results of diagnostic tests.

Most people do not keep record of minute details of their healthcare experiences and therefore find it difficult to make use of web-based PHRs. Overall, the sites selected for evaluation offered limited functionality to the public. Low adoption of web-based PHRs can be a direct result of limitations in these applications' data entry, validation and information display methods. Hence, the PHR development needs to be guided in the future by ample patient-oriented research.

6.9 Privacy and Ethical Concerns (http://en.wikipedia.org/wiki/Personal_health_record)

One of the most controversial issues for PHRs is how the technology could threaten the privacy of patient information. Network computer break-ins are becoming more common, thus storing medical information online can cause fear of the exposure of health information to unauthorized individuals. In addition to height, weight, blood pressure and other quantitative information about a patient's physical body, medical records can reveal very sensitive information, including fertility, surgical procedures, emotional and psychological disorders, and diseases, etc. Various threats exist to patient information confidentiality, example of are:

- Accidental disclosure: during multiple electronic transfers of data to various entities, medical personnel can make innocent mistakes to cause its disclosure.
- Internal leaks: medical personnel may misuse their access to patient information out of curiosity, or leak out personal medical information for spite, profit, revenge, or other purposes.
- Uncontrolled secondary usage: those who are granted access to patient information solely for the purpose of supporting primary care can exploit that permission for reasons not listed in the contract, such as research.
- External intrusion: Former employees, network intruders, hackers, or others may access information, damage systems or disrupt operations

Unlike paper-based records that require manual control, digital health records are secured by technological tools. Rindfleisch (1997) (Ball and Gold 2006) identifies three general classes of technological interventions that can improve system security:

Deterrents—These depend on the ethical behaviour of people and include controls such as alerts, reminders and education of users. Another useful form of deterrents has been Audit Trails. The system records identity, times and circumstances of users accessing information. If system users are aware of such a record keeping system, it will discourage them from taking ethically inappropriate actions

Technological obstacles—These directly control the ability of a user to access information and ensure that users only access information they need to know according to their job requirements. Examples of technological obstacles include authorization, authentication, encryption, firewalls and more.

System management precautions—This involves proactively examining the information system to ensure that known sources of vulnerability are eliminated. An example of this would be installing antivirus software in the system

The extent of information security concerns surrounding PHRs extends beyond technological issues. Each transfer of information in the treatment process must be authorized by the patient even if it is for patient's benefit. No set of clearly defined architectural requirements and information use policies is available.

6.10 Conclusions and Potential for Future Research

The introduction of EHR/PHR systems is the response to the inherent problem of the medical community in dealing with growing amount of papers and printed type of medical records. This becomes also a matter of costs as much time and money is wasted on copying, faxing, and retrieving paper files. Movement to electronically stored and managed patient records is both a simplification of the past problems, while adding new ones. Reduction of paper records helps in simplifying records and eases exchange of data, though electronic communication implies stronger focus on preventing access to such data. Hence, governments demand increasingly secure and standard-compliant health records. In today's world, it takes more than a simple document to meet national record keeping guidelines. And an increasing number of optional treatments must easily fit into today's recordkeeping systems, complicating matters even further.

Electronic Health Records are an obvious solution to all emerging problems in the medical care, offering simplification of growing patient records, stimulates easier exchange of data among medical professionals, contributes to cutting costs of medical care as a whole. Records are accessible by multiple health providers. They are also fully integrated with other office functions, and interfaces with multiple vendors and diagnostic equipment, allowing for e.g. easy storage or X-ray, MRI, and other images in a standardised way. Subject to providing sufficient safeguards at every level, a complete security of data may be achieved. Through data encryption, password protection, the electronic health record offers a peace of mind that data is kept away from unauthorised eyes. Nevertheless, although the future of e-Health has never looked so bright, there are still several concerns that needs careful attention. Growth of e-health systems inherently implies that any patient's data may be stored not in one place, but on several diverse systems implying increased risk of leaking information to unauthorised third parties. The cyber security procedures are also not consistent across various systems, implying that some of them may be easier broke into and increasing vulnerability of data stored there.

What adds to the problem is lack of seamless interoperability among e-health systems based on electronic records. Since early stages of development of HL7 standard, now one of the base reference standards for e-health, it was considered only as a set of guidelines and not a factual standard to follow. This has resulted in systems being built and deployed that had implemented only a part of the HL7 specification suited to particular needs of a given service provider. Interoperability among such restricted systems is a tiresome process, resulting in exchanging incomplete information. This deficiency has been recently recognised as critical for future e-health and the HL7 is being evolved to define the base set of interoperability criteria for ensuring smooth collaboration among different health systems. However, this process is still ongoing and requires much more research work, including interoperability at the device level and especially for mobile physiological monitoring. In conclusion we can observe a dramatic changes in the e-health domain with the

introduction of electronic health records, boosting the efficiency of medical services at a lower cost, at the same time offering still a vast range of research challenges that we may expect to be pursued and hopefully resolved in the near future.

References

ASTM WK4363 (2004) Draft standard specification for the continuity of care record, Version 1. ASTM International

Ball MJ, Gold J (2006). Banking on health: personal records & information exchange. J Healthc Inf Manag JHIM 20(2):71–83. PMID 16669591

Beale T (2002) Archetypes constraint-based domain models for future-proof in-formation systems. In: OOPSLA-2002 workshop on behavioural semantics, Seatle, WA

Beale T, Heard S (2008) Architecture overview [Online]. openEHR Foundation UCL Discovery publications repositor, http://discovery.ucl.ac.uk/66193

Beale T, Heard S, Kalra D, Lloyd D (2007) The openEHR reference model—EHR information model [Online]. openEHR foundation http://www.openehr.org/releases/1.0.2/architecture/overview.pdf

CCD. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6

CDA, HL7 clinical document architecture-release 2.0. http://xml.coverpages.org/CDA-20040830v3.pdf

CDA. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

CEN TC/251 (European Standardization of Health Informatics) ENV 13606, Electronic Health Record communication. http://www.centc251.org

CEN/TC251–ISO/TC215 I. E. (2010) Electronic Healthcare Record (EHR) communication. Parts 1: reference model, part 2: archetype model, part 3: reference archetypes and term lists, part 4: security and part 5: interface specification

CEN/TC251–ISO/TC215 I. E. (2010) Electronic Healthcare Record (EHR) communication. Parts 1: reference model, part 2: archetype model, part 3: reference archetypes and term lists, part 4: security and part 5: interface specification

Detmer D, Bloomrosen M et al (2008) Integrated personal health records: transformative tools for consumer centric care. BMC Med Inform Decis Mak 8:45

Device Driver Development Package. http://www.microsoft.com/en-us/download/details.aspx?id=26801

DICOM Standard. http://dicom.nema.org/

Dolin R et al (2006) hl7 Clinical document architecture release 2.0, Health Level 7

Eichelberg M, Aden T, Riesmeier J, Dogac A, Laleci G (2005a) A survey and analysis of electronic healthcare record standards. ACM Comput Surv 37(4):277–315

Eichelberg M et al (2005b) A survey and analysis of Electronic Healthcare Record standards. ACM Comput Surv 37(4):277–315

EN13606 Association. CEN/ISO 13606 EHR standard for healthcare providers, hospitals, IT-vendors and regional & national projects. 2009. http://www.en13606.org

European Commission (2007) eHealth priorities & strategies in European countries: eHealth ERA report

European Commission—DG INFSO/Q-REC (2007) Inventory of relevant standards for EHR systems. European Commission—DG Information Society/Q-REC

Ferranti JM, Musser RC, Kawamoto K, Ed Hammond W (2006) The clinical document architecture and the continuity of care record. J Am Med Inform Assoc 13(3):245

Goossen W, Goossen-Baremans A et al (2010) Detailed clinical models: a review. Healthc Inform Res 16(4):201

Heubusch K (2008) IT standards for PHRs: are PHRS ready for standards? Are standards ready for PHRs? AHIMA J 79(6):31–36
HL7 EHR system functional model: a major development towards consensus on electronic health record system functionality, white paper, 2004. http://www.hl7.org/ehr/downloads/dstu/EHR-SWhitePaper.zip
HL7 Template project (2007) Template project. Ballot draft
hl7 (2007) Health Level Seven International [Online]. www.hl7.org
HL7/ASTM (2007) HL7/ASTM Implementation guide for CDA Release 2-Continuity of Care Document (CCD®) Release 1. HL7/ASTM implementation guide for CDA Release 2-Continuity of Care Document (CCD®) Release 1. hl7 International/ASTM International
ICD, International Classification of Diseases (2013) World Health Organization (WHO) [Online]. http://www.who.int/classifications/icd/en
IHE Technical Framework. http://www.ihe.net/Technical_Framework
INDIVO Health. http://indivohealth.org/developer-community
Integrating Healthcare Enterprise. http://www.ihe.net
Integrating the Healthcare Enterprise (IHE) (2006) Integrating the Healthcare Enterprise [Online]. IHE. http://www.ihe.net
ASTM International (2005) A century of progress [Online]. ASTM International. http://www.astm.org/IMAGES03/Century_of_Progress.pdf
Intracom intLIFE EHR Platform. https://146.124.106.153:8181/intLIFEv1
ISO/TC 215 technical report: electronic health record definition, scope and context, 2nd Draft, August 2003
Johnson S (1996) Generic data modelling for clinical repositories. Am Med Inform Assoc 3:328–339
Kibbe D, Phillips R, Green L (2004) The continuity of care record. Am Fam Physician 70:1220–1223
Kim M, Johnson K (2002). Personal health records: evaluation of functionality and utility. JAMIA 9(2):171–180. doi:10.1197/jamia.M0978. PMC 344574. PMID 11861632
Kongstedt P, Quinn J, Shaman H (2004) Health information technology and the electronic health record: implications for healthcare organizations, CapGemini. http://www.richardsonwriting.com/uploads/Medical%20Records_4.pdf
KRM Associates Inc. http://www.krminc.com
LinkCare Alliance. www.linkcarealliance.org
LinkCare EHR. http://www.linkcare-eu.org
List of Open-Source PHR solutions. http://www.goomedic.com/open-source-emr-list#
LOINC (1994) LOINC [Online]. The Regenstrief Institute. http://loinc.org
McCay C, Kalra D, Worden R (2008) Results of investigating the transformability between HL7 V3, openEHR and EN/ISO 13606. NHS Connecting for Health, Leeds
MS Health Vault. http://www.healthvault.com/
Muñoz P et al (2011) The ISO/EN 13606 standard for the interoperable exchange of electronic health records. J Healthc Eng 2(1):1–24
National Committee on Vital and Health Statistics (2006) Personal health records and personal health record systems. U.S. Department of Health and Human Services, Washington, DC
Official Google Health WEB site. http://www.google.com/health
OpenEHR Foundation (2012) OpenEHR Foundation [Online]. www.openehr.org
openEHR Community. http://www.openehr.org
OpenEMR Project. http://www.open-emr.org
OpenMRS Project. http://www.open-emr.org
Personal Health Record (PHR). http://en.wikipedia.org/wiki/Personal_health_record
Project HER-IMPLEMENT. http://www.ehr-implement.eu
Qamar R, Rector A (2007) Semantic issues in integrating data from different models to achieve data interoperability. Stud Health Technol Inform 129(pt 1):674–678

Rector A et al (1993) A framework for modelling the electronic medical record. s.l. Methods Inf Med 32:109–119
Schloeffel P, Beale T et al (2006) The relationship between CEN 13606, HL7, and. s.l. In: Health informatics conference. Health Informatics Society of Australia, Brunswick East
SNOMED-CT (2009) SNOMED-CT [Online]. International Health Terminology Standards Development Organisation. http://www.ihtsdo.org/snomed-ct
Software Development Kits. http://msdn.microsoft.com/en-us/healthvault/jj126730
Sundvall E et al (2008) Integration of tools for binding archetypes to SNOMED CT. BMC Med Inform Decis Mak 8(suppl 1):S7
TOLVEN Patient/Clinician HR. http://www.tolven.org
UML Specification. www.uml.org
WHO ATC codes. http://www.who.int/classifications/atcddd/en/
WHO ICD-10. http://www.who.int/classifications/icd/en/
WMC on-Card. http://www.wmc-card.com/index.php?mapping=4
WMC on-Mobile. http://www.wmc-card.com/index.php?mapping=5
WMC on-WEB. http://www.wmc-card.com/index.php?mapping=6

Chapter 7
Wireless Sensor Networks: A Key Enabling Technology for Remote Healthcare

Steffen Ortmann, Peter Langendoerfer, Marcin Brzozowski, and Krzysztof Piotrowski

Abstract Recent advances in ICT and sensing technologies have created exciting options for individualised sensing and health monitoring. Wireless Sensor Networks (WSN) that are built of lightweight and autonomous devices called sensor nodes are a concrete example of such technologies. Each sensor node typically combines individual sensing, processing and wireless communication features into one small device. This chapter motivates the use of WSN as a key enabler for remote health care by introducing the manifold facilities and use cases of that technology. Based on that, it discusses the architectural basics and provides insights into practical system design issues, especially in view of reliability, energy efficiency and security of the system. After that an assessment of design goals and most critical challenges for applying WSN in health care is given. The chapter finally closes with presenting several selected solutions that successfully tackle introduced challenges.

7.1 Motivation

The prevalence of ageing in modern societies leads to an increased number of people suffering from Stroke, Cancer or Chronic Diseases (CDs) such as Cardiovascular Diseases (CVD), Diabetes, Asthma, Chronic Obstructive Pulmonary Disease (COPD) etc. Alarming numbers published recently stated that over 100 million people or 40 % of the European population aged over 15 years suffer from at least one type of a "chronic disease" whereas 2 out of 3 people aged over 65 years are already affected by at least two chronic conditions.[1,2] The World

[1] http://www.eupha.org/repository/publications/EU_HPF_Answer_to_Consult_on_CDs_Jan12.pdf.

[2] http://ec.europa.eu/research/innovation-union/pdf/active-healthy-ageing/steering-group/operational_plan.pdf and European Chronic Disease Alliance, WHO Europe.

S. Ortmann (✉) • P. Langendoerfer • M. Brzozowski • K. Piotrowski
IHP GmbH—Innovations for High-Performance Microelectronics, Frankfurt (Oder), Germany
e-mail: Ortmann@ihp-microelectronics.com; langendoerfer@ihp-microelectronics.com

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_7,

Health Organization (WHO) "considers the rise in chronic diseases an epidemic and estimates that this epidemic will claim the lives of 52 million people in the European Region by 2030".[3] As another example, Stroke is hitting about two Million people per year in Europe (Kirchhof et al. 2009) and the German Aerztezeitung predicts these numbers to increase by even 2.5 times until 2050.[4] Stroke is a leading cause of acquired disability and is also one of the main reasons for the need of care for elder people. In Germany, Stroke meanwhile is the third main cause of death (Heuschmann et al. 2010).

What all of the major diseases have in common is that their socio-economic implications are enormous. It is widely acknowledged that 70–80 % of healthcare costs are spent on chronic diseases—approximately equalling to 700 billion €—and the cost is expected to rise in the coming years.[5] According to the estimates of World Economic Forum and Harvard School of Public Health chronic diseases will cause a 47 trillion US $ global economic output loss over the period of 2011–2030.[6] Recent analysis also showed that productivity loss due to Cancer in UK alone in 2008 is approximately 15 billion[7] € and the cost associated with mental health problem is estimated to be 3–4 % of the Gross National Product (GNP).

The list of health impacts and respective social-economic implications can be easily extended by many other facts. However, the following major aspects require developing new means to cope with the effects of continuously increasing number of patients:

1. The reduction of quality and length of life of millions of people.
2. The enormous pressure for the socio-economic system in terms of health care costs, long time care as well as productivity loss.
3. According to statistics, the chance that a European citizen suffers from a major disease and/or a NCD in her/his life converges up to 100 %.
4. Constant aging of the society and increasing urbanization exacerbates the problem of adequate and good-quality care for all people, especially in rural areas.

Given these facts, there is a strong need for a radical change in disease management. Suitable approaches need to shorten or even avoid stationary hospitalisation and to improve the ambulant care model, in particular, at the home settings. The patients need to be involved into the care pathway to achieve a maximal outcome in terms of clinical treatment as well as quality of life. Wireless sensor networks are considered as a key enabler for novel telemedical concepts and implementations ensuring an adequate supply with little or no additional personal efforts.

[3] United Nations General Assembly 19 May 2011 Report by the Secretary-General on the prevention and control of non-communicable diseases (A/66/83).

[4] http://derstandard.at/1259281485986/Neuer-Gesundheitsbericht-Risiko-einer-Schlaganfall-Krise-in-Europa.

[5] See http://www.oecd.org/dataoecd/43/9/48245231.pdf and "The future of healthcare in Europe", The Economist Intelligence Unit Limited 2011.

[6] http://www.hsph.harvard.edu/news/features/features/noncommunicable-diseases-report.html.

[7] Policy Exchange, The Cost of Cancer, Featherson and Whitman, 2010.

The concept of telemedicine pursues letting patient health data "travel" around instead of the patients themselves. That means, instead of visiting the doctor directly, the patient is provided with technical equipment that monitor and assesses the patient's actual health state, for example by measuring blood pressure, heart rate or even body weight. Gathered data is submitted to the doctor via internet, who can then assess the data without the need to really "see" the patient. Thus, telemedicine offers remote health care applications for patients at home and during daily life (Ortmann et al. 2011).

The advances in ICT and sensor technologies have developed novel means for personalised sensing and health monitoring. Wireless Sensor Networks (WSN) are a prime example of such technologies. WSN consist of small sensor nodes that in principle are lightweight and autonomous mini-computers. Sensor nodes combine latest sensing, processing and wireless communication features into small autonomous systems that can be used to execute a common respectively distributed task. Given that, WSNs are ideally usable for personal health monitoring due to the following key facts:

- Customised monitoring of patients and their environment becomes feasible using sensor nodes. Sensor nodes may be attached to the patient's body, build into equipment or wearable garment or be installed in a fixed infrastructure, for example the home of the patient.
- Sensor nodes provide great flexibility due to the fact that these may be equipped with a variety of sensors, such as motion, biomedical or environmental sensors, or can be used for signal processing, assessment or forwarding of data etc.
- Due to their autonomous nature, sensor networks provide mobility by default. Equipped with portable power supply, for example a portable battery, these are enabled to independently carry out their tasks for weeks, months or even years depending on the power consumption.
- Wireless communication means on each node enable to communicate amongst sensors themselves or to public infrastructure such as the UMTS or GSM network. Thus, remote data access becomes feasible.
- Sensor nodes are low cost devices. The costs for single sensor nodes even decreases rapidly when several hundreds or thousands are used.
- WSN can provide so called self-X properties that are key for ease of use. This means sensor nodes can be programmed to react autonomously and flexibly on changing operational conditions. Therefore they are capable of learning and optimizing their behaviour during operation, e.g. to self-organise to changes in the environment, to self-heal the network in case of failing nodes, to self-coordinate distributed processing according to available resources etc.

The benefits of applying WSNs as health care technology are manifold. WSNs can play an important role while developing the next generation of health care technology by contributing to:

- *Ambient health care solutions* that enable a better and safer control almost anywhere at any time. Hence, medical monitoring will no longer be bound to a physical place such as a clinic. Instead, smart and ambient systems will overtake many of the tasks that are currently provided by health care professionals.

- *Improved well-being of patients* when being monitored or "treated" during daily life. Sensor networks worn throughout the day or installed in the daily environment, e.g. the patient's home, can control the patient's health state continuously even outside a clinical environment.
- *Reduction of stationary treatment* due to inherent flexibility of sensor nodes. Being small, lightweight and wearable, autonomous sensor nodes can amongst other things overtake patient specific monitoring tasks in the post-acute phase to shorten hospitalisation as well as for detection of future acute phases or ensuring well-being respectively to prevent future stationary treatment.
- *Reduction of costs* in many ways. In addition to the reduction of expensive stationary treatment, WSN also allow transferring regular evidence-based treatment means into home services or to support rehabilitation means and thereby empower patients to faster return home and to their job again. WSN might further reduce costs of future health monitoring devices too.
- *Enriched data sets* of patient's health states and disease progress. Monitoring solutions outside a clinical environment provide clinical experts with great long-term data sets not available today. This will lead to a better understanding of disease management and effective treatment.

Featuring such great potential and flexibility, wireless sensor networks are already used for different application fields in modern e-health technologies and telemedicine scenarios. Amongst many others, current research focusses on the following four main application areas:

Acute care: Sensor networks or lightweight sensing boards that are used to gather various bio-signals and health parameters such as ECG, EMG, blood pressure, breath rate, O_2 saturation etc. on demand or in emergency cases. Taken data is then not only accessible by the paramedic but can in parallel be submitted to the next clinic for proper preparation of equipment or surgery if necessary.

Rehabilitation: After stationary or acute care, sensor networks can still be used to monitor the recovery process. That does not only cover bio-signals. In fact, they can also be used for activity monitoring or training supervision to measure the progress of rehabilitation for example for stroke survivors[8] (Ortmann et al. 2012) or people suffering from Parkinson's disease[9] (Edwards 2012).

Prevention: Preventive collection of bio-signals mostly concerns continuous monitoring of people suffering from chronic diseases. For example, the most known case are Cardiovascular diseases (CVD), which are monitored with mobile ECG and daily taken body weight (Chen et al. 2012).

Mobile diagnostics: Mobile and lightweight diagnosis means are obviously part of all three application scenarios mentioned above. However, there exists many more

[8] www.strokeback.eu.

[9] www.cupid-project.eu.

technical solutions ranging from mobile and instantaneous measuring of blood parameters with Lab-on-Chips (2011) over monitoring of fire fighters in action (Piotrowski et al. 2010) up to implantable devices measuring the blood glucose level continuously (Birkholz et al. 2009; Fröhlich et al. 2012). In combination with communication and flexibility means provided by WSN's, these approaches can provide novel value-added e-health services.

7.2 Background

Wireless sensor networks have been introduced in research about 15 years ago. The core idea was that extremely cheap devices can be deployed in a very large number and which are then forming a network without human intervention. The sensor network is then gathering data about its environment. In order to facilitate such an idea the devices to be used need to be extremely cheap and battery powered to keep the coats at a reasonable level and to enable autonomous operation. The standard sensor node is built of a microcontroller, volatile and non-volatile memory, a wireless communication interface i.e. a transceiver and an interface to the sensing device, see Fig. 7.1.

Table 7.1 shows common sensor node platforms including their micro-controllers, available memory and radio interfaces. Most of the micro controllers come with very limited computing resources and memory, resulting in serious challenges for software developers. The other limiting factor of sensor nodes is the fact that they need to run out of a battery i.e. they may use very little energy in order to ensure long lifetime.

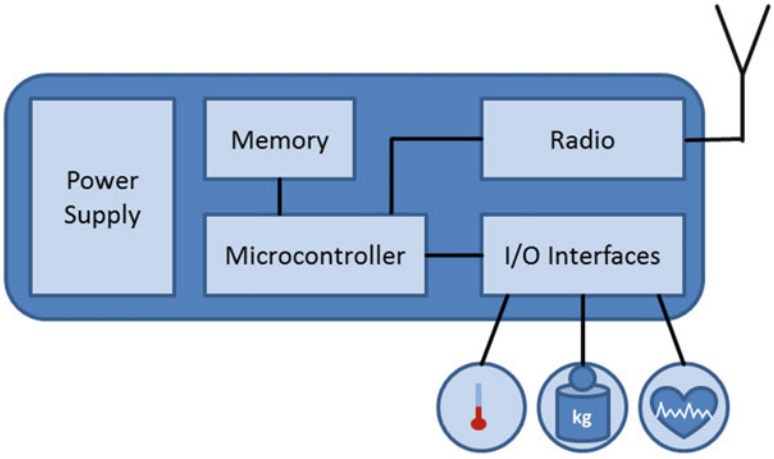


Fig. 7.1 General assembly of a sensor node that can integrate various external sensors

Table 7.1 Overview of existing sensor platforms and their basic features

Platform	μC	f_{max} [MHz]	RAM [KB]	Code [KB]	Storage [MB]	Radio
MICAz	ATmega128L (ATMEL 2011)	8	4	128	0.5	CC2420 (TI Inc. 2007a)
MICA2	ATmega128L	8	4	128	0.5	CC1000 (TI Inc. 2007c)
IRIS	ATmega1281 (ATMEL 2010b)	8	8	128	0.5	CC2420
TelosB	MSP430F1611 (TI Inc. 2011)	8	10	48	1	CC2420
IHPNode	MSP430F5438A (TI Inc. 2010c)	25	16	256	4	CC1101 (TI Inc. 2010a) CC2500 (TI Inc. 2009) CC2520 (TI Inc. 2007b)
Chronos	CC430F6137 (TI Inc. 2010b)	20	4	32	–	CC430F6137 (CC1101)
Sun Spot	AT915SAM9G20 (ATMEL 2010a)	400	1,024	8,192	–	CC2420
Imote2	PXA271 (Marvel 2010)	416	256 + 32,768	32,768	–	CC2420
Shimmer (Burns et al. 2010)	MSP430F1611	8	10	48	µSD-card	BT RN-42 CC2420

7.2.1 *Original Application Areas*

WSN's have originally been developed out of the vision of pervasive intelligent environments that may surround and serve us at any place and any time (Weiser 1991). Long time ago in the early nineties, this computing paradigm already predicted computing devices to be embedded in everyday objects, e.g. in coffee cups (Gellersen et al. 2000) or garment (Patel et al. 2012), allowing information technology to fade into the background and become nearly invisible to their users. As one of the first real world examples enabling pervasive computing, WSNs have become a rising star in this research field. Envisioned to be distributed like "Smart Dust" (Kahn et al. 1999, 2000), these networks support a broad range of applications (Aboelaze and Aloul 2005; Akyildiz et al. 2002) and may become the perfect service and surveillance tool (Bohn et al. 2004). Based on their capabilities to identify physical phenomena, sensor networks can be applied for environmental and structural control (Dikaiakos et al. 2007; Mainwaring et al. 2002; Sun et al. 2005; Werner-Allen et al. 2005), context-awareness for personal services (Robinson and Beigl 2003), military applications (Gillies et al. 2009) and ubiquitous healthcare (Morchon et al. 2009), to

mention a few. As a particular variant of WSNs, Body Area Networks (BANs) that utilise sensor nodes attached to the human body have emerged rapidly for a broad range of novel e-health applications. The following paragraphs will therefore focus on applying wireless sensor nodes for body area networks.

7.3 Architecture

Wireless body area networks can be used to monitor vital parameters of patients and even healthy persons. These networks can in addition monitor also the environmental conditions of the monitored persons. While under normal conditions such monitoring is a task that can be performed locally and even the assessment of the measured parameters can be done by the BAN itself, it is important to allow for reporting features to make full benefit out of such a solution. There are two reasons to aim for an always connected solution:

- If vital parameters are above/below a predefined threshold the BAN can send the measured values directly to a telemedicine centre so that appropriate actions can be triggered immediately.
- The measured values should be stored even while being in the allowed range of values to enable anamneses later on if necessary, or to allow the patient to do kind of self-assessment. Due to the limited resources of the BAN devices they cannot be used as long term storage i.e. they need to be connected to a more powerful device.

For providing connectivity to a BAN a lot of different networking solutions can be used. It ranges from wireless LAN and standard Internet to GSM/UMTS communication. Figure 7.2 illustrates these possibilities and depicts the communication end point i.e. the BAN and a telemedicine centre. The protocols used for ensuring the always connected feature are out of the scope of this chapter. The only issue to be considered when designing the BAN is the specification where the bridging functionality between the BAN and the cellular network or the Internet is

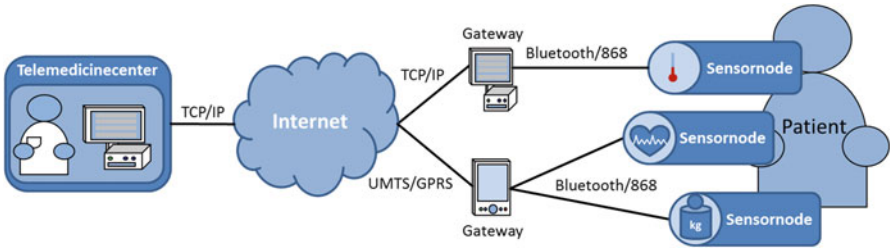


Fig. 7.2 Main principle of BAN applications. Sensor nodes attached to the patient's body wirelessly communicate to a gateway device that is connected to a global network over which it can reach the telemedicine center

to be placed. It can be allocated inside the BAN or outside the BAN, e.g. when a Smartphone is used. For practical reasons, i.e. the design of Smartphones normally won't be heavily influenced by BAN designers, the bridging functionality will be placed inside the BAN that then has to provide a communication link that is compatible with those provided by Smartphones. Here low-power Bluetooth (Bluetooth 4.0) is a highly probable candidate. Whether or not the Smartphone as the most probable gateway to the outer world then uses a cellular network or the Internet, may depend on:

- Features of the Smartphone i.e. supported communication devices
- Availability of the one or the other type of network
- Emergency consideration
- Certifiability as a medical product[10]

This short discussion of the integration of the BAN into a telemedicine service architecture should show that we are aware of those problems, and that there are solutions at hand, whereas a detailed discussion of these issues is out of scope of this chapter.

7.3.1 Centralized vs. Distributed BAN Architecture

When designing BANs a lot of decisions need to be taken. One of the first decisions and one with an extremely high impact on the final design is, whether or not the BAN will be centralized or distributed. In addition, there are a lot of design goals such as dependability of the BAN that need to be ensured independent of the type of the BAN network. But please note that the decision whether the BAN is centralized or distributed has considerable impact on the approaches to be applied for ensuring features of the BAN such as dependability, availability, power supply etc.

While there are good reasons for both approaches—centralized and distributed—which we will discuss here, we will focus on distributed architectures when going into a more detailed discussion of design goals and means to achieve these goals.

7.3.2 Centralized BANs

The basic idea of a centralized BAN solution is that there are more or less dumb sensor devices that are capable to measure and send parameters such as heart beat, temperature etc. All computation is then done on a central intelligent BAN node that comes with a processor and that knows how to assess the raw data sent by the sensing nodes. The right side of Fig. 7.3 illustrates this approach. Such a centralized model has the following benefits. The central device e.g. the Body Central Unit

[10] Designers should be aware of the simple fact that an increased number of features and capabilities provided by the device directly lead to a more complex certification procedure.

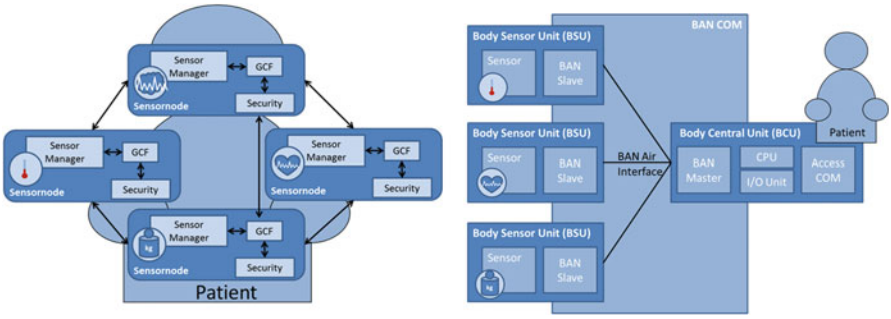


Fig. 7.3 Conceptual design of distributed and centralized BAN architectures. Whereas a distributed approach allows all nodes to communicate independently to each other (left side of the figure), the centralized approach forces all nodes into one-way streaming to a central unit (right side of the figure)

(BCU) in Fig. 7.3 can be used as a scheduler for medium access. This allows using relatively simple Medium Access Control (MAC) protocols since only the BCU can initiate communication, i.e. all other devices are actively polled by the BCU. This in addition helps to reduce the memory needed for the software that is deployed on the sensor nodes. Such a centralized approach also facilitates pretty simple management of sleep cycles that are used to reduce the energy consumption of the sensor nodes. The open issue here is the synchronization of the wake-up times. These issues will be discussed in detail later in this chapter. Last but not least, the implementations of centralized BANs may be easier to certify as medial product. This assumption is based on the fact that the majority of the devices used in the BAN are relatively simple and equipped with little software only, i.e. the amount of components—hardware and or software—is small.

But this approach comes with some drawbacks. All measured data need to be transferred to the central unit for assessment. This puts a serious communication load on the sensing devices. Here data and not information is transferred. Since transmitting data is more power consuming than processing data it seems reasonable to do evaluation of measurement data directly after measuring it instead of sending it to the central unit. The central unit is a single point of failure, i.e. if it fails the whole BAN can no longer provide its service. Such failure may put the patient's health at serious risk. The raw data will no longer be assessed and it is even impossible to communicate it to a telemedicine centre. From a security and privacy point of view the centralized approach is extremely challenging. If the sensing devices really come with extremely limited resources—less than normal sensor nodes have—it is infeasible to run cipher algorithms or other protection means. This means that the measured values are send in plain to the central unit, which in turn means that these values are accessible by any potential attacker in the vicinity of the BAN user, rising serious privacy concerns. Security questions as well as feasible solutions will be detailed later in this chapter.

7.3.3 Distributed BAN

The left side of Fig. 7.3 illustrates the distributed BAN approach in which all sensor nodes in the BAN are equally equipped with respect to processing capabilities, transceivers etc. Here the major difference in the sensor nodes are the sensing devices attached to the sensor node. Due to the concept of almost identical sensor nodes a kind of distributed intelligence can be realized. I.e. in such a system each of the sensor nodes can do an assessment of the values measured by itself, so that raw data does not need to be send so the energy consumption can be reduced compared to the centralized model. The fact that all sensor nodes are identical provides kind of resilience since the nodes can replace each other. Thus if one of the sensor nodes fails the BAN can still work and provide measurements at least for those values for which the sensors are functioning. The processing features on each of the BAN nodes enable the use of proper cipher means to protect the data from eavesdropping when being exchanged in the BAN. The realization of security and privacy features is an essential and non-trivial task, on which we will elaborate later in this chapter.

The improved dependability of the distributed BAN concept comes with some drawbacks. All sensor nodes are more or less equal, i.e. a kind of distributed control is required. This affects the protocol stack especially the MAC layer. Here no centralized scheduler, such as the BCU before, is available so that the MAC protocol needs to be able to cope with contention based medium access, while trying to allow for energy efficiency etc. This requires a more complex software to be installed an all nodes, consuming additional memory which is a scarce resource on senor nodes.

7.4 Design Goals and Challenges

In the originally addressed application areas the holy grail for designers was—and still is—energy efficiency since here the sensor nodes are expected to work unattended for years. This situation changes in the area of BAN design. The sensor nodes are easy accessible and reloading /exchanging of batteries can be afforded. But even if the intended lifetime of a BAN does not need to be years, often exchanges of batteries will have significant and negative impact on usability and acceptance of BANs, so energy efficiency is still an important design goal. Not to mention that less power consumption certainly allows using smaller and more lightweight battery packs of course. But other features are considered to be equally important for convincing BAN designs. As human being lives might depend on BANs the following features are essential:

Safety: this means the BAN may not put the patients' health at risk. To ensure this is tricky but as long as the BAN is used for monitoring only, i.e. no actuators are included, only a few features need to be considered. Radio transmission is one of

these features. Here the exposure of the patient should be kept minimal which can be achieved in the design by reducing the communication frequency and transmit power. As long as the hardware works according to its specification there should not be any problem. In order to ensure this the hardware needs to be measured from time to time. Another challenge might occur due to possible (over-)heating of electrical components and battery packs during operation or charging. Here the usage of Negative Temperature Coefficient Thermistors (NTC) is common to monitor the temperature of devices or partial components. NTCs are available as single parts but are also available as fully integrated solutions e.g. in enhanced battery packs.

Reliability/Availability: this means the BAN need to work correct at every point in time when in use. There is a certain probability for any component in the BAN to fail be it permanently or transient. Redundancy is a means to increase the overall reliability. It should be applied for all parts of the BAN. So it should be taken into account when designing individual sensor nodes, e.g. by integrating more than one transceiver and micro controller, but also when designing the BAN by integrating two or more nodes providing the same functionality. Also for software redundancy should be considered. Software can also be used to provide redundancy if replication strategies are supported by the software used in the BAN. We will introduce such a concept later on. Even though redundancy is a key concept to achieve reliability/availability, energy efficiency of the overall design is nearly as important as redundancy if not more important. If one of the nodes run out of battery supply it can be replaced by another one using redundancy, but energy efficiency can help to even avoid the failure of the node.

Security: or more precisely confidentiality and data integrity are essential in BANs used for health care applications. Here all data is extremely sensitive and needs to be protected against eavesdropping. While this might be considered as convenience feature, data integrity is required to ensure correct behaviour of the telemedicine system. If measurement values are falsified during transport, this may have significant impact e.g. emergency situations are not detected, negatively affecting the patient's health status. In Halperin et al. (2008) the authors have shown that eavesdropping and even manipulating of a pacemaker can be easily done. The problem with integrating reasonable security functions is that most cipher algorithms require significant computation effort and by that consume energy. Both are scarce resource so that the design of the cipher means i.e. selection and implementation needs special attendance.

Privacy: For acceptance of telemedicine solution ensuring privacy is a core issue. Protecting the patients' data requires strong cryptographic means to ensure confidentiality whenever data is transmitted. This holds true for the communication inside the BAN but also for data transport via open networks to the telemedicine centre. The most secure and advanced solution is having a real end-to-end security

in which the sensor nodes use the same crypto system as the telemedicine centre. Ortmann (2011a, b) introduce such an architecture. That way transmitted data is protected regardless whether or not unknown or public devices/gateways are used as communication hop (Ortmann and Maaser 2011). But the more tricky issue is to protect the data after it is received in the telemedicine centre or stored in patients' health records. Here different types of access model that take the current role of the data user into account can be applied, i.e. whether it is a nurse or physician (Maaser and Ortmann 2010). A very interesting concept for privacy protection is proposed in Scheffler et al. (2011). Here not only the patient is considered to be a data owner, but also the physician who did the examination and updated the health record is considered to be data owner. These privacy enhancing techniques are mainly related to the back end, where patient data is processed. Even though privacy protection is a major concern there might be conflicting situations in which privacy might put the patient at risk. Consider kind of emergency situation in which the patient suffers from a heart attack or similar. First aid arrives but may not access the patients vital data stored in the BAN due to privacy settings and still on-going encryption of measurements values. Conflict resolution techniques have gained some interest in the recent past (Naqvi et al. 2010). But in this chapter we will focus on the basic security techniques that need to be available inside the BAN to allow for privacy at all.

Usability: Despite clinical/therapeutical advances and opportunities, usability is by far the most crucial acceptance factor! That means a comfortable and easy way to use the sensor nodes is the key to success in the real world beyond any research trials. The idea that patients are going to use BANs at home requires usability in several aspects. First, almost none of the patients will be computer experts, which means the BAN may not need any kind of configuration, at least not after a first set-up phase. In fact the BAN design is required to ensure that a single push button is going to bring the BAN into a full operation mode. Here the self-healing and self-configuring features of wireless sensor network are providing a solid basis to do such a design. Wireless sensor network have been considered to run unattended and by that need to do kind of network bootstrapping without human intervention. Nevertheless, body worn sensors should work nearly "invisible" and must not hinder the patient in her/his daily life activities. The sensors should be small and lightweight but still allow for monitoring the patient for several hours or the whole day respectively. The latter is certainly a trade-off between power consumption and an appropriate battery supply, which grows in size and weight with the amount of energy that is needed. The second point is physical handling. Some patients may be elderly people some suffer from physical handicap so that attaching the sensor node of the BAN might be difficult to those patients. This requires that the BAN shall come with some help to apply it properly. This can be achieved for example by integrating the sensor node in clothes, or by appropriately selecting easy to reach parts of the body for attaching the sensor node etc. This aspect will not be detailed further in this chapter, since it is more an ergonomics than computer science issue.

7.5 Selected Solutions

After discussing design issues and challenges we are going to introduce some solutions for the before mentioned challenges. The selection of the solutions is done mainly on our own experiences with those solutions so we are not aiming at a full survey. There are for sure more approaches that are most probably equally suited, without being mentioned.

7.5.1 Redundancy for Improved Availability and Reliability

Redundancy is a key feature when it comes to reliability issues. Redundancy means using several equal components in parallel to detect and overcome failures. Redundancy allows substituting functionalities, components or even complete sensor nodes during operation. The latter has been discussed in the previous section of distributed BAN architectures, where more or less all nodes provide the same hardware and can therefore overtake tasks from other nodes. In the following, we will therefore focus on two approaches providing reliability in wireless communication and by data redundancy.

Reliable Communication

Wireless communication is inherently prone to errors such as link failures, packet errors during transmission, physical disturbances e.g. caused by water or metal, or blocked frequencies due to other devices such as microwave ovens at 2.4 GHz. The various radios available at the market certainly offer significant differences in power consumption, communication distances and reachable data rates. In addition, the physical destruction of radios or whole sensor nodes, e.g. in case of falls, need to be taken into account as well. In summary, there exists dozens of issues that might hinder the wireless communication to work properly at least for a certain time period.

Since we target on availability and reliability of wireless communication, i.e. ensuring that communication is always feasible, integrating more than one radio into one sensor node might be a suitable solution. To again increase resilience against failures, all nodes used in such BAN architecture should be equally equipped. To show evidence that such approach really works, we here report on the *IHPNode* sensor platform originally developed for monitoring vital and environmental of fire-fighters in action (Piotrowski et al. 2010). This means, the BAN of *IHPNodes* locally gathers all data and then needs to transfer it to the head of mission who is usually outside of the emergency area, e.g. outside of the burning building. While the pure reporting of data from moving persons is a challenge as such, we have faced almost the worst case scenario for wireless sensor networks because of

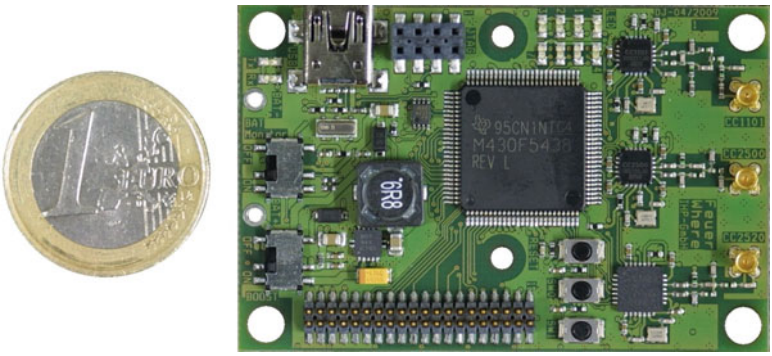


Fig. 7.4 Top view of IHPNode in comparison to the size of a 1 Euro coin

the harsh environmental conditions, i.e. heat, fire, acids, 100 % humidity (water steam) etc.

However, to cope with communication errors, we have integrated the following transceivers in parallel on each to provide redundancy in radio connectivity:

- TI CC1101: Low cost transceivers working in the 868 MHz band
- TI CC2500: Low cost transceivers working in the 2.4 GHz band
- TI CC2520: ZigBee™ Transceiver working in 2.4 GHz band

We applied two pin and logic compatible transceiver chips working in different radio frequency bands, i.e. the first in the European 868 MHz band and the second in the 2.4 GHz. The third transceiver (ZigBee) is also working in the 2.4 GHz band and provides 802.15.4 support. It was applied to be able to communicate with other known node platforms. The IHPNode is empowered by the MSP430F5438 Microcontroller from TI. The complete sensor node is depicted in Fig. 7.4.

Redundant Data Storage

Protecting from data loss by applying redundancy in data storages is of common use for decades in server applications, e.g. by using RAID systems. However, full redundancy in sensor nodes is counterproductive due to the enormous additional efforts in terms of processing power and energy consumption. On the other hand, mere redundancy at single nodes will not be helpful when complete sensor nodes fail. To still cope with potential data loss or failed sensor nodes, we have investigated how to provide redundant data storage means in a comfortable way in sensor networks. As a result, we will report on the *tinyDSM* middleware approach that introduces the concept of Distributed Shared memory (DSM) to sensor networks. The tinyDSM middleware:

- Provides a data storage model that not only supports replication of measurements but ensures also consistency of data,

- Supports storing of historical data up to a specified number to facilitate anamnesis,
- Provides an event detection mechanism based on the shared data.

Compared to passive data storage approaches, like tinyDB (Madden et al. 2005), cougar (Yao and Gehrke 2002) or tinyPEDS (Girao et al. 2007), tinyDSM provides an active data monitoring. There are also other approaches that support distributed computation and data storage for WSN, like (Abdelzaher et al. 2004; Costa et al. 2007; Gummadi et al. 2005), but they are usually tied to a specific operating system. Our middleware is implemented in pure C programming language and provides a clear interface to both, the underlying operating system and the application on top of it. Encapsulated in an OS adaptation layer (wrapper) it can work with all C based operating systems, and thus, it supports heterogeneity in both, hardware and software (or OS) dimension.

The tinyDSM middleware provides means that allow sensor nodes to autonomously share their data in an application dependent way. Moreover, tinyDSM middleware supports an event mechanism based on the shared data. Therefore a predefined set of shared variables is known to all nodes in the network. There are two types of variables—local variables and global variables. A variable defined as local has as many entities as many nodes are in the network, i.e., each node owns an entity of a local variable. Each global variable has only one entity existing in the whole network. In this context, an entity is a unique and independently addressable data unit in the shared memory space. Thus, to address a variable defined as global, only a reference to the variable is needed. In case of addressing a local variable it is necessary to point at a specific node—the owner—to address the right entity of the variable.

There are two operations that the application can perform on a variable, i.e. the WRITE operation and the READ operation. If an operation requires multiple nodes to interact, there is a need to exchange messages to fulfil the task. For example, since the sharing of the variables is based on data replication, once a new value is written to an entity of a variable in a WRITE operation, an UPDATE message is broadcast to all the nodes that hold a copy of the entity. Updates are autonomously handled by the underlying middleware layer without the need of extra programming effort. Data replication makes the system robust against data loss caused by lost nodes and reduces the overhead of the read access if the data is available locally on the nodes. The span of the data replication and its density is controlled by a policy chosen at the compile time.

The integrated event mechanism allows the application developer to define runtime events at compile time that shall be detected during operation. Each event is described as a logical equation evaluated each time any variable included in this equation changes. The result of this evaluation is stored in a variable that is automatically created for each event. From the application perspective, these event variables are read only and store only logic values, i.e., true or false. According to the event definition, the application is notified about the evaluation result. For more details about configuration of variables, events and the middleware interfaces please refer to Piotrowski et al. (2009, 2010).

Table 7.2 Percentage of tinyDSM messages delivered between node pairs inside the BAN

NodeID	1	2	3	4	5	6
1		87.2	93.7	86.3	83.2	85.5
2	86.1		94.3	92.1	79.7	89.0
3	89.3	92.0		85.9	82.8	86.6
4	85.9	89.0	87.1		80.8	88.9
5	88.6	87.2	93.9	88.1		86.4
6	86.1	87.5	88.1	94.4	79.8	

In order to verify the distributed mechanisms we ran five over 2 h-long tests of our system with focussing on the weakest point of this distributed system, i.e., the networking in general and the packet loss rate in particular. The middleware in its basic configuration does not include the network protocols and relies on external ones, i.e., the medium access controlled by the transceiver only. The BAN consisted of six nodes, each randomly generating the values of 2–7 parameters, resulting in a total of 26 updates per second. An UPDATE message was 120 bits long, while the data rate was set to 400 kbps.

Based on that we could estimate the quality of the replication in the system. In order to provide a reliable replication, the copies of the entities need to be updated frequently, i.e., it is important that the packets are not missing in series. Table 7.2 shows the percentage of middleware requests that were delivered. The columns represent the packets sent and the rows the packets received by each node. The average number of packets missing in a series was below 1 and the maximal length of a series was 18, which is less than 1 s in time.

Using the *IHPNode* and the *tinyDSM* data storage concept we have been able to provide a proof of concept by testing the nodes in a real fire drill. Therefore the sensor nodes together with a couple of sensors have been packaged into fireproofed cases. The local BAN sensors have communicated amongst each other using the CC2500 transceivers at 2.4 GHz while at the same time actual vital parameters have been reported to the squad leader outside of the burning building via the CC1011 transceiver at 868 MHz. If single BAN nodes had no direct link to the squad leader's device, one of the other BAN nodes has been used as communication hop. Similarly, nodes equipped with inactive redundant sensors observe the primary nodes and in case the value is not updated for 5 s they take over the management of the parameter variable and provide the data. Note that such self-healing feature is only feasible when all BAN nodes provide similar or redundant resources and features as previously introduced as decentralized BAN approach. The graphical user interface and the basic equipment used in the fire drill are shown in Fig. 7.5.

7.5.2 Ensuring Data Integrity and Privacy

When it comes to providing security in wireless sensor networks two core problems have to be solved, i.e. key distribution and energy efficient implementation of the selected cipher algorithms. For key distribution in WSNs a lot of proposals have

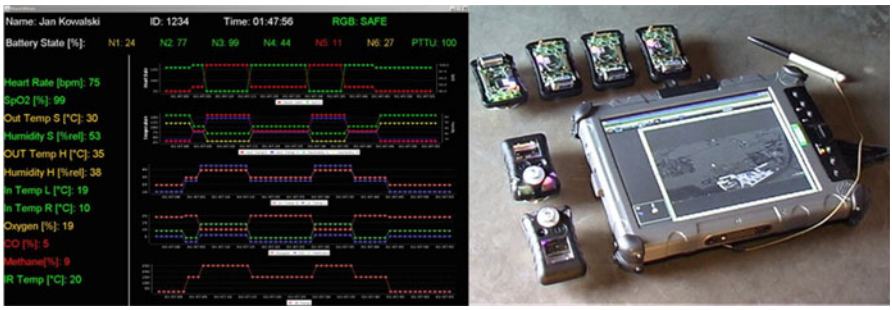


Fig. 7.5 Tactical screen of the squad leader with vital parameters of a single fire fighter in action (left side) and a set of packaged BAN sensors with the squad leader device (right side)

been made that fully rely on lightweight secret key cryptography such as SPINS (Perrig et al. 2001) and the system proposed in Chang et al. (2007), but there are also proposals that use public key cryptography such as those presented in Karlof et al. (2004), Shaikh et al. (2006). But in the field of telemedicine where the number of nodes is very limited and deployment is done in a controlled way the key distribution is not as important as it is in standard WSN application fields since keys can even be manually programmed into each device. But energy consumption of the applied crypto system is crucial for the lifetime of the BAN. Here especially the fact that data integrity is a core issue since it almost directly requires using public key cryptography to ensure integrity and non-repudiation. In principle two public key scheme could be applied, RSA and Elliptic Curve Cryptography (ECC). The resource restrictions rule RSA out since it requires a key length of at least 1024 bits which is extending ten the messages sent by the BAN extremely, requiring not only significant processing but also transmission power. In contrast to this, ECC can provide the same level of security with key length of 160–233 bits. But even ECC puts a serious burden on the micro-controller. Here serious efforts have been made to reduce the energy consumption of the ECC processing see for example (Ugus et al. 2009). But also the memory consumption needs to be taken into account on a micro controller a space optimized implementation was proposed in Uhsadel et al. (2007). The core problem with software implementations of ECC (as well as with any other algorithm) is that memory can be traded versus speed and vice versa. In other words, if the implementation is fast and by that energy efficient the memory provided by the sensor nodes might not be sufficiently large enough to install all needed software modules. If the implementation is very small it will most probably drain down the battery quickly. Here efficient hardware implementations help to cope with both issues. They require almost no energy for processing the underlying mathematical functions of the crypto system, and do not allocate memory for the implementation. To the best of our knowledge the most efficient hardware implementation of ECC with a key length of 233 bits was proposed in Dyka and Langendoerfer (2005), a preliminary but still very efficient implementation was integrated with a 16-bit microcontroller (Panic et al. 2011). Hardware

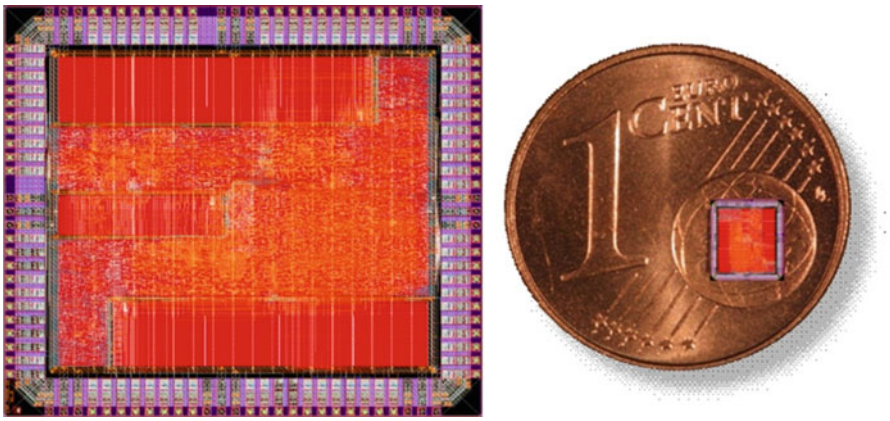


Fig. 7.6 Layout of the IHP-crypto-microcontroller (www.ihp-microelectronics.com) with a 128-pins pad frame. Additionally the size of the chip is compared to the size of a 1 Euro cent coin

implementations of specific functions such as crypto operations are often considered as too costly for sensor nodes but their impact on the total cost of a telemedicine system is negligible. Their size is about 2.5 mm^2 in an old fashioned 250 nm technology resulting in an additional prize in the range of several 10 Euro cents, and without any recognizable effect on the form factor, see Fig. 7.5 for an example of our 16 bit microcontroller enhanced with hardware accelerators for AES, ECC, SHA-1. An interesting software based approach is presented in Sojka et al. (2010). Here the authors propose a lightweight security approach based on modification of elliptic curves cryptography. The reduction of the length of the security parameters influences the security level but also helps to save the energy needed for computation and especially for communication, which is not achieved with the hardware based approaches (Fig. 7.6).

The performance and the energy consumption of the accelerators integrated into the IHP-crypto-microcontroller are given in Table 7.2. For the ECC component the elliptic curve B-233 (standardized by NIST) was hard-coded. As expected, the AES and ECC versions implemented in hardware consume orders of magnitude less power while boosting the performance at the same time. Since performance efficiency and energy consumption are two of the main factors for usage on sensor nodes and the protection of gathered private data is mandatory in the medical domain, integrated hardware accelerators provide the ideal solution for cipher means on resource constrained devices like sensor nodes.

Even with the encryption of data on their way from sensor source to the sink in the e-health centre and an according effective access control (e.g. by appropriate key management), the privacy of patients is still not fully protected (Table 7.3). The patient remains in control of who may access her/his data. Nevertheless, the patient cannot control the extent to what the data is communicated (compare to the definition of privacy in Westin (1967)) to those that may in principle be entitled. That is, with the mere access control and encryption, a doctor or admitted

Table 7.3 Comparison of the performance and the energy consumption of AES and ECC implemented in hard- and software. The software variants have been implemented in native C code

Variant	Duration	Energy consumption
AES (software) 8 MHz/3 V	9,090/19,500 cycles (Encoding/Dec.)	8.5/18.3 µWs
AES (hardware)	78 cycles	16.957 nWs
ECC (software) 8 MHz/3.3 V	82,436 ms	454.3 mWs
ECC (hardware) 33 MHz	180–810 µs	25.1–10.6 Ws

paramedic could access more data than actually required for the current treatment. A role based access control may limit potential privacy risks stemming from exploited authority but in general it is hard if not even impossible to autonomously decide about the need of accessing certain data for treatment. If it was possible to objectively determine the medical necessity to access particular (maybe historic) vital data in the current medical context, we could introduce a context-aware access control approach.

Unfortunately, in the medical domain this is not objectively possible (through an autonomous computing system) without putting the patient at risk. This is still a very important open issue that especially concerns the handling of data access in emergency cases, where it is not reasonable to obstruct the paramedic's or emergency doctor's access to current and historic vital data. In principle free access to the stored and currently measured data should be given in case of danger to life so that acute health problems can be cured by medical or even non-medical helpers. In the simplest form, an emergency situation may be detected if a health parameter exceeds or falls below a pre-configured threshold. Unfortunately that only works fine under certain conditions. For example if the heart rate falls below 10 or exceeds 200 beats per minutes (bpm), it is almost certain that a danger to life exists. However, such simple assumption can e.g. not be taken for monitoring blood pressure that significantly depends on the actual activity of the patient and many other constraints. In other words, even a high blood pressure level may be ok when doing sports but of course not while sleeping. To the best of our knowledge, a highly-reliable and useful monitoring approach that correctly correlates many various vital and physical parameters is still missing.

A means to at least discourage a privacy breaching use by an accessing entity is discussed in Maaser and Ortmann (2010). The system proposed herein logs every data access automatically. The logging includes also rejected access attempts. Therefore, each data access or access attempt is associated with a timestamp and the ID of the accessing entity retrieved from the certificate given in the credentials. In case of complaints of the affected patient, data access e.g. on a data history stored in the sensor node, is documented. The idea is based on using an incremental secure logging functionality as known from other application fields (Sandler et al. 2008; Crosby and Wallach 2009) but the implementation efforts are still too heavy to be used on sensor nodes.

7.5.3 Energy Efficient Protocols

As already mentioned, energy is a scarce resource and thought to be the most important factor when it comes to wireless sensor network lifetime and applicability. Transmitting and receiving are the most power hungry actions on a sensor node. Since most of the time nothing is to be sent or received the coordination of communication is a well-researched field. Most work has been done on Medium Access Control (MAC) protocols. Here the major challenge is to reduce the energy consumed during idle listening. The focus is on allowing sensor nodes to keep the radio powered off as often and as long as possible and to switch it on rarely to send short data frames, dubbed low duty cycle. By doing so, they can achieve a theoretical lifetime of 2 or 3 years with state-of-the-art hardware (Brzozowski et al. 2010c). However, to send data, both the sender and the receiver must be awake at the same time. The problem of synchronizing wake-up time is referred to as "rendezvous" and can be solved in the following ways (Lin et al. 2004):

1. Asynchronous: wake-up radio
2. Pseudo-asynchronous: preamble sampling and similar approaches such as BMAC (Polastre et al. 2004), STEM (Schurgers et al. 2002) and Koala (Liang and Terzis 2008)
3. Synchronous: schedule based wake-ups, e.g. S-MAC (Ye et al. 2002), Dozer (Burri et al. 2007), or DLDC-MAC (see below).

The ideal solution would be a wake up radio, i.e. a transceiver that consumes almost no energy but can be used to indicate a planned transmission. Such solutions are still under investigation.[11] The major advantage of pseudo-asynchronous protocols is their simplicity. The code size is quite small, even ten times smaller than those of synchronous protocols, and they do not suffer from clock drift problems. However, as nodes send a long preamble or plenty of wake-up beacons before data transmission, it may lead to a high collision risk in many WSN scenarios. On the contrary, synchronous approaches deal with the collisions problem well thanks to TDMA schedule. However, such protocols suffer from the clock drift problem, and must precisely synchronize wake-up times of many nodes.

Currently, more stable solutions are based on MAC protocols that rely on a "rendezvous" principle in which sensor nodes agree on a specific time when to wake up for data exchange. Well known approaches to reduce energy consumption on the MAC layer are e.g. STEM (Schurgers et al. 2002), Dozer (Burri et al. 2007), solutions taking energy as parameter for routing decisions have as well been researched HEED (Younis and Fahmy 2004), EECMT (Shemshaki and Shahhoseini 2009), MR-Leach (Farooq et al. 2010). None of these solutions take Quality of Service into account, but monitoring applications such as those to be researched in SAID are requiring QoS e.g. in the sense of minimum latency etc. First steps in this direction have been researched and realized in the FP7 project

[11] www.aet-projekt.de.

WSAN4CIP in which we developed the DLDC-MAC protocol that can be tuned to guarantee certain latency (Brzozowski et al. 2009, 2010a, b, c, 2012). The routing layer can contribute to enhance the lifetime of a sensor network if it takes energy into account when selecting the routes. Here several energy aware routing protocols and metrics have been analysed by Stecklina et al. (2013). In BANs the routing layer is not really needed all nodes should reach each other directly. Therefore we will not provide details on routing here.

Even though many solutions have been proposed already, there are still some basic questions that always need to be considered by the application developer and system designer when it comes to efficient wireless (remote) monitoring applications.

First, at least one node needs to know the gateway to the internet except the gate can become part of the BAN, which might be the most convenient solution.

Second, nodes in a BAN need to learn who else is there and when do they awake or are listening respectively. Therefore coordination of communication activities is needed, which can be executed by some kind of master device (centralized approach) or by the nodes themselves (distributed approach). The latter can be realised by neighbourhood discovery protocols. These are more complex than the centralized approaches indeed, but allows for ease of use and self-set-up means that might become a very important usability issue. Then the sensor nodes stay awake for at least one sleep period and listen on the communication channel to detect all other BAN nodes. Based on that information, they agree on wake-up periods in dependence of the availability of the next hops or the receiving node. In ideal case each sensor node listens at one period during the sleeping time only.

Third, sensor nodes have to cope with overhearing, which means they may receive packets that are not addressed to them and hence processing them is useless and wastes resources not to mention that the communication channel is blocked. This problem does not longer exist if redundant data storage is used in the WSN (BAN) as introduced before. Then all nodes shall get measurement values and store them. However overhearing caused by other BAN's in the vicinity is still an open issue. The currently applied solutions do not fit, since they are meant for cooperating nodes in a common network. A potential approach to solve that may be shifting the sleep cycles in such a way to ensure the own BAN is sleeping while others are communicating, but that is still under research.

7.6 Open Issues

In summary, even with considering all mentioned aspects properly, there is much space left for potential improvements in hardware and software design of body area networks. For example power gating technologies for integrated circuits (Panic et al. 2008) may be facilitated to optimize power consumption e.g. of the micro-controller and the communication module. The focus here lies on minimising the internal power dissipation in times the devices are not used. The software can further be optimized up to certain extend when it is fully adapted to a single hardware platform and a single task rather than using a multi-purpose operating

system. There even exist approaches where the operating system itself implements power management techniques, e.g. the IQLevel OS (Stecklina et al. 2013) where the driver implementations themselves select the deepest possible sleeping mode of components right on a single command.

Looking ahead, additional challenges will arise when the sensor nodes shall be implanted into the human body, which has already been investigated in some selected solutions such as defibrillators (Kao et al. 2010) and glucose monitoring (Fröhlich et al. 2012). However, it is quite obvious that such application has to deal with even more aggravated energy issues Basmer et al (2012) and the form factor of the senor node platform plays a very important role due to missing space inside the human body of course. Even with a fully integrated single chip sensor node, which would then feature smallest dimensions e.g. less than a 2 Euro coin (depending on its capabilities), the overall size will mainly be driven by the power consumption and hence, by the size of the battery. Nevertheless, the domain of implantable devices holds a lot more legal and biological restrictions in terms of device heating and communication. The heating of the device must be limited to around 1 °C during operation. In other words, the device temperature must not exceed 37.5 °C even if the normal body temperature yet is around 36 °C to protect from potential health risks. Also, there exist a couple of legal restrictions in particular for the use of wireless communication. In Germany, wireless communication of implantable devices is restricted to a certain mixed signal band around 403 MHz. This means, you cannot simply put some Bluetooth or any other widely used radio onto such platform, not to mention that communication in the 2.4 GHz band is useless since it will be blocked by water and other body liquids too.

Last but not least, bio stability issues need to be considered. Any kind of implantable device needs to cope with the non-ideal environment (from the technical point of view) of body liquids, blood and natural body defences of course. This may not only lead to processes harming the implanted device itself, e.g. by oxidising or short cuts, but may also cause unwanted reactions repelling the device in the end that can put the patient at serious risks. That does not only concern the sensor nodes but especially holds true for the sensing devices themselves since those most probably need to directly interfere with or contact to the body.

7.7 Conclusions

Wireless sensor nodes are lightweight and autonomous devices that can be applied for individualised sensing and health monitoring. Being used in small networks worn on a person's body, they enable the implementation of various remote health care applications. However, a lot of requirements have to be considered in WSN system design for e-health use. First, the very basic network architecture has to be determined. Centralised architectures are easier to implement but provide less flexibility and often feature single points of failures. In contrast to that, centralized architectures—in which all sensor nodes provide same or similar capabilities—are

more reliable and fail-safe while in addition providing various self-X properties such as self-organisation, self-configuration and self-healing.

These mentioned basic considerations may be sufficient to make the application running properly, but nevertheless many more design challenges have to be included for e-health use of WSN. In such application area the running system needs to meet all technical and also legal demands of the health domain. That raises further critical design challenges in particular in view of reliability, security and privacy of the system. In addition, the holy grail for WSN design is the energy efficiency of the system since the whole set of features, the size and the weight of the sensor nodes and hence, the usability of the system depend on it. Especially if the latter is not sufficiently given, even a good working system may not be used by patients and clinicians. In this chapter we have presented several selected solutions that successfully tackle the introduced challenges. Finally, we are also aware of the fact that there still exist a lot of open issues and much space for future research as we have discussed at the end of the chapter.

References

Abdelzaher TF et al (2004) EnviroTrack: towards an environmental computing paradigm for distributed sensor networks. In: Proceedings of the 24th international conference on distributed computing systems (ICDCS 04), IEEE CS Press, pp 582–589

Aboelaze M, Aloul F (2005) Current and future trends in sensor networks: a survey. In: Proceedings of second IFIP international conference on wireless and optical communications networks WOCN, Dubai, pp 551–555

Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E (2002) Wireless sensor networks: a survey. Comput Netw 38:393–422

Basmer T, Genschow G, Fröhlich M, Birkholz M (2012) Energy budget of an implantable glucose measuring system. Biomed Technol 57:259–262

Birkholz M et al (2009) Mikroviskosimeter zur kontinuierlichen Glucosemessung bei Diabetis mellitus. In: Proceedings of Mikrosystemtechnik-Kongress. VDE-Verlag, Berlin, 124p

Bohn J, Coroama V, Langheinrich M, Mattern F, Rohs M (2004) Living in a world of smart everyday objects—social, economic, and ethical implications. J Hum Ecol Risk Assess 10:763–786

Brzozowski M, Salomon H, Langendoerfer P (2009) Completely distributed low duty cycle communication for long-living sensor networks. In: Proceedings of the 2009 international conference on computational science and engineering (CSE '09), vol 2. IEEE Computer Society, Washington, DC, pp 109–116

Brzozowski M, Salomon H, Langendoerfer P (2010a) On efficient clock drift prediction means and their applicability to IEEE 802.15.4. In: Proceedings of the 2010 IEEE/IFIP international conference on embedded and ubiquitous computing (EUC '10). IEEE Computer Society, Washington, DC, pp 216–223

Brzozowski M, Salomon H, Langendoerfer P (2010b) ILA: idle listening avoidance in scheduled wireless sensor networks. In: Osipov E, Kassler A, Bohnert TM, Masip-Bruin X (eds) Proceedings of the eighth international conference on wired/wireless internet communications (WWIC'10). Springer, Berlin, pp 363–374

Brzozowski M, Salomon H, Langendoerfer P (2010c) Limiting end-to-end delays in long-lasting sensor networks. In: Proceedings of the eighth ACM international workshop on mobility management and wireless access (MobiWac '10). ACM, New York, NY, pp 11–20

Brzozowski M, Salomon H, Langendoerfer P (2012) Support for a long lifetime and short end-to-end delays with TDMA protocols in sensor networks. Paper presented at the International Journal of Distributed Sensor Networks, Article ID 651748

Burns A et al (2010) SHIMMER™—a wireless sensor platform for noninvasive biomedical research. IEEE Sens J 10(9):1527–1534

Burri N, Rickenbach P, Wattenhofer RR (2007) Dozer: ultra-low power data gathering in sensor networks. In: Proceedings of the sixth international conference on Information processing in sensor networks (IPSN '07). ACM, New York, NY, pp 450–459

Chang N, Zhang QY, Cungang Y (2007) A lightweight security protocol for wireless sensor networks. In: Proceedings of the international workshop on telecommunications—IWT/07

Chen T, Mazomenos E, Maharatna K, Dasmahapatra S, Niranjan M (2012) On the trade-off of accuracy and computational complexity for classifying normal and abnormal ECG in remote CVD monitoring systems. In: Proceeding of the IEEE workshop on signal processing systems (SiPS), Quebec, pp 37–42

ATMEL Corporation (2010a) AT91SAM9G20 preliminary. http://www.atmel.com/dyn/resources/prod documents/doc6384.pdf. Accessed May 2013

ATMEL Corporation (2010b) ATmega640/1280/1281/2560/2561 preliminary. http://www.atmel.com/dyn/resources/proddocuments/doc2549.pdf. Accessed May 2013

ATMEL Corporation (2011) ATmega128(L). http://www.atmel.com/dyn/resources/proddocuments/doc2467.pdf. Accessed May 2013

Costa P, Mottola L, Murphy AL, Picco GP (2007) Programming wireless sensor networks with the TeenyLIME middleware. In Proceedings of the eighth ACM/IFIP/USENIX international middleware conference (Middleware 2007), Newport Beach, CA, pp 26–30

Crosby SA, Wallach DS (2009) Efficient data structures for tamper evident logging. In: Proceedings of the 18th USENIX security symposium, Montreal, CA

Dikaiakos MD, Florides A, Nadeem T, Iftode L (2007) Location-aware services over vehicular ad-hoc networks using car-to-car communication. IEEE J Sel Areas Commun 25(8):1590–1602

Dyka Z, Langendoerfer P (2005) Area efficient hardware implementation of elliptic curve cryptography by iteratively applying Karatsuba's method. In: Proceedings of design, automation and test in Europe, IEEE

Edwards J (2012) Wireless sensors relay medical insight to patients and caregivers [special reports]. IEEE Signal Process Mag 3(29):8–12

Farooq MO, Dogar AB, Shah GA (2010) MR-LEACH: multi-hop routing with low energy adaptive clustering hierarchy. In Proceedings of the 2010 Fourth international conference on sensor technologies and applications (SENSORCOMM '10). IEEE Computer Society, Washington, DC, pp 262–268

Fröhlich M et al (2012) Biostability of an implantable glucose sensor chip. IOP Conf Ser Mater Sci Eng 41:012022

Gellersen HW, Schmidt A, Beigl M (2000) Adding some smartness to devices and everyday things. In: Proceedings of third IEEE workshop on mobile computing systems and applications, Monterey, CA, pp 3–10

Gillies D, Thornley DJ, Bisdikian C (2009) Probabilistic approaches to estimating the quality of information in military sensor networks. Comput J

Girao J, Westhoff D, Mykletun E, Araki T (2007) Tinypeds: tiny persistent encrypted data storage in asynchronous wireless sensor networks. Ad Hoc Netw J 5(7):1073–1089

Gummadi R, Gnawali O, Govindan R (2005) Macro-programming wireless sensor networks using kairos. In: Proceedings of the international conference on distributed computing in sensor systems (DCOSS 05), LNCS 3560, Springer, Heidelberg, pp 126–140

Halperin D et al. (2008) Pacemakers and implantable cardiac defibrillators: software radio attacks and zero-power defenses. In: Proceedings of IEEE symposium on security and privacy, Washington, DC, pp 129–142

Heuschmann PU et al (2010) Schlaganfallhäufigkeit und Versorgung von Schlaganfallpatienten in Deutschland. Aktuelle Neurol 7:333–340

Hoon CL (2009) Practical broadcast authentication using short-lived signatures in WSNs. Information security applications, Lecture Notes in Computer Science 5932:366–383

Kahn JM, Katz RH, Pister KSJ (1999) Next century challenges: mobile networking for "smart dust". In: Proceedings of international conference on mobile computing and networking (MOBICOM), Seattle, pp 271–278

Kahn JM, Katz RH, Pister KSJ (2000) Emerging challenges: mobile networking for smart dust. J Commun Netw 2(3):188–196

Kao CW, Friedmann E, Thomas SA (2010) Quality of life predicts one-year survival in patients with implantable cardioverter defibrillators. Qual Life Res 19(3):307–315

Karlof C, Sastry N, Wagner D (2004) TinySec: a link layer security architecture for wireless sensor networks. In: Proceedings of the second international conference on embedded networked sensor systems, Baltimore, MD, Nov 2004, pp 162–175

Kirchhof P, Adamou A, Knight E, Lip GYH, Norrving B, de Pouvourville G (2009) How can we avoid a stroke crisis? ISBN 978-1-903539-09-5

Lab-on-Chip-System (2011) Lab-on-Chip-System. Medizin&Technik, http://www.medizin-und-technik.de/home/-/article/27544623/35089533?returnToFullPageURL=back. Accessed May 2013

Liang CJM, Terzis A (2008) Koala: ultra-low power data retrieval in wireless sensor networks. In: Proceedings of the seventh international conference on Information processing in sensor networks. IEEE Computer Society

Lin E-YA, Rabaey JM, Wolisz A (2004) Power-efficient rendez-vous schemes for dense wireless sensor networks. In: IEEE international conference on communications, vol 7, Enschede

Maaser M, Ortmann S (2010) Remote medical treatment at home using the using the Java mobile sensor API. In: Proceedings of the third international workshop on smart homes for tele-health (SmartTel '10), Miami, USA

Madden SR, Franklin MJ, Hellerstein JM, Hong W (2005) TinyDB: an acquisitional query processing system for sensor networks. ACM Trans Database Syst 30(1):122–173

Mainwaring A, Culler D, PolastreJ, Szewczyk R, Anderson J (2002) Wireless sensor networks for habitat monitoring. In: Proceedings of the first ACM international workshop on wireless sensor networks and applications (WSNA '02). ACM, New York, NY, pp 88–97

Marvel (2010) PXA27x Specification update. http://www.marvell.com/products/processors/applications/pxafamily/pxa27x emts.pdf. Accessed April 2013

Morchon OG, Falck T, Heer T, Wehrle K (2009) Security for pervasive medical sensor networks. In: Proceedings of sixth annual international conference on mobile and ubiquitous systems (MobiQuitous 2009), Toronto, Canada

Naqvi S, Dallons G, Michot A, Ponsard C (2010) Assuring privacy of medical records in an open collaborative environment—a case study of Walloon Region™ eHealth platform. Privacy and identity management for life, vol 320. Springer, Berlin, pp 146–159. ISBN:978-3-642-14281-9. http://dx.doi.org/10.1007/978-3-642-14282-6_12

Ortmann S, Maaser M (2011) Enabling secure and privacy-aware mobile sensing and e-health applications on everybodys smartphone. In: First IEEE international conference on consumer electronics-Berlin (ICCE-Berlin 2011), Berlin

Ortmann S, Maaser M, Parandian B, Schultz M (2011) Telemedizinisch assistierte ambulante Betreuung von Patienten, vol 4. Deutscher AAL-Kongress, Berlin

Ortmann S, Langendoerfer P, Sik-Lanyi C (2012) Telemedical assistance for ambulant rehabilitation of stroke patients. In: Proceedings of the ninth world congress on brain injury, Edinburgh, Scotland

Panic G, Dietterle D, Stamenkovic Z (2008) Architecture of a power-gated wireless sensor node. In: Proceedings of 11th EUROMICRO conference on digital system design architectures, methods and tools (DSD'08), IEEE, Parma

Panic G, Basmer T, Schrape O, Peter S, Vater F, Tittelbach-Helmrich K (2011) Sensor node processor for security applications. In: Proceedings of the 18th IEEE international conference on electronics, circuits and systems (ICECS 2011), Beirut

Park T, Shin KG (2004) LiSP: a lightweight security protocol for wireless sensor networks. ACM Trans Embed Comput Syst 3(3):634–660

Patel et al (2012) A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 9(12):1–17

Perrig A, Szewczyk R, Wen V, Culler D and Tygar JD (2001) SPINS: Security protocols for sensor networks. In: Proceedings of seventh annual international conference on mobile computing and networking, Rome, Italy, Aug 2001, pp 188–189

Piotrowski K, Langendoerfer P, Peter S (2009) tinyDSM: a highly reliable cooperative data storage for wireless sensor networks. In: Proceedings of the 2009 international symposium on collaborative technologies and systems (CTS '09), Washington, DC, pp 225–232

Piotrowski K, Sojka A, Langendoerfer P (2010) Body area network for first responders: a case study. In: Proceedings of the fifth international conference on body area networks (BodyNets '10), ACM, New York, NY, pp 37–40. doi:10.1145/2221924.2221933

Polastre J, Hill J, Culler D (2004) Versatile low power media access for wireless sensor networks. In: Proceedings of the second international conference on embedded networked sensor systems. ACM, New York

Robinson P, Beigl M (2003) Trust context spaces: an infrastructure for pervasive security in context-aware environments. In: Proceedings of first international conference of security in pervasive computing. Springer, Berlin, pp 157–172

Sandler D, Derr K, Crosby S, Wallach DS (2008) Finding the evidence in Tamper-Evident logs. In: Proceedings of third international workshop on systematic approaches to digital forensic engineering (SADFE '08), Berkeley, pp 69–75

Scheffler T, Schindler S, Lewerenz M, Schnor B (2011) A privacy-aware localization service for healthcare environments. In: Proceedings of the fourth international conference on pervasive technologies related to assistive environments (PETRA '11). ACM, New York

Schurgers C, Tsiatsis V, Ganeriwal S, Srivastava M (2002) Optimizing sensor networks in the energy-latency-density design space. IEEE Trans Mob Comput 1:70–80

Shaikh, R. A., Lee, S., Khan, M. A., & Song, Y. J. (2006). LSec: Lightweight security protocol for distributed wireless sensor network. In Personal Wireless Communications (pp. 367–377). Springer Berlin Heidelberg.

Shemshaki M, Shahhoseini HS (2009) Energy efficient clustering algorithm with multi-hop transmission. In: Proceedings of the 2009 international conference on scalable computing and communications; eighth international conference on embedded computing (SCALCOM-EMBEDDEDCOM '09). IEEE Computer Society, Washington, DC, pp 459–462

Sojka A, Piotrowski K, Langendoerfer P (2010) ShortECC: a lightweight security approach for wireless sensor networks. In: Proceedings of INSTICC international conference on security and cryptography, SECRYPT

Stecklina O, Langendoerfer P, Goltz C (2013) A fair energy trade multi-hop routing in wireless sensor networks. In: Proceedings of the sixth joint IFIP wireless & mobile networking conference (WMNC2013), Dubai

Sun T, Chen LJ, Han CC, Gerla M (2005) Reliable sensor networks for planet exploration. In: Chen LJ (ed) Proceedings of IEEE networking, sensing and control, Tucson, USA, pp 816–821

Texas Instruments Inc. (2007a) 2.4 GHz IEEE 802.15.4 / ZigBee-Ready RF Transceiver. http://www.ti.com/lit/gpn/cc2420. Accessed May 2013

Texas Instruments Inc. (2007b) 2.4 GHz IEEE 802.15.4/ZIGBEE RF TRANSCEIVER. http://www.ti.com/lit/gpn/cc2520. Accessed May 2013

Texas Instruments Inc. (2007c) Single-chip very low power RF transceiver. http://www.ti.com/lit/gpn/cc1000. Accessed May 2013

Texas Instruments Inc. (2009) Low-cost low-power 2.4 GHz RF transceiver. http://www.ti.com/lit/gpn/cc2500. Accessed May 2013

Texas Instruments Inc. (2010a) CC1101 low-power Sub-1 GHz RF transceiver. http://www.ti.com/lit/gpn/cc1101. Accessed May 2013

Texas Instruments Inc. (2010b) CC430F613x, CC430F612x, CC430F513x MSP430 SoC with RF core. http://www.ti.com/lit/gpn/cc430f6137. Accessed May 2013

Texas Instruments Inc. (2010c) MSP430F543xA, MSP430F541xA Mixed signal microcontroller. http://www.ti.com/lit/gpn/msp430f5438a. Accessed May 2013

Texas Instruments Inc. (2011) MSP430F15x, MSP430F16x, MSP430F161x mixed signal microcontroller. http://www.ti.com/lit/gpn/msp430f1611. Accessed May 2013

Ugus O et al (2009) Optimized implementation of elliptic curve based additive homomorphic encryption for wireless sensor networks. arXiv preprint arXiv:0903.3900

Uhsadel L, Poschmann A, Paar C (2007) Enabling full-size public-key algorithms on 8-bit sensor nodes. Security and privacy in ad-hoc and sensor networks. Springer, Berlin, pp 73–86

Wei Y, Heidemann J, Estrin D (2002) An energy-efficient MAC protocol for wireless sensor networks. In: Proceedings of the twenty-first annual joint conference of the ieee computer and communications societies (INFOCOM 2002), vol 3, IEEE, New York

Weiser M (1991) The computer for the twenty-first century. Sci Am 265(3):66–75

Werner-Allen G, Johnson J, Ruiz M, Lees J, Welsh M (2005) Monitoring volcanic eruptions with a wireless sensor network. In: Proceedings of the second European workshop on wireless sensor networks, Istanbul, pp 108–120

Westin AF (1967) Privacy and freedom. Atheneum, New York

Yao Y, Gehrke JE (2002) The cougar approach to in-network query processing in sensor networks. ACM Sigmod Rec 31(2):9–18

Younis O, Fahmy S (2004) HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans Mob Comput 03(4):366–379

Chapter 8
System Integration Issues for Next-Generation Remote Healthcare System

Cristiano Paggetti, Carlos Cavero Barca, and Juan Mario Rodríguez

Abstract The system integration is always a terrible headache for IT technologists. Several aspects are related to a proper integration of different components and services into remote healthcare solution. Furthermore the overall integration issue cannot be regarded only from a technical point of view but it has to take into account aspects such as deployment scenario, service organization, educational and business context, resource sharing with other services. The system integration is a crucial activity and requires to be properly planned, it is based on system and service architecture design however it must be empowered taking into account use case and deployment scenario, functional and technical specification and interoperability requirements with other services. Due to the variety and complexity of system integration, in this chapter only some of the major issues related to system integration are taken into account; in particular the authors have selected the following main issues: system integration topics checklist; the interoperability and portability of data as one of the crucial aspects enabling system integration and proper deployment of solutions into the healthcare domain; structured approach for solution deployment; the user interface design as basic aspect to engage the medical professionals. Finally the critical issues are raised breaking down the lessons learnt.

8.1 Introduction

As discussed in previous chapters remote healthcare systems have been effectively employed in a number of contexts with very different and heterogeneous stakeholders. For each remote healthcare service scenario it is required a dedicated

C. Paggetti (✉)
Connected Health Unit, I+ S.r.l., Florence, Italy
e-mail: c.paggetti@i-piu.it

C.C. Barca • J.M. Rodríguez
ATOS Research and Innovation (ARI), ATOS, Madrid, Spain
e-mail: carlos.cavero@atos.net; juan.rodriguezp@atos.net

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_8,

analysis of different main crucial aspects such as information workflow, service models and business scenario; these aspects will compose the real context of system integration and solution deployment activities. Furthermore for a successful integration, some general features of the solution are really relevant, for instance:

- Conformance to the scenarios defined in the service requirements;
- Modular and open standard approach to support interoperability and scalability of the system;
- Interoperability among the system modules and data exchange interoperability between the system and existing products or services;
- Secure access/exchange of information, privacy issues;
- Compliance with solution intended use related standards and directives;
- Reliability of the system based on detailed risk analysis
- Ease of use and structured training and learning curve.

8.2 Main Integration Criteria

A remote healthcare system can be composed by different components and services, hereafter are listed some of the main components commonly integrated in a Remote Healthcare solution:

– Virtual Individual Model
– Middleware platform for Remote Healthcare
– Personal Sensing Platforms
– Personal Service Profiling engine
– Health Risk Predictions engine
– Cloud-based Data Management
– Workflow modelling and Orchestration Engine

Each of the above components is based on different building blocks and the interaction among such components must be properly analysed in order to reach a successful integration. In particular hereafter is reported a non-exhaustive checklist of activities to be undertaken to run a system integration, some of them can be also related to single components development, but the Integration Manager has to take into account them as well:

- Analyse the specific concept of the Remote Healthcare application;
- Identify diseases/health conditions addressed by the service and related indicators
- Identify and define specific physical, mental, and social activities that can be monitored and modified to prevent the above mentioned diseases/health conditions
- Identify psychosocial and human factors affecting the potential acceptance of the service
- Review trends in technology adoption and use of existing services

- Identify the appropriate Remote Healthcare solutions available on the market and perform competitor analysis
- Identify the users' group technological characteristics
- Identify external IT services and components to be integrated with Remote Healthcare service
- Crosscheck, on the basis of above elements, system configuration, communication modalities towards users and caregivers
- Based on unit test and integration test, evaluate the performance of each components and the overall integrated system
- Identify major integration risks and related contingency plans
- Elicit the main legal and ethical issues pertaining to the Remote Healthcare service
- Provide specific references concerning the legal requirements related to the Remote Healthcare service
- Describe relevant aspects of data protection and privacy laws
- Identify certification requirements based on the specific intended use of the Remote Healthcare solution
- Establish robust positions on patents, copyright and design rights to protect IP and to minimise the risk of plagiarism and loss of commercialization revenue
- Define an iterative integration and deployment process to guarantee all requirements are met
- Verify the requirements for solution deployment, based on technical and operational services and support to marketing services

8.3 Data Interoperability

Interoperability among systems in terms of data exchange is one of the major issues related to successful system integration and deployment in real settings. Healthcare Information Technologies compared to other Information Technologies sectors lags behind in terms of integration and interoperability adopted standards and guidelines. Most of the time the implementation of Remote Healthcare service is discussed in terms of technological components and information sharing issues, However for a successful solution implementation, it is crucial to work on service scenarios focused on patient-centred strategies, putting patients first and supported by an healthcare ecosystem able to scale and compose different levels of services. Openness and transparency are needed to change the paradigm, the patient-centric approach must be based on the interoperability between systems sharing multi-source data of the individuals, all the medical specialist should have access to information anytime and anywhere (Main and Slywotzky 2012). Openness and transparency can be achieved with the use of EHR standards, but also implementing open source health solutions (Murray et al. 2009). In this subsection interoperability will be addressed mainly in relation to exchange of meaningful information.

The IEEE (2007) defines interoperability as the "Ability of a system or a product to work with other systems or products without special effort on the part of the

customer" (Iakovidis et al. 2007) and considers the interoperability in healthcare "as the ability of different information technology systems and software applications to communicate, to exchange data accurately, effectively and consistently and to use the information that has been exchanged. Making healthcare information systems interoperable will reduce the cost of health care and will contribute to more effective and efficient patient care".

The exchange of information involves semantic aspects, in particular for the Semantic Health study (2009), the semantic interoperability (SIOp) addresses issues on how to best exchange seamlessly the information between health services providers and patients. Its geographic scope ranges from local interoperability (within, e.g., hospitals or hospital networks) to regional, national and cross-border interoperability. Different levels of interoperability can be identified as follows:

- Level 0: no interoperability at all
- Level 1: technical and syntactical interoperability (no semantic interoperability)
- Level 2: two orthogonal levels of partial semantic interoperability
- Level 2a: unidirectional semantic interoperability
- Level 2b: bidirectional semantic interoperability of meaningful fragments
- Level 3: full semantic interoperability, sharable context, seamless co-operability

Those interoperability levels are associated by Blobel (2010) describing the syntactic level as the messages, clinical documents and agreed vocabulary and the semantic level as the advanced messaging, common information models (standardization and terminologies).

In the domain of the organization/service, interoperability is defined as the common business deployment of all the aforementioned characteristics which is similar to the "process interoperability" mentioned by Duftschmid et al. (2010) describing the methods for optimal integration of computer systems into actual work settings.

The process interoperability is essential to reach the cross-organizational EHR communication and succeeds in the volume to value revolution and patient-centric approach which will build the future healthcare services.

According to ISO/TR 20514 (2005), there are four prerequisites for EHR semantic interoperability:

(i) Agreement on a standardized reference model.
(ii) Standardized service interface models to provide interoperability between the health services and other services such as demographics, terminology, access control and security services.
(iii) A standardized set of domain-specific concept models, e.g. archetypes and templates for clinical, demographic, and other domain specific concepts.
(iv) Standardized terminologies associated with controlled vocabularies (European Commission—DG Information Society/Q-REC 2007).

The strength of the dual model approach is twofold:

1. In the Reference Model (RM) is its static and stable implementation serving as a solid basis for an EHR architecture, but at the same time,
2. Allows to include new clinical knowledge using the archetypes modelled by the doctors and constrained to the RM. Actually the RM permits the syntactic interoperability (level 1 or human-readable format) between the systems whilst the archetypes provide the semantic interoperability (level 2 or machine-readable format).

To reach the archetype based interoperability, the healthcare systems need to comply and find the existing archetypes available based on the clinical concepts needed; compose templates from archetypes and retrieve corresponding data from the underlying medical information systems.

European countries have been giving special importance to the standards during the last decade. Between the evolution of eHealth activities and strategies (Stroetman et al. 2011) the most important national initiatives are focused on international communication (one of the reference examples is epSOS (http.// www.epsos.eu) project): sharable EHR patient summary, ePrescription, patient Id and standards.

Some initiatives at a regional level have already been deployed in Sweden (Chen et al. 2009), Denmark (Bernstein et al. 2009), UK (McCay et al. 2008), Slovenia (Think!Med Clinical 2013), Austria and Norway (Stroetman et al. 2011), with full archetype-based development. At a vendor level they continued to use their own software interfaces but linking them to the dual model solution, so describe mainly the clinical knowledge modelled and the relations with the system. Other clinical document architecture (CDA) archetype-based tools have been deployed in Japan focused on the referral letter (Yong et al. 2008) and Turkey which benefited of using XML instead of rules for the validation of exchanged medical data (Kabak et al. 2008).

The dual model approach is being used in Australia, The Netherlands, Brazil, Portugal, Russia and Slovenia with deployments of openEHR solutions into production healthcare (openEHR Healthcare Providers and Authorities 2012). It is also used in parts of the UK NHS Connecting for Health Programme and it has also been selected as the basis for the national program in Sweden. It is under evaluation in a number of countries which already have deployed tools at a regional level including Denmark, Slovakia, Sweden, Chile and Brazil (openEHR Wikipedia 2011). Recently Russia has started a large project aiming at centralising EHR for more than 11 million Moscow inhabitants (Marand 2013). It is beginning to be utilised in commercial systems throughout the world. The NHS Connecting for Health initiative defined the following steps to be followed to deploy health systems providing full semantic interoperability applicable to other systems:

(a) *Knowledge level transformation*

Publish EHR content models as archetypes and templates using the openEHR and EN/ISO 13606 reference models and data type specifications and as instances of HL7 Clinical Document Architecture Release Two (CDA R2 and possibly HL7 templates).

(b) *Semantically equivalent instance transformation*

Semantically equivalent instance transformation therefore requires that the domain of possible instances corresponding to any given clinical data structure can all be represented faithfully by each kind of reference/instance model in use.

(c) *Full interoperability transformation*

This kind of transformation requires that all of the clinical and medico-legal information included within an EHR instance in one representation can be transformed into another (McCay et al. 2008).

8.4 Data Portability

Portability is "the quality of being portable" or using other words "the ability to be carried or moved easily" (Collins 2013). In the ICT world and more specifically in the healthcare environment what it should be carried easily is the information but also the solutions to manage the data in a standardised way. Two meanings of portability should be taken into consideration in the present section: Software and data portability.

Software portability as stated in the wiki (Wikipedia 2013) is.

> "*Portability in high-level computer programming is the usability of the same software in different environments. The pre-requirement for portability is the generalized abstraction between the application logic and system interfaces. When software with the same functionality is produced for several computing platforms, portability is the key issue for development cost reduction.*"

Data portability in healthcare environments is the process by which the electronic health record can in whole or in part be exchanged between two different systems. On one hand healthcare interoperable solutions should be easily exported which means the capability of being installed and integrated with other hospital information systems. On the other hand the data should be accessible anywhere at any time. The implementation and deployment of wrap-up solutions standard-based would also provide the capability of exporting the system to other countries.

If the stakeholders and medical professionals can access shared data the reduction of costs would be considerable avoiding the duplicity of the information, maintenance and upgrades.

Information is flowing from one side to another. The use of portable devices changed the process humans exchange their personal data (Main and Slywotzky 2012). Each byte is stored locally but also through the internet and it should be seamlessly transferred in a secure way. But the information security is crucial to protect the person privacy.

The health solutions are always dealing with sensitive portable and interoperable data, thus security must be available on the platform. Scalability, modularity and transparency will facilitate the implementation of the protection mechanisms, because at the end, security must not materially impact the performance of the system.

In a service oriented environment the information should be protected using the following concepts described in the next sections (OASIS 2008): confidentiality, integrity, triple A (Authentication, Authorization and Accounting), availability but also privacy, other concept such as data fidelity are really relevant to refer to data set that have been acquired and stored according to specific protocol. Least but not last the data protection framework for Electronic Health Record (EHR) provides guidelines to protect the patient information.

8.4.1 AAA (Authentication, Authorization and Accounting)

Authentication is the way the identity of the participants in a process is verified by means of credentials in order to assure the credibility of the entity.

Authorization legitimates the entity in the interaction with the system receiving permissions from the owner of the data or resource to act accordingly.

Accounting is the capability associated with resources that allows tracking the tasks and measures applied to them.

AAA (Authentication, Authorization, and Accounting) is the abbreviation for denoting the framework for intelligently covering the control to access the resources, policy enforcement and audit trail.

8.4.2 Availability

Availability is the ability of ensuring that the services, channels, resources and other process work properly and the information is accessible in a trustworthy way. Availability must prevent against the so-called Denial of Service attack which aims at tampering the system to make resources unreachable.

8.4.3 Confidentiality

Confidentiality ensures that information cannot be accessed by unauthorised entities not allowing reading messages or parts of messages that are transmitted (Niski 2008). The sensitive data managed in the healthcare environment oblige to apply high degrees of confidentiality: third parties are not aware when an exchange happened. Encryption algorithms ensure confidentiality of the patient information and only the authorised entities can access the data (Henry 2009).

8.4.4 Integrity

Integrity refers to the protection of information that is exchanged—either by malicious agents or accident. Integrity deals with the mechanisms to provide assurance that information has not been altered from its original form (Niski 2008).

The message digest permits the application of hash algorithms to the content in order to provide verification methods in the receiving application to compare the digest generated by the transmitting entity. If one byte has changed the hash algorithm would recreate a different digest value.

8.4.5 Privacy

Privacy concerns with the proper management of personal information. Privacy is not secrecy or hiding information, it is about how to treat and ensure that people manage the disclosed information appropriately. Privacy is fundamentally contextual because it depends on the sensitive data handled and protection of the dignity of the human being (Glazer and Blakley 2009).

We can define privacy just citing the Hippocratic Oath:

> "What I may see or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself, holding such things shameful to be spoken about. If I fulfil this oath and do not violate it, may it be granted to me to enjoy life and art, being honoured with fame among all men for all time to come; if I transgress it and swear falsely, may the opposite of all this be my lot" (Edelstein 1943)

8.5 The Data Protection Framework for Electronic Health Record (EHR)

The patient information stored in the EHR systems has to provide the corresponding protection mechanisms of personal data. The Article 29 Data Protection Working Party set out in Recital 2 of the Directive says that "data processing systems are designed to serve man; (...) they must, whatever the nationality or residence of natural persons, respect the fundamental rights and freedoms, in particular the right to privacy, and contribute to economic and social progress, trade expansion and the well-being of individuals" (European Commission 2007).

It should also pay attention to the obligations to comply with the security regulations of the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which require the system in terms of integrity, confidentiality and availability.

8.6 Model Driven Architecture (MDA)

MDA (Model Driven Architecture) is an approach for automatically developing software focused on a well-defined design by means of clear specifications and guidelines for software implementation.

As defined by the Object Management Group (OMG) (2013), "*MDA is a way to organize and manage enterprise architectures supported by automated tools and services for both defining the models and facilitating transformations between different model types*"

MDA bridges the gap between analysis and implementations by achieving high levels of re-use (portability) and reduce the development effort just with the design of model diagrams and interfaces. MDA is related to multiple technologies as Unified Modelling Language (UML), the Meta-Object Facility (MOF) and XML Metadata Interchange (XMI) among others.

8.7 User Interface Design

Taking into account the patient-centric approach, adopting a single and open interoperability framework for the specification of user interfaces and application seems logical. Interoperable "User Interface" clinical documents, archetypes and templates imply that nodes share a commonly agreed granularity and terminology. For establishing equivalence between nodes that express the same meaning in different contexts and different granularities the use of common information model, data types, structures and a common terminology to share the clinical concepts at the level of user interface specification, can ease the mapping to products already stored in the hospital.

The idea is to construct dynamic mechanisms to implement the outputs in a document or message clinical statement specification, using an editor tool that allows detailed archetypes, templates or clinical documents to be defined, at a suitable granularity for user interface design. The available tools include the connection to a terminology server to guide the constraint author to the concepts that could usefully be separately displayed, collected, or constrained, but also they would also include a mapping between archetypes. User interface clinical models (templates) need to have data points that correspond to each field on a document, report, or data entry form, whereas for interoperability archetypes need to have a set of coherent and attributable record units (clinical statements or record components) that correspond in granularity to that of the terminology in use (McCay et al. 2008).

Expressivity of the standards should be taken into account. The *user-friendliness* is another main factor, which we interpret as the suitability of the standard's conceptual model for representing clinical data (e.g. observations, problems, procedures, or clinical assessments and recommendations for care plans). *Extendibility* should be considered in order to enlarge the clinical knowledge straightforwardly. The reference models mentioned provide a small set of classes with attributes that are clearly defined and new clinical concepts can be easily included using archetypes created from scratch for a specific purpose or reuse those available in the openEHR Clinical Knowledge Manager. Other important aspects are *flexibility and transparency* and the two-level modeling approach, where the storage can be selected independently of the high-level archetypes representation, which makes it more flexible by separating both layers.

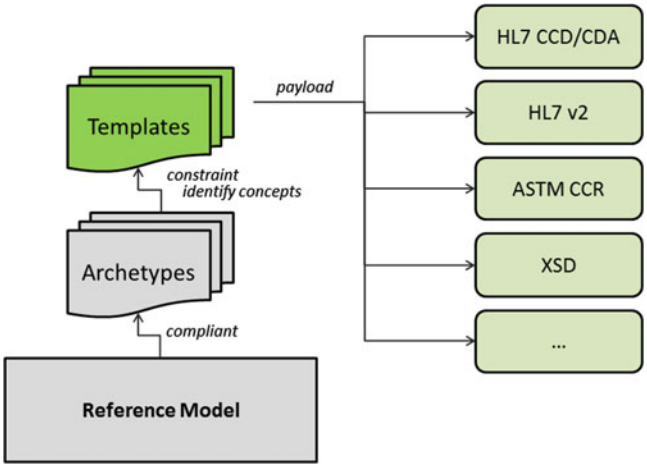


Fig. 8.1 Automatic generation of dual model templates exported to data schemas such as hl7 CDA, ASTM CCR and other message formats automatically

From a vendor point of view, the hospital interfaces used should empower the vision for commercialization and standardization. Information in a particular EMR must be easily mapped to the reference model chosen without effort. The system should support the exchanging of information such as HL7 v2.x, v3, CCD/CDA, ASTM CCR messages because they are the most used interfaces (providing also the transformation of CCD/CDA messages) (Fig. 8.1) (openEHR Foundation 2010). Providing interface transformation would ease the deployment in the hospitals environment connecting and mapping the concepts to any terminology needed and covering multiple standards. There are validated archetypes in the CKM, Clinical Knowledge Manager from openEHR which is a web-based repository allowing for archetype search, browse and download to visualize those already stored (openEHR Foundation 2012). Other repositories available are the NEHTA Clinical Knowledge Repository (NEHTA CKM 2012) and the Swedish Clinical Knowledge Repository (Swedish openEHR CKM 2010):

- **Ocean Archetype Editor**, version 2.2 Beta Release (Ocean Archetype Editor 2012)
- **linkEHR** (2011), an archetype editor not only for openEHR but also for EN 13606 archetypes, hl7, etc.
- **LiU Archetype Editor**, version 0.5.2 (LiU Archetype Editor 2007)

The information is distributed in compositions created for the EHR of the patient. The entries are stored and retrieved from the composition and possibly into sections (Fig. 8.2). See Fig. 8 extracted from Beale et al. (2007).

A methodology on how to build an archetype is outlined by Leslie and Heard (Leslie and Heard 2006). The steps to be followed to properly design the clinical concept are the following:

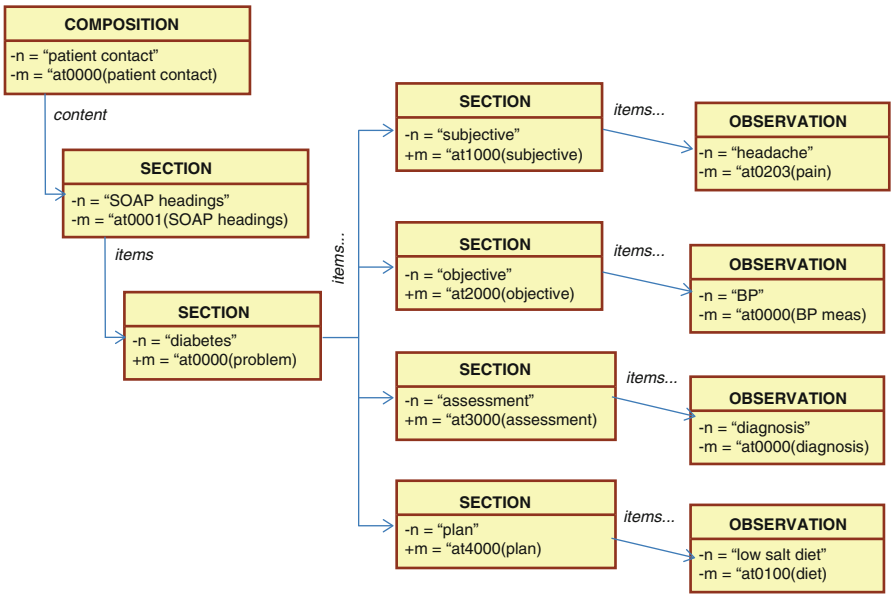


Fig. 8.2 Compositions and entries

- Identify the clinical concepts. Identifying the single and the multiple terms in order to avoid errors or overlapping.
- Identify existing archetypes. Check the repositories abovementioned and select the best candidates. Don't waste your time creating models already available and published by experts. Before publication, archetypes undergo an iterative review process to ensure that they cover as many use-cases as possible and thus constitute a sensible maximal data set.
- Only if needed, create new archetypes. The procedure for archetype creation is: Gather and organize the content (purpose, elements and terminology bindings); choose the type of archetype (composition, entry and so on); Build the archetype naming the archetype, selecting the suitable structure and data elements and types and adding the terminology; finally publish the archetype (McCay et al. 2008).

One of the targets of the EN 13606 is to exchange part or all of a health record between different manufacturer's systems, with maximum safety and without loss of meaning or of any original context. The modeling and integration of clinical information has always been very complex. Everything seems connected and inter related when we need to describe a person's health state. The way to deal with such complexity is to consider parts of a whole and to apply a delicate integration of all the parts maintaining the original clinical meaning. The Detailed Clinical Models (DCM), such as the 13606 archetypes and open EHR and the statements of the HL7 are created taking into account these problems and build upon the idea of two level modeling for the electronic records.

The data exchange interfaces between the components on a Healthcare architecture can be provided by means of a SOA (Service Oriented Architecture) using either Web Services (SOAP and WSDL) or RESTful services. For data interchange the recommendation is to use XML or JSON as a common and widespread markup language. The user interfaces will be guided by the reference model (either hl7 RIM or ISO/EN 13606) and the clinical documents or templates edited and deployed automatically in the health environment providing also the exchanging of information with other systems using hl7 XML-based standards (and following the recommendations of IHE).

8.8 Remote Healthcare Solution Deployment

To develop technological solutions for remote healthcare is often a challenge at the same time for both engineers and caregivers, indeed to deploy solutions able to support people with special needs in their daily activities requires a great knowledge of their problems and a qualified understanding how new care approaches can be managed within the healthcare service. These solutions quite often are deployed in the framework of research projects or regional trials, each context and require customized solutions both in terms of components and workflow.

Adopting design process such as Unified Process (UP) surely can help during analysis, design and develop phases, but lacks in details for deployment phase (Jacobson et al. 1999). Based on the experience and outcome collected in the deployment of services in different European countries, this paragraph focuses the discussion about the application of a methodology tailored on UP processes. As matter of fact, the Deployment and Environment workflows of the UP contain less detail than other workflows. For the deployment of a remote healthcare service is really fundamental to be able to design, track and monitor issues such as: proper involvement of stakeholders in the service workflow design, the functional and technical specifications for each system component, the contractual terms towards supplier and customers, data storage features, regulation aspects, ICT infrastructure design and features, system integration strategy, overall system installation aspects, service level of service and technical assistance, the clinical and QOL indicators, training of stakeholders and users based on structured procedures. All these aspects should be covered and defined by a solid procedure during domain analysis phase and then translated and achieved during deployment phase.

8.8.1 A Real Service Deployment Case Study

This paragraph describes the work performed in definition, implementation and deployment of specific remote healthcare services. To facilitate the understanding about the criticism encountered in their implementation, this section introduces briefly a real cases of study applied for the I-care platform.

I-care platform provides different service levels ranging from security to telecare up to remote healthcare services. Twenty apartments have been equipped with I-care platform with specific service customization according with user's profile and related care model. I-care collects data from a set of devices deployed into user environment and processes them to provide information and periodic reports to the users or to care giver staff. These services are based on a monitoring system spread on three levels and integrate different devices for monitoring ECG, blood pressure, glucose levels, weight, movement activity, entrances controls, smoke and flood alarm. All these devices are integrated by means of a universal hub capable of exchange data between a service center and the users. Such a service collects data and provides warning about user's unexpected behaviors or about healthcare metrics. The warnings can be forwarded through the DGHome "I Care" service manager platform to the user, relatives or to a care-giver team, according to the patient profile and setting. As result, the service platform enable a service for a push/pull request of intervention by the proper profile.

I-Care platform can be regarded as one of the typical modules enabling acquisition and exchange of information into a remote healthcare service. Typically it is composed by a user friendly interface, several components such as monitoring devices made by different manufacturers and communicating through different protocol with the universal hub installed into user environment. Each universal hub is connected to the DGHome platform module I-Care as a server side layer, able to integrate several service level and authorization profiles. This components operates the reasoning and context aware module able to produce specific reports and generate timely warnings. In order to achieve a proper and successful deployment, due to the complexity of real scenario of healthcare service organizations, such technological platform requires a proper deployment strategy and procedure based on a specific deployment workflow and a set of quality procedures and checkpoints.

As a matter of fact, a monitoring service can embrace different stakeholders (i.e. GP, Medical Specialist, Nurse) with different responsibilities in a specific service workflow. These stakeholders as well as customers and final user must be properly involved during deployment of the service. Instinctively, a final user has a mistrust approach to technology changes. Consequently the importance of a soundness deployment based on acceptance tests and a deeper beta-test is very important.

8.8.2 Method for Service Deployment

The methods applied for remote healthcare systems integration and deployment have been structured in a specific procedure. The methodology chosen to achieve such a goal is the UP framework. UP is an iterative integration process (Kruchten 2003) intended to be tailored by development teams in order to select those elements that best fit specification and requirements. UP can help in the process not only for implementing an object oriented analysis but also for modeling business process as well (Prince 2003).

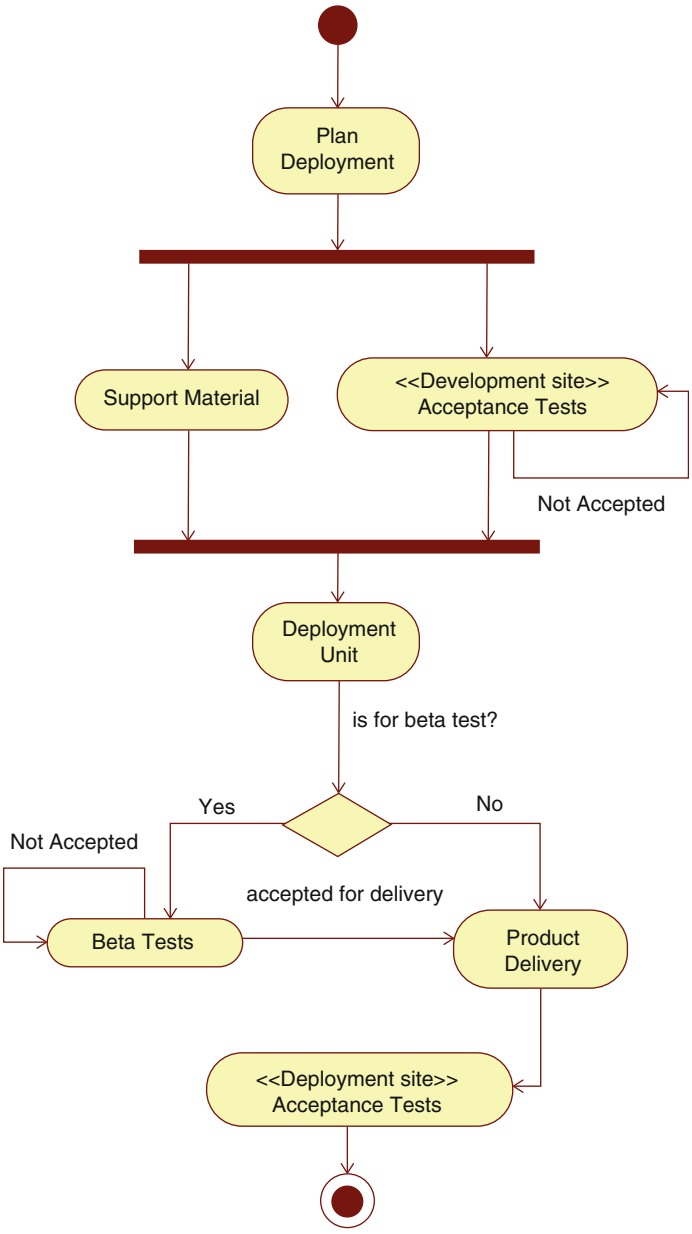


Fig. 8.3 Deployment: main workflow

In particular in this paragraph is presented the deployment process workflow in relation to some major problems encountered in deployment phases. In particular, the case study presented discusses the rationale of a process capable of scheduling activities. Figure 8.3 shows the overall process: each step will be iterated along

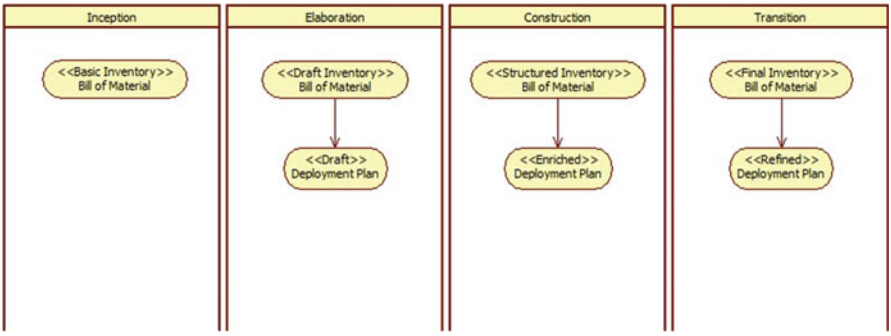


Fig. 8.4 Deployment cycle

project lifecycle through the four UP phases. The resulting output for the completion of each iteration is an artifact or series of artifacts.

This process is described in more details, using a variation of UML notation. Following activities are represented as activity diagrams, but swing lanes do not represent activities' actors as usual: such lanes mean to identify activities to be performed for each UP phase (represented as lane), so every artifact produced in a phase will be used as input for the following.

8.8.3 Deployment Plan

In a service or system integration activity, deployment planning should be considered a crucial aspects for proper validation and acceptance by the users of the proposed service. Deployment Manager has to take into account not only deliverable software and hardware components, but also the provision of proper of training material and activities, assistance coverage and define their timing through process cycle. During Inception phase, designers will collect a bill of material as a basic inventory, in order to create a list of artifact that go to make up the build/product, such as configuration items, documents, installation scripts, devices, packaging items. This document will be enriched during next phases (Fig. 8.4). If the product includes some devices, this discipline aims to list all of them, describing installation issues, considering delivery country and which communication channel will be adopted if data exchange with such devices is needed (Williams 2003).

8.8.4 Support Material

Support material covers the development of the full range of information required by end users to install, use and maintain the system produced. It also includes development of training material and assistance service definitions (Fig. 8.5).

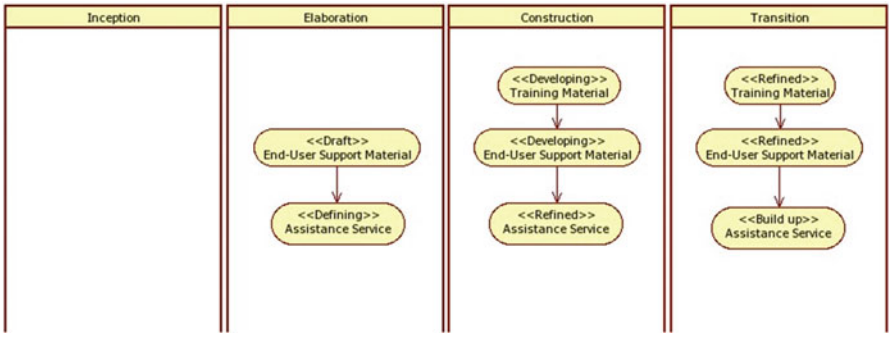


Fig. 8.5 Support material cycle

Developing training material has mostly effect in latest phases of product lifecycle, where previous phases of design are almost consolidated. Some artifact from previous phases can help in training developing: software requirement specifications, use cases, navigation maps and graphic interface stub can constitute the basic inputs for producing effective training and end users support materials (i.e. documentation, installation guides. . .).

On the other hand, assistance process definition should start in earlier iterations of Elaboration phase, producing a draft of Assistance Plan. This artifact should collect all material and consideration for defining an assistance service capable of cover product lifecycle. In latter phases of UP, this document will be enriched with proper procedures to follow in order to notify any issue or how to contact assistance service. A good service should also be capable of receipt and discern evolutionary assistance from maintenance assistance: this approach will let the application to be fixed when needed and to evolve in time accepting and validating new requests.

8.8.5 Acceptance Tests

The main purpose of this discipline is to ensure the product is acceptable to the customer prior its release: these are formal testing, supervised by the customer, performed at development site using target environment and conducted to determinate whether or not a system satisfies its acceptance criteria (Fig. 8.6).

Such criteria should be defined since Inception phase and then refined according to project development, producing an Acceptance Plan. This artifact should describe how the customer would evaluate requested specifications in order to determinate if they meet a predefined set of acceptance criteria. It's important to remind that Acceptance Plan, like most of the UP artifacts, is a "live documents", i.e. they evolve over time to best fit project development process. This is the preliminary main session of tests that aims to discovery issues and bugs since Construction phase. Much more tests activities are planned in Deployment discipline.

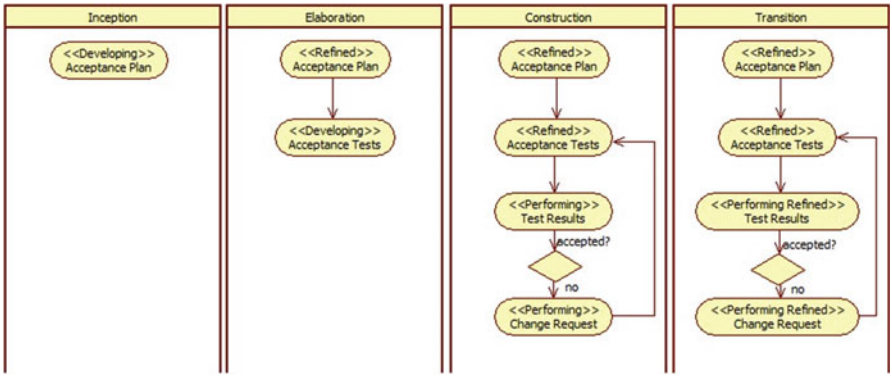


Fig. 8.6 Acceptance test cycle

In case of unacceptable test, a Change Request will be raised for the anomalies: failures will be evaluated and proper tests will be performed again. In this case, a bug tracking system (such as Bugzilla, JIRA. . .) is a good practice in order to capture, manage and store issues, defects, enhancements, new requests and changes.

8.8.6 Deployment Unit

This discipline aims to create a deployment unit ready to be effectively installed and used. It means that the produced build will consist of the software and of all necessary accompanying artifacts required. In our workflow, this unit is created as a prototype for beta testing purpose. The unit will be submitted to a subset of final users: using this unit, they should be able to install the software, configure devices, connect to the network (whereas such characteristics were requested) and use the application. The design of a Deployment Unit starts in Elaboration phase: only in further phases it will be enriched with Installation Artifacts. These artifact could be installation utilities, installation instructions, user manuals or some hardware if needed, such as adaptors for power supply or for devices, and all that things that allow the final user to be able to use the product.

8.8.7 Beta Tests

After each deployment unit activity iteration, a set of beta tests are performed. This kind of tests will be executed with a subset of final users, supervised by the customer. This activity will show application usability and stability: such tests should be executed over all functionality developed at the moment in order to

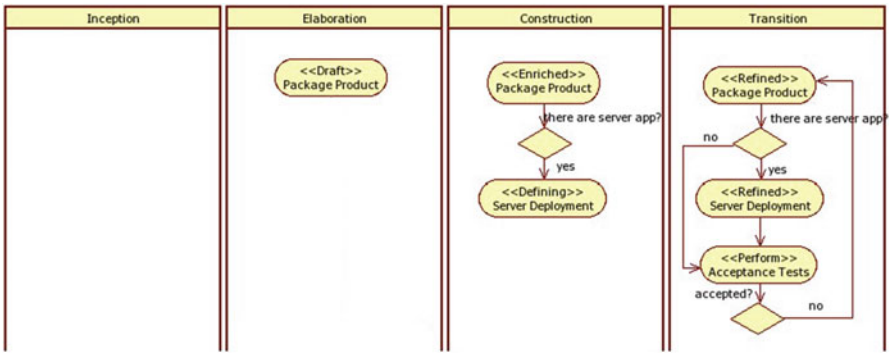


Fig. 8.7 Product delivery cycle

stress the application. The main purpose is to solicit feedback from a subset of intended users. Feedback from the beta program should be treated as refinement requests or bug fixing, by submitting a new Change Request. During construction phase, deployment manager should define some measurement parameters for validating and monitoring tests. In transitions phase, these parameters will be used to evaluate tests results.

8.8.8 Product Delivery

The basic idea of this discipline is to take the deployment unit with all its artifacts and package them for production. Since Elaboration phase the packaging issue should be taken into account, if required by the product (Fig. 8.7).

This discipline should consider all questions about product artworks, manufacturing constraints and delivery services by configuring all necessary hardware. As matter of fact, most of the remote healthcare services are enterprise applications based on communication between the end user and a specialist who cares about their patients.

So, all issues about server configuration and deployment of server side application should be considered. During Construction phase, server configuration and all related questions should be planned: only in Transition phase (in one iteration) the server will configured and the software will be deployed. Once server application is configured, if expected by project, and a unit is delivered, new acceptance tests iteration is required. This final tests will be performed at deployment site in order to ensure that all system architecture is working properly. If something goes wrong, some activities, such as hardware/software fixing, server re-deployment, have to be reiterated.

As result based on UP approach it is possible to define a pathway to ensure deployment of a working and reliable solution in the domain of remote healthcare and Ambient Assisted Living. The proposed UP approach it is of paramount

relevance not only to model deployment processes but also to empower specific workflow with the proper involvement and participation of the overall value chain of stakeholders.

8.9 Conclusion and Open Questions

As lesson learnt the correct integration of the components to successfully integrate a remote healthcare system and deploy it in a real settings should follow: (0) clear description of service model and the related advantages with an early stage involvement of stakeholders; (1) detailed integration plan addressing minimal impact and seamless integration with the existing healthcare system; (2) service deployment procedure; (3) educational framework with relevant opinion leaders; (4) compliant with current laws and regulations; (5) adoption of accepted and successful business and service models, (6) proper evolution and scalability strategy; (7) normalise data to international standards.

For each context it is important to properly fit the remote healthcare solution with the peculiarities of healthcare organization; based on such understanding plan a modular and scalable deploy of the service. In order to achieve the proper adoption. At the same time there are several open questions that it is important to take into account:

Why relevant group of stakeholders are not properly engaged in service design?

Why such solutions are usually perceived as technical skill demanding?

Why the role of the patient is always not very relevant in the service model?

Why, given a clear proof of clinical and cost benefits, decision makers do not apply any budget shift or budget allocation for the deploy of such a solutions?

References

Beale T, Heard S, Kalra D, Lloyd D (2007) The openEHR reference model—EHR information model. [Online] openEHR foundation [Cited: 15 June 2013]. www.openehr.org/releases/1.0.1/architecture/rm/ehr_im.pdf

Bernstein K, Tvede I et al (2009) Can openEHR archetypes be used in a national context? The Danish archetype proof-of-concept project. Stud Health Technol Inform 150:147–151

Blobel B (2010) Architectural approach to eHealth for enabling paradigm changes in health. Methods Inf Med 49(2):123–134

Chen R, Klein G et al (2009) Archetype-based conversion of EHR content models: pilot experience with a regional EHR system. BMC Med Inform Decis Making 9(1):33

Collins (2013) Portability. http://www.collinsdictionary.com/dictionary/english/portability. Accessed 15 Jun 2013

Duftschmid G, Wrba T et al (2010) Extraction of standardized archetyped data from electronic health record systems based on the entity-attribute-value model. Int J Med Inform 79 (8):585–597

Edelstein L (1943) The Hippocratic oath: text, translation, and interpretation. Johns Hopkins Press, Baltimore, MD
European Commission (2007) Article 29—"Data Protection Working Party". Working document on the processing of personal data relating to health in electronic health records (EHR), WP131. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp131_en.pdf
European Commission—DG Information Society/Q-REC (2007) Inventory of relevant standards for EHR systems. European Commission—DG Information Society/Q-REC, Brussels
Glazer I, Blakley B (2009) Identity and privacy strategies—privacy. Burton Group, Portland
Henry T (2009) Security and risk management strategies, information confidentiality. Burton Group, Portland
Iakovidis I, Dogac A et al (2007) Interoperability of eHealth systems, a selection of recent EU's Research Programme Development. eHealth: Combining Health Telematics, Telemedicine Biomedical Engineering and Bioinformatics to the Edge, Regensburg, Germany
Institute of Electrical and Electronics Engineers (IEEE) (2007) The role of standards in engineering and technology. http://www.ieee.org/education_careers/education/standards/standards_glossary.html. Accessed 15 June 2013
ISO/DTR 20514 (2005) Health informatics—electronic health record: definition, scope and context. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=39525. Accessed 15 June 2013
Jacobson I, Grady B, Rumbaugh JE (1999) The unified software development process—the complete guide to the unified process from the original designers. Addison-Wesley, Reading
Kabak Y, Dogac A, Köse I, Akpinar N, Gürel M, Arslan Y et al (2008) The use of HL7 CDA in the National Health Information System (NHIS) of Turkey. In: Ninth international HL7 interoperability conference (IHIC), Crete, Greece, pp 49–55
Kruchten P (2003) The rational unified process an introduction. Addison-Wesley, Boston
Leslie H, Heard S (2006) Building archetypes. In: Proceeding of learn about openEHR, CEN and SNOMED. Ocean Informatics and the National Centre for Classification in Health Australia
LinkEHR (2011) LinkEHR. [Online] Biomedical Informatics Group. Universidad Politécnica de Valencia [Cited: 15 June 2013]. http://www.linkehr.com/
LiU Archetype Editor (2007) LiU Archetype Editor. [Online] Linköpings universitet. Institutionen för medicinsk teknik [Cited: 15 June 2013]. http://www.imt.liu.se/mi/ehr/tools/
Main T, Slywotzky A (2012) The volume-to-value revolution. Oliver Wyman, New York
Marand & IBS (2013) Marand provides IBS with Think!Med EHR platform to create information infrastructure for electronic healthcare in Moscow. http://en.ibs.ru/content/eng/703/7039-article.asp. Accessed 15 June 2013
McCay C, Kalra D, Worden R (2008) Results of investigating the transformability between HL7 V3, openEHR and EN/ISO 13606. NHS Connecting for Health, Leeds
MDA Success Story, OMG (2013). M1 Global Solutions. Model driven software development and offshore outsourcing. http://www.omg.org/mda/mda_files/M1Global.htm. Accessed 15 June 2013
Murray PJ, Wright G, Karopka T, Betts H, Orel A (2009) Open source and healthcare in Europe—time to put leading edge ideas into practice. Stud Health Technol Inform 150:963–967
NEHTA CKM (2012) NEHTA CKM. [Online] NEHTA CKM [Cited: 15 June 2013]. http://dcm.nehta.org.au/ckm/
Niski J (2008) Application security frameworks. Burton Group, Portland
OASIS (2008) Reference architecture for service oriented architecture version 1.0. OASIS Standard Specification
Ocean Archetype Editor (2012) Ocean Archetype Editor. [Online] Ocean Informatics [Cited: 15 June 2013]. http://wiki.oceaninformatics.com/confluence/display/TTL/Archetype+Editor+Releases
OpenEHR Foundation (2010) Clinical Knowledge Manager [Cited: 15 June 2013]. http://www.openehr.org/ckm/
openEHR Foundation (2012) What is openEHR? [Online] openEHR Foundation [Cited: 15 June 2013]. http://www.openehr.org/what_is_openehr

openEHR Healthcare Providers and Authorities (2012) openEHR Healthcare Providers and Authorities. openEHR Foundation. http://www.openehr.org/who_is_using_openehr/healthcare_providers_and_authorities. Accessed 15 June 2013

openEHR Wikipedia (2011) Wikipedia. http://en.wikipedia.org/wiki/OpenEHR. Accessed 15 June 2013

Prince R (2003–2005) Using RUP/UP: 10 easy steps—a practical guide. Toolbox: a knowledge sharing community [Online]. http://hosteddocs.ittoolbox.com/RP092305.pdf

Semantic HEALTH (2009) Semantic interoperability for better health and safer healthcare. Semantic Health Report. European Commission. Information Society and Media Directorate-General, Luxembourg

Stroetman KA, Artmann J, Stroetman VN et al (2011) European countries on their journey towards national eHealth infrastructures. eHealth strategies report. http//www.epsos.eu

Swedish openEHR CKM (2010) Swedish openEHR CKM. [Online] Swedish openEHR [Cited: 15 June 2013]. http://sllocean.karolinska.se/ckm/

Think!Med Clinical (2013) Marand. http://www.marand-thinkmed.com/. Accessed 15 June 2013

Wikipedia (2013) Software portability. http://en.wikipedia.org/wiki/Software_portability. Accessed 15 June 2013

Williams G (2003) Telehealth—a keystone for future healthcare delivery. Health Informatics Europe [Online]. http://www.hi-europe.info/files/2003/9994.htm. Accessed 11 Mar 2003

Yong H, Jinqiu G, Ohta Y (2008) A prototype model using clinical document architecture (CDA) with a Japanese local standard: designing and implementing a referral letter system. Acta Med Okayama 62(1):15–20

Chapter 9
A Business Perspective

Silvio Bonfiglio

Abstract Major trends are impacting healthcare such as the access to a larger number of people, the growing demand for better quality of care, the ageing of the population in Europe; all of them are causing skyrocketing healthcare costs. To cope with these challenges improvements in quality and efficiency are needed and new organizational models with major changes in the way healthcare services are organized and delivered to the citizens have to be introduced.

Healthcare has to move from treatment to prevention, from a "clinical-centric" to a "patient-centric" approach where the citizen is motivated and empowered to manage his own health.

ICT will play an important role in this new scenario and this healthcare's evolution is attracting the attention of the industry. The digital revolution is impacting healthcare too. The new frontier of the mobile healthcare (m-health) will contribute to bridge the gap between the healthcare system and the citizens and will increase the level of patient/caregiver connectivity.

From a business perspective e-health represents a market still in a nascent phase with a lot of barriers and challenges but has the potential of becoming the third largest industry in the healthcare sector. Mainly in this initial phase, small and medium enterprises, technology-driven companies, innovators could have great opportunities.

This chapter deals with the specificity of the e-health business; it identifies business opportunities, investigates on the existing barriers and analyses possible business models and the methodologies for a socio-economic assessment of e-health.

S. Bonfiglio (✉)
FIMI-BARCO, via S. Banfi 1, Saronno 21047, Italy
e-mail: Silvio.Bonfiglio@barco.com

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_9,

9.1 Trends in Healthcare

Major trends are impacting healthcare; among them:

- The access to healthcare for a larger number of people,
- The growing demand for better quality of care and the shift of the economic power towards the consumers (i.e. the patients),
- The dramatic changes in demographics and a larger number of patients with age-specific, chronic and degenerative diseases (in Europe 25–60 % of older people suffer from chronic diseases such as heart disease, diabetes and COPD (WHO report—2010)).

As a consequence we see skyrocketing healthcare costs and shortfall of care staff, mainly in some regions (as an example in India 75 % of the population live in rural villages while the 75 % of Indian doctors are based in cities (Agency France Press 2005)).

To cope with these challenges major changes are needed in the organization, delivery and management of healthcare to get improvements in quality and efficiency. According to the Institute for Healthcare Improvement (2007 report) *"many healthcare systems around the world will become unsustainable by 2015. The only way to avoid this critical situation is to implement radical changes"*.

Healthcare needs to focus on prevention and to move from a clinical-centred to a person-centred approach. We expect an evolution over the next decade towards a continuous and pervasive care at home and on the move. In this approach informal caregivers (relatives, friends, volunteers) will play an important role by complementing the tasks of the medical professionals and the patient himself will get motivation and empowerment to better manage his/her own health through medical education, healthier lifestyle and better adherence to the prescribed therapy.

Some data give a clear picture of the situation:

- In Europe the number of people aged 65+ is expected to go up from approx. 20 % of the overall population (2010) to 28 % (2050) (Ageing Report 2012, European Commission 2012).
- Public healthcare spending is projected to rise—as an average—by 3 % of GDP in advanced economies and by 1 % of GDP in emerging economies with regional variations (see Fig. 9.1).
- In Europe the healthcare spending was equivalent to 8.3 % of GDP in 2008 of which 6.2 % as average public expenditure (France with 11 % of GPD, Germany 10.6 % and Belgium 10.3 % were among the countries with higher public expenditure) (OECD, StatExtracts - Health Expenditure and Financing 2012);

Analysts expect that these costs will double during the period 2005–2050 and it was highlighted how healthcare spending is rising faster than economic growth in most EU countries (e.g. in The Netherlands 3 % vs. 1.25 %).

The healthcare system is expanding with regard to the users (from episodic patients to the overall population and on a regular basis), to the caregivers (from the medical professionals to informal carers and up to the patients themselves), to

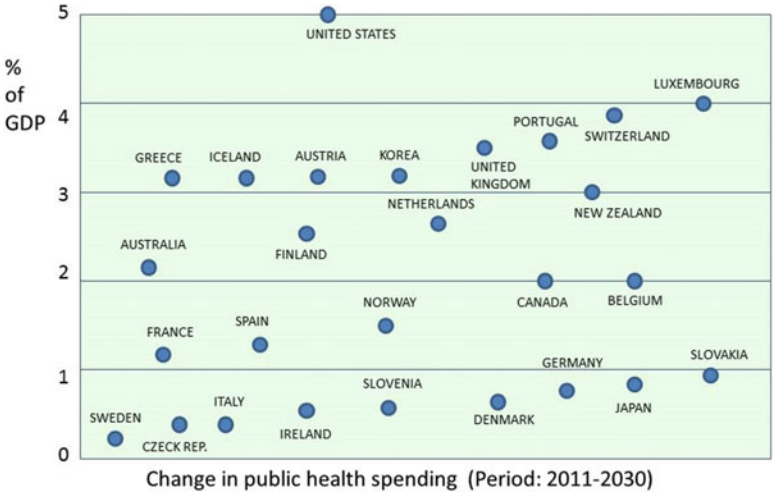


Fig. 9.1 Rising of the healthcare spending (period: 2011–2030)

the location where care is provided (from the hospital to the patient's home and everywhere) and finally with regard to the tasks and objectives (from diagnosis and treatment to prevention) (see Fig. 9.2).

9.1.1 The Role of ICT in Transforming Healthcare

New technologies and mainly ICT will play an important role; the digital revolution is impacting healthcare too and some analysts expect that healthcare will represent a new accelerator for the advance in ICT (see Fig. 9.3).

The provision of ICT-enabled healthcare services is seen as a great opportunity for the containment of the healthcare delivery costs while maintaining the expected levels of quality of care and safety.

ICT will complete the realization of the "digital hospital" by facilitating and enabling the clinician's work at the point of care (security, decision support, increased productivity, mobile access to information, reduction of errors, cost reduction). Moreover ICT

- Has the challenging opportunity of reshaping the healthcare beyond the hospital's environment by making possible "home care" and "personal healthcare" approaches;
- Has the potential of making possible the move towards a person-centric healthcare.

A better leveraging of the potential provided by ICT represents a challenge and at the same time an economic opportunity.

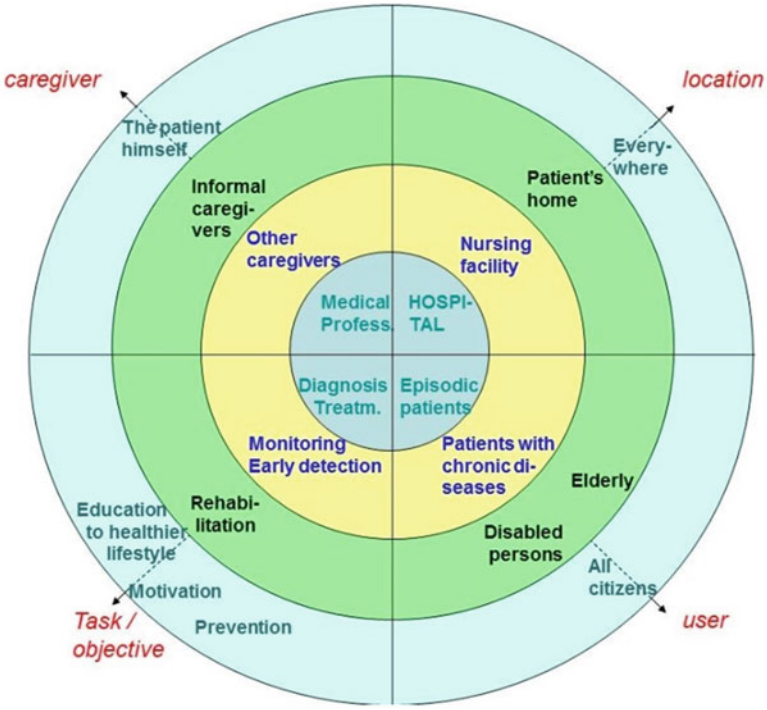


Fig. 9.2 The expanding world of healthcare

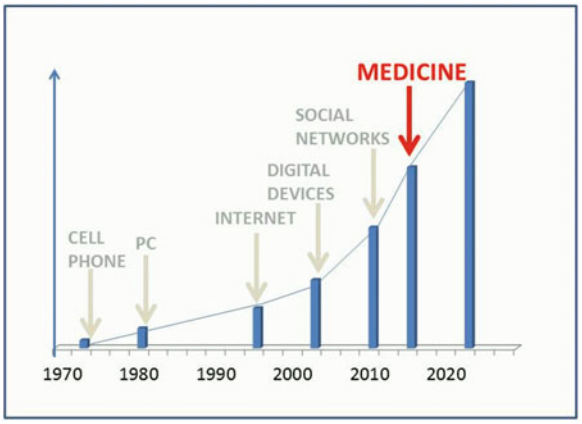


Fig. 9.3 The technological revolution

9.1.2 New Technological Opportunities

New technologies are expected to have a profound impact on the way healthcare services will be provided in the coming years.

A. The mobile health (m-Health)

Advances in sensor networks, embedded systems, such as the smart wearable systems and pervasive computing, create opportunities for treatments to be available to people when and where they need; m-health will help to bridge the gap between the healthcare system and the citizens and will increase the level of patient/caregiver connectivity.

The global m-health market is primarily driven by factors such as:

(a) The rapid advance in mobile technologies and applications and the continuing growth in the coverage of mobile cellular networks;
(b) New opportunities for the integration of mobile health into existing e-health services;
(c) A growing awareness and a more active participation of the patients to the healthcare management.

The development of smartphone applications has created new and interactive ways of communication between patients and healthcare providers. The use of smartphone applications is expected to revolutionize the m-health market by improving healthcare delivery and dissemination of medical information.

The market is attracting the interest of large companies such as IBM, GE, Qualcomm, AT&T.

According to a report (Transparency Market Research 2013) the global m-health market is expected to reach USD 10.2 billion by 2018 from USD 1.3 billion in 2012 at a CAGR of 41.5 % from 2012 to 2018. The monitoring services segment contributed with about 63 % of the global m-health market revenue in 2012.

The following table reports the outcomes of a survey performed by the World Health Organization (WHO) in 114 Member States; four aspects of m-health were analyzed: the adoption of m-Health applications, the types of applications, the status of assessment/evaluation programs and the barriers to the implementation (Table 9.1).

A1. Medical Apps in smart phones *("the doctor in your pocket")*

The development of smartphone-based applications has created new and interactive ways of communication between patients and healthcare providers. Use of smartphone applications is expected to revolutionize the m-health market by improving healthcare delivery and dissemination of medical information. Many medical applications for smartphones have been developed and widely used by health professionals and patients. The use of smartphones is getting more attention in healthcare day by day and they can play a very important role in patient education, self-management of diseases and remote monitoring of patients. According to the most recent data, some 40,000 medical apps are now available for smartphones and tablet PCs such as iPads. They were introduced typically by small/medium enterprises and by start up companies. For the SMEs and start ups working in this field a big issue is represented by the needed certification of their products before going in the market; in USA the FDA approval process adds months, if not years and it discourages many of these enterprises.

Table 9.1 m-Health in the world

WHO survey on m-health conducted in 114 member states			
Analyzed m-Health services:			
• Appointment reminders, • Community mobilization and health promotion, • Decision support systems • Emergency toll-free • Health awareness raising and • Health call centres, • Health surveys and data collection, • Information access,		• Managing emergencies and disasters, • Mobile patient records, • Mobile telemedicine, • Patient monitoring, • Surveillance, • Telephone services, • Treatment compliance.	
Investigated aspects			
1. Adoption	2. Type of applications in use	3. Realization of evaluation programs	4. Barriers to implementation
Survey outcomes			
1	83 % of Member States (95 over 114) implemented 1 m-health service at least; many countries have experienced 4 to 6 services;		
2	Most adopted services/ applications are: • Health call centres (59 %), • Emergency toll-free telephone services (55 %), • Managing emergencies and disasters (54 %) • And mobile telemedicine (49 %)	Least frequently adopted m-Health services are: • Surveillance, • Health awareness raising and • Decision support systems.	
3	A part from health call centres, emergency toll-free telephone services and managing emergencies and disasters, all the other services are in a pilot or informal stage only. The European countries are currently the most active ones in m-health		
4	**Main barriers** • Economic situation: health systems worldwide are under increasing pressure to perform under multiple health challenges, such as chronic staff shortage and limited budgets; • Lack of evaluation programs and solid evidence on the benefits (only 12 % of the Member States executed or are running well structured evaluation programs) • Data security is still a main concern;	• Typically only small-scale pilot projects addressing single issues were realized; there are just few large scale implementations primarily supported by public-private partnerships; • Lack of policies and strategies integrating e-health and m-health into conventional healthcare services; • Lack of standardization; • Lack of suitable telecommunication infrastructures.	

Obviously challenging aspects remain the limited resources in terms of power, computational capability and memory; moreover the "clinical validity" of some proposed solutions is still to be proven.

Patient self-monitoring could be one of the most important emerging categories of applications for Android and iOS phones.

B. **Mobile imaging**

Medical images can now be visualized in a smart phone or in a i-Pad. BARCO developed a solution for easy and effective DICOM calibration of images on the i-Pad screen. This makes possible new interesting applications such as:

(a) Remote referral and image-based screening

Images can be acquired with a mobile device (e.g. a mobile ultrasound system such as the Vscan of General Electrics), digitized and uploaded to the Internet cloud; doctors on the move can download the images with their smart phone or i-Pad and assess them. It will avoid unnecessary traveling to specialized medical centers and reduce examinations at the hospital with costly equipment.

(b) Consultation among doctors

It will help to reduce errors through exchange of opinions and will facilitate sharing of experiences;

(c) Remote monitoring of patient at risk for early detection of degenerative trends.

It will allow early detection of degenerative trends, avoid unnecessary traveling to specialized medical centers.

C. **Cloud computing and e-health**

According to NIST (National Institute for Standards and Technology) *"cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."*

Cloud computing has the potential to solve many pressing issues in the application of IT in health. It offers opportunities for innovative IT-enabled approaches by leveraging distributed health resources in organized ways. Electronic Medical Records (EMRs) and Personal Health Records (PHRs), e-referral, clinical decision support, chronic diseases management and remote monitoring could be the entry point for cloud computing into healthcare (Canada Health Infoway - Cloud Computing in Health 2012). Furthermore the cloud allows providers and more importantly health care delivery organizations to focus less on managing IT and more on delivering care to their patients.

Cloud computing promises a future where IT resources are unshackled from day-to-day tasks such as system maintenance and introduces the opportunity of focusing scarce staff on more important, challenging and rewarding initiatives like

improving services and increasing innovation. Obviously several issues need to be addressed, including concerns about privacy and security.

D. **Social networks in healthcare**

Some recent data give the confirmation of a growing penetration of healthcare-based social networks mainly in the United States of America:

- 33 % of American citizens going online for healthcare topics use social networks to find fellow patients and discuss their conditions;
- 36 % of social network users evaluate and leverage other consumers' knowledge before making health care decisions;
- 55 % of surveyed Americans get information about a therapy or condition online;
- 60 % of surveyed physicians and 65 % of surveyed nurses are interested in using social networks for professional purposes.

Although social networks are considered primarily a recreational tool, they are becoming increasingly important also to businesses and organizations. Social networks have the potential to offer value not only to the patients but also to health care organizations; in fact they can be used to reach stakeholders, aggregate information and leverage collaboration.

Public, Internet-based social networks can enable communication, collaboration and collection and sharing of information in the health domain.

Nevertheless—as noticed by some analysts—the networks' commercial value to specific health care organizations and the industry as a whole remains elusive. Unlike Internet advertisements and traditional marketing tactics, many businesses have not yet developed robust metrics to measure success and return on investment related to social networks. In addition, regulations limit what industry organizations can say to physicians and consumers. For these and other reasons, health care lags other industries in its social network presence. Healthcare industry stakeholders will face inevitable challenges in adapting to this new technology, complying with regulations and evaluating/acting upon the potential clinical and business opportunities inherent in social networks.

As the regulatory environment becomes more defined and innovative organizations demonstrate measurable commercial value from social networks, increasing numbers of health care stakeholders are expected to recognize and leverage this transformational technology's role in information acquisition and access.

Expected main use of social networks in healthcare for the patients are:

- Health coach and Health advices,
- Disease management,
- Connection with other patients having similar diseases,
- Sharing and commenting health-related news,
- Address of hospitals, medical services providers and opinions on their services.

Examples of potential social network use by organizations are:

- For public health organizations to quickly reach the public and alert them with regard to policy debates and key decisions such as vaccination, quarantines, evacuations and in general public health campaigns;
- For hospitals to help consumers understand their treatment options and obtain feedback on their services;
- For medical research organizations to recruit patients for clinical trials;
- For medical device companies to train physicians remotely on their products.

9.2 e-Health, a Market Still in Its Nascent Phase

According to a definition of Pagliari et al. (Pagliari et al. 2005 - What is e-health?) e-health can be defined as "a broad range of medical informatics applications for facilitating the management and delivery of healthcare"; telemedicine (or telehealth) can be considered as a part of e-health characterized by the physical distance between the patient and the healthcare provider.

According to another definition (Eysenbach, What is e-health? J Med Internet Res 2001) e-health is "not only a technical development but also a state-of-mind, a way of thinking, an attitude and a commitment for networked, global thinking to improve health care locally, regionally and worldwide by using ICT".

Regardless of the definition, from a business perspective e-Health has been considered by the European Commission as a "lead market" i.e. a "market for innovative products and services or technological solutions with high growth potential, where EU industry can develop competitive advantage and that requires action by the public authorities to deal with regulatory obstacles" (European Commission, DG SANCO 2009).

The e-health Lead Market Initiative of the European Commission identified four market segments of major interest for Europe (Table 9.2):

According to a report issued by Capgemini Consulting and related to 2008 the shares of these segments are as indicated in the following diagram (Fig. 9.4).

In the poor 0.9 % share of telemedicine the 70 % is related to cardiac diseases. (Frost and Sullivan, European Market for Remote Patient Monitoring 2013).

Between 2008 and 2012 the situation evolved with a major shift from Secondary Usage Systems to Clinical Information Systems.

This indicates that e-health systems are currently targeted more towards supporting the operational processes of healthcare professionals rather than in changing the fundament approach on the way to deliver healthcare services.

Although a considerable range of promising services and systems has emerged from research efforts, wider mainstreaming of ICT-enabled solutions has to a large extent yet to occur (Ageing Report 2012, European Commission). The situation is similar in the North America; also there e-health market does not yet fully ensure the availability and uptake of the necessary ICT-enabled solutions coming out from the research laboratories.

Table 9.2 e-health market segments

	Market segment	Description
1	Clinical Information System (CIS)	(a) Specialized solutions for health professionals within healthcare institutions (e.g. hospitals). Examples are radiology information systems, nursing information systems, medical imaging, computer-assisted diagnosis, surgery training and planning systems; (b) Tools for primary care and/or for outside care institutions, such as general practitioners and pharmacy information systems.
2	Secondary Usage Non-clinical Systems (SUNCS)	(a) Systems for health education and health promotion of patients/citizens, such as health portals or online health information services, enabling patients to have greater control of their health; (b) Specialized systems for researchers and public health data collection and analysis, such as bio-statistical programs for infectious diseases, drug development and outcomes analysis; (c) Support systems, such as supply chains management, scheduling systems, billing systems, administrative and management systems, which support clinical processes but are not used directly by patients or healthcare professionals.
3	Telemedicine	Personalized health systems and services, such as (a) Disease management services, (b) Remote patient monitoring (e.g. at home) (real time or store-and-forward approaches; video-based or web-based), (c) Teleconsultation, (d) Telecare, (e) Telemedicine and teleradiology.
4	Integrated Health Clinical Information Network	(a) Distributed electronic health record systems and (b) Associated services such as e-prescriptions or e-referrals.

A study (Empirica & WRC 2008) realized on behalf of the European Commission gives an overview of the worldwide market situation (see Table 9.3):

In a report issued by Gartner (Gartner, eHealth for a healthier Europe 2008) the estimated time to mainstream adoption of e-health products is still long and forecasted in the range of 10–15 years.

A main characteristic of the e-health market in Europe is its differentiation across the various countries with United Kingdom, Germany, France, Spain, Italy, The Netherlands as the principal European e-health regions (see Fig. 9.5).

In spite of the current situation all the analysts agree that the market potential is huge (World Health Organization (WHO) 2009). The growing number of older adults and the consequent rise of the number of patients suffering from chronic diseases represent a great opportunity for e-health devices/solutions/services. Market participants agree that the e-health market in Europe is set for

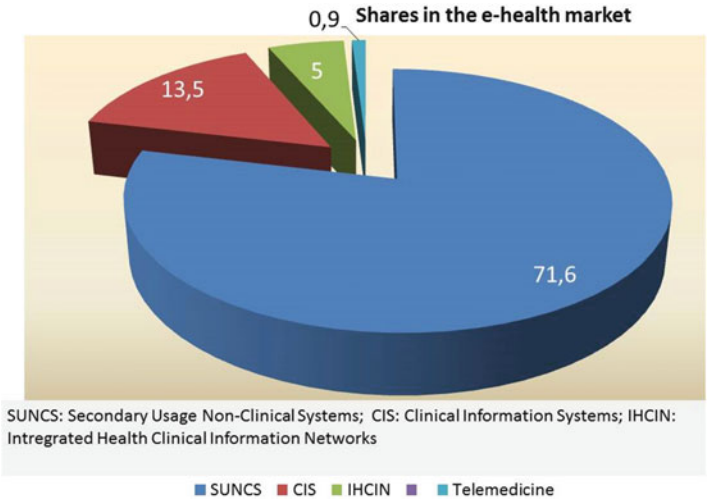


Fig. 9.4 Shares of the various e-health market segments

Table 9.3 Availability of e-health products/solutions in the market

Fully mainstreamed products	Partially mainstreamed products	Localized instances	Only pilots and trial activity	Little or no activity
None	U.S.A.	Denmark, Spain, Finland, United Kingdom, Germany, Sweden	Italy, France, Hungary, The Netherlands, Poland, Japan	Other countries

exponential growth and it is also believed that the e-health worldwide market has the potential to be the third largest in the health sector with a global turnover of Euro 50 to Euro 60 billion, of which Europe represents one third (Price Waterhouse 2005; HINE 2007).

9.2.1 Value and Sustainability of e-Health

e-health services are requested to provide value and sustainability where:

- 'Value' refers to socio-economic and health-related benefits for the patients and/or medical professionals (better clinical care, safety, timeliness of care, quality, effectiveness and efficiency);
- 'Sustainability' refers to the overall financial/economic performance in line with a set of predefined benchmarks and indicators.

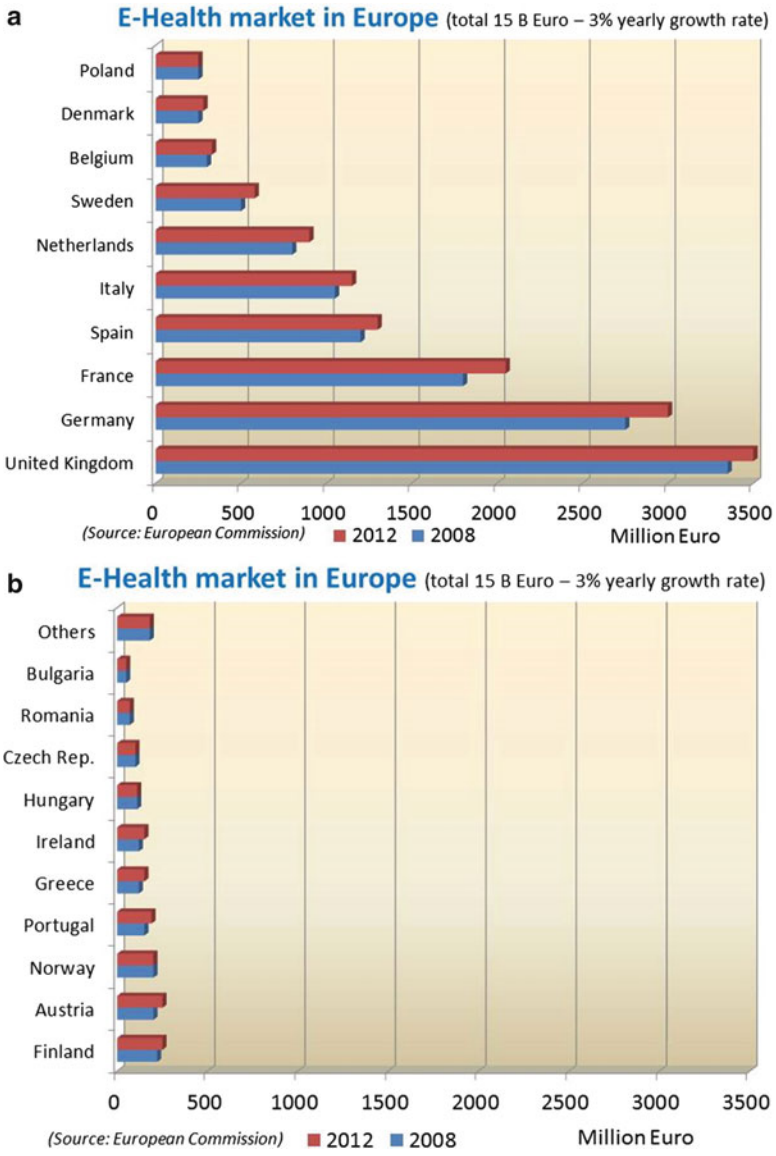


Fig. 9.5 Growth of the e-health market in Europe (2012 vs. 2008)

Value creation and sustainability require e-health services to be supported by business models reflecting the interests of all the involved stakeholders and public policies establishing the right conditions for devising these business models.

Furthermore it has become evident that market forces alone are insufficient to ensure the realization of this potential. "An effective technology will fail if resources are lacking" (Prof. John Cleland).

9.2.2 The External Environment

There are various external factors having a relevant impact on the uptake of the e-health market; we analyze some of them here below:

A. **The political issues**

A1. Lack of coordinated polices

Currently each EU Member State is responsible for the operational delivery and financial management of healthcare. In some countries the healthcare services are managed by the regional and local administrations. This situation conflicts with the possibility of EU citizens of receiving equal quality treatment at equal conditions irrespective of their geographic location. The European Commission is working towards the establishment of a common framework for the application of patients' rights in cross-border healthcare.

Moreover there is a lack of actions for the integration of e-health into the conventional healthcare system and into the hospital's workflow (often linked not only to technical problems).

A2. Lack of a clear reimbursement policy

The funding of healthcare among EU Member States varies; generally they all rely on a combination of resource funding, with the majority of funds directly or indirectly controlled by national state administration.

Overall, in Europe there are four different methods of healthcare financing:

- A system centred on public taxation;
- An approach based on compulsory social insurance;
- A model based on private finance through voluntary insurance that operates on top of standard social insurance;
- Other cost sharing mechanisms through which patients contribute to healthcare financing.

Often the financing of ICT-enabled healthcare services is difficult because of the increasing mismatch between societal needs and available financial resources.

Other political issue to be considered are the l**ack of a clear medical, regulatory and legal framework,** the **lack of infrastructures** (at least in some European countries) and the d**ifficulty in implementing joint public-private initiatives**

B. **Market barriers and drivers**

***B1*. Market fragmentation**

Currently the market is characterized by the presence of local participants; it makes difficult the reaching of a market globalization driving cost reduction

B2. High initial costs and doubtful return of the investments at least in the short/medium terms;

B3. Market drivers

On the contrary as main drivers for the e-health market we can list:

- The ageing population and the growth of chronic diseases requiring continuous monitoring,
- The technological advance,
- The demand for early hospital discharge,
- The shortfall of medical resources and the long waiting lists,
- The demand for a personalized healthcare.

9.2.3 The Need for New Approaches

New approaches are needed to establish the enabling conditions for a successful deployment and exploitation of e-health in Europe and for its integration into the clinical workflow.

- **Political initiatives**

Public policy actions are needed to support the development of value-creating business models for e-health. Options include:

– Bringing legal clarity mainly in the areas of access of personal health data, safe exchange of medical data across national borders, personal integrity of the patient and right of patients to give a consent to the use of their medical data;
– Facilitating the market developments by fostering interoperability, common medical terminologies and technical standards.
– Launching pilot actions of e-health-related projects where different business models can be tested or simulated using appropriate modeling approaches and defining benchmarking parameters to monitor and compare business models for e-health;
– Supporting the development of best practices and fostering their sharing across Europe for the design of sustainable and value-creating business models.

- **A synergic effort**

The synergic effort of all the stakeholders towards a common objective (a better quality of care with sustainable costs) is needed. It has to involve national healthcare authorities, health professional associations, healthcare delivery organizations, industry and the research community.

Co-operation and alliance are needed between manufacturers, system integrators, telecom companies and small-medium enterprises; it includes the commercialization by large companies of innovative products developed by smaller companies. In this regard it will be of high importance to understand the real

commitment of the telecom companies to incorporate remote patient monitoring services in their offer to the market.

- **Community-supported services**

The gap between health and social care services is identified as an impediment to better and more personalized care. Deploying systems that link up additional services from third parties and the independent sector creates opportunities for interoperable and unified services that are uniquely focused on patient's needs.

In this way innovation can support collaboration between health- and social care professionals improving services and outcomes for patients and at the same time allowing providers to work more efficiently and productively. Telehealth can be used as a tool to enable more integrated care.

- **A cultural change**

Often awareness is lacking at all levels (by the patients, by the medical community, by the institutions and the policy makers).

The reluctance of the medical community to accept the entrance in the medical world of new technologies such as ICT is often a "cultural matter", a mindset problem. To avoid this barrier, telemedicine should definitely be considered and used by health professionals as a tool and should definitely be at the service of the patients. It should not change the way clinicians practice; it has to fit with the existing framework. The use of technology "should be to support medicine not the other way around".

- **Interoperability of the solutions and reference architectures**

To allow the growth of the market and its sustainability all the solutions must be developed with a fast time to market, must be interoperable and easily integrated. A contribution could be represented by the focus on reference architectures for person-centric health management and the related subsystems and tools. It will offer opportunities for a sustainable business.

Standardization initiatives are needed to fill the existing gaps with regard to interoperability.

- **User-centred design**

A user-centred design approach must be followed to foster the acceptance of the final users through the unobtrusiveness of the developed solutions, the ease of use and the awareness of the resulting benefits. Moreover the robustness (technical reliability) and the medical validity of the proposed solutions represent key features to increase the trust of the patients.

It has been observed that the "the failure of e-health initiatives in many cases was due to a too technologically driven approach; the focus was to exploit new and advanced technologies rather than to meet the real user's needs".

- **The need of proving the real benefits**

Claiming that a e-health system and its underlying business model are value-creating is not the same as being able to prove it. This is a complex process also due to the fact that available literature provides unclear guidance in actually quantifying or qualifying the benefits achieved with the introduction of a e-health system.

9.3 Designing Effective Business Models for e-Health

9.3.1 A Definition of "Business Model"

In the literature we can find several definitions of "business model"; they can be grouped into three main categories:

– The first interprets business models solely in terms of the company's *economic* model i.e. the emphasis is on the firm's mechanism to make money;
– A second group of *operational* definitions focuses on the internal processes and infrastructure, enabling the company to create value;
– Finally, *strategic* definitions emphasize the firm's market positioning, its interactions across organizational boundaries and the company's growth opportunities.

These categories represent a hierarchy in that the understanding of business models becomes more comprehensive as one moves from the economic to the operational and finally to the strategic level.

Irrespective of the different definitions, there seems to be consensus on the core components of a business model; according to Müller-Stewens and Lechner (2005) a business model consists of four parts, each addressing a particularly important question:

– Product and service offerings (Which goods and services shall we offer to which customer segments?)
– Processes and organization (How and in which structure shall we create the products and services?)
– Marketing (How can we acquire and retain customers?)
– Revenue model (How shall we design our profit mechanism?)

The first three components (product and service offerings (i.e. value proposition), processes and organization, marketing) determine the costs of the value created by the company whereas the revenue model defines how the money is earned. The comparison of these two perspectives results in the business model's profitability (Fig. 9.6).

A business model structures and orchestrates the interests of all the involved stakeholders; it maps the interactions among them based on their interests, so as a starting point it requires the identification of each stakeholder, its role and value.

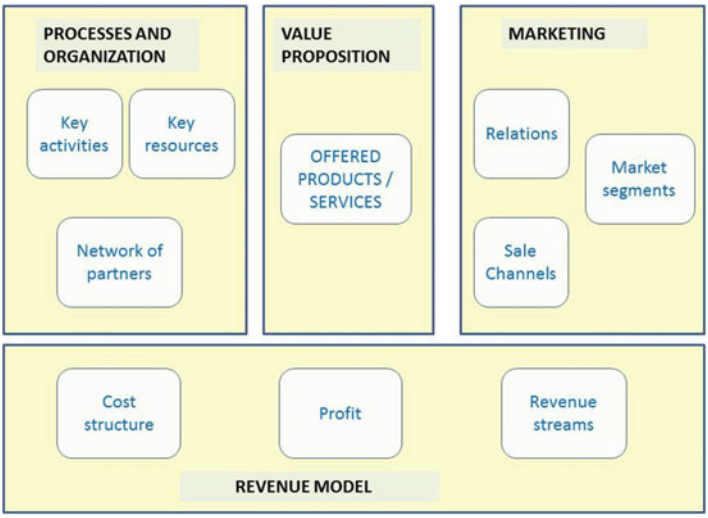


Fig. 9.6 The components of a business model

Furthermore—mainly in a dynamic market—it is important to build a business model in a flexible way in order to make possible its dynamic adaptation according to the changes of the context and to allow to catch new opportunities.

9.3.2 The Specificity of a Business Model in the Health Care Sector

In the healthcare domain the society as a whole is involved. Therefore, when devising a business model for healthcare it is necessary to identify the value to be gained not only by the commercial chain and by the buyers but also by the overall community (the patient, the carer, the family, the other citizens, the institutions, etc).This represents a key difference with the consumer market segments.

Also the exact connotation of value differs between business models supporting commercial objectives and those associated with e-health. While in the first case the value is linked to a pure financial return, in the case of e-health the overall context varies: the sustainability of social or healthcare systems cannot be measured exclusively according to a financial perspective; it must include 'soft' elements (intangible internal and external elements).

For internal elements we consider the specific benefits associated with activities within the healthcare delivery organizations while in the external elements we include the social benefits, such as quality of life of the persons , peace of mind of the family's members. Often we notice an excessive emphasis to e-health solutions focusing on efficiency instead of value for the patients. In the business world the overall financial assessment is usually described in an S-curve model,

giving an overview of the investments/costs, revenues, profit and maturity during the entire development and implementation phases of the individual system/service.

The structuring of an S-curve is not always possible, especially when a system—such as the healthcare one—is not aimed at providing direct positive financial returns but rather at socio-economic benefits and the sustainability of social or healthcare systems cannot be measured exclusively according to a financial perspective.

9.3.3 *A Challenging Business: The Need of a Collaborative Approach*

As previously mentioned the e-health market is in its early stage; it represents a segment still unconsolidated and with uncertainties (political, regulation, financial conditions, doubtful and long term return of the investments, etc.).

Still, the adoption of e-health solutions has not been as rapid as expected, since the positive financial returns depend on different factors ranging from implementation challenges, evolving legislative and procurement processes to perceptions of the expected positive results among all the involved stakeholders.

Mainly in the initial phase there is the need of sharing opportunities and risks. Collaborative business models can mitigate the risks linked to the business in the e-health sector.

For this reason it has been proposed (Osl et al. 2007) to add in the business model cooperation not as a separate component but rather as an integrated element emphasizing a collaboration perspective within the other items (see Fig. 9.7).

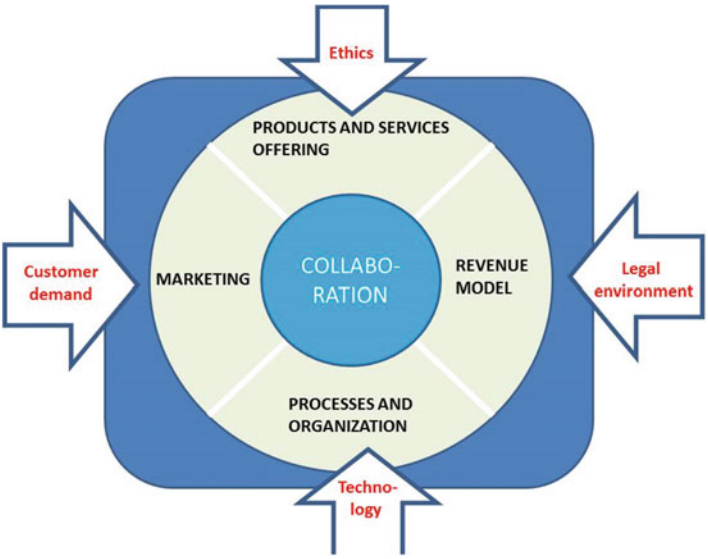


Fig. 9.7 Collaborative business model

In the reported diagram the structure of the business model is completed by including also some important external factors such as the customer demand, the legal and technological environment as well as ethical aspects. Compliance with these external factors highly influences the success of a business model.

In general, collaborative business models need to be in accordance with the strategic goals of participating companies and built according to guidelines such as:

- **Value proposition to the market**: added value for the customers (in our case the citizens) as a result of the cooperation and of the enterprise network including a shift of the complexity from the user/customer to the enterprise network.
- **Revenue model**: win-win situation for all participants and risk dependent revenue distribution.
- **Marketing**: efficient use of the customer bases and the sale channels of all partners, leveraged reputation of participating companies, focus on value for the customer.
- **Processes and organization**: clear distribution of roles and competences, defined and commonly agreed goals, clear decision making structure and defined conflict management process, mutual trust, transparency and open communication and information culture, defined exit possibilities for partners, buffer capacity in resources and competencies (due to high uncertainties in long term cooperation).

In our approach collaboration goes beyond the industrial/business sector and involves all the stakeholders also through the implementation of synergic and harmonized joint initiatives of public institutions and private entities such as User Associations for the delivery of the healthcare services mainly when related to medical education/prevention/social support (community-supported healthcare).

The vision is a participatory model of care that requires public and private sectors, profit and no-profit organizations to work in partnership to create value and sustainability in healthcare and to realize integration between healthcare and social care services.

In the e-health value chain an important role is that of the "mediator" who has a high level of trust (e.g. a medical professional, an expert doctor) and can promote the use of a e-health service by the patients.

9.3.4 An Overview of Possible Business Models in e-Health

The new and challenging market segments in e-health are those not directly linked to the clinical setting such as the personal healthcare and the home-based care (e.g. telemonitoring). Typically the services related to them are not covered by the National Health Organizations and business models have to be based on private or mixed funding.

Among the business models already experienced in other sectors, we analyze here below some of them which could be successful also in the healthcare domain (see Table 9.4).

Table 9.4 Different business models for e-health

Business model	Short description	Example
FREEMIUM	Basic services are offered for free, while a premium is charged for the advanced services	Linkedin, skype
MULTI-SIDED	The service is not funded by the direct users but by organizations having collateral interests	
RAZOR and BLADES	A customer is lured with a special deal; by relying on lock-in effects the profit is made from sales of complementary goods	Printers and ink cartridge
CROWD SOURCING	Commitment of motivated individuals produces value for the organization for free (e.g. crowd creation of contents, crowd funding, etc.)	Wikipedia
Inverted Freemium	Consumers pay a premium for being entitled to consume free services	Insurance
Inverted Razor and Blades	Complementary goods attract the customer and motivate him to buy the core product	iPhone and Apps
AS A SERVICE	Only the usage of the service is charged but not the product itself.	

An application of the FREMIUM model in e-health could be represented by a health community portal and a free social network service connecting patients with similar diseases and allowing them to exchange opinions; as a premium they could subscribe a consulting service with expert doctors.

The assumption is that more people use a service for free, the more valuable the service becomes for all consumers and that greater is the network and higher is the lock-in effect.

A multi-sided business model could be used for telemonitoring and medical education services offered free of charge to the customers of insurance companies and funded by them; in exchange the insurance companies will get a better insight on the health status of their patients and—by fostering health management—reduce also the economic risk.

Another example is represented by a telemonitoring service offered free of change but in exchange of the authorization given to the service provider to sell—in an anonymous way—the collected health data.

9.4 A Socio-Economic Assessment of e-Health

9.4.1 Cost-Benefit Analysis (CBA), Cost-Effectiveness Analysis (CEA) and Cost-Utility Analysis (CUA)

Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) are techniques widely used in various sectors to support decision making processes.

In CBA and CEA the outcomes of a project/program are compared with the involved costs; in CBA benefits and costs are both measured in money, while in

CEA the outcomes are evaluated in non-monetary terms (e.g. in years of life saved, degree of blood pressure lowered, etc.).

In healthcare CEA and CBA are used to assess and control resources allocation; given current problems the continued interest in this matter would not be surprising.

In the period of a decade, the society's principal goal for the health system has shifted from increasing the access to healthcare to controlling the rapidly inflating costs of care.

The dilemma today is in containing costs without scarifying desired benefits.

Moving the healthcare from the hospital to home through the use of the technology represents a promising and important step.

Cost-Benefit Analysis methods were preferred up to 1975; more recently Cost Efficiency Analysis is becoming more popular. It is a preferred approach in healthcare since researchers have not yet been able to quantify the social benefits such as quality-of-life in monetary terms.

Quality-of-Life (QOL) is a subjective parameter and is calculated by using surveys; in healthcare popular tools of Health-related Quality of Life (HRQOL) are the following ones:

- Short form health survey (SF-36) (1993),
- Sickness Impact Profile (1975),
- Symptom States Scales such as the Faces Pain Scale (1988), the Fatigue Scale (1993) or the Functional Independence Measure Scale (1987).

All include emotional, social and physical well being.

A specific case of CEA is the Cost Utility Analysis (CUA) where the costs are compared with benefits in terms of their utility to the patient usually evaluated in quality of life measures.

A challenge is to measure the quality of life of users and the socio-economic benefits of the care systems in a more scientifically acceptable way. There is no clear and formal identification of specific socio-economic indicators of outcomes (both quantitative and qualitative). No consistency was found within and between studies; often comparison of results of one study with another can be misleading.

It was observed (European Commission, ICT PSP 2009) that "solid evidence on common indicators and statistically significant measurement methodologies are essential to convince policy makers and industry to invest further in wide deployment and uptake of solutions".

In 2005 Tsai, Chen and Yin introduced the concepts of "effectiveness score" and "cost-effectiveness ratio (c/e ratio)" to compare home healthcare with usual, hospital-based care. The effectiveness score has a multi-dimensional perspective (social interaction, satisfaction, outcome indicators of the disease, etc.).

Tsai, Chen and Yin estimated a c/e ratio of 4.3 for home healthcare (cost = 1,420$; effectiveness score of 320.8) and 13.5 for the usual (hospital-based) care (cost = 3,209$; effectiveness score of 238.0).

In a similar way Intel and the Center for Aging Services Technologies CAST) highlighted the cost reduction combined with a better quality of care by moving from acute care at the hospital to residential care and finally to home care.

Another important indicator is the "opportunity cost". Due to the limitation of the available resources, the investment in a program causes often the missed opportunity of getting the benefits offered by another program i.e. causes the loss of a cost opportunity. The objective is to minimize the "opportunity cost" or to maximize the benefits of the chosen program.

9.4.2 The Overall Process of a Socio-Economic Assessment

The following diagram shows the overall process for the socio-economic assessment (Fig. 9.8).

The **Cost Analysis** will include direct costs and indirect costs.

Direct Costs to be considered are:

- Costs of the healthcare service,
- Equipment/devices (purchasing cost, amortization, technical duration of the equipment vs. the clinical duration of its use),
- Communication costs (Internet access),
- Service fee,
- Training costs,
- Installation costs,
- Maintenance and technical assistance costs.

An indirect cost could be the time lost for undertaking the treatment.

It is important to perform **discounting** i.e. to adapt the costs to their evolution in time; various discount rates could be considered (both for costs and benefits).

For the **Costs Assessment** it is recommended:

- To identify all the stakeholders and allocate the costs incurred by each of them,
- To include all the costs even if some of them could be hardly quantifiable,
- To list and analyze the uncertainties and perform a "sensitivity analysis" (see below).

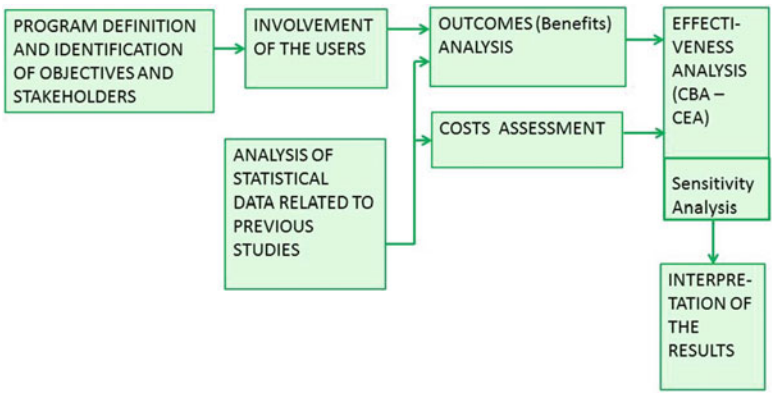


Fig. 9.8 Socio-economic assessment process

The **Benefit Analysis** will follow a similar approach.

The economic impact or the 'savings' that can be attributed to various ICT interventions are of considerable interest to both policy makers and industry. In health care benefits are frequently gained from cost avoidance, cost reduction, control of resource allocation, improvement in service quality rather than from revenue enhancement activities.

9.4.3 The Sensitivity Analysis

Sensitivity analysis is the examination of an uncertain event under different assumptions; applied to CEA it assesses the impact on the cost-effectiveness ratio of varying the baseline assumptions across a range of plausible values. Sensitivity analysis thus provides insight into the stability of the cost-effectiveness ratio, identifies those baseline assumptions that have the greatest impact on the overall costs and defines boundaries beyond which a specific project/program may no longer be cost-effective.

The following ones are some of the aspects to be considered in a sensitive analysis:

- Changing of personal habits and of interactive environmental conditions;
- Financial situation of the users;
- General economic situation blocking the investment of the Governments in healthcare;
- Unforeseeable future technological developments;
- The value of the discount rate;
- Evolution and costs of new healthcare approaches.

It is important to place the results obtained from the analysis in perspective, to examine closely the assumptions on which the analysis is based and to test the sensitivity of the results to reasonable changes in these assumptions.

Uncertainty can be classified into that which is due to random events and that which is due to lack of information (e.g. uncertainty linked to future events over which we have no control); in this last case the best we can do is to examine trends and to use expert opinion.

For random events, probability theory can be used.

When probabilities are not known, expert judgment can be substituted. Thus, without knowing the cause or even the dynamics of a given random process, the analyst can attempt to predict the likelihood of an outcome.

As an example, in the case of discounting, we could have doubts in calculating the precise discount rate and a consensus could never be reached. Under this uncertainty, one logical course for the analyst is to test the sensitivity of the results to several discount rates (a low, high, and middle value).

9.4.4 Weaknesses Inherent to CEA/CBA

Experts evidenced some heavy weaknesses of Cost-Benefit and Cost-Effectiveness analyses in general and specifically of those applied to e-health (Table 9.5):

Table 9.5 Critical issues in CEA/CBA applied to healthcare

Issue	Description
Quantifiable objectives (outcomes)	(a) For a good CEA/CBA the principal objectives should be quantifiable; when that is not possible, reasonable proxies should be available; (b) Unfortunately some objectives are intangible; it is important to avoid the temptation to forget them and to measure only the quantifiable objectives. (c) As an example it is difficult to correctly quantify the costs generated by organizational changes as a result of the adoption of ICT solutions by organizations (e.g. telehealth service).
Technological changes	(a) Costs and benefits of new and not yet existing programs/technologies are difficult to predict; (b) Changes in the technological scenario have to be taken into account.
Immaturity of the methodology	In some aspects CEA and CBA methods are still immature.
Attribution problem	Sometimes it is difficult to separate the changes attributable to the interventions being studied and those that are the result of other influences. It is therefore important to provide evidence that ICT solution makes an independent contribution to the outcome.
Selection/eligibility criteria of trials/ interviews/participants to surveys	Typically volunteers are used (also for ethical reasons); these likely are positive about ICT interventions. The information about refusal rates at point of initial referral is not always recorded; thus it is not possible to take into account the selection bias in the results.
Often Quality of Life is a too generic outcome indicator	(a) There is a general agreement that quality of life requires a comprehensive assessment, meaning that measurement should not be limited to physical and functional capacities alone. Physical health and day-to-day functioning, psychological well-being, social relationships and environment should be taken into account when determining the impact of ICT on Quality of Life. Health status, safety, independence, loneliness, anxiety, relationships with family members are only a few examples of quality of life predictors that ICT may have an impact on. (b) Subjective approach: to take into account the users' feedback raises problems due to the subjective judgments of quality of life made by people involved in trials: for example users' expectations influence appraised quality of life; also individuals may feel constrained because of courtesy or intimidation from actually expressing their views while taking the survey. The intimidation is more likely if the person is in vulnerable health and perceives himself or herself as dependent on care providers.

9.5 Conclusions

As evidenced by several analysts the healthcare system is experiencing a critical phase: demand is dramatically growing both in terms of the enlargement of the healthcare users and in terms of higher expectations with regard to the quality of care.

ICT has the great opportunity to foster the transformation of healthcare toward a "patient-centric" focus and to realize better quality of care even at a lower cost. Nevertheless several barriers have to be removed and often they are not linked to the shortage of suitable and advanced technologies but to political, legal and cultural issues. All the stakeholders (medical professionals, policy makers, institutions, patients and their families, the overall community) need to find the way to cooperate in a common effort.

References

Agency France Press (2005) India to launch satellite exclusively for telemedicine. Agency France Press, 17 March 2005

Ageing Report (2012) Europe needs to prepare for growing older, 2012—European Commission, economical and financial affaires. http://ec.europa.eu/economy-finance/publications

Canada Health Infoway, Emerging Technology Group (ETG) white paper: cloud computing in health (2012) (ETG@infoway-inforoute.ca)

Empirica & WRC (2008) ICT and ageing: European study on users, markets and technologies—preliminary findings: October 2008. www.ict-ageing.eu/Project documents

European Commission, DG SANCO (2009) The contribution of health to the economy in the European Union

European Commission, ICT PSP (2009) Developing indicators and impact assessment methodologies on ICT for ageing well—workshop results, Brussels, 19 Oct 2009

Eysenbach G (2001) What is e-health? J Med Internet Res 3(2):e20. doi:10.2196/jmir.3.2.e20

Gartner, E-health for a healthier Europe!—a study conducted by Gartner on behalf of the Sweden Government. http://www.government.se/content/

HINE (2007) Accelerating the development of the ehealth market in Europe (European Commission, report of the ehealth task force on preparation of the lead market initiative. http://ec.europa.eu/enterprise/policies/innovation

Müller-Stewens G, Lechner C (2005) Strategisches Management—Wie strategische Initiativen zum Wandel führen. Schäffer Poeschel, Stuttgart

Osl P, Sassen S, Oesterle H (2007) A guideline for the design of collaborative business models in the field of ambient assisted living (Deutscher AAL-Kongress mit Ausstellung. http://works.bepress.com/philipp_osl/1)

Pagliari C, Sloan D, Gregor P et al (2005) What is eHealth ?: a scoping exercise to map the field. J Med Internet Res 7(1):e9

Price Waterhouse Coopers (2005) HealthCast 2020—creating a sustainable future: January 2005. www.pwc.com/gx/en/healthcare

OECD StatExtracts (2012) Health expenditure and financing. http://stats.oecd.org. Accessed 23 March 2013

Frost, Sullivan (2013) European market for remote patient monitoring applications: gaining momentum but uncertain times ahead, (8 Jan 2013—M8D5-01)

Transparency Market Research (2013) m-Health market-global industry analysis, market size, share, growth, trends and forecast, 2012–2018: 8 January 2013. Transparency Market Research, Albany, NY

World Health Organization (WHO) (2009) Telemedicine—opportunities and developments in member states (Report on the 2nd global study in e-health). WHO, Geneva

Chapter 10
Case Studies

Evangelos B. Mazomenos, Juan Mario Rodríguez, Carlos Cavero Barca, Gennaro Tartarisco, Giovanni Pioggia, Božidara Cvetković, Simon Kozina, Hristijan Gjoreski, Mitja Lustrek, Hector Solar, Domen Marincic, Jure Lampe, Silvio Bonfiglio, and Koushik Maharatna

Abstract Information and Communication Technologies—as analyzed in this book—could allow a radical change in the way healthcare services are delivered to the citizens and could represent an effective tool to cope with the today's healthcare challenges.

E.B. Mazomenos (✉) • K. Maharatna
School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
e-mail: ebm@ecs.soton.ac.uk; km3@ecs.soton.ac.uk

J.M. Rodríguez • C.C. Barca
ATOS Research and Innovation (ARI), ATOS, Albarracín 25, Madrid 28037, Spain
e-mail: juan.rodriguezp@atos.net; carlos.cavero@atos.net

G. Tartarisco • G. Pioggia
Institute of Clinic Physiology, National Research Council (CNR), Via Moruzzi 1, Pisa 56124, Italy
e-mail: gennaro.tartarisco@ifc.cnr.it; giovanni.pioggia@ifc.cnr.it

B. Cvetković • S. Kozina • H. Gjoreski • M. Lustrek
Department of Intelligent Systems, Jožef Stefan Institute, Jamova cesta 39, Ljubljana SI-1000, Slovenia
e-mail: boza.cvetkovic@ijs.si; simon.kozina@ijs.si; hristijan.gjoreski@ijs.si; Mitja.lustrek@ijs.si

H. Solar
CEIT, Parque Tecnológico de San Sebastián Paseo Mikeletegi, N° 48 20009, Donostia/San Sebastián, Spain
e-mail: hsolar@ceit.es

D. Marincic • J. Lampe
Mobili d.o.o, Abramova 12, Ljubljana 1000, Slovenia
e-mail: domen.marincic@mobili.si; jure.lampe@mobili.si

S. Bonfiglio
FIMI-BARCO, via S. Banfi 1, Saronno 21047, Italy
e-mail: Silvio.Bonfiglio@barco.com

K. Maharatna and S. Bonfiglio (eds.), *Systems Design for Remote Healthcare*,
DOI 10.1007/978-1-4614-8842-2_10,

In this chapter we introduce two European research projects where large part of the concepts addressed in this book are applied; they are the MICHELANGELO project of the seventh Framework Program and CHIRON of the ARTEMIS JU Program.

The CHIRON project (Cyclic and person-centric health management: Integrated approach for home, mobile and clinical environments) focuses on prevention i.e. on a move away from 'health care' towards 'health management', from 'how to treat patients' to 'how to keep people healthy', from a "reactive care" to a "proactive care". CHIRON designed a system's architecture making possible a "continuum of care" i.e. an integrated health management approach in which health is patient-centric at home, in the hospital and in nomadic environments. Care is moved from the hospital to the home and the healthcare staff is enlarged by adding informal carers to the medical professionals and by motivating and empowering the patient himself to manage his own health. Moreover the CHIRON system builds a personalized risk assessment of the patient by integrating personal information, data gathered at home and in a mobile environment through an innovative set of wearable sensors and data available at the hospital including outcomes of image-based tests. The expected results are a reduction of the healthcare costs and a better quality of care.

MICHELANGELO addresses a specific category of patients i.e. the autistic children; the aim is to use ICT to promote and facilitate the assessment of autism within the home setting, away from the traditional clinical environments and to provide personalized "home-based" intervention strategies. This is achieved through the provision of cost-effective, patient-centric home-based intervention remotely controlled by the therapist (remote rehabilitation). The proposed method aims at enhancing the effectiveness of the treatment through its "intensiveness" and "personalization" matching the individual characteristics of autistic children and the involvement of the parents in their "natural" home environment in the role of "co-therapists".

Both projects offer interesting inputs on how Information and Communication Technology could help in "revolutionizing" healthcare. It is worthwhile to highlight that both projects keep the doctors at the core of the healthcare process and in both of them technology is not replacing the experience and the competences of the medical professionals and is not removing the needed physical contact between them and the patients but it supports the doctors in executing their tasks in a more effective and better way.

This chapter is split into two parts: in the first we will introduce the two projects mainly from a strategic perspective in line with the current efforts towards "radical changes" needed to cope with the heavy challenges the healthcare system is facing.

The second part gives a technological insight of the CHIRON project and shows how this project is deploying several of the concepts analyzed in the previous chapters of this book.

This part presents the architecture of an integrated continuous monitoring system for Cardiovascular Disease (CVD) patients in nomadic settings developed under the ARTEMIS-JU CHIRON Project. The proposed sensor platform constitutes of commercially available subsystems effectively integrated into a single multi-sensor non-invasive wearable solution. To enable medical experts to assess the patient's condition remotely, a number of analysis algorithm were developed and implemented into an Android application in order to provide the desired medical information. The key challenge in the development of these algorithmic solutions, was to balance the expected performance while maintaining a low level of power consumption, thus facilitating the continuous monitoring purpose of the system. Furthermore, a web-server based framework provides medical experts with an interactive analysis and monitoring interface and provides the infrastructure for storing the obtained data.

10.1 Part 1: Two European Research Projects for the Healthcare of the Coming Years—The CHIRON and the MICHELANGELO Projects

Silvio Bonfiglio

10.1.1 Introduction

As an effective response to the current challenges in healthcare we need a radical change of the way healthcare services are delivered to the citizens; we need to move the care of the citizens from the hospital to home to take benefit from the availability of informal caregivers and enable a more affordable healthcare service by reducing the need of hospitalization. This will translate into early intervention when the economic impact of the disease is still low.

The previous chapters of this book have shown how advanced ICT-based solutions applied to medicine (what we call e-health) could contribute to the achievement of the ultimate goal i.e. a better quality of care at more affordable costs.

In this chapter we analyze two European research projects representing interesting case studies, the CHIRON project of the ARTEMIS JU Program, and the MICHELANGELO project of the FP7-ICT. Both projects emphasize the active role of patients and citizens, promote a person-centred healthcare with the patient at the core of the overall process and recognize the individual, specific characteristics of each patient.

CHIRON addresses the entire healthcare chain and proposes a new approach and a reference architecture fostering a continuum of care; the MICHELANGELO project focuses on the treatment of children suffering from Autism Spectrum Disorders and introduces ICT and cost-effective, patient-centric home-based intervention remotely controlled by the doctors to promote the assessment and the therapy of autism within the home setting, away from the traditional clinical environment.

The overall chapter is divided in two parts. In the first part the strategic views of the CHIRON and MICHELANGELO projects are described to give the reader a flavor of the possible broad spectrum of applications and how a remote healthcare

S. Bonfiglio (✉)
FIMI-BARCO, via S. Banfi 1, Saronno 21047, Italy
e-mail: Silvio.Bonfiglio@barco.com

system could be applied in practice within that spectrum. The second part is more technology orientated. Here we restrict ourselves within the detailed description of the CHIRON system only because of the fact that it represents an excellent example of an advanced remote healthcare system developed by the joint effort of Industry, Academia and Research Institutes. In this case we start from the basic functional requirements of the system and show in step-by-step fashion how the complete end-to-end system is developed. Without loss of generality, such an approach could be followed for developing any applications orientated remote healthcare system.

10.1.2 The CHIRON Project

Healthcare is facing demographic and socio-economic challenges: from an ever ageing population suffering from chronic diseases and various handicaps to the need for affordable "global" healthcare provided by fewer and fewer professionals and medical infrastructures for critical, often mobile, patients.

Moreover medicine is confronted by the "age of globalization": the movement of people all over the world along with societal and economic developments means that medical professionals have to deal with new diseases, new symptoms and new contextual factors relevant to the health of the citizens, such as changed environments, new nutritional habits and lifestyles.

Healthcare—traditionally focused on institutional care and on curing diseases (diagnosis, treatment, rehabilitation)—needs to shift towards monitoring, risk assessment and early detection to effectively manage chronic diseases and to reduce complications at least. We need a shift from "health care" (how to treat patients) to "health management" (how to keep people healthy).

CHIRON (Cyclic and person-centric health management: Integrated approach for home, mobile and clinical environments) is a European research project of the ARTEMIS JU Program and is addressing these issues; it has received funding from the ARTEMIS Joint Undertaking under grant agreement #100228 and from the National Governments participating to the project.

Twenty-seven Organizations from eight European member states and representing the industry (large companies and SMEs), the research and the academic communities and medical institutions are involved.[1] Having started in March 2010, the project has a term of 45 months and a budget of 18 million Euro.

[1] The following Organizations are participating at the CHIRON project: FIMI Srl, Universities of Bologna, Genova, Roma, Trieste, CNR, I + Srl, ITS, SELEX ES, W-LAB Srl from Italy; ALMA, ATOS, CEIT, CIMNE, Ibermatica,Tecnalia from Spain; Athena ISI and Intracom from Greece; BARCO from Belgium; Philips Healthcare and ZorgGermak from The Netherlands; University of Budapest from Hungary; Jozef Stefan Institute and Mobili from Slovenia; Cardionetics, Southampton University Hospitals NHS Trust and University of Southampton from United Kingdom.

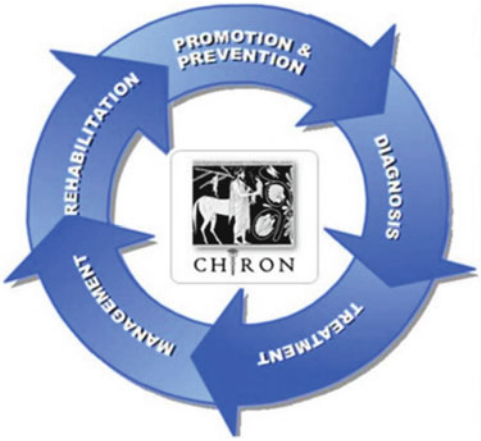


Fig. 10.1 CHIRON addresses the complete healthcare cycle

The project addresses one of the today's societal challenges i.e. "effective and affordable healthcare and wellbeing". CHIRON combines state-of-the art technologies and innovative solutions into an integrated framework of embedded systems for effective and person-centred health management throughout the complete healthcare cycle, from primary prevention (person still healthy) to secondary prevention (initial symptoms or discomfort) and tertiary prevention (disease diagnosis, treatment and rehabilitation) in various domains: at home, in nomadic environments and in clinical settings (Fig. 10.1).

It applies several of the approaches and techniques discussed in the previous sections of this book.

In currently applied healthcare practice, a disease is commonly discovered after symptoms have emerged. Only then, people become patients and apply for medical care, a diagnosis is made and a treatment is proposed and initiated. CHIRON fosters a change in this approach i.e. a shift towards an earlier diagnosis based on the risk assessment of the still healthy person (Fig. 10.2).

In CHIRON, detection and diagnosis are based on the analysis of heterogeneous medical data that originate from personal data of the patient, medical past knowledge, imaging systems, collection of vital sign data through sensors and handheld devices. Medical doctors having remote access to the person's health record are able to take a more proactive role in prevention of diseases.

This is a significant innovation compared to the currently used approaches that are based on health-related parameters only, often monitored *instantaneously* and *episodically* (during clinical center visits). In addition CHIRON intends to explore the correlation across physiological, psycho-emotional and environmental parameters.

The CHIRON system involves modeling at different structural and functional scales, combining inputs from sensor data, medical imaging, clinical diagnoses,

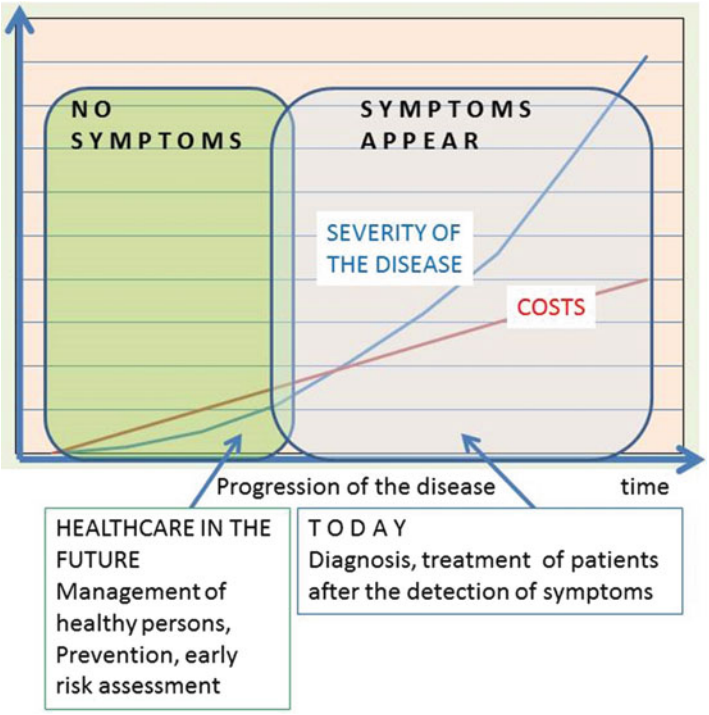


Fig. 10.2 From treatment of patients to early risk assessment and management of still healthy persons

patient history and population-level statistical information. Continuous and close-to-the patient remote monitoring of physiological and psycho-emotional state and trends are integrated and contextualized with environmental parameters, information on patient activity and lifestyle related factors. The CHIRON reference architecture ensures interoperability between heterogeneous devices and services, a reliable and secure patient data management and a seamless integration with the clinical workflow.

A new way of looking at impending risk is constructed, based on simple measures obtained on a continuous basis with time-frame acquisitions, processing and relation with future events depending on the type of variable considered: heart rate, skin and ambient temperature and humidity, daily weight, water and urine variations and in house movements (Puddu et al. 2012).

It is worthwhile to notice that such a correlation model does not exist in practice; presently diagnosis of a disease depends largely on the experience of the individual doctors. The typical models for diseases only predict mortality or morbidity in long term but never identify "when actually" the person is at risk.

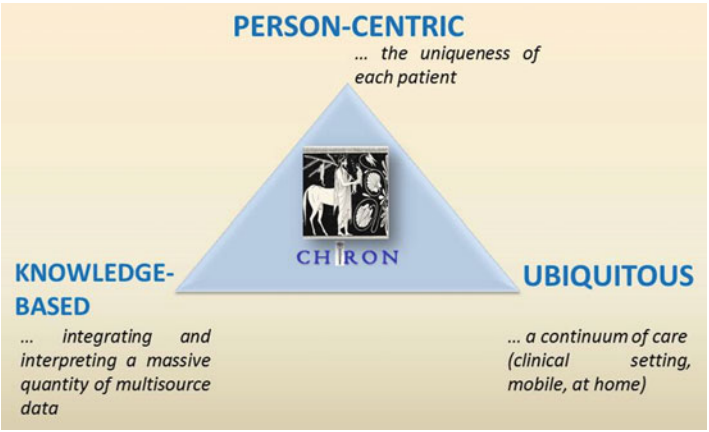


Fig. 10.3 The basic concepts of the CHIRON approach

10.1.2.1 The CHIRON Concepts

In the field of e-health CHIRON represents an interesting case study since it applies innovative approaches made possible by advanced ICT solutions of which some already addressed in this book. More specifically the CHIRON research work is based on three main concepts (Bonfiglio 2010; ARTEMIS Joint Undertaking 2010):

(a) The realization of a "continuum of care" i.e., addressing in an integrated approach health management at home, in the hospital and in a nomadic environment.
(b) A "person-centred approach" where the needs of the citizens, the medical professionals and the whole community are at the core of the design and the uniqueness of each user is fully recognized ("personalized healthcare") including the use of proactive solutions anticipating the needs of people, adapting themselves to context and taking the required actions and thus enriching quality of life and fostering independence.
(c) A knowledge-based system integrating past and current data of each patient together with statistical data related to the whole community in a large and distributed repository (in CHIRON called *Virtual Data Repository*) and promoting a move towards prevention and early risk assessment (Fig. 10.3).

A. Care at home and everywhere ("continuum of care")

A continuous multi-parametric monitoring of physiological and psycho-emotional state, environmental parameters, patient activity and lifestyle related factors has been realized and the best compromise between unobtrusiveness and accuracy of the gathered data was pursued in designing it.

Measurements provided by the close-to-the patient remote monitoring are incorporated into physiological models and fused with anatomical and

functional information derived from MRI and 3D ultrasound data, from which organ-level models are built and updated.

All these data—integrated with those available in the Hospital Information System—contribute to the building of an evolving Patient Health Profile and to the definition of a personalized and constantly updated risk assessment model (identified in CHIRON as *Alter Ego*).

In the CHIRON architecture, high emphasis is given to the seamless integration of the personal health platform with the clinical workflow and the integration of the personal and medical data gathered in a "non-clinical setting" into the electronic health record of the patient. Adherences to standards with regard to the exchange of the medical information (HL7 v.3) as well as reliable and secure patient data management are key issues. Semantic interoperability is realized.

B. A person-centric healthcare management

CHIRON provides personalization in the delivery of the healthcare services. The "one size fits all" paradigm will no longer apply, especially when it comes to medicine. Personalization will also feature in a coaching system to help the patient to reduce immediate risk and improve long term recovery. Patients are empowered and motivated to self-manage their wellbeing. This person-centred approach puts the needs of the citizens, the medical professionals and the whole community at the core of the CHIRON design. Health is considered a value the single person and the whole community have to build and preserve.

The user profile represents an evolving model of the relevant aspects related to the health of a user and provides a continuously up-to-date risk assessment of his health status.

C. A knowledge-based medicine—Support to the doctors

To make a diagnosis, to define the most suited treatment for the patient, the physician needs to make use of a massive quantity of data, aggregated from different large data sets, interpreted and integrated with community related statistical data and past knowledge. Moreover he needs to take timely decisions.

All these data represent the outputs of multiple and heterogeneous devices and subsystems ubiquitously distributed: physiological signs and activity information are gathered through unobtrusive on-body sensors and are contextualized with environmental parameters measured by discrete devices; results of lab analyses including image-based examinations are available in the information systems and PACS systems of the clinical centres and hospitals, epidemiological/statistical datasets and past knowledge are dispersed everywhere all across the world and sometime organized and stored in various repositories. In CHIRON all of them—integrated in the *Virtual Data Repository* (*VDR*)—provide medical professionals with an effective support for early diagnosis and personalized treatment planning.

A critical aspect CHIRON addresses, is the effectiveness and the intuitiveness of the feedback provided to the medical professionals, avoiding the risk of overloading the doctors with a plethora of data. CHIRON interprets and

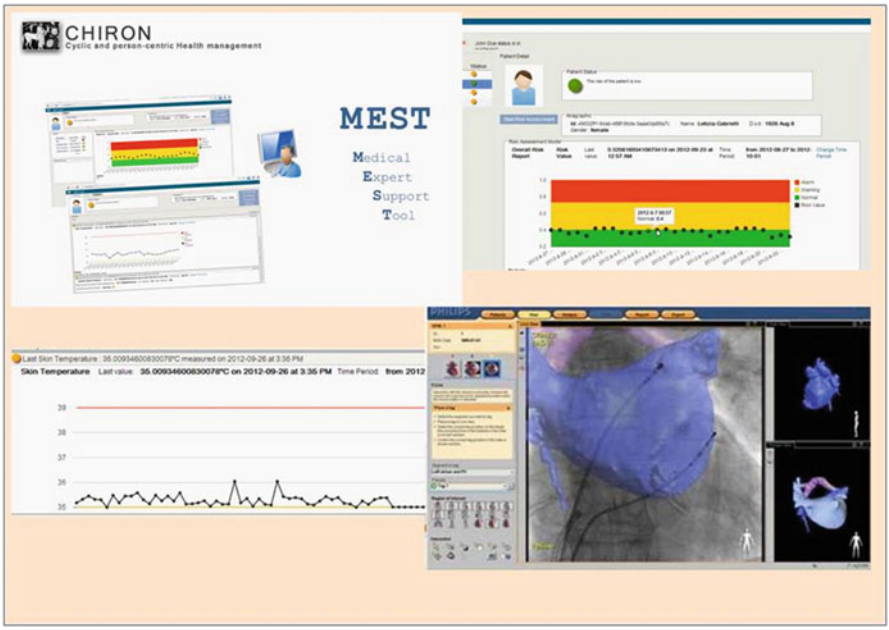


Fig. 10.4 Intuitive and effective feedback to the doctor

translated those data into features. An ontological approach is used to retrieve information from multiple distributed content repositories and present it in a structured and aggregated manner giving the doctor an easy and powerful tool for immediate understanding of the status of the patient (Fig. 10.4).

We are aware that technology cannot replace the experience of the medical doctors but can support them in taking the most appropriate and timely decision. In CHIRON the doctor remains the protagonist of the healthcare process and the only owner of the clinical decision.

10.1.2.2 The CHIRON Architecture

In the CHIRON architecture all the devices/subsystems generating, interpreting and storing data are embedded modules of an overall complex system; each of them has the computing and memory capabilities to execute a specific task for which the outputs of other sub-systems are needed. The task of each of these embedded blocks could be very simple such as processing the raw data, removing artefacts and extracting features. Tasks may also be complex such as that of a medical image acquisition system or an image-guided surgical system.

In contrast to the typical architectures and keeping in mind the energy constraint of a sensor node, specialized low-complexity algorithms for on-board signal

denoising, artifact removal, intelligent classifier and feature extraction were realized for each of the electrophysiology sensors. It was done without compromising the medically trustworthiness quality of the captured signals.

In this network of embedded systems Internet provides the communication infrastructure for the exchange of information among them and the weaknesses of Internet in terms of quality, reliability and real time interchange of the information are compensated by a higher autonomy and robustness of each of the modules of the system. Each block is capable of operating independably even in the presence of network degradation or temporary failure and is able of executing a minimal, essential task autonomously (e.g. the prompt warning of the patient if a vital sign is above the allowed threshold or the activation of an emergency call in the case of a critical event).

In this regard, the CHIRON project realizes the vision driving the European ARTEMIS JU Program i.e. a world in which smart "objects" have "a presence in cyber space, exploit the digital information and services around them, communicate with each other, with the environment and with people, and manage their resources autonomously" (ARTEMIS Joint Undertaking 2011). CHIRON intends to contribute to the realization of this vision in a so important and complex domain such as the healthcare one. Being the citizen and his quality of life at the core of the healthcare model proposed by the project, the CHIRON system—even if specifically designed for healthcare—interoperates across many application domains. CHIRON uses a middleware which allows to achieve not only interoperability between heterogeneous devices and services and exchange of multisource information across the whole healthcare cycle but also a multi-domain compatibility.

10.1.2.3 Technological Challenge and Innovation in CHIRON

The project promises to progress beyond state-of-the art in several disciplines (Bonfiglio 2011) such as healthcare architecture, sensor nodes with enhanced capabilities, new methods and algorithms for the measurement of physiological parameters, advanced solutions for the analysis and the visualization of medical images.

The technological challenges of the full integration of a Personal Health System into the conventional hospital-based Healthcare System are related to the realization of the best synergy among conflicting requirements such as:

- Data gathering and data processing in resource-constrained systems (the wireless mobile sensors network),
- Non-invasiveness of the monitoring combined with accuracy of the measured data,
- Continuous monitoring vs. the acceptance of the solution by the patients,
- Automation or minimal human intervention vs. the reliability of the feedback,

- Availability of a massive quantity of data vs. the easy and fast interpretation of these data by the medical professional,
- Privacy and security of data vs. easy and ubiquitous access to these data by authorized persons, etc.

The advanced solutions proposed by CHIRON represent a solid contribution to solve these issues. More specifically CHIRON developed:

(a) Sensor nodes with enhanced capabilities (multi-parametric monitoring, local processing, energy harvesting and power management),
(b) Innovative methods for the measurement of biological parameters (ECG, blood pressure with enhanced accuracy, serum potassium concentration in the blood),
(c) Advanced solutions for the analysis and the visualization of medical images,
(d) A dynamic and personalized risk assessment model (*Alter Ego*),
(e) A distributed data storage approach (Patient's Virtual Data Repository) with enhanced data security and privacy provisions.

In the current age of image-centred medicine the new advanced tools developed in the CHIRON project are expected to facilitate real time processing, computer-aided detection and accurate visualization of medical image as well as to support doctors in making accurate diagnosis by reducing the risk of a "false negative" or the need for additional and costly examinations due to a "false positive".

10.1.2.4 Conclusions

The CHIRON research project contributes in promoting a new way of delivering healthcare services and promotes a paradigm shift from diagnosis and treatment of patients based on symptoms to diagnosis of patients based on risk assessment of healthy persons.

Moreover CHIRON intends to contribute to the uptake of a market still in its nascent phase, the e-health one, identified by the European Commission as a "*lead market*" where "innovative products/services/solutions have high growth potential, where EU industry can develop competitive advantage and where action by the public authorities to deal with regulatory obstacles is needed".

CHIRON contributes to bridge the gap between R&D and market readiness.

The healthcare societal challenges will promote the creation of new and innovative services/products and the solutions and the technological advances made in CHIRON will contribute to the uptake of applications such as medical imaging, personal healthcare (home and mobile monitoring), health management and coaching, wellbeing, clinical decision support systems and healthcare informatics.

The move towards "global, open networks of multi-domain embedded systems" will lead to radical changes in the business processes, by asking for cross-company collaborations.

10.1.3 The MICHELANGELO Project

10.1.3.1 Background

The MICHELANGELO project is co-financed by the European Commission (Framework Program 7—Grant Agreement # 288241). It aims to use emerging technologies to design personalized home-based intervention strategies for children with Autism Spectrum Disorders. MICHELANGELO brings together an interdisciplinary team of eight partner organizations in Italy, Malta, United Kingdom and France and bridges academia, industry, research and clinical practice.[2]

In the diagnostic manual used to classify disabilities (the DSM-IV of the American Psychiatric Association), 'Autistic Disorder' (AD) is listed as a category under the heading of 'Pervasive Developmental Disorders (PDD)'. It affects thought, perception, attention and social interaction and—rather than one disorder with a well-defined set of symptoms—it represents a broad spectrum of disorders that ranges from mild to severe (Autistic Spectrum Disorders—ASD).

According to a recent study (Pediatrics 2009) the prevalence of Autism Spectrum Disorders (ASD) in the young population has increased in USA from the 0.6–0.7 % of the previous 2003 estimate in excess of 1 % (i.e. from 1 in every 150 children to 1 in every 91 children). A similar situation has been reported in Europe where the prevalence of autism in children increased from the 0.5–0.7 % (Atladottier et al. 2007) to 1.16 % as reported by Kuehn (2007) and by several other studies. These trends have considerable social and economic impacts. The average lifetime costs for a person with autism were evaluated at approximately 5.7€ million for people with low functioning autism and 1.5€ million for people with high functioning autism (Knapp et al. 2007).

Consolidated data on the prevalence of autism in Europe is lacking. Variability in reported data is often due to methodological factors; the result is that 'there is no comprehensive or comparable data at EU level concerning the incidence or prevalence of this disease' (report of the European Commission dated 2004). In reviewing various prevalence studies, Williams JG at al. found that 50 % of the variation among study estimates could be explained by factors such as age of the screened children, diagnostic criteria used, country, rural or urban location, etc.

Equally in Europe the social and economic burden of ASD has not been adequately recorded, as epidemiological figures are unreliable and inconsistent. In a recent report, the economic consequences of autism in the UK were calculated and the findings reveal that children with autism cost £2.7 billion (3.8€ billion) annually, yet for adults the figure is £25 billion (36.2€ billion)—over eight times as much.

[2] The following Organizations are participating to the MICHELANGELO Projects : FIMI S.r.l., Italian National Research Council (CNR), Fondazione Stella Maris and I+ Srl (Italy); University of Southampton and University of Ulster (United Kingdom); University Pierre et Marie Curie, and Assistance Publique-Hopitaux de Paris (France); Across Limits Ltd. (Malta).

It is well accepted in the scientific community the importance of an early diagnosis of autism and that prognosis is greatly improved if a child is placed into an early intensive and highly structured educational program by age of 3 years old. Early treatment is a crucial step to ameliorate some of the symptoms associated with autism and some investigators have reported that this is a critical developmental period in a child's life during which brain's plasticity is maximal and environmental influences can interfere with neural connections (Dawson 2008; Oberman et al. 2008; Altemeier et al. 2009; Dawson et al. 2010). Moreover, researchers agree on the inadequacy of one single treatment approach for all the children with ASD due to the heterogeneity and the developmental nature of the disorder. Earlier identification of children with ASD and the development of personalized and evolving protocols for intervention could increase the effectiveness of the treatment. Both these objectives could be achieved with the support of the technology as developed in the MICHELANGELO project. Moreover researchers are starting to develop **prosthetic technologies** able to help children with autism by supplementing or replacing the natural capacities. Advances in key technology areas such as affective computing and wearable computing offer a lot of potential for computer-based prosthetic tools.

10.1.3.2 The Michelangelo Project: Concepts and Objectives

Typically, the assessment of the status of the child with autism and the delivery of a number of therapeutic interventions are executed within a clinical setting. Often this approach implies a reduced effectiveness mainly in consideration of:

(a) The "artificial" context of the lab-based environment, which generates the potential for artifacts and adds systematic and non-systematic bias to the findings, therefore potentially producing results that do not reflect behaviors in real life.
(b) The lack of "intensiveness" of the therapeutic intervention (just a few hours per week) limiting its beneficial effects, as demonstrated in some recent studies (Remington et al. 2007).
(c) The poor "individualization" or "personalization" of the intervention protocol, due to a lack of the *a priori* knowledge of which treatment method will be most effective for a specific child.

The MICHELANGELO Project aims to address and overcome these aforementioned considerations by moving, as much as possible, the assessment and the therapeutic interventions from a clinical setting to a more "natural" home environment and by using non-obtrusive or minimally invasive techniques. As an example, quantitative EEG (QEEG) analysis has proved its capability of identifying dysfunction in various regions of the brain in children with autism. Electroencephalogram (EEG) measures brain waves; it shows the variations in electrical potentials at a number of scalp sites. Inside the brain neurons produce their own electrical fields; it is thought that an unhealthy brain will have large changes in the electrical potential

compared to the potentials produced by a healthy brain. A brain map (quantitative QEEG) could give useful information.

Unfortunately the invasiveness of the currently used systems induces both systematic and non-systematic biases to the experimental outcome eluding the actual nature of the brainwave behaviour and connectivity.

MICHELANGELO intends to minimise the biases and modulation effects by making the recording system pervasive in nature so that the patients becomes "unaware" of its presence; it will allow the patients to move freely in their "own" environment while at the same time offering them naturally occurring stimuli and continuously monitoring their EEG activities.

In the pervasive wearable EEG system (a cap of just 60 g weight) being used within MICHELANGELO, the risk for poor accuracy due to the limited number of the employed electrodes (19 electrodes), and to the portability of the system is reduced as much as possible by calibrating and adjusting the system according to the specific clinical child to be monitored. Throughout the Project, wearable EEG recordings has been correlated with those of a dense-array EEG system used as "golden standard" and the wearable model adapted and "customized" through the optimization of the processing algorithms. In this way, a suitable characterization of the brain connectivity of the child and a timely detection of abnormal brain waves activity become possible.

Similar studies (Laibow et al. 2001) have reported a change of parameters such as heart rate, systolic and diastolic pressure, pulse rate, sweat index, body temperature, fingertip temperature, electro-dermal activity (EDA) as an effect of the progress in the rehabilitation program of patients with brain injury. In the MICHELANGELO 'system' a set of wearable and non-invasive sensors will facilitate the monitoring of ECG parameters also.

Equally important to the quality and accuracy of the interventions and therapies is the fact that the MICHELANGELO approach will be extremely cost effective as a result of the limited involvement of the medical professionals in spite of the intensiveness of the treatment.

Another important objective of the MICHELANGELO Project is to open new opportunities in the field of 'personalized' autism research and treatment. Although most researchers agree on the fact that it is necessary to establish an individualized treatment for each child, very little is known about how to individualize treatment protocols, there have been very few studies in which different methods have been validated and compared and there is a little knowledge on how to determine *a priori* which intervention is more likely to benefit individual children.

ICT can provide a rich understanding of the requirements of each individual and allows to monitor them as they develop. It is about realizing a system that recognizes that every child is unique and exhibits unique symptoms and needs and that interventions and therapies need to be tailored according to that.

The MICHELANGELO Project aims to increase the knowledge in these fields by comparing the effect of different methods.

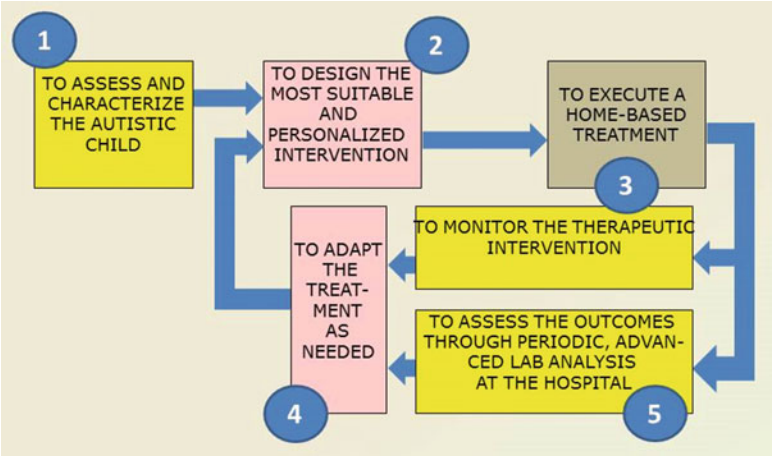


Fig. 10.5 The MICHELANGELO approach

10.1.3.3 The Five Steps of the Michelangelo Approach

In the diagnosis, characterization and treatment of the children with autism the Project proposes an approach that will include the following core steps (Fig. 10.5):

1. The characterization of the child with autism

In the specialised medical center a multimodal and unobtrusive system including video observation, audio signal analysis, QEEG and other physiological signals monitoring (such as heart rate) synchronized all together is used in a controlled environment (the "*Michelangelo room*") to observe reactive behaviours of children with autism. The developed system allows the analysis and annotation (automatic, semi-automatic or manual annotations) of recorded video captured during the observations. Advanced solutions allow to maximize the automation of the system in detecting "events" i.e. behaviours of the observed child indicating a response to the proposed stimuli. The annotations referring to specific behaviours of a child are correlated to changes in the characteristics of the captured signals. In such a way personalized indicators of the brainwave functionalities under different naturally occurring stimuli are identified. Real time feedbacks are given to the therapist to allow him to adapt—when needed—the therapeutic approach and the interaction with the child during the session.

Furthermore multi-modal coordination and interactional synchrony will be analyzed to get useful insights into the level of rapport between social partners. We employ both audio and video analysis to monitor the rhythm and the convergence of actions (including communicative and motor functions). The system enables to evaluate statistically significant correlations between actions (from both intra- and inter-personal points of view) and produces an interpretable representation of the coordination during interactions (Delaherche and Chetouani 2010).

The stimuli able to generate a response in the child's brain are analyzed further on through the combined use of the wearable Q-EEG system and of an eye-tracking system. Each one of them is presented to the child on the screen of a PC, equipped with an eye tracker which tracks the movement of the child's eyes and records their fixations. At the same time, the brain wave activities are recorded via Q-EEG. Different image processing schemes can be initiated on the acquired scenes and this will make it possible to determine the nature of the stimulus resulting in the brain wave behaviour.

2. Developing an appropriate patient-centric intervention protocol

Once the above mentioned activities are completed, the findings can be used for developing appropriate patient-centric intervention protocol. MICHELANGELO intends to explore the benefits of a heterogeneous intervention strategy combining neuro-feedback and developmental/behavioural interventions.

Joint attention and imitation tasks are incorporated in a set of serious games with a level of difficulty that evolves depending on the progress of the intervention and on the response of the child. The project is developing a methodology for easy adaptation and personalization of the games.

3. Therapeutic intervention in a natural environment and the unobtrusive observation of the child

The child is monitored during his/her normal daily activities, moving freely within a controlled environment while wearing the "QEEG cap" and the other physiological sensors. The QEEG system detects abnormal brain waves and trigger a microcamera integrated into the QEEG cap with the objective of tracking the specific objects, situations generating this specific brain activity or a change in the monitored physiological parameters. Data monitored at home are transmitted periodically to the remote Medical Center for further analysis.

Serious games are proposed to the child and the quality of the execution of the planned tasks is automatically assessed by the system and made available to the doctor.

4. Adaptation of the therapeutic intervention

On the basis of the gathered information, the doctor assesses the intermediate outcomes of the therapeutic intervention and adapts it by giving information and training about the change to the parents of the child. In addition to the periodic sessions at the medical center, remote meetings are planned between the therapist and the parents to share observations, opinions and to strengthen the cooperation between them.

5. Periodic assessment at hospital

In MICHELANGELO the "at-home monitoring" is combined with periodic advanced lab examinations. The project plans to develop new methods for a better analysis of the anatomical and functional brain's connectivity based on magnetic resonance images (DTI and fMRI techniques). They allow a better assessment of the evolution of the disorder and of the effect of the therapy.

10.1.3.4 Expected Advance Beyond the State-of-the Art

The MICHELANGELO Project intends to go beyond the state-of-the art in the assessment of ASD and the intervention protocols by making possible a remote and effective "at-home management" of the autistic children. The recent introduction of new equipment such as QEEG-, eye-tracking-, diffusion tensor imaging-based systems and of wearable biomedical devices is providing differential frames to seize invaluable insights, to test hypotheses and collect data that shed new light into the nature of typical and atypical developmental processes of social cognition.

MICHELANGELO project will contribute by introducing advanced solutions such as:

- A pervasive wearable EEG system in conjunction with an eye-tracking system to record brain wave behaviour when the patient is presented with a number of naturally occurring stimuli. According to our knowledge it is the first time that these technologies are used together and that eye tracking is used to better define the therapeutic intervention rather than to assess and/or diagnose the disorder;
- A wearable, unobtrusive multi-sensorial platform for the continuous monitoring of vital signs able to detect behavioural changes in the autistic children;
- An audio- and video-based system monitoring interactive patterns of the child at both motor and communicative levels to detect multi-modal coordination and interactional synchrony and get useful insights into the level of rapport between social partners;
- A multimodal solution for the patient's treatment where neuro-feedback and developmental/behavioural intervention techniques are used at home and designed in such way to combine an intensive intervention and the avoidance of heavy involvement of the clinicians and therefore to be cost effective;
- A set of specifically developed serious games and techniques for the assessment of the quality of execution of the game and for the adaptation of its level of difficulty and its personalization;
- A personalization of the overall therapeutic intervention and its timely adaptation—in a closed loop approach—according to the feedbacks from the autistic child and the intermediate outcomes of the intervention.

Moreover MICHELANGELO intends to explore and to assess emerging ICT-based therapeutic techniques such as 3D and virtual reality, computer vision, and robotics-based methods.

Progress beyond the state of the art is expected also in specific technological areas such as:

(a) An advanced system for the analysis and the annotation of recorded videos and synchronized QEEG, physiological and audio signals.
(b) Sophisticated signal processing algorithms to remove the redundancies present in the traditional Principal Component Analysis (PCA) used to characterize the brain connectivity.

(c) A decision support system helping the therapist in assessing the status of the disorder and the response of the child to the presented stimuli and in adapting and personalizing the therapeutic intervention accordingly.
(d) Advanced techniques and image processing solutions for the analysis of anatomical and functional brain connectivity made respectively with Diffusion Tensor Imaging (DTI) and Functional Magnetic Resonance Imaging (fMRI). The use of these techniques for the assessment of Autism Spectrum Disorders and specifically applied to autistic children represents a new field of research.

10.1.3.5 Medical, Social and Economic Impact

A. Medical Impact

The MICHELANGELO project aims at improving the management of ASD in children by transferring the treatment from the clinical setting to the more comfortable environment of their own homes. There, an intensive and monitored treatment and more "natural" conditions will ensure a more accurate assessment of the evolution of the disease and will allow a timely adaptation and personalization of the therapeutic treatment. The clinical exploratory study planned in the last year of the project will provide a rich set of information that will enhance the medical knowledge related to ADS and to the role ICT could play in autism. It will include the use of state-of-the art ICT solutions (3D, virtual reality, robotics) for enhanced developmental/behavioral interventions combined with neurofeedback-based protocols. Moreover, the outcomes of the project will contribute to a research in the ASD field that is still in its infancy i.e. the "a priori" definition of the right treatment for each individual patient ("individualization of the treatment") and will allow to collect a large quantity of data related to ASD.

B. Social Impact

The project intends to contribute to the management of ASD, a disease with serious socio-economic impacts and heavy consequences affecting not only the patient but all the family members. The MICHELANGELO solution will enable children with autism to stay in the comfort of their homes, thus reducing costs related to therapeutic interventions in external infrastructures with clear benefits for the families and the overall community. The final objective will be to enhance the chances for a gradual rehabilitation of the autistic child and his inclusion in the community. The user-centric approach of the MICHELANGELO Project will foster a user-friendly design and minimise the intrusiveness of technology. By transferring the therapeutic intervention from the hospital to home—it emphasizes the role of the parents and of the educators and fosters the cooperation between them and the medical professionals; they will participate all together to the care process of the autistic child.

C. Economic Impact

The MICHELANGELO approach combines intensiveness of the treatment (no longer limited to few hours a week at the therapist's office or at the hospital) with cost effectiveness due to the automation of the treatment asking for a minimal intervention by the doctors, even if they are enabled to seamlessly supervise, follow-up and monitor the autistic children during the evolution of their therapeutic program at home. The MICHELANGELO project helps to put in place early therapeutic interventions and it will result in cost saving. As it was observed in the EAIS (European Autism Information System) final report (European Autism Information System (EAIS) (2006–2008)) "the latter is the intervention in the affected child's life, the greater the time and costs involved in providing healthcare and support services".

10.1.3.6 Conclusions

The MICHELANGELO project addresses main societal needs related to health as highlighted by the European Commission such as "affordable healthcare through cost reduction, new technologies for patient monitoring, diagnostics and treatment that allow for better, faster and more cost effective care". It opens new horizons and exploits ICT and other technologies in assessing and treating pervasive developmental disorders with specific focus on autistic children.

10.2 Part 2: The CHIRON Project—A Technological Insight

Evangelos B. Mazomenos, Juan Mario Rodríguez, Carlos Cavero Barca, Gennaro Tartarisco, Giovanni Pioggia, Božidara Cvetković, Simon Kozina, Hristijan Gjoreski, Mitja Lustrek, Hector Solar, Domen Marincic, Jure Lampe, and Koushik Maharatna

10.2.1 Introducing the Proactive Approach in CVD

World Health Organization (WHO) reports indicate that Cardiovascular diseases (CVD) is the most common cause of death in the world accounting for 30 % of all-cause deaths. Recent estimates predict that CVD dominance will continue to raise leading to over 20 million deaths in 2015 (Frost and Sullivan 2009). The already high associated cost for the treatment and long-term care of CVD patients is also expected to follow an incremental trend. As an example the cost for

E.B. Mazomenos (✉) • K. Maharatna
School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
e-mail: ebm@ecs.soton.ac.uk; km3@ecs.soton.ac.uk

J.M. Rodríguez • C.C. Barca
ATOS Research and Innovation (ARI), ATOS, Albarracín 25, Madrid 28037, Spain
e-mail: juan.rodriguezp@atos.net; carlos.cavero@atos.net

G. Tartarisco • G. Pioggia
Institute of Clinic Physiology, National Research Council (CNR), Via Moruzzi 1, Pisa 56124, Italy
e-mail: gennaro.tartarisco@ifc.cnr.it; giovanni.pioggia@ifc.cnr.it

B. Cvetković • S. Kozina • H. Gjoreski • M. Lustrek
Department of Intelligent Systems, Jožef Stefan Institute, Jamova cesta 39, Ljubljana SI-1000, Slovenia
e-mail: boza.cvetkovic@ijs.si; simon.kozina@ijs.si; hristijan.gjoreski@ijs.si; Mitja.lustrek@ijs.si

H. Solar
CEIT, Parque Tecnológico de San Sebastián Paseo Mikeletegi, N° 48 20009, Donostia/San Sebastián, Spain
e-mail: hsolar@ceit.es

D. Marincic • J. Lampe
Mobili d.o.o, Abramova 12, Ljubljana 1000, Slovenia
e-mail: domen.marincic@mobili.si; jure.lampe@mobili.si

CVD-related treatment in the E.U. countries is 169€ billion per year, of which 62 % accounts for direct healthcare costs (Dickstein et al. 2008).

By taking into account the related productivity loss and informal care represent 21 and 17 % of the overall CVD costs respectively it is evident that CVD poses a formidable socio-economic burden for developed countries. Due to these facts, current conventional care delivery systems are under serious strain with the risk of reaching unsustainability in the future unless a pioneering shift in the care delivery approach of disease management is initiated. The concept of *Proactive Care* instead of the conventional *Reactive Care* approach has been highlighted as the most effective model for care delivery in years to come which will enable pre-emptive medical intervention and management by prognosticating an impending episode even before the full manifestation of the symptoms, leading to a significant decrease in death and hospitalization rates.

For the implementation of such a proactive care system, capturing, processing and information fusion of heterogeneous physiological data on a continuous manner in nomadic settings, is a key element. This will be followed by the constant interpretation of the acquired data in a "context-aware" way utilizing established medical knowledge. Recent achievements in Information and Communication Technology (ICT), in particular in the area of Wireless Sensor Networks (WSN), mobile communication and sensor technology have the potential of developing the necessary technological background of such a remote healthcare system that is capable of delivering cost-effective high-quality ubiquitous healthcare. The key challenge in the design of such a system is associated with the resource constrained nature of wireless embedded nodes, which are typically battery-powered devices with only basic processing capabilities. On the contrary, the requirements of a remote healthcare monitoring system involve significant computational tasks, since almost continuous radio communication and signal processing of data is essential. Hence, in order to achieve long-term, uninterrupted, operation of such a system the limited resources of the sensor platforms must be utilized effectively.

This chapter discusses the development and evaluation of such a remote healthcare solution focused on the uninterrupted monitoring of CVD patients. The entire system was developed under the Cyclic and person-centric Health management: Integrated appRoach for hOme mobile and clinical environments (CHIRON) Project. The proposed system constitutes of on-body Electrocardiogram (ECG), temperature, sweating index and activity sensors while an Android-based smart phone acts as the communication gateway and computational platform. In order to facilitate the desired long-term system operation, the designed algorithms for the analysis of the physiological data are characterised by their low-complexity which maximizes the system's lifetime. Dedicated firmware has been developed for implementing the sensors operation and their communication with the smartphone.

According to a review on remote monitoring solutions for CVD patients previous approaches present several limitations (Klersy et al. 2009). The majority of them makes extensive use of home devices, or ignores parameters that have been medically established as indicators of an impending critical cardiovascular episode, like skin sweating or ST-segment shift with heart rate changes. These documented shortcomings prompted the development of the CHIRON remote monitoring system.

In the rest of this chapter, the clinical requirements of the proposed system are initially presented in Sect. 10.2.2, with the overall system architecture analysed in Sect. 10.2.3. The design of the individual system components are described in Sect. 10.2.4 while the system integration and its interaction with the Information processing platform are discussed in Sect. 10.2.5. The Medical Analysis and Storage subsystem is analyzed in Sect. 10.2.6 and conclusions are drawn in Sect. 10.2.7.

10.2.2 Clinical Requirements of a Proactive *Remote CVD Monitoring System*

The fundamental purpose of a *Proactive* CVD monitoring system is to provide clinicians with the necessary physiological data and medical parameters for remotely assessing the patient's overall cardiovascular condition with an acceptable degree of confidence. Clearly since the work presented here focuses on CVD, capturing and analyzing the ECG is of paramount importance. Of the same importance as the ECG trace itself, is a number of clinical parameters directly derived from the ECG signal. These metrics, are extensively used in clinical cardiology and their value is extracted from the temporal position and amplitude value of the constituent waves of the ECG waveform. Other vital signs proven to be significant for cardiovascular monitoring and thus included in the CHIRON system, are the patient's skin temperature and skin sweating index. Moreover, in the CHIRON system it was decided that an activity recognition mechanism coupled with the calculation of the equivalent metabolic energy value, consumed during activities has to be included. Activity data can play a key role in assessing and managing the patient's health status. The inclusion of activity and metabolic energy expenditure information, is also expected to facilitate the correlation analysis among the monitored parameters which may ultimately provide novel predictive features that can act as early warnings and lead to medical intervention before the full manifestation of a critical episode (e.g. myocardial infarction). To the best of our knowledge the CHIRON sensor platform is the first attempt in developing a fully-integrated solution capable of remotely monitoring all physiological signals required for the evaluation of CVD (Pantelopoulos and Bourbakis 2010).

In order for the system to satisfy its purpose, the aforementioned physiological parameters should be monitored in a continuous manner over extended periods of time during the day. This is the pivotal point of the *Proactive Care* approach as it allows to capture the variation in the value of these parameters over time, which is considered by medical experts to be the key approach in the attempt to capture the decline of the patient's health and to a further level the prediction of an imminent critical episode.

These clinical requirements posed by the *Proactive Care* approach yield certain engineering and technological challenges that must be addressed. Firstly, extensive uninterrupted monitoring necessitates the development of a non-invasive, having a maximum of three straps attached to the body's patient, low-power solution able to

effectively and robustly communicate the vital signs and relevant parameters from the point of acquisition to the clinician's remote location. In essence, long-term refers to a system capable of acquiring and processing information for up to 6–8 h during the day. This timeframe is dictated from the fact that CVD patients will not tolerate to bear such a system for more time during the day. To accommodate the clinical requirements described in this section, a customized sensor-platform and its corresponding system architecture was developed and presented in details in the following sections.

10.2.3 System Architecture

The sensor platform of the CHIRON system comprises of four sensor devices. Specifically, one ECG single-lead sensor, one temperature/sweating index (TS) sensor and two 3-axis accelerometers are employed. From the data captured by the two accelerometers the patient's activity and energy expenditure are derived. The four sensor devices are embedded in three sensor modules, powered with rechargeable batteries, attached to two elastic straps, resulting in a total of three separate platforms positioned in the patient's body. The first strap is a chest belt designated to be attached in the patient's upper abdomen/lower thorax area containing the ECG module and an integrated accelerometer-temp/sweat (ATS) module. The second strap places the second accelerometer module on the patient's thigh area. The three sensor modules collect readings from the four sensors and transmit them wirelessly, using the wireless Bluetooth (BT) standard, to the information processing platform (IPP). A pictorial representation of the wearable sensor system is provided in Fig. 10.6.

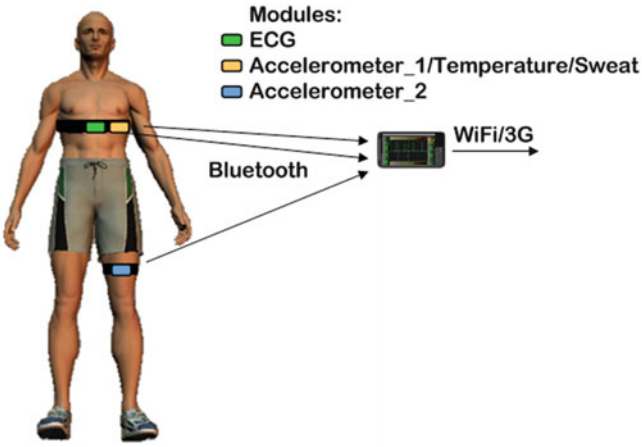


Fig. 10.6 The configuration of the integrated wearable sensor platform

The IPP constitutes an Android-based smartphone, tasked with fusing and processing the incoming sensor data while also providing the user interface. Apart from the TS data the rest of the sensor readings are subjected to specific signal processing for the extraction of the clinical ECG parameters, the identification of the patient's activity and finally the calculation of the energy expenditure. For these operations a customized Java-Android application (pIPP) was developed which includes the processing routines for the ECG and accelerometer data. In addition, to provide cardiologists, apart from the value of the EGC parameters, with the ECG trace itself, a compression scheme has been included to decrease the amount of ECG data that need to be communicated from the IPP to the medical analysis and storage subsystem (MASS). The MASS is a server-based framework providing medical experts with visualization and analysis tools for effectively monitoring the captured data. The raw physiological readings and the analysis outcomes are communicated to the MASS from the IPP through FTP protocol over wireless 3G data connection. This choice was dictated by the fact that 3G offers much wider coverage than any other wireless standard, allowing for data transmission to be taking place without the need of the patient being constrained to a specific area. To provide cardiologists with the ability to monitor the patient's status "anytime" a second Java-Android application (mIPP) which connects to the MASS and upon request downloads data and displays them in a handheld device (tablet/smartphone) was built.

Based on the above architecture, two operational modes were defined. In the real-time streaming mode, data are transmitted to the IPP in real-time while the processing outcomes are uploaded on the MASS at the moment of their production. Data and medical parameters are then accessed from medical experts with minimum latency. Under this mode, clinicians obtain a real-time assessment of the patient.

Although energy demanding, due to the continuous BT transmission, the real-time mode is ideal for patients in high risk of an impending episode or in situations where the clinician wishes to perform a remote stress test, by instructing the patient to walk on a treadmill while monitored. The second operational mode is the one expected to be utilized under normal conditions. In this mode, sensor data are not transmitted to the IPP in real-time. Instead, data are time stamped and locally stored to the communication modules during sensing. Once sensing is completed data transmission and processing on the IPP takes place in batch mode. The final outcomes are then uploaded and stored to the MASS. Medical experts can then access data through queries at will and thanks to timestamping, correlate and combine readings and outcomes obtained from different sensors. The IPP governs which operational mode is in use by commands issued to the sensor modules, which upon request can set their mode of operation.

10.2.4 Sensor Components

The sensor modules utilized in the CHIRON sensor platform are based on the Shimmer WSN module. The standard Shimmer platform is based on a

T.I. MSP430F1611 microcontroller, which operates at a maximum frequency of 8 MHz and is equipped with 10 Kb RAM and 48 Kb of Flash. An 8-channel 12-bit A/D converter enables connectivity with various sensors (ECG, EEG, accel, gyro, etc. . .). Wireless communication is achieved either with class 2 BT (RN-42 module) or through IEEE 802.15.4 (T.I. CC2420 module.). The standard BT v2 was elected as the sensor modules communication standard. For storage purposes the Shimmer platform is equipped with an integrated 2GB microSD card which is used in normal operation mode to store sensor readings (Burns et al. 2010). The power supply is comprised of a 450 mAh rechargeable Li-ion battery. Finally the Shimmer platform is designed to be programmed with the open-source TinyOS operating system (Levis et al. 2005).

10.2.4.1 ECG Sensor

The standard Shimmer ECG platform uses four snap electrodes attached through wires and medical patches to the body. It was thus deemed to be too invasive for the long-term monitoring purposes of the CHIRON system. The Shimmer ECG module was redesigned by replacing the four snap electrodes with two dry electrodes that can be attached to a commercially available chest strap (i.e. Polar®, Adidas®) and provide a single ECG lead. The resulting solution, depicted in Fig. 10.7, is lightweight (~80 g) with minimal invasiveness without any need for skin preparation or adhesive gels.

An analog front end with high gain and low cut-off frequency is necessary to condition the raw ECG signal and remove the noise for digital conversion and processing. The Shimmer ECG board has a low power front-end data acquisition circuit, composed by analog amplifiers and filters able to reduce the artefacts of movement, respiration and muscle contraction and to reach the desired dynamic range. The frequency response is 0.05–150 Hz with an amplifier gain of 175. The collected analog signal is then sampled through an A/D converter with 12-bit resolution. The digitized data are then passed to the microcontroller (MSP430) for processing and storage through a USART serial connection bus.

It was determined that in order to extract all necessary medical parameters, the ECG signal must be sampled at a frequency of at least 500 Hz. The quality of the acquired ECG signal from the customized platform was evaluated in hospital bedside settings as well as during a marathon run. Additional experiments were performed and the ECG was captured by both the CHIRON ECG chest strap and the Holter ELA device (considered as the golden standard for ambulatory ECG)

Fig. 10.7 The customized ECG sensor

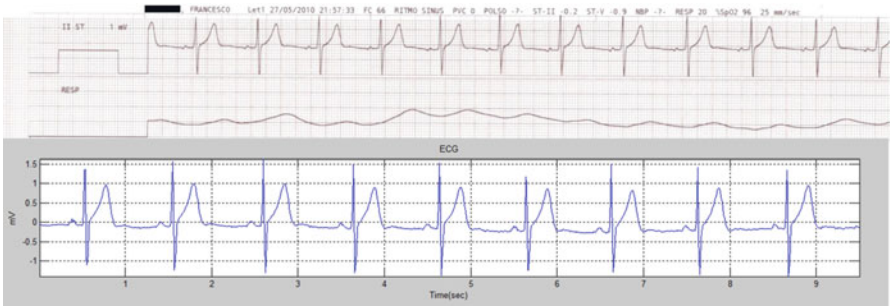


Fig. 10.8 Comparison of ECG chest strap recordings with a standard ECG machine at bedside settings

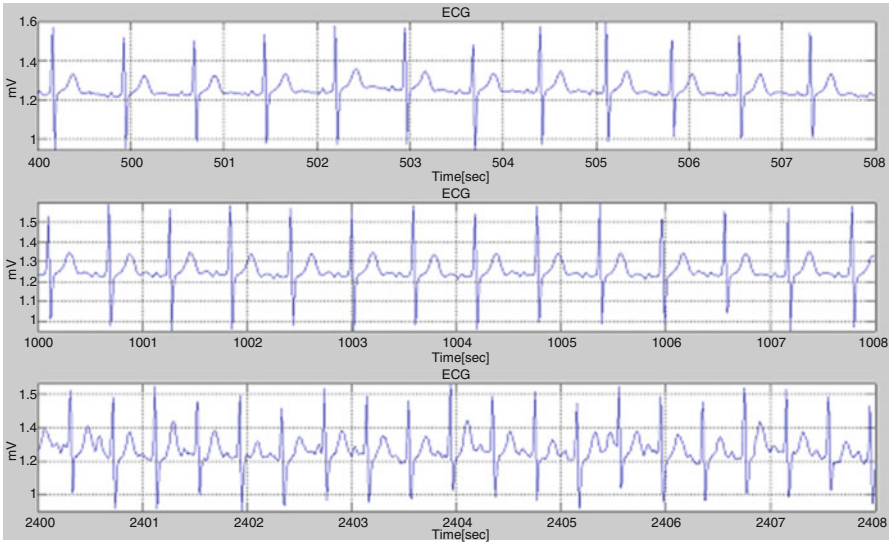


Fig. 10.9 Three portions of ECG chest strap recordings of a runner during a marathon with different heart rates

simultaneously. The obtained waveform confirmed that the signal quality was comparable to the one acquired by the gold standard Holter. A comparison of bedside recordings demonstrated that the CHIRON chest strap recordings produce a similar waveform to the one captured by standard ECG machines as shown in Fig. 10.8.

The CHIRON ECG chest strap was also used to obtain the ECG of runners participating in a marathon. As depicted in Fig. 10.9 the ECG signal appears stable and with a negligible number of artifacts.

Additional tests were performed in five subjects having their ECG captured at bedside settings for 3 h using both the CHIRON ECG chest strap and the Holter

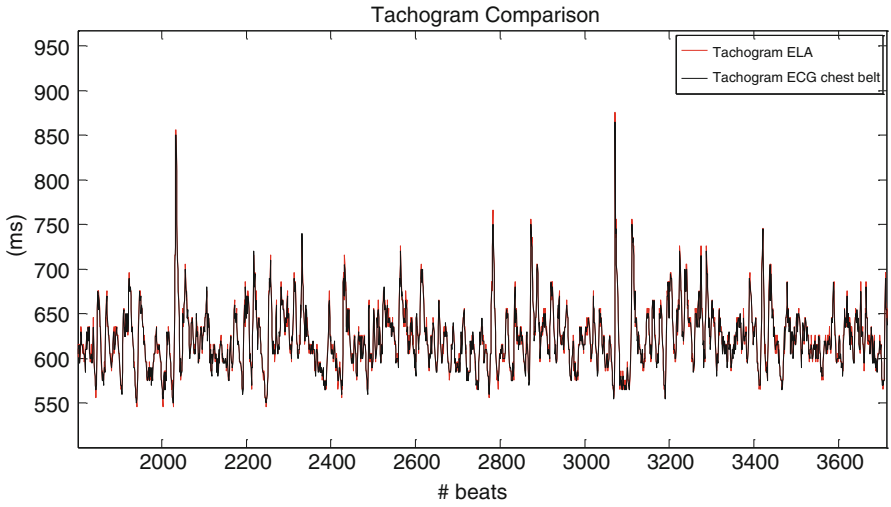


Fig. 10.10 Comparison of the tachograms produces from the ECG chest strap and the ELA Holter

ELA. The ECG chest strap provided readable signal for more than 95 and 99 % of the time of acquisition while the subjects were working and lying supine at bedside respectively. Following, the ECG signals from both devices were analyzed and the heart rate (HR) of the subjects was derived from both devices. As it can be seen in Fig. 10.10, the two tachograms are almost identical and the low value of the mean error distribution, (~0.01 s), as depicted in Fig. 10.11 confirms the similarity of the ECG signals acquired by the two devices. In addition, this investigation justifies the ability to derive medical important parameters from the CHIRON ECG chest strap with a similar level of accuracy as with golden standard devices.

10.2.4.2 Accelerometers

Accelerometers are utilized in the CHIRON system in order to provide the required data for activity recognition and metabolic energy calculation. Two Shimmer platforms, each containing a 3-axis accelerometer (Freescale MMA7361) are used in the CHIRON sensor platform. The two accelerometers are placed on the chest and the thigh. These positions were chosen based on the achieved activity recognition performance on preliminary tests (analyzed in Sect. 10.2.3). The sampling frequency for the accelerometers was set at 50 Hz.

10.2.4.3 Temperature/Sweat Index Sensor (TS)

Clinical specifications for skin temperature and sweating index measurements dictate not to use more than one patches for collecting both temperature and sweat index data. Subsequently, a sensor combining both capabilities was

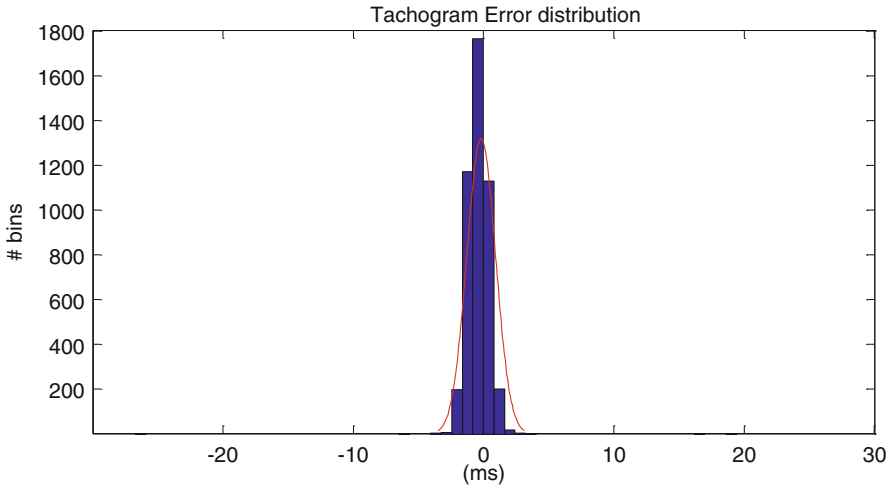


Fig. 10.11 ECG chest strap tachogram error distribution with respect to Holter ELA

Table 10.1 Temperature and sweat index sensor characteristics

Description	Value	Units
Sensor size	3 × 3	mm^2
Active power consumption	1	mW
Sleep power consumption	1.2	μW
Humidity range	0–100	%RH
Humidity accuracy	±5	%RH
Humidity resolution	0.04	%RH
Temperature range	0–60	°C
Temperature accuracy	0.3	°C
Temperature resolution	0.04	°C

developed. The accuracy of the sensor should be better than ±0.5 °C error with a measurement rate of one temperature and sweat index sample per 5 min. In order to satisfy these specifications, the Sensirion SHT21 digital sensor was chosen (http://wwww.sensirion.com). This sensor is a suitable choice for temperature and sweat index measurements, as it integrates both a temperature and a humidity sensor, used to obtain sweat index data by measuring the relative skin humidity. The Sensirion SHT21 sensor is also suitable for the purposes of the CHIRON system in terms of size, accuracy and low-power consumption in both active and sleep mode. Table 10.1 summarizes the TS sensor characteristics.

The TS sensor is placed under the patient's armpit a common place for acquiring temperature and sweat index measurements (Sund-Levander et al. 2004). In the CHIRON system the TS sensor is connected to the Shimmer accelerometer module attached to the chest strap (through the I2C bus) as shown in Fig. 10.12. Temperature and sweating index values are sampled once every 5 min. Samples are

Fig. 10.12 Prototype of the temp sweat index sensor connected to the Shimmer acceleration module

combined with the accelerometer data in a single message format and either stored or sent to the IPP. The platform was tested at a higher rate, i.e. every 20 s, in order to also check the autonomy of the sensor together with BT Shimmer module communications. Experiments were conducted with a volunteer wearing the sensor while performing different activities. Figure 10.13 shows the temperature and relative humidity of the volunteer who performed low physically demanding activities (e.g. sitting down or walking) until 7 p.m. when the volunteer changed to a moderate activity, (e.g. climbing stairs) for 15 min that resulted in sweating.

Results in Fig. 10.13 show a noticeable increment on humidity with a subsequent decrease when sweating stops. Concerning the results of the temperature, they are coherent with auxiliary temperature values and with temperature daily rhythms (Sund-Levander et al. 2002; Monk et al. 1995).

The sensor platform was also tested in a DYCOMETAL CCM81 climatic chamber with the chamber humidity being varied from 40 to 90 % in 2-h cycles for a period of 20 h while keeping the temperature constant at 36 °C. From Fig. 10.14, it is clear that the sensor follows the humidity variations almost instantaneously. This experiment also reveals that the autonomy of the platform exceeds 20 h at 20 s sample rate.

10.2.4.4 System Integration

Modules Packaging

The four sensor modules were packaged into a single wearable system. The POLAR chest belt was used and customized to provide encapsulation for both the ECG module and the combined ATS module placed on the chest. Also the two Shimmer

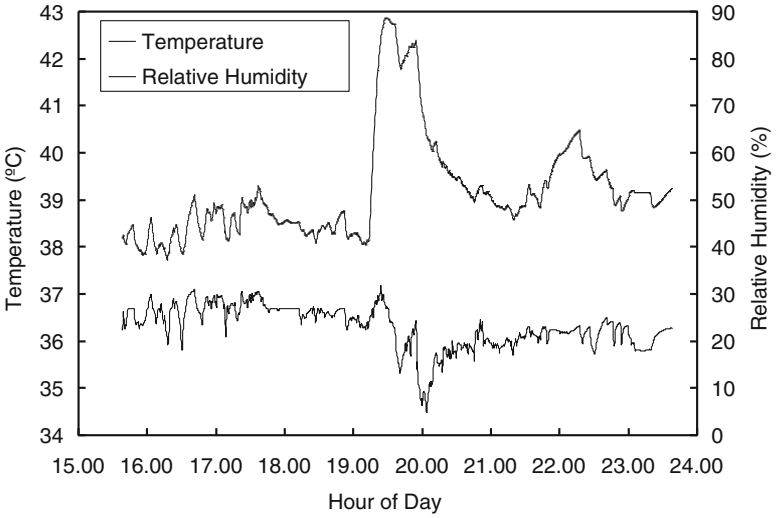


Fig. 10.13 Skin temperature and relative humidity (sweat index) results for 8 h of continuous monitoring and for different physically demanding activities

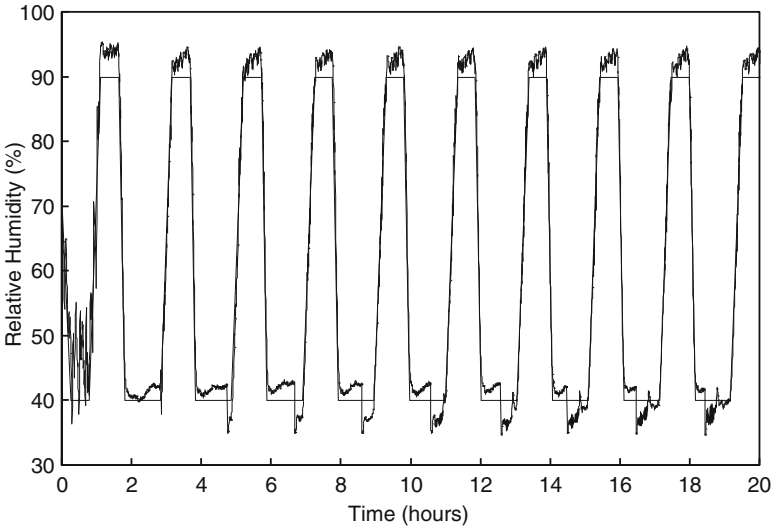


Fig. 10.14 Sensor test for 20 h of continuous data acquisition. The humidity varies from 40 to 90 % in 2-h cycles while the temperature is kept constant at 36 °C

modules in the chest belt had their charging ports connected so the user can recharge them without removing them from their encapsulation. Figure 10.15 illustrates the enclosure of the two sensor modules and the placement of the strap on the lower thorax area. The ECG module and the ATS module weight, 28 and

Fig. 10.15 The chest strap encapsulating the sensor modules

22 g respectively. The overall weight of the chest strap is less than 100 g, resulting in a non-invasive lightweight solution.

Firmware Development: Operational Protocol

To achieve the dual-mode functionality, imposed by the clinical requirements, all three Shimmer modules were programmed with both the ability to store sensor readings in their SD-card (with the BT turned off) or stream the readings through BT to the IPP. The user operational protocol was designed with two things in mind. The first was to minimize, as much as possible, the actions needed to be performed by the user and secondly to provide an energy-efficient strategy that would ensure a smooth uninterrupted operation. Subsequently the operational protocol was formed as following. Initially the sensor platforms are being charged on their docking stations with the BT "on" awaiting commands from the IPP. The required BT pairing of sensor modules to the IPP has taken place before the deployment of the system. The user removes the platforms and places them on his/her body to begin sensing.

Following the IPP sets the mode of operation to real-time and performs a sanity check on the correctness of the sensor placement. If this check fails the user is instructed to reposition the sensors appropriately. Once the IPP recognizes that the

sensors are correctly placed, broadcasts a command with the current IPP time to all modules to swap to normal operation mode. The sensor modules then turn the BT radio "off" and start logging data in their SD card. Once the logging period is over the user places the platforms at the docking stations again and the platforms turn "on" their BT. At this point the IPP requests the log, from each sensor, which contains information on the size of stored data. Then the IPP requests the stored data samples which are transmitted in batch mode. During data transmission, the platforms battery is recharged and a new operational cycle begins the next time (e.g. following day) that sensing must take place.

The sensor platforms customized firmware was developed in the TinyOS operating system. For the ECG sensor and the thigh accelerometer the typical TinyOS sensor components and interfaces were used to develop the sensor's firmware. On the other hand, the functionality of the ATS module had to be implemented. The TS sensor was integrated in the Shimmer module and connected to the I2C bus of the MSP430 microcontroller. A TinyOS interface was built in the stack layer to receive the TS data when necessary. The firmware for the ATS module employs both the accelerometer and TS interfaces in different sampling times.

In order to minimise the overall power consumption of the sensor platforms the BT module was kept "off", apart from when data are transmitted to the IPP either in real-time mode or in batches from the SD-cards at the end of sensing during the normal-mode. Predefined external events and commands produced by the IPP and formulated as TinyOS commands were used in order to trigger the swap, from one mode of operation to the other, initiate sensing and communication and govern the overall operation of the sensing platforms. Figure 10.16 illustrates the block diagram of an operational cycle and the commands and events that are generated. Figure 10.17 shows an analytical flow chart of the firmware that was developed for the ATS sensor platform. The description of the commands is as following.

Commands initiated by the IPP:

- Real-time operation: The sensor platform sample data at various frequencies (depending on the sensor) and sends each sample to the IPP.
- Normal operation: Platform stops Real-time transmission and wait for the absolute real timestamp from the IPP. The timestamp is written at the beginning of each log file and used as a reference for reconstructing the timestamp of each data sample. Following, the sensor platform turns its BT off and start logging data. No communication link exists between the sensor platform and the IPP during the logging process.
- Is Log Available: The platform checks if the log file is available and sends the status to the IPP.
- Send Log: The platforms send the log file. First, the absolute timestamp is sent, and then the data samples
- Delete Log: Platform deletes the log file.

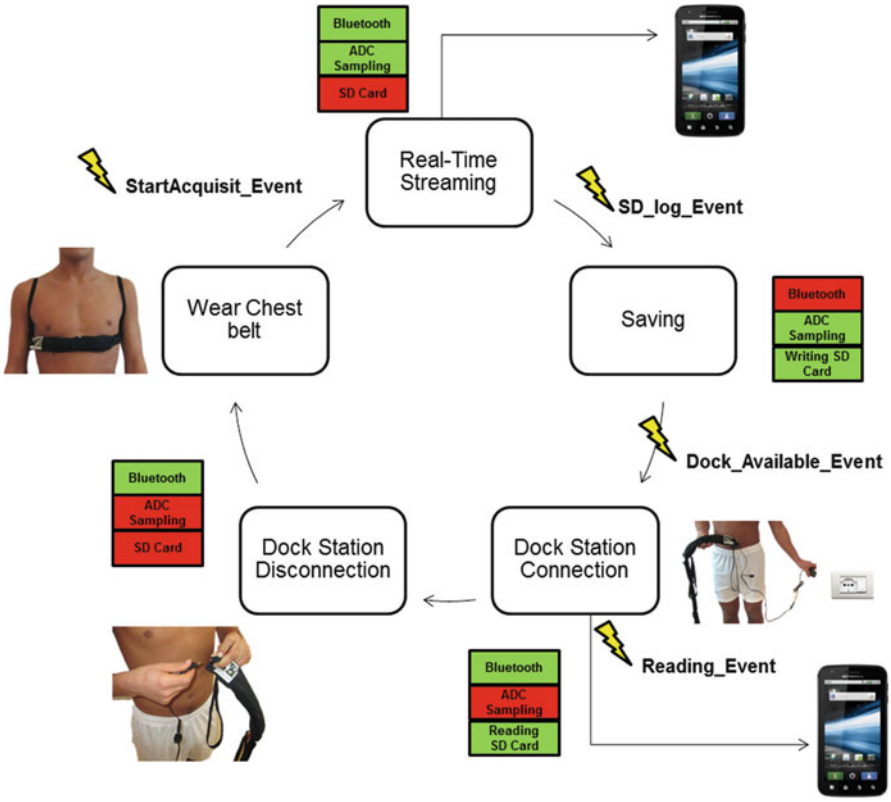


Fig. 10.16 The operational cycle of the normal (SD-logging) mode of operation

Sensor Event:

- Docking event (the sensor is put on the charger after normal operation): The platform stops logging and turns on the BT. Yellow LED is turned on, indicating that the log file is ready to be sent.

Data Generation

Considering the Shimmer platforms specifications an analysis follows on the amount of data expected to be generated on a daily basis. Firstly, the MSP430 A/D channels perform 12-bit (2 bytes of storage) digitization and that a 16 bit (2 bytes) timestamp is stored for each sample, Table 10.2 summarizes our projections based on the sampling frequency of every sensor. Based on these calculations the total amount of data for 6–8 h of daily use should be 57.5–76.2 Mb. This amount of data does not pose any issue in any operational mode, since in the real-time scenario BT can achieve datarates up to 300 kbps

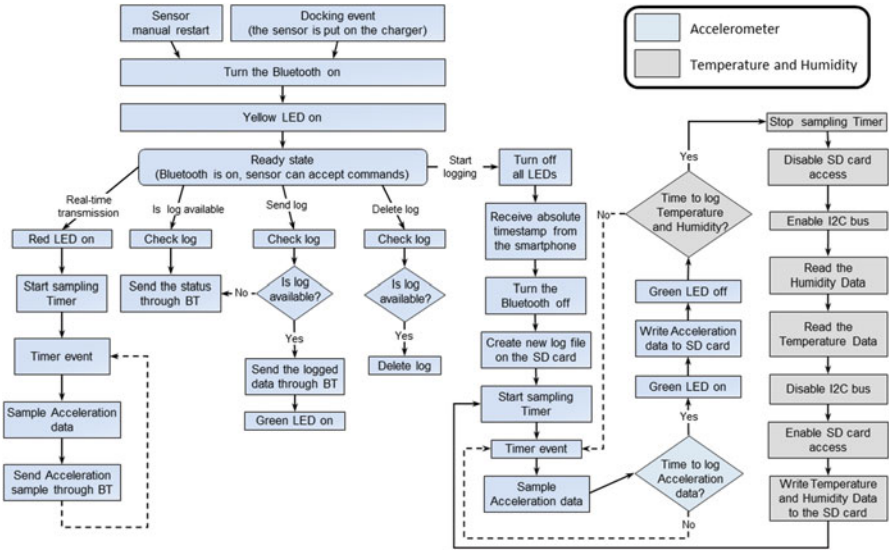


Fig. 10.17 ATS platform firmware flow chart

Table 10.2 Projected sensor data generation on a daily basis

Sensor module	Sampling frequency (Hz)	Data per second (KB/s)	Data per hour (MB/h)
ECG	500	1.95	6.85
2*Accelerometer (3-axis)	50	0.78	2.74
Temp/sweat index	0.0033	0.0264 B/s	0.0009
Total		3.89	9.59

which is more than adequate for the amount of data generated per second and in the logging operating mode the micro-SD cards on the modules have more than enough capacity to store the generated data.

Energy Profiling

A crucial part, of the CHIRON system is to ensure that the energy consumption of the various system components is such, that under any mode of operation the clinical requirements are fully met. The energy consumption analysis of the Shimmer platform, presented in (Burns et al. 2010), designates the ECG to be the most energy demanding module in the SD-logging mode drawing an average of 4.5 mA at 500 Hz, compared to the accelerometer which draws 1.6 mA at 50 Hz. When the modules stream data in real-time through the BT the consumption increases to 21.1 and 15.9 mA for the ECG at 500 Hz and the accelerometer at 50 Hz, respectively.

Table 10.3 Maximum battery lifetime for the three sensor components

	Maximum battery lifetime	
Sensor module	Real-time mode	SD-logging mode
ECG	8 h	100 h (4 days)
Accelerometer	18 h 30 m	14 days
ATS module	–	8 days

The consumption of the T/H sensor is negligible since the Sensirion SHT21 sensor draws only 300 μA during operation and its sampling frequency is very small. The same analysis also states that an ECG module operating at 500 Hz can log data generated at the rate of Table 10.2, in the micro-SD card or stream them in real-time via BT for at least 48 and 12 h respectively considering a 280 mAh battery without any energy management protocol. From this analysis, it is safe to assume, that since the same hardware equipped with a 450 mAh battery is used the clinical requirement of 6–8 h data logging (in the storing mode) or an adequate amount of time (~1 h) for live streaming (streaming mode) can be easily met, provided that the module's batteries are fully charged at the beginning of sensing. Table 10.3 lists the maximum battery lifetime of the three sensors components as obtained from a series of experiments.

10.2.5 Information Processing Platform (IPP)

The IPP has a central role in the CHIRON monitoring system. Sensor data are fused in the IPP and processed accordingly. The IPP also governs the operation of the sensor modules and provides the user interface. It was decided that a latest generation Android-based smartphone would be the best selection for the IPP. Smartphones are universal devices with which users feel familiar. This is particularly useful since it will result in a small learning curve for both types of users. In addition, users normally carry smartphones almost all the time, which is ideal for the continuous monitoring scenario under consideration. Latest generation smartphones are equipped with powerful microprocessors and wireless connectivity chipsets which guarantee that the desired IPP operations can effectively run on such devices.

10.2.5.1 Analysis Algorithms

The data analysis algorithms comprise of the ECG analysis, which leads to the extraction of several medical parameters from the ECG signal and the activity recognition and metabolic energy calculation (ARMEC) algorithms, which use the accelerometer data as inputs. The temp/sweat-index data are not subjected to any processing and are simply fused to the MASS.

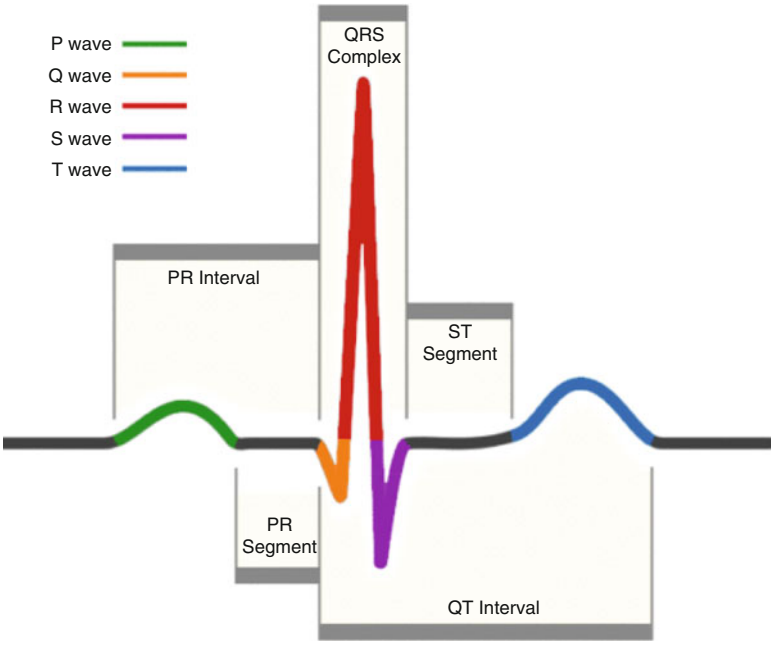


Fig. 10.18 The ECG parameters

ECG Feature Extraction and Compression

For the ECG analysis the main demands are to provide clinicians with the required ECG-related parameters as well with the ECG waveform itself. Since the ECG provides a direct representation of the heart's electrical activity, cardiologists have defined a number of ECG parameters that are used for the evaluation of the heart's condition. These features are either morphological (e.g. R-height) or temporal (e.g. QRS-duration, QT-interval, ST-segment, etc). Figure 10.18 illustrates a typical PQRST-complex and the relevant ECG features. Evaluation takes place by examining the morphology of the waveform and the value of the clinically important parameters. For example, a prolonged QT-interval is a biomarker of arrhythmia, while an elongated QRS may be indicative of bundle branch block. The estimation of the aforementioned parameters requires the detection of a total of 11 time instances within a single ECG heartbeat (PQRST-complex). In details, the onset and offset instances of the P-wave, the QRS-complex and the T-wave, as well as the time instance of the peak of each wave (P, Q, R, S, T) must be extracted. By obtaining this set of 11 time points, all the relevant temporal ECG parameters can be calculated. Additionally, parameters defined between successive heartbeats (e.g. P-P interval, R-R interval) can also be approximated.

A novel algorithmic solution, was subsequently developed which not only achieves the approximation of the ECG parameters with comparable to the state-

of-the-art accuracy, but also allows for an effective ECG compression scheme to be applied in parallel. This dual feat is achieved by employing the Haar Discrete Wavelet Transform (DWT) as the primary analysis tool. The DWT is realised as a cascaded filter bank using high and low pass filters defined according to the mother Wavelet and its scaling function. The output of the high pass filters provides the detailed coefficients (cD) and the output of the low pas filter the approximate coefficients (cA). The structure is known as Mallat's algorithm and allows for multi-level DWT decomposition to be performed by using the cA coefficients of a level as the output to the next pair of filters. It is established that due to the nature of the Haar function, potential extrema points in the input signal, will be represented as zero-crossing points in the cD coefficients and deflection points will be mapped into extrema in the DWT coefficients. By considering a single PQRST-complex and within it, the boundaries of the ECG waves as deflection points and their peaks as local extrema points, DWT multilevel decomposition in the dyadic space is performed and by using the aforementioned principle temporal points of the constituent waves of the ECG and the temporal location of the peak of each wave was extracted. Analytical details and performance results of the HFEA algorithm that was developed specifically for the CHIRON system can be found in (Mazomenos et al. 2013). The key novelty of the HFEA algorithm is the fact that it achieves comparable performance to the state-of-the-art ECG delineators while its computational complexity, in terms of mathematical operations required for its completion, is approximately five orders of magnitude smaller than that of the-state-of-the-art ECG delineators. It therefore provides an ideal solution for the continuous CVD monitoring nature of the CHIRON system.

Of all the ECG parameters (see Fig 10.18) the ST-segment (starting at the QRS-offset and finishing at the T-onset) carries significant clinical information and is of great value for the diagnosis of ischemic episodes and for the evaluation of the amount of oxygen supplied by the coronary arteries to the myocardium. Therefore, following the cardiologists guidelines, the ST-segment after being localised by the HFEA was subjected to additional processing for the calculation of a number of specific ST-segment medical parameters. Namely the ST-slope, the slope of the line connecting QRSoff and Ton points, the ST-deviation, the amplitude deviation of the ST-segment from the isoelectric line, the ST-area1, the area under the ECG waveform between the two ST-segment temporal boundaries when considering the isoelectric line as reference and the ST-area2, the area under the ECG waveform between the ST-segment temporal boundaries when considering the ST-line (line connecting QRSoff and Ton) as a reference were calculated using the Simpson rule. In total 20 ECG parameters are produced from the HFEA and the ST-analysis routines per heartbeat. These are either stored on the SD-card of the ECG module or transmitted directly to the IPP depending on the operational mode.

Finally the raw ECG data are also stored/transmitted in a compressed form. As analysed in Sect. 10.2.2 the majority of the generated data will be a result of the high sampling frequency (500 Hz) of the ECG Sensor. Moreover the clinical requirements dictate that clinicians must have access to the full ECG waveform apart from the value of the ECG parameters. This posed a requirement for an

efficient compression technique to be implemented alongside the HFEA algorithm. Since the DWT is utilized in the HFEA the inherent ability of DWT for compression and reconstruction of non-stationary signals was also exploited. Compression is achieved inherently in DWT because the frequency spectrum of the original signal is halved after each level of the DWT filter bank. Subsequently the output of the high and low-pass filters is subsampled, by a factor of 2 which discards half of the samples of the input signal and eliminates redundancy. Combined with a thresholding technique, based on the investigation of the Energy Packing Efficiency (EPE) of the resulting DWT coefficients, the compression scheme used in the CHIRON system maintains and store only those DWT coefficients that have a significant contribution to the energy of the signal. This is achieved by storing only the coefficients that have higher amplitude than a percentage of the maximum coefficient at that level. Reconstruction is achieved in a similar way that DWT decomposes the signal in cA and cD coefficients. The inverse DWT (IDWT) can reconstruct the signal from cA and cD coefficients using synthesis filters. Previous work reveals that the aforementioned compression method can achieve compression ratios of 16.5:1 and a percentage root mean square difference of 0.75 (Biswas et al. 2012).

Activity Recognition and Metabolic Energy Calculation (ARMEC)

The Activity Recognition and Metabolic Energy Calculation subsystem recognizes basic activities and estimates the patient's energy expenditure. To develop and test these methods, recordings of physiological signals during activities with different energy expenditures were made. The recordings consisted of acceleration, heart rate and skin temperature data used for activity recognition and energy expenditure estimation. Reference energy expenditure was provided by indirect calorimetry, and true activities were labelled by an observer. The reference energy expenditure was measured by Cosmed K4b portable gas analyser (http://www.cosmed.it/index.php?option=com_content&view=article&id=249&Itemid=174&lang=en), which computes the metabolic activity through the analysis of inhaled and exhaled gases. The recordings were made by ten subjects performing the basic activities to be recognized as listed in Table 10.4. The energy expenditure values are expressed in MET (metabolic equivalent of task, 1 MET corresponds to the average energy expenditure at rest), and are averaged over all ten test subject to provide an indication of the mean energy expenditure of each activity. Both the activity recognition and energy expenditure calculation were essentially tackled in the same way. The stream of acceleration measurements was split into windows, each window overlapping with the previous one by one half of its length. Several attributes were computed from the acceleration data within each window. These attributes formed a vector, which was fed into a machine learning algorithm. The algorithm constructed a model which either recognized the activity within the window or estimated the energy expenditure based on the average energy

Table 10.4 Activities in the test scenario

Activity	Basic	Energy (MET)
Lying	Y	1.19
Sitting	Y	1.26
Standing	Y	1.26
Walking	Y	
Slowly (4 km/h)		3.50
Quickly (6 km/h)		5.03
Running slowly (8 km/h)	Y	7.80
Stationary cycling	Y	
Lightly (1 W/kg of body mass)		4.91
Vigorously (1 W/kg of body mass)		7.22
Kneeling	Y	1.30
On all fours	Y	1.77
Lying doing light exercise		1.28
Sitting doing light activities		2.30
Walking doing light chores		2.30
Scrubbing the floor		2.65
Shoveling snow, digging		3.40

expenditure of Table 10.4. The output of the activity recognition model was the used as input to the energy expenditure estimation model.

To reduce the number of attributes thus the computational complexity and power consumption, attribute selection was applied to the attribute sets of each of the two models. First, the quality of all the attributes was estimated using the ReliefF method (Kononenko 1994). The attributes were then ranked by their ReliefF score. Second, the attributes were removed from the attribute set one-by-one, starting with the lowest-ranked one. After each attribute was removed, the classification/regression accuracy was tested using cross-validation. The removal of low-ranked attributes slightly increased the accuracy, then the accuracy plateaued, and finally the removal of high-ranked attributes decreased the accuracy. The final attribute set consisted only of the attributes whose removal decreased the accuracy. The models were tested using the leave-one-person-out method. This means that the model was trained (including attribute selection) on nine test subjects and tested on the tenth. The procedure was repeated ten times, using a different subject for testing each time. This procedure estimates the performance on a previously unseen person.

A. Activity Recognition

Since the user's activity is changing over time, it should be recognized with an appropriate frequency. Prior work demonstrated that 2 s windows are a reasonable trade-off between the overall recognition accuracy (increases with window length) and the ability to recognize very short activities (decreases with window length), so considering the 50 Hz sampling frequency of the two accelerometers each window contains approximately 100 acceleration measurements (Žbogar et al. 2012).

From these, 128 attributes per window were computed, drawn from previous works (Gjoreski 2011; Tapia 2008). Attribute selection using the RelieF method reduced the number of attributes to 41. The attributes were computed from data processed by either a low-pass filter (to single out the acceleration component related to the user's posture, which changes infrequently) or a band-pass filter (to single out the acceleration component related to the user's motion, whose frequency is above the frequency of posture changes and below the frequency of sensor noise).

Prior work on activity recognition with accelerometer data showed that the Random Forest algorithm is the most suitable for this task (Gjoreski 2011). The attributes were thus used to train and test a Random Forest activity-recognition model, using the Weka machine. With the reduced attribute set (41 attributes), the classification accuracy was 92.0 %, which was a 0.8 % point increase compared to the full attribute set with 91.2 % classification accuracy. Although the difference in accuracy is not statistically significant, the computational complexity was reduced, due to both reducing the number of attributes in the model and by removing the calculation of attributes with complexity above linear in each window (energy, entropy, ...) which require the computation of Fourier transform. This improves the battery life and speeds up the recognition process.

B. Metabolic Energy Calculation

Since longer time is needed to estimate the energy expenditure, windows of 10 s were used (Luštrek et al. 2012). Each window contained approximately 500 acceleration measurements and 10 heart rate and 10 skin temperature measurements. For each window, 137 attributes were computed, 135 from accelerometer data, one was the heart rate and one the skin temperature. Attribute selection reduced the number of attributes to 29. Among these 29 attributes 5 require the calculation of Fourier transform, which was removed to decrease the computational complexity.

It was verified, through experimentation that the Support Vector Regression (SVR) algorithm, yields optimum results in estimating the energy expenditure estimation (Luštrek et al. 2012). Again the Weka framework was used to train and test the SVR model. The accuracy of the model in terms of mean absolute error (MEA) was 0.67 MET. The analysis of the error showed that the largest error occurred in the cases of running and cycling. To overcome this problem, two additional regression models were trained specifically for these two activities. This resulted in three regression models: two trained on running and cycling and used to estimate the energy expenditure on these two activities, and one trained on all activities and used on all activities except running and cycling (training on all activities increased the robustness in case running and cycling were misclassified by the activity-recognition model). All models were again trained with the SVR algorithm. The approach with three models resulted in MEA of 0.56 MET, a decrease of by 0.11 MET. Figure 10.19 illustrates the true vs. the estimated energy expenditures. On the x-axis are the true MET values and on the y-axis the estimated MET values.

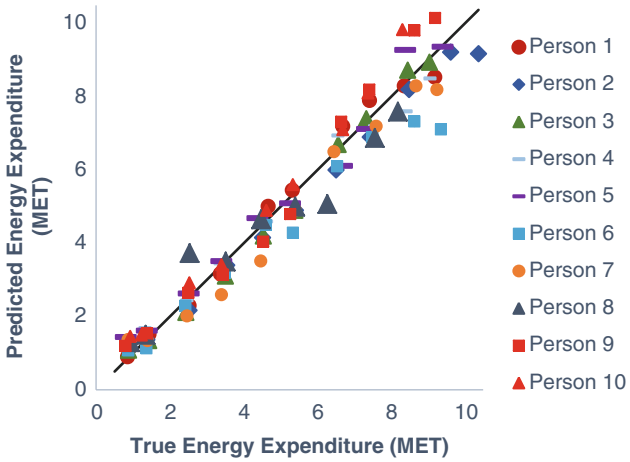


Fig. 10.19 True vs. estimated energy expenditure

10.2.5.2 IPP Application

The IPP provides the user interface to both groups of users; medical experts and CVD patients. Subsequently two separate Android-based applications were developed that implement the functionality of the IPP for each user group—patient/user and medical experts.

Patient IPP (pIPP)

The IPP of the patient/user (pIPP) has the role of the gateway of the sensor plane analyzed previously. It is responsible for setting the desired mode of operation, synchronizing the data from the sensor modules, fusing the sensor data, executing the analysis algorithms and ultimately transferring outcomes and compressed data to the MASS. To implement these operations an Android-based application was developed and its overall architecture is given in Fig. 10.20.

To begin sensing, the patient attaches the two straps (chest and thigh) to his/her body, and initiates the system by pressing the "start sensing" button on the pIPP. The pIPP establishes communication with the three sensor modules through BT and depending on the mode of operation acquires the sensor readings either in real-time or off-line, from the Shimmer modules SD-card at the end of sensing. Latest BT modules attached on Android smartphones are more than capable of simultaneously maintaining the link and communicating to multiple BT modules. The pIPP sets the appropriate operational mode by transmitting commands through the BT link to the sensor modules. The core of the pIPP contains the two analysis packages of the ECG and ARMEC, both implemented in Java language as per the Android OS

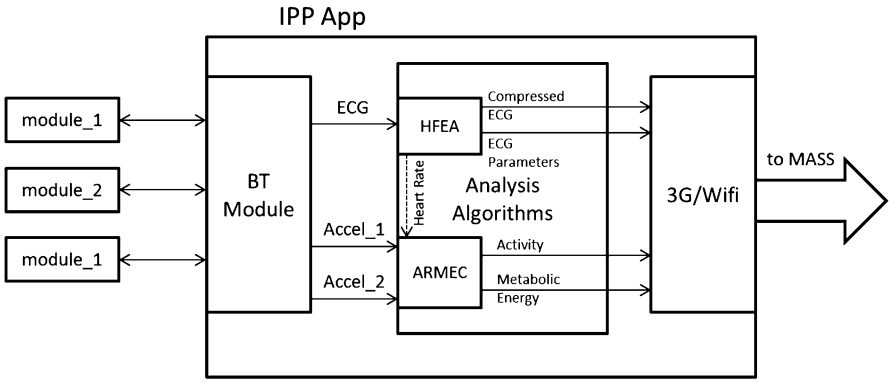


Fig. 10.20 Schematic diagram of the IPP-app

requirements. Again the operational mode dictates the execution of the two algorithms either in real-time or off-line. Moreover the IPP-App is also tasked with conveying the algorithm's outcomes to the MASS through 3G wireless communications.

The analysis algorithms are also used to evaluate the quality of the received data and affirm the correct placement of the sensor modules before the beginning of sensing. This sanity check stage takes place in the pIPP, after the patient "starts sensing", by setting the system in real-time operation and evaluating the algorithm's outcomes. The patient is instructed to lie down for 30 s and then stand for another 30 s in order to confirm the output of the activity recognition subsystem. The ECG signal is evaluated by investigating the soundness of the HFEA results for 2 min. If one of the modules is found to be misplaced the user is then instructed to reattach it.

The pIPP uses the smartphone's real-time clock to inform the patient when monitoring must stop. In addition the pIPP is tasked with synchronizing the data received from the three sensor modules. Synchronization is critical as the ARMEC algorithms require the data from the two accelerometers to be synchronized. In addition, data synchronization facilitates the correlation analysis of different physiological parameters to take place (ECG parameters, activity) a requirement posed by the medical experts. Synchronization is achieved achieved using the "start sensing" command as reference and the Shimmer counter timestamp stored with every data sample. Each data sample is then labelled with a global timestamp by the IPP. This method will of course suffer from the inaccuracy of the Shimmer internal crystal clock (Epson FC-135 32.7680KA-A3) which has a tolerance of ± 20 ppm, which results in 1.8 s maximal drift in 24 h. In the worst case scenario, when two sensors have different drift directions, the time difference is 3.6 s. However, in the CHIRON system a maximum of 6 h continuous logging is expected, which would result in a maximum drift of 0.9 s, which is acceptable for the task of activity recognition and energy expenditure. Experiments carried out showed that no

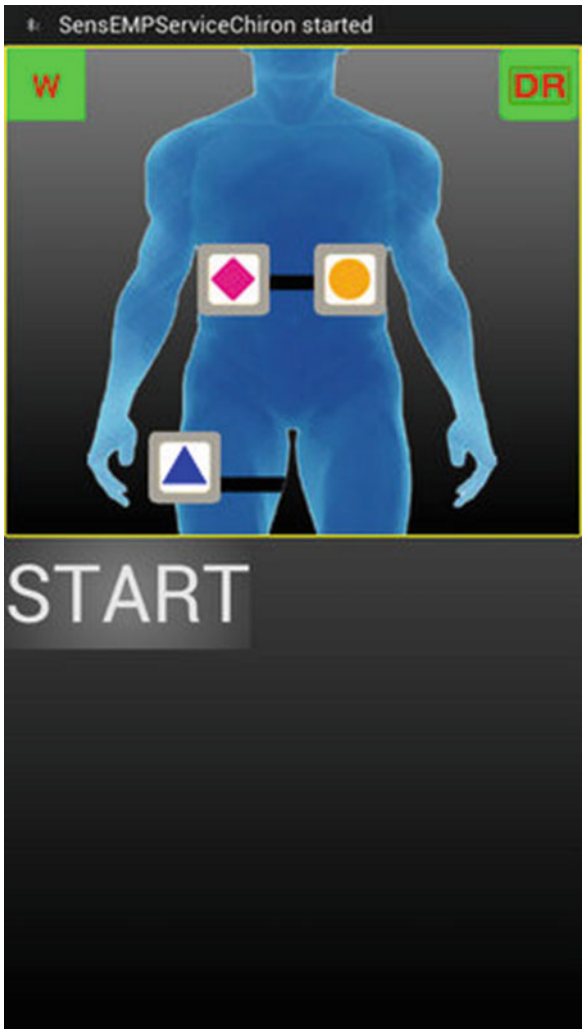


Fig. 10.21 The user interface of the pIPP

significant drift was present after 6 h and a time difference of 0.3 s was observed among the two accelerometers after 24 h of continuous logging.

The user interface of the pIPP was designed in such a way as to require minimum intervention from the user. In essence the user need to simply click the "START" and the pIPP takes over and instructs the user to go through the sanity check and then the normal sensing session begins. Moreover, the pIPP offers a report interface where the patient/user can manually input the value of a number of medical parameters (blood pressure, weight, etc..) not measured by the CHIRON system. An illustration of the pIPP user interface is provided in Fig. 10.21.

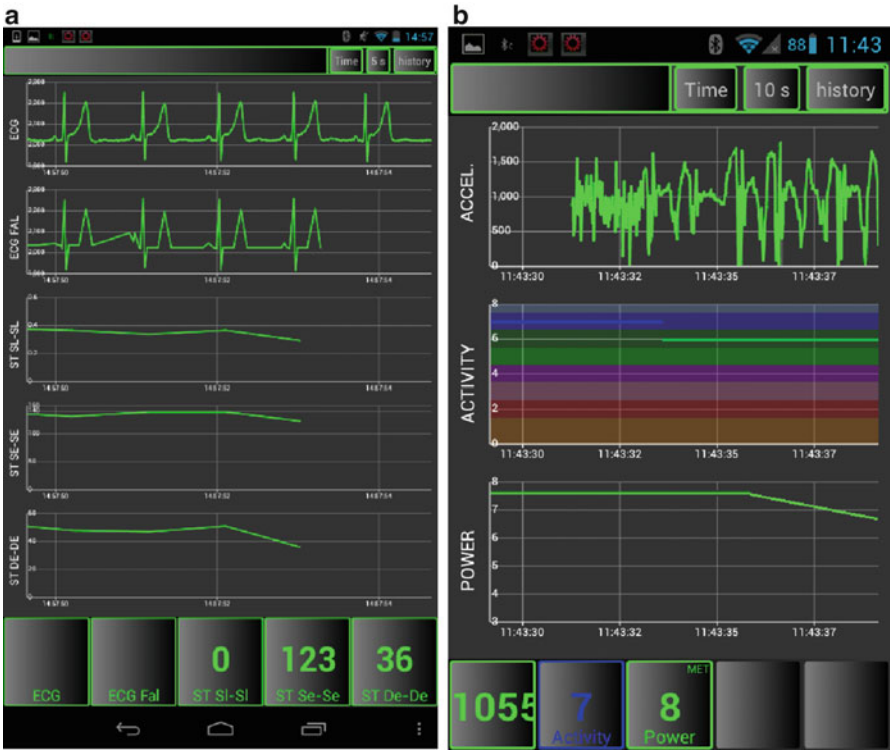


Fig. 10.22 The interface of the mIPP application; (**a**) ECG analysis (**b**) ARMEC analysis

Medical Expert IPP (mIPP)

The Android-based application designed for the medical expert user of the CHIRON system constitutes a monitoring interface in which medical experts can obtain the values of all parameters, measured by the sensors or produced from the analysis of the physiological signals. Data and algorithmic outcomes are first transferred to the MASS and the mIPP either obtains these with minimum latency if the system is operating in real-time mode or on a query-based approach where the expert specifies the time window to obtain data for. Figure 10.22 shows an example of the mIPP where the ECG analysis and the ARMEC analysis take place in the medical expert's handheld device.

10.2.6 Medical Analysis and Storage Subsystem (MASS)

The MASS is in essence a server-type database that receives data, through an ftp link from the pIPP. In the normal mode of operation, data are stored locally on the

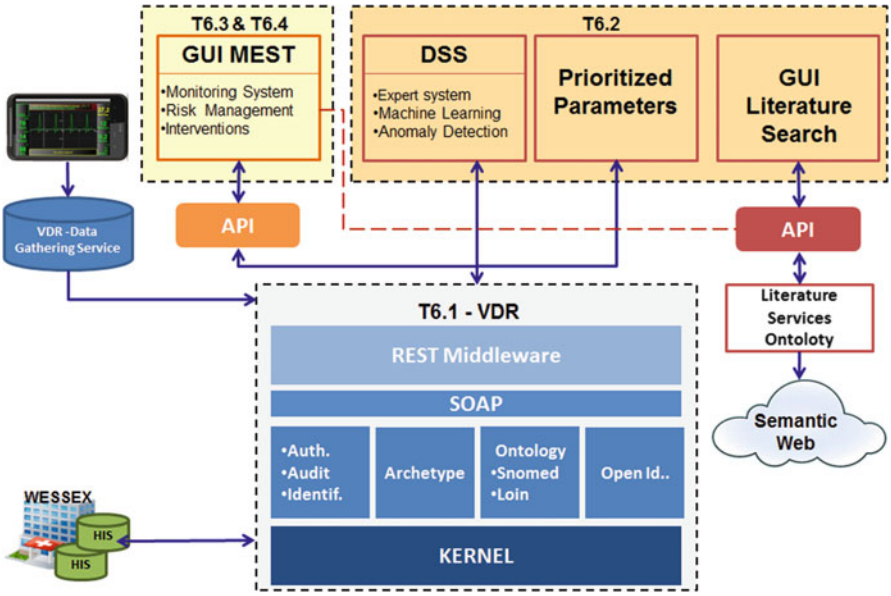


Fig. 10.23 CHIRON integration of multisource information

MASS and upon a request from the mIPP, data that correspond to the specific timeframe, defined by the expert, are sent to the mIPP. In the real-time mode data are directly fused to the mIPP from the database with minimum latency. In addition to the mIPP, a web-based interface (Medical Expert Support Tool) was built on the top of the MASS to provide doctor with an enhanced tool for the monitoring of patients through the CHIRON system. In addition the MASS was built in such a way to facilitate the integration of multisource medical information and the analysis and feedbacks generation. Under CHIRON specialized software was developed to analyze and integrate all medical data obtained from heterogeneous sources (CHIRON sensor platform, imaging systems, laboratory values, epidemiologic data, lifestyle information, family history, etc.) and the data already existing in the Hospital System (HIS).

In the core of the MASS, the Virtual Data Repository (VDR) (see Sect. 10.2.6.1) provides an archetype-based solution to represent patients' Electronic Health Records (EHR). As represented in components integration schema, VDR is commonly accessed by all the other component of the MASS. All the flow of information is done through the VDR accepting and delivering data in several standards. Figure 10.23 illustrates the overall architecture of the MASS.

The integration done during the CHIRON project has followed several steps:

1. All the information is accessed using archetypes, but other standards are available (for instance HL7 CDA).
2. Decision Support systems are communicated with the Advance or Slim Medical Expert Support Tool (MEST) Graphical User Interface.

3. DSS communicates with the VDR directly. Once a day the three modules compute all the parameters stored in the database.
4. Retrieve the risk assessment values stored and the value of the parameters. The DSS component gives the results and writes them in the VDR using the archetypes.
5. An average of the parameters extracted from the raw data information coming from the tele-monitoring (VDR Gathering service) should be stored once a session (three sessions a day) in order to provide to the doctors a general view of the health status of the patient.

The information inside the archetype is retrieved by means of a GUID (Global Unique Identifier) which permits to identify each object. For CHIRON project the communication is done using RESTful web services through JSON which is compatible with the MEST UI (Fielding 2000). The VDR is accessed using a proprietary middleware (REST API).

10.2.6.1 Virtual Data Repository (VDR) Concept

Heterogeneous medical information comes from multiple systems. The Virtual Data Repository is the middleware, which permits the connection with the data management transparently for the user. The VDR covers all the storage components, the input to the modules of the system and will be the data on the patients, which will be gathered by the CHIRON sensor platform and contained in the patients' EHRs. However, data from other sources can also be included, for instance to train machine-learning classifiers and to demonstrate that the system works before the real data from CHF patients is available. The VDR is a set of repositories located somewhere and each element of the system could access to the data using the openEHR Kernel interface.

The VDR is physically in the hospital and has some very bright and clear aspects:

- It complies to the EU standard for medical data transport
- It has an interface model that "talks" all standardised languages (XML, Edifact, etc.)
- It can use the approved and authorized NHS Archetypes for this study
- It can connect to other Interfaces (like the Southampton General Hospital or Sapienza University Hospital).

10.2.6.2 Decision Support System

As mentioned previously, apart from the mIPP, with which the doctor is able to monitor the patient's physiological data, a more elaborate tool is required to provide doctors with the information and analysis mechanisms needed to make clinical

decisions regarding the patient. Therefore a web-based decision support system (DSS) was developed and linked to the VDR to accommodate this need.

The risk assessment component is divided into three sub-components using three different approaches. The expert system sub-component incorporates existing medical knowledge. The machine learning sub-component learns the relation between the parameters characterizing the patient's health and his/her risk using an Artificial Neural Network. The anomaly detection sub-component detects anomalies in the values and relations between the values of the parameters characterizing the patient's health, using the "Local Outlier Factor algorithm" (Barca et al. 2012). The literature search component is used when a doctor wishes to consult the medical literature on any issue that comes up during his/her decision-making process. It chiefly accesses data sources external to the CHIRON system, but the search can be contextualized by information from the patients' EHRs. The prioritized parameters component attempts to identify the parameters important for the risk assessment, and interesting relations between them. It can thus provide suggestions for the configuration of the risk assessment.

10.2.6.3 VDR Bridge

The openEHR REST API offers a set of functions to store and retrieve the data. In order to facilitate the access to the rest of components (DSS, MEST, etc.) a Java VDR library has been implemented called VDR Bridge. The library is a set of functions, which allow storing and retrieving the data from the openEHR REST API, but instead of using the long paths, the inputs and outputs are translated into meaningful clinical concepts. Figure 10.24 shows a snapshot of the main functionalities of the Java Library.

For the CHIRON Project several archetypes have been created or reused to cover 67 medical parameters in order to store the information in the VDR. The modified OBSERVATION archetypes have been used to retrieve the information (coming from the Monitoring Settings component). The CHIRON data is coming from the sensors and is retrieved using the archetypes included as entries inside one composition of the patient EHR. For each parameter OBSERVATION and EVALUATION archetypes downloaded from the CKM (see Chaps. 6 and 8 for further details) or created ad-hoc have been used to show the measurements, diseases, medications and so on.

The way to configure the clinical concepts and the archetypes is using the openEHR configurator dragging and dropping the paths read from the archetypes. This tool allows the user associating the parameters needed by the doctors to the archetypes stored in the openEHR kernel. The Graphical User Interface read the configuration file and builds the screens dynamically. Figure 10.25 shows an example of concepts and archetypes.

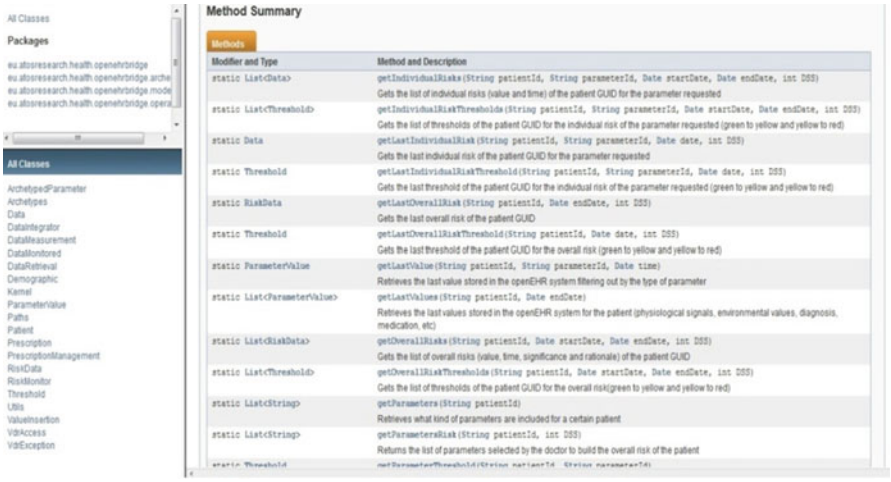


Fig. 10.24 VDR bridge API doc

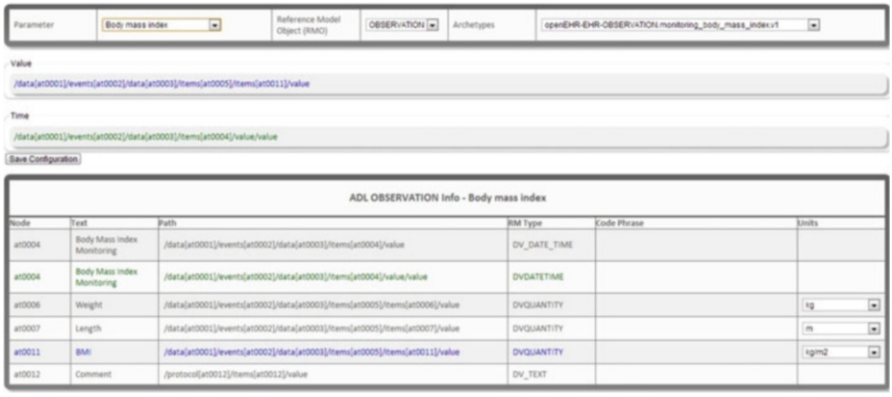


Fig. 10.25 openEHR configurator

10.2.6.4 Medical Expert Support Tool (MEST)

Advanced MEST

"The Medical Expert Support Tools (MEST) is a set of UI-based tools built on top of Computer Based Clinical Decision Support System (CDSS). They have the purpose of helping the medical professionals to detect risk from chronic disease monitoring and subsequently to manage the risk in terms of assessment and intervention to mitigate the dangerous situation related to the detected risk factor" (Barca et al. 2012).

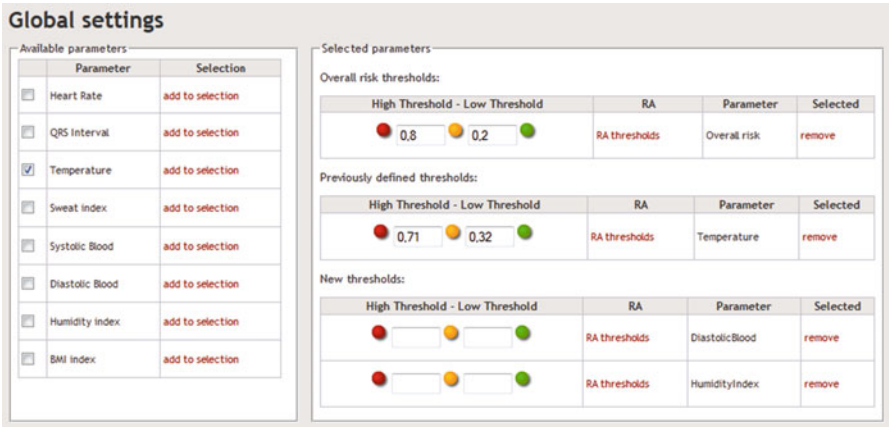


Fig. 10.26 Monitoring settings GUI

The Advanced MEST provides the medical professionals with a multi-parametric view and an overall status of the patient indicating the risk factors with meaningful colours (green, yellow and red). The assessment of the person's risk based on the analysis of the information already stored in the system offers the doctor a clear insight of the current situation of the individual. The detection of critical situations will activate the feedbacks to the patient and to the clinicians. The MEST API is designed and developed as a RESTful API to communicate between MEST applications and adapt data from CHIRON backend services and components exchanging the clinical data.

The risk assessment procedure (Fig. 10.26) starts when the clinician configures the system communicating to the CDSS modules which parameters have to be measured and their corresponding thresholds.

Once the system is configured MEST presents the list of his/her patients with the overall risk coming from the CDSS components and stored in the VDR (see Fig. 10.27). The overall risk is calculated taking into account all the computed parameters shown by means of a green/yellow/red mark.

For each parameter configured an individual assessment is calculated (see Fig. 10.28) to provide the doctors with the specific evaluation of the values producing the risk factor along with clinical literature articles related with the disease of the patient and his/her monitored clinical parameters (left down in Fig. 10.28).

Finally the clinician can access the causes (parameter values) that led to the evaluation of a risk factor, displaying the complete history of the monitored clinical parameters (see Fig. 10.29). The thresholds of the parameters are also shown in order to see if the values exceed the limits (above or below).

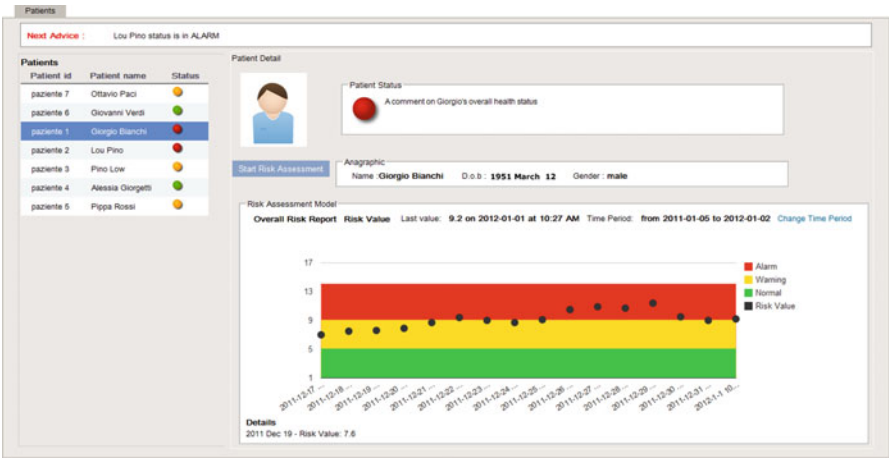


Fig. 10.27 Overall risk assessment view

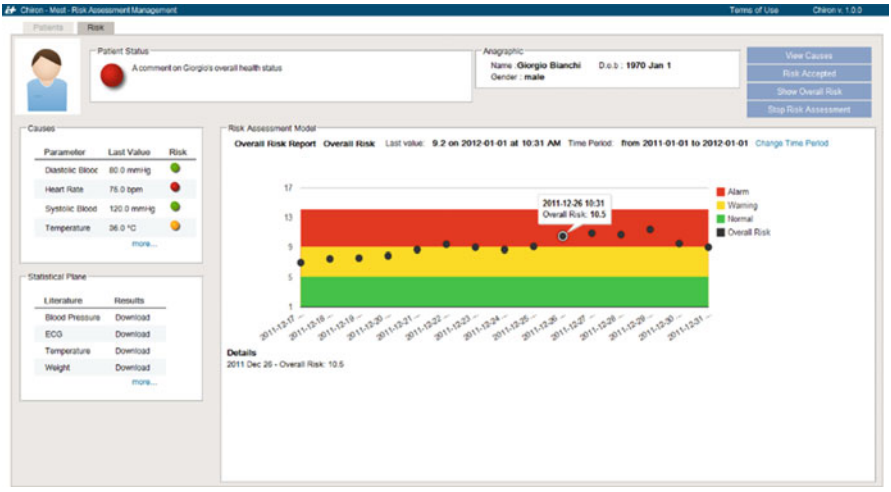


Fig. 10.28 Individual risk assessment view and articles contextualised to this patient

Refinement of Functional Aspects for Observational Study: Slim MEST

For the CHIRON Observational study the risk assessment modules were not included due to the lack of relevant data, which will be gathered during this period. For this reason clinical partners defined a new set of requirements focused mainly on Observational Study requisites enabling the medical user to insert whatever clinical data associated to a patient in openEHR format.

A new Graphical User Interface was proposed called "**Slim MEST**" which includes some parts of the MEST and allows the doctor to analyse the data of the

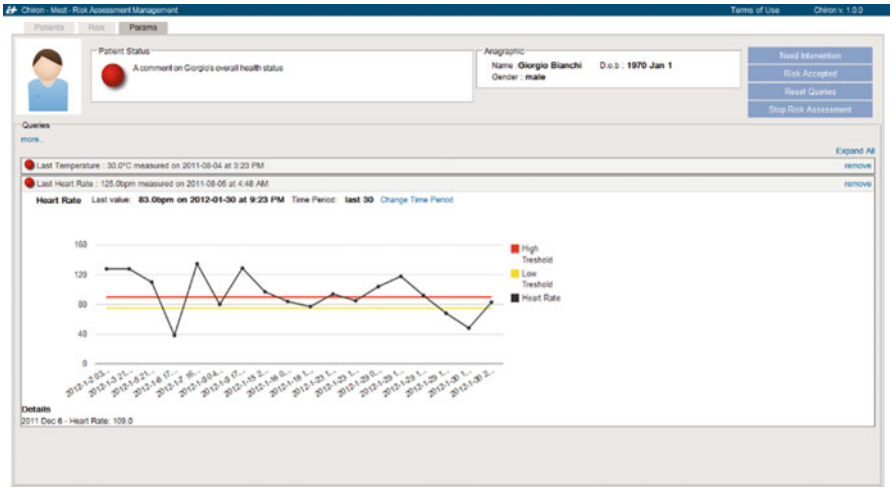


Fig. 10.29 Clinical data and thresholds view

patient (monitored data) collected in the overall observation period plus some data generated by the doctor himself. The risk assessment is not included since the "rules" are still to be defined.

The **Slim MEST** includes:

- **Visualization of health status parameters**
  It shows the last health status of the patient coming from the sensors. After a correct access, main screen is loaded. The visualization option offers an overview of parameters that have been measured for selected patient. A period of time can be pre-selected both general and individual for each of the parameters. Data is displayed as a self-expandable table, which shows in every row parameter name beside its last value. Selecting one of the numeric parameters and choosing a period of time with different measurements produces a graph (Generate button) that illustrates parameter value variation. It is possible to generate a pdf file (Fig. 10.30).
- **Insert parameters health status**
  This option brings out a paginated list of all available parameters in CHIRON system. An empty text box lets the user update the value of each parameter. As a reference, last value stored is shown accompanied by the date when the measurement was taken. In case the user wants to repeat the value, a copy button provides this functionality. It will update the value with the current day and time. Controls section in the bottom of the page, lets the user:
  - Move back and forward between pages.
  - Print a PDF report with inserted values.
  - Store values.
  - Exit the insertion module (Fig. 10.31).

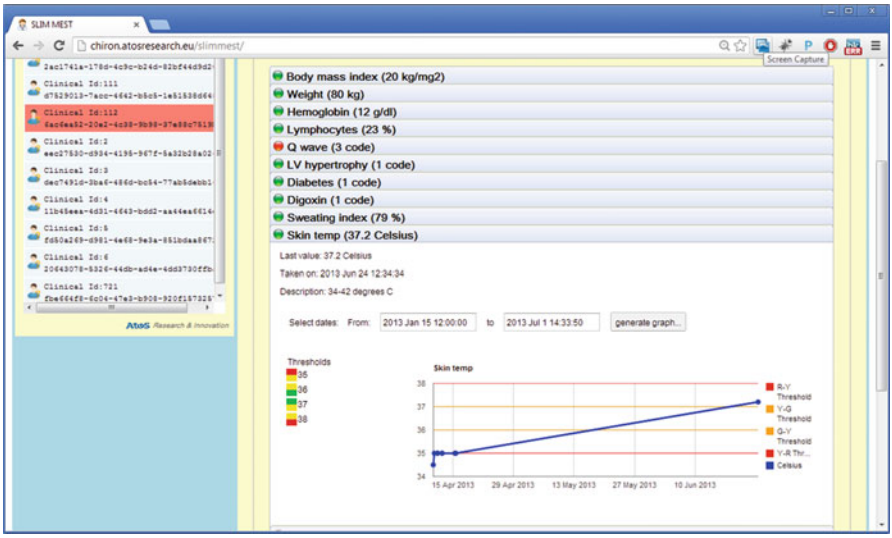


Fig. 10.30 Slim MEST: visualization

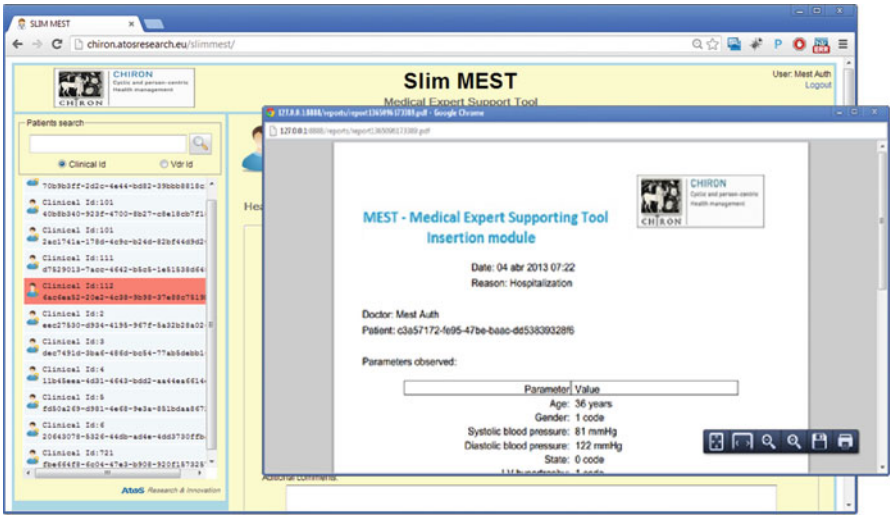


Fig. 10.31 Last health status reports generation

Security Aspects

MEST has been designed to be embedded in a security system. From the point of view of the Graphical User Interface the main aspects to be considered are the authorization and authentication aspects. An authorization mechanism that allows the identification and access levelling of a caregiver is highly recommended, both

for security purposes and for privacy concerns. The REST services are designed to expect that the UIs add a textual (String) representation of an authorization token to every privacy-concerned (or potentially risky) action the user wishes to perform (e.g. request of the clinical data of a patient, or the notification of a risk acceptance). Slim MEST requires a Google account to get access to the tool (due to authentication process). Users that own a Google account can write their access details. The rest, can Sign up for a new account. After entering the system, the application asks for permission to access through your Google account. Accessing through Google API produces a token, which is required by VDR API which contains a list of medical users and patients associated to provide authorization.

10.2.7 Conclusions

This chapter provided the reader with analytical details of the CHIRON Project case study. The CHIRON Project focused on combining state-of-the art technologies and innovative solutions into an integrated framework designed for an effective and person-centric health management along the complete care cycle of CVD. Within the areas of e-health and tele-medicine, the CHIRON case study proposed a holistic system architecture constituting of a wearable sensor system, its corresponding mobile application and an advanced interface for medical evaluation and analysis bundled into a single integrated framework for the continuous evaluation and assessment of CVD patients. To the best of the author's knowledge the CHIRON system is the very first attempt of an e-health tele-monitoring system towards the *Proactive Approach* for the care of CVD. The CHIRON system is envisioned as a tool to provide medical experts with the desired technological support for analyzing the variation and the correlation of heterogeneous medical data in an attempt to introduce novel biomarkers for the preemptive intervention in CVD.

Acknowledgements The CHIRON research programme has received funding from the ARTEMIS Joint Undertaking under grant agreement no 100228 and from the Governments of the eight European countries participating to the project.
The MICHELANGELO project is co-financed by the European Commission within the seventh Research Framework Program (FP7-ICT) (Grant Agreement #288241).

References

Altemeier WA., Altemeier LE. (2009), How can early, intensive training help a genetic disorder? Pediatr Ann, 38:167–170
ARTEMIS Joint Undertaking (2010) The book of projects, pp 56–61. www.artemis-ju.eu/publications
ARTEMIS Joint Undertaking (2011) ARTEMIS strategic research agenda. www.artemis-ju.eu

Atladottier HO et al (2007) Time trends in reported diagnoses of childhood neuropsychiatric disorders: a Danish cohort study. Arch Pediatr Adolesc Med 161(2):193–198
Barca CC, Rodriguez JM, Rugnone A et al (2012) Medical expert support tool (MEST): a person-centric approach for healthcare management. In: 2012 international conference on smart homes and health telematics (ICOST), Artiminio, pp 99–106
Biswas D, Mazomenos EB, Maharatna K (2012) ECG compression for remote healthcare systems using selective thresholding based on energy compaction. In: 2012 international symposium on signals, systems, and electronics (ISSSE), Seoul, 3–5 Oct 2012, pp 1-6
Bonfiglio S (2010) The CHIRON project. Artemis Mag 7:35–38
Bonfiglio S (2011) CHIRON. . .Fostering a continuum of care. Artemis Mag 10:24–27
Burns A, Greene BR, McGrath MJ, O'Shea TJ, Kuris B, Ayer SM, Stroiescu F, Cionca V (2010) SHIMMER™—a wireless sensor platform for noninvasive biomedical research. IEEE Sens J 10(9):1527–1534
Cosmed K4b [Online] http://www.cosmed.it/index.php?option=com_content&view=article&id=249&Itemid=174&lang=en
Dawson G. (2008) Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev Psychopathol, 20:775–804
Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, Donaldson A, Varley J (2010) Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics, 125:17–23
Delaherche E, Chetouani M (2010) Multimodal coordination: exploring relevant features and measures. Presented at *SSPW'10*, 29 Oct 2010, Firenze, Italy
Dickstein K, Cohen-Solal A, Filippatos G et al (2008) ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the task force for the diagnosis and treatment of acute and chronic heart failure 2008 of the European society of cardiology. Eur Heart J 29:2388–2442
European Autism Information System (EAIS) (2006–2008) Final project report (Note: The EAIS Project was promoted by the European Autism Alliance (EAA) and funded by the European Commission)
Fielding RT (2000) Architectural styles and the design of network-based software architectures. Ph.D. Dissertation, University of California, Irvine
Frost & Sullivan (2009) Preparing for an aging society: challengers faced by healthcare system in European Union, Japan and United States, 2009 [Online] http://www.frost.com/
Gjoreski H (2011) Adaptive human activity recognition and fall detection using wearable sensors. MSc Thesis, Jožef Stefan International Postgraduate School
Klersy C, De Silvestri A, Gabutti G et al (2009) A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 54:1683–1694
Knapp M, Romeo R, Beecham J (2007) The economic consequences of autism in the UK. Foundation for People with Learning Disabilities, London
Kononenko I (1994) Estimating attributes: analysis and extensions of RELIEF. In: European conference on machine learning, Catania, pp 171–182
Kuehn BM (2007) Autism spectrum disorders common. JAMA 297(9):940
Laibow RE, Stubblebine AN, Sandground H, Bounias M (2001) EEG NeuroBioFeedback treatment of patients with brain injury—Part 3: cardiac parameters and finger temperature changes associated with rehabilitation. Presented at the 2001 Society for Neuronal Regulation, ninth annual conference, Monterey, CA
Levis P, Madden S, Polastre J, Szewczyk R, Whitehouse K, Woo A, Gay D, Hill J, Welsh M, Brewer E, Culler D (2005) TinyOS: an operating system for sensor networks. In: Weber W, Rabaey JM, Aarts E (eds) Ambient intelligence. Springer, Berlin, pp 115–148, Chapter 7
Luštrek M, Cvetković B, Kozina S (2012) Energy expenditure estimation with wearable accelerometers. IEEE international symposium on circuits and systems, Seoul, pp 5–8

Mazomenos EB, Biswas D, Acharyya A, Chen T, Maharatna K, Rosengarten J, Morgan J, Curzen N (2013) A low-complexity ECG feature extraction algorithm for mobile healthcare applications. IEEE Trans Inf Technol Biomed 17(2):459–469

Monk TH, Buysse DJ, Reynolds CF III, Kupfer DJ, Houck PR (1995) Circadian temperature rhythms of older people. Exp Gerontol 30(5):455–474

Oberman LM., Pascual-Leone (2008). Cortical plasticity: A proposed mechanism by which genomic factors lead to the behavioral and neurological phenotype of autism spectrum and psychotic-spectrum disorders. Behavioral Brain Sciences, 31:276–277

Pantelopoulos A, Bourbakis NG (2010) A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans Syst Man Cybern C Appl Rev 40(1):1–12

Pediatrics. October 2009 issue. http://pediatrics.aappublications.org/content

Puddu PE, Morgan JM, Torromeo C, Curzen N, Schiariti M, Bonfiglio S (2012) A clinical observational study in the CHIRON project: rationale and expected results. Full paper accepted at ICOST 2012

Remington B, Hastings RP, Kovshoff H, Degli Espinosa F, Jahr E, Brown T, Alsford P, Lemaic M, Ward N (2007) Early intensive behavioral intervention: outcomes for children with autism and their parents after two years. Am J Ment Retard 112(6):418–438

Sensirion [Online]. http://wwww.sensirion.com

Sund-Levander M, Forsberg C, Wahren LK (2002) Normal oral, rectal, tympanic and axillary body temperature in adult men and women: a systematic literature review. Scand J Caring Sci 16(2):122–128

Sund-Levander M, Grodzinsky E, Loyd D et al (2004) Errors in body temperature assessment related to individual variation measuring technique and equipment. Int J Nurs Pract 10:216–223

Tapia EM (2008) Using machine learning for real-time activity recognition and estimation of energy expenditure. Ph.D. Thesis, Massachusetts Institute of Technology

Žbogar M, Gjoreski H, Kozina S, Luštrek M (2012) Improving accelerometer based activity recognition. In: 15th International multiconference information society, Pisa, pp 167–170