Annals of Information Systems 19

Ashish Gupta
Vimla L. Patel
Robert A. Greenes *Editors*

Advances in Healthcare Informatics and Analytics

Springer

Annals of Information Systems

Volume 19

Series Editors
Ramesh Sharda
Oklahoma State University
Stillwater, OK, USA

Stefan Voß
University of Hamburg
Hamburg, Germany

More information about this series at www.springer.com/series/7573

Ashish Gupta • Vimla L. Patel
Robert A. Greenes
Editors

Advances in Healthcare Informatics and Analytics

Springer

Editors
Ashish Gupta
College of Business
University of Tennessee Chattanooga
Chattanooga, TN, USA

Robert A. Greenes
Department of Biomedical Informatics
Arizona State University
Scottsdale, AZ, USA

Vimla L. Patel
Center for Cognitive Studies in Medicine and Public Health
New York Academy of Medicine
New York, NY, USA

ISSN 1934-3221 ISSN 1934-3213 (electronic)
Annals of Information Systems
ISBN 978-3-319-23293-5 ISBN 978-3-319-23294-2 (eBook)
DOI 10.1007/978-3-319-23294-2

Library of Congress Control Number: 2015954173

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

The field of biomedical informatics is undergoing rapid transformations due to recent developments in healthcare information technologies (HIT), new healthcare acts, federal regulations, new healthcare financing model, and the growth of data science. This is collectively having a tremendous impact on various care delivery processes, clinical decision-making, and outcomes. Due to the complexity of healthcare operations, data, and multiple stakeholders (e.g., patient, provider, payers), the impacts of recent innovations and transformations in this area require increased systematic study and research.

This volume addresses these issues and contains an important collection of articles that present recent research in healthcare information technology and analytics. In particular, the following key thematic areas are addressed.

A well-designed HIT system can have a multifaceted impact on various stakeholders and can require a deep-dive understanding of behavioral, cognitive, and perceptual factors. For example, Kato-Lin et al. investigate the impact of technology failure on electronic prescribing behavior in primary care; Dinev et al., on the other hand, focus on understanding user attitudes toward electronic health records from a privacy perspective; Thrasher studies the impact of information technology integration on integrated delivery systems; and Petter et al. evaluate emergency response medical information systems.

It is critical to reduce errors and improve patient safety. This can be done through various mechanisms. One approach is to design new and improved clinical systems or procedures and validate them. An excellent example of this is the paper by Myneni et al., which proposes a new risk assessment framework for critical care units and validates it in clinical settings. Another approach is to improve medical training through the use of persuasive technologies. For example, Khanal and his colleagues provide a comprehensive review of the historical work, as well as the state-of-art virtual world application for improving medical training.

Systematically evaluating various decision outcomes and using modeling approaches to gain better understanding of different care scenarios also play a vital role in reducing various medical errors. For example, Frezee et al. utilize a latent growth modeling approach to understand lifestyle decisions based on patient historical data, Ramsey et al. investigate the dynamic decision-making tasks of primary care physicians treating patients with chronic conditions such as type 2 diabetes, and Bosire et al. design an integrated surgical care delivery system using axiomatic design and petri net modeling approaches.

Various healthcare processes and operations can significantly benefit from automated processes. The study by Loy and his colleagues validates the use of robotic operations for better inventory management and supply chain control in a health system pharmacy.

Big data and analytic approaches hold great promise for the biomedical field. The adoption of such approaches should lead to the development of next-generation innovations in healthcare and technology and serve as a guide to policy making. As our ability to process complex data increases due to next-generation computational infrastructure and the field of data sciences grows, we will see increased applications of real-time applications (such as in stroke) and smart systems such as cognitive systems integrated seamlessly into healthcare processes and technology. This clearly warrants a more aggressive research agenda. Montero et al. studied analytics of decision-support theoretic assistants based on contextual gesture recognition. This is a good example of integrating affective computing and analytics for patients requiring rehabilitation. The study by Scotch et al. uses natural language processing for understanding contraceptive use at the VA and is a good application of processing unstructured healthcare data. Scalability of such data sets requires shifting toward big data platforms and will likely facilitate the effective discovery of patterns.

Papers such as those collected in this volume indicate the kinds of research that are needed as healthcare transformation continues. Finally, we thank all the authors for making important contributions to this collection and the large pool of reviewers who provided valuable time to review the manuscripts.

Chattanooga, TN, USA — Ashish Gupta
New York, NY, USA — Vimla L. Patel
Scottsdale, AZ, USA — Robert A. Greenes

Contents

Contributors

Khalid Almoosa M.D. Divisions of Critical Care Medicine, UT Health, Houston, TX, USA

Valentina Albano Department of Business Studies, Roma Tre University, Rome, Italy

George Biltz School of Kinesiology, University of Minnesota, Minneapolis, MN, USA

Joshua Bosire Cooper University Hospital, Camden, New Jersey, USA

Cynthia Brandt M.D. M.P.H. VA Connecticut Healthcare System, West Haven, CT, USA

Yale Center for Medical Informatics, New Haven, CT, USA

Jesse Breidenbach Pharm.D. Sanford Health Medical Center Fargo, Fargo, ND, USA

Trevor Cohen M.B.Ch.B., Ph.D. Center for Cognitive Informatics and Decision Making, School of Biomedical Informatics, UT Health, Houston, TX, USA

Alessandro D'Atri Late Director of the Research Center on Information Systems (CeRSI), LUISS University of Rome, Rome, Italy

Tamara Dinev Department of Information Technology and Operations Management, College of Business, Florida Atlantic University, Boca Raton, FL, USA

Toni Fera Independent Healthcare Consultant, Pittsburgh, PA, USA

Ronald Freeze University of Arkansas, Fayetteville, AR, USA

Ann Fruhling School of Interdisciplinary Informatics, College of Information Science & Technology, University of Nebraska at Omaha, Omaha, NE, USA

Tejas Gandhi Navicent Health, Macon, Georgia, USA

Ashish Gupta Ph.D. College of Business, University of Tennessee Chattanooga, Chattanooga, TN, USA

Paul Hart Department of Information Technology and Operations Management, College of Business, Florida Atlantic University, Boca Raton, Florida, USA

Paul E. Johnson Carlson School of Management, University of Minnesota, Minneapolis, MN, USA

Keith T. Kanel M.D., M.H.C.M. Pittsburgh Regional Health Initiative, Pittsburgh, PA, USA

Prabal Khanal Ph.D. 3D Systems Corporation, Andover, MA, USA

Mohammad Khasawneh Department of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, NY, USA

Pratap Kotala M.S. College of Science and Mathematics, North Dakota State University, Fargo, ND, USA

Sylvia Leung VA Palo Alto Health Care System, Palo Alto, CA, USA

Yi-Chin Kato-Lin M.S. Frank G. Zarb School of Business, Hofstra University, 134 Hofstra University, Hempstead, NY 11549, USA

Maari L. Loy Pharm.D., M.B.A. Sanford Health Medical Center Fargo, Fargo, ND, USA

Miriam Martínez Acapulco Institute of Technology, Av. Instituto Tecnológico S/N, Crucero del Cayaco, C.P. 39905, Acapulco, Guerrero, Mexico

Debra McGinnis R.R.T. Memorial Hermann Hospital, Houston, TX, USA

José Antonio Montero Acapulco Institute of Technology, Av. Instituto Tecnológico S/N, Crucero del Cayaco, C.P. 39905, Acapulco, Guerrero, Mexico

Sahiti Myneni M.S.E. Center for Cognitive Informatics and Decision Making, School of Biomedical Informatics, UT Health, Houston, TX, USA

Robert Nelson Pharm.D., B.C.P.S. Sanford Health Medical Center Fargo, Fargo, ND, USA

Patrick J. O'Connor HealthPartners Institute for Education and Research, Minneapolis, MN, USA

Rema Padman Ph.D. The H. John Heinz III College, Carnegie Mellon University, Pittsburgh, PA, USA

Bela Patel M.D. Divisions of Critical Care Medicine, UT Health, Houston, TX, USA

Vimla L. Patel Ph.D., D.Sc. Center for Cognitive Studies in Medicine and Public Health, New York Academy of Medicine, New York, NY, USA

Stacie Petter Department of Information Systems, Hankamer School of Business, Baylor University, One Bear Place #98005, Waco, TX 76798, USA

T.S. Raghu Arizona State University, Tempe, AZ, USA

Gregory W. Ramsey The Earl G. Graves School of Business and Management, Morgan State University, Baltimore, MD, USA

Monte Roemmich Pharm.D. Sanford Health Medical Center Fargo, Fargo, ND, USA

William A. Rush HealthPartners Institute for Education and Research, Minneapolis, MN, USA

Matthew Scotch Ph.D., M.P.H. VA Connecticut Healthcare System, West Haven, CT, USA

Department of Biomedical Informatics, Arizona State University, Samuel C. Johnson Research Bldg, 13212 East Shea Boulevard, Scottsdale, AZ 85259, USA

Marshall Smith Banner Health, Phoenix, AZ, USA

JoAnn M. Sperl-Hillen HealthPartners Institute for Education and Research, Minneapolis, MN, USA

Krishnaswami Srihari Department of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, NY, USA

Luis E. Sucar Department of Computer Science, National Institute of Astrophysics, Optics and Electronics, Tonanzintla, Mexico

Evelyn Thrasher Gordon Ford College of Business, Western Kentucky University, Bowling Green, KY, USA

Rodney D. Traub Ph.D. College of Business, North Dakota State University, Fargo, ND, USA

Ajay Vinze Arizona State University, Tempe, AZ, USA

Shengyong Wang Department of Mechanical Engineering, University of Akron, Akron, OH, USA

Julie Womack C.N.M., A.P.R.N., Ph.D. VA Connecticut Healthcare System, West Haven, CT, USA

Heng Xu College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802, USA

Limin Zhang Ph.D. College of Business, North Dakota State University, Fargo, ND, USA

Chapter 1
The Impact of Technology Failure on Electronic Prescribing Behavior in Primary Care: A Case Study

Yi-Chin Kato-Lin[1], **Rema Padman**[2], **Keith T. Kanel**[3], **and Toni Fera**[4]

[1]Frank G. Zarb School of Business, Hofstra University, 134 Hofstra University, Hempstead, NY 11549, USA

[2]The H. John Heinz III College, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

[3]Pittsburgh Regional Health Initiative, Pittsburgh, PA, USA

[4]Independent Healthcare Consultant, Pittsburgh, PA, USA

Abstract Electronic Prescribing (e-Rx) has significant potential to improve quality of care and reduce medication errors. However, its adoption rate in primary care has been slow for a variety of reasons. We examine the adverse impact of an information technology (IT) failure on the prescribing process as a critical reliability barrier to adoption. Data from Allscripts TouchWorks® database containing prescriptions written by six physicians in two primary care settings were analyzed using a statistical change-point detection algorithm to identify the tipping point in actual usage and subsequent trends in usage behavior. Physicians overwhelmingly switched from electronic transmission of prescriptions to print option in the presence of such a failure. We propose an approach for a control system that will allow for early detection of system failures and rapid process improvement, and discuss implications for handling such failures in the rapidly evolving IT-enabled healthcare delivery context.

Keywords Electronic prescribing • Reliability of information technology • Change point detection • CUSUM control chart

---

R. Padman, Ph.D. (✉)
The H. John Heinz III College, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA
e-mail: rpadman@cmu.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*, Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_1

1.1 Introduction

Electronic Prescribing (e-Rx) technology enables physicians to transmit prescriptions directly to the pharmacy of choice in electronic form rather than submitting handwritten prescriptions carried by patients (Bell et al. 2004). It is widely perceived that this technology reduces preventable medication errors and improves the safety, efficiency and costs related to medication prescribing and dispensing processes (Bell et al. 2004; McKibbon et al. 2011; Donyai et al. 2008; Eslami et al. 2007; Jani et al. 2008; Moniz et al. 2011; Fischer et al. 2008a; van Doormaal et al. 2009; Ammenwerth et al. 2008; Cusack 2008; eHealth Initiative 2004). A study conducted by SureScripts and Walgreens also revealed that the use of e-Rx increased medication pick-up rate by 11 % (as opposed to handwritten or fax prescriptions) (Walgreens and Health Press Release 2007). However, despite these potential benefits, the adoption rate of electronic prescribing remains low and slow, especially in outpatient settings, where the bulk of prescriptions are generated. In 2005, 22 % of physicians were using e-Rx (Grossman and Reed 2006). At the end of 2010, a half decade later, the adoption rate was at 36 % (Surescripts 2011). After almost 4 years of the incentive program established by the Health Information Technology for Economic and Clinical Health (HITECH) Act in 2009, the adoption rate among prescribers, including physicians, nurse practitioners, and physician assistants, has reached only 54 % (Surescripts 2013).

To understand barriers to adoption, Grossman et al. interviewed current e-Rx prescribers and concluded that unreliable connectivity between prescribers and pharmacies is a major barrier to electronic transmission for renewals and mail-orders, but not for new prescriptions for which transmission failures were rarely found (Grossman et al. 2012). However, as will be discussed later, the latter conclusion might still be arguable because opinions from users who opt-out due to transmission problems were excluded from the study. In addition to underlining the importance of developing a reliable mechanism to report "adverse effect", Miller et al. also advocate strengthened coordination among stakeholders, such as government, vendors, and providers, to achieve the goal of e-Rx (Miller et al. 2005). However, no actual impact of adverse effect has been demonstrated, and coordination mechanisms are yet to be developed and implemented.

Other studies have also identified or analyzed the major barriers to adoption, such as deployment cost, workflow concerns, perceived value, end-user constraints, organizational culture, and technology reliability and capability (Cusack 2008; Grossman et al. 2007, 2012; Miller et al. 2005; Friedman et al. 2009; Hollingworth et al. 2007). Yet too little is documented in the literature about the impact of technology failures on adoption and continued use to draw this industry's attention. Corresponding to the conclusion drawn by Fischer et al. (2008b), this study also finds slowly increasing pattern and low uptake of electronic prescribing, and further shows that this uptake is vulnerable to technology failures and could collapse very quickly.

This case study examines the overwhelming impacts of an actual communication failure in the transmission of electronic prescriptions on physicians' usage behavior and subsequent technology de-adoption in the ambulatory care setting. We report insights from a retrospective analysis of e-Rx data to determine *when* the failure event could have happened and suggest a process monitoring approach to mitigate the negative impacts of technology failures on adoption of information technologies. We anticipate that, by highlighting technology problems and suggesting possible solutions, this study can contribute to higher adoption rates for electronic prescribing and a more reliable electronic transmission environment resulting in enhanced patient safety.

1.2 Case Description

1.2.1 Study Setting

Our study sites were two primary care practices affiliated with a major academic medical center in Pittsburgh, with a total of six physicians. Both practices deployed the stand-alone Allscripts TouchWorks® e-Rx application in 2005, one on June 4 and the other on July 16. The schematic of the implementation is shown in Fig. 1.1.

Physicians either wrote prescriptions using an HP iPaQ Personal Digital Assistant (PDA) as the client interface to the Allscripts TouchWorks® e-Rx application or used a paper prescription pad. Furthermore, using the PDA, one of two transmission options, *Send* or *Print*, was executed for each prescription, primarily based on patient preferences, physician perceptions or technology constraints in the use of the system. While *Send* transmits the prescription electronically to the selected retail pharmacy, *Print* generates a paper copy of the prescription that patients can take to the pharmacy. The electronic *(Send)* prescription is routed from the wireless PDA via Allscripts' database to the destination pharmacy via the retail pharmacy consortium, SureScripts, network. Figure 1.2 depicts the prescribing and transmission processes in the two medical practices.

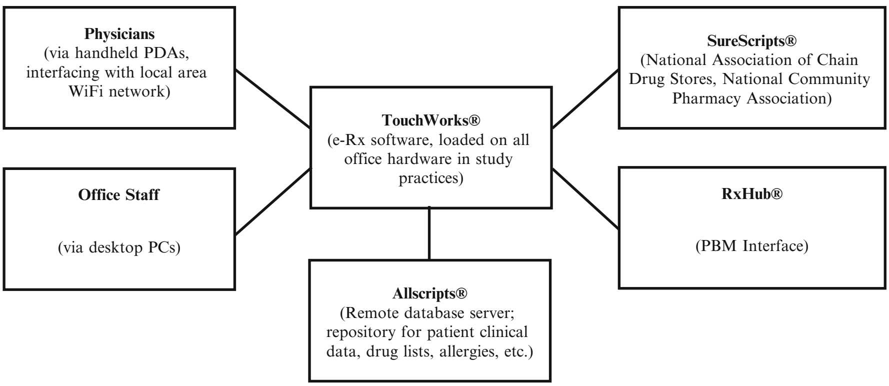


Fig. 1.1 Implementation architecture in study

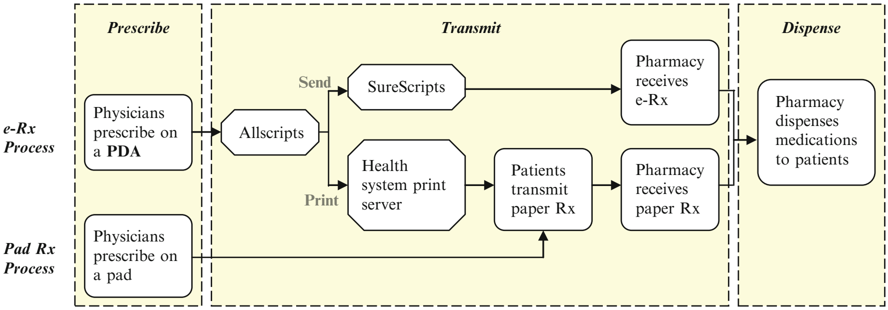


Fig. 1.2 The prescribing and transmission processes at the study sites

1.2.2 The Failure Event

In Fall 2006, a series of patient complaints alerted the practices to a failure in the e-Rx transmission flow, primarily to a high volume retail pharmacy chain (Pharmacy A), which handled 38 % of the prescriptions transmitted electronically by the practices. Interviews of clinicians and staff supported this early anecdotal evidence that patients had difficulty getting their prescriptions filled at the concerned pharmacy because their prescriptions were found to be "missing". However, due to the absence of confirmed failure reports, physicians continued to submit prescriptions electronically to Pharmacy A for a while longer, but patient complaints soon resulted in declining use of the electronic transmissions as physicians switched to either printing or writing the prescriptions.

1.2.3 Hypotheses

Given the above scenario, we hypothesize three possible effects or user responses resulting from the connectivity problem (failure mode): (1) decline in the use of PDA by switching to writing the prescriptions on a paper pad (de-adopt PDA), (2) decline in the use of PDA-initiated prescriptions for electronic transmission by switching to printing these prescriptions (de-adopt e-Rx), and (3) decline in the use of both PDA and electronic transmission (de-adopt PDA and e-Rx). Our hypotheses regarding the user response to the failure mode is summarized in Table 1.1.

We test these hypotheses to determine the specific scenario applicable to this case study using data from Allscripts®. We further apply this knowledge about the scenario to develop our process monitoring approach. We analyzed the changing trends in the time series using two evaluation metrics, *%Send* (calculated as the ratio of number of prescriptions transmitted electronically to a retail pharmacy to the number of PDA-initiated prescriptions) and *#PDA* (the number of PDA-initiated prescriptions).

Table 1.1 Trend hypotheses and effects of technology failure

User response	Trend hypotheses		Example effects on prescribing
	%Send	*#PDA*	
Scenario 1: De-adopt PDA	→	↘	Physician declines use of PDA, but maintains same % of electronic transmission
Scenario 2: De-adopt e-Rx	↘	→	Physician continues to use PDA to print and track prescriptions, but does not transmit electronically to the pharmacy
Scenario 3: De-adopt PDA and e-Rx	↘	↘	Physician manually writes prescriptions
No change or increase	→ ↗	→ ↗	Process in control or increase in adoption over time

1.2.4 Data

Our data set included 36,352 prescriptions generated via desktop (PC) and PDA transactions using Allscripts TouchWorks® from June 8, 2005 to January 31, 2007. Each prescription record included the drug ID, drug name, prescription date, de-identified physician ID, whether the prescription originated on a PC or PDA, whether the prescription was sent electronically or printed, and destination pharmacy ID. Interviews with physicians at our study sites revealed that when writing electronic prescriptions, they use a PDA for prescribing during the encounters and a PC for prescription refills signed off at the end of the day. Since our focus is on new prescriptions, we excluded 21,158 non-PDA generated prescriptions. We also excluded 76 prescriptions written by physicians who left the practices sometime during our study period. We aggregated the pharmacy categories from 91 to 5 to represent the four largest chain pharmacies, named A, B, C, and D, respectively, and the remaining pharmacies were grouped as Other.[1] We also aggregated the daily prescription records to weekly data. Each week in our data set starts on a Thursday and ends on a Wednesday. Finally, we excluded the weeks in Summer 2005, when only one practice had implemented e-Rx, eliminating 1049 prescriptions. This data preprocessing resulted in an evaluation dataset of 14,069 prescriptions written by six physicians using the PDA over 80 weeks, from July 21, 2005 to January 31, 2007. Figures 1.3 and 1.4 display this time series data of physician prescribing decisions at the point of care for *#PDA* and *%Send* over the 80-week study period.

[1]Pharmacy A accounted for 38.0 % of all the prescriptions transmitted electronically to the pharmacies, with 17.2 %, 11.7 %, 4.4 %, 28.7 % for Pharmacies B, C, D, and Other, respectively.

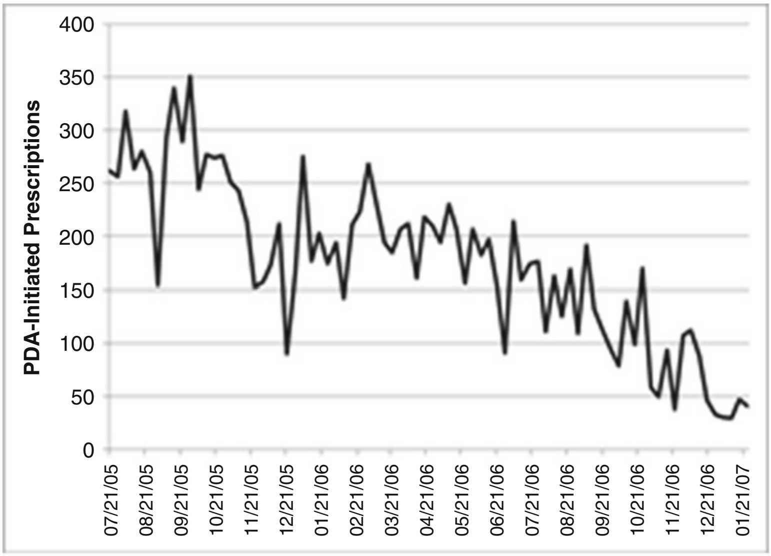


Fig. 1.3 Number of PDA-initiated prescriptions (*#PDA*) over 80 weeks

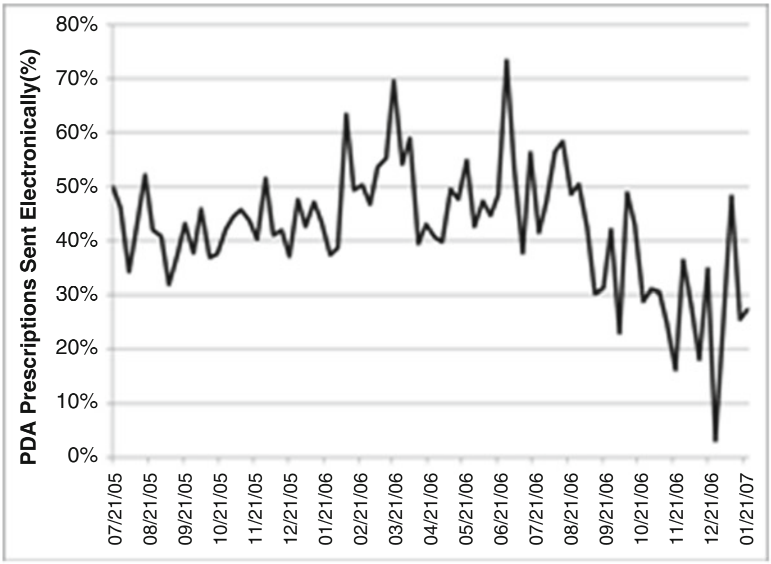


Fig. 1.4 Proportion of PDA prescriptions sent electronically (*%Send*) over 80 weeks

1.3 Methods

Physicians and staff at the study sites all agreed that the failure event happened in Fall 2006 but the exact, or even an approximate, date was not known. We first estimate this date by applying a statistical breakpoint detection method developed by Bai and Perron (1998, 2003) to identify the week with the most significant shift in prescribing behavior using the time series data associated with *%Send* and *#PDA*. Furthermore, using the prescribing process prior to the breakpoint as an "in-control" process for parameter estimation, we develop a Cumulative Sum (CUSUM) control chart approach to monitor *#PDA* and *%Send.* These control charts are subsequently used to detect abnormal usage behavior in the two practices in the post-breakpoint period.

1.3.1 Change Point Detection Technique

We applied Bai and Perron's model (Bai and Perron 1998, 2003) to determine multiple structural changes. The original model involves multiple linear regressions with partial structural change, meaning that only some of the coefficients are subject to shifts. The model is reduced when we apply it to our data, since we only have one independent variable in our linear regression. Therefore, with m breaks ($m + 1$ regimes), the pure structural change model is presented as:

$$y_t = z_t'\delta_j + u_t \qquad t = T_{j-1} + 1, \ldots, T_j \quad j = 1, \ldots, m + 1 \tag{1.1}$$

and the matrix form of (1.1) as

$$Y = \overline{Z}\delta + U$$

where $Y = (y_1, \ldots, y_T)'$, $U = (U_1, \ldots, U_T)'$, $\delta = \left(\delta_1', \ldots, \delta_{m+1}'\right)'$, and $\overline{Z}$ is the matrix which diagonally partitions Z at $(T_1, \ldots, T_m)$, i.e., $\overline{Z} = diag\,(Z_1, \ldots, Z_{m+1})$ with $Z_i = (z_{T_{i-1}+1}, \ldots, z_{T_i})'$. The estimate of δ_j is then obtained by minimizing the residual sum of squares (*RSS*) for each m-partition $(T_1, \ldots, T_m)$:

$$\left(Y - \overline{Z}\delta\right)'\left(Y - \overline{Z}\delta\right) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} \left[y_t - z_t'\delta_i\right]^2$$

Thus, denoting the *RSS* of the estimated break points $\left(\widehat{T}_1, \ldots, \widehat{T}_m\right)$ as $S_T(T_1, \ldots, T_m)$, the overall estimated break points are obtained by

$$\left(\widehat{T}_1, \ldots, \widehat{T}_m\right) = argmin_{T_1, \ldots, T_m} S_T\,(T_1, \ldots, T_m)$$

We apply this model to determine both the *number* and *location* of breaks, if there are any, in the coefficients of linear regression data. However, in this study, both of the *#PDA* and *%Send* time series violated one of the key assumptions of linear models, that of independent observations. The value of each metric in any week was correlated with the value in the previous week. Therefore, an *AR*(*1*) model (autoregressive model with order 1) was used to account for autocorrelation before applying Bai and Perron's model on *#PDA* and *%Send*. Hence, the $z_t^{'}$ in (1.1) is substituted with y_{t-1}' and Z_i represents $(y_{T_{i-1}}, \ldots, y_{T_i-1})'$.

This model was then used to indicate where to split a series in order to fit individual *AR*(*1*) models in each regime, separated by breakpoints, such that the *RSS* can be minimized. By applying the procedure, we first determined the break point weeks in the entire process by minimizing *RSS* given each possible number of break points. Then we chose the optimal number of break points by using the *Bayesian Information Criterion (BIC)*, which selects the model that best describes our time series, subject to a minimal number of break points (Ramsey and Schafer 2002).

1.3.2 Cumulative Sum Control Chart

CUSUM is a statistical process control technique that is superior to the original control chart, such as Shewhart chart, because it improves the sensitivity of detection by accumulating deviations prior to the evaluated break point (Devore 2003; Page 1961). It is commonly used to monitor a process and send an alert when the process experiences abnormal patterns, or is out of control. To demonstrate how *CUSUM* can be used to prospectively detect the abnormal process of transmitting *e-Rx*, we applied the method to determine the out-of-control events over the duration of the study period, by assuming the process before the breakpoint as the in-control process.

A process is defined to be out-of-control when an event significantly deviates from the target mean, μ_0, and the sensitivity of *CUSUM* is determined by two parameters, α and δ. δ denotes the amount of shift in the target mean that we want to detect as a multiplier of the sample standard deviation, σ, and α is the tolerance in *Type I error*. Such a scheme is able to raise an alarm when the observed event is $\delta(\sigma)$ away from μ_0, with the probability of a false alert being α. The more sensitive the designed scheme, the faster the shifts can be detected, but higher is the occurrence of false alerts. Therefore, it is a decision makers' choice to determine the target mean, or the in-control process, and the two parameters. Finally, a *V-Mask* was used to visualize the monitored process (Devore 2003). Observations falling below the lower arm or above the higher arm of the *V-Mask* suggest that they are significantly higher or lower than the target value. Both situations indicate the process is out-of-control.

However, the serially correlated *#PDA* and *%Send* series both violate the independence assumption of CUSUM chart. To account for this problem, a residual-based control chart was used where the serially uncorrelated residuals were monitored rather than the original observations.

1.4 Results

1.4.1 Prescribing Behavior Change Points

The week of September 14, 2006 was determined by the Bai and Perron model as the single optimal break point in *%Send* time series ($BIC = 589.53, RSS = 5779.78$), visualized in Fig. 1.5. The results indicate that electronic transmission of PDA prescriptions as a proportion of the total prescriptions on the PDA had reached a high of 50 % prior to the failure, but dropped to less than 30 % after September 14, 2006. This was identified as the most significant decline in electronic transmissions over the study period.

Using *#PDA* as the evaluation metric, the change point technique identified two optimal break points, namely, the weeks of November 17, 2005 and September 14, 2006 ($BIC = 844.93$, $RSS = 124{,}126.64$), as shown in Fig. 1.6. The total number of

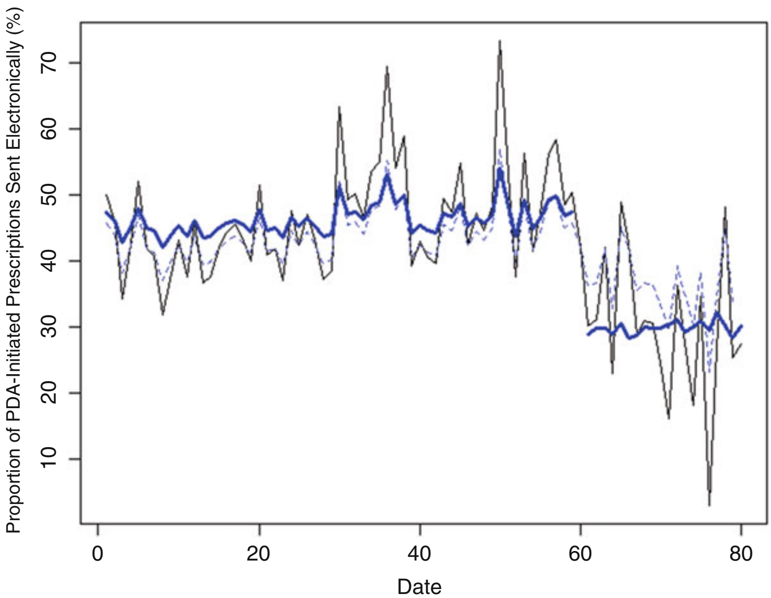


Fig. 1.5 *Dotted lines* represent the AR(1) model of the *%Send* time series. The *bold solid lines* are models of the two periods, before and after the change point, at week 61 (September 14, 2006), determined by change point detection technique

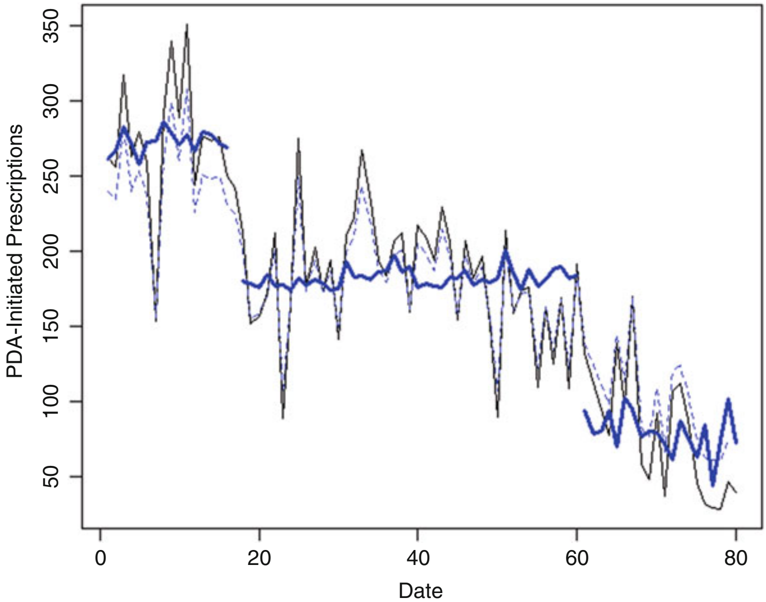


Fig. 1.6 *Dotted lines* represent the AR(1) model of the *#PDA* time series. The *bold solid lines* are models of the three periods, before and after each of the two break points, at weeks 18 (November 17, 2005) and 61 (September 14, 2006), respectively, determined by change point detection technique

PDA-initiated prescriptions written by the six physicians was around 275 per week at the beginning of the study period; it declined to about 180 per week in Fall 2005, and reached a steady state at the same level until the failure event. After September 14, 2006, it dropped significantly to around 80 per week, which was only 30 % of the volume in the beginning. Thus there was an overwhelming shift in prescribing behavior from PDA-initiated to a paper pad in Week 61, indicating a break point in early September of year 2006. Since records were not maintained regarding all failure events over the study period, it was not possible to identify the causes of the first break point, on November 17, 2005. However, in this study, supported by the same turning point revealed from both metrics, we argue that it is highly likely that the reason for the change in the week of September 14, 2006 was the technology failure incident. Furthermore, since the results indicate that both *#PDA* and *%Send* time series declined after the break point, our third hypothesis, de-adopt PDA and de-adopt e-Rx as a reaction to the IT failure, is validated.

1.4.2 Pharmacy Market Share

Although this communication failure was caused by the server problem of one single chain pharmacy, the physicians hesitated to submit *e-Rx* to all of the pharmacies in

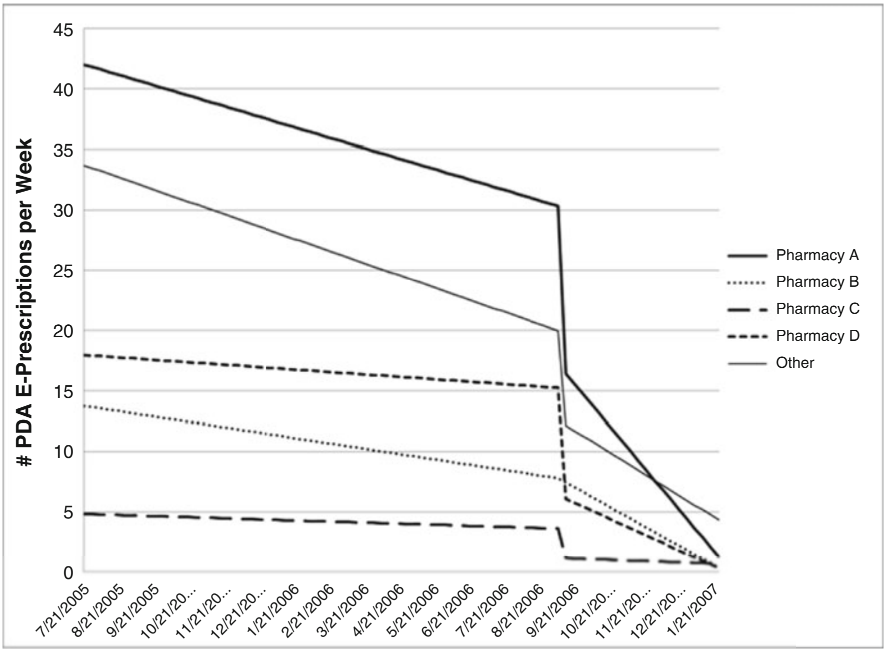


Fig. 1.7 Number of PDA electronic prescriptions to each pharmacy chain, by week (fitted values by simple linear regressions)

the region. Subsequent inquiry determined that there was a server error at Pharmacy A that caused the transmission error. As shown in Fig. 1.7, the drastic decline in the number of electronic prescriptions submitted by physicians after the break point (the week of September 14, 2006) occurred to not only Pharmacy A, but also to Pharmacies B, C, and Other. Ironically, when we further examine the *e-Rx* market share across the pharmacies, the share of Pharmacy A was not affected but instead, the shares of Pharmacy B and D slumped (Fig. 1.8). Although the decrease in *e-Rx* market share does not necessarily indicate a decrease in revenue, the 11 % additional revenue normally generated by *e-Rx* would possibly be lost (Walgreens and Health Press Release 2007). Thus, although caused by a single pharmacy, the transmission failure did change the structure of *e-Rx* market share and potentially diminished pharmacies' revenue as a whole.

1.4.3 Monitoring System

For the purposes of demonstrating the methodology, we regarded the period from March 9, 2006 to September 13, 2006, a 6-month period prior to the breakpoint, as the in-control process. By taking the process average during the in-control process,

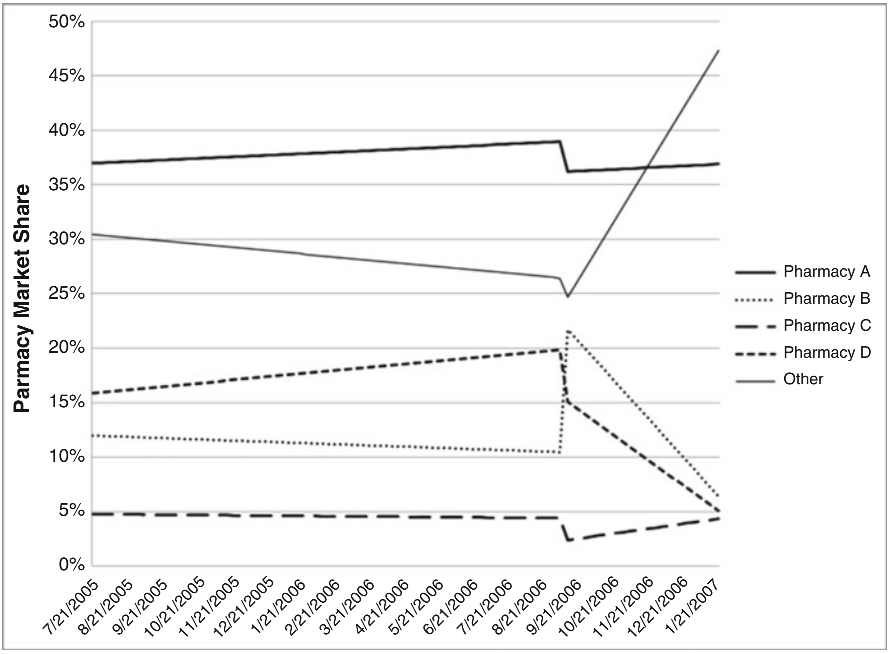


Fig. 1.8 Pharmacy market share of PDA electronic prescriptions over study period (fitted values by simple linear regressions)

the target means (of residuals) were found to be 5.14 % for the *%Send* and 15.74 for the *#PDA* process. The *CUSUM* scheme in this study was designed to detect an abnormal value 3σ away from the target mean, while allowing the probability of *Type I error* to be 0.01. Figures 1.9 and 1.10 demonstrate the *CUSUM V-Masks* for *%Send* and *#PDA*, respectively, using this scheme.

We used this approach to detect out-of-control events in the post-breakpoint period, from September 14, 2006 to January 31, 2007. Out of control events in *%Send* time series were detected in the weeks of October 5, 2006, November 23, 2006, December 14, 2006, December 28, 2006, January 4, 2007, and January 18, 2007, and all of them were downward shifts. The *#PDA* process was out-of-control only in the week of November 9, 2006, also in a downward shift. This suggested that the physicians significantly increased the proportion of printed prescriptions in the week of October 5, 2006, followed by a dramatic decrease in the use of PDA on November 9, 2006.

1.5 Discussion

Currently, the electronic prescribing process is composed of several data transmission steps managed by different independent organizations. As shown in Fig. 1.11, after being prescribed by a physician, an *e-Rx* needs to be processed by provider's

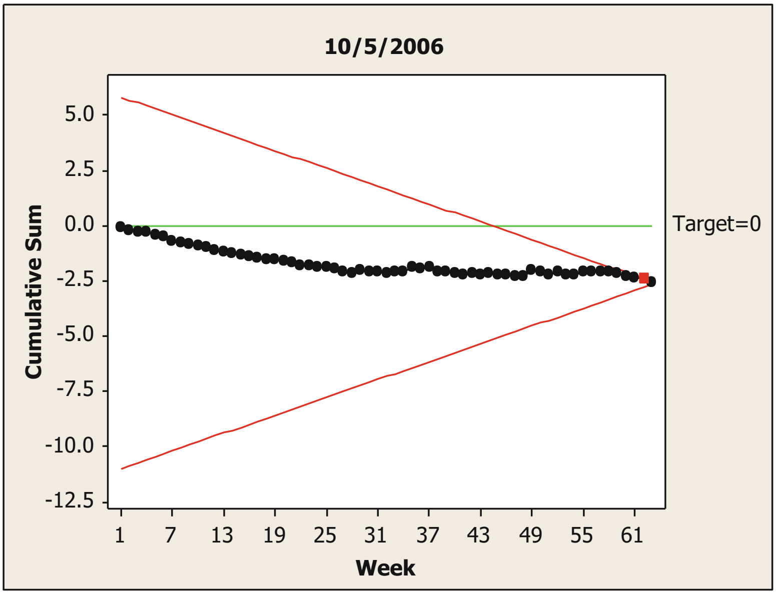


Fig. 1.9 The first out-of-control event in *%Send* series was in week 64 (October 5, 2006). Then several alerts were found between November 23, 2006 and January 18, 2007 (not shown in this graph)

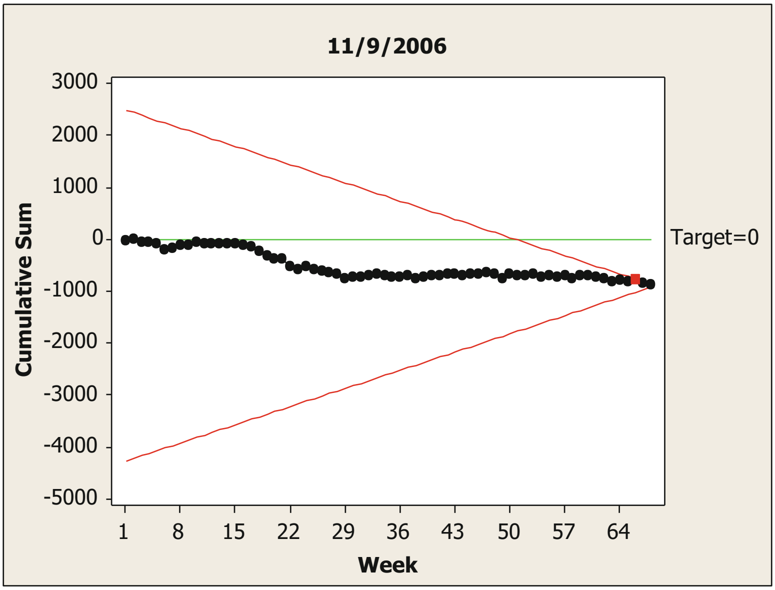


Fig. 1.10 The only out-of-control event in *#PDA* series was in week 69 (November 9, 2006)

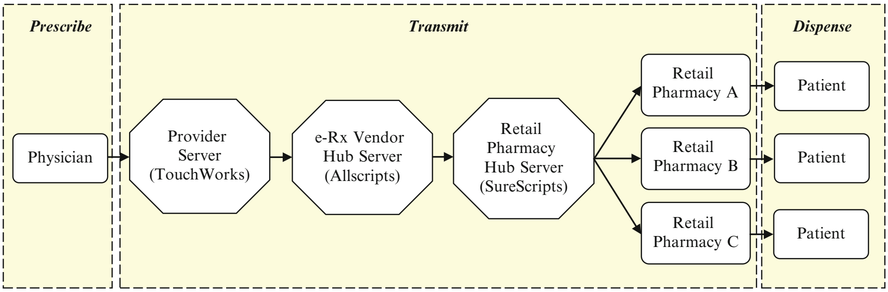


Fig. 1.11 e-Rx flow from medication prescribing to dispensing

server, *e-Rx* vendor hub server, retail pharmacy hub server, and the servers of each individual retail pharmacies before it reaches the patient's hands. Although these transaction data are automatically collected in web logs and audit trail databases, they exist in disparate systems accessible only to the organizations that collect them and are not easily monitored for integrity across the entire system. Mechanisms must be in place to ensure reliability and quality of the service locally. Therefore, if one e-Rx is found to be missing in a pharmacy, it would be very challenging to identify the problematic step in the process in a timely manner. The organizations may not even realize this problem if no patient reports such a failure.

This study demonstrates that any *e-Rx* technology breakdown, even if it is just a partial technology communication failure, can have an overwhelming impact on the extent to which physicians utilize the technology subsequently, if no correcting action is taken to minimize the impact of the failure. As observed in the data, after being alerted by patient complaint calls, the physicians switched to printing prescriptions in the week of October 5, 2006. Since there was no way to confirm where the problem occurred, physicians chose a "safer" way by printing out the prescriptions. Subsequently, a distrust of electronic prescribing started to grow and finally, without receiving any information about corrections, physicians' faith in electronic prescribing collapsed in the week of November 9, 2006, when they stopped using the PDA. And without a correction plan, both of the physician offices completely stopped using electronic prescribing and de-adopted the respective devices. Resulting from the physicians' de-adoption of *e-Rx*, patient safety is potentially jeopardized and pharmacies' revenue likely decreased. Consequently, the lack of appropriate mechanisms to detect and resolve problems with electronic prescribing technology can significantly affect reliability and performance.

Aligned with the conclusion drawn by Miller et al. (2005), we argue that unless the data is shared amongst these organizations and combined with a reporting system to monitor system failures and provide feedback to appropriate stakeholders, from prescribing clinicians to *e-Rx* vendors, pharmacy consortia and information technology support staff, end-to-end quality and reliability of the electronic prescribing process cannot be ensured. Therefore, if the process can be integrated or if each of the organizations can create mechanisms to monitor the process and detect any exception event before it affects medication dispensing, health IT adoption can eventually become sustainable. However, such a substantial change that involves different stakeholders may be very challenging to implement. A suboptimal alternative would be to let the *e-Rx* vendor take the role of the "prescribing process manager" and apply control charts to monitor the key metrics such as *%Send* and *#PDA*. In addition, sufficient documentation of all failure events must be recorded by providers in order to analyze root causes and prevent future failures. Although this strategy cannot detect the transmission error before it affects medication dispensing, it can detect, diagnose, and resolve the error before medication dispensing is overwhelmingly affected adversely and physicians' trust in electronic prescribing is impacted.

This case study has some limitations. The small sample size, limited time frame of the study and a single instance of failure limit both generalizability and our ability to draw causal inferences. Given the lack of detailed records, we can only conclude that it is highly likely that *e-Rx* de-adoption was caused by the transmission failure. Additionally, the lack of documented failure events after September 14, 2006 prevents a detailed validation of the accuracy and consistency of the proposed *CUSUM* scheme. Nevertheless, we believe that a monitoring system that uses the detection and feedback methods described in this paper can help re-build clinicians' faith and boost the adoption rate of electronic prescribing.

1.6 Conclusion

In this study, we investigate the impact of a technology failure on electronic prescribing as a critical barrier to the sustained use of electronic prescribing in two ambulatory care practices. Anecdotal evidence from physicians and clinic staff indicating a failure in prescription transmission to a dominant pharmacy in Fall 2006 is confirmed by a statistical breakpoint estimation method demonstrating a turning point in prescribing behavior in the week of September 14, 2006. The downward trends of both *%Send* and *#PDA* reveal that after the transmission failure, physicians' use of the PDA declined significantly, and even when a PDA was used, they printed the prescriptions for the patient instead of transmitting electronically. Without an error report or feedback system in place, physicians de-adopted the technology entirely, preventing patients, providers, pharmacies, and the practices from taking advantage of the safety, efficiency, and productivity benefits of e-Rx applications.

Therefore, this study suggests that the *e-Rx* vendor employ control charts, such as CUSUM, to monitor and quickly detect abnormal patterns in the use of IT interventions and prescribe action plans to mitigate the effects. With a CUSUM-based approach in use, adjusted for autocorrelation and parameterized for different application settings, it would be possible to detect and react to the unusual downward trends of *#PDA* and *%Send*, and with a feedback loop in place, prevent the breakdowns from impacting operations significantly. Statistical process control mechanisms, such as control charts and control procedure designs, have the potential to help identify system breakdowns quickly, provide the necessary feedback to establish meaningful action plans, and facilitate high technology utilization.

The fragmented nature of the e-Rx transmission process makes it extremely vulnerable to IT failures. These failures have the potential to impact patient safety and health, clinician adoption and sustained use of the technology, and benefits for the pharmacies. As the scope of deployments continues to expand to a large number of medical practice settings, system reliability will be critical. By incorporating a failure detection mechanism such as statistical process control, the electronic prescribing and similar information technology enabled care delivery processes will be more robust and reliable, and failure may be controlled before it causes provider de-adoption and loss of potential financial and clinical benefits of the technology, including patient safety and health outcomes.

Acknowledgements This research was supported in part by a grant from the University of Pittsburgh Medical Center eRecord program. We thank Dr. C. Shalizi of Carnegie Mellon University and Dr. W. Vogt of University of Georgia for valuable suggestions on the analytical models used in this paper and Allscripts® for providing us with the de-identified prescription data analyzed in this study. We are also grateful to the clinicians and staff of the two medical practices for participating in this study through interviews, feedback, and data.

References

Ammenwerth E et al (2008) The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. J Am Med Inform Assoc 15(5):585–600

Bai J, Perron P (1998) Estimating and testing linear models with multiple structural changes. Econometrica 66:47–78

Bai J, Perron P (2003) Computation and analysis of multiple structural change models. J Appl Econ 18(1):1–22

Bell DS et al (2004) A conceptual framework for evaluating outpatient electronic prescribing systems based on their functional capabilities. J Am Med Inform Assoc 11(1):60–70

Cusack CM (2008) Electronic health records and electronic prescribing: promise and pitfalls. Obstet Gynecol Clin North Am 35(1):63–79

Devore JL (2003) Probability and statistics for engineering and the sciences, 6th edn. Duxbury, Pacific Grove

Donyai P et al (2008) The effects of electronic prescribing on the quality of prescribing. Br J Clin Pharmacol 65(2):230–237

eHealth Initiative (2004) Electronic prescribing: toward maximum value and rapid adoption. eHealth Initiative, Washington, DC

Enthoven AC, Tollen LA (2005) Competition in healthcare: it takes systems to pursue quality and efficiency. Health Affairs W5:420–433

Eslami S, Abu-Hanna A, de Keizer NF (2007) Evaluation of outpatient computerized physician medication order entry systems: a systematic review. J Am Med Inform Assoc 14(4):400–406

Etchen LL, Boulton JL (2000) Designing the IDS business portfolio. Healthcare Financial Management 29–33

Fischer MA et al (2008a) Effect of electronic prescribing with formulary decision support on medication use and cost. Arch Intern Med 168(22):2433–2439

Fischer MA et al (2008b) Uptake of electronic prescribing in community-based practices. J Gen Intern Med 23(4):358–363

Friedman MA, Schueth A, Bell DS (2009) Interoperable electronic prescribing in the United States: a progress report. Health Aff 28(2):393–403

Grossman JM, Reed MC (2006) Clinical information technology gaps persist among physicians. Issue Brief Cent Stud Health Syst Change 106:1–4

Grossman JM et al (2007) Physicians' experiences using commercial e-prescribing systems. Health Aff 26(3):w393–w404

Grossman JM et al (2012) Transmitting and processing electronic prescriptions: experiences of physician practices and pharmacies. J Am Med Inform Assoc 19(3):353–359

Hollingworth W et al (2007) The impact of e-prescribing on prescriber and staff time in ambulatory care clinics: a time–motion study. J Am Med Inform Assoc 14(6):722–730

Jani YH et al (2008) Electronic prescribing reduced prescribing errors in a pediatric renal outpatient clinic. J Pediatr 152(2):214–218

McKibbon KA et al (2011) Enabling medication management through health information technology. Agency for Healthcare Research and Quality, Rockville

Miller RA et al (2005) Clinical decision support and electronic prescribing systems: a time for responsible thought and action. J Am Med Inform Assoc 12(4):403–409

Moniz TT et al (2011) Addition of electronic prescription transmission to computerized prescriber order entry: effect on dispensing errors in community pharmacies. Am J Health Syst Pharm 68(2):158–163

Page ES (1961) Cumulative sum charts. Technometrics 3(1):1–9

Ramsey F, Schafer D (2002) The statistical sleuth: a course in methods of data analysis, 2nd edn. Duxbury, Pacific Grove

Surescripts (2011) The national progress report on e-prescribing and interoperable health care: 2010. Surescripts LLC: Minneapolis

Surescripts (2013) The national progress report on e-prescribing and safe-rx rankings year 2012. Surescripts, St Paul

van Doormaal JE et al (2009) The influence that electronic prescribing has on medication errors and preventable adverse drug events: an interrupted time-series study. J Am Med Inform Assoc 16(6):816–825

Walgreens and IMS Health Press Release (2007) New research suggests that, when sent electronically, more new prescriptions make it from doctors office to pharmacy to patient

Chapter 2
Individuals' Attitudes Towards Electronic Health Records: A Privacy Calculus Perspective

Tamara Dinev[1], Valentina Albano[2], Heng Xu[3], Alessandro D'Atri[4], and Paul Hart[5]

[1]Department of Information Technology and Operations Management, College of Business, Florida Atlantic University, Boca Raton, FL 33431, USA

[2]Department of Business Studies, Roma Tre University, Rome, Italy

[3]College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802, USA

[4] Late Director of the Research Center on Information Systems (CeRSI), LUISS University of Rome, Rome, Italy

[5]Department of Information Technology and Operations Management, College of Business, Florida Atlantic University, Boca Raton, Florida, USA

Abstract National adoption of Electronic Health Records (EHRs) is considered an essential component of the health care system overhaul sought by policy makers and health care professionals, in both U.S. and Europe, to cut costs and increase benefits. And yet, along with the technological aspects, the human factor consistently proves to be a critical component to diffusion of any IT system, and is even more so regarding health care. The highly personal and sensitive nature of health care data and the associated concerns about privacy impede even the most efficient and technologically perfect system. Our objective is to investigate individuals' attitudes towards EHR and what factors form these attitudes. If we understand individuals' attitudes regarding EHR and the factors that influence them, we will be in a better position to take responsive measure to facilitate Privacy by Design for EHRs. A positivist research model is empirically tested using survey data from U.S. and Italy and structural equation modeling techniques. We find that perceived effectiveness of regulatory mechanisms positively impact trust; perceived effectiveness of technological mechanisms positively impacts perceived privacy

T. Dinev, Ph.D. (✉)
Department of Information Technology and Operations Management, College of Business, Florida Atlantic University, Boca Raton, FL 33431, USA
e-mail: tdinev@fau.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*, Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_2

control and trust; the latter two help reduce privacy concerns which, along with perceived benefits, convenience, and Internet experience, play the privacy calculus-type formation of attitudes towards EHR.

Keywords Health care • Electronic health care records • Privacy • Attitudes • Structural equation modeling

2.1 Introduction

The health care sector has a significant societal impact, as it affects individuals' quality of life as few other sectors do. Concerns about the quality of health care and economic sustainability have existed for years in all developed economies. Government agencies and business executives who are involved in providing health coverage for workers and citizens have long called for cost control. In United States, a recently published report from the Business Roundtable, which represents CEOs of major companies, has concluded that US health care system has become a liability that hinders companies' competitiveness in a global economy (Alonso-Zaldivar 2009). The unsustainable costs demand system overhaul. As additional twist, the report found that higher U.S. spending fails to deliver a healthier work force, thus creating the largest "value gap" between cost and benefits among the developed economies' health care systems.

While the extent of the value gap is less for the European countries, they also struggle with increased health care costs (Pagliari et al. 2007). Moffit et al. (2001) found that a major factor pushing recent health care policy changes in Europe, including privatization efforts, is the rapid aging of the European population and the stagnated demographics. The unfavorable shift in the demographic balance imposes tremendous financial pressures on the health care and neutralizes the overall higher efficiency of the European nations' health care systems as compared to the Unites States', thus introducing noticeable urgency for reforms (Moffit et al. 2001).

There is broad agreement that accelerated Information Technology (IT), and specifically Electronic Health Records (EHR) adoption will be critical to close the value gap and reduce the health care costs in general (Landro 2004). Per the definition introduced by Angst and Agarwal (2006a), EHR is a software system that health care providers use to *create, store, update and/or share patient information in electronic format* (see also HHS 2006). Health care is an information-intensive industry, in that a large percentage of its activities are enabled by the storage, processing, transfer, and analysis of data. Quick access to a patient's medical record, often integrated from various sources, can reduce medical errors help diagnosis, and facilitate the communication with related agencies and businesses. Electronic forms and data management, electronic prescription filling, and electronic managed care all increase health care quality and safety, cut costs, and improve efficiency and precision of diagnosis and operations. Thus, digitizing patient records is an essential part of the IT overhaul of the system seeking to improve the care and curbing costs.

Recognizing the importance of the information access and sharing in health care and the slow rate of IT adoption in this sector (Angst and Agarwal 2006a), governments, policy makers, advocacy groups, and individuals have invested extensive efforts to induce more rapid digitization and sharing of medical data both in Europe and North America. In United States, the recently adopted stimulus package dedicates $50 billion over 5 years to spur the adoption of EHR. In November 2005, the U.S. Senate passed with unanimous consent the Wired for Health Care Quality Act (S. 1418), a bill to enhance the adoption of a nationwide health information technology and to improve the quality and reduce the costs of health care in the United States. In Britain, The NHS (National Health Service), through Connecting for Health, is introducing two types of online health record for everybody in England—the summary care record and HealthSpace (Kidd 2008) that are set to become the world's first fully national system, according to Pagliari et al. (2007). In Italy, a new model of cooperative governance targeting to define national roadmaps coherently with the European guidelines was embraced. The Department of Technological Innovation and the Ministry of Health established an eHealth board called TSE (Tavolo di Sanità Elettronica), charged with coordinated implementation of interoperable IT infrastructures and applications. Since 2005, the TSE has developed a general and comprehensive eHealth conceptual framework and architectural guidelines for a software infrastructure supporting distributed health care processes (TSE 2005, 2006). Coherently with this new approach many projects have been developed fostering the implementation and adoption of new TSE compliant ICT solutions (e.g. the General Practitioners' Network Pilot Program (Rete di Medici di Medicina Generale—RMMG) as well as the interoperability at both national and European level.

Although few question the potential benefits of digitized health care information and it would seem it is ripe and technologically ready for widespread, global adoption, a realistic assessment of the current state of affairs is more sobering. As with every technology, beyond the technical challenges looms the people factor—the individuals' attitudes and readiness to adopt EHR.

There is a significant level of complexity when adoption of EHRs is concerned. On one level, EHR adoption takes place at institutional level: the healthcare provider such as doctors' or dentist's office, a hospital or insurance company adopting a digitized medical record system. It may seem that once the adoption is complete at that level, the patient would not have a choice and would need to automatically comply by letting the provider enter all the necessary medical and personal information in the computer system. However, while patients and individuals cannot directly influence the institutional adoption of a technology such as EHR by healthcare providers, they may resist digitization of their own data until confident that security and privacy policies and practices are in place. As in any case of forcing a technology into individuals, the risk of unintended consequences looms. The importance of the patients' attitudes towards EHRs takes precedence when their weight as political and social capital is considered as an undisputable factor for the success of the nation's transition to electronic healthcare data. Even if providers successfully adopt EHRs, patients can demand to opt-out of digitized systems, a provision present in all national policies. If controversies and negative attitudes escalate, political pressures may amount to levels that will render any well intentioned policies and institutional and national efforts unachievable. Thus, it is important to understand that patients' cooperation and willingness to allow their medical information in digital form is crucial to the success of EHR (Angst and Agarwal 2009; Bansal et al. 2010).

So far, citizens' attitudes and intentions towards electronic medical systems are not well understood in any country. Angst and Agarwal (2009) point that there is limited knowledge of how patients will be involved in delivery, monitoring, and the dissemination of the information related to their healthcare. However, all the countries that are moving towards the EHRs systems introduce an option for patients to opt-out from having their medical records in electronic form, including Italy and United States. As Angst and Agarwal (2009) point out, patients may demand that their records remain non-digitized (Kauffman 2006) and thus they have a central role in the EHRs diffusion process. The resistance of patients and society in general, the unaddressed and escalating controversies regarding the use of such a sensitive personal information may reach such levels "as to render any national efforts unachievable" (Angst and Agarwal 2009, p. 360). Our study focuses on individual patient's or consumer's attitudes towards EHR as opposed to institutional adoption attitudes by healthcare providers.

As evidenced by many studies, among the major impediments for wide adoption of EHR by patients and consumers are privacy concerns (Angst and Agarwal 2006a, 2009; Bansal et al. 2007; Bodenheimer and Grumbach 2003; Cantor 2001; Earp and Payton 2006; Harris 2002; Masys et al. 2002; Shortliffe 1999; Westin 2003). For example, Bansal et al. (2007) argue that individuals' trust, privacy concern, and information sensitivity are factors in the success of electronic delivery of health services.

Thus, parallel to the efforts for digitization of the health care sector, there is a sense of urgency associated with efforts to ensure strong patient privacy and personal data protection. By focusing on electronic transactions, the privacy regulation required by HIPAA (Health Insurance Portability and Accountability Act of 1996) attempted to assure consumers that as their health records become fully electronic and networked, the information would be protected. Initiatives in the health care sector have led to privacy guidelines and standards for health websites (eHealth) and Health Information Technology (HIT) in general (Choy and Goldman 2001). HIPAA explicitly does not overrule state level privacy rules that often provide stronger protection. It serves instead as a minimum national standard. The core principles and practices of these voluntary efforts are in accordance with the U.S. Federal Trade Commission's (FTC) Code of Fair Information Practices. In Italy, TSE cooperated with the Italian Privacy Authority to the enactment of a set of Privacy Guide Lines on the implementation of an EHR system (Gazzetta 2009), in order to ensure the citizens' control of their personal data. In general, however, self-regulatory efforts differ substantially in focus and comprehension across nations, health care organizations, and providers (Choy and Goldman 2001; Earp and Payton 2006).

In spite of these efforts, many consumers are reluctant to take advantage of the potential benefits associated with electronic health care data due to privacy concerns. The consumers are concerned that misuse of their health information may result in undesirable social or financial consequences (Luck et al. 2006). A recent national study by the California Health care Foundation found that 67 % of the respondents were "somewhat" or "very concerned" about the privacy of their medical records (Bishop et al. 2005). The study also found that recent reports in the media of privacy breaches have raised the level of the concerns, and although consumers are willing to share personal medical information in exchange for better coordination of their medical treatment, privacy protection behavior persists. One in six Americans engages in some form of privacy protective behavior to prevent themselves from a harmful or intrusive use of health information. These privacy protective responses may include: falsify or withhold information; pay out-of-pocket for care; or see multiple providers to avoid a consolidated record; (Rindfleish 1997). Such 'privacy-protective' behavior can compromise both individual care and public health initiatives (Lo et al. 1999, p. 3).

It is obvious, then, that the relationships between individuals' perceptions and behaviors; health care organizations' policies and practices; their sectors' guidelines; national regulatory frameworks; and global factors form a complex framework. Individuals—whose acceptance of, and cooperation with, a digitized health care system is critical—form their perceptions from within this complex framework. A fully functional health care IT environment, such as EHR, would lead to individuals' acceptance only if the individuals first form an overall positive attitude towards that environment. The attitude formation is in its part governed by cost-benefit type of analysis an individual makes about EHR. That is, a positive attitude is formed when positive perceptions outweigh the concerns and risks the individual associates with that environment. Thus, a full understanding of the privacy dynamics regarding the digitization of the health care industry can only be attained by looking across various factors, both positive and negative, that affect the individual's attitudes towards EHR.

Drawing on attitude and attitude persuasion literatures, Angst and Agarwal (2009) showed that individuals *can* be persuaded to change their attitudes towards EHRs and opt-in behavioral intentions, even in the presence of significant privacy concerns, if appropriate messages about the value of EHR are imparted to the recipient, with argument framing and issue involvement. Angst and Agarwal (2009)'s findings are in accordance with previous research on privacy calculus (Culnan and Armstrong 1999; Dinev and Hart 2006; Laufer and Wolfe 1977) that treated privacy concerns not as acting in isolation, but rather as part of an individual's cost-benefit analysis in economic, and social contracts. The privacy calculus stream of research showed that high privacy concerns *can* be balanced off by strong positive perceptions about a certain activity that individuals perceive as beneficial (e.g., online shopping, receiving quality health care, etc.). Coming from the rich perspective of privacy calculus, our goal is to identify the factors that, along with privacy concerns, play role in the formation of the overall positive attitude toward EHR. Thus, our research question is: *What are the factors that drive individuals' attitudes towards EHR*?

While the overarching framework in our study is privacy calculus, we will also be guided by the psychological control theories and trust theories which will help us in the conceptualizations of privacy concerns in the health care context and identify their antecedents and consequences. We argue that perceived privacy control over EHR and consumer trust toward health care providers are two key factors that determine individuals' privacy concerns. We will evaluate the effectiveness of regulatory and technological mechanisms of privacy enforcement on enhancing the perceived privacy control and building consumer trust. We will further investigate how individuals weigh the costs and benefits of potentially compromising some degree of privacy for the possibility of getting better health benefits.

The rest of the paper is organized as follows. We first present the theoretical foundations and develop the conceptual model. Next we discuss the methodology used in this study for the empirical testing of the model. The results and findings are summarized in a discussion section, followed by limitations and suggestions for future research, implications for theory and practice, and conclusion.

2.2 Theoretical Foundations and Research Hypotheses

The proposed research model shown in Fig. 2.1 is anchored in multiple theories. The dependent variable of interest is attitudes toward EHR, and privacy concern is posited as the major factor negatively impacting the attitudes. Attitude toward EHR is determined through an assessment of one's beliefs regarding the consequences arising from use of EHR and an evaluation of the desirability of these consequences (Ajzen and Fishbein 1980). We define the attitude toward digitization of EHR as the individual's positive or negative feelings about providing EHR. We further define *information* privacy concerns (hereafter referred to as privacy concerns) as beliefs reflecting the extent to which individuals are disturbed about the information collection practices of others and how the acquired information will be used (Culnan 1993; Smith et al. 1996; Stewart and Segars 2002).

The overarching theoretical framework for this model is privacy calculus. As a concept, privacy calculus was first considered in the seminal paper of Laufer and Wolfe (1977) and further elaborated by Culnan and Armstrong (1999). Dinev and Hart (2004, 2006) and Xu et al. (2009) developed a quantitative and empirically testable model based on privacy calculus framework. They empirically explored the simultaneous effect of positive and negative personal beliefs, including privacy concerns, all associated with inhibiting or driving behavior or behavioral intention. The set of inhibitors was shown to hinder decisions to get involved in e-commerce transactions. The set of drivers, such as trust and perceived privacy control, was shown to positively influence e-commerce transactions. The important concept in this model is the *cumulative* influence of the inhibitors and drivers, forming the so-called 'privacy calculus' (Culnan and Armstrong 1999)—a mental calculation as to

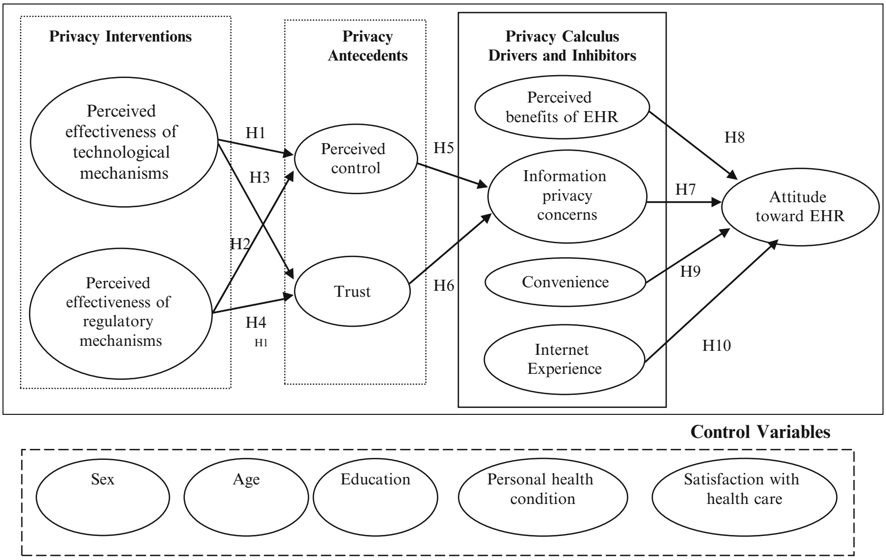


Fig. 2.1 Research model

which beliefs are strong enough to override the contradictory ones. Each set can outweigh the other, determining the user's final decision on whether to perform a certain behavior or not. If the cumulative effect of the drivers is higher than the cumulative effect of the inhibitors, the user will more likely make a decision to perform the behavior. Otherwise, the behavior is less probable (Chellappa and Sin 2005; Xu et al. 2009).

Dinev et al. (2008) adapted the privacy calculus framework through the lens of the Theory of Planned Behavior (Ajzen 1991) and showed strong and reliable connection between the driving and inhibiting factors and attitudes which in turn are strongly associated with behavior. Thus they argued that attitude formation is also influenced simultaneously by two sets of contrary beliefs. For each individual, these beliefs are weighed and the strength of one may over-ride the influence of another. The relative strength of the inhibitors and the drivers provides insight into a complex process that leads to attitude formation and eventual behavior intention. Higher levels of negative beliefs would suggest individual's predominant negative attitude that would to resistance to adopt EHR, and vice versa. The fact that our model includes contrary antecedents indicates that their relative influence needs to be taken into full consideration when attempting to understand the resulting attitudes.

To explore the antecedents to privacy concerns relevant in the context of health care, we draw on the theories of psychological control and trust. We hypothesize that the perceived effectiveness of privacy interventions such as the regulatory and technological mechanisms will affect an individual's perceived privacy control toward EHR and trust toward health care providers.

2.2.1 Privacy Interventions and Perceived Privacy Control Over EHR

The two basic types of privacy interventions in the context of health care and digital medical records are regulation and technology. In this section we posit that the individuals' perceptions of their effectiveness will lead to higher perceived privacy control. Perceived privacy control has been defined as a psychological construct reflecting the belief regarding the extent to which an agent can produce desired outcomes (Skinner et al. 1988, p. 11). One very important psychological perspective views privacy as grounded in the *control* of personal information. For example, Stone et al. (1983) viewed privacy as the ability of the individual to control personally information about one's self. This control perspective of privacy is also found in prior privacy studies that posited that loss of control over information is central to the notion of privacy (Dinev and Hart 2004; Phelps et al. 2000; Sheehan and Hoy 2000). Following this perspective, control (interpreted as psychological control) is identified as one major factor that is closely related to consumers' privacy concerns (Xu 2007; Xu and Teo 2004; Xu et al. 2012). According to Yamaguchi (2001, p. 226), people should feel greater autonomy when they exercise direct personal control, that is, they act themselves as the control agent.

In the health care context, due the heightened concern and perceived risks associated with EHR, privacy interventions—institutional technological and regulatory mechanisms to ensure privacy—have been heavily embarked upon in the design of EHR systems. EHRs are records which may not necessarily be available on the World Wide Web (only providers to access them) but once medical records are digitized, they are practically online. Automatic email generation among providers and to patients will mean that health records are sitting in email servers, routers etc. They can be hacked, used surreptitiously against the individuals and be even more available on the Internet. These risks are very much in people's mind when they think about digital health records and therefore they would want to assert control through control mechanisms. The latter may include ability of individuals to access and modify their own health records or to block others from accessing (and modifying) this information.

A number of technological mechanisms which enhance security of the data and prevent unauthorized use or accidental disclosure (Rindfleish 1997) have been implemented in the recent years. For example, Cimino et al. (2002) proposed to adopt two-factor authentication to enhance user control to their own EHR. In general, the technological trend in the recent years has been to develop tools for the users that will help them enhance their privacy in the EHR. These tools go beyond what is required by policy and security assurances standards the extent of which the users may not be aware of. But small things go long way in that respect. For example, even if policies and mandatory practices do not require it, EHRs may provide the option for the individuals to control what information he or she will make available to health insurance companies. Do private insurance companies who determine the health insurance premium and out of pocket expenses need to know what contraceptive a woman uses? Or, that 20 years ago someone broke an arm? This is becoming particularly relevant issue given the private insurance exchange markets given in the U.S. health insurance bill. If the users perceive that they have well implemented and effective technological privacy enhancing mechanisms they will feel empowered to exert direct control over their EHR. Hence, they would also perceive they have more control over their EHR. For the purpose of this study, we define perceived effectiveness of privacy enhancing technological mechanisms (for parsimony, also referred to in this paper as perceived effectiveness of technological mechanisms) as the extent to which individuals believe that the privacy protective technologies implemented and used in EHRSs are able to provide effective and reliable protection against privacy breaches on their medical data. Therefore, we hypothesize that:

Hypothesis 1: *The perceived effectiveness of privacy enhancing technological mechanisms will lead to higher perceived privacy control over personal EHR.*

When exercise of direct personal privacy control is not readily available or sufficient, people might relinquish their direct privacy control and seek "security in proxy control" (Bandura 1982, p. 142). *Proxy control* is an attempt to gain control through powerful or skillful others when people do not have enough skills, resources, and power to bring about their desired outcome or avoid an undesired outcome (Yamaguchi 2001; Xu et al. 2012). In the health care context, when users perceive that they lack the requisite resources to directly control their EHR, they may reshape their decisions by considering the availability of powerful others (e.g., regulators and legislators) who can act on their privacy preferences. Prior sociology and legal literature lend strong support to the effectiveness of regulatory mechanisms on individuals' control over their personal information (Bandura 1986; Faden et al. 1986). In the health care context, the regulatory mechanism such as HIPAA or country and state-specific laws should be powerful in terms of the exercise of social control (Spiro and Houghteling 1981) since it requires that offenders be punished in order to maintain the deterrent effectiveness of the legal system (Tittle 1980). For the purpose of this study, we define perceived effectiveness of privacy enhancing regulatory mechanisms (for parsimony, also referred to in this paper as perceived effectiveness of regulatory mechanisms) as the extent to which individuals believe that the privacy regulations regarding EHR are able to provide effective and reliable protection against privacy breaches on their medical data.

Hence, viewing the deterrent effectiveness of regulation, individuals would believe that the legal assurance of their privacy control should enable them to exercise proxy control over their EHR. Therefore, we hypothesize that:

Hypothesis 2: *The individual's perceived effectiveness of privacy enhancing regulatory mechanisms will lead to higher perceived privacy control over personal EHR.*

2.2.2 Privacy Interventions and Trust in EHR

Trust is a crucial enabling factor in relations where there is uncertainty, interdependence, risk, and fear of opportunism (Dan et al. 2008; Mayer et al. 1995). Trust has been defined as the willingness to depend on another person or institution based on the belief in the integrity, ability, and benevolence of this other (Mayer et al. 1995; McKnight et al. 2002). In the context of health care, because of security breaches and potential unauthorized or accidental disclosure of sensitive personal medical information (Dixon 2005), there is a heightened risk perception, and the latter is a salient antecedent to any innovative technology acceptance (Xin et al. 2010). Trust comes to be a critical element in helping individuals overcome their perceptions of uncertainty and risk especially in the context of digitization of EHR (Bansal et al. 2010). Similar to Internet and e-commerce trust beliefs, there are several facets of trust in the context of EHR—trust in the healthcare institutions (Doctors' offices, drug companies, insurance companies etc.) that will use EHR (called in the MIS research Institutional trust (McKnight et al. 2002)) and trust in the dependability and reliability of the EHR Information system itself (McKnight et al. 2002). An individual who trusts a health care provider may not necessarily trust EHR systems used by the health care provider. Alternatively, and individual may believe that an EHR information system is well built but the institutions that use and access it may have intentions that are against the interest of the individual. While the complex landscape of the trusting beliefs warrants a separate research and model that will incorporate all the facets of trust, our research is interested in the trust in the EHR systems rather than the trust in the healthcare providers. We also recognize that the two are related—if the healthcare provider effectively implements the privacy enhancing technological mechanisms such as enhanced access control (Cimino et al. 2002), this should directly build individuals' trust beliefs (Hu et al. 2010; Xu et al. 2005) toward a health care provider using EHR. Indeed, a substantial and nontrivial investment of time and resources are required to design and implement the technological privacy-enhancing mechanisms that go beyond the state laws requirements and mandatory protection policies and practices. Such action should be interpreted as a signal that the health care provider is proactively addressing users' privacy concerns and that it will undertake the responsibility to manage users' personal information properly. In other words, effective implementation and use of privacy enhancing technological mechanisms will create user trusting beliefs in the EHR systems used by trusted healthcare providers. Therefore, we hypothesize that:

Hypothesis 3: *The perceived effectiveness of privacy enhancing technological mechanisms will lead to higher users' trust in EHR.*

A health care provider's compliance to and enforcement of the privacy regulatory mechanism should directly build consumer's trust toward the particular health care provider for two reasons. First, the privacy regulatory mechanism could limit the EHR providers' and users' ability to behave in negative or opportunistic ways, allowing patients to form and hold beliefs about expectations of positive outcomes (Johnson and Cullen 2002). Second, when violations do occur, the privacy regulatory mechanism could provide mechanisms of voice and recourse for the betrayed (Johnson and Cullen 2002), which could create strong incentives for health care providers to refrain from opportunistic behavior and behave appropriately and invest the time and resources to build sound, reliable EHR systems. Therefore, we hypothesize that:

Hypothesis 4: *The perceived effectiveness of privacy enhancing regulatory mechanisms will lead to higher users' trust in EHR.*

2.2.3 Perceived Privacy Control and Trust as Antecedents to Information Privacy Concerns

Overwhelming empirical evidence and theoretical arguments reveal that control is one of the key factors that provide the greatest degree of explanation for privacy concerns (Phelps et al. 2000; Sheehan and Hoy 2000; Xu 2007; Xu and Teo 2004). In the e-commerce context, for example, it was shown that individuals have fewer privacy concerns when they have a greater sense that they control the disclosure and subsequent use of their information (Culnan 1993; Culnan and Armstrong 1999; Milne and Boza 1999; Stone and Stone 1990). In this study, we expect a similar negative relationship between perceived privacy control and privacy concerns in the health care context. Therefore, we hypothesize that:

Hypothesis 5: *Perceived privacy control over personal EHR will have a negative effect on privacy concerns.*

The literature on information privacy in the e-commerce context suggests that trust could play an important role in alleviating consumers' privacy concerns (e.g., Caudill and Murphy 2000; Culnan and Bies 2003). Various studies revealed the strong relationship between the two constructs although with various causality direction (Bansal et al. 2007; Dinev et al. 2006). In health care, privacy concerns are a major factor responsible for negative attitudes and resistance to any new way that may risk disclosing personal data. The characteristics of EHR are such that there is an expected increase in the likelihood of privacy violations and misuse of information (Angst and Agarwal 2009). Additionally, the highly sensitive nature of personal medical data adds even more to the uneasiness individuals feel about the violations and misuse. These are general privacy concerns that do not relate to particular systems and practices. People do not want their medical data to land in wrong hands and used against them in any possible way. Whether there will be security breaches, hacking, or insiders snooping into the personal records for financial, political, or legal gains (for example giving them to the media); whether there will be surreptitious gathering of information by private detectives, insurance companies, or the government to be used in unimaginable and unforeseen for the individuals way, people are uncomfortable with the possible implications of all their records available in one place (rather than folders in various Doctors' offices and hospitals) and used against them. These concerns can never go away completely since they are general and since there so many ways that information can leak. But what can help alleviate them is the trust (a wholesome trust—in institutions and technology) that a particular system, well thought out, well designed and implemented will bring the possibility of such incidents to a minimum. An example of such particular system is EHR.

Based on Hypotheses 3 and 4, we argue that the formation of trust in EHR, by means of the above discussed privacy interventions, will alleviate to a certain extent privacy concerns. If individuals trust their doctors and hospitals in their professionalism and best intentions, they will also trust that the doctors will want reliable, dependable, good EHR systems that will be free of errors, will not be breached, will not be used against the individuals. Therefore, individuals may lower their privacy concerns. We thus view, for the context of EHR in particular, that trust is a necessary precursor that will set the stage for an individual to lower his or her privacy concerns regarding EHR. Without the individual having built trust in the system in the first place, there is no rational for him or her to lower his or her privacy concerns. Therefore, we hypothesize that:

Hypothesis 6: *Trust in EHR will have a negative effect on privacy concerns.*

2.2.4 Drivers and Inhibitors in Privacy Calculus

Information Privacy Concerns

We defined information privacy concerns in the beginning of the theoretical section of this study. As revealed by the cited in the Introduction surveys, the lengths to which consumers will go to hide health information because of privacy concerns are surprisingly great (Bishop et al. 2005). Thus, Angst and Agarwal (2009) argued that the EHR offers a different context than other information technologies with respect to privacy concerns. Following Angst and Agarwal's (2009) arguments, we posit that the multidimensional view of privacy concerns as developed by the Smith et al.'s (1996) concern for information privacy (CFIP) construct is the most appropriate treatment of privacy concerns when digitization of personal medical records and history is concerned. CFIP as composed of four distinct, yet correlated latent factors—(privacy concerns regarding) collection, errors, unauthorized access, and secondary use of information. Stewart and Segars (2002) showed that a second order factor structure is empirically valid and can be used in subsequent empirical models of privacy.

Attitude toward EHR is determined through an assessment of one's beliefs regarding the consequences arising from use of EHR and an evaluation of the desirability of these consequences (Ajzen and Fishbein 1980). Privacy concerns, viewed as a negative antecedent belief, could negatively affect a person's attitude toward EHR. Chellappa and Sin (2005) have similarly argued that when people have stronger concerns about information privacy, their attitudes about using a technology will be more negative, as also confirmed recently by Angst and Agarwal (2009) in the context of EHR. Hence, we hypothesize that:

Hypothesis 7: Individual's *privacy concern will have a negative effect on attitude toward EHR.*

Perceived Benefits of EHR

Perceived benefits are an important component of cost-benefit analysis. IS researchers have utilized perceived benefits as an independent variable that affects adoption of IT innovations (Beatty et al. 2001; Forsythe et al. 2006; MacKay et al. 2004; Teo and Yeong 2003). The literature on perceived benefits points to the important role of this construct in rational decision making. In the context of the research model, perceived benefits are defined as the expected relative advantage associated with using EHR. Analogous to perceived usefulness in the Technology Acceptance Model framework, the perceived benefits are an important antecedent to attitudes (Goodwin 1991; Milne and Gordon 1993). Hence, in the health care context, individuals are more likely to have a positive attitude toward digitization of health care data when they perceive a significant level of benefits in having EHR. Indeed, Song and Zahedi (2007) found that users' beliefs about the ability and benevolence of the health infomediary critically affect their behavior intentions. Therefore:

Hypothesis 8: *Perceived benefits of EHR will have a positive effect on attitudes toward EHR.*

Convenience

Convenience has been considered in e-commerce transactions as a utilitarian motivator that positively affects attitudes and behaviors of consumers (Childers et al. 2001). Convenience is defined as the individual's perception of presence of a set of opportunities that facilitate the accomplishment of a task, and also make the process of accomplishing the task more appealing. The convenience in interactive use of medical records increases search efficiency through the ability to do that at home (Angst and Agarwal 2006a, b), by eliminating such frustrations as fighting traffic and looking for a parking space, and avoiding long lines at medical offices for simple documentation handling, scheduling appointments, or gathering recent test results from various offices. Convenience is manifested in the single "stop" for all your medical records (Angst and Agarwal 2006b) that eliminates travel to and from a variety of offices and long telephone tags with the doctors, clinics, and insurance companies. Thus, convenience includes both the elements of when and where a patient can operate with his or her medical records. Indeed, speed, efficiency, ease of finding the information or performing tasks 24/7 increase substantially the advantage and uniqueness, as was shown in the case of online shopping (Childers et al. 2001). We should expect the same relationship to be valid for EHR as well, as also found by Angst and Agarwal (2006b). Indeed, the convenience of the medical clinical data to be available electronically: the patients and their doctors and pharmacists to pull them fast and error-free when needed, without the risk of loss or misplacement, are all pointing to the high utilitarian value of EHR. Therefore:

Hypothesis 9: *Convenience will have a positive effect on attitudes toward EHR.*

Internet Experience

Significant extant research has shown that technology experience can affect attitudes towards the various technologies (e.g. Taylor and Todd 1995; Venkatesh et al. 2003). Experience and continued usage of technology informs the individual's expectations about its capabilities, benefits, and drawbacks, as well as his or her familiarity with that technology. Through greater experience and familiarity, individuals form more positive reactions to new technologies (Bansal et al. 2007; McKnight et al. 2002). Relating these findings to the context of EHR which is an online technology, we posit that Internet experience and frequent internet shopping will help to build positive attitude formation towards EHR. Therefore:

Hypothesis 10: *Internet experience will have a positive effect on attitudes toward EHR.*

Control Variables

Based on prior research on adoption and consumer behavior, a number of additional factors may influence privacy concern, trust, and attitudes toward EHR. Because there is no sufficient theoretical argument that we can draw on to include them in our model as antecedents, we include them as control variables, to eliminate the variance explained by them. They are *sex, age, education, personal health condition, and satisfaction with health care*.

2.3 Research Methodology

2.3.1 Construct Operationalization and Scale Development

Most of the measurement items of the survey instrument were adapted from extant instruments in the literature (Table 2.1) and a pilot study was conducted to test for clarity, consistency, and validity with 27 undergraduate and graduate students in a Southern U.S. university acquainted with EHR. Following standard practices, scales were purified and refined. In general, however, the pilot test resulted in only minor changes to the initial instrument. Several items were dropped for parsimony of the model (Pavlou and Fygenson 2006). The instrument is given in the Appendix.

2.3.2 Data Collection

Since the issues discussed in this study are truly global, the researchers sought to eliminate U.S. biased treatment of the theory and the findings, with this enhancing the generalizability of their model by collecting data from U.S. and Europe. With this we believe that the implications of the model will inform in a more convincing way the policy makers, health care executives and providers and transcend national specifics of the health care model. The researchers chose Italy as a typical European style of health care provider—by and large, the Italian government provides universal health care and is committed to implementing electronic health care records in the nearest future. Since the health care model in most European countries

Table 2.1 Sources of construct operationalization

Construct	Source of measurement items
Attitude toward EHR	Pavlou and Fygenson (2006), Taylor and Todd (1995)
Convenience	Angst and Agarwal (2006a)
Perceived privacy control over HER	Xu (2007)
Privacy concerns—collection, errors, unauthorized access, secondary use	Angst and Agarwal (2009)
Perceived benefits	Iacovou et al. (1995), Liu (2007)
Perceived effectiveness of regulatory mechanisms	Pavlou and Gefen (2004), Xu et al. (2008)
Perceived effectiveness of technological mechanisms	Xu et al. (2008)
Trust	Jarvenpaa et al. (2000); McKnight et al. (2002); Pavlou and Fygenson (2006)
Health status	Bansal et al. (2007)
Experience with internet	Angst and Agarwal (2006b), Bansal et al. (2007)

is substantially different from the one in US, we introduced a control variable, satisfaction with current health care, to capture the citizens' attitudes towards their current health care system, and to eliminate possible influences if differences exist. To ensure that the subjects understand the use of the term and the nature of EHRs, we provided detailed description of the systems and how they can be used by the doctors and by the patients. The questionnaire was translated from English to Italian by native Italian speakers and then back to English following a generally accepted practice to ensure consistency in cross-lingual surveys (Karahanna et al. 2002). It was pretested with multiple respondents from Italy and U.S. with diverse age, gender, and education. While no major problems were identified, instrument was further refined and a few modifications in the Italian translation were made.

The goal in the data collections from both countries was to reach as diverse sample as possible that would closely follow the representation of the demographic categories of the general population. For that purpose, individuals were approached in various settings, including hospital and doctors' waiting rooms, neighborhoods, small business and public meeting places such as parks and transport stations. The individuals were asked to participate completely voluntary and if they wished, a preaddressed and prepaid envelope was given to them, so they can fill in and return the survey by mail at a later time. After eliminating several responses due to multiple empty entries, we used total of 217 responses from United States and 188 from Italy. The demographic distribution (Table 2.2) shows that both Italian and U.S. samples are diverse, comprising a wide range of age, education, with equal representation of genders.

2.3.3 Model Testing

The model developed in this study (Fig. 2.1) was tested by Structural Equation Modeling's (SEM) Partial Least Square (PLS) Method using SmartPLS 2.0 (Ringle et al. 2005). A two-stage approach, as recommended by Gefen et al. (2000), was used to first assess the quality of the measures through the confirmatory factor analysis (CFA) stage and then test the hypotheses through the structural model, the SEM stage. The CFA was performed on the entire set of items simultaneously with each observed variable restricted to load on its a priori factor. Validation and reliability assessment of the measurement model were conducted following the widely used SEM heuristics Gefen et al. (2000).

Measurement Model

The measurement model was assessed for the following three criteria: (1) item reliability and convergent validity, (2) internal consistency, and (3) discriminant validity. The factors loadings of the measurement items are presented in Table 2.3. To assure measurement adequacy for both nations, we ran the measurement model separately

Table 2.2 Demographic characteristics of the Italian and U.S. samples

	Sex		Age			Education					Knowledge of EHR	Health status	Satisfaction with current health care	Internet experience		
											Mean (Std Dev)	Mean (Std Dev)	Mean (Std Dev)	Mean (Std Dev)		
	Male	Female	<30 years	30–60 years	>60 years	High school or less	GED or equivalent	Some college degree	Bachelors degree	Graduate degree				exp1	exp2	exp3
Italy	98	90	64	86	38	13	71	34	53	17	1.74 (.99)	3.41 (.88)	2.85 (1.13)	1.81 (1.18)	2.32 (1.46)	2.86 (1.46)
U.S.	119	98	82	78	41	59	47	45	57	9	1.71 (1.08)	3.17 (1.11)	2.99 (1.43)	2.97 (1.43)	4.00 (1.23)	4.10 (1.28)

Notes: Knowledge of EHR: How much are you informed about the EHR (Not at all—Very well)

Table 2.3 Item loadings and cross-loadings, Italy and U.S.

	Italy										U.S.									
	ATT	CO	CTRL	LIKE	PCON	PERBEN	REGUL	TECHN	EXP	TRUST	ATT	CO	CTRL	LIKE	PCON	PERBEN	REGUL	TECHN	EXP	TRUST
ATT1	**.92**	.04	−.01	.13	−.27	.28	.22	.14	−.02	.15	**.89**	.28	−.02	.29	−.13	.21	.04	.15	.36	.25
ATT2	**.93**	.08	.07	.09	−.32	.27	.25	.18	−.07	.23	**.89**	.20	.10	.15	−.27	.28	.10	.28	.22	.33
ATT3	**.89**	.01	.04	.10	−.29	.29	.16	.22	−.04	.22	**.89**	.24	.04	.22	−.15	.31	.14	.30	.27	.28
CO1	.04	**.89**	.08	−.03	.03	−.09	−.04	.11	−.01	−.10	.29	**.96**	.02	.22	−.10	.10	.05	.12	.21	.14
CO2	.05	**.93**	.11	.04	−.02	−.05	.01	.11	.00	−.06	.18	**.88**	.00	.13	−.09	.04	−.03	.07	.17	.12
CTRL1	.01	.07	**.76**	−.11	−.23	−.03	.19	.23	.00	.17	.01	.00	**.82**	−.07	−.26	.11	.22	.33	−.11	.29
CTRL2	.07	.15	**.82**	−.08	−.16	.03	.08	.27	−.07	.11	.05	−.04	**.85**	−.07	−.21	.18	.25	.36	−.09	.22
CTRL3	.09	.04	**.87**	−.16	−.22	.08	.18	.39	−.18	.22	.06	.09	**.81**	−.01	−.22	.09	.23	.31	−.06	.19
CTRL4	−.06	.09	**.76**	−.09	−.03	.11	.21	.38	−.02	.19	.02	.01	**.79**	.00	−.28	.05	.39	.40	−.02	.29
LIKE1	.10	.01	−.19	**.84**	.06	−.01	.00	.07	.40	.09	.21	.17	−.04	**.83**	.03	.15	.08	.06	.42	.18
LIKE2	.08	.00	−.01	**.71**	.00	.17	.19	.09	.35	.11	.20	.16	−.03	**.81**	.09	.06	.19	.18	.37	.09
PCONC	−.25	−.01	−.05	.00	**.75**	−.14	−.21	−.24	.09	−.20	−.15	−.16	−.24	.07	**.82**	−.06	−.18	−.39	.12	−.36
PCONE	−.13	.00	−.26	.09	**.78**	−.17	−.27	−.27	.08	−.28	−.05	−.09	−.24	.17	**.76**	.02	−.17	−.24	.18	−.29
PCONSU	−.37	.01	−.12	−.02	**.83**	−.34	−.37	−.30	−.01	−.30	−.29	−.08	−.28	−.01	**.82**	−.13	−.19	−.43	.00	−.29
PCONUA	−.25	.01	−.22	.09	**.84**	−.21	−.28	−.29	.09	−.32	−.13	.03	−.18	.02	**.74**	−.11	−.20	−.24	.06	−.29
PERBEN1	.26	−.09	.08	.07	−.17	**.84**	.29	.20	.07	.26	.26	.05	.04	.18	−.04	**.83**	.11	.22	.00	.27
PERBEN2	.28	−.05	.03	.10	−.28	**.86**	.37	.21	.05	.29	.19	.04	.08	.06	−.06	**.79**	.06	.08	−.05	.19
PERBEN3	.26	−.05	.05	.05	−.26	**.87**	.38	.31	.07	.44	.28	.10	.19	.08	−.13	**.85**	.20	.28	−.03	.30
REGUL1	.22	−.02	.17	.16	−.27	.37	**.88**	.29	.15	.39	.05	.01	.28	.20	−.17	.09	**.85**	.35	.17	.27
REGUL2	.21	−.01	.18	.08	−.18	.42	**.92**	.36	.08	.40	.14	.04	.34	.07	−.28	.17	**.90**	.39	.09	.37
REGUL3	.20	.00	.21	.06	−.33	.31	**.91**	.38	.04	.34	.07	.00	.27	.19	−.15	.14	**.87**	.38	.13	.30
TECHN1	.15	.17	.38	.09	−.30	.28	.32	**.91**	.06	.44	.29	.13	.40	.22	−.39	.23	.39	**.90**	.21	.46
TECHN2	.20	.09	.33	.11	−.34	.26	.36	**.91**	.10	.40	.27	.12	.40	.09	−.42	.22	.36	**.91**	.08	.44
TECHN3	.18	.07	.38	.08	−.28	.22	.34	**.86**	.06	.40	.18	.05	.42	.08	−.33	.21	.40	**.88**	.10	.37
EXP1	−.04	.02	−.07	.26	.09	.09	.07	.01	**.87**	−.07	.31	.17	−.10	.34	.17	−.03	.14	.10	**.92**	−.01
EXP2	−.04	−.03	−.09	.28	.03	.03	.09	.13	**.84**	.08	.29	.20	−.04	.37	.02	−.03	.12	.17	**.91**	.10
TRUST1	.17	−.06	.15	.03	−.26	.29	.35	.34	−.14	**.75**	.22	.17	.15	.26	−.19	.21	.23	.27	.20	**.67**
TRUST2	.17	−.08	.13	.21	−.24	.33	.40	.44	.14	**.83**	.29	.07	.27	.11	−.35	.31	.37	.44	.05	**.87**
TRUST3	.19	−.06	.24	.07	−.34	.30	.41	.42	−.01	**.83**	.26	.13	.28	.08	−.36	.22	.26	.46	−.06	**.82**

Notes: Item codes are given in the Appendix. Privacy concerns PCON are analyzed as a second-order latent variable (see Wetzels et al. 2009) constructed by the blocks of the underlying first-order latent variables PCONC, PCONE, PCONUA, and PCONSU themselves constructed by the 15-item CFIP scale (see Appendix)

Table 2.4 Construct reliability and validity criteria, Italy and U.S.

	Italy			U.S.		
	AVE	Composite reliability	Cronbach alpha	AVE	Composite reliability	Cronbach alpha
ATT	.83	.94	.90	.79	.92	.87
CO	.83	.91	.80	.85	.92	.83
CTRL	.64	.88	.81	.67	.89	.84
LIKE	.66	.75	.34	.67	.80	
PCON	.64	.88	.82	.62	.87	.80
PERBEN	.73	.89	.82	.68	.86	.77
REGUL	.82	.93	.89	.77	.91	.85
TECHN	.80	.92	.87	.81	.93	.88
EXP	.74	.85	.64	.84	.91	.81
TRUST	.74	.84	.72	.63	.83	.71

for US and for Italy. All factor loadings (given in bold in Table 2.3) are well above the generally accepted cut-off value of .7 and exhibit generally low cross-loadings, with which individual item reliability is met. This finding suggests that the indicators accounted for a large portion of the variance of the corresponding latent constructs and therefore provided support for the convergent validity of the measures. The item reliability and internal consistency is established by the composite reliabilities of the measured constructs (Table 2.4). Compared to Cronbach's alpha (Table 2.4), which provides a lower bound estimate of the internal consistency, composite reliability is known to be a more rigorous estimate of reliability. As evident, all composite reliabilities and Cronbach alpha are high thus indicating high internal consistency. Finally, discriminant validity refers to the extent to which the measures of the different model dimensions are unique. It is generally assessed by testing whether the square root of the Average Variance Extracted (AVE) of any latent variable is greater than the correlations shared between that latent variable and other latent variables. That is, if we place the square roots of AVEs in the construct correlation matrix, the diagonal element should be greater than the off-diagonal ones for a specific construct (Table 2.5). The numbers from Table 2.5, as well as the cross-loadings in Table 2.3 clearly demonstrate adequate discriminant validity.

Structural Model

After confirming the measurement validity and reliability for both Italian and U.S. samples, we ran the structural model with all data points from both samples, consistent with the model generalizability assumption. The results of the structural model are reported in Fig. 2.2. We also provide the models for United States and Italy in Appendix 2. Since none of the tested control variables had significant effect on the attitudes, we omitted them from the figure to make it more clear and succinct.

Table 2.5 Correlation Matrix. Diagonal elements are constructs' AVEs

	Italy										U.S.									
	ATT	CO	CTRL	LIKE	PCON	PERBEN	REGUL	TECHN	EXP	TRUST	ATT	CO	CTRL	LIKE	PCON	PERBEN	REGUL	TECHN	EXP	TRUST
ATT	.83										.79									
CO	.05	.83									.27	.85								
CTRL	.04	.10	.64								.04	.01	.67							
LIKE	.12	.01	−.14	.60							.25	.20	−.04	.67						
PCON	−.32	.00	−.20	.05	.64						−.21	−.10	−.30	.07	.62					
PERBEN	.31	−.08	.07	.09	−.28	.73					.30	.09	.13	.13	−.09	.68				
REGUL	.23	−.01	.21	.11	−.36	.41	.82				.10	.02	.34	.16	−.24	.16	.77			
TECHN	.20	.12	.41	.10	−.35	.28	.38	.80			.27	.11	.55	.14	−.42	.25	.43	.81		
EXP	−.05	.00	−.09	.40	.07	.07	.10	.08	.74		.32	.21	−.08	.46	.10	−.03	.14	.14	.84	
TRUST	.22	−.08	.22	.13	−.35	.39	.50	.52	.00	.74	.33	.14	.30	.17	−.39	.32	.36	.51	.05	.83

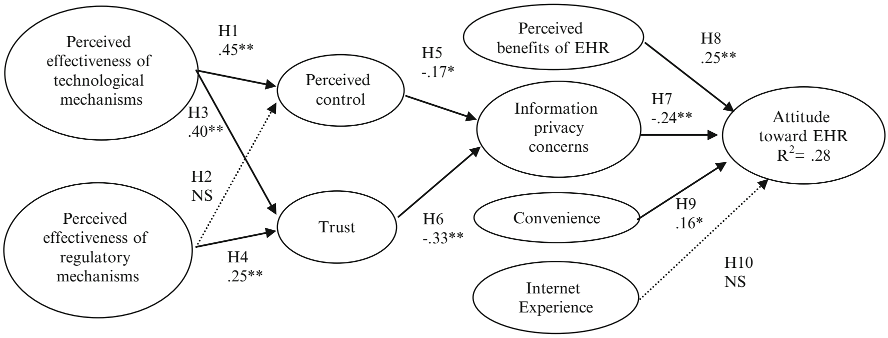


Fig. 2.2 Results of structural model testing. Notes: Path coefficients' statistical significance: *p < .05; **p < .01, *NS* not significant, *bold arrow* significant path; *dashed arrow* insignificant path

2.4 Discussion of Results

The results indicate overall support for our hypotheses. Eight of our ten hypotheses are supported when we test the model with the Italian and American data together. The perceived effectiveness of technology privacy mechanisms shows significant positive effect on both perceived privacy control and trust, thus supporting H1 and H3. The perceived effectiveness of privacy enhancing regulatory mechanisms, however, show positive effect only on trust (H4 supported) but not on the perceived privacy control over EHR (H2 not supported). Both trust and perceived privacy control can significantly reduce the privacy concerns towards EHR (H5 and H6 supported), and the privacy calculus components (privacy concerns, perceived benefits, and convenience) provide the significant competing influence on attitudes towards EHR (H7, H8, and H9 respectively supported). Finally, the Internet experience had no significant effect on the attitudes and thus H10 was not supported in the model testing with both nations' samples.

As we mentioned above, the structural model was run with both nations' data as one data set. Even though some national differences can be expected in the strengths of the hypothesized relationships, the general framework of the model should be valid for both nations, per the generalizability assumptions. Indeed, as discussed in the Introduction and the theoretical section, the basic arguments for the relationships in the model are not nation biased: the urgent need to implement digital health care record system; the prevalent suspicion of both American and European (Italian) citizens of possible security breaches or errors in the systems that may result in disclosure of personal health data, possibly the most sensitive type of personal data; and with this substantial loss of privacy with negative consequences for the individual. These common foundations provide the common background on which the cultural-independent view of the developed model was advanced.

In order to get deeper insight into the results of the model, and especially to interpret the lack of support for H2 and H10, we also ran two separate models with the Italian and U.S. data. H2 remained to be insignificant for both samples. Thus, we can conclude that individuals do not tend to rely on regulatory mechanisms in building their perceived privacy control over the EHR. Only the technological privacy mechanisms enhance the perceived privacy control. A possible explanation is that individuals look at their ability to control how their health care information is collected, used, and distributed as technology empowerment rather an empowerment given by the proxy from the top. Indeed, the technological advances provide institutions with ever more sophisticated technology to empower individuals to regulate themselves how the storage and flow of their personal information is done, with which a true regulatory mechanism would not further enhance control beliefs.

We therefore find that in the technological age, users tend to look at control as a tool that they "can see and touch" and use by themselves—that is, if there is technology in place that helps them enhances their privacy, they would tend to feel more in control. The lack of support for H10 will be addressed in the limitations section.

Consistent with the trust theories (see for example (Dinev et al. 2006)) in which government regulations and institutional self-regulations play an important role in building trust, we find that, indeed, the relationship between regulatory privacy mechanisms and trust is strong and even stronger for Italy (.36 for Italy vs. .18 for US) where there is substantial more reliance on government regulations than in U.S.

Finally, it is important to note that the calculus perspective of the formation of attitudes towards EHR is evident in the model' results. As seen from the path coefficients in Fig. 2.2, perceived benefits along with convenience *can* override the very strong impact of privacy concerns and provide an overall significant positive attitude towards EHR. Thus, the results of our model are in accordance with Angst and Agarwal (2009) who also found that "privacy concerns, while a salient barrier, may not be enough to halt the acceptance of electronic healthcare records" (p. 358) and that "through proper messaging and education, attitudes can be changes, even in the presence of great privacy concerns" (p. 360). As long as users are informed, educated, and convinced in the convenience and benefits of EHR, they would be able to form positive attitude (that would lead to adoption) even if their privacy concerns are high.

2.5 Limitations and Suggestions for Future Research

There are other model relationships that exhibit some differences between Italy and U.S. and these may be significant and due to cultural differences. While investigating the reasons and significance of these differences is beyond the limits of the current paper, they provide substantial opportunities for future research. The other insignificant relationship in the model, the impact of Internet Experience on Attitudes towards EHR (H10), demands a closer look since this is the only relationship the significance of which differs across the two nations. Indeed, the separate model runs for both countries show a non significant relationship for Italians, but a strong (at level $p < .01$) relationship for the U.S. individuals—.35**. We believe that the explanation probably can be found with the slower rate of adoption of e-commerce in Italy in comparison to United States (Dinev et al. 2006) and the nature of the measurement that we adopted. We now realize that the measure for Internet experience (see Appendix) is biased towards Internet shopping rather than other activities. Not as many Italians shop on the Internet as Americans as evident from Table 2.2. Italians are known to actively use the Internet for social networking and communication activities rather than shopping (Dinev et al. 2006). This is a significant limitation of the study that would have not shown as such if the sample had been only American. In order to really understand how Internet and technology experience shape attitudes, further studies that lack the bias of ours are needed. We posit that the relationship will be significant for the Italians as well, thus confirming previous studies that showed positive influence of experience on attitudes.

The current study is also limited by its research question—it only looks at the individuals', citizens' calculus and cost-benefit analysis of attitudes toward EHR—and thus is a one-sided perspective. Like in every major economic development, the other side of the perspective—the institutions and their cost-benefit analysis for adoption of HER—has also to be taken into account if full understanding is sought. Recent studies have investigated this side and found that privacy protection may inhibit adoption if hospitals cannot benefit from easily exchanging patient information (Miller and Tucker 2009). According to the researchers, state privacy regulation restricting hospital release of health information reduces aggregate EMR adoption by hospitals by more than 24 % which can prevent achieving the government goal of having a national health IT network by 2014 (Miller and Tucker 2009).

2.6 Implications for Theory and Practice

Our study has several practical and theoretical implications. First, it offers a comprehensive theoretical model based on cost-benefit analysis that explains the formation of individuals' attitudes in health care context. We focused on the antecedents of information privacy concerns as well as the consequences, thus filling a gap in the privacy literature and potentially leading to a more complete understanding of this important construct. The coherently developed and rigorously tested empirical map of antecedents in this model reflects the complexity of privacy concerns formation. The study highlights the roles organizational interventions in increasing privacy control and trust, which are important in diverse sectors. These contributions lead to important implications for both practice and research.

The study's results and implications align well with the conclusion of Angst and Agarwal (2009) who found that, through message framing and persuasion, positive attitudes can be promoted even in the presence of very high privacy concerns. Through the lens of the privacy calculus, our study shows similar path to overcoming the highly negative effect of privacy concerns in the health care and EHR context, and in addition it shows how to reduce the privacy concerns: (1) By empowering users with advanced technology to control on their own their privacy preferences, and by building trust through smart regulation and technology, we can substantially reduce privacy concerns, and (2) We can beat the negativity through convincing individuals to see the other very important components in the equation—the benefits, the convenience, the opportunity for technological advancements when implementing EHR. The results of this investigation should be of interest to government agencies which have oversight in establishing laws and regulations related to digitized health information. Individuals involved in any future efforts to evaluate the consequences of privacy policies and practices, for example, or to initiate new rulemaking procedures in this area should find this study useful. Our study shows that in order to overcome the negativity and suspicion rooted in high privacy concerns, broad and comprehensive raising of public awareness through education and communication should be implemented as to the benefits of EHR and the costs to the society and individuals if we do not implement them. The study is also very timely and relevant in the era of a concerted government and institutional effort to implement in health care comprehensive information technology solutions and digitization of records, so cost can be substantially cut and benefits to the patients enhanced.

The importance of privacy enhancing IT tools are emphasized in the recent study about the institutional EHR adoption impediments (Miller and Tucker 2009). The authors call for further research to find the optimum privacy protection that can be put in place that will minimize disruption to the diffusion and use of interdependent technologies. They refer to IT-based privacy protection (the same technological privacy control mechanisms that we are incorporating in our study) as a possible tool that can serve the balancing role and expedite adoption. We completely agree with the authors, especially in light of the recent findings by Agarwal et al. (2010) who posit that for most providers in-house development of EHR applications is neither feasible nor economical. Thus, commercial off-the-shelf solutions, especially Web-based services, will be the mainstream of digital medical records' adoption in the next few years. The adopted systems come with predefined interfaces and functionality, which may be incompatible with existing practices and the best intentions of the providers. So at this crucial moment, it is even more important for the providers to demand from the IT developers and designers a set of privacy-enhancing mechanisms and IT rules that will enable optimum control and trust by the individuals and will help advance the adoption of the digital health records.

2.7 Conclusion

Both U.S. and Europe see EHR adoption as strategic twenty-first century step to making heath care more efficient, modern, with lower costs and more benefits. The research objective of this study is to investigate individuals' attitudes towards EHR and what factors form these attitudes, themselves a strong predictor of behavior. We found that perceived effectiveness of regulatory mechanisms positively impact trust and that perceived effectiveness of technological mechanisms positively impacts perceived privacy control and trust. Both trust and perceived privacy control can substantially reduce privacy concerns which, along with perceived benefits, convenience, and Internet experience, play an important role in the privacy calculus-type formation of individual's attitudes towards EHR. Our model showed that two paths have to be undertaken by policy makers and executives to ensure wide adoption of EHR: (1) Reduce privacy concerns by enhancing trust and control beliefs (through technological and regulatory mechanisms), and (2) Overcome the negative impact of privacy concerns through raising awareness and message impact on the benefits and convenience of EHR.

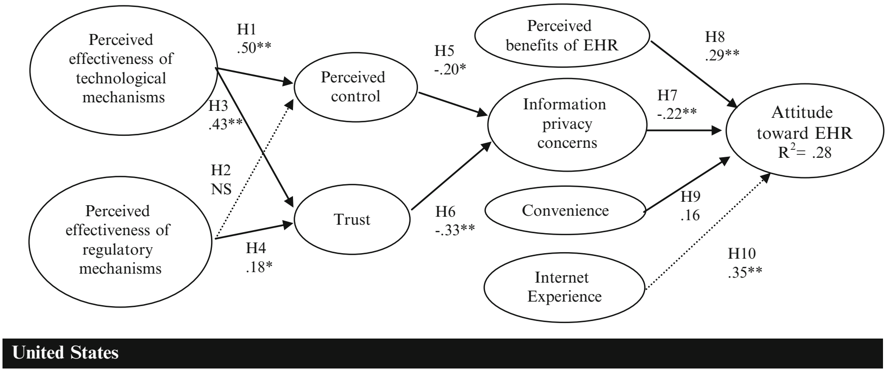
Appendix 1: Survey Instrument. All Items Were Measured on 5-Point Lickert Scale, Lowest "Completely Disagree" to Highest—"Completely Agree" Unless Specified Otherwise

Construct	Code	Item (Indicate to what extent you agree with the following:) Bold items were used in the model test
Attitude toward EHR **ATT**	**att1**	I believe it is a good idea to have electronic health records
	att2	I believe that electronic health records is a good thing to do
	att3	I have a favorable opinion about electronic health records
Convenience **CO**	**co1**	My own online medical record would help me get all my doctors on the same page when they treat me
	co2	I'd like to have all my health information in one place—and get to it with the click of a mouse
	co3	I can access my bank account online. Why not my medical records?
	co4	I've often felt the health care system has all the power. Having my own online health record seems to even it out a little bit
	co5	I'm tired of playing 'telephone tag' with doctors and filling out the same forms. Why can't I do some of this stuff online?

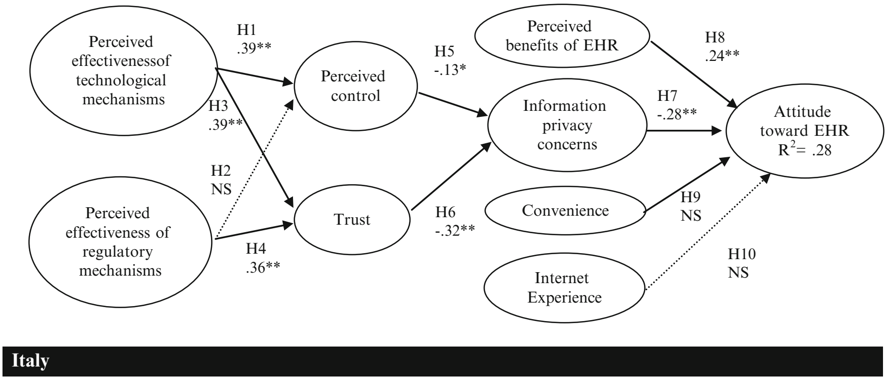
Construct	Code	Item (Indicate to what extent you agree with the following:) Bold items were used in the model test
Perceived privacy control **CTRL**	**ctrl1**	I think I will have control over what personal information is released by the electronic health records systems
	ctrl2	I believe I will have control over who can get access to my personal information collected by the electronic health records systems
	ctrl3	I believe I will have control over how personal information is used by the electronic health records systems
	ctrl4	In general, I believe I will be able to control my personal information provided to health care
	ctrl5	I believe I will have control over the amount of my personal information collected by the electronic health records systems
Satisfaction with current health care**LIKE**	**like1**	I like our current health care system
	like2	I believe the current health care system is good enough
Privacy concerns—collection **PCONC**	**pconc1**	It usually bothers me when health care entities ask me for personal information.
	pconc2	When health care entities ask me for personal information, I sometimes think twice before providing it
	pconc3	It bothers me to give personal information to so many health care entities
	pconc4	I'm concerned that health care entities are collecting too much personal information about me
Privacy concerns—errors **PCONE**	**pcone1**	All the personal information in computer database should be double-checked for accuracy—no matter how much this costs
	pcone2	Health care entities should take more steps to make sure that the personal information in their files is accurate
	pcone3	Health care entities should have better procedures to correct errors in personal information
	pcone4	Health care entities should devote more time and effort to verifying the accuracy of the personal information in their databases
Privacy concerns—unauthorized access **PCONUA**	**pconua1**	Health care entities should devote more time and effort to preventing unauthorized access to personal information
	pconua2	Computer databases that contain personal information should be protected from unauthorized access no matter how much it costs
	pconua3	Health care entities should take more steps to make sure that unauthorized people cannot access personal information in their computers

Construct	Code	Item (Indicate to what extent you agree with the following:) Bold items were used in the model test
Privacy concerns—secondary use **PCONSU**	**pconsu1**	Health care entities should not use personal information for any purpose unless it has been authorized by the individuals who provided the information
	pconsu2	When people give personal information to a company for some reason, the company should never use the information for any other reason
	pconsu3	Health care entities should never sell the personal information in their computer databases to other health care entities
	pconsu4	Health care entities should never share personal information with other health care entities unless it has been authorized by the patient who provided the information
Perceived benefits **PERBEN**	**perben1**	I believe that it is beneficial for me to have my health records electronically
	perben2	I believe electronic health records will improve health care
	perben3	Electronic health records will generate positive results for the health care in our society
Perceived effectiveness of regulatory mechanisms **REGUL**	**regul1**	I believe that the law is effective in protecting me from misuse of my personal electronic health care data
	regul2	I believe that the law effectively governs the practice of how my electronic health care records are collected, used, and protected
	regul3	I believe that the current regulations are able to address violations in usage of my electronic health care records
Perceived effectiveness of technological mechanisms **TECHN**	**techn1**	I think that the electronic health records will use effective technologies
	techn2	I think that the electronic health records will use reliable technologies
	techn3	I think that there are a lot of good technologies that will help the electronic health records
Trust	**trust1**	I think that electronic health records are dependable
	trust2	I think that electronic health records are trustworthy
	trust3	I trust that electronic health records provide good service
Health status **HS**	**hs**	In general, would you say your health is (Poor—Excellent)
Experience with internet**EXP**		How frequently did you use the Internet for the following activities (Never-Always)
	exp1	Shopped or purchased something
	exp2	Made travel arrangements or bought an airline ticket
	exp3	Paid bills or managed financial accounts

Appendix 2: Structural Model for U.S. and Italy



Notes: Path coefficients' statistical significance: *p < .05; **p < .01, *NS* not significant, *bold arrow* significant path, *dashed arrow* insignificant path



Notes: Path coefficients' statistical significance: *p < .05; **p < .01, *NS* not significant; *bold arrow* significant path, *dashed arrow*—insignificant path

References

Agarwal R, Gao G, DesRoches C, Jha AK (2010) The digital transformation of healthcare: current status and the road ahead. Inform Syst Res 21(4):796–809

Ajzen I (1991) The theory of planned behavior. Org Behav Hum Decis Process 50(2):179–211

Ajzen I, Fishbein M (1980) Understanding attitudes and predicting social behavior. Prentice-Hall, Englewood Cliffs

Alonso-Zaldivar R (2009) U.S. health care lags in bang for the buck. San Francisco Chronicle, March 12, p. 1

Angst MC, Agarwal R (2006a) Overcoming personal barriers to adoption when technology enables information to be available to others. In: Proceedings of 27th international conference on information systems, Milwaukee

Angst MC, Agarwal R (2006b) Getting personal about electronic health records: modeling the beliefs of personal health record users and non-users. Working Paper, Robert H. Smith School Research Paper No. RHS-06-007. Available at SSRN: http://ssrn.com/abstract=902904

Angst MC, Agarwal R (2009) Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS Quarterly 33(2):339–370

Bandura A (1982) Self-efficacy mechanism in human agency. Am Psychol 37:122–147

Bandura A (1986) Social foundations of thought and action. Prentice-Hall, Englewood Cliffs

Bansal G, Zahedi F, Gefen D (2007) The impact of personal dispositions on privacy and trust in disclosing health information online. In: Proceedings of AMCIS, Paper 57. http://aisel.aisnet.org/amcis2007/57

Bansal G, Zahedi FM, Gefen D (2010) The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decis Support Syst 49:138–150

Beatty RC, Shim JP, Jones MC (2001) Factors influencing corporate Web site adoption: a time-based assessment. Inform Manage 38(6):337–354

Bishop LS, Holmes BJ, Kelley CM (2005) National Consumer Health Privacy Survey 2005. California Health Care Foundation, Oakland. http://ehealth.chcf.org/topics/view.cfm?itemID=115694

Bodenheimer T, Grumbach K (2003) Electronic technology: a spark to revitalize primary care? J Am Med Assoc 290(2):259–264

Cantor JD (2001) Privacy protections for cybercharts: an update on the law. J Am Med Assoc 285(13):1767

Caudill ME, Murphy EP (2000) 2000 consumer online privacy: legal and ethical issues. J Public Policy Mark 19(1):7–19

Chellappa RK, Sin RG (2005) Personalization versus privacy: an empirical examination of the online consumer's dilemma. Inform Technol Manage 6(2):181–202

Childers TL, Carr CL, Peck J, Carson S (2001) Hedonic and utilitarian motivations for online retail shopping behavior. J Retail 77(4):511–535

Choy A, Goldman J (2001) Comparing eHealth Privacy Initiatives, Health Privacy Project. Georgetown University. http://www.chcf.org/topics/view.cfm?itemID=12739

Choy A, Hudson Z, Pritts J (2002) Exposed online: why the new federal health privacy regulation doesn't offer much protection to internet users, the Pew Internet and American Life Project. http://medicalrecordrights.georgetown.edu/publications.html

Cimino JJ, Patel VL, Kushniruk AWT (2002) The patient clinical information system (PatCIS): technical solutions for and experience with giving patients access to their electronic medical records. Int J Med Inform 68(1–3):113–127

Culnan MJ (1993) 'How did they get my name'? An exploratory investigation of consumer attitudes toward secondary information use. MIS Quarterly 17(3):341–364

Culnan MJ, Armstrong PK (1999) Information privacy concerns, procedural fairness and impersonal trust: an empirical investigation. Organ Sci 10(1):104–115

Culnan MJ, Bies JR (2003) Consumer privacy: balancing economic and justice considerations. J Soc Issues 59(2):323–342

Dan JK, Ferrin DL, Rao HR (2008) A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risk, and their antecedents. Decis Support Syst 44:544–564

Dinev T, Hart P (2004) Internet privacy concerns and their antecedents—measurement validity and a regression model. Behav Inform Technol 23(6):413–423

Dinev T, Hart P (2006) An extended privacy calculus model for e-commerce transactions. Inform Syst Res 17(1):61–80

Dinev T, Bellotto M, Hart P, Russo V, Serra I, Colautti C (2006) Privacy calculus model in e-commerce—a study of Italy and the United States. Eur J Inform Syst 15(4):389–402

Dinev T, Hu Q, Yayla A (2008) Is there an online advertisers' dilemma? A study of click fraud in the pay-per-click model. Int J Electron Commerce 13(2):29–59

Dixon P (2005) Electronic health records and the National Health Information Network: patient choice, privacy, and security in digitized environments, presented before The National Committee on Vital and Health Statistics (NCVHS) Subcommittee on privacy and confidentiality, San Francisco, CA, August 2005. http://www.worldprivacyforum.org/pdf/pamdixonNCVHStestimonyfinal.pdf

Earp JB, Payton FC (2006) Information privacy in the service sector: an exploratory study of health care and banking professionals. J Organ Comput Electron Comm 16(2):105–122

Faden RR, Beauchamp LT, King PMN (1986) A history and theory of informed consent. Oxford University Press, New York

Forsythe S, Liu CL, Shannon D, Gardner LC (2006) Development of a scale to measure the perceived benefits and risks of online shopping. J Interact Market 20(2):55–75

Gazzetta Ufficiale (2009) Garante per la protezione dei dati personali. Linee guida in tema di Fascicolo sanitario elettronico (Fse) e di dossier sanitario—16 luglio 2009. n. 178 03-08-2009. http://www.garanteprivacy.it/garante/doc.jsp?ID=1634116

Gefen D, Straub DW, Boudreau MC (2000) Structural equation modeling and regression: guidelines for research practice. Commun Assoc Inform Syst 4(7):1–78

Goodwin C (1991) Privacy: recognition of a consumer right. J Public Policy Mark 10(1):149–166

HHS (2006) Personal health records and personal health record systems: a report and recommendations from the National Committee on Vital and Health Statistics. Department of Health and Human Services, Washington DC, February, 2006. http://www.ncvhs.hhs.gov/0602nhiirpt.pdf

Harris AF (2002) Interactive and Westin, The Harris Poll: #46. Harris Interactive, New York

Hu X, Wu G, Wu Y, Zhang H (2010) The effects of Web assurance seals on consumers' initial trust in an online vendor: a functional perspective. Decis Support Syst 48:407–418

Iacovou CL, Benbasat I, Dexter AS (1995) Electronic data interchange and small organizations: adoption and impact of technology. MIS Quarterly 19(4):465–485

Jarvenpaa SL, Tractinsky N, Vitale M (2000) Consumer trust in an internet store. Inform Technol Manage 1(12):45–71

Johnson LJ, Cullen BJ (2002) Trust in cross-cultural relationships. In: Gannon MJ, Newman KL (eds) The Blackwell handbook of cross-cultural management. Blackwell, Oxford, pp 335–360

Karahanna E, Evaristo JR, Strite M (2002) Methodological issues in MIS cross cultural research. J Global Inform Manage 10:13–23

Kauffman T (2006) Your health records online: OPM: benefits will outweigh risks. Federal-Times.com, May 15 2006. http://federaltimes.com/index.php?S=1771146

Kidd MR (2008) Personal electronic health records: MySpace or HealthSpace? Br Med J 336:1029–1030

Landro L (2004) Electronic medical records are taking root locally. Wall Street J 22:D1

Laufer RS, Wolfe M (1977) Privacy as a concept and a social issue: a multidimensional developmental theory. J Soc Issues 33(3):22–42

Liu CL (2007) Modeling consumer adoption of the internet as a shopping medium: an integrated perspective. Cambria, Youngstown

Lo B, Clayton P, Crowley J, Glaser J, Hunter N, Koss S, Koyanagi C, Nielsen J, Shelton L, Van Amringe M (1999) Best principles for health privacy. Report for the Health Privacy Project, Georgetown University. http://www.healthprivacy.org/usr_doc/33807.pdf

Luck J, Chang C, Brown ER, Lumpkin J (2006) Using local health information to promote public health. Health Aff 25(4):979–991

Masys D, Baker D, Butros A, Cowles KE (2002) Giving patients access to their medical records via the internet: the PCASSO experience. J Am Med Inform Assoc 9(2):181–191

Mayer RC, Davis JH, Schoorman FD (1995) An integrative model of organizational trust. Academy Manage Rev 20(3):709–734

MacKay N, Parent M, Gemino A (2004) A model of electronic commerce adoption by small voluntary organizations. Eur J Inform Syst 13(2):147–159

McKnight DH, Choudhury V, Kacmar C (2002) Developing and validating trust measures for e-commerce: an integrative typology. Inform Syst Res 13(3):334–359

Miller AR, Tucker C (2009) Privacy protection and technology diffusion: the case of electronic medical records. Manage Sci 55(7):1077–1093

Milne GR, Boza ME (1999) Trust and concern in consumers' perceptions of marketing information management practices. J Interact Market 13(1):5–24

Milne GR, Gordon EM (1993) Direct mail privacy-efficiency trade-offs within an implied social contract framework. J Public Policy Mark 12(2):206–215

Moffit R, Manière P, Green DG, Belien P, Hjertqvist J, Breyer F (2001) Perspectives on the European health care systems: some lessons for America. Heritage Lecture #711. http://www.heritage.org/Research/Healthcare/HL711.cfm

Pagliari C, Detmer D, Singleton P (2007) Potential of electronic personal health records. Br Med J 335:330–333

Pavlou PA, Fygenson M (2006) Understanding and predicting electronic commerce adoption: an extension of the theory of planned behavior. MIS Quarterly 30(1):115–143

Pavlou PA, Gefen D (2004) Building effective online marketplaces with institution-based trust. Inform Syst Res 15(1):37–59

Phelps J, Nowak G, Ferrell E (2000) Privacy concerns and consumer willingness to provide personal information. J Public Policy Market 19(1):27–41

Rindfleish TC (1997) Privacy, information technology, and health care. Commun ACM 40(8):92–100

Ringle CM, Wende S, Will S (2005) SmartPLS 2.0, Hamburg, Germany. http://www.smartpls.de

Sheehan KB, Hoy GM (2000) Dimensions of privacy concern among online consumers. J Public Policy Market 19(1):62–73

Skinner EA, Chapman M, Baltes PB (1988) Control, means-ends, and agency believes: a new conceptualization and its measurement during childhood. J Personality Social Psychol 54:117–133

Shortliffe EH (1999) The evolution of electronic medical records. Acad Med 74(4):414–419

Smith HJ, Milberg JS, Burke JS (1996) Information privacy: measuring individuals' concerns about organizational practices. MIS Quarterly 20(2):167–196

Song J, Zahedi FM (2007) Trust in health infomediaries. Decis Support Syst 43:390–407

Spiro WG, Houghteling LJ (1981) The dynamics of law, 2nd edn. Harcourt Brace Jovanovich, New York, pp 2–10

Stewart KA, Segars AH (2002) An empirical examination of the concern for information privacy instrument. Inform Syst Res 13(1):36–49

Stone EF, Gueutal GH, Gardner DG, McClure S (1983) A field experiment comparing information—privacy values, beliefs, and attitudes across several types of organizations. J Appl Psychol 68(3):459–468

Stone EF, Stone DL (1990) Privacy in organizations: theoretical issues, research findings, and protection mechanisms. Res Personnel Human Resources Manage 8(3):349–411

Taylor S, Todd PA (1995) Understanding information technology usage: a test of competing models. Inform Syst Res 6(3):144–176

Teo TSH, Yeong YD (2003) Assessing the consumer decision process in the digital marketplace. Omega—Int J Manage Sci 31(5):349–363

Tittle CR (1980) Sanctions and social deviance: the question of deterrence. Praeger, New York

TSE (2005) Tavolo di Sanità Elettronica—Italian National Board for eHealth. A shared policy for e-Health. http://www.sanitaelettronica.gov.it/se/documenti/TSE-A_Policy_for_eHealth-ENG-DEF.pdf

TSE (2006) Tavolo di Sanità Elettronica—Italian National Board for eHealth, Architectural Strategy for Electronic Health care. http://www.sanitaelettronica.gov.it/se/documenti/TSE_IBSE_Architectural_Strategy-ENG-DEF.pdf

Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: toward a unified view. MIS Quarterly 27(3):425–478

Westin AF (2003) Social and political dimensions of privacy. J Soc Issues 59(2):431

Wetzels M, Odenkerken-Schroder G, van Oppen C (2009) Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration. MIS Quarterly 33(1):117–195

Xin L, Li H, Zhang J, Shim JP (2010) Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: an empirical study of mobile banking services. Decis Support Syst 49:222–234

Xu H (2007) The effects of self-construal and perceived control on privacy concerns. In: Proceedings of the 28th annual international conference on information systems (ICIS 2007), Montréal, Canada, Paper 125

Xu H, Teo HH (2004) Alleviating consumer's privacy concern in location-based services: a psychological control perspective. In: Proceedings of the twenty-fifth annual international conference on information Systems (ICIS 2004), Washington, DC, pp 793–806

Xu H, Teo HH, Tan BCY (2005) Predicting the adoption of location-based services: the roles of trust and privacy risk. In: Proceedings of 26th annual international conference on information systems (ICIS 2005) Las Vegas, pp 897–910

Xu H, Dinev T, Smith J, Hart P (2008) Examining the formation of individual's privacy concerns: toward an integrative view. In: Proceedings of the international conference of information systems (ICIS 2008), Paris, Paper 6

Xu H, Teo HH, Tan BCY, Agarwal R (2009) The role of Push-Pull technology in privacy calculus: the case of location-based services. J Manage Inform Syst 26(3):137–176

Xu, H., Teo, H. H., Tan, B.C.Y., and Agarwal, R. (2012). Effects of Individual Self-Protection, Industry Self-Regulation, and Government Regulation on Privacy Concerns: A Study of Location-Based Services, Information Systems Research, 23(4): 1342–1363.

Yamaguchi S (2001) Culture and control orientations. In: Matsumoto D (ed) The handbook of culture and psychology. Oxford University Press, New York, pp 223–243

Chapter 3
Understanding Lifestyle Decisions Based on Patient Historical Data: A Latent Growth Modeling Approach

Ronald Freeze[1], T.S. Raghu[2], and Ajay Vinze[2]

[1]University of Arkansas, Fayetteville, AR, USA
[2]Arizona State University, Tempe, AZ, USA

Abstract Healthcare issues related to chronic disease conditions and management does not have easy or immediate solutions. Evidence based decision making in such contexts requires long-term tracking and analysis of patient data in order to provide patient choices that produce extended quality of life. Using Latent Growth Modeling (LGM), we present a planning perspective to analyze underlying patterns of long-term chronic data related to the progression of Multiple Sclerosis (MS). Using the North American Research Committee on Multiple Sclerosis (NARCOMS) patient driven initiative that collects survey data on a biannual basis for the purpose of clinical trial recruitment and epidemiological research, this study analyzes three temporal data points spanning 3 years. Two LGM models are presented that identify patient traits correlating with disease progression. The traits analyzed are both patient and physician controlled.

Keywords Health outcomes • Structural equation modeling • Latent growth modeling • Healthcare • Multiple sclerosis • Lifestyle decisions

---

R. Freeze (✉)
University of Arkansas, Fayetteville, AR, USA
e-mail: rfreeze@uark.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_3

3.1 Introduction[1]

Health Care and the medical profession have historically focused its attention and efforts toward ameliorating the patient's current conditions. The collection of data and information across a wider spectrum has prompted a shift to the emphasis on preventive measures that improve individual chances to forestall or even eliminate the emergence of adverse conditions. For MS patients, the ability to track long-term disease progression factors can aid their current and future condition by providing guidance as to what activities forestall the emergence of conditions or advancement of most debilitating symptoms of the disease. The increased use of the World Wide Web for the provision of information that patients consume (Song and Zahedi 2007) and as a method for sharing knowledge about specific diseases (Liang et al. 2006) highlights this shift to a more proactive patient-centered approach. Individual patients, or even proactive healthy individuals, need access to health care information in a concise and useful presentation that is relevant to their individual behaviorally focused decision-making needs. However, even though much information is stored in databases and data warehouses, evidence based decision-making still relies upon a fraction of the data (Pedersen and Larsen 2001) and analytic techniques are necessary to improve this decision-making (Berndt et al. 2003). The application of healthcare information technology must recognize the proactive shift in focus to prevention for individual patients.

Knowledge development, as the first step in the Healthcare Value Chain presented by Porter and Teisberg (2006), emphasizes the results measurement and tracking necessary to capture and display evidence in a patient focused manner. The development of this proactive knowledge includes the improvement of technology and processes within the healthcare community to accurately track patients over extended periods. The second step in the healthcare value chain emphasizes the informing of patients through education and counseling. The promotion of this evidence based value creation requires a planning perspective that will initiate behavior modifications, improve health maintenance awareness that can proceed to improving patient life-style quality. This development of self-care activities for MS can follow the success of diabetic patients who participate in their own health promotion planning (Kelley et al. 2011).

A planning perspective on healthcare necessitates expertise in data collection practices that have the ability to track individual health care status over time. A systematic method of capturing health status at the community-level exemplifies this planning perspective (Berndt et al. 2003). The development of a community health profile, comprising socio-demographic characteristics, health status, quality of life indicators, health risk factors, and health resource measures, provides an example of the data collection necessary to empower change for the improvement

[1] A prior version of this paper was presented at the Hawaiian International Conference of System Sciences in January, 2009.

of health. This knowledge sharing of accumulated information is desirable for the practice of healthcare (Liang et al. 2006). The resulting success of this form of health promotion relies heavily on behavior change. This means that people should adapt to a lifestyle that helps them maintain an optimal health status. To reach these goals requires planning that incorporates the ability to track progress over time for each individual (Berndt et al. 2003).

Chronic disease management can especially benefit from long-term tracking and analysis of individual patient conditions. In this context, the application of analytic techniques capable of assessing the long-term nature of chronic disease progression for individuals is still an emerging area of research (Serva et al. 2011). While representation and measurement of change over time is a fundamental concern to practically all scientific disciplines (Duncan and Duncan 2004), chronic disease tracking and analysis presents some unique challenges and requirements for analysis. Analysis of disease progression requires the incorporation of individual differences in conditions, disease history, treatment options and lifestyle choices. These conditions incorporate individual change over time and allow the identification of factors that make individuals perceive their health status as "better" or "worse". The analysis period for chronic diseases could include periods ranging from weeks to years between assessments.

In this research, we have selected Latent Growth Modeling (LGM) as the method to analyze chronic disease data (MS) due to its unique ability to identify changes in individuals over time (Serva et al. 2011). Insights related to individual differences can help to assist and plan disease management issues. By analyzing differences in individual conditions, critical factors can be identified that differentiate individuals that have a perceived higher quality of life. LGM uses longitudinal experiments when assessing a specific condition and factors that change those conditions. The base LGM model assumes the specific condition under investigation is a linear function (with other representations possible). The individual factors (called covariates) modify the specific condition being analyzed. The different conditions analyzed for the patients over time include the disease progression rate (referred to as the slope of the LGM) and the disease starting point or initial data point of the study (referred to as the intercept). By incorporating individual differences analysis, we can investigate a large dataset of multiple sclerosis (MS) patients to investigate whether life style choices affect the long-term quality of life for these patients. MS fits well within this type of analysis due to the life-long nature of the disease. Our goal is to empirically analyze the impact of patient activity, treatments and other demographic variables such as age, race and gender, and associated life style choices on patient self-reported health condition. Patient success at choosing certain lifestyle and treatment programs through understanding of the long-term effects of MS can motivate other MS patients to enact similar activities to improve their own quality of life.

Section 2 details the context of the multiple sclerosis study in terms of the dataset and the patient participants. In Sect. 3, we present a review of the basic application of LGM to multiple sclerosis. This section also reviews the results of the LGM models. Finally, in Sect. 4 we conclude with a discussion of the broader health care context and the potential benefits of using LGM as an analytical tool. We finish with some challenges associated with the usage of LGM (Sect. 5) and our final study conclusion (Sect. 6).

3.2 Multiple Sclerosis Context

Multiple sclerosis (MS) is a neurological disease that can be diagnosed in early childhood or infancy. The rate of disease progression for patients extends for many years and does not generally shorten an individual's lifespan. While MS is not the cause of death, the quality of life for patients in terms of mobility is affected. Mobility is severely affected in the late stages of the disease and patients progress more rapidly to the final disease stages once reaching a certain mobility stage (Achiron et al. 2003).

The North American Research Committee on Multiple Sclerosis (NARCOMS) has tracked long-term disease progression over several years through a patient registry (Vollmer et al. 1999). The survey conducted by NARCOMS is administered every 6 months with a respondent population ranging from 8000 to over 10,000 MS patients for the 3-year study period. The NARCOMS database includes patients as young as age 12 with 25 % of the MS population below the age of 32. One of the key data points of the survey is a 9-point scale indicating the Patient Determined Disease Steps (PDDS), a proxy for progression of disease (Table 3.1). A significant sign of MS progression is the mobility of the patient; the PDDS scale measures the decreased mobility using an ordinal scale. The use of patient rated experiences has a rich history within MS (Freeman 1999), is considered normal in MS research (Hamalainen et al. 1999) and the NARCOMS registry has been used and validated in prior research (Marrie et al. 2005).

The extended temporal nature of MS progression requires techniques that can incorporate patient self-reported disease status longitudinally or as a repeated

Table 3.1 PDDS question

Value	Meaning
0	**Normal**: mild symptoms, mostly sensory
1	**MildDisability**: some symptoms, minor lifestyle effect
2	**ModerateDisability**: No walking limitation, some daily activity limit
3	**GaitDisability**: Walking and activity interference, athletic activities are demanding
4	**EarlyCane**: Cane, Crutch or other support used, can walk 25′
5	**LateCane**: Must have support to walk 25′
6	**BilateralSupport**: 2 canes/crutches or walker to walk 25′
7	**Wheelchair/Scooter**: Main form of mobility
8	**Bedridden**: Unable to sit in wheelchair for 1 h

This scale focuses mainly on how well you walk measures analysis. Latent Growth Modeling (LGM) provides the benefits of representing the disease progression in multiple periods and can adjust to the unique differences of the individuals in the data set (Serva et al. 2011). Certain demographic, life style, treatment and fixed effects can then act as factors that indicate a modification of the disease starting point (intercept) or progression over time (slope) in the model. The advantages to LGM include greater flexibility in specification, explicitness of all aspects of the model and the use of latent variable score estimates (Cohen et al. 2003).

This study parallels prior research where tailored interventions motivated enhancement of health-related behavior change (Liang et al. 2006). The primary context presented here is in the investigation of MS and applies to other long-term illnesses. The modeling of two base LGM models provides a pathway for planning life style modifications. In the first model, the basic disease progression marker, PDDS, is used as part of a latent variable termed Overall Health (OH). The second base model will evaluate another latent variable that represents the Emotional Health (EH) of patients.[2] Each of these base models can identify patient differences in fixed effects (age, sex, and race), life style (employment, income level) and treatment (on drug therapy). Together, covariates affecting the physical and emotional health of the patients provide a focus for health promotion.

3.3 Application of LGM for MS

Latent Growth Modeling is a technique used in longitudinal experiments when assessing specific conditions and factors that change those conditions over time (Serva et al. 2011). The use of LGM requires a dataset that has been collected consistently over the course of multiple assessments. These assessment periods can be over the course of a few months to multi-year assessments that span decades. The innovative use of LGM in the context of MS and the patient reported NARCOMS dataset provides a contribution of how certain self-reported factors are related to the MS progression. Patients and physicians can benefit from a view of how different factors effected the disease progression over time.

Some examples of LGM application include a change in alcohol use that was assessed every 6 months for participants that were randomly assigned to three separate relationship treatment programs (Longabaugh et al. 1995). The results indicated that the extended relationship treatment needed to be calibrated against a participant's relative levels of investment in supportive and unsupportive relationships. Adolescent substance use has been assessed yearly for 7 years as an associative model of participant and participant peer groups (Wills et al. 2001). Conclusions of the study indicated that coping dimension correlations (behavioral,

[2] See Appendix for the questions used to construct the Overall Health and Emotional Health latent Variables.

cognitive and avoidant) needed to be understood in order to predict adolescent substance abuse. Rates of autism identification in schools have been tracked over a 7-year time period based on the district's available resources (Palmer et al. 2005). Additional resources were identified as necessary to provide an increasing ability of districts to identify children with autistic spectrum disorders. Psychologists have tracked the progression of adolescent Antisocial Behavior through the middle school years (4th thru 8th), assessed the covariates of Monitoring, Discipline, Wandering Change and Peer Change and their effect on the children as they progressed through their school years (Patterson 1993). The intent of this study was to define the development of antisocial traits and resulted in the identification of an orderly progression of individuals in their development of anti-social behavior.

With respect to the application of LGM to MS patients, Latent Growth Models hypothesize a relationship across time that is represented by a linear equation (Fig. 3.1). The following assumptions are the basis for testing the base model that includes the latent variables of interest. Time 1 through Time 3 of OH and EH, represents the measurement points, and subsequently the latent variables. Respondents must participate in each of the data collection periods. The linear equation hypothesized is assumed to have the same starting point for the disease (Intercept) in the initial period of the survey. This starting point (Intercept) is represented in the model by the coefficient of 1 that is attached to each path going from the Intercept factor to each survey time period (Time 1, 2 and 3). The model's linear representation across the time periods, indicating the disease progression as either better or worse is referred to as the Slope. The initial assumption is that the disease progresses in a linear fashion over time, as opposed to a quadratic (escalating) fashion. This assumption is represented by the path coefficients of 0, 1 and 2 leading from the Slope to each time period. The model specified here follows the PDDS question and OH latent variable that there is an increasing slope (greater number of disease symptoms) that would be tested by the model. Base models have no Covariate included and represent the starting point at which covariate factors modify disease progression. When the Covariates are added to the model, the path coefficients leading from the Covariate would represent either positive or negative directional changes to the starting point of the disease (Intercept) and how rapidly the disease progresses (Slope). Model fit statistics provide a measure of how good the assumptions are represented by the LGM.

NARCOMS collects data on MS patients twice per year on a "spring" and "fall" schedule (Vollmer et al. 1999). The dataset provided by NARCOMS included 3774 patient respondents that completed the questionnaire across six survey periods and have followed the collection schedule in separating the data: (1) Fall 2004 (F04), (2) Spring 2005 (S05), (3) Fall 2005 (F05), (4) Spring 2006 (S06), (5) Fall 2006 (F06), and (6) Spring 2007 (S07). The population demographics are presented in Table 3.2 and parallel other studies on MS (Vollmer et al. 1999; Marrie et al. 2005). Note that the demographic covariates have been modified to show the study splits based on the PDDS disease stage determination. On average, the patient respondents have been living with MS for 15 years since diagnosis and 23 years since the onset of

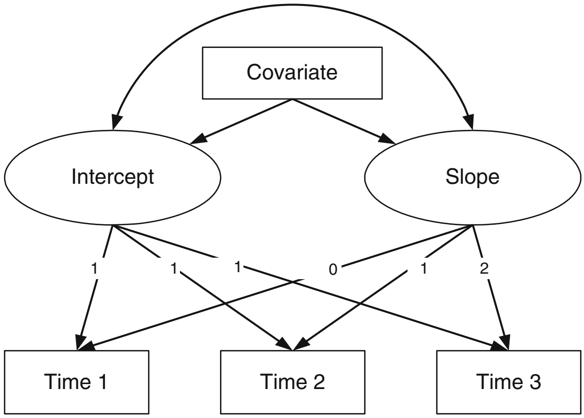


Fig. 3.1 Latent growth model

Table 3.2 MS patient demographics

N	% Female	% White	Employed	Health Ins
3767	74.9 %	93.6 %	~35 % F05	~98 %
			~35 % S06	
	Age	@ Diag.	@ Onset	
Average	55.21	38.75	30.69	
N	3767	3623[a]	3509[a]	
Min	20	12	2	
Max	90	74	65	
	Early PDDS	Mid PDDS	Late PDDS	Drug therapy
Fall 05	~32 %1199	~40 %1519	~28 %1049	~35 %
Spring 06	~31 %1169	~40 %1520	~29 %1078	~35 %
Income	Fall 05		Spring 06	
1	432	~11 %	424	~11 %
2	654	~17 %	618	~16 %
3	762	~20 %	767	~20 %
4	1130	~30 %	1144	~30 %
5	528	~14 %	553	~15 %
6—Did not answer	261[a]	~7 %	261[a]	~7 %

[a]Reduced N is from non-respondents to these questions. Those respondents were removed from the associated models

symptoms. Females in the population outnumber males three to one. MS seems to be racially discriminant with 93.6 % of the study participants reporting their race as white.

The timeline for disease progression (Slope) is known to be different for early vs. later stages of MS from both prior research (Achiron et al. 2003) and the collaborating physicians for this research. As an exploratory aspect of this analysis, one of the goals is to investigate differences in disease progression (Slope) indicative of the segregated stages. Table 3.2 provides a division of the population that may be in either an early, middle or late stage of MS. The early stage is defined as anyone on the PDDS scale of 0, 1 or 2. The middle stage is defined as answers of 3, 4 or 5 and the late stage consists of answers of 6, 7 or 8. This division was based on the respondent answers to the PDDS question for the survey administered in Fall of 2005 and, for the confirmatory analysis, Spring of 2006. The resulting three groups represent different sample sizes of the original 3774 respondents. The three PDDS groups partitioned on the Fall of 2005 and Spring of 2006 response total 3767 respondents. Seven respondents had missing data for part of the LGMs tested. The data for these respondents was removed from the analysis. Approximately 35 % of the respondents are on some form of drug therapy. In addition, 35 % of the population is employed with 98 % of the population having some form of Health Insurance. The ∼symbol indicates that these percentages are approximated across the six survey timelines since the PDDS answers, as well as drug therapy, employment and health insurance, can change from survey to survey.

The latent variables, OH and EH, each include four scale items. In addition to the 9-item PDDS scale, the remaining three scale items included were labeled *mobility* (*Mobility*), *moderate activity* (*ModAct*) and *climbing stairs* (*ClimbStrs*). *Mobility* is a 7-point scale ranging from 0—Normal to 6—Bedridden. *Moderate activity* and *climbing stairs* are both a 3-point scale ranging from 1—limited a lot to 3—no limitations. *PDDS* and *Mobility* are reverse coded, compared with *Mod Act* and *Climb Strs*, with higher numbered responses indicating disease progression or negative effects. Detailed wording of these questions have been included in the Appendix.

The EH latent variable included: *Accomplished Emotional* (*AccEmot*), *Careful Emotional* (*C Emot*), *Calm* (*Calm*) and *Depression* (*Depress*). All of the EH items are on a scale of 1–5 with 1 indicating "All of the time" and 5 indicating "None of the time". *AccEmot* asks if the patient "accomplished less than they would have liked" and *C Emot* asked if the patient "Didn't do work as carefully as usual". The *Calm* scale item asked if the patient felt "calm and peaceful" while the *Depress* item asked, "Have you felt downhearted and depressed". (See Appendix for exact wording.) *Calm* is the only scale item that is reverse coded where higher numbered responses indicate negative effects or conditions.

3.3.1 Base LGM Model

To validate the increasing progression between the three groups and set a base model for assessing the covariates, two LGMs were modeled to assess these changes. The first LGM consisted of the OH latent variables from the surveys of S05, S06 and S07. The same surveys were used to run the LGM for EH latent variables. The presentation of the fit statistics for the base models can be found in Table 3.3. The OH base models both indicate a more rapid increase in disease progression in-group 3 when compared to group 1. The EH base model provides a more stable picture of the patients (Fig. 3.2). The base models provide a baseline comparison when different covariates are included in the LGM.

Table 3.3 Base model results

LGM	Type	Group	Split	N	df	Chi^2	RMSEA	NFI	NNFI	CFI	SRMR	GFI
Base	Overall health	1	Fall	1192	33	155	0.056	0.99	0.99	1.00	0.027	0.98
			Spring	1169	33	144	0.054	0.99	0.99	1.00	0.027	0.98
		2	Fall	1512	33	260	0.067	0.99	0.98	0.99	0.061	0.97
			Spring	1520	33	269	0.069	0.99	0.98	0.99	0.057	0.97
		3	Fall	1036	33	110	0.047	0.99	0.98	0.99	0.047	0.98
			Spring	1078	33	122	0.050	0.99	0.98	0.99	0.008	0.98
	Emotional health	1	Fall	1192	33	110	0.044	0.99	0.99	1.00	0.029	0.98
			Spring	1169	33	104	0.043	0.99	0.99	1.00	0.032	0.99
		2	Fall	1512	33	283	0.055	0.99	0.99	0.99	0.032	0.98
			Spring	1520	33	165	0.051	0.99	0.99	0.99	0.030	0.98
		3	Fall	1036	33	252	0.080	0.98	0.97	0.99	0.047	0.96
			Spring	1078	33	246	0.077	0.98	0.97	0.99	0.044	0.96

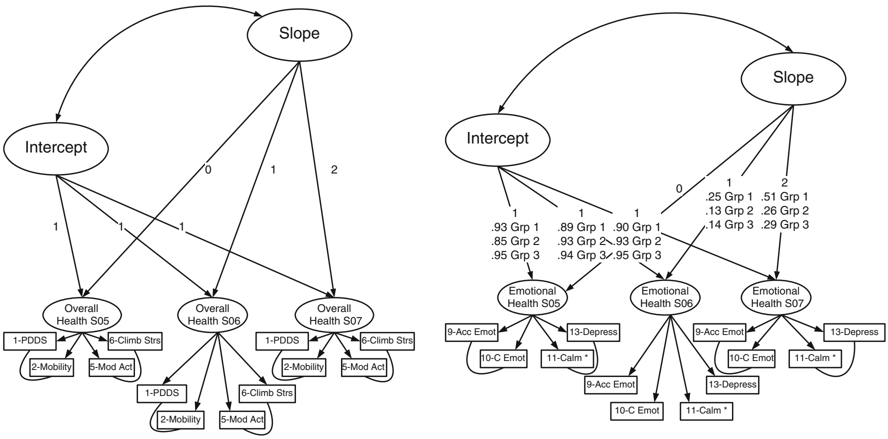


Fig. 3.2 LGM base models

Examples of the covariates for this study include age, employment status and income. The inclusion of independent variables, and their effect on the dependent variable, is only assessed after the relationship between the covariates and dependent variables are removed (Tabachnick and Fidell 2001). Covariates are used in LGM to determine modifications to either intercept or slope. Figure 3.3 demonstrates the anticipated modifications to the linear progression of MS based on a single covariate. The decreasing line labeled OH base represents the OH Base model discussed previously. All intercept and slope modifications due to the covariate are premised from this base line. A covariate that adjusts the base model line may have the effect statistically of modifying the slope or disease development of OH either positively or negatively. The covariate may also be associated with a change in intercept, or starting point, either positively or negatively. Note that these associations do not necessarily imply causation. The hypothesized representational changes is shown in

Fig. 3.3 Covariate modifications

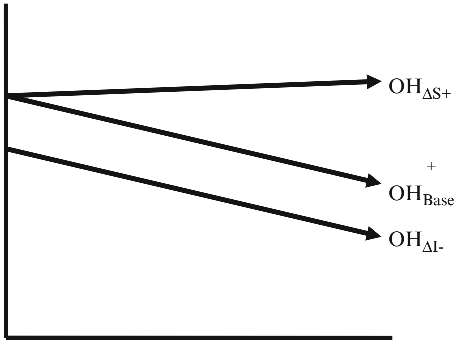


Fig. 3.3. In this Figure, a negative intercept change, line $OH_{\Delta I-}$, from the base OH linear disease progression line (OH_{Base}), represents a lower OH starting point with no slope change or rate of disease progression. A positive slope change, or reduced rate of OH disease progression, is represented by the line $OH_{\Delta S+}$.

When assessing the effects of covariates on MS disease progression, a consistent modification of either the disease starting level or disease progression advancement is searched for across the three groups of MS assessment respondents. Tables have been constructed to assist in the identification of the consistency of covariate-associated adjustments to the base model characteristics. Tables 3.4 and 3.5 are constructed to present the results of covariates on the EH and OH of each group. The consistency of the results is determined when a single covariate has the same effect across all three groups for either OH or EH. These effects may be either a significant positive/negative slope change or a significant positive/negative intercept change.

3.3.2 Age Effects of MS

The expectations of disease progression may vary based on several views of a patient's age. Since MS is not a curable disease, part of the goal is to provide MS patients with a life style that is as close to normal as possible. Three views of the patient's age are used as covariates: Age, Age @ Diagnosis and Age @ Onset. For these LGMs, Age is simply defined as the patient's birth age. The Age @ Diagnosis is the age at which a medical professional has confirmed that the patient does in fact have MS. Once this confirmation has occurred, life prior to the Age @ Diagnosis is reviewed. This review allows the patient to determine at which point in their life they can recognize the onset of symptoms of MS. This will then determine the covariate

Table 3.4 Age covariate effects

			Age		Age @ Onset		Age @ Diagnosis	
Type	Group	Split	Intercept adjust	Slope adjust	Intercept adjust	Slope adjust	Intercept adjust	Slope adjust
Overall health	1	Fall	Positive***	Positive*	No Effect	Negative*	Positive***	No Effect
		Spring	Positive***	Positive*	No Effect	Negative*	Positive***	No Effect
	2	Fall	Positive***	Positive*	No Effect	No Effect	Positive**	No Effect
		Spring	Positive***	Positive*	No Effect	No Effect	Positive**	No Effect
	3	Fall	Positive***	No Effect	No Effect	No Effect	No Effect	No Effect
		Spring	Positive**	No Effect	No Effect	No Effect	No Effect	No Effect
Emotional health	1	Fall	Positive**	No Effect	Positive*	Negative*	No Effect	No Effect
		Spring	Negative*	No Effect	Positive*	No Effect	No Effect	No Effect
	2	Fall	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
		Spring	No Effect	Positive*	Positive*	No Effect	No Effect	No Effect
	3	Fall	DNC	DNC	No Effect	No Effect	No Effect	No Effect
		Spring	No Effect	No Effect	No Effect	No Effect	DNC	DNC

Significance levels - * $-$ $p < .10$ ** $-$ $p < .05$ *** $-$ $p < .01$

Table 3.5 Employment/income and drug therapy covariates

		Split	Employment status		Income		Drug therapy	
Type	Group	Intercept adjust	Slope adjust	Intercept adjust	Slope adjust	Intercept adjust	Slope adjust	
Overall health	1	Fall	Pos***	No Effect	Neg***	No Effect	No Effect	Neg**
		Spring	Neg***	No Effect	Pos***	No Effect	No Effect	No Effect
	2	Fall	Pos***	No Effect	Neg***	No Effect	Neg***	No Effect
		Spring	Neg***	Pos***	No Effect	Neg*	Pos***	No Effect
	3	Fall	Pos***	Neg**	No Effect	No Effect	Neg***	Pos*
		Spring	Neg**	Neg*	No Effect	No Effect	Pos***	No Effect
Emotional health	1	Fall	Neg***	No Effect	Pos***	No Effect	Pos**	No Effect
		Spring	Neg***	No Effect	Pos***	No Effect	Pos***	No Effect
	2	Fall	Neg***	No Effect	Pos***	Neg*	Neg*	No Effect
		Spring	Neg***	No Effect	Pos***	Neg*	Neg**	Pos**
	3	Fall	Neg***	No Effect	Pos***	No Effect	Pos**	No Effect
		Spring	Neg***	No Effect	Pos***	No Effect	Pos***	No Effect

Significance levels - * — $p < .10$ ** — $p < .05$ *** — $p < .01$

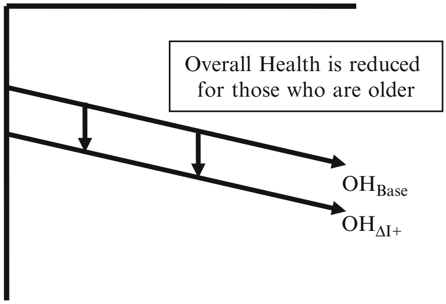


Fig. 3.4 Age intercept modification

referred to as the Age @ Onset. Each of these covariates is used to assess their adjustment on the base LGM models.[3]

An analysis of these covariates for all three groups using sequential Fall and Spring responses is summarized in Table 3.4. The initial covariate (Age), across all six models, indicates that, Age is associated with OH decrease as depicted by a positive increase in the Intercepts. This covariate is the only one that indicates a consistent directional adjustment across all six models for both the Fall and Spring groups. A graphical representation of this modification is provided in Fig. 3.4. The change over time or slope across the Fall and Spring split shows a positive increase only for Groups 1 and 2, that is, for the MS patients with the highest degree of mobility. This positive increase for "younger" patients indicates that the disease progression is slower for younger patients and more rapid for older patients.

For the Age @ Diagnosis, the individuals for Group 1 and 2 that were diagnosed later in life exhibit a more advanced disease starting point. The starting point for OH of these older patients does not indicate any change in the rate of their disease progression. The only significant covariate for Age @ Onset was with Group 1, which indicated a reduction in the disease progression rate for the most capably active Group 1 patients. In other words, the older the patients are before "getting" the disease symptoms, the slower the disease progresses through the rest of their life.

With the EH assessment, positive increases in slope and intercept would indicate an improvement in the EH of the patient. None of the Age associated covariates exhibited consistent results across the three groups or from the Spring and Fall split of the respondents. Group 1 had the most significant results across the models, with three models showing a positive EH adjustment for the intercept of Age @ Onset. These positive adjustments were for group 1, both spring and fall splits, and the spring split in group 2. This modification would indicate that the EH for older patients was better than for younger patients in earlier stages of the disease. Two models of all of the Age covariates did not converge (DNC). These two models were

[3]Only the Slope and Intercept changes, along with their significance, have been provided due to length considerations. Fit statistics for the Age Effect models and the remainder of the models outlined can be requested from the authors.

the group 3 Spring Age @ Diagnosis model and the group 3 Fall Age @ Diagnosis model. In most of the models, the three Age covariates did not show a significant change in either the Slope or Intercept which indicates that the EH of an MS patient is generally not affected by age.

3.3.3 Life Style and MS

The early stages of MS are noted for how a patient is willing to continue to be active and productive in their lives. There are two data points collected through the NARCOMS survey that best indicates the activity and productivity of MS patients. For these two LGMs, one covariate is provided through the patients Income level, where increasing value of this categorical variable indicates higher levels of income, and the other covariate, Employment Status, is operationalized as a dichotomous variable of either Yes (0) or No (1).[4]

Neither the Employment Status nor the Yearly Income level indicated consistency across the three groups for changes in the Intercept or Slope (Table 3.5). There was a consistent adjustment to Slope between the fall and spring split in group 3 for the Employment Status covariate. The adjustment to the Slope was negative which would indicate that those employed individuals in advanced stages of the disease progress less rapidly through to the final stages.

The EHof all three groups in both the Spring and Fall patient splits demonstrated a negative Intercept adjustment across all six models. There was consistently no effect on Slope for all models for the Employment Status covariate. This negative adjustment for the Intercept in EH indicates that those who were not employed experienced lower EH than those that were employed (Fig. 3.5). There was also a consistently positive EH improvement for patients with increasing levels of income

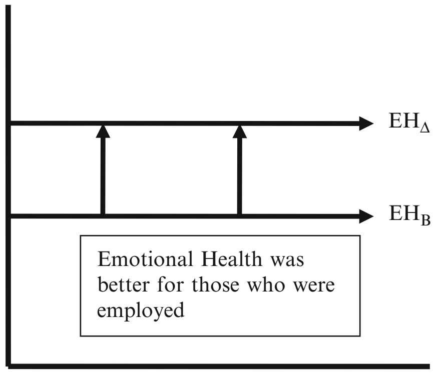


Fig. 3.5 Emotional health income/employed covariate

[4]The fit statistics for the models of these Life Style effects can be requested from the authors.

across all six models. The EH Slope adjustment was inconsistent across the three groups, but was consistent for the Fall and Spring splits between the groups. Group 1 and 3 were consistent in that No Effect was detected for the Income Covariate Slope adjustment. Group 2 had a negative Slope adjustment for the Income Covariate.

3.3.4 Drug Therapy Effectiveness

Whether a patient enters into a drug-treatment therapy is a decision made under the consultation of a physician. There are multiple categories of drug therapies available. However, the ability to model these differences is difficult due to the multitude of variables for this category. As a starting point, data utilized for this investigation was a dichotomous variable indicating only whether the patient was participating in a drug treatment therapy or not. Several factors that justify this methodology were gleaned from the researchers interaction with practicing physicians associated with NARCOMS.

Generally, when a MS patient begins a drug treatment therapy, they continue for extended periods. The data was checked to ensure that patients on drug therapy continued for the portion of the study in question. The patient data used for these LGMs were checked to ensure that patients included as on Drug Therapy for the target survey of spring 2006 were also on drug therapy for all other survey time periods.

The significance of this covariate (Table 3.5) was inconsistent across the groups for OH. Even those models that indicated a significant relationship were not consistent in the direction between the Spring and Fall models. This would indicate that whether the patients are, or are not, on some form of drug therapy, their OH did not degrade or improve due to these therapies. This surprising result prompted the researchers to question the physician contacts as to why drug therapies did not improve the OH of the patients. It was noted that switching types of drugs during the study period was common which could confound the results. However, the physician collaborators indicated that the drug treatments available during the analysis period are equivalent in efficacy and therefore this result was not surprising.

Being on Drug Therapy did not modify the EH stability in a consistent manner. However, there was some within-group consistency for the Spring and Fall splits with respect to the Intercept or starting points of EH. For Group 1 and 3, being on drug therapy was positively related with improved EH. Group 2, in contrast, showed a negative relationship in EH when they were on drug therapy. Five of the six models indicated that there was no effect in Slope for those patients on some form of Drug Therapy.

Several observations and guidelines can be drawn from this analysis and are summarized below.

- Each of the groups studied indicate a different linear trajectory for the progression of MS

- Those patients diagnosed with MS later in life experience better EH and a slower OH decline
- Patients that stay employed and procure higher incomes will experience better EH and a slower OH decline
- The participation in a drug therapy treatment will have mixed results

3.4 Implications for the Health Care Context

The application of LGM to the NARCOMS dataset brings out the importance of this technique, not only for MS, but also for multiple health care contexts. This initial analysis of MS growth factors highlight individual patient differences that are controllable (job status and drug therapy) and are not controllable (age). Each of the covariates attempted to explain individual differences within the three groups of patients analyzed. Confirmation of these changes was sought through the patient split to the Fall and Spring groups. Specific improvements in OH and EH could be anticipated with the patient's determination of staying employed or determining not to seek some form of drug therapy. These results were reached with an investigation using three data points across a 3-year time span of assessment and making a basic assumption of linearity in the dataset. The possible insights to be gained of the NARCOMS dataset, and other health care datasets, can be extended with more complex specifically targeted LGM models.

The assumption of linearity requires at least three data points for analysis (Serva et al. 2011). The Base models indicated a different slope for each of the three groups and could indicate that grouping all three study groups would require more data points in order to accurately represent the MS disease progression over time. LGM allows the ability of a nonlinear slope function to be determined by the dataset under investigation. A simple model determining the slope would allow the calculation of additional time-period slope functions. Any result indicating a non-linear function would then more accurately be termed a shape function (Duncan and Duncan 2004). The determination of the true shape of the curve may require more data points for this unspecified growth function. Guidance on the number of data points required is available from several prior studies (Tisak and Meredith 1990; Burchinal and Appelbaum 1991). An acceleration of growth, or MS progression, over time can be specified as either quadratic, cubic or left unspecified. A future study could assess a patient's disease progression in order to identify certain "triggers" that hastens the advancement of the disease.

The change over time or shape of the curve, of any disease may vary between two multiple consecutive assessment periods. A Piece-wise LGM may be applied to determine a linear growth that exists in the first assessment period (three data points) while a quadratic growth may exist in the second subsequent assessment period (three data points) (Bryk and Raudenbush 1992). These periods may be separated by age, treatment or some other distinguishable variable. LGM also allows an Accelerated Design model in which cohorts can be sequentially modeled in order to form a common developmental trajectory of growth. This "cohort-sequential" design (Nesselroade and Baltes 1979), in the case of MS and the NARCOMS dataset, can allow the three groups (early, middle and late disease progression) to be sequentially modeled to provide a clearer interpretation of the rapid advance of the late stages of the disease. This design is referred to as accelerated since the data from the different cohorts was collected during a shortened or identical period.

The inclusion of additional covariates and application of models with greater complexity can provide even greater insights into the shape of MS progression and how predictors can modify that shape. The importance of treatment and life style interventions that improve the overall and emotional health of MS patients, as identified through the various LGM techniques, can allow patients to enact and their physicians to recommend positive life style changes.

3.5 Challenges to LGM Usage

Several challenges to future research using LGM present themselves to the health-care community. These challenges include patient privacy, patient mobility and maintenance of technology. The use of LGM for MS patients is a multi-year task and the NARCOMS dataset has been one of the most consistently gathered patient reported studies. As indicated, MS patients can make choices that affect the progression of the disease. Additional patient behavioral choices or personal environmental conditions may exist that retard the disease progression that NARCOMS does not currently collect. Consistent acquisition of this information may be difficult due to perceptions of patient privacy. Collecting this information over an extended period may also be problematic.

Patient mobility to a different geographical location needs to be accomplished without loss of historical information. The evidence already indicates that continued employment is beneficial to retarding the disease progression. Was potentially illuminating data lost when individual patients relocated to new geographical locations due to job movement? Were these relocations due to the patient's or their spouse's employment opportunities? The NARCOMS dataset was predominantly focused in the southwest United States. An expansion of this study would include the tracking of individual MS patients nationally to further expand this current finding and allow a geographical perspective on MS progression.

Two technological challenges exist in future research applications of LGM for the study of MS patients. The initial challenge is the merging of clinical data with patient survey data. Clinical data was not available for any of this research due to the difficulties in merging these two data sources. The difficulties exist procedurally, through concerns of patient privacy, and technologically, through the collection and storage of the information on separate systems. The second challenge, especially with MS, is the technological changes over time. NARCOMS has been conducting the survey for this study over a much longer time than the 3-year study period. However, the movement of records to a new system prevented the acquisition of all available records. An enterprise approach to all patient historical records is a major challenge for healthcare providers, but one that can pay enormous benefits to improving patient and physician evidence based decisions.

3.6 Conclusions

The analysis of the NARCOMS dataset for MS patients included three cohorts at different disease stages and four covariates in the investigation of the linear progression of MS. Each covariate provided an indication of some linear modification of either Overall or Emotional Health of the patient and resulted in the observations and guidelines provided at the end of Sect. 3. The extended nature of the disease, an individual's lifetime, indicates that a single LGM study over a 3-year assessment period is small compared to the length of time individuals live with the disease. Even in the abbreviated time of this study, LGM has provided some evidence that controllable influences in an individual's life can restrain progression of the disease and improve the overall health of patients. Though a control group of the normal population was not analyzed, extending this study to assess the overall health differences between MS patients and a normal population could bring greater understanding of the potential intervention effects for MS patients. This assessment would need to track comparable data from a control group of the normal population that also includes the covariates identified through the NARCOMS survey. The "normal" populations would answer their PDDS questions for comparisons to the MS population.

Achieving these goals does require a long-term view for the collection of data. NARCOMS has a dataset much larger than was currently available for study. Recent efforts are underway to consolidate all of the NARCOMS data in a central repository in order to mine their data for up to a 10-year period. Improvements in the ability to assess additional covariates would require the inclusion of specific treatments. These treatments could be physical therapy, drug therapy or social interventions. Data cleansing techniques should be used when scales have been adjusted for survey questions. Finally, care must be taken to ensure the privacy of all participants. Yet, for LGM to be useful to health care research, each assessment must continue to capture and link the clinical and self-reported patient data in order to shape the progression of any long-term disease and assess effects of predictive variables.

Appendix: Survey Scale Items

Overall Health Latent Variable[5]—PDDS, Mobility, Moderate Activities, Climbing Stairs

Mobility*: Compare your current condition to your mobility before you developed MS.

Value	Meaning	
0	**Normal**:	Functionally normal walking and running
1	**MinimalGaitDisability**:	Minor but noticeable effects on mobility
2	**MildGaitDisability**:	Noticeable effects on mobility
3	**OccasionalUseofCane**:	May use cane for greater distances
4	**FrequentUseofCane**:	Unable to walk 25 ft without cane support
5	Severe **GaitDisability**:	Requires bilateral support to walk 25 ft
6	**TotalGaitDisability**:	Essentially confined to a wheelchair

Does your health now limit you in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No not at all
Moderate Activities: such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
Climbing several flights of **stairs**	1	2	3

Emotional Health Latent Variable—Accomplished Emotional, Careful Emotional, Calm, Depressed

During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems?

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
Acc Emot: Accomplished less than you would like	1	2	3	4	5
C Emot: Didn't do work or activities as carefully as usual	1	2	3	4	5

[5]See Table 3.1 for PDDS Question.

How much of the time during the past 4 weeks . . .

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
Calm: Have you felt calm and peaceful?	1	2	3	4	5
Depress[a]: Have you felt downhearted and depressed?	1	2	3	4	5

[a]*Note*: reverse coding

References

Achiron A, Barak Y, Rotstein Z (2003) Longitudinal disability curves for predicting the course of relapsing—remitting multiple sclerosis. Mult Scler 9:486–491

Berndt DJ, Hevner AR, Studnicki J (2003) The Catch data warehouse: support for community health care decision-making. Decis Support Syst 35:367–384

Bryk A, Raudenbush S (1992) Hierarchical linear models in social and behavioral research: applications and data analysis methods. Sage, Newbury Park

Burchinal M, Appelbaum MI (1991) Estimating individual developmental functions: methods and their assumptions. Child Dev 62:23–43

Cohen J, Cohen P, West S, Aiken L (2003) Applied multiple regression/correlation analysis for the behavioral sciences, 3rd edn. Lawrence Erlbaum, Mahwah

Duncan TE, Duncan SC (2004) An introduction to latent growth curve modeling. Behav Ther 35:333–363

Freeman JA (1999) Rehabilitation in multiple sclerosis: outcome selection in clinical trials. Int J MS Care 2:20–26

Hamalainen P, Seinela A, Ruutiainen J (1999) Adaptational approach to cognitive rehabilitation in multiple sclerosis: description of three models of care. Int J MS Care 2:7–12

Kelley H, Chiasson M, Downey A, Pacaud D (2011) The clinical impact of eHealth on the self-management of diabetes: a double adoption perspective. J Assoc Inform Syst 12(Special Issue):208–234

Liang H, Xue Y, Berger BA (2006) Web-based intervention support system for health promotion. Decis Support Syst 42:435–449

Longabaugh R, Wirtz PW, Beattie MC, Noel N, Stout R (1995) Matching treatment focus to patient social investment and support: 18-month follow-up results. J Consult Clin Psychol 63:296–307

Marrie RA, Cutter G, Tyry T, Hadjimichael O, Vollmer T (2005) Validation of the NARCOMS registry: pain assessment. Mult Scler 11:338–342

Nesselroade JR, Baltes PB (1979) Longitudinal research in the study of behavior and development. Academic, New York

Palmer RF, Blanchard S, Jean CR, Mandell DS (2005) School district resources and identification of children with autistic disorder. Am J Public Health 95:125–130

Pedersen MK, Larsen MH (2001) Distributed knowledge management based on product state models—the case of decision support in health care administration. Decis Support Syst 31:139–158

Patterson GR (1993) Orderly change in a stable world: the antisocial trait as a chimera. J Consult Clin Psychol 61:911–919

Porter ME, Teisberg EO (2006) Redefining health care: creating value-based competition on results. Harvard Business School, Boston

Serva M, Kher H, Laurenceau J-P (2011) Using latent growth modeling to understand longitudinal effects in MIS theory: a primer. Commun Assoc Inform Syst 28(Article 14):213–232

Song J, Zahedi FM (2007) Trust in health infomediaries. Decis Support Syst 43:390–407

Tabachnick BG, Fidell LS (2001) Using multivariate statistics, 4th edn. Allyn & Bacon, Needham Heights

Tisak J, Meredith W (1990) Descriptive and associative developmental models. In: von Eye A (ed) Statistical methods in developmental research, vol 2. Academic, San Diego, pp 387–406

Vollmer T, Ni W, Stanton S, Hadjimichael O (1999) The NARCOMS patient registry: a resource for investigators. Int J MS Care 1:12–15

Wills TA, Sandy JM, Yaeger AM, Cleary SD, Shinar O (2001) Coping dimensions, life stress, and adolescent substance use: a latent growth analysis. J Abnormal Psychol 110:309–323

Chapter 4
Designing an Integrated Surgical Care Delivery System Using Axiomatic Design and Petri Net Modeling

Joshua Bosire[1], Shengyong Wang[2], Mohammad Khasawneh[3], Tejas Gandhi[4], and Krishnaswami Srihari[3]

[1]Cooper University Hospital, Camden, New Jersey, USA

[2]Department of Mechanical Engineering, University of Akron, Akron, OH, USA

[3]Department of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, NY, USA

[4]Navicent Health, Macon, Georgia, USA

Abstract Over the last decade, a wide range of philosophies and management strategies have been embraced by an increasing number of healthcare organizations, with the principal goal of improving productivity and the quality of care while reducing the associated costs. These have had a significant impact on various healthcare services, including the delivery of surgical care. These philosophies and strategies are composed of scientific techniques and methods, a majority of which can be classified under the Industrial and Systems Engineering (ISE) field of study. For the most part, these techniques have proved to be instrumental in engineering radical changes and improvements in the ailing healthcare industry. This research seeks to contribute to the growing efforts on the application of ISE tools in the healthcare industry. First, an attempt is made to illustrate how various techniques can be used to engineer tangible benefits into the processes for surgical care delivery. Axiomatic Design (AD) concepts are leveraged to synthesize redesign changes, which are then applied to a conceptual model of the surgical system. The model is developed using a Petri Net (PN) technique and subsequently validated, verified and analyzed using PN heuristics and Dependency Structure Matrix (DSM). The outcomes of this redesign depict process improvements with the potential to enhance the quality of surgical care while enabling close to 70 % productivity gains. As a result, this research proposes a framework for the systematic redesign of processes, especially when changes need to be tested and verified before

S. Wang (✉)
Department of Mechanical Engineering, University of Akron, Akron, OH, USA
e-mail: wangs@uakron.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_4

implementation. The techniques identified within this framework are ideal for the successful completion of the '*design*' and '*plan*' phases when pursuing operations management philosophies like Six-Sigma (SS) and Total Quality Management (TQM).

Keywords Surgical system • System redesign • Axiomatic design • Petri net • Integrated care delivery

4.1 Introduction

Surgical care forms a significant proportion of the healthcare industry in the US, with millions of Americans opting for surgical interventions every year. Surgical services are offered in both the inpatient and outpatient (ambulatory) settings, with recent trends indicating an inclination toward the latter. In either case, the surgical procedures are performed in operating rooms suites, which could be either stand-alone or part of other hospital facilities. As part of the hospital, surgical services are estimated to account for up to two thirds of the total earnings (Pham and Klinkert 2008). Operating rooms have been established to be one of the leading avenues of expenditure in many hospitals due to the costs of facilities, resources and personnel (Pham and Klinkert 2008; Dexter et al. 1999; Nivarthi 2004).

Surgical service is a fully interdependent system of multi-level healthcare systems, where each level is operationally dependent on the others, often tightly coupled, with information distributed across people, units, services, and devices. Modern surgeons, proceduralists (interventionists) and anesthesiologists are among the most highly trained professionals and technicians in any industry; the technology and resources they employ are often extraordinarily advanced, expensive, and relatively scarce. Optimal utilization of these limited human, technological, and financial resources is of paramount importance to a safe, effective, timely and efficient U.S. health care delivery system.

In the face of diminishing reimbursements along with the escalating costs of providing health services, healthcare organizations have embarked on pursuing philosophies of continuous improvement with the hope that these will reverse the current predicament. The realization that the US healthcare system is such a complex and fragmented system has led to the increased adoption of systems engineering approaches as an enabler in the pursuit for operational excellence. This is in line with the observation by the IOM (Institute of Medicine 2001) that the US healthcare system (both the structure and culture) is in dire need for improvements appertaining to the six dimensions of healthcare quality.

The IOM (Institute of Medicine 2001) further observes that the ineffective structure of the US health system can be attributed as being the root-cause of the observed inherent wastage of resources. For instance, redundancies and duplicities in care delivery processes along with the protracted waits for services are some of the common wastes observed. Narrowing down to surgical care, the

Healthcare Financial Management Association (Healthcare Financial Management 2003) observes that surgical care delivery in many hospitals lacks even the basic structures for obtaining standardized and comprehensive data on the operations of their Operating Rooms (ORs). Since data forms the basis for most continuous improvement techniques, it follows that surgical care delivery systems need an integrated approach that facilitates the capture and access to critical information regarding their performance.

With a similar flow of thought, Ransom et al. (2005) reemphasize the need for healthcare organizations to design healthcare delivery systems focused on enhancing the quality of outcomes. Besides the culture of healthcare practitioners and staff, Ransom et al. (2005) also note that the associated systems and structures must be redesigned, with focus being placed on aspects including the following.

1. Designing systems by anticipating and circumventing failures and errors,
2. Streamlining processes not only to eliminate wastes, but also to simplify them in the bid to eliminate potential causes of process failures,
3. Standardize processes, tools, technology and equipment as a means to minimize variations and complexity to result in quality and productivity gains,
4. Adopting workplace standardization techniques and philosophies, and
5. Embrace technologies that automate critical but manual activities (e.g., automated data entry procedures), and those that induce ‘forced functions’ to circumvent process failures and error generation.

This research was initiated with the premise that the productivity and quality of surgical care can be enhanced by integrating the flow of information in a manner that directly supports the coordination of resources and execution of surgical processes. The surgical system of a multi-facility hospital was studied to demonstrate the impacts of integrated surgical delivery system redesign. The goal of these efforts was to develop a conceptual framework to illustrate how tasks, resources and information can be integrated to drive operational efficiency in a surgical system. The specific research objectives are as follows.

1. Incorporate concepts of business process improvement to facilitate the redesign of an existing surgical care delivery system,
2. Use a systems design approach to develop process definitions that support the integrated flow of information through the surgical system,
3. Show benefits of the redesigned process definitions by using analytical techniques to evaluate the surgical system models; both qualitatively and quantitatively, and
4. Demonstrate the application and suitability of Business Process Management (BPM) concepts in the analysis and redesign of healthcare delivery processes.

To achieve these research objectives, the Business Process Management (BPM) framework (zur Muehlen 2004; Lusk et al. 2005; Leymann et al. 2002; Shaw et al. 2007; Miers 2005; Smith and Fingar 2003) was elected for the redesign of a surgical system. This choice was guided by the observation that BPM is associated with a wide range of techniques that enable the analysis of systems, generation of redesign ideas, evaluation and consequent enactment of design changes. Uniquely, this endeavor adopts Axiomatic Design (AD) (Suh 1990) into the BPM framework to facilitate the effective realization of the surgical system redesign. Among other benefits, AD extends the capabilities of design techniques like Quality Functional Deployment (QFD) by providing a systematic and formal approach for generating and verifying systems redesign ideas based on sound scientific axioms. Additionally, all design parameters synthesized through the AD process form part of the system design (Yang and El-Haik 2003), eliminating the need for prioritization using supporting techniques like Analytical Hierarchy Process (AHP) (Su et al. 2003; Oddershede et al. 2007; Gerdsri and Kocaoglu 2007; Forman and Gass 2001; Montazar and Behbahani 2007). This is unlike a redesign pursued through QFD, where the resulting critical to quality (CTQ) characteristics have to be prioritized to eliminate the weakly ranked parameters.

To enable the eventual evaluation of system redesign ideas prior to implementation, this research utilized Petri Nets (PNs) (van der Aalst and Van Hee 2002; René and Hassane 2005; van der Aalst 1998) to create and analyze models of the surgical system. PNs were chosen based on their underlying characteristic of ‘formality’, which gives them a unique mathematical edge over the rudimental flowcharting techniques. The choice of PNs enables the use of system design parameters to create models of their corresponding process definitions and subsequently analyze their logical correctness as well as their performance characteristics. PNs boast as the lone technique with the ability to support verification heuristics (evaluation of logical correctness), while simultaneously supporting simulation techniques that allow for performance analysis.

The remainder of this paper is structured as follows. Section 2 explores the studied existing surgical system to illustrate its operations and discover any inherent failures with respect to information flow. This is followed by a systematic synthesis of the architecture of a redesigned surgical system, as detailed in Sect. 3. This architecture is subsequently modeled and analyzed in Sect. 4, after which the key redesign achievements are identified and summarized in Sect. 5.

4.2 Baseline System Analysis

This research was conducted in conjunction with a healthcare provider whose health system consisted of four hospitals, with surgical facilities (OR suites) in each. The four facilities are located within a fifteen mile radius of each other, and their OR suites bear different capabilities and characteristics. The scope of this research was limited to within ‘*surgical system*’, which was defined as being comprised of (1) surgical scheduling, (2) pre-admission testing, (3) hospital admissions, (4) surgical preparatory, (5) surgical operation, and (6) surgical recovery. This section explores the flow of processes (*sequence of activities*) and information through the surgical system.

4.2.1 Flow of Processes

Surgical procedures are normally initiated outside the surgical system, primarily in the physician or surgeon's office where a decision is made to pursue surgery as a mode of care delivery. Once the decision is made, a series of activities is triggered throughout the surgical system with the primary aim of ensuring that all the required support is provided for the surgical procedure. In most cases, this series of activities will commence with scheduling the surgical case (reserving OR time) and terminate once the patient leaves the surgical care delivery facilities. These activities are further discussed in this section.

Case Scheduling

Once a decision is made to pursue surgical care, the surgeon's office will contact the preferred surgical facility to request for OR time (request to use surgical facilities and resources). This entails the process of 'case scheduling', which serves to manage and coordinate the utilization of surgical resources. For this surgical system, there were four surgical facilities to be contacted, alongside a '*central scheduling office*'. Typically, up to 75 % of the cases were scheduled through the '*central scheduling office*' (*routine cases*) whereas the remaining 25 % were scheduled by contacting the individual facilities (*add-on cases*). As a '*rule of thumb*', cases were considered as 'add-ons' if they were to be performed within 24 h of being requested.

Routine cases are scheduled by (1) faxing a '*reservation form*' to the scheduling office, (2) calling the scheduling office with the case details, or (3) contacting the scheduling office by both fax and telephone. It was estimated that 10 % of the routine cases were scheduled based on faxed request forms. The remaining 90 % involved communication by telephone, with the number of phone calls ranging from one to five. Generally, once the necessary case information was received, a scheduling clerk would use it to locate an opening in the surgical schedule, allocate the requested time (and other necessary resources), reserve time for pre-admission testing and finally provide a confirmation that the case has been scheduled. This sequence of activities is illustrated in Fig. 4.1.

On the other hand, all add-on cases are scheduled by communicating directly with the selected surgical facility via telephone. Once the telephone call is received, the facility manager will evaluate the current surgical workload to establish whether there was enough capacity for an additional case. The case is then appended to the printed-out copy of the schedule, after which the case details are updated in the surgical scheduling system. Finally, the surgical preparatory unit is informed to expect an '*add-on patient*'. This sequence of activities is illustrated in Fig. 4.2.

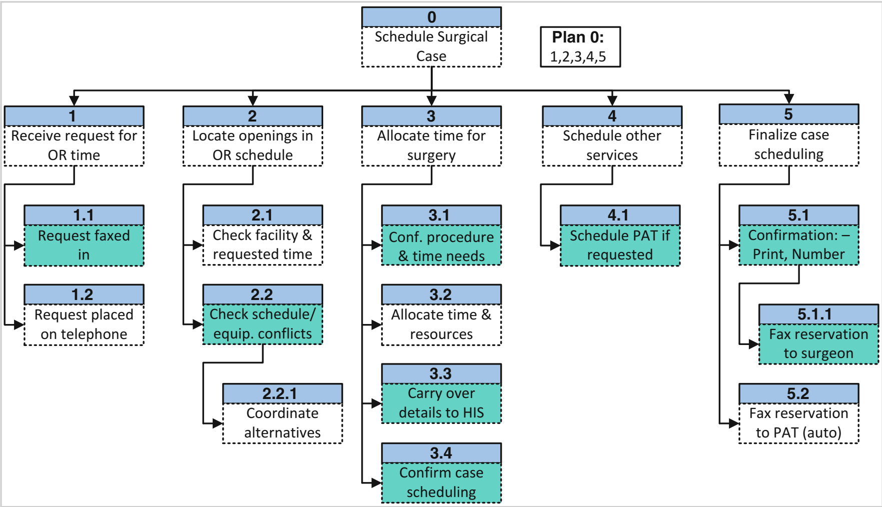


Fig. 4.1 Scheduling for routine surgical cases

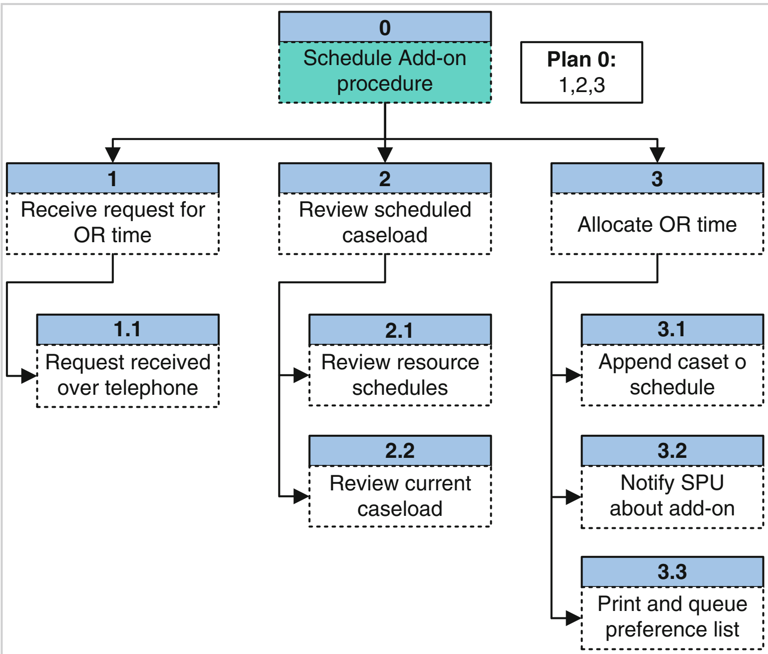


Fig. 4.2 Scheduling for add-on surgical cases

Pre-admission Testing

Before a 'clearance for surgery' is issued, patients seeking elective surgery need to undergo a 'Pre-Admission Testing' (PAT) session aimed at ascertaining their fitness for surgery. This does not include emergency and in-house patients, whose pre-surgery assessments and tests are normally conducted in the Surgical Preparatory Unit (SPU). PAT is normally conducted a few weeks to a few days before the '*day of surgery*', thereby availing ample time for any health-issues to be properly addressed. The sequence of PAT activities can be depicted as shown in Fig. 4.3.

Patients arriving for PAT could either be scheduled or unscheduled with the later being the dominating category. On arrival, they will undergo a '*mini-registration*' (to acquire basic patient information) after which they will wait for subsequent assessment by a PAT nurse. The nurse will conduct a medical history interview, followed by a physical examination. Based on this, the nurse will determine any other diagnostic tests that will be required to declare the patient 'clinically fit' for surgery. The nurse will then 'educate' the patient regarding the impending surgical procedure, and consequently dispatch him/her for the relevant diagnostic tests in the appropriate testing department.

Typically, the patient will depart the facility after completing the required tests. After analysis, test results are uploaded into the relevant information system, from

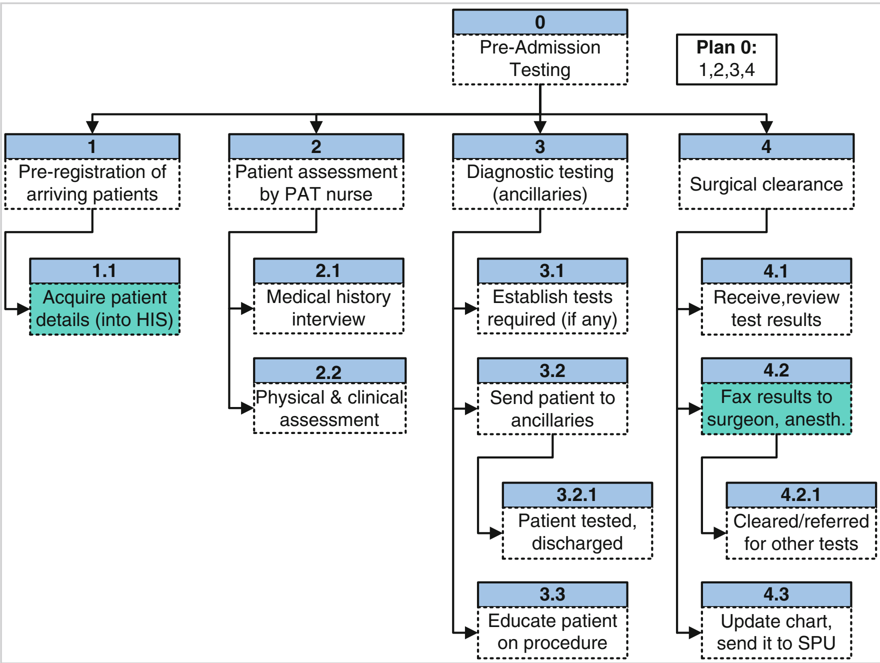


Fig. 4.3 Sequence of pre-admission testing activities

where they can be accessed via a centralized system repository (*Net Access*®). The PAT nurse will review the results and inform the anesthesiologist, who will also review them and consequently issue or decline a clearance for anesthesia administration. The results will also be viewed by the surgeon, who will issue a clearance and complete the patient's 'history and physical' assessment (a surgical requirement).

Hospital Admissions

The admission process is intended to capture patients' personal profiles and a summary of their financial and medical details. This information (1) provides a basis for the hospital to bill for services, and (2) enables the maintenance of patients' medical records. This information is captured into the Hospital Information System (HIS). Generally, incoming patients will be directed to the admissions-desk where they will undergo an admission-interview (5–15 min long). The patient's personal and clinical profile will be created or updated (for returning patients) in the HIS, after which a proof of health insurance will be requested and its copies made. Finally, the patient will sign the consents for care and information privacy before being issued with an ID band.

The patient's chart (file with clinical documentation) will be created and the relevant admissions forms updated. A '*patient transfer aide*' is then called to take the patient (and the patient chart) to the SPU. Meanwhile, the patient's insurance information is used to perform an '*insurance verification*'. Based on the outcomes, the necessary action will be initiated. For patients incapable of paying for surgical services, plans are made to educate them about the available 'payment options', including applying for 'charity care'. The hospital admissions process is illustrated in Fig. 4.4.

Surgical Preparatory

Normally, surgical patients (except emergency and some in-house patients) require some amount of preparation (SPU) before proceeding to the ORs. Emergency and in-house patients might undergo the same process or they may be prepared at a 'holding unit'. On the day of surgery, elective surgery patients are expected to arrive at least 3 h prior to their scheduled time of surgery. On arrival, they will undergo the admission process, after which they are directed to the SPU or guided there by the patient transfer aide. Here, the patients are received and given an orientation of the SPU unit by the SPU-nurses. The patients will then be assigned to a room, in which they will stay until they get transferred to the ORs.

Once settled in their rooms, patients will undergo a clinical assessment conducted by the nurse, followed by an evaluation by the anesthesiologist. The nurse will then update the relevant medical documentation and proceed to execute the surgeon's

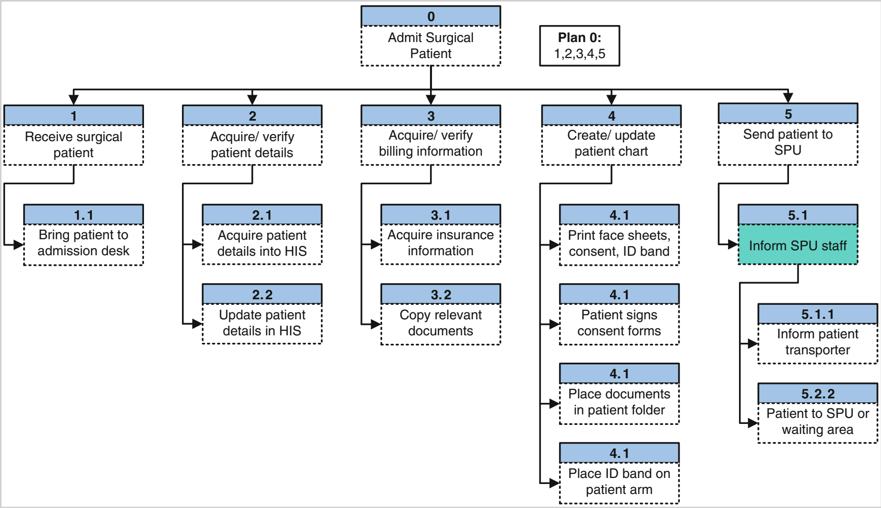


Fig. 4.4 The sequence of activities associated with hospital admission

instructions (surgical preparatory orders) to prepare the patient for surgery. When all preparations are completed, the '*surgical team*' will be informed and the patient's family will be allowed into the SPU to spend time with the patient prior to surgery. Finally, the surgeon will visit the patient to verify his/ her identity and mark the anatomical part to be operated on. With this completed, the patient is ready for transfer into the ORs. This sequence of activities is illustrated in Fig. 4.5.

Surgical Procedure

All surgical procedures are performed in the ORs, and can involve people including the surgeon, circulating nurses, OR technologists or 'scrub nurses', anesthesiologists, and any other individuals that may be required. At the bare minimum, a surgical procedure will require a 'surgical team' composed of the surgeon, a circulating nurse, a nurse anesthesiologist and a scrub nurse. Once the team is alerted of their next patient, the nurse anesthesiologist and circulating nurse will proceed to the SPU to verify the patient's identity and initiate the 'hand-over' process.

By this time, the OR will already be set-up with all the equipment and supplies required for surgery. The patient will be brought in and transferred onto the operating table, where he/she will be properly positioned before being sedated. The surgeon will then commence the surgery while the nurse anesthesiologist keeps track of the patient's vital signs. At the same time, the circulating nurse will oversee all OR activities and subsequently document them in the surgical information system. Eventually, the surgeon will complete the operation and depart the OR.

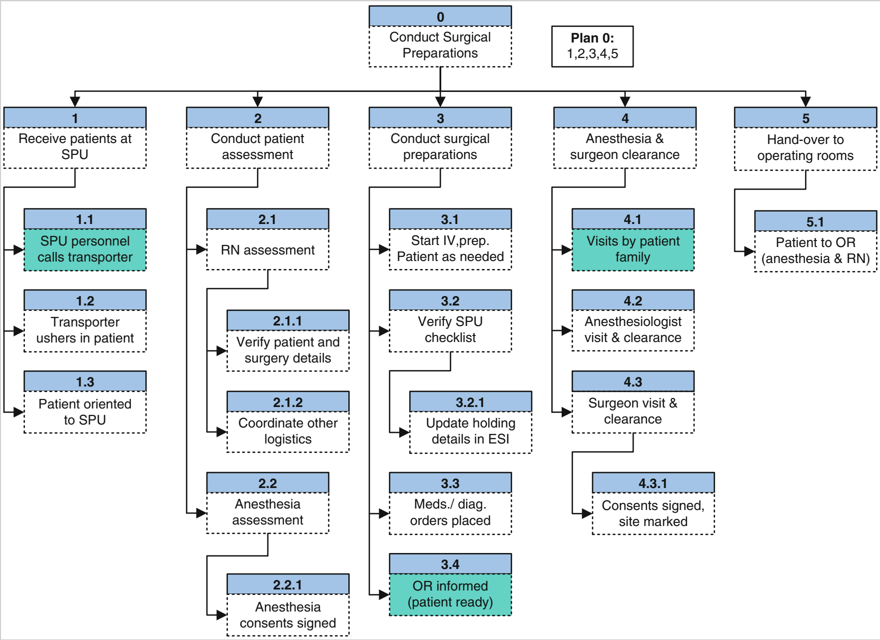


Fig. 4.5 The sequence of activities associated with surgical preparatory

The patient will be cleaned-up and bandaged before being brought out of sedation and subsequently transported to the Post-Anesthesia Care Unit (PACU) by the nurse anesthesiologist and the circulating nurse.

The circulating nurse will conduct a '*supplies count*' to document all the surgical supplies utilized during the surgical procedure. He/ she will also document the OR time utilization and complete any other applicable surgical documentation. Concurrently, support staff will be called in to clean-up the OR in preparation for the next procedure. Eventually, the nurse will hand-over copies of the documentation to the OR unit secretary for subsequent 'charge entry' (billing for time and supplies utilization). This sequence of activities is illustrated in Fig. 4.6.

Surgical Recovery

All surgical patients require some amount of recovery or observation before they can depart the surgical system—this is the core function of the PACU. As illustrated in Fig. 4.7, a surgical patient will be received by a PACU nurse, who will also sign-off the relevant hand-over documentation. The patient will then be assigned to a recovery bed, where they will be connected to health-monitoring devices and

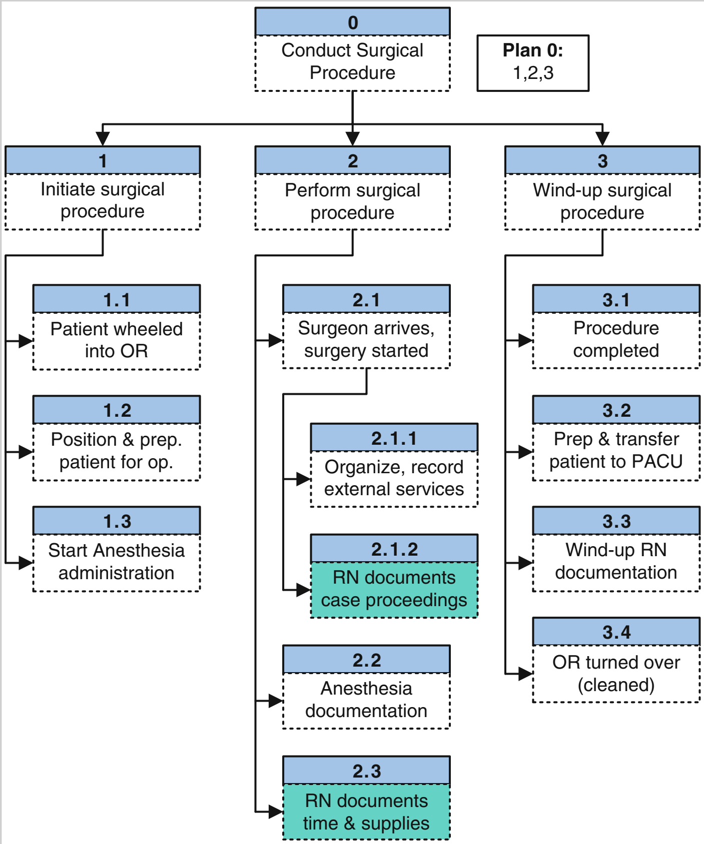


Fig. 4.6 Sequence of activities during a surgical operation

consequently kept under close observation. Typically, recovery patients are cared for based on the '*recovery orders*' which are issued and signed by the surgeon after the surgery is completed.

Some facilities have a '*single recovery level*' while others have '*two recovery levels*'. For those with a single recovery level, their SPU spaces are often utilized as the second recovery level. All patients will commence their recovery at the first level, after which they may be admitted to extended care units (e.g., *nursing floors, ICU*) or transferred to the second recovery level for further recovery and subsequent discharge (allowed to go home). Typically, the second recovery level will only handle the Same-Day-Surgery (SDS) category of patients, as they are eventually discharged to depart the facility.

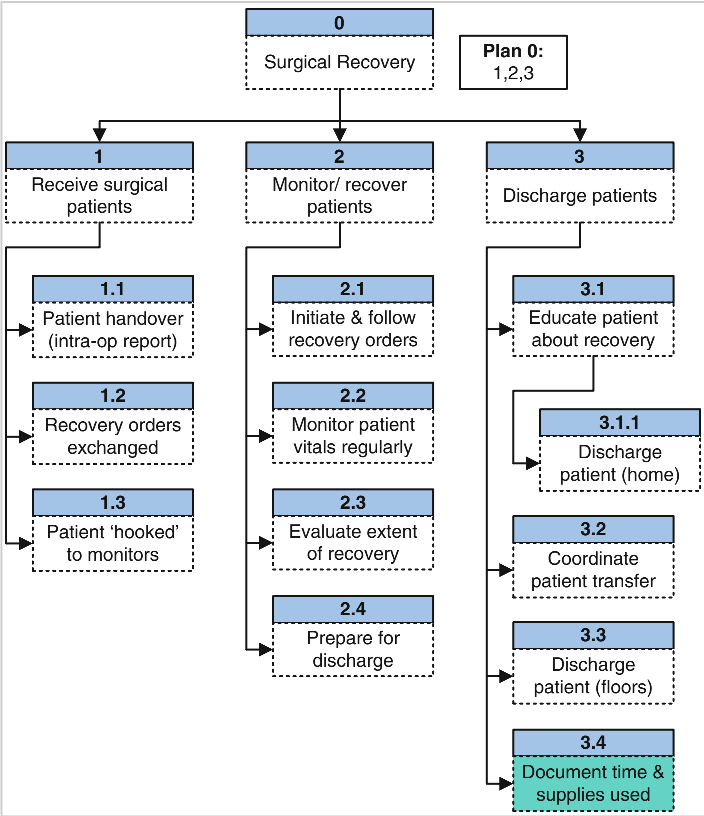


Fig. 4.7 The sequence of activities associated with surgical recovery

4.2.2 *Failures in Information Flow*

Continuous communication is often required to coordinate the smooth flow of patients through the surgical system. This may be verbal (*telephone, pagers or word-of-mouth*), written (*fax and email*) or intuitive (*mental judgment*). Figure 4.8 makes an attempt to capture the key forms of information flow in the surgical system alongside the patient flow during the process of surgical care delivery. The various modes of communication exhibited form the primary means by which the various stakeholders are kept informed of the progress of the surgical process. Therefore, it is relatively important to have the right information at the right time as this enables the right operational and care decisions.

The main purpose of conducting a baseline evaluation on the existing surgical system was to identify and elucidate modes of failure inherent in its information flow structure. The analysis was successful in revealing key failures in information

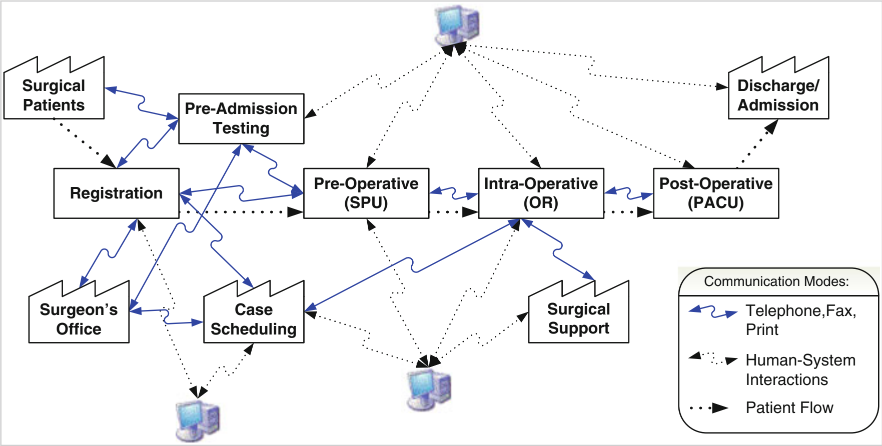


Fig. 4.8 Flows of information and operations in the surgical system

flow within the surgical system, hence setting focus for the subsequent redesign efforts. Some of the main failures that were identified are briefly discussed below.

1. There is a lack of 'real-time knowledge' of the operating status of the OR suites in the surgical system. When needed, tacit knowledge of the OR status was obtained by contacting the staff at specific facilities and inquiring about the proportions of cases completed, cases in progress and cases pending. This implies that processes like case scheduling have to be based on general estimates, rather than explicit knowledge of the surgical operations. This directly affects the effectiveness of case scheduling and consequently the utilization of OR facilities.
2. Often, there was insufficient information regarding the resource needs for various surgical cases (i.e., trends in OR time utilization, equipment needs and personnel preferences). To some extent, this was impacted by the lack of systematic tracking for the status of surgical equipment and facilities. For the purpose of proactive planning, the lack of such information often incapacitated the ability to accurately forecast surgical case requirements. As a result, it was often impossible to have a complete 'bill-of-materials' for a surgical case. This subsequently impacted the surgical case progress, even leading to delays and cancellations.
3. It was common to experience various forms of conflicts during a typical surgical case (e.g., multiple facility reservations, equipment unavailability, and multiple-scheduling of personnel). The existing system for conflict identification did not bear detailed information regarding the location and state of equipment, trend of reservations and/or cancellations of facilities and specific details on the calendars of OR team personnel. Such conflicts often led to case delays, and at times even resulted to cancellations.

4. Similar to resource conflicts, there were also schedule clashes which were typical for the operating surgeons. The surgeons' calendars were often dynamic and inaccessible during scheduling. Subsequently, it was often difficult to identify a 'conflict free' time-slot for a surgeon's case without calling the surgeon's office to inquire about his/ her calendar. This resulted to avoidable and redundant communication (rework) during the surgical scheduling process.
5. The existing process for tracking the utilization of surgical supplies and managing their replenishment was purely manual, depending on a physical inventory-count and the associated paper-documentation. Besides being laborious, this system was also inaccurate and was often associated with delays in billing for the supplies utilized during a surgical procedure. This ineffective tracking of surgical supplies utilization was also associated to delays in the surgical procedure, especially those attributed to the timely availability of the appropriate surgical supplies.
6. A persistent trend of lateness and/or uninformed personnel was also observed in the existing surgical system. Notably, this was a result of the busy schedules for the OR team though it could also be explained by the observation that all personnel beepers and alarms were not triggered by explicit events in the surgical process. Besides being a major cause of delays in surgical cases, lateness also impacted the patient's perception of the quality of care as well as their overall satisfaction with the surgical care services.

These failures were noted amid the myriad of communication instances associated with answering patient calls and inquiries, coordinating and scheduling add-on cases, locating and alerting surgical personnel, updating personnel on the progress of surgical cases, etc. These observations were supported by user opinions collected as part of the redesign process for the surgical system. Therefore, it was apt that the various forms and modes of communication be redesigned in a manner that would streamline the associated flow of information, a primary undertaking of this research effort.

4.3 Surgical System Redesign

Redesign of the surgical system was addressed using Axiomatic Design (AD). AD commences in the customer domain, where the relevant Customer Needs (CNs) are identified. Information from earlier 'transformational sessions' was utilized to frame a detailed list of unambiguous CNs. These were then assigned a 'customer importance' rank in addition to a 'direction of improvement'.

The QFD technique was then utilized to relate the CNs to their corresponding '*critical to quality*' (CTQ) characteristics, equivalent to the desired Functional Requirements (FRs) of the functional domain. Only one phase of the QFD was performed, as this was sufficient to convert the CNs into corresponding CTQs. A total of eight FRs were realized from the initial list of eighty CNs. Based on their

0	FR	Faciltiate efficient surgical operations	DP	Integrated surgical processes
1	FR	Improve software features	DP	Customized software compontents
1.1	FR	Convenient use of software applications	DP	Integrated navigational wizards
1.2	FR	Accurate tracking for surgical data	DP	Automated data capture
1.2.1	FR	Auto-logging of patient status data	DP	Patient tracking module
1.2.2	FR	Auto-logging of supplies/ equipment data	DP	supplies/equipment tracking module
2	FR	Provide alerts and alarms	DP	Automated system reminders
2.1	FR	Generate alerts on resource status	DP	State-triggered alerts
2.2	FR	Trigger alarms (timed reminders)	DP	Event-triggered alarms
3	FR	Manage resources better	DP	Resource management system
4	FR	Support decision making	DP	Decision support system
4.1	FR	Establish current case requirements	DP	DSS - BOM generator
4.2	FR	Catch/Resolve scheduling conflicts	DP	DSS - Scheduling
4.3	FR	Evaluate status of the OR (utilization)	DP	DSS - OR status
5	FR	Improve access to information	DP	Distributed computer applications
5.1	FR	View relevant information centrally	DP	Relational databases
5.2	FR	Access information across departments	DP	Robust computer network
5.2.1	FR	Access relevant information online	DP	Web portal (surgical information)
6	FR	Streamline scheduling activities	DP	Effective processes
6.1	FR	Consistent procedures	DP	Standardized business processes
7	FR	Complete actions on time	DP	Coordinated tasks
8	FR	Archive surgical data	DP	Extended archiving modules

Fig. 4.9 Snapshot of the architecture for the redesigned system

relative ranking, some of the CTQs appeared to be weak, though not completely unrelated to the stipulated customer needs.

'*FR-DP (Design Parameter) decomposition*' was used to facilitate a transition from the functional to the physical domain. Starting with the system-level FR ($\boldsymbol{FR_0}$), the corresponding DP ($\boldsymbol{DP_0}$) was established; after which the zigzagging approach was followed to decompose all FRs into their atomic sub-FRs and corresponding DPs and sub-DPs. This yielded a *system architecture*, a snapshot of which is illustrated in Fig. 4.9. This depicts a summary of all the system elements that are required to redesign the surgical system as stipulated by the customer needs. This set of elements (design parameters) is further discussed below.

1. **Workflow Support and Standardization**: The DPs that support the realization of this objective include DP1.2.1, DP1.2.2, DP6.1 and DP7.1 (patients tracking module, supplies/equipment tracking module, standardized business processes and scheduled preparatory activities). These are designed to satisfy the established need for standardized and predictable surgical processes. They also make it possible to monitor the flow of surgical patients as well as the status and utilization of surgical equipment and supplies. Together, these system components serve to enhance patient safety, information accuracy and resource management.

The patient tracking component will bear the capability to identify the location of a patient in the surgical system at any given time. Ideally, this will be comprised of an electronic tracking mechanism interfaced with appropriate software to monitor patient flow through various stages of surgical care. The supplies/equipment tracking component will work in a similar manner, only that it will be monitoring the location and utilization of surgical equipment and supplies. The functionality of these system components will be supported by standardized surgical processes, which among other things will include a scheduling mechanism for surgical preparatory activities.

2. **Resource Management**: This objective will be satisfied by the system components defined by DP3.1, DP3.2.1 and DP3.2.2 (materials management applications, OR scheduling module and personnel scheduling module). Together, these system components will facilitate the improved use of resources (materials and time) within the surgical system. The materials management applications are intended to provide the '*logic processing unit' for the equipment a*nd supplies tracking modules defined earlier (***DP1.2.2***). On the othe***r hand,*** the scheduling modules will coordinate the logic associated with (1) allocating OR time to surgical cases, (2) allocating clinicians to surgical cases, and (3) providing information links to external schedules (e.g., surgeons' office schedules).
3. **Enhanced Data Warehousing**: The need for extended information warehousing will be satisfied by the system components defined by DP5.1 and DP8 (use of relational databases and extended archiving modules). Well developed relational databases will be utilized to enhance the storage and integrity of the various datasets gathered throughout the surgical system. This will ensure that a well defined knowledge-base of surgical information is created, providing a basis for the enactment of '*decision support systems*' (*DSS). On the other hand,* extended archiving modules will provide the capacity necessary to meet the growing demands for data to meet the analysis and information needs of modern day organizations.
4. **Informed Decision Making**: The ultimate goal of surgical information integration is to further enhance the culture of informed decision making, both for direct patient care and support activities within the surgical system. This realization will be enabled by the system components defined by DP4.1, DP4.2 and DP4.3 (DSS: bill of materials, DSS: scheduling and DSS: OR status). Essentially, these components will form a knowledge based system that will facilitate/ support surgical decision making. This system will utilize the data already collected by other system components (workflow support, resource management and data archiving) to extract business intelligence and recommend or prompt the appropriate actions.

A decision support system can be viewed as an interactive computer-based system that is designed to enable decision makers to better utilize communication technologies and data to identify and solve problems or make decisions. For the surgical system, the '*DSS: bill of materials*' was desirable as it would provide the intelligence necessary to coordinate the allocation of surgical supplies and equipment to cases being scheduled. This would utilize the archived surgical procedure profiles, alongside surgeon preferences, to establish the precise requirements for all cases being scheduled. On the other hand, the '*DSS: scheduling*' and '*DSS: OR status*' would work hand in hand to identify and circumvent/ eliminate potential conflicts in personnel and OR suite schedules.

5. **Activity Coordination Based on Cues and Alerts**: In order to enforce compliance and timely actions, automated reminders are incorporated into the design of the surgical system using the system features defined by DP2.1 and DP2.2 (state triggered alerts and event triggered alerts). State triggered alerts are designed to prompt for action whenever '*undesirable*' system states are attained (e.g., l*ow inventor*y levels, misplaced surgical equipment and resource conflicts among others). On the other hand, event triggered alerts are aimed at attracting attention to timed surgical activities (e.g., scheduled surgical or PAT appointments). These reminders are generally aimed at enhancing the smooth coordination of surgical activities.
6. **Enhanced Access to Information**: To maximize the effectiveness of the information attained from the integrated surgical system, the system components defined by DP1.1 and DP5.2.1 (navigational wizards and web portals for surgical information) are relied upon to provide intuitive and flexible access to surgical information. Web portals are considered as an ideal approach to providing a wide array of users with flexible and scalable access to information. These portals will be designed with integrated 'wizards' (software functionalities incorporated into the graphical user interface to simplify the use of a software application) as a means to enhance their usability and effectiveness.

To ensure that the architecture depicted a sound and robust system, the design was analyzed based on the 'independence axiom'. Using the 'Design Structure Matrix' approach, the design was evaluated and ascertained to exhibit no coupling among the system FRs. The surgical system design was established to be '*decoupled*', suggesting that its architecture represented a system, free of conceptual and operational vulnerabilities. The design matrix corresponding to this system architecture is illustrated in Fig. 4.10.

$$\begin{Bmatrix} FR_1 \\ FR_2 \\ FR_3 \\ FR_4 \\ FR_5 \\ FR_6 \\ FR_7 \\ FR_8 \end{Bmatrix} \begin{bmatrix} A_{11} & A_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{22} & A_{23} & A_{24} & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{33} & A_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{44} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_{55} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{66} & A_{67} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{77} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{88} \end{bmatrix} \begin{Bmatrix} DP_1 \\ DP_2 \\ DP_3 \\ DP_4 \\ DP_5 \\ DP_6 \\ DP_7 \\ DP_8 \end{Bmatrix}$$

Fig. 4.10 A design matrix evaluating the redesigned system

Table 4.1 'Atomic' FR-DP combinations for the integrated surgical system

FR1.1	Convenient use of software applications	DP1.1	Integrated navigational wizards
FR1.2.1	Auto-logging of patient status data	DP1.2.1	Patient tracking module
FR1.2.2	Auto-logging of supplies/equipment data	DP1.2.2	Supplies/equipment tracking module
FR2.1	Generate alerts on resource status	DP2.1	State-triggered alerts
FR2.2	Trigger alarms (timed reminders)	DP2.2	Event-triggered alerts
FR3.1	Control supplies and equipment	DP3.1	Materials management applications
FR3.2.1	Control use of OR time	DP3.2.1	OR scheduling module
FR3.2.2	Manage time for clinical personnel	DP3.2.2	Personnel scheduling module
FR4.1	Establish current case requirements	DP4.1	DSS—BoM generator
FR4.2	Catch/Resolve scheduling conflicts	DP4.2	DSS—Scheduling
FR4.3	Evaluate status of the OR (utilization)	DP4.3	DSS—OR status
FR5.1	View relevant information centrally	DP5.1	Relational databases
FR5.2.1	Access relevant information online	DP5.2.1	Web portals for surgical information
FR6.1	Consistent procedures	DP6.1	Standardized business processes
FR7.1	Finish preparations for surgery earlier	DP7.1	Scheduled preparatory activities
FR8	Archive surgical data	DP8	Extended archiving modules

Table 4.1 presents a summary of the system components that formed the synthesized architecture. Consecutively, Fig. 4.11 illustrates how these identified system components will relate to each other in the envisioned integrated surgical care delivery system. As it can be observed, the changes introduced by this system redesign are mainly driven by information technology. As a result, their impact will be primarily noticed on the '*technology layer*' of the service oriented architecture for the surgical care delivery system, whereas the business and service layers will remain almost intact.

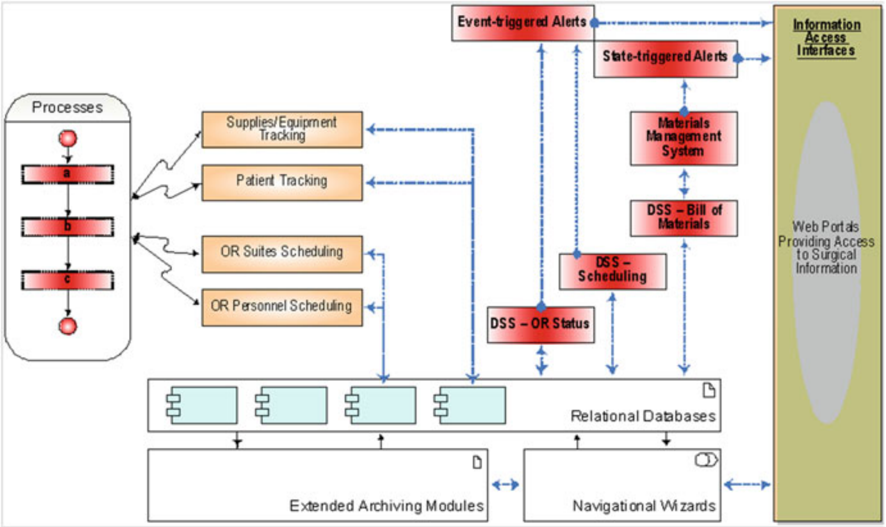


Fig. 4.11 Integration of information in the conceptual surgical system

4.4 System Modeling and Analysis

It is essential to ensure that errors do not exist in process definitions representing models developed for the purpose of systems redesign, as this may compromise the quality and outcomes of the redesign effort (van der Aalst and Van Hee 2002). For this research, it was necessary to verify the system architecture that was realized from the redesign efforts that were focused on the surgical system. The architecture was first subjected to verification using Petri Nets (PNs) and Dependency Structure Matrix (DSM). Surgical scheduling processes were then separated and analyzed quantitatively to address the time, cost, quality and flexibility dimensions. This analysis was focused on quantifying the improvements attained through the redesign efforts.

4.4.1 Qualitative Evaluation Using Petri Nets

By using the various PN constructs, tasks associated with all the key activities of the existing surgical system were modeled into the lifecycle of a surgical case from initiation (surgical scheduling) to termination (discharge from surgical care). This provides an integrated systems view of the surgical operations, which makes it possible to analyze the relations exhibited between the various elements of the surgical system. The baseline model was consequently modified to depict all the system redesign features that were synthesized through the redesign effort based on Axiomatic Design concepts (discussed earlier).

The PN models representing the surgical system (before and after redesign) are shown in Figs. 4.12 and 4.13. PN-based heuristics were used for system analysis, which consisted of a validity check followed by a verification of the associated process definitions. Validation was accomplished by utilizing the '*token-game*' technique, a form of interactive simulation availed by PN modeling tools. This enabled a validation of the behavior of the surgical system (i.e., ensuring that the sequence of execution of the modeled activities conformed to the known flow of surgical operations as well as the expected process logic of the integrated surgical system).

The inspection of process definitions for errors forms the basis upon which heuristics for process verification are defined. Various properties have been proposed in the bid to establish a '*precise definition of correctness*' as relates to process definitions. Important among these is the soundness property, which is considered to be satisfied when a set of process definitions meet the following three requirements (van der Aalst and Van Hee 2002): (1) for each token introduced into the net, only one token results in its '*end*' place, (2) once a token reaches the '*end*' place, there should be no token in any other place in the net, and (3) it is possible to move from the initial state of a task to one in which it is enabled. Computerized heuristics verifying process definitions will normally use '*short-circuited*' PNs (with an extra transition: '***t***') to test for two main properties—liveliness and boundedness (van der Aalst and Van Hee 2002). A '*live*' net is viewed as one in which each transition can fire an arbitrary number of times, whereas a '*bounded*' net is one which does not exceed a given limit of tokens per place.

4.4.2 Qualitative Analysis Using Dependency Structure Matrix

Due to the complexity illustrated by information flow and the various interactions within the surgical system, it was necessary to focus on ensuring that the proposed system integration would enable the realization of streamlined information flow. Ideally, process activities should exhibit information flow in a linear and non-redundant fashion which helps to minimize process wastes (i.e., the energy or resources involved with rework in a process). Thus, the goal of this analysis was to compare the structure of information flow realized in the conceptualized system with that observed in the existing surgical system. Based on this comparison, was possible to identify the system configuration with the 'best' information flow characteristics.

To accomplish this, the Dependency Structure Matrix (DSM), a presentational and analysis tool that is invaluable for the modeling of systems, especially in situations where decomposition and integration are necessary (Browning 2001), was utilized to identify and evaluate the dependencies exhibited by the various flows of information identified within the surgical system. The analysis of DSM

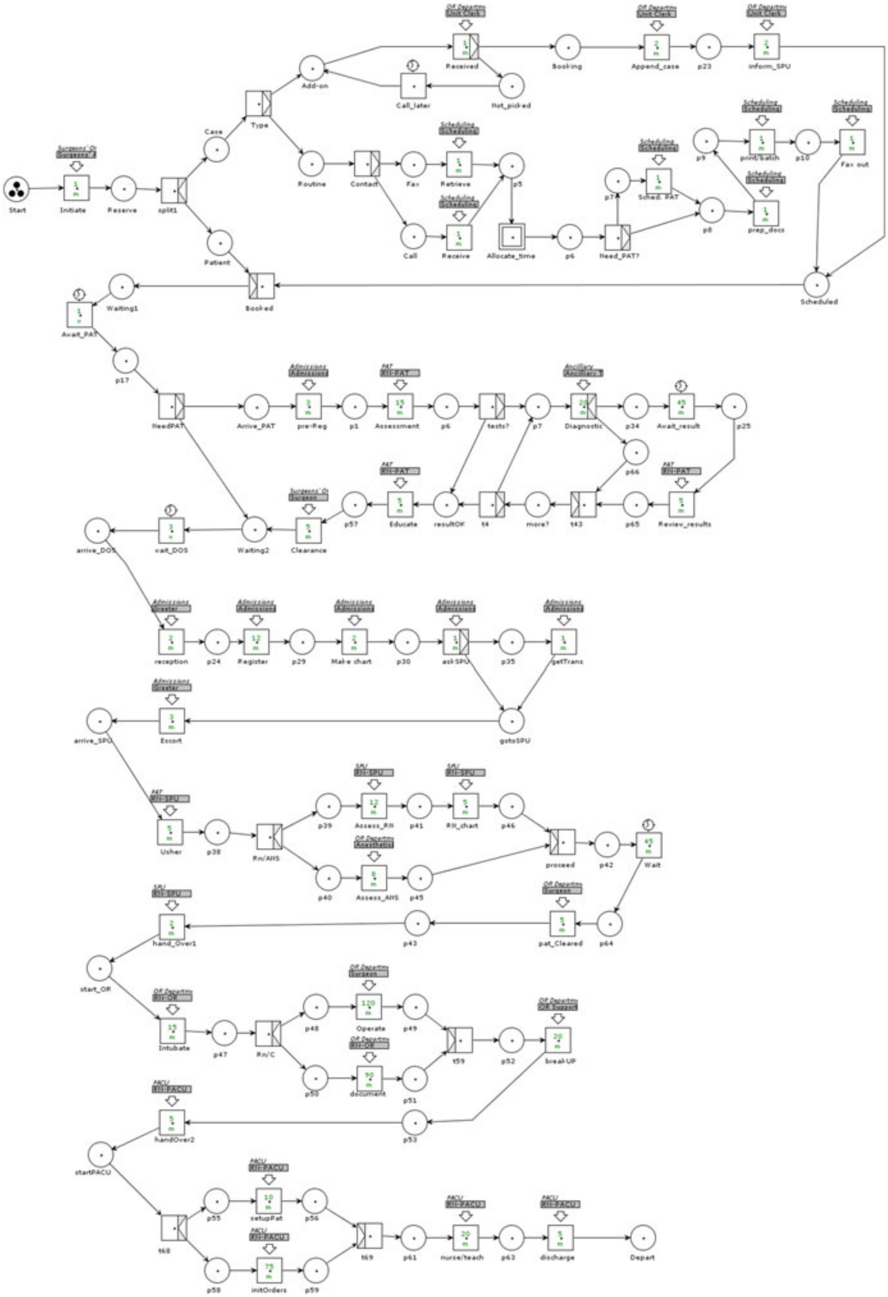


Fig. 4.12 Petri net model of processes of the baseline surgical system

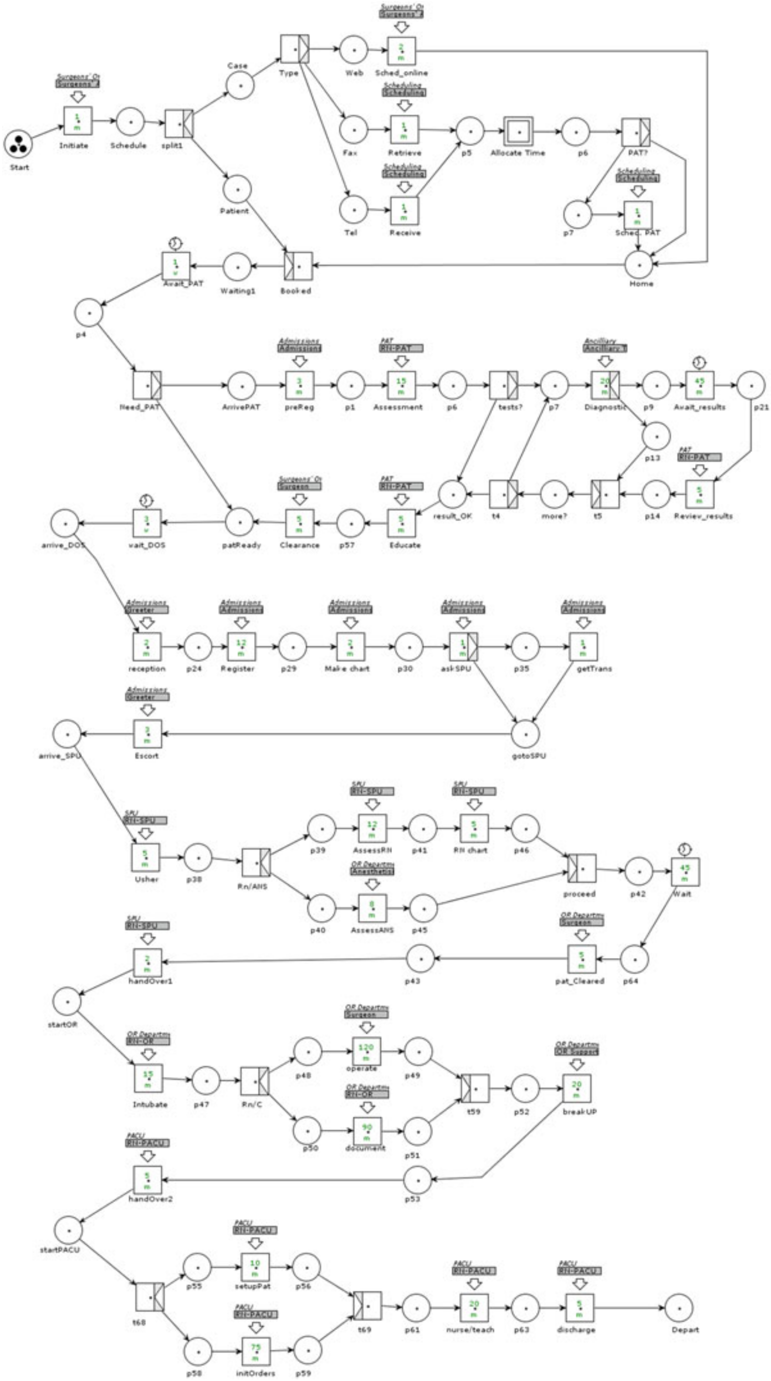


Fig. 4.13 Petri net model of the redesigned surgical system

models is mainly conducted through a '*partitioning*' process which involves the rearrangement of the tasks presented in the matrix in a manner that will cluster all dependency marks in the lower diagonal or group them together into squares around the main diagonal. This kind of analysis will identify loops/cycles of information (*coupled tasks*) and try to cluster them close to the main diagonal. Even though coupling is an undesired feature in such models, it is often tolerated especially where it depicts a planned iteration. Generally, the partitioned matrix presents an effective tool for systems integration as it emphasizes the interfaces between activities (i.e., their dependencies). Used effectively, such a matrix will enable processes to be dissected into their constituent activities, which are in turn stratified into layers and reorganized in an execution order that enhances the efficiency of information flow.

DSM models were created to depict the flow of information in both the existing and the conceptualized designs of the surgical system. Emphasis was placed on evaluating the activities associated with surgical scheduling as these were more clearly defined into a single process. The process was stratified into its atomic activities, after which their relations were identified and mapped out in a DSM. These were then analyzed using the DSM partitioning algorithm (based on the '*As-Early-As-Possible*' rule), which yielded the matrices illustrated in Figs. 4.14 and 4.15 for the baseline and the conceptualized designs respectively. By analyzing these models, it is observed that the dependencies among the activities in the conceptualized model are more streamlined than those in the existing system. This is depicted by the reduced levels of dependency highlighted in the latter model, compared to the former. It suggests that the redesigned scheduling process has fewer opportunities for information integration than the existing one. This can be attributed to the redesign, which enabled the integration of information in a manner that minimizes the layers of information dependency.

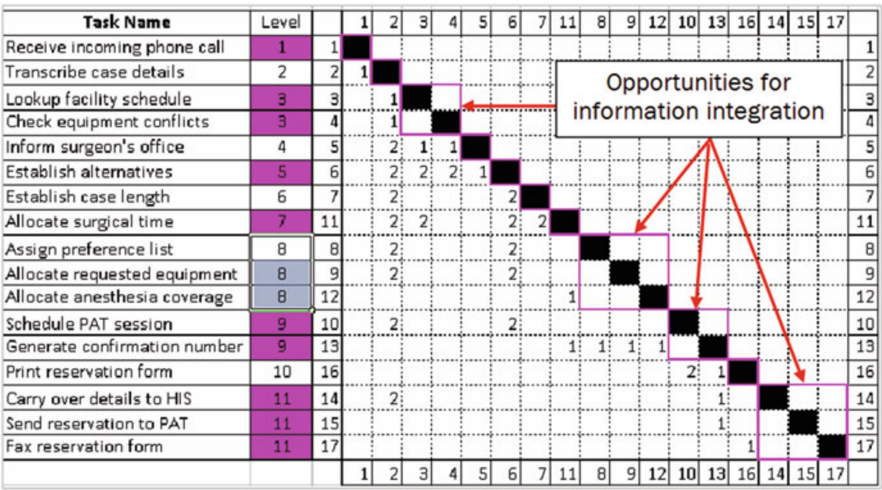


Fig. 4.14 Partitioned model for scheduling activities (baseline system)

Task Name	Level		1	2	3	4	5	6	7	8	
Receive phone-call/ fax	1	1									1
Transcribe case details	2	2	1								2
Indentify resource conflicts	3	3		1							3
Inform surgeon's office	4	4		2	1						4
Establish alternatives	5	5		2	2	1					5
Allocate surgical resources	6	6		2			1				6
Allocate time requirements	6	7		1			2				7
Finalize case scheduling	7	8		2				1	1		8
			1	2	3	4	5	6	7	8	

Fig. 4.15 Partitioned model for scheduling activities (conceptualized system)

4.4.3 Quantitative Evaluation

To quantify the impacts that the proposed redesign would bear on the surgical system, the verified PN models were subjected to a quantitative analysis. This analysis focused on the surgical scheduling operations, for which the 'before' and 'after' Petri Net flow-logic states are shown on Figs. 4.16 and 4.17. As the modeled tasks exhibited various categories of interactions, including parallelism, iterations and sequential orderings, it was important to establish the actual number of times that each task would be executed if the modeled system was enacted. This was achieved by mapping-out the reachability graph based on the constructs employed in the Petri Net models.

With the execution of 120 cases (the estimated daily workload for the surgical system), it was possible to estimate the '*processing*' and '*resource engagement'* times in both models of the surgical system. The outcomes of this analysis are illustrated in Tables 4.2 and 4.3 for the baseline and modified models of the surgical system respectively. Here, '*task*' gives a description of a particular scheduling activity, which requires the duration '*service time*' to execute based to the order depicted by its '*ID*'. The '*resource class*' simply defines the category of resources responsible for the various tasks. '*Items/case*' is the number of times that any given task will be executed in order to complete the overall execution of N cases ($N = 120$). Therefore, the time needed for each task to execute a single case (*time/case*) is given by [*(service time)* * *(items/case)*] whereas the time needed for each case to execute the overall workload (N) is given by [*(service time)* * *(N* (items/case))*].

Outcomes of the analysis indicated that scheduling of a typical surgical case in the existing system required three resources classes and consumed an estimate of 17.55 min. On the other hand, the idealized concept of an integrated surgical system gave the promise of scheduling a surgical case in about 5.37 min while engaging two categories of resources. This suggests a time saving of 12.18 min per surgical case, translating to a 69 % decline in the work-hours dedicated to surgical case

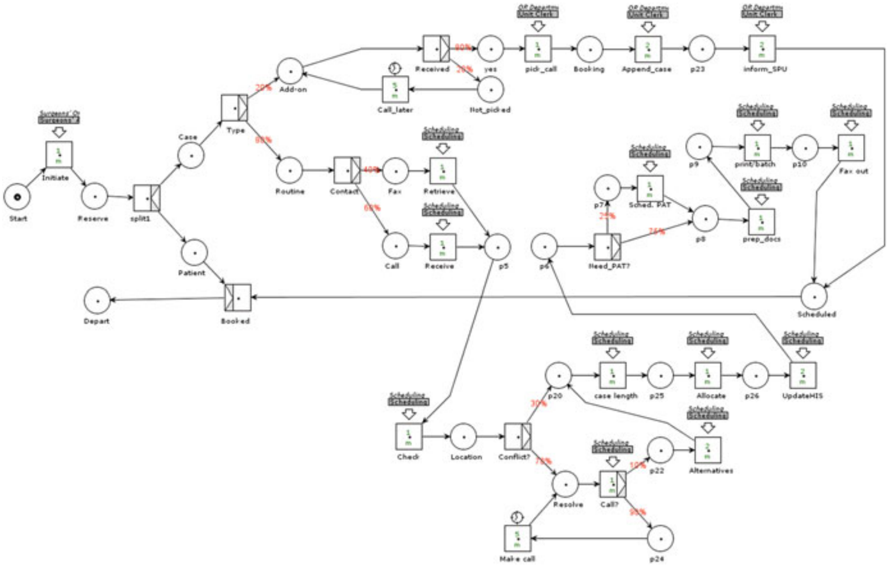


Fig. 4.16 Petri net model of the existing scheduling process

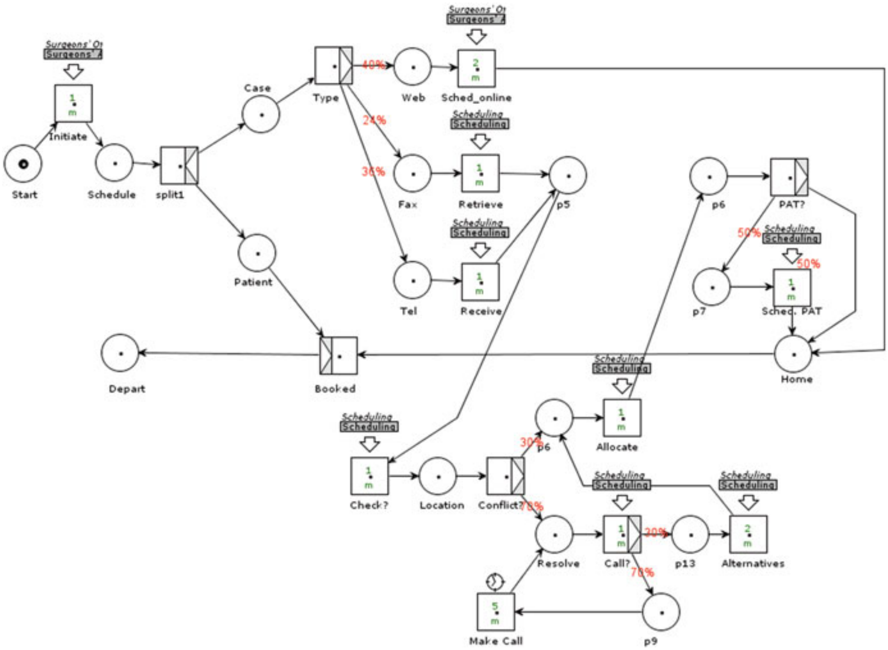


Fig. 4.17 Petri net model of the redesigned scheduling process

Table 4.2 Quantitative evaluation of the baseline scheduling activities

ID	Task	Resource class	Service time	Items/case	Time/case	Time/period
2	Retrieve faxed request	Scheduling	0.5	0.32	0.16	19.2
2	Receive request call	Scheduling	0.5	0.48	0.24	28.8
3	Check for conflicts	Scheduling	1.0	0.32	0.32	38.4
4	Establish need, call surgeon	Scheduling	1.0	2.24	2.24	268.8
4	Follow up call (scheduler)	Scheduling	5.0	2.02	10.1	1212
5	Establish alternatives	Scheduling	2.0	0.22	0.44	52.8
6	Establish, confirm case length	Scheduling	1.0	0.32	0.32	38.4
7	Allocate OR time	Scheduling	0.5	0.32	0.16	19.2
8	Update case details in HIS	Scheduling	2.0	0.32	0.64	76.8
9	Schedule PAT	Scheduling	0.5	0.08	0.04	4.8
10	Finalize case documentation	Scheduling	1.0	0.32	0.32	38.4
11	Print confirmations	Scheduling	0.5	0.32	0.16	19.2
12	Fax case confirmation	Scheduling	0.5	0.32	0.16	19.2
1	Request case time	Surgeon Office	1.0	1.00	1.00	120
2	Follow up call (OR)	Surgical	5.0	0.05	0.25	30
2	Receive call (OR)	Surgical	1.0	0.20	0.20	24
3	Inform SPU of add-on	Surgical	2.0	0.20	0.40	48
4	Append add-on case	Surgical	2.0	0.20	0.40	48
					17.55	2106

scheduling. It also translates to a workload reduction for scheduling clerks, since the amount of time dedicated to scheduling surgical cases in a day is reduced from 1836 min (average utilization of 95 %) to 428 min (average utilization of 22 %).

Table 4.3 Quantitative evaluation of the redesigned scheduling activities

ID	Task	Resource class	Service time	Items/case	Time/case	Time/period
2	Retrieve faxed request	Scheduling	0.5	0.24	0.12	14.4
2	Receive request call	Scheduling	0.5	0.36	0.18	21.6
3	Check for conflicts	Scheduling	1.0	0.24	0.24	28.8
4	Establish need, call surgeon	Scheduling	1.0	0.56	0.56	67.2
4	Follow up call (scheduler)	Scheduling	5.0	0.39	1.95	234
5	Establish alternatives	Scheduling	2.0	0.17	0.34	40.8
6	Allocate OR time	Scheduling	0.5	0.24	0.12	14.4
7	Schedule PAT	Scheduling	0.5	0.12	0.06	7.2
1	Request case time	Surgeon Office	1.0	1.0	1.0	120
2	Schedule online	Surgeon Office	2.0	0.4	0.8	96
					5.37	644.4

4.5 Summary and Conclusions

In elucidating the achievements of this research, it is necessary to take a retrospective view at its objectives. By embracing an AD approach, this research is able to illustrate a systematic and scientific methodology for ‘mapping conceptual ideas of systems redesign into process improvement’ (first objective). Features of the redesigned system support the improved coordination of surgical activities and resources, thereby enhancing the operational efficiency of the surgical system. This is a stipulation of the second research objective. On the other hand, the introduction of Petri Nets to facilitate in quantifying the redesign impacts serves to satisfy the third objective.

The wider scope of the fourth objective is satisfied by concatenating the various techniques used throughout this research into a reusable framework, shown in Fig. 4.18. This shows how ideas of process improvement (stakeholder and customer requirements) can be synthesized into informal and formal parameters for systems redesign. In the process, concepts of Business Process Redesign/ Reengineering (BPR) are introduced to ensure that the conceptualized system design leverages upon all available opportunities to realize an overall improvement within the system in focus. At this point, the conceptual system can be tested to ascertain that it

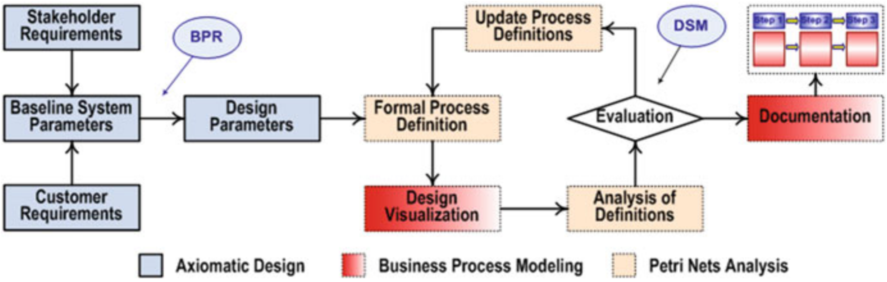


Fig. 4.18 Conceptual framework for systems (re)design

adheres to sound Systems Engineering principles. This system can then be modeled using PNs and subsequently analyzed to verify its logical correctness as well as its expected performance measures.

This conceptual framework presents a formal and systematic means by which redesign efforts can be channeled when deploying process improvement philosophies. It illustrates how various techniques can be utilized to support each other throughout the system design cycle. The incorporation of AD into this framework ensures that the design conceptualization is well rounded (encompassing all stakeholder views) yet precise (devoid of the characteristic design-verify loops). This is attributed to the unique features of the AD approach. The PNs theory, by far the most popular framework for workflow modeling (Gido 2000), then becomes an attractive technique for modeling, validating and verifying workflow descriptions of the conceptual system.

This framework has been utilized within the scope of this research and illustrated to be effective. Its application has been illustrated based the redesign of an existing surgical system, where various ISE techniques have been used to illustrate process improvements including (1) a streamlined surgical scheduling process with vast information access capabilities, (2) automated resource management capabilities, (3) real-time patient tracking abilities, and (4) integrated alarms and reminders. By focusing on the surgical scheduling process, it was further illustrated that productivity improvements (reduction in work-hours) of up to 69 % were feasible. Additionally, this research was also able to establish and illustrate a formal and explicit design approach that could be adopted in BPR projects, especially when undertaking redesign using philosophies like BPM, TQM, and Six-Sigma. Thus, this framework is recommended as an invaluable tool for future applications on the redesign of systems and processes.

References

Browning T (2001) Applying the design structure matrix to system decomposition and integration problems: a review and new directions. IEEE Trans Eng Manag 48(3):292–306

Dexter F, Macario A, Traub R, Hopwood M, Lubarsky D (1999) An operating room scheduling strategy to maximize the use of operating room block time: computer simulation of patient scheduling and survey of patient's preferences for surgical waiting. Anesth Analg 89:7–20

Forman E, Gass S (2001) The analytic hierarchy process: an exposition. Oper Res 49(4):469–486

Gerdsri N, Kocaoglu D (2007) Applying the analytic hierarchy process (AHP) to build a strategic framework for technology roadmapping. Math Comput Model 46:1071–1080

Gido B (2000) Queuing models with admission and termination control: monotonicity and threshold results. Doctorial Dissertation, Eindhoven University of Technology, The Netherlands

Healthcare Financial Management Association (2003) Achieving operating room efficiency through process integration. Expert Paper. http://www.hfma.org. Accessed Aug 2010

Institute of Medicine (2001) Crossing the quality chasm: a new health system for the 21st century. National Academy, Washington, DC

Leymann F, Roller D, Schmidt M (2002) Web services and business process management. IBM Syst J 41(2):198

Lusk S, Paley S, Spanyi A (2005) The evolution of business process management as a professional discipline. White Paper. http://www.bptrends.com. Accessed May 2010

Miers D (2005) BPM: driving business performance. White Paper. http://www.bptrends.com. Accessed Oct 2010

Montazar A, Behbahani S (2007) Development of an optimized irrigation system selection model using analytical hierarchy process. Biosyst Eng 98:155–165

Nivarthi S (2004) Enhancing the utilization of multiple suites of OR using a simulation based approach. Masters Thesis, State University of New York, New York

Oddershede A, Arias A, Cancino H (2007) Rural development decision support using the analytic hierarchy process. Mathematical and Computer Modeling 46:1107–1114

Pham D, Klinkert A (2008) Surgical case scheduling as a generalized job shop scheduling problem. Eur J Oper Res 185(3):1011–1025

Ransom S, Joshi M, Nash D (2005) The healthcare quality book: vision, strategy and tools. Health Administration, Chicago

René D, Hassane A (2005) Discrete, continuous, and hybrid petri nets. Springer, Berlin

Shaw D, Holland P, Kawalek P, Snowdon B, Warboys B (2007) Elements of a business process management system. Bus Process Manag J 13(1):91–107

Smith H, Fingar P (2003) Workflow is just a pi process. Computer Sciences Corporation. http://www.bpm3.com/picalclus. Accessed Aug 2010

Su J, Chen S, Lin L (2003) A structured approach to measuring functional dependency and sequencing of coupled tasks in engineering design. Comput Ind Eng 45:195–214

Suh P (1990) The principles of design. Oxford, New York

van der Aalst W (1998) The application of petri nets to workflow management. J Circuit Syst Comput 8(1):21–66

van der Aalst W, Van Hee K (2002) Workflow management: models, methods and systems. MIT, Cambridge

Yang K, El-Haik B (2003) Design for six sigma: a roadmap for product development. McGraw-Hill, New York

zur Muehlen M (2004) Workflow-based process controlling: foundation, design and application of workflow-driven process information systems. Logos, Berlin

Chapter 5
Examining Failure in a Dynamic Decision Environment: Strategies for Treating Patients with a Chronic Disease

Gregory W. Ramsey[1], Paul E. Johnson[2], Patrick J. O'Connor[3], JoAnn M. Sperl-Hillen[3], William A. Rush[3], and George Biltz[4]

[1]The Earl G. Graves School of Business and Management, Morgan State University, Baltimore, MD, USA

[2]Carlson School of Management, University of Minnesota, Minneapolis, MN, USA

[3]HealthPartners Institute for Education and Research, Minneapolis, MN, USA

[4]School of Kinesiology, University of Minnesota, Minneapolis, MN, USA

Abstract In this paper we investigate the dynamic decision-making task of primary care physicians treating patients with type 2 diabetes to achieve a blood glucose goal. The focus of the study is on developing and testing an information processing theory that can explain why some physicians more often succeed and others more often fail to achieve desirable clinical goals. The developed theory is represented in the form of two types of computational models, one employing a feedback decision-making strategy and the other a feedforward strategy. The models were implemented in software and tested using data from a previously reported experiment where physicians treated simulated patients with type 2 diabetes. The physician data were scored for a defined set of treatment errors. Computational processes were systematically examined to identify and specify processes to perturb in order to generate the observed errors. Models were created for each physician by introducing perturbations in computational processes based on errors that each physician committed during the experiment. These models treated the same simulated patients that the physicians treated; results from each model treating the patients were compared with the represented physician's results to test the sufficiency of the models to explain observed errors. Process perturbations which explained observed errors took two characteristic forms, both of which resulted in delayed treatment action: (1) elevated thresholds for triggering action and (2) overestimating delayed effects

G.W. Ramsey (✉)
The Earl G. Graves School of Business and Management, Morgan State University, Baltimore, MD, USA
e-mail: gregory.ramsey@morgan.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*, Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_5

of medications. Physician models made predictions for types and timing of subjects' treatment errors: physician models generated 79 % of the same types of treatment errors as committed by physicians. As demonstrated by this study, developing task specific information processing theories (expressed as computational models) are useful for investigating patterns of decision making that lead to errors of performance. Studies of this nature can support the design of decision support systems intended to reduce errors associated with dynamic tasks, such as treating a chronic disease.

Keywords Computational models • Physician decision making

5.1 Introduction

Some decision agents when performing dynamic tasks are able to achieve an intended goal but many fail to reach that goal. This paper examines the decision processes that primary care physicians (i.e., general internists and family practitioners) use to treat patients with the chronic disease of type 2 diabetes to achieve a blood glucose goal. Treating a chronic disease is an instance of a dynamic decision task. Such tasks have four distinguishing characteristics (Edwards 1961; Brehmer 1990, 1992): (1) a series of decisions must be made over the course of time to reach a goal, (2) the environment changes because of agent actions as well as independent of agent actions, (3) feedback on actions taken is delayed, and (4) prior actions constrain later actions.

Treating type 2 diabetes is considered a dynamic task because a patient's condition changes over time as the disease progresses. Additionally, effects of treatments rendered (or lack of treatment) can have delayed effects on a patient's condition; consequently treatments have to be adjusted over time for a patient's changing needs. The focus of this paper is on identifying and modeling physicians' clinical decision-making processes and resulting actions to explain why some physicians fail to attain a blood glucose goal. The developed models represent a theory for how information is processed to achieve a particular goal in a dynamic environment and reasons why those goals may not be reached.

The physician as a decision agent is tasked with taking action to move toward a goal while responding to changes in the environment (i.e., the patient being treated). Several factors make treating these patients a non-trivial task, two such factors are: (1) accounting for delayed effects of previously prescribed medications, failure to do this can result in patients reaching perilous states of health, for example, experiencing dangerously low blood glucose levels (hypoglycemia), and (2) detecting and treating issues that impair a patient's adherence with treatment orders, failure to do this can make prescribed treatments ineffective.

Thirty to sixty percent of type 2 diabetes patients who have been treated by primary care physicians for at least 2 years are not at goal (Stanton 2001; Karter et al. 2007, www.mncm.org). The consequences of patients not reaching evidence-based blood glucose goals are increased risks for developing diabetes-related complications such as renal failure, amputations, blindness, heart attacks, or stroke, which impair quality of life and result in potentially avoidable healthcare costs (American Diabetes Association 2008; Holman et al. 2008). This study uses process control as a paradigm to show that when decision agents over-estimate delayed effects of actions taken, elevate thresholds for triggering actions (i.e., delaying action), or both, then patients fail to reach clinical goals.

This paper is organized in the following manner. First, a dataset from an experiment where physicians treated simulated patients with type 2 diabetes is presented; this dataset is the focus of investigation. Second, a brief background on the treatment of type 2 diabetes is provided to establish a context for the study. Third, a methodology is presented for creating and testing computational models for a group of individual physician's treatment actions. Fourth, results of the modeling effort are presented in two parts: (1) models representing decision processes for the idealized treatment of type 2 diabetes, and (2) models of individual subject physicians created by introducing perturbations in idealized treatment processes. Finally, conclusions are discussed.

5.2 Dataset

In a published study conducted by O'Connor et al. (2009) 19 primary care physicians each treated three simulated patients with type 2 diabetes. Physicians were tasked with treating these patients to bring them to clinical goals for blood glucose, blood pressure, and lipids. For this investigation we focus exclusively on control of blood glucose levels, where control can be assessed using a blood test. The test measures glycated hemoglobin, abbreviated A1c; patients with A1c levels of 7 % or lower are deemed to have reached an evidence-based goal for blood glucose control (Mazze et al. 2005).

The three patients selected for the simulation cases each represented a type of patient that is typically encountered in the clinic by primary care physicians: (Case 1) a patient not taking any medication, initial A1c of 9.5 %, requiring initiation and titration of oral medications; (Case 2) a patient taking oral and insulin medications, initial A1c of 10.5 %, exhibiting symptoms of depression and requiring adjustments to prescribed medications; and (Case 3) a patient taking medication, initial A1c of 9.8 %, requiring the initiation/titration of insulin. A record of the physicians' treatment actions and patient responses are the dataset for which computational models were developed. Each physician treated the same three simulated cases using an electronic medical record-like interface. The term *titration* as used refers to the process of progressively increasing medication doses in a step-wise manner until a desired clinical effect is achieved (Nathan et al. 2009).

Figure 5.1 below shows the final A1c values that each physician achieved with each patient. The numbers of visits that physicians treated the patients are indicated.

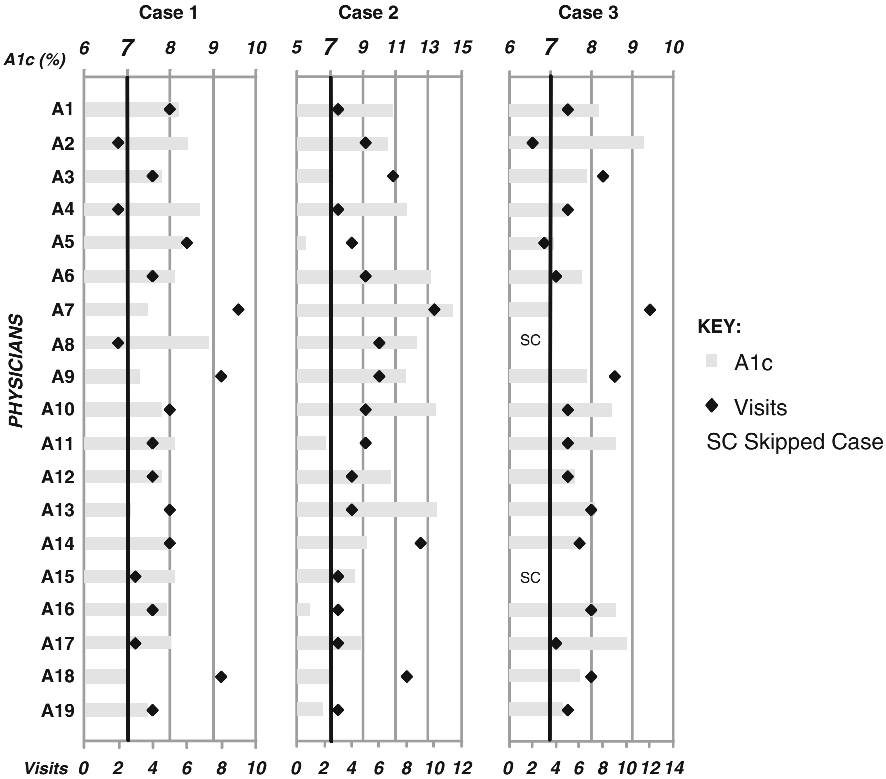


Fig. 5.1 A1c outcomes obtained by each physician on the three cases. *Horizontal bars* represent final A1c values (*top scale*), *diamonds* represent number of visits physicians used to treat the patient (*bottom scale*). *Horizontal bars* ending to the *left* of the *bold vertical lines* indicate physicians who achieved an evidence-based A1c goal level

On Cases 1, 2, and 3 relatively few (2, 5, and 2 respectively) physicians reached the A1c goal. How and why do some physicians reach goal while most do not? Answering this question is the objective of this study. We investigate decision processes that lead to success, where success is defined as a decision agent taking actions that result in obtaining a clinical goal without violating the norms of treatment. Once an understanding of how success is achieved then attention is focused on identifying and explaining the causes for failures in treatment. In this investigation failures are identified and measured in terms of physicians committing errors in treatment.

5.3 Background

Type 2 diabetes is a chronic disease characterized by elevated blood glucose level and related metabolic perturbations such as cholesterol (lipid) abnormalities and elevated blood pressure (BP) that usually require lifestyle and pharmacologic treatment (Gale 2006). The treatment can take the form of a controlled diet, exercise, medication, or some combination of these. Currently the disease cannot be cured and most patients require on-going pharmacologic treatments to control blood glucose levels (Hunt and Arar 2001). Control of blood glucose is assessed by measuring glycated hemoglobin (A1c) values. A1c values are a moving average of a patient's blood glucose values over the previous 8–12 week period (Guven et al. 2005).

Physicians who treat chronic diseases engage in a process control task where they cycle through the following steps over the course of time: obtain information on the patient's state, make decisions about actions to take, take action, observe changes in patient state, and decide on next actions based on observed state and history of actions taken. In this study we restrict our investigation to the treatment of blood glucose only. Blood glucose is difficult to treat in part due to the delay before effects of treatments can be measured in A1c (commonly referred to as dose response times), whereas treatments for BP and lipids rapidly affect patient state measures (Mazze et al. 2005).

Because multiple decisions are required to reach a goal in a dynamic environment the process of decision making has been likened to gaining control over a system (Bainbridge 1981, 1997; Broadbent et al. 1986; Brehmer 1990). Brehmer (1990) proposes control theory as a means for describing these decision-making processes. Within a control theoretic framework people use one of two types of decision strategies for achieving goals in dynamic environments (Brehmer 1990). These strategies are feedback and feedforward. A feedback strategy examines the difference between a current state and a desired state and then makes decisions to move closer to the goal (Johansson 2003). A feedforward strategy uses knowledge of how sequences of actions over time can affect the system to anticipate changes resulting from past and current actions, and then makes decisions to move toward the goal (Johansson 2003).

Conant and Ashby (1970) established that for a regulator (decision maker) to control a system it needs a model of the system, this is known as the model principle. Brehmer (1990) applies the model principle and concludes that a decision strategy has as part of it a mental model of the environment, where the mental model is a representation of important aspects of the system being controlled (Doyle and Ford 1998). Mental models have been functionally defined as mechanisms for describing, explaining, and predicting anticipated states of a system (Rouse and Morris 1986).

Brehmer (1990) suggests that feedback and feedforward strategies require different types of mental models. A feedback strategy uses a relatively simple mental model for making decisions, whereas a feedforward strategy requires a more complex mental model to make decisions to reach expected future states. The complexity of a mental model is relative to types of decisions being made.

A feedback strategy does not require an elaborate mental model because in its simplest form it only has to gauge if decisions are moving the system toward the goal. On the other hand a feedforward strategy requires a more elaborate mental model to anticipate future states of the system so that decisions can be made to reach some expected future state (Brehmer 1992). To anticipate future states requires that the mental model account for time-dependent properties of the system (Sterman 1989; Brehmer 1990, 1992), such as time delays between actions taken and effects showing in the system being controlled, and how a series of actions taken over time affects the system.

Feedback strategies are best fit to environments where delays in feedback are minimal. When delayed feedback is used to make decisions without accounting for delays then the decision agent will make decisions that are relevant for past states of the system (Brehmer 1992). Feedforward strategies are best fit to stable environments, where stability is defined as relationships between dynamic aspects of the environment are known and those relationships do not change. When the environment is not stable then future states of the system cannot be estimated reliably (Brehmer 1992).

Feedback and feedforward strategies are typically represented as closed-loop systems where some portion of the output is fed back to the input (Leigh 1997). These closed-loop systems compare the current state of the system with a goal state; the difference between these two states forms an error signal. The closed-loop system takes action in a cyclical manner to minimize the error signal, that is, take action to reduce error signal, check error signal, take additional action to further reduce error signal (Leigh 1997). Feedback and feedforward closed-loop systems have been expressed using various frameworks from control theory, such as the internal model control, model predictive controller, or smith predictor (Garcia and Morari 1982; Seborg et al. 1986; Garcia et al. 1989; Lee et al. 1996; Ungar et al. 1996; Leigh 1997; Lin and Su 2000).

The specific representation of interest in the present study is the internal model control framework (Garcia and Morari 1982; Ungar et al. 1996). This representation segments the control problem into two major parts that are solved by different components within the system (Jordan and Wolpert 2004). These components are an inverse model and a forward model. The inverse model is so named because it solves an inverse problem where the order of reasoning is reversed, that is, for a given state of the system what is the action that is required to move the system toward the goal (Garcia and Morari 1982; Cheney 1997). The forward model is so named because it solves a forward problem, which is for a given action what is the expected state of the system after the action takes effect (Karniel 2002; Jordan and Wolpert 2004). In the realm of decision making in dynamic environments the forward model is the same as the mental model of the system being controlled (Brehmer 1990).

Most people when faced with the task of controlling a process in a dynamic environment tend to under-control the process (Martin et al. 2004). For the treatment of type 2 diabetes under-treatment (i.e., lack of control) is considered a failure in treatment.

5.4 Methodology

Modeling Process. The precursor to modeling errors within computational processes requires defining processes that generate actions for successfully treating patients to reach a blood glucose goal. These processes were defined via a task analysis for the task of treating type 2 diabetes to reach an evidence-based goal. A task analysis entails determining what a decision agent is required to do to achieve a specific goal for a given task (Kirwan and Ainsworth 1991). The analysis was performed from the perspective of a primary care physician being tasked with bringing a patient to a blood glucose goal. The knowledge base for the task was obtained from guidelines for treating type 2 diabetes and from primary care physicians who were expert in treating type 2 diabetes.

Using evidence-based guidelines as a norm, a set of errors were defined for identifying where physicians did not take treatment actions that were consistent with those of an idealized physician. Physicians' data were scored for these errors. Physicians were scored on each case as using either a feedback or feedforward strategy. The default strategy classification was feedback (Brehmer 1992), physicians were reclassified as using feedforward if they took anticipatory action such as prescribing medications without waiting for prior dose changes to take full effect. Processes within the computational models were examined to create a mapping between each error and a computational perturbation that could generate an observed error. Based on the error mapping and strategy scoring, individual models were particularized for each physician by activating process perturbations within a strategy appropriate computational model. The particularized models were run to test whether the models generated the same errors under the same conditions as those generated by the modeled physicians.

Assessing Goodness of Fit. A metric was defined for evaluating the goodness of fit between a physician's behavior and the model of his/her behavior. Physician models are used to explain treatment paths, that is, generate similar behavior (moves made and errors committed) observed at specific points during a course of treatment.

Physicians were evaluated on each visit during the simulation experiment to determine errors committed and treatment moves made. Physicians' treatment moves and errors were encoded as strings of sequentially concatenated codes representing moves and errors made while completing a case.[1] Similarly the errors and moves made by the physician models were encoded as strings, one string representing errors made on a case and another string representing treatment moves made on the case.

Differences between patterns of treatment moves and errors generated by physicians and physician models were measured using string distances. Measuring string distances is common in fields such as genomics and computational linguistics

[1]For error strings, visits without errors receive a code indicating that no error occurred; this was to enable measuring differences in the timing and quantity of errors committed.

for assessing similarity between symbolic patterns. Encoded strings for physicians and physician models were compared to measure the similarity of the patterns in order to assess the goodness of model fit. String distance measures enabled assessing partial matches of errors committed by a physician and her representative model for a given visit, as opposed to using a courser measure of either errors on a visit were replicated or not.

Differences between patterns were computed as Levenshtein string distances (Navarro 2001). The Levenshtein distance is a measure of the minimum distance between two strings based on a set of editing primitives (*insertions, deletions, and substitutions*). For two character strings the minimum string distance (MSD) is measured as the minimum number of primitive manipulations to transform one string into the other (Soukoreff and MacKenzie 2001).

For example, if string 1 is ‘ABBD’ and string 2 is ‘ADCBD’ then the minimum string distance between the two strings is 2. Two primitive operations are required to transform the second string into the first string. The first primitive operation is the second character of string 2 is deleted. The second primitive operation is the third character of ‘C’ in string 2 is substituted with the character ‘B.’ By performing these operations the two strings become equivalent.

Soukoreff and MacKenzie (2003) proposed an MSD error rate for measuring differences in a typed string intended to replicate a string of text. The error rate is given below in (5.1):

$$\text{MSD Error Rate} = \frac{\text{MSD (P, T)}}{\max(|\text{P}| \cdot |\text{T}|)} * 100\% \tag{5.1}$$

where P and T are the presented and typed character strings, |•| represents the length of the string. MSD error rate ranges from 0 to 100 %. The MSD Accuracy Rate is defined as:

$$\text{MSD Accuracy Rate} = 100\% - \text{MSD Error Rate} \tag{5.2}$$

MSD accuracy rate is the metric used to evaluate how well physician models’ error and treatment move patterns match those of the physicians’ error and move patterns.

5.5 Results

From a task analysis it was determined that the physician-patient dyad progresses through four stages: (1) assessment of condition, (2) initiation of treatment, (3) adjustment of treatments to reach goal, and (4) maintenance of the goal. Based on these four stages a dynamic decision (computational) model was developed for controlling blood glucose levels in type 2 diabetes patients. The computational

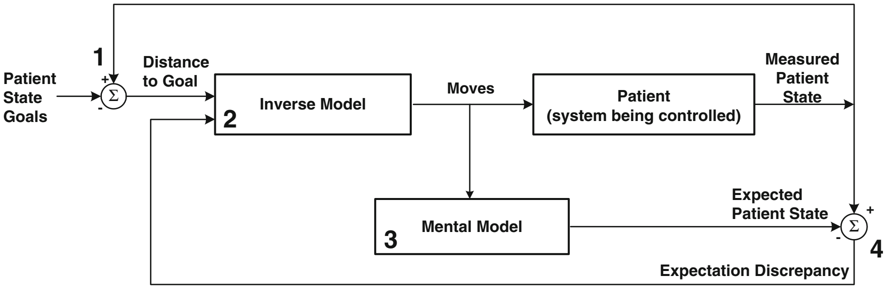


Fig. 5.2 Physician decision model for treating patients with type 2 diabetes

model is represented using a modified internal model control (IMC) framework as shown in Fig. 5.2 (Garcia and Morari 1982).

The physician decision model (as shown in Fig. 5.2) is a dynamic decision model for controlling a type 2 diabetes patient's blood glucose levels. The focus of this modeling activity is on the treatment of blood glucose levels using feedback and feedforward control strategies.

The framework is composed of the following parts. The *inverse model* in Fig. 5.2 solves the inverse problem of determining appropriate actions to take for a given point in time. Computing a solution to this type of problem entails using observations of a system to infer the state of that system, from which actions are determined for moving the system toward a goal. The *patient* with type 2 diabetes is the system being controlled, the patient is the recipient of actions computed in the inverse model (the patient is represented by a patient model that provides clinically appropriate responses for actions taken). The *mental model* is a physician's representation of the type of patient being treated (e.g., adherent or non-adherent to treatment regimens) and expected changes in patient states for actions taken (such as medication treatments). The *expectation discrepancy* is an error signal that denotes the difference between a measured patient state and an expected patient state. Expectation discrepancy is used to determine if treatment moves are changing blood glucose values by expected amounts. The *distance to goal measure* is an error signal that is computed as the difference between a patient's current state and a goal state. Distance to goal provides information to the inverse model for the purpose of determining if treatment moves need to be made.

Inputs to the decision model are patient state goals. Goals are achieved by making moves. There are three types of moves: treatment, information seeking, and scheduling. Treatment moves are actions that affect the patient's blood glucose; the two types of treatment moves are prescribing medications and making referrals to health professionals. Information seeking moves are the ordering of tests for determining patient state and effects of past treatment actions; five types of tests are ordered, which are A1c, plasma glucose (measured as self-monitoring blood glucose, abbreviated SMBG), BP, lipids (LDL), and creatinine tests. Scheduling moves are actions to set when a next visit will occur, i.e., opportunities for making additional moves. All moves are computed by the inverse model using distance to goal and expectation discrepancy as inputs.

The *patient* receives moves (outputs from the inverse model) and responds with changes in A1c accounting for delayed responses of medications. In this study the patient model is composed of a rule-base of equations of dose–response times for simulated medications used to treat the patient (Dutta et al. 2005).

The physician decision model computes decisions that answer two questions (Jordan and Rumelhart 1992; Gibson et al. 1997): (1) for a current state of the system (patient) what are the actions that need to be taken to move toward goal?, and (2) for actions taken what are the expected effects on the system? By the model cycling through computations to answer these questions to make decisions a goal can be achieved and maintained. Decision agents using a feedback strategy wait until effects of past actions have finished taking effect on blood glucose levels. This is computed using a representation of a mental model for maximum times required for medications to take effect as published in dose–response curves. Decision agents using a feedforward strategy compute a rate of progress toward goal and make medication moves to maintain that rate without waiting for effects of past actions to completely show in blood glucose values. A feedforward strategy computes time-dependent effects of medications.

The computational models are stated as a set of functional equations (Newell 1982). There were five types of functional equations defined: (1) goal setting, (2) treatment moves, (3) information seeking moves, (4) scheduling moves, and (5) expected next states. Each equation is expressed as a tabular mapping between inputs and outputs. Below in Table 5.1 are examples of equations for each type of functional equation.

Computational models for feedback and feedforward decision strategies were implemented in software for testing. The models were given the same three simulated cases that the subject physicians treated. Figures 5.3 and 5.4 below show the resulting A1c traces as a function of time when the two computational models treated the three simulated cases. Both feedback and feedforward models brought the patients in the three cases to an evidence-based A1c goal of 7 % within 1 year. These computational models as defined made treatment decisions in the same manner as that of an idealized physician.

5.6 Defining and Modeling Errors

Subjects' process traces that were recorded when they performed each case were analyzed to determine ways in which physicians actions differed from those of the models of idealized treatment. Differences that violated the evidence-based guidelines of Staged Diabetes Management, 2nd edition (Mazze et al. 2005) were

Table 5.1 Samples of each type of functional equation that were used to define physician decision models for treating patients with type 2 diabetes

Type of equation	Input	Processing rule(s)	Output
Goal setting	• Organizational policies	• Given an evidence-based guideline and patient state select a goal. Compare goal with organizational policy, if goal is not consistent with policy then follow policy	Blood glucose goal
	• Patient state		
	• Evidence-based guidelines		
Treatment move	• Current meds and doses	• If a med is contra-indicated exclude that med from set of usable meds	Medication/dose change
	• Distance to goal	• If A1c distance to goal >0.5 then increase dose of current med by 1 tablet size. If current med at maximum dose start next med in set at minimum dose	
	• Contra-indicator variables		
	• Time since last dose change		
Info seeking move	• Days since last test	• If days since last test > min. testing frequency then order test	A1c test order
	• Testing requirements	• If prescribed med has contra-indicator variables then order test to determine if contra-indicator present	
	• Days to next visit		
	• Medications prescribed		
	• Contra-indicator variables		
	• Patient state value last test		
Scheduling move	• Days since last visit	Feedback strategy:	Days to next visit
	• Time since last dose change	• Days to next visit = Estimated time required for meds to show full effect on A1c values [a constant value, say 90 days]	
	• Patient state	Feedforward strategy:	
	• Medications prescribed	• Days to next visit = Estimated time before next dose increase required to maintain a constant rate of progress toward the goal	
	• Expectation discrepancies		
Expected next state	• Current BG value	Feedback strategy:	Expected change in blood glucose level by next visit
	• Adherence type (low or high)	• (Current BG value) *minus* (Expected Change in BG for medications given)	
		Feedforward strategy:	
		• (Current BG value) *minus* (Expected change in BG as a function of time for medications given)	

The abbreviation *med* is used in place of medication

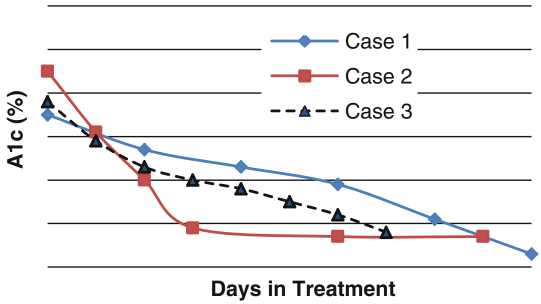


Fig. 5.3 Feedback models treating cases 1–3

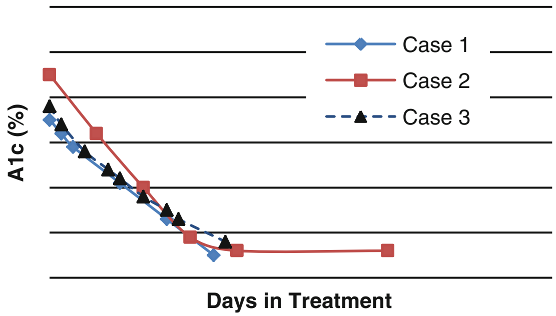


Fig. 5.4 Feedforward models treating cases 1–3

scored as errors.[2] Two types of errors were defined: errors of *omission*, not taking an action that should be taken considering a patient's current state; and errors of *commission*, taking action that should not be taken considering a patient's current state. Errors of omission were labeled *no move* errors; errors of commission (specifically prescribing contraindicated medications) were labeled *improper use of meds* errors.

Physician traces were grouped according to A1c levels. A1c levels above 8 % were considered to require physician intervention to move the patient toward the goal (Beaton et al. 2004). A1c levels between 7.5 and 8 % were close to goal and ideally should be acted upon to reach the goal. A1c levels of less than 7.5 % were considered to be within the goal range. Table 5.2 presents a summary of the findings. On Case 1 where physicians were using oral medications *no move* errors were committed mostly when A1c levels were below 8 %. On Case 2 for all A1c levels primarily *improper use of meds* errors were committed. On Case 3 both *no move* errors and *improper use of meds* errors were committed.

[2]The rules for error scoring were defined and programmatically applied to the data set. A check for agreement of the scoring rules was made by randomly selecting and hand scoring process traces from the data set. Results of the hand scoring were compared with the errors scored programmatically and determined to match in every case.

Table 5.2 Proportion (count) of physicians that committed the different types of errors

Error case	A1c range	No. phys	Class of errors	
			No move	Improper use of meds
Case 1	A1c >8 %	8	0.13 (1)	0 (0)
	A1c between 8 and 7.5 %	6	0.67 (4)	0 (0)
	A1c <= 7.5 %	5	0.8 (4)	0.2 (1)
Case 2	A1c > 8 %	13	0.08 (1)	0.54 (7)
	A1c <= 7.5 %	6	0 (0)	0.67 (4)
Case 3	A1c > 8 %	7	0.29 (2)	0.71 (5)
	A1c between 8 and 7.5 %	6	0.5 (3)	0.33 (2)
	A1c <= 7.5 %	4	0.25 (1)	0.75 (3)

Errors of omission were modeled by perturbing dose effect times within the representation of the mental model. Errors of commission were modeled by perturbing titration processes.

Subject physicians were divided into two groups using a stratified sampling approach: (1) a modeling set for refining the form of the perturbations and (2) a hold-out set for testing how well errors mapped to perturbations could predict individual physician performance for controlling A1c. Models were evaluated by computing minimum string distance accuracy rates for each physician-model pair. Tables 5.3, 5.4, and 5.5 display the results of running individual physician models on the three cases.

Case 1: Six physicians' treatment traces were selected for modeling, three used feedback (FB) and three used feedforward (FF) strategies while treating the case. The physician models were able to replicate the pattern of errors generated by the physicians when treating Case 1. Physicians A6 and A12 did not commit errors when treating the case, the models for these physicians did not commit errors however the models matched the moves of the physicians with 75 % MSD accuracy. These physicians treated the patients in a manner that differed from the specified idealized form of treatment, these variations in treatment moves were not classified as errors.

Thirteen physicians' treatment traces were reserved for the hold-out set, eight used FB and five used FF strategies while treating the case. Four of the FF physician models and four of the FB physician models replicated the error patterns of the modeled physicians. Three FB physician models did not generate the errors as the modeled physicians. One FB and one FF physician models generated the same errors but on different visits from those where the modeled physicians committed the errors.

On this case a total of 19 errors were committed: (a) 15 errors of omission and 1 error of commission were committed by physicians who used a FB strategy, (b) 3 errors of omission were committed by physicians who used a FF strategy. The FB physician models correctly generated 9 (60 %) of the omission errors and 2 (67 %) of the commission errors. The FF physician model correctly generated the single

Table 5.3 Summary of modeling outcomes for Case 1

Subject	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
Strategy (FB/FF)	FF	FB	FB	FF	FB	FB	FB	FB	FF	FB	FF	FF	FB	FF	FB	FB	FF	FF	FB
Final A1c (%)	8.2	8.4	7.8	8.7	8.3	8.1	7.5	8.9	7.3	7.8	8.1	7.8	7.1	8	8.1	7.9	8	7	7.5
Model A1c (%)	7.7	8.3	7.8	8.5	8.2	8	7.9	8.9	7.5	7.9	8.1	8.1	7.2	8	8.1	8.4	7.9	6.7	7.4
Visits with patient	5	2	4	2	6	4	9	2	8	5	4	4	5	5	3	4	3	8	4
Days in treatment	42	60	90	30	360	172	309	10	120	300	56	56	284	125	88	180	88	91	126
No of errors—subject	0	0	2	0	3	0	6	0	1	2	0	0	1	1	0	1	0	1	1
Type of errors	–	–	O	–	O	–	O	–	O	O	–	–	O	O	–	O	–	O	C
MSD—errors	100	100	100	100	33	100	33	100	63	100	100	100	80	100	100	20	100	100	100
MSD—moves	60	100	100	50	50	75	78	100	75	100	50	75	100	100	100	50	100	100	100

Treatment actions of subject physicians were compared with actions of their respective models on a per visit basis. Highlighted subjects were part of the modeling set; all others were in the hold-out set. MSD is an abbreviation for minimum string distance for patterns compared from physician and model. Type of errors: O = error of omission, C = error of commission

Table 5.4 Summary of modeling outcomes for Case 2

Subject	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
Strategy (FB/FF)	FF	FB	FB	[a]	FB	[a]	[a]	[a]	[a]	[a]	FB	FF	FB	FB	FB	FB	FF	FF	FB
Final A1c (%)	10.9	10.5	6.9		5.5						6.7	10.7	13.5	9.2	8.5	5.8	8.9	7.2	6.6
Model A1c (%)	10.8	8.8	6.9		6.4						6.8	11.1	12	9.5	8.5	7.5	8.8	7.4	6.8
Visits with patient	3	5	7		4						5	4	4	9	3	3	3	8	3
Days in treatment	21	56	186		222						134	28	150	344	44	120	37	154	112
No of errors—subject	3	5	7		4						5	0	0	11	0	3	1	0	0
Type of errors	C	C	C		C						C	–	–	O,C	–	C	C	–	–
MSD—errors	33	100	100		100						100	100	100	100	100	33	100	100	100
MSD—moves	33	80	71		50						60	75	75	67	33	67	67	0	33

Treatment actions of subject physicians were compared with actions of their respective models on a per visit basis. Highlighted subjects were part of the modeling set; all others were in the hold-out set. MSD is an abbreviation for minimum string distance for patterns compared from physician and model. Type of errors: O = error of omission, C = error of commission

[a]Physician did not treat depression and was unable to control blood glucose levels

Table 5.5 Summary of modeling outcomes for Case 3

Subject	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
Strategy (FB/FF)	FF	FB	FB	FB	FB	FB	FB	Skipped	FB	FB	FF	FB	FB	FB	Skipped	FB	FF	FB	FB
Final A1c (%)	8.2	9.3	7.9	7.4	7.1	7.8	7		7.9	8.5	8.6	7.6	8.1	7.7		8.6	8.9	7.7	7.3
Model A1c (%)	7.9	9.1	6.8	7	7.7	9.1	6.8		8.3	8.6	8.8	8.1	8	7.5		7.4	9.3	8.3	7.2
Visits with patient	5	2	8	5	3	4	12		9	5	5	5	7	6		7	4	7	5
Days in treatment	91	60	128	270	120	174	664		194	178	95	72	326	247		264	44	105	154
No of errors—subject	9	0	0	5	4	1	0		11	3	8	5	0	0		13	4	1	6
Type of errors	C	–	–	C	C	O	–	–	O,C	C	O,C	C	–	–	–	O,C	C	O	O,C
MSD—errors	67	100	100	60	100	100	100	–	71	60	57	100	100	100	–	14	100	100	80
MSD—moves	60	100	100	40	100	100	50	–	56	60	100	100	86	83	–	29	100	100	80

Treatment actions of subject physicians were compared with actions of their respective models on a per visit basis. Highlighted subjects were part of the modeling set; all others were in the hold-out set. MSD is an abbreviation for minimum string distance for patterns compared from physician and model. Type of errors: O = error of omission, C = error of commission

commission error that was committed. Table 5.3 above presents a summary of the results.

Case 2: The model of the patient used in the simulation experiments was created to account for a patient's level of adherence to prescribed treatment regimens. Within the model of the patient depression was represented in a manner so that depression reduced the adherence level. Case 2 contained a depressed patient that at the beginning of the case had prescribed medications which were not being taken. Six physicians did not diagnose and treat depression; as a result treatment moves were ineffective in reducing blood glucose levels because the patient did not take prescribed medications. These physicians when performing this case took actions that were inconsistent with modeled actions and appeared to resemble a trial and error method of treatment. Since these physicians were not following modeled processes they were excluded from the analysis set.

Five physicians were selected for modeling, three used FB and two used FF strategies on this case. The three FB physician models replicated the error patterns of the modeled physicians. One of the FF physician models generated errors in the same pattern as the modeled physician and the other FF physician model generated the same errors but the visits on which the errors occurred differed from those of the modeled physician. Case 2 is more complicated to model because physicians have two different types of treatment moves that they can make on a single visit, which are oral medication and insulin. While the physicians models were able to generate errors resembling those of physicians, the treatment moves differed between physician and modeled physician.

Eight physicians' treatment traces were reserved for the hold-out set, six used FB and two used FF strategies while treating the case. Five of the FB physician models and two of the FF physician models replicated the error patterns of the modeled physicians. One FB physician model generated the same errors but on different visits from those where the modeled physician committed the errors.

On this case a total of 38 errors were committed: (a) 2 errors of omission and 33 errors of commission were committed by physicians who used a FB strategy, (b) 3 errors of commission were committed by physicians who used a FF strategy. The FB physician models correctly generated 2 (100 %) of the omission errors and 28 (85 %) of the commission errors. The FF physician model correctly generated 2 (67 %) commission errors that were committed. Table 5.4 above presents a summary of the results.

Case 3: Five physicians' treatment traces were selected for modeling, three used feedback (FB) and two used feedforward strategies while treating the case. Two FB and 1 FF physician models were able to replicate the pattern of errors generated by the physicians when treating Case 3. Models for physicians A4 (FB) and A11 (FF) committed the same errors as the physicians but they occurred on visits different from those where the physicians committed them, this resulted in MSD accuracy rates of approximately 60 %. These physicians treated the patients in a manner that differed from the specified idealized form of treatment. The physicians when performing these cases followed treatment processes that differed from those associated with the modeled idealized treatment processes, these variations in treatment moves were not classified as errors and could not be replicated by the physician models.

Two physicians did not complete Case 3. Twelve physicians' treatment traces were reserved for the hold-out set, 11 used FB and 1 used FF strategies while treating the case. Seven of the FB physician models replicated the error patterns of the modeled physicians. One FF and two FB physician models generated the same errors but on different visits from those where the modeled physician committed the errors. Two FB physician models did not generate the same errors as the modeled physicians.

On this case a total of 70 errors were committed: (a) 10 errors of omission and 39 errors of commission were committed by physicians who used a FB strategy, (b) 1 error of omission and 20 errors of commission were committed by physicians who used a FF strategy. The FB physician models correctly generated 5 (50 %) of the omission errors and 36 (92 %) of the commission errors. The FF physician models correctly generated one omission error and 14 (70 %) of the commission errors. Table 5.5 above presents a summary of the results.

5.7 Analysis of Physicians Across Cases

We examined the results of the strategy classification and error scoring to better understand interactions amongst strategy selection on a given case, errors committed, and final A1c outcomes. Cases 1 through 3 were selected from different areas of the possible space of type 2 diabetes patient types, consequently variation was expected in strategy selection, errors committed, and A1c outcomes for individual physicians across cases. The 17 physicians that completed all three cases committed at least one error on one of the cases, varied in the deployment of strategies across cases, and had varying degrees of success in reducing A1c values.

A task physicians faced was matching a decision-making strategy to a case. Physicians were grouped by the combination of strategies each physician deployed across the three cases. Physicians A1 and A17 used FF strategies on all three cases. These two physicians committed errors of commission, but did not commit errors of omission. Physicians A2, A3, A5, A13, A16, and A19 used FB strategies on all three cases. These physicians all committed errors of omission, and all but one (A13) committed errors of commission. The remaining physicians used some combination of FF and FB strategies on the cases, and committed some combination of errors of omission, commission, or both.

Physicians A3, A18, and A19 were consistently able to bring patient A1c values to 8 % or lower on all three cases. Physicians A3 and A19, who used FB strategies for all cases, committed 9 and 7 errors in total for the cases; whereas A18, who used a combination of strategies on the cases, committed a total of 2 errors across the cases.

Physicians were able to bring patients to goal using FB or FF strategies. A distinguishing feature of the two strategies was the amount of time required to reach the goal. Those physicians who used FB strategies required more time to achieve similar results obtained by physicians who used FF strategies. The analysis of the physicians' actions suggests that clinical goals can be reached by applying a single decision making strategy regardless of patient type, but a single strategy may not be appropriate for achieving goals with minimal errors for those patient types (i.e., achieving an appropriate outcome using inappropriate means). Physicians who used only FB strategies were more likely to under-treat patients (committing errors of omission) compared to physicians who used FF strategies. Physicians who used only FF strategies were more likely to use inappropriate medication regimens (committing errors of commission) than those who used a combination of strategies.

5.8 Discussion and Conclusion

The three cases that were examined in this study constitute a fair sampling from the range of type 2 diabetes patients that a physician (decision agent) may encounter in a clinical setting. These decision agents are faced with the task of determining what portion of the treatment problem space a patient falls into and what are the appropriate actions to take to move the patient toward a blood glucose goal. The explanation, in the form of physician models, for why some decision agents fail to achieve a goal in a dynamic environment accounted for: (1) 81 out of 100 (81 %) errors committed by physicians who used a feedback decision strategy, and (2) 19 out of 27 (70 %) errors committed by physicians who used a feedforward decision strategy. Overall the models were able to generate 79 % of the observed errors.

To model omission errors dose response times of medications represented within the mental model were extended by 150–200 % of the times used in the computational models that delivered care in an idealized manner. The decision for how much to extend the response time of medications was determined as a function of blood glucose level, that is, response times were increased as blood glucose levels decreased. Extending dose response times resulted in physician models committing errors of under-treating the patient cases.

To model commission errors threshold values for taking action were increased by double. These threshold values were used to compute when to cease the use of selected types of medications because of a patient's health state. The elevated threshold values resulted in the physician models prescribing medications that were inappropriate for a patient's condition, such actions could adversely affect a patient.

Each case presented challenges for modeling and for the physicians who performed them. A balancing act was performed to prevent over-fitting processes for generating errors. Case 1 necessitated the use of oral medications which have delays before results of actions can be observed. On Case 1 95 % of the errors committed were errors of omission. This suggests that uncertainty from delay of feedback from past actions causes people to delay taking additional actions, particularly when those actions could cause harm to the system (patient). The physician models were able to generate the right types of errors, but these models were only able to generate the errors on the same visits where the physicians committed them in 11 out of

18 (61 %) instances. Determining that an agent will delay action is easier than determining specifically when that action will be delayed.

Case 2, regardless of the decision strategy used, required acting on a condition that was preventing treatment actions from being effective, i.e., diagnosing and treating depression. The physician models were able to generate the two omission errors on the same visits where the physicians committed them. On this case 95 % of the errors committed were errors of commission. This was an easier modeling task because once a physician committed an error of commission she or he persisted with making the error across remaining visits with the patient (hence raising a threshold for initiating this action was an effective way to model the error).

Case 3 was a more difficult case to model because there were fewer constraints on treatment actions that physicians could take, specifically there were fewer restrictions on dosing when insulin was prescribed. Insulin is dispensed in multiples of single units, computing doses that a given physician would prescribe was more difficult compared to computing the dose that a physician could prescribe for oral medications (which are prescribed in multiples of fixed tablet sizes). Prescribing insulin is riskier because there is higher potential for causing unsafe patient states, as a result some physician do not use it when it is clinically appropriate. This case required the use of insulin, decision agents that used any other type of treatment committed errors of commission. The physician models were able to reproduce 85 % of the errors associated with physicians who failed to use insulin. It was more difficult to generate errors of omission that occurred on the same visits where physicians committed them; the models correctly predicted these errors in 55 % of the instances where they were committed.

Modeling studies such as this one help to further our understanding of decision making processes and the types of support that decision makers may need in order to make better decisions. Computational modeling is a useful method for developing and understanding information processing problems. As demonstrated in this study, we use computational models as a means of identifying why failures were occurring in a specific dynamic decision task. Using the developed models, we tested mechanisms for generating observed behavior, including error-laden behavior. A conclusion drawn from the context studied is primary care physicians who are treating type 2 diabetes would benefit by having decision support to indicate time dependent dose effects for prescribed medications.

References

American Diabetes Association (2008) Economic costs of diabetes in the U.S. in 2007. Diabetes Care 31(3):596–615

Bainbridge L (1981) Mathematical equations or processing routines? In: Rasmussen J, Rouse WB (eds) Human detection and diagnosis of system failures. Plenum, New York, pp 259–286

Bainbridge L (1997) The change in concepts needed to account for human behavior in complex dynamic tasks. IEEE Trans Syst Man Cybern Part A Syst Humans 27(3):351–359

Beaton SJ, Nag SS et al (2004) Adequacy of glycemic, lipid, and blood pressure management for patients with diabetes in a managed care setting. Diabetes Care 27(3):694–698
Brehmer B (1990) Strategies in real-time dynamic decision making. In: Hogarth R (ed) Insights in decision making: a tribute to Hillel J. Einhorn. University of Chicago Press, Chicago
Brehmer B (1992) Dynamic decision-making—human control of complex-systems. Acta Psychol 81(3):211–241
Broadbent D, Fitzgerald P et al (1986) Implicit and explicit knowledge in the control of complex systems. Br J Psychol 77:33–50
Cheney M (1997) Inverse boundary-value problems; from oil prospecting to medicine, the science of remote sensing benefits from interaction between mathematicians and computers. Am Scientist 85:448–455
Conant R, Ashby WR (1970) Every good regulator of a system must be a model of that system. Int J Syst Sci 1(2):89–97
Doyle JK, Ford DN (1998) Mental models concepts for system dynamics research. Syst Dyn Rev 14(1):3–29
Dutta P, Biltz GR et al (2005) SimCare: a model for studying physician decision making activity. In: Henriksen K, Battle J, Marks E, Lewin D (eds) Advances in patient safety: from research to implementation: programs, tools, and products, vol 4. Agency for Healthcare Research and Quality, Rockville, pp 179–192
Edwards W (1961) Behavioral decision theory. Ann Rev Psychol 12:473–498
Gale EAM (2006) Declassifying diabetes. Diabetologia 49:1989–1995
Garcia C, Morari M (1982) Internal model control. 1. A unifying review and some new results. Ind Eng Chem Process Des Dev 21:308–323
Garcia C, Prett D et al (1989) Model predictive control: theory and practice—a survey. Automatica 25(3):335–348
Gibson FP, Fichman M et al (1997) Learning in dynamic decision tasks: computational model and empirical evidence. Organ Behav Human Decis Process 71(1):1–35
Guven S, Kuenzi J et al (2005) Diabetes mellitus and the metabolic syndrome. In: Porth CM (ed) Pathophysiology: concepts of altered health states. Lippincott Williams & Wilkins, Philadelphia
Holman R, Paul S et al (2008) 10-year follow-up of intensive glucose control in type 2 diabetes. N Engl J Med 359(15):1577–1589
Hunt LM, Arar NH (2001) An analytical framework for contrasting patient and provider views of the process of chronic disease management. Med Anthropol Q 15(3):347–367
Johansson B (2003) Feedforward control in dynamic situations. PhD thesis. Linkoping University
Jordan M, Rumelhart D (1992) Forward models: supervised learning with a distal teacher. Cognit Sci 16:307–354
Jordan M, Wolpert D (2004) Computational motor control. In: Gazzaniga M (ed) The cognitive neuroscience III. MIT, Cambridge
Karniel A (2002) Three creatures named 'forward model'. Neural Netw 15:305–307
Karter A, Moffet H et al (2007) Glycemic response to newly initiated diabetes therapies. Am J Manag Care 13:598–606
Kirwan B, Ainsworth LK (1991) A guide to task analysis: the task analysis working group. Taylor & Francis, London
Lee TH, Wang QG et al (1996) Robust smith-predictor controller for uncertain delay systems. AIChE J 42(4):1033–1040
Leigh JR (1997) Control theory: a guided tour. Peter Peregrinus, London
Lin C-L, Su H-W (2000) Intelligent control theory in guidance and control system design: an overview. Proc Natl Sci Counc ROC(A) 24(1):15–30
Martin M, Gonzalez C et al (2004) Learning to make decisions in dynamic environments: ACT-R plays the beer game. In: Sixth international conference on cognitive modeling. Carnegie Mellon University/University of Pittsburgh, Pittsburgh
Mazze R, Strock ES et al (2005) Staged diabetes management: a systematic approach, 2nd edn. Wiley, Hoboken

Nathan DM, Buse JB et al (2009) Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy. Diabetes Care 32(1):193–203
Navarro G (2001) A guided tour to approximate string matching. ACM Comput Surv 33(1):31–88
Newell A (1982) The knowledge level. Artif Intell 18:87–127
O'Connor PJ, Sperl-Hillen JM et al (2009) Simulated physician learning intervention to improve safety and quality of diabetes care: a randomized trial. Diabetes Care 32(4):585–590
Rouse W, Morris N (1986) On looking into the black box: prospects and limits in the search for mental models. Psychol Bullet 100(3):349–363
Seborg DE, Edgar TF et al (1986) Adaptive control strategies for process control: a survey. AIChE J 32(6):881–913
Soukoreff RW, MacKenzie IS (2001) Measuring errors in text entry tasks: an application of the Levenshtein string distance statistic. In: ACM conference on human factors in computing systems. ACM, New York
Soukoreff RW, MacKenzie IS (2003) Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, Ft. Lauderdale
Stanton MW (2001) Improving care for diabetes patients through intensive therapy and a team approach. Research in action. U. S. D. o. H. H. Services. Agency for Healthcare Research and Quality, Rockville. November 2001, pp. 1–12
Sterman JD (1989) Misperceptions of feedback in dynamic decision making. Organ Behav Hum Decis Process 43:301–335
Ungar L, Hartmann E et al (1996) Process modeling and control using neural networks. In: First international conference on intelligent systems in process engineering. American Institute for Chemical Engineers, Central Florida, pp 312–318

Chapter 6
An Empirical Investigation of the Impact of Information Technology Integration in Healthcare Integrated Delivery Systems

Evelyn Thrasher

Gordon Ford College of Business, Western Kentucky University, Bowling Green, KY, USA

Abstract Healthcare alliance networks, also known as integrated delivery systems (IDSs), have developed rapidly in the US. Yet, research has not kept pace. This study examines two interrelated needs of healthcare alliance network research—the value of information technology (IT) integration for the IDS and the value of IT integration for the participating hospitals. Hypotheses are developed that relate IT integration to quality and performance outcomes at both the IDS level and at the hospital level. In addition, a taxonomy of hospitals is developed based on size; and the value of IDS IT integration is examined for each hospital quartile. The results of the study suggest that IT integration does indeed have a significant impact on performance in the healthcare industry. In addition, the results suggest that the performance improvements realized at the IDS level of analysis do not extend equally to all hospitals in the IDS.

Keywords Electronic survey methods • Empirical research • Healthcare • Healthcare integrated delivery systems • Partial least squares

6.1 Introduction

Strategic alliances, defined as "clusters of organizations that make decisions jointly and integrate their efforts to produce a product or service" (Alter and Hage 1993, p. 2), continue to receive attention as a potentially valuable means of achieving performance improvement and competitive advantage (Payton and Ginzberg 2001;

E. Thrasher (✉)
Gordon Ford College of Business, Western Kentucky University, Bowling Green, KY, USA
e-mail: evelyn.thrasher@wku.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_6

Lapiedra et al. 2004; Straub et al. 2004). The key to these alliances is the integration of otherwise autonomous organizations established to pursue a common goal or purpose (Alter and Hage 1993; Young and McCarthy 1999; Deluca and Enmark 2002; Schumaker 2002).

Researchers argue that strategic alliances have been enabled largely by information technology (IT) advances and innovations (Al-Mashari and Zairi 2000; Schumaker 2002; Straub et al. 2004; Simoens and Scott 2005). IT enhances the formation and growth of the alliance network through its ability to facilitate information flow across organizational boundaries (Al-Mashari and Zairi 2000). Further, IT has been posited to encourage innovation in business and decision making processes (Hsiao and Ormerod 1998; Subramani 2004; Straub et al. 2004) and may play a more critical role in the success of these network arrangements than it does in individual firms (Straub et al. 2004). The increased importance of IT is due, in part, to the additional complexities of structure, form, administration, and function of alliance networks that are not necessarily present at the firm level (Page 2003; Straub et al. 2004).

Alliance networks have developed at a rapid rate across some industries, such as the healthcare industry which has over 450 alliance networks currently active in the United States (HIMSS 2004). In the healthcare industry, these networks are known formally as “integrated delivery systems (IDSs).” IDSs are composed of different healthcare stakeholders including hospitals, insurers, physician practices, laboratories, pharmacies and other organizations that have agreed through formal arrangements to work together to deliver better healthcare to patients at a lower cost. The need for IDSs is predicted to continue to grow in response to new governmental policies and increasingly complex issues that should be more effectively addressed through tighter integration and coordination (Schumaker 2002). One prominent feature of IDSs has been the use of IT to help facilitate the operational integration of the various organizations in the network and to help provide the benefits for which the IDS was formed (Killian 1999; Kaplan et al. 2001; Enthoven and Tollen 2005). However, while IDSs continue to surface, research into the role and value of IT in these organizational forms has not kept pace.

This research has two primary objectives. The first objective is to examine the value of IT integration at the IDS level of analysis. IT helps with the integration of processes and decision-making by providing a conduit that offers almost transparent communication and coordination media regardless of time and distance which should, in turn, result in improved quality and financial performance for IDSs. The second objective is to examine how the value of IT integration across the IDS is passed down to the participating hospitals. Hospitals in an IDS invest a great deal of money for IT resources that will enable and enhance communication and collaboration with other entities in the IDS. However, it is not likely that the value achieved for the IDS as a whole is equally distributed across the entities within the IDS. While it would certainly be positive and favorable for the IDS as a whole to realize quality and financial improvements from IT integration, it is not realistic to expect that all firms existing underneath the IDS umbrella will realize equal value. Thus, this study will offer a starting point in the investigation and discussion of IT value distribution across the IDS.

6.2 Background

Healthcare organizations, like firms in most industries, face significant pressure to provide higher quality service at lower costs. According to Etchen and Boulton (2000), researchers and market analysts in the early 1990s predicted a dramatic increase in the consolidation of healthcare providers. Much of the consolidation in the healthcare industry prior to 1990 involved only hospitals and was caused by a decreased demand for inpatient care. However, as demand for alternative services such as outpatient and home health services has increased over time, further consolidation has occurred among a variety of healthcare entities to provide a diversified, comprehensive continuum of care, or healthcare IDS (Conrad and Shortell 1997; Young and McCarthy 1999; Page 2003).

IDSs take on various forms, including contract arrangements, partnerships, strategic alliances, and corporate ownership; and multiple arrangements may be present within a single IDS (Conrad and Shortell 1997; Page 2003; Weiner et al. 2004). Comprised of diverse combinations of insurers, hospitals, physician practices, laboratories, pharmacies, and other medical care organizations, IDSs are arguably enabled through IT capabilities and innovation (Young and McCarthy 1999). To illustrate the general impact of technologies, traditional versus IT-enabled healthcare network information flow is briefly discussed.

6.2.1 Traditional Healthcare Network Information Flow

The purpose and nature of work in the healthcare industry has always required the sharing of information, knowledge, resources, and people. Traditionally, this has been accomplished through hybrid arrangements of healthcare providers, each with different roles, organizational structures, and information needs. Figure 6.1 depicts the traditional healthcare network as a hub and spoke arrangement linked together by the patient (Rosow and Grimes 2003). Information exchange was accomplished primarily through manual processes that forced the patient to relay information numerous times to a number of different organizations. While the risk associated with multiple data entry points is an important issue among firms in most industries, this risk becomes even more crucial in the healthcare industry where diagnosis and treatment depend upon accurate, timely, comprehensive patient data (Dowling 1997; Friedman and Wyatt 1997). To further complicate the issues, at any given point in time, a network entity could be both a partner and a stakeholder depending upon the circumstances of the transaction or the nature of the patient service being provided (Friedman and Wyatt 1997).

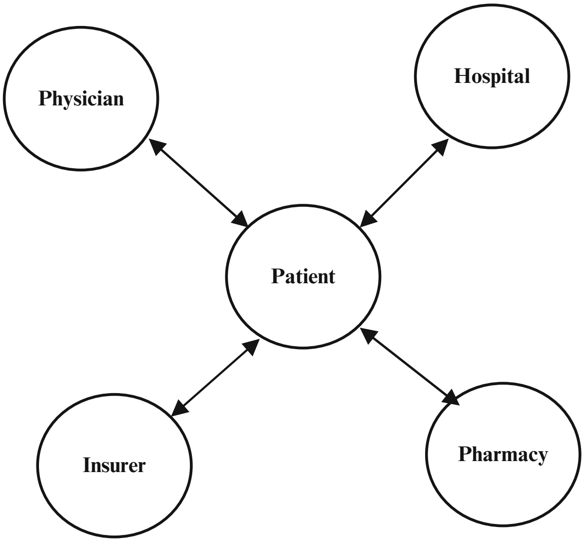


Fig. 6.1 Traditional healthcare network information flow

6.2.2 IT-Enabled Healthcare Network Information Flow

IT has the capability to act as an electronic intermediary, thus, acting as a surrogate for the patient in the healthcare network as shown in Fig. 6.2. The intent of the electronic intermediary is to promote the exchange of information, medical data, and discoveries among a number of combinations of patients, providers, and payers in a timelier, more efficient manner (Venkatraman 1994; Friedman and Wyatt 1997; Kaplan et al. 2001; Rosow and Grimes 2003). IT's role in the healthcare industry should allow for direct links among all entities of the healthcare network, regardless of function, purpose, or geographic location (Venkatraman 1994; Kaplan et al. 2001). Information is still seen as the key driver of the healthcare industry and serves as the link among the various entities in the network. However, IT capabilities such as IT flexibility and accessibility, IT and business alignment, and IT modularity are making information exchange more relevant, consistent, and accurate. In addition, these new applications, along with continued IT innovation, are leading to the establishment of logical integrated healthcare systems with more fluid boundaries and less reliance on the patient (Kaplan et al. 2001; Beaudoin and Bogaert 2003).

6.3 Conceptual Development

6.3.1 Quality Performance Defined

Patient-centered measures are generally used in studies of organizational performance in the healthcare industry (Dowling 1997; Devaraj and Kohli 2000; Smith

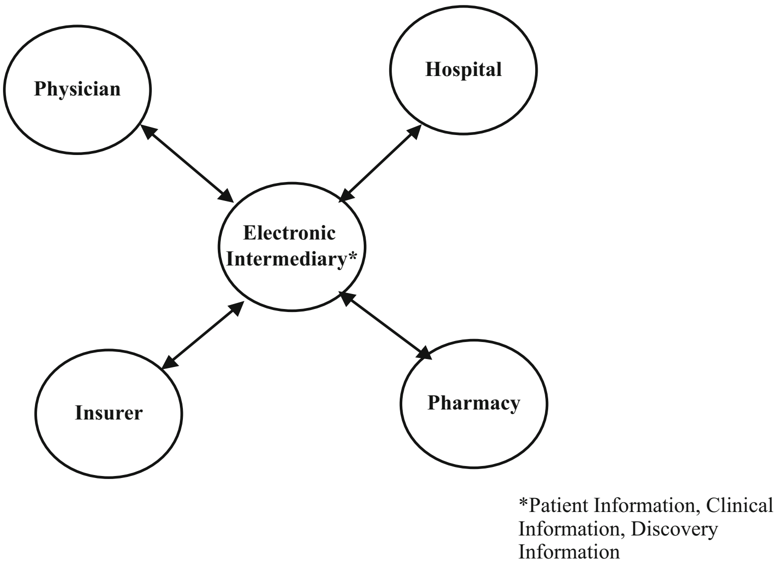


Fig. 6.2 IT-enabled healthcare network information flow

and Swinehart 2001). Because the patient is the primary focus of any healthcare entity, measures related to patient care are appropriate for the evaluation of quality performance. Patient days in the hospital, hospital admissions, and mortality rate were chosen as measures of quality for this study.

6.3.2 *Financial Performance Defined*

Cost measures are often used as financial performance indicators in the healthcare industry (e.g., Dowling 1997; Killian 1999; Kim 2000; Friedman and Goes 2001). Use of cost measures in the context of healthcare research addresses two issues. First, the sample selected may include both for-profit and not-for-profit organizations—however both types are concerned with reducing costs despite their differences regarding profit-centered measures. Second, the healthcare industry is facing increasing pressure to reduce costs while continuing to improve the quality of patient care (Snyder and Paulson 2002); thus, it is important to examine the potential impact of integration on cost measures as well as quality. For the purposes of the current study, operational cost, administrative cost, and clinical cost have been chosen as measures of financial performance, as these represent the majority of costs incurred by healthcare organizations.

6.3.3 IT Integration to Performance Outcomes

This paper utilizes theory from the learning and action literature that states that actions in firms can be divided into either for *exploitation* or *exploration* (Subramani 2004). Exploitation is a modification or extension to existing capabilities and resources in organizations. The results of the actions from exploitation would include matters such as improvements in operational efficiencies, reduction of organizational costs, improvements in quality, and better communication and collaboration across functions. Exploration is related more to risk taking and involves experimentation, reengineering, or other major renovations or innovations that are introduced into the business processes and organizational environment. Tallon et al. (2000) have found evidence that most organizations still use IT for exploitation with far less using it for exploration. This may be especially true in the healthcare industry which lags behind most industries in its use of IT. Since the effects of IT are likely to be those related to exploitation not exploration at this point in the healthcare industry, we focus our study around the concept of how IDSs are using IT for exploitation via integration.

Waring and Wainwright (2000, p. 137) defined integration as “a result or an effect of correct commonality.” Extending the idea of correct commonality to IT, IT integration may be defined as seamless, boundariless, access to information from anywhere anytime (Knoedler 2003). Successful IT integration will deliver IT resources and solutions that effectively support the roles and functions of workers (Rockart et al. 1996; Skipper et al. 2008). Furthermore, IT integration extends to all facets of the IT infrastructure, including telecommunications, hardware, software, and data (Byrd and Turner 2000).

Tapscott and Caston (1993) discussed the changing role of IT in the organization and the accompanying improvements that should result from increased IT integration. These authors described the digital ticketing system implemented by Northwest Airlines in the late 1980s. In addition to implementing the new system, Northwest also integrated the ticketing system with other applications, such as the reservation system, flight operation system, and revenue collection system. In doing so, Northwest was able to implement expert systems that could identify backlogs, discrepancies in ticket revenue collection, and other information regarding the data attached to the ticket. As a result, the processing time for ticket revenue accounting was decreased by half, cash collections were significantly accelerated, and customer inquiry time was shortened. The combination of these improvements resulted in significant cost reductions for Northwest (Tapscott and Caston 1993).

An investigation of an electronic health record implementation and integration was conducted at Florida Hospital (Hamilton et al. 2004). This was an intensive project aimed at re-engineering current clinical information processes and creating an integrated interdisciplinary electronic health record. In doing so, an advanced clinical application was integrated with existing financial and patient-management applications, thereby providing a single point of data access for personnel from a variety of functional areas within the hospital. Following completion of the project, a follow-up survey was done to measure overall satisfaction with the system and to assess potential performance improvements. Respondents reported an 88 % approval rating, and hospital personnel identified specific tangible improvements. For instance, one nurse estimated patient-care savings of 30 min per patient per shift. Another reported a 50 % decrease in physician to pharmacy communication regarding patient medicine (Hamilton et al. 2004). As Hamilton et al. (2004) noted, these improvements should translate into a shortened hospital stay for the patient and reduced costs for the hospital over time.

In another study from the healthcare industry, researchers examined the link between electronic entry of physicians' orders and operational cost (Bates 2002). By entering the orders electronically, the system gave the hospital the ability to transmit electronic orders to wherever they were needed instantaneously, thereby providing an initial level of IT integration throughout the hospital. Thus, the pharmacy, lab, nurses' station, and other areas of the hospital received instant notification of patient needs and, in addition, could respond electronically with updates and additional information. The order entry implementation resulted in cost savings of 13.1 % for the hospital as opposed to the use of paper orders (Bates 2002).

Similarly, an experimental study of electronic patient order entry was conducted at a large hospital in Indiana (Tierney et al. 1993). Employees of the hospital were divided into an experimental group and a control group. The experimental group was to use electronic order entry, while the control group continued with the traditional paper system. The link to average length of hospital stay was tested to determine potential differences between the two groups. Results of the experiment revealed that average length of hospital stay was 0.89 days shorter for patients admitted and treated by the experimental group than for those of the control group. This equates to a 10.5 % reduction in hospital stay as a result of the use of electronic patient order entry, which should result in greater perceived quality of care and reduced costs for the hospital over time (Tierney et al. 1993).

Based on previous research efforts, the following hypotheses are proposed:

H1. IT integration will have a significant influence on IDS quality performance improvement.
H2. IT integration will have a significant influence on IDS financial performance improvement.
H3. IT integration across the IDS will have a significant influence on quality performance in participating hospitals.
H4. IT integration across the IDS will have a significant influence on financial performance in participating hospitals.

Extending hypotheses 3 and 4 further, this study proposes that hospital level benefits achieved through IT integration across the IDS will not be equally realized in each participating hospital. Almost certainly, there will be disparities in value achieved for the hospitals across the IDS.

6.4 Method

6.4.1 Unit of Analysis and Participants

The population of interest was the approximately 450 active US healthcare IDSs, as identified in the 2005 HIMSS Analytics Database. These IDSs represent a broad spectrum of diversity, size, geographic reach, and comprehensiveness of patient care. Because of the use of email to recruit participants and to distribute the link to the survey instrument, the CIOs for all IDSs in the population with valid email addresses were invited to participate in the study and to complete the survey instrument. The CIO was chosen due to the need for survey participants to have a broad level of understanding and knowledge regarding the structure of the IDS and the current state of IT integration within the IDS.

6.4.2 Secondary Data

Two important sources of data for this study were the HIMSS Analytics Database and the American Hospital Directory. This secondary data was used to measure the six quality and financial performance variables. Quality performance outcomes were measured using patient days in the hospital, mortality rate, and hospital admissions. Financial performance outcomes were measured by operational cost, administrative cost, and clinical cost. Secondary data was collected for the 99 IDSs and also for the 732 hospitals that are part of the IDSs in the sample.

6.4.3 Primary Data

A survey instrument was used to collect primary data to represent the independent variables. The survey was developed primarily from multiple existing scales rather than introducing a new instrument (Byrd and Turner 2000; Lewis and Byrd 2003). The items, shown in Table 6.1, were measured on a five-point Likert scale that ranged from strongly disagree to strongly agree.

Measurement–IT Integration. In line with Kilbridge's (1998) suggestions, three facets of IT integration were measured (see Table 6.1). First, IT flexibility and accessibility was measured using a scale developed by Byrd and Turner (2000). Consisting of six items, this scale was designed to assess the flexibility of an organization's IT resources to provide integration and seamless sharing of information throughout the organization. Second, IT modularity was measured using a scale developed by Byrd and Turner (2000). This scale was designed to assess the perceived modularity and ease of application development within the organization. IT planning was measured using a scale developed by Lewis and

Table 6.1 Primary data—measurement

Information technology integration
Information technology flexibility and accessibility
ITI1-Flexible electronic links exist between all organizations of the IDS
ITI2-Our IDS utilizes open systems network mechanisms to boost connectivity
ITI3-Our user interfaces provide transparent access to all platforms and applications
ITI4-Data received by our IDS from electronic links (e.g., EDI, EFT) are easily interpretable
ITI5-Information is shared seamlessly across our IDS, regardless of the location
ITI6-Our IDS provides multiple interfaces or entry points (e.g., Web access) for external end users
Information technology planning
ITP1-A plan exists for IDS-wide information technology
ITP2-Our information technology plan reflects business goals
ITP3-An information technology steering team/advisory committee is active within our IDS
ITP4-Senior management participates in the information technology steering team/advisory committee
ITP5-Users participate in the IT steering team/advisory committee
ITP6-A formal methodology for information technology development is in place in our IDS
Information technology modularity
ITM1-Reusable software modules are widely used in new systems development in our IDS
ITM2-Our databases are able to communicate through many different protocols
ITM3-The development of new applications is not restricted by legacy systems within our IDS
ITM4-A common view of our IDS's patient data is available to all authorized personnel in the IDS
ITM5-Data captured in one part of our IDS are immediately available to everyone in the IDS

Byrd (2003). This scale was designed to assess the existence and appropriateness of organizational IT plans and the degree of widespread participation in, and effectiveness of, an IT steering team.

6.4.4 Survey Administration

The survey was administered electronically, using an online survey instrument. Invitations to participate in the study were distributed via email containing a hyperlink to the survey instrument. Thus, from the total population of 450 IDSs contained in the HIMSS Analytics 2005 Database, those with an email address for the CIO were identified, resulting in an initial sample of interest of 394 healthcare supply networks. Of the 394 healthcare supply networks sent the first email invitation, 25 were returned due to invalid email addresses or electronic refusals of email delivery. Five email addresses were corrected and the email invitation was resent, resulting in a final sample of interest of 374 healthcare supply networks. After the initial email invitation and three subsequent email reminders, 104 IDSs

Table 6.2 Descriptive statistics for respondents

Demographic characteristic	Mean	Std Dev
Number of facilities	41.87	46.47
Types of facilities	4.53	1.08
Number of hospitals	7.40	11.27
Cities served*16 IDSs serve cities in multiple states	14.81	19.20
Number of enterprise applications	30.77	44.24
Number of application types	9.11	2.32
Number of wireless access points	101.45	222.92
Number of hospitals with intranet	5.01	8.16
Patient revenue (in millions)	$2775.75	$3481.56
Operational cost (in millions)	$1110.20	$1310.61
Clinical cost (in millions)	$901.71	$1067.58
Administrative cost (in millions)	$191.54	$220.84
Hospital admissions	26,277.04	34,637.75
Patient days in the hospital	340,235.88	415,482.39
Mortality rate	1041.51	1391.86

responded, for a response rate of 28 %. Of those, five had incomplete data and were removed, for a final sample size of 99. Descriptive statistics for the respondent IDSs are provided in Table 6.2.

6.4.5 Data Analyses

Empirical examinations of the hypotheses were conducted in two steps. First, confirmatory factor analysis was conducted for all latent constructs in the model. Goodness of fit for the measurement model was assessed using several fit indices (Hair et al. 1998). In addition, the strength of factor loadings was examined. Potential adjustments to the measurement model were considered only if appropriate and theoretically supported. Second, hypotheses testing was conducted using Partial Least Squares (PLS) (Chin 2001).

Tests for Sample Bias. Two tests for bias were conducted using t-tests. First, from the final sample of interest of 374 healthcare supply networks, the non-respondent networks were compared to the respondents based on several variables obtained from the HIMSS Analytics Database (e.g., number of facilities, facility types, cities served, operational cost, number of enterprise applications, application types). No significant differences were found. Second, from the population of 450 healthcare supply networks, the same variables from HIMSS Analytics were used to compare those networks who could not be contacted via email to the respondent networks. Again, no significant differences were noted. The results of the bias tests

Table 6.3 Confirmatory factor analysis

Measurement	Range of standardized factor loadings	NFI	IFI	CFI
IT flexibility and accessibility	0.48–0.81	0.93	1.00	1.00
IT planning	0.71–0.83	0.86	0.90	0.90
IT modularity	0.63–0.83	0.86	0.90	0.89

Table 6.4 Scale reliabilities and correlations

Scale	1	2	3
1. IT flexibility and accessibility	(.76 .83 .46)		
2. IT planning	.46	(.85 .89 .58)	
3. IT modularity	.77	.40	(.70 .84 .57)

Note: Cronbach's Alpha, composite reliability, and AVE designated in the diagonal

suggest no cause for concern with regard to non-response or the use of only email for survey administration.

Confirmatory Factor Analysis. Confirmatory factor analysis was conducted in AMOS 5 to assess the goodness of fit for each construct. ITM3 was removed from the IT Modularity scale due to a low factor loading of .26. As presented in Table 6.3, the goodness of fit indices indicated acceptable fit for all constructs.

Scale Reliability and Validity. The reliability and validity measures are summarized in Table 6.4. Cronbach's alpha, obtained from AMOS 5, and the composite reliability, obtained from PLS Graph 3, were used to assess scale reliabilities. A coefficient alpha of at least .70 is indicative of strong covariance among the items and suggests a satisfactory sampling domain (Hair et al. 1998; Hinkin 1995; Nunnally 1978). All factors demonstrated acceptable Cronbach's Alpha and composite reliability levels, ranging from .70 to .85 and .83 to .89, respectively. A measure of factorial validity common to PLS is the Average Variance Extracted (AVE). While strict thresholds for this measure are yet to be determined, an AVE of approximately .50 is desirable (Gefen and Straub 2005; Hair et al. 1998). As such, each factor demonstrated AVEs very near or above the suggested threshold, ranging from .46 to .58.

Analysis of Statistical Power. Statistical power is "the probability that a statistical test will correctly reject a null hypothesis" (Baroudi and Orlikowski 1989, p. 87). Statistical power for the current study was conducted according to the recommendations of Baroudi and Orlikowski (1989) and the standards established by Cohan (1988). Because effect sizes for the scales were not readily available, we assumed a medium effect size, deemed to be appropriate for IS research (Cohan 1988; Baroudi and Orlikowski 1989). Based on Cohan (1988) tables, a sample size ranging from 68 ($p < .05$) to 108 ($p < .01$) is necessary to produce a desired power level of at least .80 for medium effect sizes. Our sample of 99 falls within this range, thereby suggesting that our hypotheses tests demonstrate acceptable power levels.

Table 6.5 Summary of hypotheses tests

	Hypotheses	Results
H1	IT integration will have a significant influence on IDS quality performance improvement	Supported –Mortality rate –Admissions
H2	IT integration will have a significant influence on IDS financial performance improvement	Supported –Administrative cost –Clinical cost
H3	IT integration across the IDS will have a significant influence on quality performance in participating hospitals	Supported –Patient days –Mortality rate –Admissions
H4	IT integration across the IDS will have a significant influence on financial performance in participating hospitals	Supported –Operational cost –Clinical cost

Table 6.6 Measurement model path coefficients

	IT integration	
	IDS level	Hospital level
Admissions	-0.182^{**}	-0.084^{**}
Mortality rate	-0.158^{*}	-0.152^{***}
Patient days	-0.128	-0.138^{***}
Operational cost	-0.121	-0.124^{***}
Clinical cost	-0.135^{*}	-0.135^{***}
Administrative cost	-0.145^{*}	-0.045

$p<.10^{*}$; $p<.05^{**}$; $p<.01^{***}$

6.5 Results

Partial Least Squares conducted in PLS-Graph 3.0 was used to analyze the path model for both IDSs and hospitals, respectively (Sambamurthy and Chin 1994; Byrd et al. 2005). For the hospital taxonomy, hospitals were divided into quartiles based on size, as determined by the number of beds. Tables 6.5, 6.6, and 6.7 provide an overview of the data analyses results.

6.5.1 IT Integration: IDS Level

Hypotheses 1 and 2 regarded the relationship between IT integration and both quality and financial outcomes at the IDS level of analysis. Hypothesis 1, which stated that IT integration will have a significant influence on IDS quality performance improvement, was supported for mortality rate ($b = -0.158$, $t = 1.4588$, $p < .10$) and admissions ($b = -0.182$, $t = 1.6713$, $p < .05$). Although the relationship was negative, the hypothesis was not supported for patient days in the hospital ($b = -0.128$, $t = 1.2126$). Additionally, hypothesis 2, which stated that IT integration will

Table 6.7 Taxonomy of hospital level results

	Lower quartile	Second quartile	Third quartile	Upper quartile
	n = 191	n = 182	n = 178	n = 181
Number of beds	≤57	58–126	127–242	≥243
Admissions	0.1603**	0.0913	−0.2416***	0.0504
Mortality rate	−0.0605	0.0338	−0.2726***	0.0415
Patient days	0.0999	−0.1074	−0.2788***	0.1060
Operational cost	0.1178*	−0.1026	−0.3169***	0.0887
Clinical cost	0.0209	−0.1559***	−0.3465***	0.1169
Administrative cost	0.4120***	0.1960*	−0.0110	0.1086*

$p < .10^{*}$; $p < .05^{**}$; $p < .01^{***}$

have a significant influence on IDS financial performance improvement, was supported for administrative cost ($b = -0.145$, $t = 1.4815$, $p < .10$) and clinical cost ($b = -0.135$, $t = 1.4076$, $p < .10$). The hypothesis was not supported for operational cost ($b = -0.121$, $t = 1.1597$). The direction of the path between IT integration and operational cost was negative, but the relationship was not significant.

6.5.2 IT Integration: Hospital Level

Hypothesis 3, which stated that IT integration across the IDS will have a significant influence on quality performance improvement for participating hospitals, was supported for mortality rate ($b = -0.1523$, $t = 3.5631$, $p < .01$), admissions ($b = -0.084$, $t = 1.7682$, $p < .05$), and patient days ($b = -0.1375$, $t = 3.1839$, $p < .01$). Hypothesis 4, which stated that IT integration across the IDS will have a significant influence on financial performance improvement for participating hospitals, was supported for operational cost ($b = -0.1242$, $t = 3.0912$, $p < .01$) and clinical cost ($b = -0.1348$, $t = 3.2398$, $p < .01$). The hypothesis was not supported for administrative cost ($b = -0.0449$, $t = 1.2058$). The direction of the path between IT integration and administrative cost was negative, but the relationship was not significant.

6.5.3 IT Integration: Hospital Taxonomy

IT integration across the IDS demonstrated the most significant influence on performance for hospitals in the third quartile. IT integration to administrative cost was negative but not significant. However, all other relationships were significant—admissions ($b = -0.2416$, $t = 2.1322$, $p < .01$), mortality rate ($b = -0.2726$, $t = 2.2197$, $p < .01$), patient days ($b = -0.2788$, $t = 3.1348$, $p < .01$), clinical costs ($b = -0.3465$, $t = 2.6646$, $p < .01$), and operational costs ($b = -0.3169$, $t = 2.3977$, $p < .01$).

6.6 Discussion

6.6.1 Theoretical Contribution

Supply networks are frequently being adopted in many industries as a means to improve performance and create competitive advantage (Payton and Ginzberg 2001; Straub et al. 2004). Healthcare IDS participants are organized jointly to enhance quality and financial performance for network participants (Alter and Hage 1993; Payton and Ginzberg 2001; Lapiedra et al. 2004; Straub et al. 2004). The integration of business processes, functions, and tasks, enabled largely by IT, appears to be critical to IDS success; yet empirical evidence of this relationship is lacking in the IT literature (Schumaker 2002; Straub et al. 2004). Therefore, this study has provided some much needed attention to this area by examining the relationships of IT integration to quality and financial performance improvement, both for the IDS and for the participating hospitals. Patient days in the hospital, mortality rate, and hospital admissions were used as measures of quality that are appropriate for the healthcare industry (Dowling 1997; Bates 2002); operational cost, administrative cost, and clinical cost represented financial performance outcomes (Devaraj and Kohli 2000; Snyder and Paulson 2002). Primary survey data gathered from IDS CIOs was combined with secondary data from HIMSS Analytics and the American Hospital Directory to empirically test the hypothesized relationships of IT integration to quality and financial performance improvement.

6.6.2 IT Integration to Performance

Hypotheses 1 and 2 predicted that IT integration would have a significant impact on both quality and financial performance. Regarding quality, IT integration was linked to reductions in hospital admissions and mortality rate. Financial improvements in terms of reduced clinical cost and administrative cost as a result of IT integration were also demonstrated. Hypotheses 3 and 4 predicted that hospitals participating in IDSs would also achieve quality and financial improvements through IT integration with other IDS entities. Significant reductions in hospital admissions, patient days in the hospital, and mortality rate were observed, while both operational and clinical costs were significantly reduced. These findings, both at the IDS level and hospital level, are in line with prior healthcare research, which has suggested a significant improvement in both cost and quality attributed to the use of IT and greater integration of IT resources (Tierney et al. 1993; Bates 2002; Hamilton et al. 2004).

The results of the present study support the work of Tierney et al. (1993) and Hamilton et al. (2004), who noted that as IT integration leads to improvements in patient-centered measures, such as mortality rate and admissions, patients should perceive a higher quality of care, a primary goal of healthcare organizations (Dowling 1997; Laplante 2005). Further, these results are in line with Bates (2002), who demonstrated a significant cost savings associated with physicians' use of electronic order entry. Finally, these findings lend support to the firm-level evidence within other industries. For example, in their investigation of IT integration in the airline industry, Tapscott and Caston (1993) also found a significant relationship between increasing levels of IT integration and cost reductions. Thus, the findings of the present study further emphasize the need for continual efforts to increase and improve IT integration for greater supply network success, both in terms of quality and financial performance (Weiner et al. 2004).

6.6.3 Hospital Taxonomy

As anticipated, the impacts of IT integration at the IDS level of analysis were not equally realized across all participating hospitals. While the performance benefits were significantly improved for the sample of hospitals in total, the taxonomy of hospitals characterized by size demonstrated that those in the third quartile seemed to realize the most significant benefits of their investments in IT integration with other IDS entities. Perhaps these results are indicative of an optimal hospital size for those who are part of a larger healthcare network. Or, perhaps size is an indicator of maturity, suggesting that growth is a part of the maturation process. This explanation is in line with the IT-Enabled Business Transformation Framework, which suggests the real value of IT resources is in their ability to enhance and integrate business processes (Venkatraman 1994). The Framework proposes that as companies integrate their IT resources increasingly tighter with their business processes, the value of this integration increases substantially. In fact at the highest level of integration between these organizational resources, Venkatraman (1994) states that companies begin to move from being internally focused to being more externally focused. Finally, those that reach a maximum level of maturity tend to see IT value flatten or even decline, perhaps explaining the lack of significant results for those in the upper quartile. It is quite possible that the number of hospital beds is more of an indicator or proxy for some other characteristic, such as maturity, expansion, or innovation, as opposed to merely an indicator of size.

6.6.4 Implications for Managerial Decision-Making

Since the early 1990s, healthcare organizations have actively formed IDSs with two primary goals of reduced costs and improved patient care quality. Researchers and practitioners alike have attributed much of the formation and expansion of IDSs to the employment of appropriate IT resources and innovations (Dowling 1997; Boone and Maley 2000; Weiner et al. 2004). To that end, healthcare administrators have tended to focus on the cost benefits of these arrangements as justification for further growth and expansion. Yet, others argue that these administrators must also consider the importance of improvements to patient care quality in their decisions (Beaudoin and Bogaert 2003). Recently, however, many IDS administrators have expressed a hesitance to make additional expensive investments in IT given the pressures to control costs. In fact, some have even considered the dissolution of IDSs due to an inability to quantify the anticipated improvements in either quality or financial performance (Parker et al. 2001; Laplante 2005); and these same concerns have been raised regarding supply networks across other industries. Therefore, the results of the present study are timely and should aid in the justification of decisions by management to implement additional innovations, integration, and IT-enabled change within the IDS.

IT has been demonstrated to be a critical enabling force for IDS integration through its enhancement of process integration, communication, and collaboration. This argument supports past researchers who have touted IT as one of the most critical enabling resources in the formation and success of supply networks (Payton and Ginzberg 2001; Melville et al. 2004; Straub et al. 2004). Therefore, IDSs may seek even higher levels of IT integration as an important and valuable component of integration across all aspects of the organization. Additionally, managers of participating hospitals would be well served to conduct a thorough examination of the value of IDS IT integration for their organization. Unfortunately, quite often the costs of integrative technologies are an unavoidable expense for being an active member of the IDS, as integration across all entities is of critical importance to the success of the IDS as a whole. With that said, perhaps the participating hospitals should look internally to see if the technologies they are using to connect to other IDS entities might also allow for similar connectivity among the units inside the hospital. It is likely that there are redundant applications and processes in place inside the hospitals that may be reduced or eliminated by making more effective use of the IDS-level IT resources.

One important contribution of the present study is the use of patient-centered, objective measures of quality. The results of the present study suggest that IT integration in the healthcare industry serves not only to improve administrative costs, but also directly and favorably impacts the medical care of patients. Of those measures examined, healthcare administrators have identified mortality rate as the most critical measure of patient care quality, noting the importance of identifying innovations, techniques, and medical discoveries that can aid in reducing the number of patient deaths (Dowling 1997). Thus, the results of the present study suggest that IT integration has a direct and significant influence on reductions in patient deaths. These results alone should spur the interest of IDS and hospital administrators and should serve as a catalyst to promote the diligent pursuit of increased IT integration.

Rogers and Lutz (2003) observed slowed or stalled progress in terms of the quality and financial performance benefits of IDS formation. The findings of the current study suggest that IDS formation alone is insufficient to bring about noticeable change. It is rather IT's ability to enhance communication and collaboration that affords an IDS the greatest performance improvement, both in terms of quality and financial performance. Thus, as IDS administrators contemplate further investments in IT resources and integration, perhaps this research and the associated results can aid in the justification of these investments.

Based on the findings and discussion of the present study, instead of dissolution, IDS administrators would be well-served to adopt a practice of continual dynamic integration and innovation. Even where relationships were insignificant within the present study, these links demonstrated a desirable path direction, suggesting the potential for greater benefit in the future. In addition, a closer examination of the disparities in IT value across hospitals may also point to opportunities to improve hospital performance that will, in turn, translate to IDS performance improvements. If the participating hospitals can significantly improve their financial and quality performance through effective IT integration, these organizations will be more supportive of the integration efforts at the IDS level.

Additionally, as the lag effect of IT suggests, the maximum quality and financial benefits of integration may not have been fully realized due to the young age of many of these networks (Barua et al. 1995; Byrd et al. 2005). Consider the argument of the lag effect which suggests that new IT investments may require at least 2–3 years to generate significant performance benefits. Then, allow ample time for IDSs to expand, diversify, and integrate the IT resources of all participants. This leaves very few years for the maximum performance benefits to have materialized. In fact, Beaudoin and Bogaert (2003) suggest that within the healthcare industry, organizational transformations of the magnitude required by IDS expansion and sustained by IT integration may take as long as 5–7 years to produce tangible financial benefits. Yet, the results presented in this study are significant and favorable and should be even more prominent in the future, thereby encouraging IDSs to stay the course, to continue to pursue appropriate integration, and to exercise patience in terms of quantifying the benefits of these investments.

Finally, as IDS and hospital administrators make the choice to expand and integrate, they must recognize that neither the formation of the IDS nor the purchase of an ERP system is sufficient in isolation to bring about extensive improvements. As this study has demonstrated, IT integration includes the hardware and software, planning processes, and modularity and flexibility of the chosen solutions. Thus, all aspects of IT integration must be adequately addressed if substantial improvement is to be achieved and sustained. Similarly, as participants join the IDS, the contract between providers is simply not enough to spawn performance improvement. Administrators must also encourage an atmosphere of effective communication, collaboration, and business process integration among all participants. The synergy brought about by sufficient attention to integration efforts, and the recognition and emphasis on IT to promote and enable information sharing, is ultimately what brings about performance improvement.

6.6.5 Future Research

This study has laid the groundwork around IT value for IDSs and for those hospitals participating in the IDS. Additional research is needed to bring depth to this broad investigation, beginning with a more intensive study into the characteristics of those hospitals who realize significant performance improvements through IDS IT integration. In fact, a thorough taxonomy is needed to identify the optimum set of characteristics that will allow a hospital to achieve significant performance improvements by effectively utilizing the IT resources of the IDS.

In addition, while quality measures such as mortality rate, patient days in the hospital, admissions, and costs will certainly continue to be important measures for investigation and improvement, a need exists to explore the value of more complex treatments and medical procedures, which are largely enabled by IT but the results of which cannot be readily obtained from a database (Bates 2002). Further, as patients continue to take a more active role in their care, satisfaction with the information available to them electronically may also play a more significant role in studies of IT value to healthcare. Additional research will be needed to better assess this factor (Bates 2002).

References

Al-Mashari M, Zairi M (2000) Creating a fit between BPR and IT infrastructure: a proposed framework for effective implementation. Int J Flex Manuf Syst 12(4):253–274

Alter C, Hage J (1993) Organizations working together. Sage, Newbury Park

Baroudi JJ, Orlikowski WJ (1989) The problem of statistical power in MIS research. MIS Quarterly 13(1):87–106

Barua A, Kriebel CH, Mukhopadhyay T (1995) Information technologies and business value: an analytic and empirical investigation. Inform Syst Res 6(1):3–23

Bates DW (2002) The quality case for information technology in health care. BMC Med Inform Decis Making 2(7):1–9

Beaudoin D, Bogaert M (2003) Measuring return on investment. HCT Project 119–122

Boone B, Maley R (2000) Integrated health care delivery systems' challenges. IRMI. http://www.irmi.com/expert/Articles/2000/Boone06.aspx. Accessed 23 July 2006

Byrd TA, Turner DE (2000) Measuring the flexibility of information technology infrastructure: exploratory analysis of a construct. J Manage Inform Syst 17(1):167–208

Byrd TA, Thrasher ET, Lang T, Davidson N (2005) A process-oriented perspective of IS success: examining the impact of IS on operational cost. Omega 34(5):448–460

Chin WW (2001) PLS Graph User's Guide: Version 3.0. Soft Modeling, Inc.

Cohan J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum, Hillsdale

Conrad DA, Shortell SM (1997) Integrated health systems: promise and performance. In: Conrad DA (ed) Integrated delivery systems: creation, management, and governance. Health Administration, Chicago

Deluca JM, Enmark R (2002) The CEO's guide to health care information systems. Wiley, San Francisco

Devaraj JM, Kohli R (2000) Information technology payoff in the health-care industry: a longitudinal study. J Manage Inform Syst 16(4):41–67

Dowling WL (1997) Strategic alliances as a structure for integrated delivery systems. In: Conrad DA (ed) Integrated delivery systems: creation, management, and governance. Health Administration, Chicago
Enthoven AC, Tollen LA (2005) Competition in healthcare: it takes systems to pursue quality and efficiency. Health Affairs W5:420–433
Etchen LL, Boulton JL (2000) Designing the IDS business portfolio. Healthcare Financial Management 29–33
Friedman L, Goes J (2001) Why integrated health networks have failed. Front Health Serv Manage 17(4):3–28
Friedman CP, Wyatt JC (1997) Evaluation methods in medical informatics. Springer, New York
Gefen D, Straub D (2005) A practical guide to factorial validity using PLS-graph: tutorial and annotated example. Commun Assoc Inform Syst 16:91–109
Hair JP Jr, Anderson RE, Tatham RL, Black WC (1998) Multivariate data analysis. Prentice-Hall, Upper Saddle River
Hamilton C, Jacob JM, Koch S, Quammen RL (2004) Automate best practices with electronic healthcare records. Nurs Manage 35(2):40E–40F
HIMSS 15th Annual HIMSS Leadership Survey (2004) http://www.himss.org/2004survey/ASP/healthcarecio_home.asp/. Accessed 1 Nov 2005
Hinkin TR (1995) A review of scale development practices in the study of organizations. J Manage 21(5):967–988
Hsiao RL, Ormerod RJ (1998) A new perspective on the dynamics of information technology-enabled strategic change. Inform Syst J 8:21–52
Kaplan B, Brennan PF, Dowling WL, Friedman CP, Peel V (2001) Toward an informatics research agenda: key people and organizational issues. J Am Med Inform Assoc 8(3):235–241
Kilbridge PM (1998) The role of information systems in IDS-physician relationships. Healthc Financ Manage 52(6):31–34
Killian RW (1999) Providers eye integration to improve care, lower costs. Manag Health 9(5):36–37
Kim YK (2000) An impact of integrated delivery system on hospital financial and quality performance under managed care. Dissertation.
Knoedler J (2003) IT integration transforms the care continuum. Nurs Manage 34:8–10
Lapiedra R, Smithson S, Alegre J, Chiva R (2004) Role of information systems on the business network formation process: an empirical analysis of the automotive sector. J Enterp Inform Manag 17(3):219–228
Laplante A (2005) Integrated systems improve medical care, control costs, according to Enthoven. Stanford Report, October 26, pp 1–3
Lewis BR, Byrd TA (2003) Development of a measure for the information technology infrastructure construct. Eur J Inform Syst 12:93–109
Melville N, Kraemer K, Gurbaxani V (2004) Information technology and organizational performance: an integrative model of IT business value. MIS Quarterly 28(2):283–322
Nunnally JC (1978) Psychometric theory. McGraw-Hill, New York
Page S (2003) Virtual health care organizations and the challenges of improving quality. Health Care Manag Rev 28(1):79–92
Parker VA, Charns MP, Young GJ (2001) Clinical service lines in integrated delivery systems: an initial framework and exploration. J Healthcare Manage 46(4):261–275
Payton FC, Ginzberg MJ (2001) Inter organizational health care systems implementations: an exploratory study of early electronic commerce initiatives. Health Care Manage Rev 26(2):20–32
Rockart JF, Earl MJ, Ross JW (1996) Eight imperatives for the new IT organization. Sloan Manage Rev 38(1):43–55
Rogers J, Lutz S (2003) Cost of caring: key drivers of growth in spending on hospital care. PricewaterhouseCoopers pp 1–19
Rosow E, Grimes SL (2003) Technology's implications for health care quality. Nurs Admin Quarterly 27(4):307–317

Sambamurthy V, Chin W (1994) The effects of group attitudes toward GDSS designs on the decision-making performance of computer-supported groups. Decis Sci 25(2):215–242
Schumaker AM (2002) Inter organizational networks: using a theoretical model to predict effectiveness of rural health care delivery networks. J Health Hum Serv Adm 25:372–406
Simoens S, Scott A (2005) Integrated primary care organizations: to what extent is integration occurring and why. Health Serv Manage Res 18(1):25–40
Skipper JB, Craighead CW, Byrd TA, Rainer RK (2008) Towards a theoretical foundation of supply network interdependence and technology-enabled coordination strategies. Int J Phys Distrib Logist Manag 38(1):39–56
Smith AE, Swinehart KD (2001) Integrated systems design for customer focused health care performance management: a strategic service unit approach. Int J Health Care Qual Assur 14(1):21–28
Snyder KD, Paulson P (2002) Healthcare information systems: analysis of healthcare software. Hosp Top 80(4):5–12
Straub D, Rai A, Klein R (2004) Measuring firm performance at the network level: a nomology of the business impact of digital supply networks. J Manage Inform Syst 21(1):83–114
Subramani MR (2004) How do suppliers benefit from information technology use in supply chain relationship? MIS Quarterly 28(1):45–73
Tallon PP, Kraemer K, Gurbaxani V (2000) Executives' perceptions of the business value of information technology: a process-oriented approach. J Manage Inform Syst 16(4):145–173
Tapscott D, Caston A (1993) Paradigm shift: the new promise of information technology. McGraw-Hill, New York
Tierney WM, Miller ME, Overhage JM, McDonald CJ (1993) Physician inpatient order writing on microcomputer workstations: effects on resource utilization. J Am Med Assoc 269(3):379–383
Venkatraman N (1994) IT-enabled business transformation: from automation to business scope redefinition. Sloan Manage Rev 35:73–87
Waring T, Wainwright D (2000) Interpreting integration with respect to information systems in organizations—image, theory and reality. J Inform Technol 15:131–148
Weiner BJ, Savitz LA, Bernard S, Pucci LG (2004) How do integrated delivery systems adopt and implement clinical information systems. Health Care Manage Rev 29(1):51–66
Young DW, McCarthy SM (1999) Managing integrated delivery systems: a framework for action. Health Administration, Chicago

Chapter 7
Beyond the Use of Robotics: Operations and Supply Chain Control for Effective Inventory Management in a Health System Pharmacy

Maari L. Loy[1], Rodney D. Traub[2], Limin Zhang[2], Pratap Kotala[3], Monte Roemmich[1], Jesse Breidenbach[1], and Robert Nelson[1]

[1]Sanford Health Medical Center Fargo, Fargo, ND, USA

[2]College of Business, North Dakota State University, Fargo, ND, USA

[3]College of Science and Mathematics, North Dakota State University, Fargo, ND, USA

Abstract This study describes a process improvement initiative conducted at Sanford Health Medical Center—Fargo an academic tertiary hospital that recently implemented an inventory management system. The objective of this project is to identify opportunities for improvement in inventory management and use of various drug dispensing technologies. Data was collected from wholesaler purchases, patient charge histories, as well as reporting from a robot, carousel system, and automated dispensing cabinets. Ultimately, the initiative uses supply chain management techniques to identify and implement appropriate inventory levels through utilization of a periodic inventory system. This reveals inventory cost history, cost upon initiation of automation, and forecasted costs with appropriate inventory levels upon implementation. The primary outcome upon implementation showed a 25.96 % decrease in cycle stock. Secondary outcomes included an increase frequency of drug being ordered (116.7 orders/week vs. 200 orders/week for top 100 drugs), supporting evidence showing 0.95 % of drugs have a weekend rate greater than one unit larger than the weekday rate and a decrease in whole orders sent to the wholesaler from 5/week to 4/week.

This study provides critical insight and practical guidelines to improve operational efficiency and cost effectiveness in a health system pharmacy. Such improvement efforts are becoming common as companies work to improve their operational efficiencies (Interfaces, 41(1):66–78, 2011).

M.L. Loy, Pharm.D., M.B.A. (✉)
Sanford Health Medical Center Fargo, Fargo, ND, USA
e-mail: maari.loy@sanfordhealth.org

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*, Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_7

Keywords Pharmacy operations • Continuous inventory system • Periodic inventory system • Pharmacy automation • Operational efficiency

7.1 Introduction

The use of drug distribution technologies has grown substantially over the last decade. In 2007, 82.8 % of hospitals used automated dispensing cabinets, 10.1 % utilized robots, and 12.7 % reported using carousel systems for medication delivery. The adoption and implementation of automation is a costly process and is acquired on the premise that these costs will be offset by enhanced patient safety, greater staff productivity, and reduced inventory expenses (Pedersen and Gumpper 2008, Farasyn et al. 2011).

Automated drug distribution combined with inventory management systems present a number of new challenges including building confidence in automation reports, integration of newer systems with existing inventories that are difficult to automate, and transfer of inventory levels within automation to the wholesaler for appropriate reorder (Skibinski et al. 2007). Furthermore, the system does not address the critical nature and varying demand of life-saving drugs. With 211 new drug shortages in 2010, the recent drug shortage crisis which nearly mandates stockpiling is also not addressed (Critical Connections 2011) within standard automation procedures.

(A background on drug shortages: over the last few years, many common medications have become shortage items. Many operating pharmaceutical plants were/are sole providers of some of these life-saving medications and have run into regulatory problems. The crisis continues due to reasons described as the Food and Drug Administration lacking the authority to manage drug shortages and the need for stronger regulation of the "distribution of pharmaceutical products" (Institute for Safe Medication Practices 2011).)

7.2 Daily Operations

Sanford Health Medical Center—Fargo is a 550 bed licensed facility spread across two campuses. Drug delivery is accomplished through a robot, one carousel unit, an automated tracking shelf and 31 automated dispensing cabinets. All of these utilize barcode technology for increased patient safety (Skibinski et al. 2007). Barcoding is also accomplished at the bedside of the patient within the facility.

The main campus robot and carousel system is responsible for the daily cart-fill. The cart-fill is described as all scheduled doses for 24 h starting at 1100 h. The cart-fill is accomplished nightly starting at approximately 0300 h. The scheduled doses are delivered to the patient care units at approximately 0800 h. The automated dispensing cabinets are utilized for controlled substance IIs and as needed medications. These are restocked daily and on an immediate basis as medications are depleted from the cabinet.

Upon dispensing of the cart-fill, the medications are charged to the patient. The automated dispensing cabinets are charged to the patient upon withdrawal from the cabinet. Many health system pharmacies charge upon administration for more efficient inventory control and ease of patient charging (Troiano 1999).

Currently robot and carousel are restocked daily based on usage reports. A 7 day supply of pre-packaged medications is ready to be placed on the robot at any time. The robot is restocked with the intent of having 3 days supply of hospital usage within the robot and is restocked daily a few hours prior to cart-fill based on the usage reports. Medication is also ordered and prepackaged based on automated usage reports to the wholesaler. The carousel is currently restocked daily based on usage reports and minimum and maximum levels identified by the automation.

The multiple types of technologies that are currently in use greatly facilitate the distribution of drugs. However, as growth continues, there was a motivation to look for other systematic changes that could improve internal efficiencies. This was a primary motivation for this study.

7.3 Methods

A retrospective study was undertaken with a date range of January 1, 2010 to December 31, 2010. Areas of interest included weekend vs. weekday volume, drug shortages and critical medications. Rationale for the study was to address current inventory cost carried by the pharmacy department and to maximize the rewards with the implementation of new pharmacy automation. The top 100 medications out of 3437 with the highest average usage rate and pricing data available from the wholesaler were included in the study. The study did not include medications and vaccines that primarily were dispensed to clinics.

To analyze the current system, information was drawn from a variety of sources. As medical databases are often large, complex, heterogeneous, and time-varying (Kaur and Wasan 2010), the data used in this study were drawn from four databases and were restricted to the year 2010 and early 2011. They included reports from the main wholesaler (88 % of drug volume) detailing all drugs purchased during the year 2010 (39,400 entries). This report provided data on quantities ordered, shipped, and shortages. It also provided costs. This data was used in estimating annual savings.

A second database was the purchases made during the year 2010 (3131 entries). This has similar data to the wholesaler database but is in aggregate form. For example, if a particular drug was ordered each week, the wholesaler database would have 50 entries whereas this second database would generally have a single entry. This database also included purchases from multiple wholesalers.

A third database was the carousel medication daily usage report. This report provided detailed data from December 23, 2010 to March 21, 2011 for 1328 drugs and provided the current ordering information for these drugs as well as a moving average for daily usage.

Table 7.1 Distribution of demand for baclofen

Daily net demand	Frequency (days)
Less than 0	44
0	99
1–10	167
11–20	46
21–30	8
Over 30	1

A fourth database which was the main source of data for an inventory analysis came from the daily database. The daily database has an entry for a drug on any day when there is a transaction such as a dispensing of the drug or a return of the drug (1,803,597 entries). Recall that the robot completes the cart-fill early in the morning. It is not unusual for an order that was intended to be given to the patient later in the day to be changed or discontinued prior to it being administered. When this occurs, the drugs are routed back to the pharmacy and eventually re-loaded into the robot. In such case, they are recorded as a negative demand. If there is neither an order nor a return for a drug on a particular day, then there is no entry in the database for that drug for that day.

Consider the medication baclofen as an example. For the year 2010, there were 276 entries in the daily database. That means on the remaining 89 days, there was neither positive demand nor restocks. The database had 10 days with an entry of zero. On those days, demand would have been equal to restocks. Thus on 99 days, the inventory at the end of the day was equal to the inventory at the start of the day. However, on 10 of those days, there was activity. The distribution of net demands is given in Table 7.1. The 44 negative demands reflect days in which restocks were larger than any demand that occurred.

The average daily (net) demand was 5.46 units with a standard deviation of 7.05 when considering only those 276 days where the database contained an entry. When the 89 days with no activity are included, the average falls to 4.13 units and the standard deviation to 6.56. The maximum demand on any day was 32 units and the minimum demand was -24 units.

As illustrated in the above example, the daily demand of a drug may vary significantly. In the current process, an order is manually placed when the drug stock drops to a certain level. Thus, orders may be placed at any time for any quantity without considering the demand pattern of each drug.

Critical medications were addressed by surveying operating room pharmacists, critical care pharmacists and internal medicine pharmacists. These pharmacists made recommendations of medications that were absolutely needed and could not be forgone as an inventory cost saving measure. This was accomplished with the history of current utilization of little used medications.

7.4 Inventory Systems

Two basic types of inventory systems were considered: A continuous system and a periodic system (Krajewski et al. 2010). In a traditional continuous system, an order is placed when inventory has dropped to a certain level. Thus an order can be placed at any time. Comparatively, in a traditional periodic system, an order is placed after a fixed amount of time has passed. For example, an order might be placed every Wednesday or every other Wednesday. Two main factors influenced the choice towards a periodic system. First, the periodic system allows orders to easily be grouped. That is, a single order for multiple drugs can be placed at the same time. In a continuous system, these orders may occur at different times. To group them together would require additional procedures to be implemented. Second, there are staffing and wholesaler limitations contributing to the avoidance of ordering on weekends. Since the continuous system could easily result in weekend orders, it was decided that a periodic system would be best for this health system pharmacy.

In a periodic system, the order quantity will vary from period to period based on the demand that has occurred since the last order. For each product, a target value is established. When it is time to place an order, the on-hand inventory is subtracted from this target value and the result is the quantity of the product that is to be ordered. Considerable research has been done on periodic models. For a good review, the reader is referred to Federgruen (1993). A periodic system for guaranteed service supply chains has been studied by Bossert and Willems (2007).

One of the challenges faced was establishing this target value while making the process convenient for employees. In some environments, demand may follow a normal distribution. When this is true, the target is relatively easy to establish. When demand is not normal, the process of establishing the target can be rather involved. A common approach is to use simulation bootstrapping (Davison and Hinkley 1997) to approximate the distribution. This allows the target value to be approximated from the historical demand rather than a known distribution.

To address this issue, a number of drugs were evaluated and a target was established in two ways. First, normality was assumed and the classical approach was used. Then the bootstrapping method was used to establish that target. If these two targets provided similar results, it was determined that implementation of the most understood one (assumption of normality) would be suitable.

Assuming that demands were normally distributed, the target value is given by the following formula: Target $= d * (LT + P) + Z * \sqrt{LT + P} * S_d$

Definition of the variables and discussion of how this information for the drug baclofen is utilizable.

d is the average daily demand and S_d is the daily standard deviation of demand. For baclofen, these values are 4.13 and 6.56 as discussed earlier.

LT is the lead time or the time it takes for an order to arrive. For this health system pharmacy this is 1 day. Generally orders are received the next day, the exception being that Friday, Saturday, and Sunday orders are received on Monday.

P is the time between orders. For the purposes of this example, we assume that an order is placed every 4 days.

Z is drawn from the standard normal table and establishes the desired cycle service level or the percent of time that a shortage will not be incurred. A Z-value of 2.0 sets the cycle service level at 97.7 %. This means that in about 97.7 % of the cycles, the next order will arrive before a shortage occurs. That is, in about 2.3 % of the cycles, a shortage will have occurred before the next order arrives.

Using this data, the target value for baclofen assuming a normal distribution will be 64.64 or about 65. Now whenever an order is to be placed, the on-hand inventory is subtracted from this number to obtain the order size. So if the on-hand inventory were 35, an order for 30 would be placed.

One might point out that with an average daily demand of less than 5, this seems like rather large values when the order is only to cover 4 days of demand. But the volatility of daily demand is quite large and desiring to have a high cycle service level dramatically increases the target value. In the above formula, the first term $d * (LT + P)$ is the average units needed to cover the lead time and the period length of 4 days. The total for this term is about 20.6 or less than a third of the target value. The second term (starting with Z) is safety stock. It is inventory carried to protect against uncertainty. In this instance, it is a large portion of the target value.

The above analysis assumes that demand is normally distributed. However, this is not the case. To verify what the proper up-to-quantity should be, a simulation bootstrapping method was employed. Demands were drawn randomly from the year 2010 data and simulated to estimate the demand distribution over a 5-day period. (Note that this is a 1 day lead time and coverage over 4 days, similar to previous approach.) The result was the histogram provided in Fig. 7.1. Approximately 97.5 % of the replications resulted in demands of 50 or less. Note that in 7.2 % of the replications, total demand over 5 days was less than or equal to zero.

For baclofen, using a target value of 50 would appear to provide about a 97 % cycle service level and a target value of 60, a cycle service level in excess of 99 %. Thus, the value derived by assuming normality (65) actually provides a high

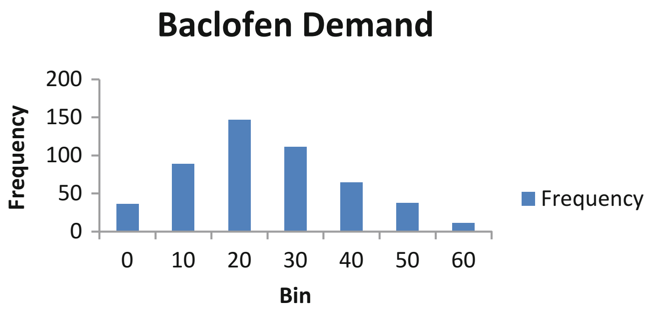


Fig. 7.1 Frequency of demand over 5 days for baclofen

cycle service level in this instance. It would appear that assuming normality sets a high service level at least for some drugs and might be useful for initial testing purposes.

7.5 Recommended System

As stated earlier, this study focused on the top 100 volume drugs. Those drugs which had characteristics such as high demand, unit dose pre-packaged medications, non-refrigerated medications, and medications that could be prepared and bar-coded much in advance were included. These can also be identified as carousel and robot items. Excluded were intravenous admixtures, pediatric oral doses needing to be manually drawn up, vaccinations, refrigerated medications, and other manually dispensed medications. The resulting set of drugs is expected to provide a template which can ultimately be expanded to other drugs in the pharmacy. Furthermore, since these are all high volume drugs, the potential impact, both in terms of inventory and costs could be substantial.

For these drugs, orders can be placed from the wholesaler on any day, but there is a surcharge for emergency orders placed on either Saturday or Sunday. Thus, there is a desire to only order during weekdays. A policy of ordering each drug once a week was examined, but having a week's demand in an order suggested a large increase in cycle stock (half the order size). So ordering for a shorter period was also found to be preferable. The recommendation was made to order each drug twice a week.

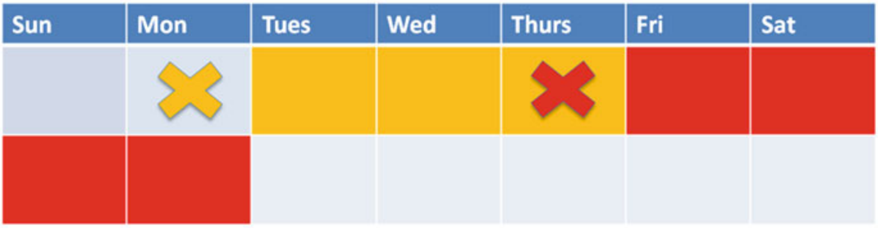
To obtain the twice a week order, one order is expected to last for 3 days and the other for 4 days. Two separate ordering patterns were presented. In pattern A (Fig. 7.2), orders are placed on Monday and Thursday. The Monday order arrives Tuesday and is intended to last through Thursday, covering 3 days. The Thursday order arrives Friday and is intended to last through the following Monday, covering 4 days and including weekends. The periodic system chosen provides coverage for weekend orders due to the 3-day/4-day ordering pattern.

Pattern B performs in a similar manner except that orders are placed on Wednesday and Friday. Here the Wednesday order arrives Thursday and is intended to last through Sunday. The Friday order arrives on Monday and is intended to last through Wednesday. Note that both patterns A and B would need to be modified to accommodate holidays.

This system would require that two target levels be identified, one for the 3-day period and one for the 4-day period. One concern with this approach was that the two different target values could cause confusion. As a result of this concern, orders would be placed as though they were always to cover demand for 4 days. This puts slightly more inventory in the system than is necessary during the week, but it greatly simplifies the procedure and also reduces the risk of having a shortage during the week.

Initially concern existed that for some drugs, the weekend demand may be higher. However an analysis showed that only 0.95 % of drugs have an average weekend

Pattern A



Pattern B

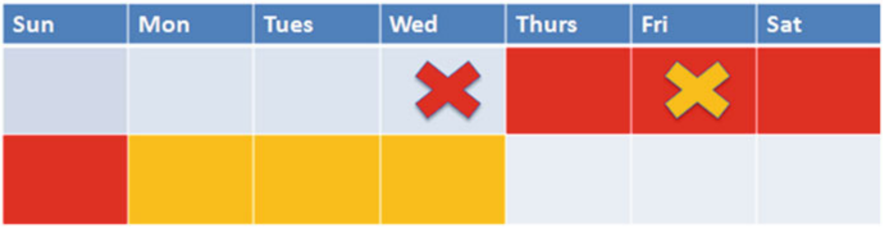


Fig. 7.2 Pattern A displays the Monday, Thursday ordering pattern and time period covered. Pattern B displays the Wednesday, Friday ordering pattern and time period covered

rate that is more than one unit higher than the weekday rate. This is attributed to a smaller amount of elective surgical procedures, rotation of internal medicine service providers, and an overall desire to have patients discharged prior to the weekend. Since demand for some drugs is lower on the weekend, the risk of running out over the weekend is lower as well, providing a buffer. Note that when demand is lower on the weekend, it is the weekday demand that is used to determine order quantities.

7.6 Outcomes

Using the described system, the primary outcome, defined as the change in cycle stock (half the order size) was estimated to decrease by 25.96 %. Orders would also be placed on fewer days. In the current system, orders are placed every weekday and sometimes on weekends. In the proposed system, orders would be placed 4 days per week. The proposed system does place more total orders. The current system averaged 116.7 orders per week for the top 100 drugs utilized. The proposed system would average 200 orders per week for the top 100 drugs. The increased number of orders is the primary reason why the cycle stock in the system has decreased.

The proposed system allows the health system pharmacy to adjust the safety stock level to achieve the best cycle service level for their needs simply by changing the Z-value. By changing the Z-value, they can increase or decrease the amount of safety stock in the system. For the drugs we examined, using a Z-value of 3 provides a high service level. For baclofen the service level was in excess of 99 %. If a particular drug can also be used as a substitute for another drug, it may be desirable to carry even more of the drug to satisfy that need. If a drug has a ready substitute, less could be carried to save on inventory costs.

Many drugs come in package sizes larger than one and this provides an additional buffer against shortages. For example, duloxetine 30 mg capsules are supplied in package sizes of 30 units. If the system target value indicates a need to order 50 units in that cycle, two packages or 60 units would actually be ordered. Of course, these extra 10 units would ultimately serve to reduce the size of a future order, but they do have the effect of increasing the average on-hand inventory. In the case of duloxetine, each time an order is placed, 0–29 extra units are ordered. This will have the effect of raising average on-hand inventory levels by 29/2 or 14.5 units.

7.7 Discussion of Recommended System

The recommended system allows the pharmacy to improve their performance in a number of ways. As noted earlier, the cycle stock is reduced by 25.96 %. This occurred primarily due to an increased number of orders. The reduced cycle stock translates to a reduction in costs as well. However, since not all drugs have the same cost, the resulting cost savings could be either larger or smaller than 25.96 %.

This system also lends itself to automation. Once target values specific to the health system pharmacy are established, it would be straightforward to further automate the ordering process. Procedures could be implemented that provide for minimum order sizes, package sizes greater than one, and adjustments for long weekends and holidays. Of course, there would need to be a level of confidence that the inventory levels are being properly reported. It may be reasonable to implement a cycle-counting program (Piasecki 2003) in order to verify inventory levels on a regular basis. A cycle counting program verifies the inventory count of a small number of items each day. Over a period of time, eventually all items are examined. With the technology being used in the pharmacy, cycle counts could be conducted in a number of ways. For example, the robot could be programmed to move a drug from one location to another, verifying the inventory level in the process. Or a portion of a drawer in the carousel could be selected and those inventory levels could be verified.

Drugs could be categorized as to how willing the pharmacy is to have a shortage occur. This could allow different Z-values to be used for different categories of drugs. If a drug is critical and has no ready substitutes, the Z-value can be increased to help insure that a shortage does not occur. The Z-values can then be adjusted as circumstance warrants. However, caution should be used in simply increasing Z-values as the extra safety stock will increase inventory costs.

The proposed system is also easy to learn. Change management can be a challenge, but the systematic placing of orders should be easy to understand and implement. If orders are placed manually, the on-hand inventory need only be determined. As procedures are automated, quantities ordered can easily be seen and an intervention can occur if necessary. Currently the intravenous admixture area and manual fill sections of the pharmacy are not automated. Using a similar system throughout the pharmacy should make it more straightforward for inventory management.

An initial shortcoming of the approach is that it is only using one year of data to establish order points. However, here again, automation could be used to update the average demand values as time passes. It would be necessary to track weekday vs. weekend demand rates to avoid the problem of understating average weekday demand just because weekend demands tend to be lower.

Another limitation is that the data used came from a single wholesaler. The particular wholesaler analyzed provided 88 % of the pharmacy usage, so the impact of other wholesalers may be limited. However, other wholesalers are used and some adjustments may be necessary as a result.

The proposed system assumes that the patient census is relatively stable and that seasonality is not an issue. Both of these factors will need to be considered. The demand rates used were an annual average. Changes in patient census or the seasonality of a drug can cause the demand for a particular month to be higher or lower than the average annual demand. If these issues are not properly addressed, the pharmacy will be at risk of shortages during certain times of the year and will be carrying unnecessary inventory during other times. The periodic system could be adjusted to handle these unstationary demands (Neale and Willems 2009).

This study focused primarily on drugs which were of relatively high volume and were ordered for use in the hospital. The health system pharmacy also serves as an intermediate step for local clinics to obtain drugs. Approximately 102 health system owned clinics place drug orders through the pharmacy. This provides for economies of scale in ordering. However, it also causes certain challenges for the pharmacy. For example, say that a particular clinic ordered a small quantity of baclofen. Rather than placing an individual order from the wholesaler, the pharmacy may instead choose to take it from their stock and just increase the next order by the appropriate amount. This increases the chance that the pharmacy will run short of baclofen in the interim for the in-hospital patients. The proposed system does not currently include this potential source of demand in setting the target values. Caution must be used when considering these other sources of demand.

Finally, this study focused on a relatively small percentage of the total drugs the pharmacy handles (100 out of 3437). As the system is implemented, technology such as data mining will facilitate the inclusion of additional drugs in a similar system by categorizing drugs based on their shared characteristics. Such techniques can also be used in the pharmacy to improve internal efficiencies (Chen et al. 2011). The system can be further enhanced with RFID technology that has been shown to achieve incremental cost-saving over conventional barcodes in the context of drug inventory management (Çakıcı et al. 2011).

7.8 Conclusion

As health system pharmacy pushes forward for improved operational efficiency and cost effectiveness in an ever-changing environment, a periodic inventory system will well serve automated inventory management systems. This will provide greater confidence in automation reports, integration of newer systems with existing inventories that are difficult to automate, and will address the critical nature and varying supply and demand of life-saving drugs specific to each health system pharmacy.

References

Bossert JM, Willems SP (2007) A periodic-review modeling approach for guaranteed service supply chains. Interfaces 37(5):420–435

Çakıcı ÖE, Groenevelt H, Seidmann A (2011) Using RFID for the management of pharmaceutical inventory—system optimization and shrinkage control. Systems, Decision Support

Chen LC, Chen CH, Chen HM, Tseng VS (2011) Hybrid data mining approaches for prevention of drug dispensing errors. J Intell Inf Syst 36(3):305–327

Critical Connections. Drug Shortages and Quality Care in the ICU (2011). 10: 2, 1

Davison AC, Hinkley DV (1997) Bootstrap methods and their application. Cambridge University Press, Cambridge

Farasyn I, Humair S, Kahn JI, Neale JJ, Rosen O, Ruark J, Tarlton W, Van de Velde W, Wegryn G, Willems SP (2011) Inventory optimization at Procter & gamble: achieving real benefits through user adoption of inventory tools. Interfaces 41(1):66–78

Federgruen A (1993) Centralized planning models for multi-echelon inventory systems under uncertainty. In: Graves SC, Rinnooy Kan AH, Zipkin PH (eds) Handbooks in operations research and management science, vol 4, Logistics of production and inventory. North-Holland Publishing Company, Amsterdam, pp 133–173

Institute for Safe Medication Practices (2011) Gray Market, Black Heart: pharmaceutical gray market finds a disturbing niche during the drug shortage crisis. 16:17

Kaur H, Wasan SK (2010) An integrated approach in medical decision-making for eliciting knowledge. Ann Inf Syst 7:215–227

Krajewski LJ, Ritzman LP, Malhotra MK (2010) Operations management: processes and supply chains. Hall, Prentice

Neale JJ, Willems SP (2009) Managing inventory in supply chains with nonstationary demand. Interfaces 39(5):388–399

Pedersen C, Gumpper K (2008) ASHP national survey on informatics: assessmentof the adoption and use of pharmacy informatics in U.S. hospitals—2007. Am J Health Syst Pharm 65:2244–2265

Piasecki DJ (2003) Inventory accuracy: people, processes and technology. Ops Publishing, Kenosha WI

Skibinski K, White B, I-Kuei Lin L, Dong Y, Wu W (2007) Effects of technological interventions on the safety of a medication-use system. Am J Health Syst Pharm 64(1):90–96

Troiano D (1999) A primer on pharmacy information systems. J Healthc Inf Manag 50:41–52

Chapter 8
Decision-Theoretic Assistants Based on Contextual Gesture Recognition

José Antonio Montero[1], **Luis E. Sucar**[2], **and Miriam Martínez**[1]

[1]Acapulco Institute of Technology, Av. Instituto Tecnológico S/N, Crucero del Cayaco, C.P. 39905, Acapulco, Guerrero, Mexico

[2]Department of Computer Science, National Institute of Astrophysics, Optics and Electronics, Luis Enrique Erro #1, Tonanzintla, Puebla, Mexico

Abstract This paper presents a novel approach that combines computer vision and decision theory for building intelligent assistants. It considers situations in which a person interacts with surrounding objects, where the system determines the most probable activity and based on it selects an action according to certain parameters. This framework is applicable to situations in which decisions are based on human activities and their interactions with objects in the environment. Examples of this type of situation include a caregiver that helps a handicapped person or an automatic video conference system that selects the best view according to the speaker's actions. The system assumes that the human activity can be recognized based on hand gestures and their interaction with relevant objects present in the environment. The proposed approach combines contextual-based gesture recognition with a decision theoretic model for selecting the best action in uncertain conditions. Gesture recognition is based on hidden Markov models, combining motion and contextual information, where the context refers to the relative position of the hand to a nearby object. The posterior probability of each gesture is used in a Partially Observable Markov Decision Process (POMDP) to select the best action according to a utility function. The POMDP is implemented as a dynamic decision network (DDN). Experiments in two settings, videoconference and human care giving, show promising results in both gesture recognition and action selection. The experiments show that the proposed framework is robust to changes in the parameters (lookahead, probabilities and rewards), and shows that the performance is similar to that of a human assistant.

J.A. Montero (✉)
Acapulco Institute of Technology, Av. Instituto Tecnológico S/N, Crucero del Cayaco, C.P. 39905, Acapulco, Guerrero, Mexico
e-mail: jamontero1@prodigy.net.mx

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*, Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_8

Keywords POMDPs • Dynamic decision networks • Intelligent assistant • Gesture recognition

8.1 Introduction

In this paper we propose a novel framework for building intelligent assistants that combines gesture recognition and decision theory. For instance, consider a videoconference room in which a speaker is giving a talk to a remote audience and for this task he is using a computer, pencil, written notes and books. In this setting a human assistant selects the best view to show to the audience according with the speaker's activity. However, after some time this task becomes very tedious for the human assistant, thus an automatic system that replicates (or does) this work would be a great benefit. Another case is one in which an elderly or handicapped person needs to be guided to complete a task, for instance hand washing. This elderly person interacts with common objects in a certain order presents on a washstand (water faucet, soap and towel); when the person can't complete the task, a human caregiver must guide the person. If the human caregiver is supported by an automatic system that guides the elderly to complete the task, and only call her when necessary, then the human caregiver could attend to other responsibilities (maybe more people). Intelligent assistants built according to the proposed framework can support human activities in many situations with minor changes in their structure. In these situations we distinguished two main aspects: (1) a person interacts with surrounding objects to do a task while she is being observed by a human assistant and, (2) based on the behavior observed, the human assistant makes decisions according to certain parameters or objectives.

In this work we propose an automatic decision-theoretic controller that supports the tasks performed by human assistants in dynamic and uncertain environments, and those tasks which can easily be adapted to different domains. This system includes two main components: (1) a computer vision component that observes the behavior of a person and determines the most probable activity, (2) a decision-theoretic controller component that selects the best[1] action according to certain preferences based on the goals of the system. Both components are probabilistic graphical models which are adequate for decision making in uncertain conditions. Activity recognition is done using hidden Markov models (HMMs) (Rabiner and Juang, 1993) which combines motion and context information. The outputs of the HMMs are incorporated into a dynamic decision network (DDN) (Howard and Matheson, 1984; Russell and Norvig, 2002). DDNs provide a formalism for capturing the various types of knowledge involved in a decision problem and offer reliable algorithms for computing preferred decisions. Additionally, DDNs provide a much more natural way to specify the dynamics of a system, including the effects

[1] Action with the highest expected utility.

of actions and observations, by exploiting problem structure in a utility framework. This combination of HMMs and DDNs provide a powerful framework for building decision-theoretic assistants which can be easily adapted for different situations. This combination is also robust to changes in the parameters (probabilities, rewards) as shown in the experiments.

Automatic vision-based systems for monitoring the behavior of people have been the focus of recent research. The majority of these works are based on human motion recognition in the context of gait recognition, or simple activity detection (Arseneau and Cooperstock, 1999; Ren and Xu, 2002; Zobi et al., 2003). Few have attempted to explicitly include aspects of actions and utility (Hoey and Little, 2004; Trevor and Alex, 1996) for activity recognition. In order to make decisions, previous work have used Markov decision processes (MDPs) (Bandera et al., 1996; Levner et al., 2003). Others have used partially observable Markov decision processes (POMDPs) (Hoey and Little, 2004; Hoey et al., 2005; Trevor and Alex, 1996). In contrast with previous work, we approximate POMDPs as DDNs with limited lookahead (number of time periods), which allow us to obtain approximate solutions in real time. Additionally, this structured representation facilitates the integration with the HMMs used for activity recognition.

Activity recognition is based on hidden Markov models (HMMs) (Rabiner and Juang, 1993) that integrate motion and context features. In this paper, we consider *context* as the relative position of the user, in particular the user's hand, with respect to the relevant objects in the environment. This extends our previous work (Montero and Sucar, 2004, 2006a,b) where we showed that gesture recognition is more reliable by combining motion and context features. In this paper we include a more comprehensive experimental evaluation, and incorporate the gesture recognition component to the decision-theoretic controller.

The main contribution of this work is a general framework for building decision-theoretic assistants based on human gesture recognition. This framework is based on the integration of HMMs and DDNs, and can be *easily* adapted to different domains. While achieving this, we make two additional contributions:

1. A gesture recognition system based on hidden Markov models that combines motion and contextual information.
2. A decision theoretic controller based on DDNs for selecting the best action based on human gestures, which approximates a POMDP, and obtains and approximate optimal policy based on a limited lookahead.

We have tested this approach in two complex environments: (a) A videoconference setting in which the system selects the best view according to the speaker's gestures, and (b) A caregiver system that helps an elderly or handicapped person to complete a task: hand washing. Both are challenging environments in which the decisions are based on human activities and their interaction with objects in the environment, and require a real-time response. Experimental results are presented in terms of (1) gesture recognition, (2) a user study, (3) the effects of different stages of lookahead in the utility of the decisions, and (4) an empirical analysis of the response time and sensitivity of the model. Results obtained indicate that, at least in these environments, a few stages in the DDNs are sufficient to approximate optimal decisions allowing to the proposed model to operate in real-time applications. They also demonstrate that the proposed framework is robust to changes in the parameters, that is rewards and probabilities. In terms of the user study, the system shows a similar performance to that of an experienced human assistant in both domains.

The rest of the paper is organized as follows. In Sect. 8.2 we contrast our approach with related work. Section 8.3 presents an overview of the gesture recognition component, and Sect. 8.4 describes the decision-theoretic controller. Description of domains are explained in Sect. 8.5. Experimental results are presented in Sect. 8.6. Finally in Sect. 8.7 we give conclusions and directions for future work.

8.2 Related Work

Few approaches exist that integrate vision techniques with decision-theory for building intelligent assistants. The majority of these approaches are used separately in different applications. Levner et al. (2003) use Markov decision processes for the recognition of buildings in aerial images. The goal is to learn a control policy to choose the next action (image processing operator) in each step, so the image quality is optimal for interpretation. A vision-based, adaptive and decision theoretic model of human facial displays for interaction is proposed by Hoey and Little (2004). They integrate vision and decision theory, using POMDPs as a stochastic method to make decisions. These works apply vision techniques in which decision-theory is incorporated as feedback so that the human activity recognition process is more reliable, or to enhance the quality of the images.

Regarding the use of context for human activity recognition, Ayers and Shah (1997) use context information to recognize human activities performed in a room. The recognition process is based on prior information, and modeled using a deterministic automaton. In another approach to action recognition in an office environment, Moore et al. (2000) combines the analysis of sensory data and symbolic contextual information provided by a scene model; the goal is the recognition of unknown objects interacting with the hand. Ayers considers that an activity is recognized when the hand of a user is over a known object, for example, the user is using the computer if its hand is over the mouse. In general, this is not sufficient to recognize an activity, because the hand of the person can be over the object and do nothing, so we extended this consideration and added motion (the gesture trajectory). Moreover, we considered contextual information and motion as features which perform more reliably the activity recognition process.

Some works related to intelligent assistants for guiding persons to complete tasks follow. Pollack et al. (2003) create *Autominder*, a new reasoning formalism based on Quantitative Temporal Bayesian Networks (QTBNs), intended to help older adults adapt to cognitive decline and continue the satisfactory performance of routine activities. Autominder uses a range of AI techniques to model an individual's daily plans, observe and reason about the execution of those plans, and make decisions about whether and when it is most appropriate to issue reminders. In the same sense, Mihailidis et al. (2001) create COACH, a cognitive aid for Alzheimer's patients that actively monitors an user attempting a hand washing task and offers assistance in the form of task guidance (prompts or reminders) when it is most appropriate. These systems are focused on a particular application, while we propose a general framework for gesture recognition. Additionally, they do not incorporate context for gesture recognition and do not use DDNs for decision making.

In the proposed framework we use a Dynamic Decision Network (DDN) model for decision making for several reasons. This model allows to represent explicitly the different observation, state and decision variables, which provides a transparent structured representation which is easier to build and adapt to different domains. Additionally, algorithms for solving influence diagrams (Bhattacharjya and Shachter, 2010; Cooper, 1990) can be adapted to solve DDNs to provide approximate optimal policies in real-time. Finally, the outputs of the HMMs used for gesture recognition are easily incorporated into the DDN.

Our approach differs in three main aspects from previous work: (1) it integrates context and motion information in hidden Markov models for gesture recognition, (2) it uses a dynamic decision network with few time-steps or lookahead (POMDP with finite horizon) to decide the best action based on gesture information, and (3) it can be applied to different complex, real environments with minor changes in its structure.

8.3 Gestures Recognition Component

The overall architecture of the proposed system is illustrated in Fig. 8.1. It has three main components: (1) image processing and feature extraction, (2) gesture recognition and (3) decision-theoretic controller. Components (1) and (2) are briefly described in this section (for a detailed explanation see Montero and Sucar 2004, 2006a,b). The third component is presented in the following section.

8.3.1 Image Processing and Feature Extraction

Image processing and feature extraction for gesture recognition consists of three main parts: (a) detection of human body parts, (b) hand tracking, (c) recognition of relevant objects.[2] Next we describe briefly each component.

[2]Objects of interest in this setting are those objects in the environment with a high possibility of being used by the user.

Detecting human body parts. The recognition of some body parts (i.e. face, hands) is performed using a color-based approach. The human skin color is usually more distinctive and less sensitive to illumination changes if we use the *rgy* normalized color space obtained by Martinez and Sucar (2006). Next, to determine if a blob in the image contains skin pixels we apply the technique known as histogram intersection proposed by Swain and Ballard (1991). However, to improve the constraint of a fixed threshold value indicated by Ballard, we are using the Otsu algorithm (Otsu, 1979). Based on Otsu's algorithm, our method incorporates adaptive[3] thresholding, so it is able to tolerate changes in lighting conditions. Initial detection of the hand combines the color-based approach with motion information to make it more robust with respect to occlusions and illumination changes.

Hand tracking. Once we have detected skin regions in an image sequence, the next step consists on tracking the hand using only color information. Currently we assume that the user performs all the gestures mainly with one hand and the other hand is basically static. For hand tracking, we are guided by two basic rules. The first rule considers that only the hands of the person cause a significant movement in the image sequence. The second rule establishes a minimum threshold (number of skin labeled pixels) that a region must have to be considered a human hand. Then the system starts tracking it. During tracking, we obtained the centroid of the right hand region.

Object detection and tracking. We recognize manipulative gestures (human interacting with surrounding objects) performed by humans in an environment. In this work, the objects presented in a domain are modeled using a color-based approach, for this our work is based on Swain and Ballard (1991) and Bradski (1998). Objects are represented using color histograms in the *hsv* color space, and

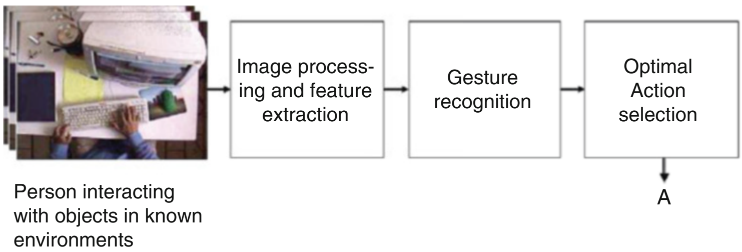


Fig. 8.1 General architecture of the system. It has three main components: (1) image processing and feature extraction, (2) gesture recognition, and (3) selection of the optimal action

[3]For to obtain a binary image from a video operating in not controlled environmental conditions, fixed value for threshold is not recommendable, so, to use a method that obtain the best threshold value based on histogram values is a better choice, Otsu algorithm applies statistical measures to do this.

Fig. 8.2 Example of object detection in two domains. *Left image*: washstand. *Right image*: office environment. Each object detected is *highlighted* and closed in a *rectangle*

they are detected using histogram intersection (Swain and Ballard, 1991) so we obtain an image in gray scale where pixels close to 255 represent the detected object. Figure 8.2 illustrates the object detection process in two domains: a washstand and an office environment. In the washstand environment (left), the objects detected are: a towel (red rectangle), a soap (green rectangle), and a faucet (blue rectangle). For the office situation, the relevant objects are: a book (red rectangle), a mouse (green rectangle), the monitor's screen[4] (light blue rectangle), and a sheet of paper (dark blue rectangle). Once an object is detected, we use the tracking algorithm proposed by Bradsky (1998) in an appropriate search window for each object, where the relative distance between objects and the hand centroid represents contextual information. The gesture recognition process integrates this information with the features extracted from the hand trajectory.

Feature extraction. An adequate selection of visual features is very important for the success of the gesture recognition process. In this work, we considered the set of basic features (angle (ϕ), direction (ρ) and velocity (v) of the gesture trajectory in polar coordinates) proposed by Yoon et al. (1999). Then we added a chain code (Gonzalez and Woods, 1992) with a set of 16, 32 and 64 direction numbers as chain code elements to represent the *shape* of the hand gesture trajectory. This set of four features was combined in order to obtain the minimal set of features that best represent the gesture trajectory. Based on an experimental evaluation we demonstrated that the combination of angle and distance (ϕ, ρ) in polar coordinates, is the minimal set of features needed for the gesture recognition process, it is obtained using a HMM with ten states (Montero and Sucar, 2004). Next we incorporate context information as an additional component to the features vector, so the feature vector is determined by:

$$F = (\rho, \phi, Context) \qquad (8.1)$$

[4]Because of the difficulty of detecting the screen of the computer based on histogram color technique. The screen is detected using a color mark around it.

8.3.2 Gesture Recognition Using Context

We model gestures using the hidden Markov models (Rabiner and Juang, 1993) framework, widely used for modeling temporal structures. However, HMMs for gesture recognition do not incorporate contextual information, so we propose an approach that integrates motion and contextual features to support the gesture recognition process. To recognize the activity performed by a human in dynamic domains, we focus our attention on human gestures performed by the hand when it interacts with surrounding objects. Context information to support gesture recognition has been used in others works (Ayers and Shah, 1997; Moore et al., 2000; Thonnat and Rota, 2000), but their methods impose many restrictions or are computationally complex, so it is difficult to apply them to real environments. We propose a simple approach based on the relative position of the hand with respect to relevant objects in the domain. At each time period, the Euclidean distance (in 2-D) from the right hand centroid to all relevant objects is computed and considered as an additional feature incorporated in the HMMs for gesture recognition. Then, the complete set of features used by the HMMs include: the gesture trajectory parameters (ρ, ϕ) and the distance between the hand centroid and all relevant objects $(d(O_i))$:

$$F = (\rho, \phi, d(O_1), d(O_2), \ldots) \tag{8.2}$$

O_i indicates any of the relevant objects presents in the domain. Given that the features are continuous, we made them a discrete finite set of symbols using vector quantization (Gray, 1984). We then used these symbols to construct the codebook to train HMMs to perform recognition.

Experimental results. The type of gestures considered in the experiments are those performed in two domains: an office environment and a washstand room. The gestures that the recognition system will try to identify are the following: *erasing, writing, using the mouse, turning the leaves of a book and speaking* for the office environment (see Fig. 8.2 right); and *opening the faucet, taking the soap, soaping the hands, taking the towel, drying the hands, and closing the faucet* for the caregiver situation (see Fig. 8.2 left). We used two different data sets, one set for each domain. In the office domain, a database with 4200 data points was generated for training and testing. The data was obtained from 100 gesture sequences with five different gestures performed by a person in a sitting position and interacting with objects in the office environment. In the washstand domain we used a database with 3600 data points generated for training and testing. These data were obtained from 50 gesture sequences with 6 different gestures realized by a person in a standing position and interacting with surrounding objects. In this case, we used 2560 data points for training and 1040 data points for testing. In both cases, gestures data were captured by a ceiling mounted camera pointed downwards, under normal—not controlled—illumination conditions in an office and a bathroom situation, respectively. We tested

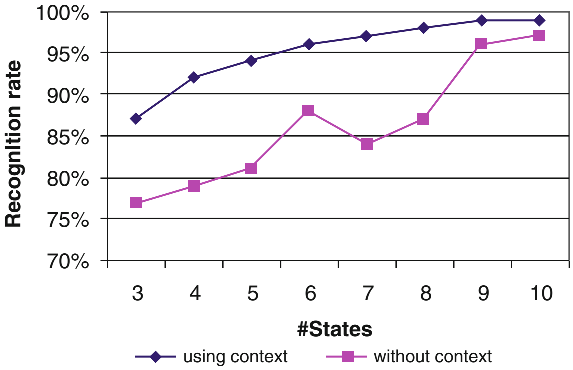


Fig. 8.3 Gesture recognition rates with and without context information vs. number of hidden states for the videoconference domain

the gesture recognition system for 11 different types of gestures, related to the manipulation of each object in the two situations.

Trials performed in the domains cited demonstrate that the gesture recognition process is more reliable if we incorporate contextual information as part of the features. Results illustrated in Fig. 8.3 for the videoconference situation indicate that using only the motion feature, that the recognition rate varies from 75 % for three hidden states to 97 % for ten hidden states. In the other case, if we integrate motion and contextual features the recognition rate varies from 88 % using three hidden states to 99.47 % using ten hidden states. Thus, there is a significant improvement by incorporating context. Similar results were obtained in the caregiver situation where the average of recognition reached was of 94 %.

The high recognition rates obtained in both domains suggest that our gesture recognition model is appropriate to be incorporated to the decision-theoretic controller, which is described in the next section.

8.4 Decision-Theoretic Controller

Once the gesture recognition component (see Fig. 8.1) is operating in reliable way, we integrate it in a model that can make decisions based on this information. Our objective is to design a system that can support human tasks in dynamic environments. This means that the controller models scenarios in which decisions, variables or preferences can change over time. The proposed model must take the *best* action based on the activities made by a person according the system goals. Actions are selected by the system according to the utilities or preferences which are directed by the domain, so, sometimes the model operates as a reactive system and other times—with minor changes in its structure—operates as instructional system. Most of the time, user behavior is affected by the actions selected by the model, but there is uncertainty about this behavior; so we are not certain about the next state of the user when an action is selected. The current state of the user is also uncertain, given that the gesture recognition process is not perfect. Thus, we have uncertainty on the current state and on the results of an action, at same time we have to reach certain goals directed by the domain. Then, the suggested technique for modeling this type of problems are Partially Observable Markov Decision Processes (Puterman, 1994).

8.4.1 Partially Observable Markov Decision Process

Partially observable Markov decision processes (POMDPs) are the principled framework for agent planning under uncertainty (Cassandra et al., 1994; Kaelbling et al., 1996). Uncertainty in the action effects and the state of the world are modeled probabilistically. Objectives are encoded with utility functions whose magnitude are indicative of their relative importance. Despite the considerable expressivity of POMDPs, their use in real-world systems remains limited due to the intractability of the solution algorithms for finding optimal control policies. Nevertheless, real-world POMDPs tend to exhibit a significant amount of structure, which can often be exploited to improve the scalability of solution algorithms. In particular, many POMDPs have simple policies of high quality. Hence, it is often possible to quickly find those policies by restricting the search to some class of compactly representable policies. In addition, when states correspond to the joint instantiation of some random variables (features), it is often possible to exploit various forms of probabilistic independence (e.g., conditional independence and context–specific independence), decomposability (e.g., additive separability) to mitigate the impact of large state spaces. In a POMDP, the transition, observation and reward functions all use the state space to define their domain. For problems with very large state spaces, several techniques allow us to represent these functions in a compact way. Despite the possibility of specifying very large POMDPs in a compact form, it is not straightforward to design solution algorithms that also work directly with a compact representation without ultimately enumerating all or most states.

Formally, a discrete time POMDP is a tuple $\langle S, A, T, R, O, B\rangle$, where S is a finite set of states, A is a finite set of actions, $T : S \times A \rightarrow S$ is a transition function which describes the effects of the agent's actions upon the world states. $R : S \times A \rightarrow \Re$ is a reward function which gives the immediate reward for taking action A in state S, O is a set of observations, and $B : S \times A \rightarrow O$ is an observation function which gives the probability of the observations for each state-action pair. Given a POMDP, the goal is to find a *policy* that maximizes the expected discounted sum of rewards. Since the system state is not known with certainty, a policy maps either belief states (i.e., distributions over S) or action-observation histories into choices of actions.

In this work we used a special class of POMDP representation called Dynamic Decision Networks (DDNs) (Russell and Norvig, 2002), that approximates optimal solutions to POMDP with finite horizon. A DDN is like a Dynamic Bayesian Network (Dean and Kanazawa, 1989) except that it has decision and utility nodes in addition to chance nodes.

8.4.2 Modeling the Problem

DDNs are composed of three types of nodes: (1) chance nodes which represent random variables. In this work we consider two chance nodes (a) state (node S in Fig. 8.4) that indicates the set of activities realized by a user in some domain, and (b) evidence (node O in Fig. 8.4), that represents the information proportioned by the vision component (most probable activity); (2) decision nodes which indicate the set of actions (node A in Fig. 8.4) that the decision maker can select; and (3) utility nodes which represent the preferences (node U in Fig. 8.4) of the model according to certain goals. The components of the DDN model for the types of environments considered in this work are the following:

States: The state space is characterized by a random variable that represents the activities (hand gestures) realized by a person in the domains considered (office and washstand).

Evidence: Represents the information obtained by the visual gesture recognition process, modeled in this work by hidden Markov models (HMMs). It consists of the most probable activity realized by the user according to the HMMs.

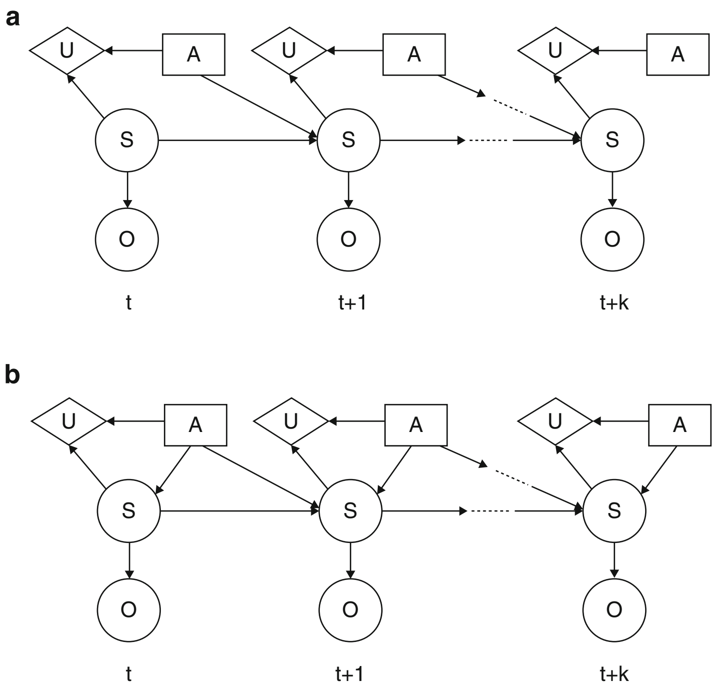


Fig. 8.4 The *K* stage dynamic decision network allows us to model different domains with minor changes in its structure, (**a**) DDN that model the caregiver setting and (**b**) is the DDN model that represents the videoconference domain

Actions: The actions selected by the system are defined using expert knowledge and depend on the domain. Actions are directed by the activities realized by the user in each domain.

Utilities: Utilities are associated with the preferences of the different actions selected by the system in the corresponding domain. Utilities are based on the objectives of the controller. The utility value depends mainly on the gestures performed by the user and indicates the preferences of selected actions when a specific activity is observed. The system must make an adequate balance between several objectives when selects an action in each setting.

Transition function: The transition function defines the probability that the system goes to the next state S_{t+1} (expected activity in time $t+1$) given that in the current state S_t (activity recognized in current time), that the given action a_i is selected. In the type of settings considered, this transition can be predictable with some degree of certainty. For example, for the video conference application, in general we expect that the speaker continues doing the same activity (gesture) for some time, so changes of activity are not frequent. This implies that in this scenario the probability of staying in the same state is high, with low probabilities of changing to another state. However, in the caregiver setting, we expect that the person continues doing the same activity (e.g., *open the faucet*) for some time and then after changes to another activity (e.g., *soap the hands*) in an expected order.

The specific model for the two scenarios considered in this work are detailed in Sect. 8.5.

8.4.3 Solving the DDN

As we said before, we solve the POMDP by approximating the result with a dynamic decision network (DDN) (Russell and Norvig, 2002) representation with a finite number of steps or lookahead (see Fig. 8.4). In this model, the state variables (S) represent the activity of the user at each time step, which are not directly observable. The observation nodes (O) represent the information obtained from the vision system—modeled by the HMMs used for gesture recognition —. The action nodes (A) correspond to the different actions that can be selected by the controller at each time. Finally, the rewards (R) represent the immediate reward that depends on the current state and preferred action.

In the DDN representation of a POMDP, every chance node (random variable), S_i, is associated with a set conditional probabilities $P_i = P(S_t|S_{t-1}, A_{t-1})$. The utility node, U, is associated with a set of $u_i = u(S_{t-1}, A_{t-1})$, specifying for each *action–state* pair, a number expressing the desirability of this combination for the decision maker.

The optimal action is obtained after analyzing k time-steps advancing (lookahead) in the evaluation of the DDN (see Fig. 8.4).

A DDN represents a decision problem. A solution to problem is a decision or, in the case of multiple decision nodes, a sequence of decisions that maximizes the sum of rewards in the future. To compute a solution for each sequence of actions, the utilities of outcomes are weighted by the probabilities that these outcomes will occur. The expected utility of an action sequence, a_i, is thus computed from:

$$\hat{u}(a_i) = \sum_i u(\pi_i(r))Pr(\pi_i(r)|a_i) \tag{8.3}$$

where $\pi_i(r)$ is a combination of values for the parents of the reward node r, $u(\pi_i(r))$ is its utility, and $Pr(\pi_i(r)|a_i)$ is the probability of $\pi_i(r)$ given that the decisions a_i are made. In this work, the preferred sequence of actions is a sequence with the maximum expected utility. We are using a clustering technique proposed by Shatcher and Peot (1992) to obtain this policy. The optimal action is selected according to the most probable activity (state) computed by the vision system. In this framework, we approximated an optimal solution for a POMDP by solving the DDN for k time steps. The number of stages depends on the domain, in this work we use two settings: (1) video conference and (2) caregiver. For each scenario we evaluated different look aheads, and selected the one for which there is a *good* compromise between optimality and computation time.

8.5 Evaluation Scenarios

Trials were realized in two settings:

- a caregiver that helps a simulated elderly to complete the task: hand washing.
- an automatic video conference system that selects the best view to show to the audience.

Both experiments were performed in realistic conditions with some minor restrictions to simplify the early vision component; in particular the persons used long sleeves and were right handed.

We first describe each scenario, and then we describe how the parameters of the models were defined.

8.5.1 Caregiver Setting

The objective for the caregiver setting is to guide the user to complete a task using an adequate selection of prompts. We consider the particular task of *hand washing*. The objective of the proposed model in this scenario is that a caregiver guides an elderly or handicapped person to complete this task. We select the hand washing task because this task is better adapted to our vision system, but we could have selected any other task.

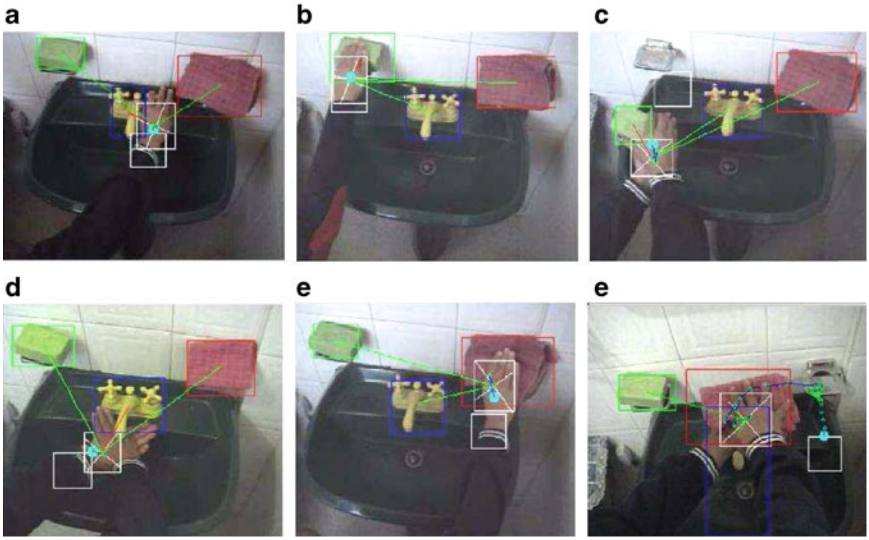


Fig. 8.5 Sequence of images showing different steps performed by the user to complete the task *cleaning the hands*. The gestures recognized are: (**a**) opening/closing the faucet, (**b**) taking the soap, (**c**) using the soap, (**d**) washing the hands, (**e**) taking the towel, and (**f**) drying the hands

The set of subtasks (activities) that must be realized in this domain are the following: (a) *Using the faucet*, this activity is detected when the hand is moved over the faucet which implies that the faucet is opened or closed, (b) *Washing the hands* when the user hands are below the faucet, the vision system recognizes this activity, (c) *Taking the soap*, is recognized when the hand is moved near the soap position, (d) *Soaping the hands* this activity is recognized when the soap is in the hands and the hands are below the faucet, (e) *Taking the towel* this is recognized when the hand is moved to towel position; and (f) *Drying the hands* if the towel is in the hands and a characteristic motion is detected, this subtask is recognized.

The relevant objects in the washstand environment are the soap, faucet, and towel. During its operation, the system detects the behavior of the user when he or she is interacting with these objects. Then it selects an action (audible prompts) to guide the user to complete a task if (he or she) is not following the correct sequence of subtasks, or it may simply do nothing (null action) if the user is performing correctly the sequence of steps required. There is an additional action (*Call for help*), which is selected when the controller detects that the same instruction is repeated in three consecutive times. To define the correct sequence of steps to complete the task, we used a *task graph*, similar to the one used by Boger et al. (2005).

To apply our model to this scenario, we define the next state and action variables. State variable has 6 possible values: s_1 = *opening the faucet*, s_2 = *closing the faucet*, s_3 = *using the soap*, s_4= *drying the hands*, s_5 = *taking the towel* and s_6 = *washing the hands*. Figure 8.5 shows some typical images for each one of the gestures considered in the mentioned setting. Also, we considered eight actions that correspond with the

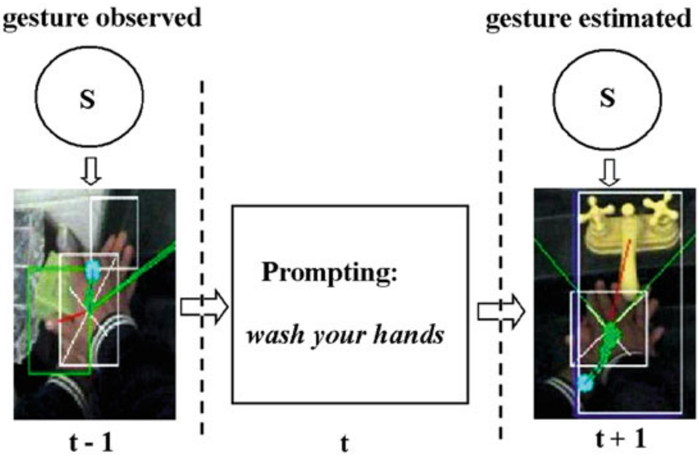


Fig. 8.6 An example of the caregiver setting. It shows the action with the maximum expected utility selected by the system in time t, while observing the gesture realized at time $t-1$ (*using the soap*)

possible prompts selected by the system: a_1 = *open the faucet*, a_2 = *close the faucet*, a_3 = *soap the hands*, a_4 = *wash the hands*, a_5 = *take the towel*, a_6 = *dry the hands*, a_7 = *Null* and a_8 = *call for help*. These prompts are designed to get the user's attention.

For example, in Fig. 8.6 we show the action selected by the system when the most probable gesture at the time $t-1$ was *using the soap*; then, the system must guide the user to the most desirable action in the sequence selecting the prompt *wash the hands*. The idea is to select the appropriate prompt (audible instruction) to help the user to complete the task (cleaning the hands). One possible application of this model could be to guide elderly people to perform different tasks and to reduce the need for human assistance.

This experiment was conducted in a bathroom of the home of one of the authors. The environmental (lighting) conditions were not controlled. Participants were "normal"[5] persons that *simulated* to be elderly and in need of caregiver to realize and complete some task: *washing the hands*.

8.5.2 *Video Conference Setting*

The second scenario in which we tested the proposed model was a video conference setting, in which a speaker uses different resources to present a conference to a remote audience. We assume that there are several cameras, and the goal of the

[5]Users with good health and without disabilities.

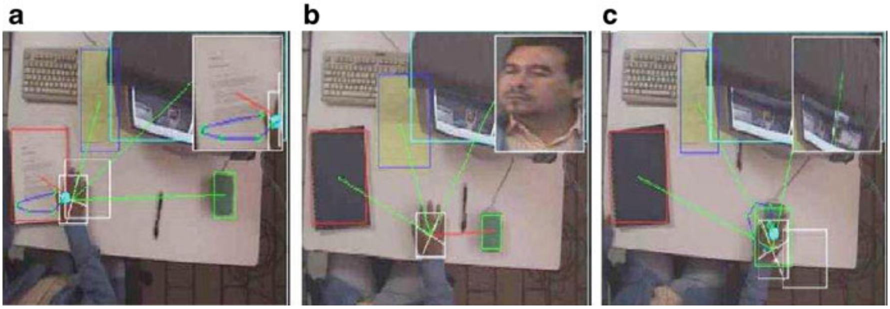


Fig. 8.7 A sequence of images showing the speaker's activity at different times during a presentation: (**a**) the speaker is turning the pages of a book, (**b**) the speaker is explaining something (no relevant gesture), and (**c**) the speaker is interacting with the computer. For each case, the selected view is shown in the *top-right corner* of the image

system is to select the best view of the videoconference environment and to show it to the audience. In this case, the model operates as a reactive system, this means that the action selected is to send the best view to a remote audience according to the current activity.

The activities in this domain are the following: (a) *Writing*, this activity is recognized when the speaker takes a pencil and her hand is over the notebook or a piece of paper; (b) *Turning the pages of a book*, activity realized by the speaker when she is turning the pages of a book to clarify the exposition; (c) *Speaking*, this is the default activity, this means that is recognized when the right hand of the speaker is not interacting with known objects; and (d) *Mouse*, in this activity the user is using the computer, which is recognized when the right hand of the speaker is over the mouse and he or she moves it.

The actions defined in this domain are: (a) *ShowFace*, this action is selected when the systems detects that the activity done by the speaker is speaking (default activity); (b) *ShowScreen*, in this case the screen of the computer is shown to the audience, this action is selected when the speaker is using the mouse; (c) *ShowNotebook,* one of the objects used by the speaker to clarify an explanation is the notebook, so when the system detects that the speaker is writing on this object, it selects this action; and (d) *ShowBook*, this action is chosen by the system when it detects that the speaker is realizing the interleaving activity.

Like the setting described earlier, Fig. 8.7 illustrates different time-steps considered in this scenario. In the right top corner of each image appears the best view selected by the system during the activities performed by the speaker. Figure 8.7a shows the speaker interleaving a book, so the best action selected by the model is to show the book to the audience. In Fig. 8.7b the speaker is explaining a theme, so the system selects the *default action* which is to show the speaker's face to the audience. In Fig. 8.7c the speaker is using the mouse, then the systems shows the computer screen to the audience.

The gestures and actions applied in the video conference environment are the following.

States: s_1 =*Writing*, s_2 = *Turning the pages of a book*, s_3 = *Speaking*, and s_4 = *Using the computer*. The default state is *Speaking*, which is considered when no relevant gesture is recognized.

Actions: a_1 = *Show face*, a_2 = *Show piece of paper*, a_3 = *Show book*, and a_4 = *Show screen*.

The experiments for this setting were performed at a university (Acapulco Institute of Technology), the speaker was situated in a room and a group of students representing the audience were situated in another room. In this scenario, like in the previous one, the lighting conditions were not controlled; the speaker is right handed, seating at a desk with a computer and several cameras.

8.5.3 Model Definition

The structure of the DDN model is basically the same for both applications and is defined in Sect. 8.4, so in this section we will focus on the model parameters.

The parameters of the DDNs for both domains were defined subjectively with the aid of domain experts. There are basically three sets of parameters that need to be specified which are utilities (or rewards), transition probabilities, and observation probabilities. The observation probabilities are basically related to the gesture recognition component, and obtained directly from the HMMs which were learned from data as described earlier. Next we describe how the utilities and transition probabilities were defined.

8.5.3.1 Utilities

Utilities are associated with the preferences of the different actions selected by the system in the corresponding domain (selects the adequate "prompt" in the caregiver setting or the "best view" in the video conference environment). Utilities are based on the objective of the controller. The utility value depends mainly on the gestures performed by the user and indicates the preferences of selected actions when a specific activity is observed. The system must make an adequate balance between several objectives when it selects an action in each setting. For example, in the caregiver setting, we must consider the details of each prompt, be directed by the clarity of the prompt and user's response to it. Those objectives are clarity, detail and effectiveness. In the environments considered in this work, we are using two different utility values. The utility values were selected experimentally. We tested different parameters, ranging from positive values (+10) to negatives values (-10), but the impact on the decision node was not important for different values (see the experimental results), so we selected the following values: $+3$ indicates a

Table 8.1 Utility values for the hand washing task, positive value of +3 indicates the preference of the system to certain action when a specific activity is recognized, and a negative value of −3 indicates the opposite

Action/activity	Manipulating the faucet	Washing the hands	Taking the soap	Soaping the hands	Taking the towel	Drying the hands
Open the faucet	+3	−3	−3	−3	−3	−3
Wash your hands	−3	−3	−3	+3	−3	−3
Take the soap	+3	−3	−3	−3	−3	−3
Soap your hands	−3	−3	+3	−3	−3	−3
Take the towel	−3	+3	−3	−3	−3	−3
Dry your hands	−3	−3	−3	−3	+3	−3
Close the faucet	−3	−3	−3	−3	−3	+3

preference; −3 indicates a penalty; and a *large penalty* has a value of −6 (used, for instance, in the case of *call for help* in the caregiver scenario). The rewards for the caregiver scenario are shown in Table 8.1, the rewards for the other domain are similar.

8.5.3.2 Transition Function

The transition function defines the probability that the system goes to the next state S_{t+1} (expected activity in time $t + 1$) given that in the current state S_t (activity recognized in current time) that the action a_i is selected. In the type of the settings considered, this transition can be predictable with some degree of certainty. For example, for the video conference application, in general we expect that the speaker continues doing the same activity (gesture) for some time, so changes of activity are not frequent. This implies that in this scenario the probability of staying in the same state is high, with low probabilities of changing to another state. However, in the caregiver setting, we expect that the person continues doing the same activity (*open the faucet*) for a certain time and then after changes to another activity (*soap the hands, take the towel, dry the hands*) in an expected order. An estimate of the transition probabilities was obtained experimentally based only on the state variables without any action from the system. This is a rough estimate as the actions are not taken into account, however, as demonstrated in the experiments, the model is robust to variations in these parameters. Next we briefly describes the procedure for both situations.

In the videoconferencing task we recorded the exposition of five experienced speakers (universities professors) with a duration of 10 min each; they exposed the same topic. We suggested to each one that they use the different expositions aid available, such as the objects located on the table the notebook, book, mouse and pencil. In this experiment, the controller was not acting, nor was the recognition system. In the case of caregiver domain, the trial conducted was slightly different, because the participants (young persons in good health) had to simulate the behavior

Table 8.2 Transition matrix for handwashing task, probabilities are the average of values manually obtained using recorded simulations of participants for to do the handwashing task

t / t+1	Using the faucet	Washing the hands	Taking the soap	Soaping the hands	Taking the towel	Drying the hands
Using the faucet	0.25	0.170	0.20	0.13	0.13	0.12
Washing the hands	0.18	0.28	0.14	0.22	0.06	0.12
Taking the soap	0.11	0.18	0.29	0.19	0.13	0.10
Soaping the hands	0.07	0.17	0.09	0.30	0.21	0.16
Taking the towel	0.21	0.12	0.17	0.15	0.27	0.08
Drying the hands	0.19	0.14	0.17	0.21	0.07	0.22

of an elderly who required some help to complete the hand washing task. Like in the previous domain, the gesture recognition component and controller were not considered. The transition probability estimates in this case are depicted in Table 8.2. We can observe that the probability that the subtask performed in S_t continues in S_{t+1} is slightly larger than the probabilities for other subtasks.

Next we describe the experiments and results obtained by the proposed model, during its operation on the environments previously described.

8.6 Experimental Results

The viability of the proposed model was evaluated in the two scenarios considering quantitative and qualitative criteria. Aspects considered for the quantitative criteria are (1) total reward, (2) response time, and (3) sensitivity. Qualitative evaluation consists on comparing the results obtained by the system versus the results obtained by a human. The experiments are divided in four aspects: *lookahead*, *response time*, *sensitivity* and *user satisfaction*.

8.6.1 Lookahead

The optimal policy is approximated by evaluating the DDN for N stages into the future (lookahead). As more stages are considered, we obtain a better approximation, but the computation time increases. Thus we experimentally selected this parameter.

For each scenario, we evaluated different lookaheads, and selected the one for which there is a *good* compromise between optimality and computation time. We consider, in both scenarios, a lookahead ranging from 3 to 10 stages. We observed that after 10 time steps the increment in accumulated utility value was not significant. Each scenario was tested under different conditions, and we averaged the results. The results obtained for the video conference scenario are summarized

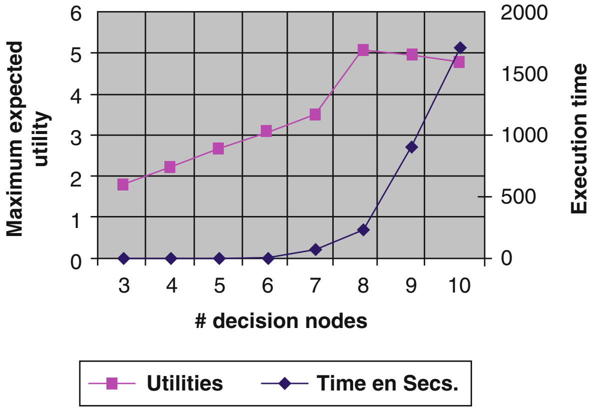


Fig. 8.8 Expected utility (*left*) and solution time in seconds (*right*) for different number of stages (decisions) in the DDN for the video conference domain

in Fig. 8.8; this illustrates the maximum expected utility (left vertical axis) versus the number of stages. It also shows the execution time (right vertical axis) reached for each lookahead. We can observe that the average expected utility reaches its maximum value at eight stages; after this stage it stabilizes. However, the execution time increases exponentially with the number of time steps, and that is very slow for real-time operation considering there are eight stages. Although there is a significant difference in utility between 5 and 8 time steps, if we analyze the actions selected in both cases, they are almost the same (consistent in more that 90 % of the cases). This phenomena is well known in MDPs, as when we solve a model, usually the policy converges before the utility values (Puterman, 1994). Thus, we considered that a good compromise between efficiency and optimality is to select a lookahead of 5 in the video conference application, which we used in the experiments for this scenario. The caregiver scenario was analyzed in similar manner and the lookahead selected for this scenario was 4.

8.6.2 Appropriate Response Time

The system must provide real-time responses in order to keep the user engaged. Thus we have to decide what is the adequate time for giving responses in the type of tasks considered. In both settings, caregiver and video conference, a reasonable time to respond is in the order of a few seconds.[6]

To evaluate the response time we considered the operation of the complete system (see Fig. 8.1), this means: (1) image processing, (2) gesture recognition

[6]Response times in the order of 1–2 s seem appropriate to maintain the interest of the audience in the videoconference domain (Tsykin and Landshow, 1998), and between 2 and 7 s to obtain the appropriate response after an audible instruction is given (Murray and VanLehn, 2000) in the caregiving application.

and (3) decision making. Because the response time of the vision component is around a fraction of a second, we are interested in the latter phase (3), which is the most expensive in terms of processing time. We know that the response time depends mainly on the inference algorithm applied to the DDN model. Therefore, we compared three algorithms: (1) the clustering algorithm proposed by Huang and Darwiche (1994), which provides exact results, (2) the logic sampling algorithm proposed by Henrion (1988), this is an stochastic sampling algorithm that provides approximate results, and (3) the likelihood sampling algorithm proposed by Shatcher and Peot (1992), which also gives approximate results. We implemented each one of the algorithms using SMILE.[7]

The response times reported correspond to the average time in seconds obtained for ten execution trails with a Personal Computer, Intel Pentium IV, 3.1 GHz processor, with 1 GB RAM; and the SMILE implementation. Results are summarized in Table 8.3. Setting A corresponds to the DDN model that represents the videoconference domain and Setting B to the DDN model that represents the caregiver domain. We observe that the different algorithms generated the following average response times for the videoconference application: (1) clustering algorithm: 5 s, (2) logic sampling with 1000 samples: 3 s and with 10,000 samples: 3 s, and (3) likelihood sampling, with 1000 samples: 3 s, and with 10,000 samples: 4 s. The response times reported on the caregiver domain, were the following: (1) clustering algorithm: 32 s, (2) logic sampling with 1000 samples: 11 s and with 10,000 samples: 12 s, and (3) likelihood sampling, with 1000 samples: 12 s and with 10,000 samples: 12 s. These response times were obtained considering the number of stages where the maximum value for expected utility was reached, this means a lookahead of 8 for the videoconference domain and a lookahead of 6 for

Table 8.3 Response times

	Response time (s)	
Algorithm	Setting A	Setting B
Exact	5	32
Approximate		
Logic sampling		
1000 samples	3	11
10,000 samples	3	12
Likelihood sampling		
1000 samples	3	12
10,000 samples	4	12

The average response time is shown for the two settings: *A* videoconference with a lookahead of 8 stages, *B* caregiver with a lookahead of 6 stages, using different inference algorithms for the DDN at the decision stage

[7]Library for probabilistic and decision-theoretic graphical models developed at the University of Pittsburgh Decision Systems Laboratory (http://genie.sis.pitt.edu/).

the caregiver domain. But in this work we are considering a compromise between efficiency and optimality, so as discussed before, we select a lookahead of 5 for the videoconference domain and a lookahead of 4 for the caregiver domain. With these conditions for each domain, we obtained an average response time of 1 s for the videoconference setting and an average response time of 3 s for the caregiver domain, using the clustering algorithm (as there is no significant difference with the approximate algorithms). These are reasonable times responses for both domains.

We can conclude that a few stages of lookahead are enough to provide near-optimal policies in reasonable times for this type of domains. We also observe that there is not significant saving in response time by using approximate algorithms, so the exact algorithm is preferable.

8.6.3 Sensitivity Analysis

In this test we are interested in the effects of the variation of the model parameters on the system performance. To evaluate this aspect we performed a sensitivity analysis varying two components of the decision model: (1) the observation probabilities provided by the vision system (system beliefs), and (2) the utility values (system goals). Next we report the results obtained for both experiments.

1. **Variation of the system beliefs**. We consider a variation of the system beliefs (probabilities) in relation with the user's behavior. In this work, system beliefs are represented by the probability values provided by the gesture recognition component (HMMs). We selected in this trial two values related to the probability of the *recognized* gesture: $p(gesture_1) = 0.7$, and $p(gesture_1) = 0.9$. Then, using the SMILE library, we computed the value for the maximum expected utility varying the number of step-times in the DDN, next we select the optimal action obtained in this way. Results are illustrated in Fig. 8.9, where the performance of the DDN operating under this condition for the video conference setting is depicted. We observe that the value for the maximum expected utility is gradually increased, reaching its maximum value between the stages 7 and 8, after this value it decreased and increased again in stage 11. Those values were obtained with utility values between -3 and +3. We observe that the behavior of both curves is very similar; and more important that the optimal action selected by the model is the same in both cases. The gesture observed in Fig. 8.9 is *writting* on a piece of paper and the maximum expected utility value was assigned to the action: *Show-piece-of-paper*, which is the best decision in this case. The procedure was repeated for each one of the gestures and the results obtained were consistent.
2. **Variation of the system goals**. In this case, the reward function (utility values in the DDN) that defines the system goals was replaced with different values. We considered for the video conference domain 3 sets of values for the rewards: $[-3, +3]$, $[-1, +1]$ and $[0, +1]$, positive values are rewards or preferences, and

negative or zero values represent penalties. For this trial we considered that the probability value for the recognized gesture was of 0.9. In Fig. 8.10 we compared the values of the maximum expected utility versus the number of time-steps or lookahead. We observed that the performance of the system is stable after a certain number of time-steps. Differences are observed when we consider negative and positive values $((-3,+3), (-1,+1))$ in the range of the values reached. In this case the maximum values are reached in the time-step 7, after the behavior is stable. In the case when we consider the interval value of (0, +1), the maximum value for the expected utility is reached in the time-step 11, then it maintained stability. In the three cases the optimal action selected by the system was the same: *show the screen*; which corresponds to the best action for the gesture *using the mouse*. Similar results were obtained for each one of the others gestures.

These experiments show that the performance of the proposed system is reliable under variations of its components: outcomes of the vision system (gestures probabilities), reward function, and the lookahead (number of time-steps). In this experiment we only evaluated the videoconference domain; since the DDN model is almost the same in the caregiver domain, we expect that the results will be similar.

8.6.4 Performance: Automatic Assistant vs. Human Assistant

Even if we solve the complete POMDP model, and obtain the optimal policy, this does not guarantee that the system will have an optimal performance in practice, as this depends on several factors including the vision and gesture recognition components, and the decision model. Thus, to evaluate how *good* the proposed

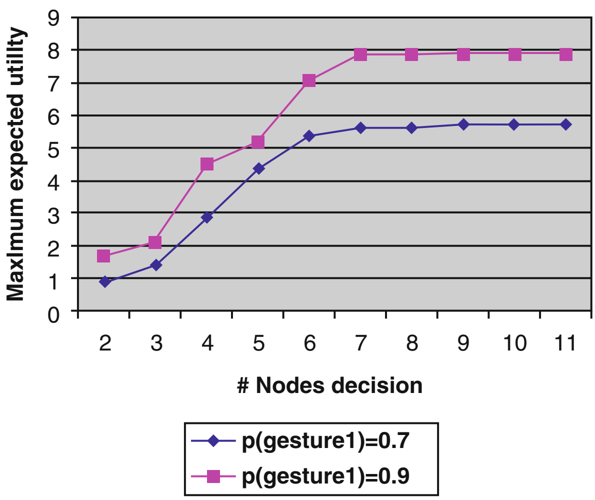


Fig. 8.9 Graphs that illustrate the behavior of the maximum expected utility vs. lookahead when changing the probability of the recognized gesture. The gesture in this case is *writing* in a video conference setting

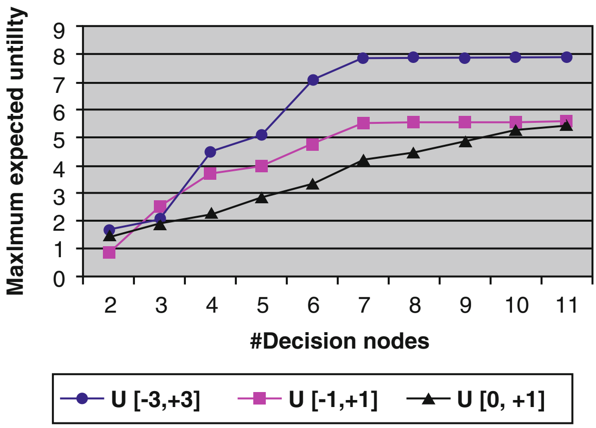


Fig. 8.10 Graphs that illustrate the behavior of the maximum expected utility vs. lookahead, varying the reward function. The gesture considered in this case is *Using the mouse* in a video conference setting

framework is in practice, we compared the decisions made by the proposed model versus the decisions of a human assistant when both perform the same task. Although there is not a guarantee that the human decisions are optimal, it is reasonable to assume that an experienced person will take "good" actions in the type of scenarios considered. Thus the system was evaluated based on a user study comparing the decisions of the automatic assistant vs. a human assistant. Results obtained in this test were supported by the Likert scale. The Likert indicator is an additive scale structured for a series of items. This scale is an indicator of the intention of the user, where the user indicates its degree of agree or disagree in relation with a measured phenomenon (Likert, 1932), in the range 1–5 (where 1 = strongly disagree, 2 = disagree, 3 = undecided, 4 = agree, 5 = strongly agree). Next, we described the aspects evaluated for each scenario.

Video conference environment. We wanted to know how the final user perceives a video conference when the view is selected by a human versus when the view is selected by the proposed model. In this experiment, an experienced university professor gave a short exposition of 10 min (see Fig. 8.7) to an audience formed by ten students at the Bachelor level. The speaker was not directly visible to the viewers, only through a computer screen. We repeated the experiment for two groups: (1) Test group: in this group, the view that was sent to the audience was selected by an automatic assistant (proposed model), (2) Control group: in this group, the view was selected by a human assistant who used a computer that allowed to him select the actions manually before beginning we indicated to him the procedure for selection. This computer was connected through a local network to a computer situated in a contiguous room, so, the selected view was sent to a remote computer as the speaker's activities were performed. After the exposition, a questionnaire was given to each student in both groups, to evaluate the following

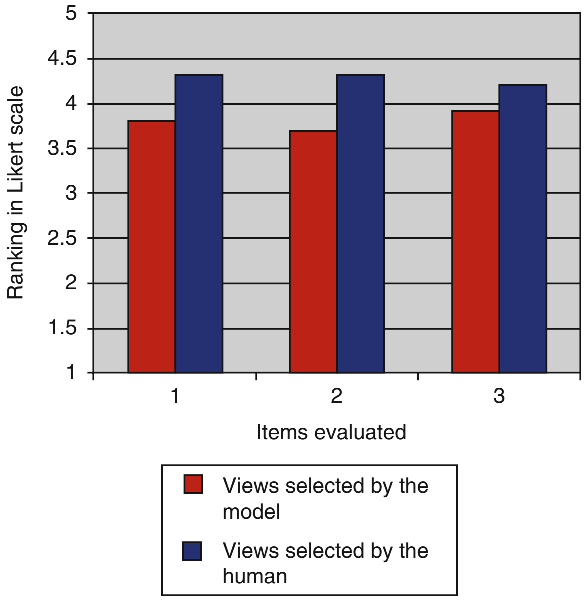


Fig. 8.11 Results obtained for the video conference setting when the decisions made by a human operator were compared with the decisions made by the proposed system. We compared three aspects: (1) consistency, (2) stability, and (3) sensibility

aspects: (1) consistency: the screen shows something related to the activity of the speaker, (2) stability: the scene must stay fixed during an adequate time so the observer can appreciate the exposition, (3) sensibility: the system must respond in a short time to changes in the speaker's behavior. Results obtained in this experiment are illustrated in Fig. 8.11.

According to the Likert indicator, there is a small advantage when the best view is selected by the human assistant, however the difference between the automatic assistant and the human counterpart is not significant. Numerical values obtained using the Likert scale indicate a difference of less than 0.5 (the recommended statistical values are < 1.0) for all the aspects evaluated. Then we can consider that the decisions of our model are satisfactory.

Caregiver scenario. In this setting, the objective is to guide a person to complete a task using an adequate selection of prompts, we did not specified a maximum time to complete the task. To evaluate the system we conducted a user study similar to the one for the videoconference scenario.

In this experiment ten normal[8] persons participated. An experienced caregiver (nurse) gave instructions to the persons that simulated the behavior of an elderly that needs some help to complete the *handwashing* task.

Persons "simulated" problems when washing their hands, such as forgetting to use the soap or doing the subtasks in the wrong order.

We divided them in two groups: five participants were used in the test group and five composed the control group. The first group was guided to complete the

[8]Persons with good health and without disabilities.

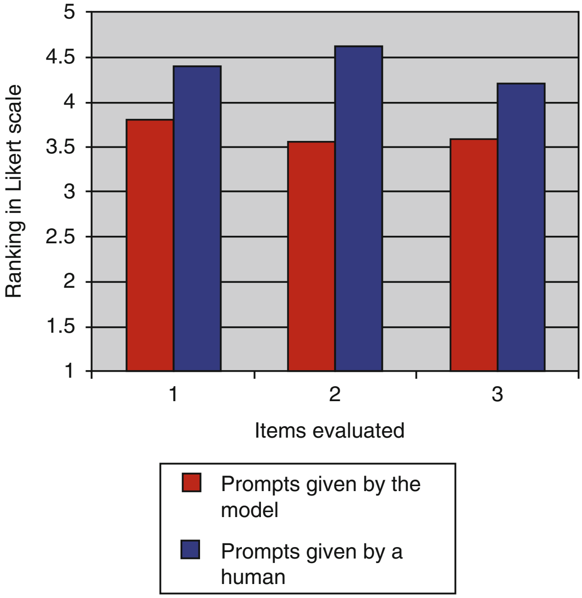


Fig. 8.12 Results obtained in the caregiver setting, where a person simulating memory problems is guided to complete the activity: *cleaning the hands*. It compares the decisions of a human assistant and our system, for three aspects: (1) clarity, (2) detail, and (3) effectiveness

task by audible prompts given by the proposed model, and the control group was guided by verbal instructions given by a human assistant. The aspects evaluated were included in a questionnaire that was completed by each one of the participants. These aspects were the following: (1) clarity of the prompt, (2) detail of the prompt, and (3) effectiveness. Detail of the prompt refers to the level of specificity of the instruction, in this case we used a general level of specificity. For instance: *take the towel* or *dry your hands*, other works have used more levels (Boger et al., 2005). Clarity considers the ease of understanding the user's message. In this work we used short phrases that can be easily understood by humans. Results of the evaluation of these aspects are illustrated in Fig. 8.12.

Based on the answers obtained from the questionnaire completed by the participants and illustrated in Fig. 8.12, the Likert scale indicates a small advantage when the prompt was selected by a human assistant to guide a person to complete the task. However the difference between the automatic assistant and human assistant was not considerable. In this scenario, two of the aspects evaluated show a low difference (less than 0.5), and one aspect, detail of the prompt, shows a more significant difference. However, this aspect does not depend on the decision model, as the prompts were predefined and just selected by the system. In the future we will improve this aspect considering the suggestions indicated by the nurse.

In general we consider that results obtained in this experiment, where our proposed model was compared versus a human, are satisfactory. That is the decisions made by the intelligent assistant were—according the Likert scale—very close to decisions of human assistants; in these domains this represents a high standard.

In current experiments, the parameters of the models were based on human expertise related to these domains. However, these parameters can be automatically learned from observations using machine learning techniques like reinforcement learning (Sutton and Barto, 1998). In the future we plan to learn the parameters for the proposed models.

8.7 Conclusions and Future Work

In this paper we presented a novel approach that combines computer vision and decision theory for building intelligent assistants. It provides a general framework for developing intelligent assistants for applications in which the decisions are based on human activities. We tested the system in two domains, but it can be adapted to other domains such as sports coaching, other types of tasks for elderly people, etc. An advantage of this framework is that it can be easily adapted to different domains with the same structure and few changes in the parameters. Some of the novel aspects of the system are: (1) A gesture recognition system based on hidden Markov models that combines motion and contextual information; (2) A decision theoretic controller to select the best action to implement using a dynamic decision network. The model proposed approximates the solution of a POMDP using an structured representation with finite horizon and Bayesian network inference; (3) An intelligent assistant that integrates the gesture recognition and decision theoretic controller, which could be applied in multiple scenarios.

We have tested our approach in two realistic environments: (1) a video conference controller that selects the best view to show to the audience based on the speaker's activities, and (2) a caregiver environment, that guides an elderly person with audible prompts to complete the task *hand washing*. We evaluate the model using several quantitative criteria and compare its performance with a human operator. Experimental results in both scenarios show that the proposed approach is robust under variations in gesture recognition, reward functions and lookahead. Even with a few time-steps of lookahead, the actions are near-optimal according to the approximate model, and this allows for real time response (few seconds) in the class of scenarios considered. A user study comparing the decisions made by the proposed model versus a human assistant indicates a satisfactory performance.

In future work, we will integrate other capabilities to extend our model. For example, the use of HMMs only allows modeling the interaction between the right hand of the user with surrounding objects, but replacing the HMMs by CHMMs, will allow us to consider both hands during the interaction, so we can export our model to a greater variety of domains. Also, we are also planning to test our system for the caregiver domain in a nursing home.

Currently we did not use the depth dimension in the vision component. In the near future we are planning to use stereo vision to solve this restriction, so we can detect movements in 3 dimensions. An important aspect that we have to solve is to learn the model (structure and parameters) on line, using machine learning techniques.

Appendix: Implementation

The system was implemented on a Personal Computer, Intel Pentium IV, 3.1 GHz processor, with 1 GB RAM. The video capture card is a Pixel-View, this card allows us to capture 30 fps, we used 320 × 240 images. Images were processed at a rate of between 12 and 15 fps. We are using a Sony TRV-19 CCD color video camera, the parameters were manually tuned. Software was implemented using Visual C++ 6.0, over Windows XP platform, also, we integrate functions of OpenCV and SMILE (Structural Modeling, Inference, and Learning Engine), a library developed in the Decisions Laboratory of the University of Pittsburgh.

References

Arseneau S, Cooperstock JR (1999) Presenter tracking in a classroom environment. In: AAAI symposium on intelligent environment, vol 1, pp 145–148

Ayers D, Shah M (1997) Monitoring human behavior in an office environment. Pattern Anal Mach Intell 7:780–794

Bandera C, Vico FJ, Bravo JM, Harmon ME (1996) Residual q-learning applied to visual attention. In: Proceedings of the thirteenth international conference on machine learning, vol 1, pp 3–6

Bhattacharjya D, Shachter RD (2010) Solving influence diagrams: exact algorithms. In: Proceedings of the twenty sixth conference on uncertainty in artificial intelligence

Boger J, Hoey J, Boutilier C, Fernie G, Mihailidis A (2005) A decision-theoretic approach to task assistance for persons with dementia. In: Proceedings of the international joint conferences on artificial intelligence

Bradski G (1998) Computer vision face tracking as a component of a perceptual user interface. In: Workshop on applications of computer vision, vol 1, pp 214–219

Cassandra AR, Kaelbling LP, Littman LM (1994) Acting optimally in partially observable stochastic domains. In: Proceedings of the twelfth national conference on artificial intelligence, vol 2

Cooper GF (1990) The computational complexity of probabilistic inference using Bayesian belief networks. Artif Intell 42:393–405

Dean T, Kanazawa K (1989) A model for reasoning about persistence and causation. Comput Intell, vol 5, 142–150

Gonzalez R, Woods RE (1992) Digital image processing. Addison-Wesley, Boston

Gray RM (1984) Vector quantization. IEEE ASSP Mag 7:407–467

Henrion M (1988) Propagation on uncertainty in Bayesian networks by probabilistic logic sampling. Uncertain Artif Intell 2:149–163

Hoey J, Little J (2004) Decision theoretic modeling of human facial displays. Technical Report TR-04-02 Department of Computer Science

Hoey J, Poupart P, Boutilier C, Mihailidis A (2005) Semi-supervised learning of a pomdp model of patient-caregiver interactions. In: Proceedings of the international joint conference on autonomous agents and multi agent systems (AAMAS), pp 1–9

Howard RA, Matheson JE (1984) Influence diagrams. In: Howard RA, Matheson JE (eds.) The principles and applications of decision analysis, Strategic Decision Group, Palo Alto, CA, USA pp 719–762

Huang C, Darwiche A (1994) Inference in belief networks: a procedural guide. Int J Approx Reason 1:1–158

Kaelbling LP, Littman ML, Moore AW (1996) Reinforcement learning: a survey. J Artif Intell Res 4:237–285

Levner I, Bulitko V, Li L, Lee G, Greiner R (2003) Automated feature extraction for object recognition. University of Alberta, Alberta

Likert R (1932) A technique for the measurement of attitudes. Ph.D. thesis, Columbia University

Martinez M, Sucar LE (2006) Learning an optimal naive Bayes classifier. In: 18th international conference on pattern recognition (ICPR'06), vol 4, pp 958–962

Mihailidis A, Fernie GR, Barbanel JC (2001) The use of artificial intelligence in the design of an intelligent cognitive orthosis for people with dementia. Assist Technol 13:23–39

Montero JA, Sucar LE (2004) Feature selection for visual gesture recognition using hidden Markov models. In: Fifth Mexican international conference on computer science, vol 1, pp 196–203

Montero JA, Sucar LE (2006a) Context-based gesture recognition. In: 11th Iberoamerican congress on pattern recognition CIARP'2006

Montero JA, Sucar LE (2006b) A decision-theoretic video conference system based on gesture recognition. In: 7th IEEE international conference automatic face and gesture recognition, FG2006, pp 387–392

Moore DJ, Essa IA, Hayes MH (2000) Exploiting human actions and object context for recognition tasks. In: Proceedings of IEEE of the 7th international conference on computer vision, vol 1

Murray R, VanLehn K (2000) DT Tutor: a decision-theoretic, dynamic approach for optimal selection of tutorial actions. In: 5th international conference of intelligent tutoring systems ITS 2000, pp 153–162

Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1:62–66.

Pollack ME, McCarthy CE, Ramakrishnan S, Tsamardinos I, Brown L, Carrion S, Colbry D, Orosz C, Peintner B (2003) Autominder: a planning, monitoring, and reminding assistive agent. Robot Auton Syst 44:273–282

Puterman ML (1994) Markov decision process discrete stochastic dynamic programming. Wiley, New York

Rabiner L, Juang B (1993) Fundamentals of speech recognition. Signal processing series. Prentice Hall, Upper Saddle River

Ren H, Xu G (2002) Human action recognition in smart classroom. In: Proceedings of the fifth IEEE international conference on automatic face and gesture recognition, vol 5, pp 54–60

Russell S, Norvig P (2002) Artificial intelligence: a modern approach, 2nd edn. Prentice Hall, Englewood Cliffs

Shatcher RD, Peot MA (1992) Decision making using probabilistic inference methods. In: Proceedings of 8th conference on uncertainty in artificial intelligence, pp 276–283

Sutton R, Barto A (1998) Reinforcement learning: an introduction. MIT Press, Cambridge

Swain MJ, Ballard DH (1991) Color indexing. Int J Comput Vis, vol 7, 11–32

Darrell T, Pentland A (1996) Active gesture recognition using learned visual attention. In: Advances in neural information processing systems (NIPS), vol 1

Thonnat M, Rota N (2000) Video sequence interpretation for visual surveillance. In: Third international workshop on cooperative distributed vision, Advances in Neural Information Processing Systems, NIPS, Denver, CO, USA, pp 1–9

Tsykin M, Landshow CD (1998) End-to-end response time and beyond: direct measurement of service levels. Work sponsored by Fujitsu Australia Limited, pp 1–13

Yoon HS, Soh J, Ming BW, Yang HS (1999) Recognition of alphabetical hand gestures using hidden Markov model. IEEE Trans Fundam 82:1358–1366

Zobi M, Wallhoff F, Rigoll G (2003) Action recognition in meeting scenarios using global motion features. In: ICASSP proceeding, vol 1, pp 115–119

Chapter 9
Developing A Method to Evaluate Emergency Response Medical Information Systems

Ann Fruhling[1] **and Stacie Petter**[2]

[1]School of Interdisciplinary Informatics, College of Information Science & Technology, University of Nebraska at Omaha,, 6001 Dodge Street, PKI 280A, Omaha, NE 68182-0392, USA

[2]Department of Information Systems, Hankamer School of Business, Baylor University, One Bear Place #98005, Waco, TX 76798, USA

Abstract Emergency response medical information systems (ERMIS) are a specific type of medical information system used for communication and decision making during a crisis. Yet given the dependence on ERMIS during a crisis, these information systems are rarely evaluated to ascertain if the system is indeed successful. This research develops a method to evaluate the success of an ERMIS using a well-established research model as a guiding framework. We explain this method in the context of an ERMIS used in the diagnosis of pathogens in hospitals and state public health laboratories. We describe the insights obtained when using this method to evaluate emergency response medical information systems.

Keywords Emergency Response Medical Information Systems • Information system evaluation • IS Success

9.1 Introduction

This paper describes the development of an evaluation process used to evaluate the success of emergency response medical information systems (ERMIS). Medical information systems, including ERMIS, are developed to address specific objectives, such as improving patient care, increasing efficiency, or reducing costs, yet these systems are rarely evaluated to determine if the system is providing the intended benefits. Yet, if organizations would evaluate the success of medical information systems, and more specifically ERMIS, potential improvements to the

A. Fruhling (✉)
School of Interdisciplinary Informatics, College of Information Science & Technology, University of Nebraska at Omaha,, 6001 Dodge Street, PKI 280A, Omaha, NE 68182-0392, USA

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_9

information system can be identified to ensure the objectives of the information system are being realized.

Information systems success is a multifaceted concept that considers many aspects when measuring the success of an information system, such as quality, use, user satisfaction, and benefits (DeLone and McLean 1992). DeLone and McLean (1992) developed a model that encompasses the multiple dimensions of information systems success which has been applied in a variety of contexts (Petter et al. 2008) including healthcare and medical information systems (Pare et al. 2005; Sicotte et al. 2009; Petter and Fruhling 2011).

Prior literature has called for the evaluation of medical information systems (van der Meijden et al. 2003) and this research responds to this call by demonstrating how to evaluate the success of a specific medical information system, an emergency response medical information system. Therefore, the purpose of this paper is *not* to evaluate the DeLone and McLean model, but rather to develop a method to formally evaluate the success of an ERMIS using a model of information system success. We further demonstrate the value of measuring the success of an ERMIS by describing the evaluation of a specific ERMIS.

This paper is organized as follows. First, we discuss the importance of emergency response medical information systems (ERMIS) and identify unique features of these systems. Then, we offer background on the information systems evaluation model used in this paper to examine the success of ERMIS. The following section describes the process used to develop the instrument and approach used to gather information from the users regarding the success of an ERMIS. Then, we explain how we used this method to obtain information from users about the success of a specific ERMIS, named STATPACK™, and use this information to evaluate the success of the system. We share our insights about the process and what was learned about the success of STATPACK™ as a result of the evaluation. We conclude with limitations of our approach and future research opportunities for evaluating the success of medical information systems.

9.2 Background

9.2.1 Emergency Response Medical Information Systems

An Emergency Response Medical Information System (ERMIS) is a specialized group decision support system that includes a systematically organized group communication system in which protocols and communication structure are provided, but there is little content about a particular crisis except in integrated electronic databases (Turoff and Van de Walle 2004). ERMIS provide necessary information for medical first responders and decision makers to determine a course of action in the event of a man-made or natural disaster. Emergency response medical information systems often require multiple decision makers to work collaboratively in a time sensitive situation where lives are at risk.

The environment during an emergency is chaotic and volatile. The critical problem in an emergency is where people focus and what resources are expended (Turoff 2002). In many situations, the process of responding to a crisis is unpredictable since almost everything in a crisis situation is an exception to the norm. Managing the exceptions to planned responses is critical in determining minute-to-minute operations during the crisis (Turoff 2002). To manage these exceptions, relevant, accurate, and up-to-date information is necessary to instill confidence in decision making, particularly since lives are often at risk. Individuals responding to emergencies work long hours around the clock and have neither tolerance nor time for distractions or information overload. During a crisis situation, people from different organizations may need to freely exchange information, delegate authority, and conduct oversight. ERMIS are used throughout the emergency to support these activities.

ERMIS have unique and challenging system requirements that affect their success. These systems must be predictable, reliable, and usable. ERMIS require real-time immediate responses in a highly secure environment. Successful ERMIS should also be used without hesitation or frustration during an emergency in that they are easy to learn and use. These systems require dynamic interaction of data, multi-level statuses and notifications, and real-time up-to-date information, but the information presented must not overload the user. The systems must be accurate, reliable and process at peak performance. Therefore, successful ERMIS are critical to optimizing the response during emergency crises situations (Turoff 2002).

9.2.2 Information Systems Evaluation

Within the information systems literature, many different criteria have been used to evaluate information systems. One criterion is technology acceptance, or the willingness of the users to adopt and use an information system (Davis 1989). Technology acceptance has been used to identify the willingness of users to accept an information system *prior to* the system's implementation in the medical information systems context (Wilson and Lankton 2004). Users' technology acceptance of a system, however, does not fully evaluate if an information system is successful (Ammenwerth et al. 2003).

After an information system has been implemented within an organization, evaluating the success of an information system can offer insight beyond whether or not the users will continue to use the information system. Within the information systems literature, DeLone and McLean (1992, 2003) suggested that the evaluation of information system success is more complex in that seven interdependent variables define success: system quality, information quality, service quality, intention to use, use, user satisfaction, and net benefits. Table 9.1 defines these variables of information systems success and identifies measures that can be used to evaluate these different aspects of information systems success.

Table 9.1 Information systems success (DeLone and McLean 1992, 2003; Petter et al. 2008)

Success dimension	Definition	Potential measures
System quality	Desirable characteristics of the technical components of the information system	Ease of use, accessibility, functionality, reliability of the system
Information quality	Desirable characteristics of the system outputs (e.g., content, reports)	Format, accuracy, relevance of the system outputs
Service quality	Quality of the support that system users receive from the information systems organization and IT support personnel	Responsiveness, reliability, technical competence, empathy of support staff
Intention to use	Likelihood that an individual will use the system the next time it is needed	Single or multiple measures of intent to use the system again
Use	Degree and manner in which staff utilize the capabilities of an information system	Nature of use, appropriateness of use, extent of use, purpose of use
User satisfaction	Users' level of satisfaction with the information system	Overall measures of user satisfaction or multiple items measuring different aspects of satisfaction
Net benefits	Extent to which the information system contributes to the success of individuals, groups, organizations, industries, and/or nations	Improved decision-making, improved productivity, cost reductions, improved profits, consumer welfare, economic development

DeLone and McLean's model of information systems success has been explored, expanded, debated, and tested in a variety of contexts (Petter et al. 2008), including medical information systems (e.g., Pare et al. 2005). Because this model considers success in a variety of dimensions, there is the potential to use this model to evaluate an information system throughout its lifecycle, which is valuable in a medical information systems context (van der Meijden et al. 2003). Furthermore, because multiple dimensions of success are examined as part of the evaluation, insight into areas requiring improvement can be identified as part of the evaluation. A final benefit of the DeLone and McLean model is that there is flexibility to adapt the measures based on the context of the system under study. Each success variable can be measured in a variety of approaches; therefore, the most relevant measures can be selected based on the system under study.

Previous research in the medical information systems field has focused on validating the DeLone and McLean information systems success model in terms of examining relationships between different success variables (e.g., Pare et al. 2005; Petter and Fruhling 2011; van der Meijden et al. 2003). Yet, research has not described how to use this model to evaluate an information system to identify potential improvements for the system and identify if the system is indeed meeting the needs of the users. Therefore, this paper develops a method for evaluating an information system using the DeLone and McLean information system success model. Considering the unique features of an ERMIS, we adopt, adapt, and create measures and practices to create a method to evaluate the success of ERMIS.

9.3 Evaluation Method

This section of the paper describes a method to evaluate the success of ERMIS. In this section, we discuss how items are identified to develop a survey to evaluate the success of an ERMIS, considering the unique characteristics of an ERMIS and its users, as well as considerations when administering the survey to ERMIS users.

9.3.1 Survey Development

To evaluate the success of ERMIS, we developed a survey instrument to capture each of the variables of information systems success identified by DeLone and McLean (2003): System Quality, Information Quality, Service Quality, Use, Intention to Use, and Net Benefits. To develop the survey, we used a multi-step approach to ensure the measures were comprehensive and appropriate given the context of an ERMIS.

1. Phase I: Creating the Question Pool

   Our first phase of survey development was the development of an initial pool of items using items from the literature (Davis 1989; Briggs et al. 2008; Wixom and Todd 2005; Doll et al. 2004; Gable et al. 2008; Chang and King 2005; Sedera and Gable 2004; Ives et al. 1983; Jiang et al. 2002; Seddon and Yip 1992). The first goal was to develop as comprehensive of a list of items as possible from the literature. Therefore, studies with validated measures of various dimensions of information system success were identified. At this point in the survey development, items were not eliminated due to lack of relevance to ERMIS.

   This first phase to develop a pool of questions that could be used to evaluate the success of an ERMIS consisted of 38 items for System Quality, 41 items for Information Quality, 22 items for Service Quality, 4 items for User Satisfaction, and 20 items for Net Benefits. All items identified in this phase of the study are listed in Appendix A.

   At this point in creating the survey pool, we specifically chose not to identify items for Use. In other contexts, such as information systems in a call center or medical information systems used for billing or patient history, Use can be measured by items such as frequency of use or duration of use (DeLone and McLean 1992). However, ERMIS are used on a non-routine basis and Use does not have the same meaning or intent. Frequency or duration of use would not be relevant in an ERMIS because this type of use would most likely be related to the frequency of emergencies, not the success of the information system. Therefore, we did not incorporate these measures of Use from the literature in this first phase of survey development.

2. Phase II: Narrowing the Question Pool

   In the second phase of survey development, we selectively removed redundant items. Our next task was to identify and remove items that are not relevant to ERMIS due to the nature of this type of information system. In this phase, we also developed new items specific to ERMIS, such as items to measure the Use construct. Since frequency of use is correlated with the number of emergencies, we wanted to measure Use in a richer format by considering the functionality and depth of system use (Burton-Jones and Straub 2006). In this stage, the list of items was reduced from 125 items to 81 items. The tables in Appendix A identify which items were retained or developed in this phase of survey development.

3. Phase III: Reviewing the Question Pool

   In the third phase we asked a small pool of ERMIS users and first-line support staff for an ERMIS to examine the survey for clarity and redundancy. We purposefully asked individuals that would offer critical feedback to examine the questionnaire. Based on this feedback, the authors revisited the survey and carefully examined each item to determine if (a) it was relevant to ERMIS; (b) if there was redundancy with any other items; and (c) if the removal of the item would affect content validity. Using these criteria, the survey items related to ERMIS success were shortened to 48 items (9 items for System Quality; 6 for Information Quality; 8 for Service Quality; 2 for User Satisfaction; 3 for Intention to Use; 10 for Use; 10 for Net Benefits). This resulted in a survey that could be completed by users in 15 min or less. Appendix A denotes the items that were retained for the final survey.

Figure 9.1 diagrams the process used to develop the evaluation questionnaire.

9.3.2 Survey Administration

After the development of the survey, multiple considerations are still necessary when administering the survey to users of an ERMIS.

The first choice is the scale for the measurement items. For example, the scale may be a 5-point, 6-point, or 7-point Likert scale (or another type of scale, such as semantic differential). Another choice in the developing the scale could be the inclusion of an additional option of "Unable to Evaluate." Given that ERMIS are not routinely used, some individuals may feel that they are unable to answer one or more questions about the system. A null response in an evaluation questionnaire could suggest that the person was unable to evaluate the ERMIS for that particular item or it could mean that the person overlooked the question. By including an

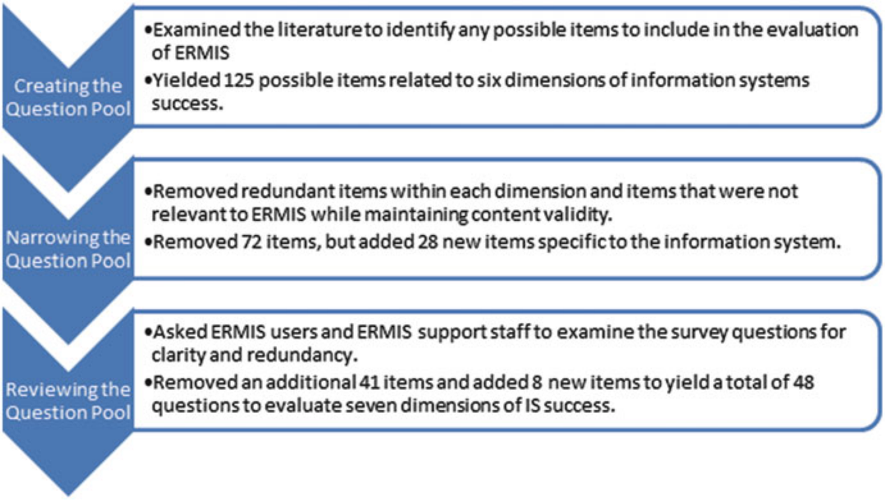


Fig. 9.1 Survey development process

"Unable to Evaluate" option, users decide whether or not they are able to evaluate the success of the ERMIS on an item-by-item basis. This increases confidence that the responses provided in the evaluation reflect the user's perceptions accurately.

Another choice in the survey administration is if whether to only gather quantitative data or if open-ended questions would provide insight into the success of the information system. If qualitative data is also solicited from users, when the results are analyzed, one can confirm that the quantitative results are consistent with the qualitative data as a form of triangulation (Jick 1979). Furthermore, with the use of qualitative data, it is possible to ask additional questions that provide insight useful to researchers, system designers, and project sponsors. These questions could ask users about useful or needed functionality to help ERMIS designers to identify the most important and necessary functionality based on the user's perspective.

Questionnaires can be administered in a variety of ways including mail, fax, telephone, and online, which each approach having advantages and disadvantages. When evaluating an ERMIS, one should consider the context to determine the evaluation method. Depending on how the system is used and the type of environment, one survey administration approach could be a better choice than others. In laboratory or clinical environments, mail, fax, or online surveys allows users to complete the survey at a time that is less hectic and more convenient for them and requires little time to complete.

Another consideration in survey administration is the hierarchical nature of medical profession. Obtaining permission from supervisors can be critical to obtaining credibility to the evaluation process and access to the users of the system. When conducting an evaluation of any information system, particularly a medical information system, it is important to receive constructive criticism about the system

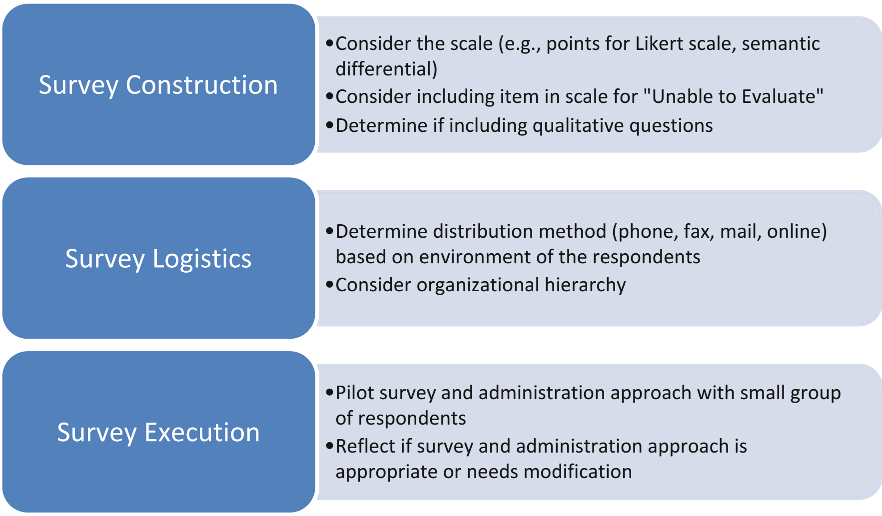


Fig. 9.2 Survey administration considerations

to know how to further improve the system. Furthermore, ERMIS are often used by a small group of people, so it is also important to identify an approach in soliciting participation that will encourage as many responses as possible to fully evaluate the system.

A final consideration in the administration of the evaluation survey is the need to pilot the survey before conducting a large-scale evaluation. Even with validated measures, a pilot is necessary to ensure that the survey development phase is valid before proceeding (Straub 1989). A pilot study ensures the questions are clear to the ERMIS users, to eliminate any redundancies in items, and to ensure the information system success variables are fully captured in the context of an ERMIS. A second benefit of conducting a pilot study is to examine the survey administration methods. The pilot allows you to identify procedures to examine choices made regarding the length of the survey, the supervisors' role in administering the survey, and to ensure the type of responses, both qualitative and quantitative, are helpful in evaluating the success of the ERMIS. Figure 9.2 illustrates considerations when administering an information systems success survey to evaluate an ERMIS.

9.4 Example of Evaluating Success of an Ermis

9.4.1 System Context: STATPACK™

A more specific type of ERMIS, a public health emergency response system, is an important component of the national information infrastructure for bioterrorism preparedness. State and local Public Health Laboratories (PHLs) are at the core of the United States public health delivery system, linking almost every facet of public health infrastructure: disease control and prevention, maternal and child health, environmental health, epidemiology, and emergency preparedness and response. As a result, PHLs interact with a wide range of local/state/federal agencies and individuals, including local hospitals/laboratories/clinics, environmental/agricultural/wildlife institutions, academic institutions/health sciences centers, and law enforcement agencies.

The ERMIS evaluated in this study is a public health emergency response diagnostic consultation system that links local and state public health laboratories. This system, STATPACK™, is an interactive computerized system that utilizes the Internet infrastructure to provide secure microbiology diagnostic consultation to hospital laboratorians and send alert notifications to hospital laboratories in the case of a bioterrorism event or public health emergency. The STATPACK™ system is a secure, patient-privacy compliant, web-based network system that supports video telemedicine and connectivity among clinical health laboratories (Fruhling 2006, 2010).

Specifically, the STATPACK™ concept involves taking macroscopic (gross) as well as microscopic digital images of culture samples and sending them electronically for consultation with experts at state public health laboratories. STATPACK™ enables microbiology laboratories around the region to send pictures of suspicious organisms to the state public health laboratory, instead of the samples themselves, thus lessening the risk of spreading potentially deadly bioterror agents or infectious diseases. The system includes an alert system that is bi-directional and has various levels of priorities (emergency, urgent, and routine).

At the time of the evaluation of the success of STATPACK™, the system was present in 55 locations in clinical hospital laboratory locations across three states. STATPACK™ systems are also located in numerous food, water, environmental and veterinary science diagnostic testing laboratories. The system is regularly updated to address user needs, but no formal measurement of the success of STATPACK™ had been conducted to determine its ability to meet the needs of first responders during emergencies.

9.4.2 Survey Administration

We chose to use a 7-point Likert scale for the questions identified in Phase 3 of the survey creation process. We wanted users to have the ability to more finely discriminate in their responses as well as have a neutral response (i.e., a 4 on a 7-point scale). We also included an "Unable to Evaluate" option on the scale for each question since users do not routinely use the information system or may have concerns about expressing their opinion about one or more questions in the scale. The questionnaire also included open ended questions about the functionality and value of STATPACK™ to allow us to gain additional insights about the success of the system.

To administer the survey, the laboratory supervisor for each STATPACK™ system was contacted personally via email to explain the purpose of the survey and to determine how the supervisor wanted the survey distributed. Since laboratorians would be completing the survey at work, it was important to have their supervisor understand the value and need for the assessment as well as the supervisor's approval and endorsement. The supervisors were contacted and were provided a web link that both the supervisor and his/her staff could use to take the survey online. In the email, we also provided an option to receive a paper-based version of the survey if that was supervisor's preference.

We performed a pilot of the survey administration procedure with four clinical microbiology laboratories. After receiving positive feedback about our method of administration, we contacted the remaining laboratory supervisors, requesting them to please have their staff complete the survey within a 3 week period. As part of this email, each supervisor was asked to inform the research team about the number of laboratorians that use the STATPACK™ system in his or her laboratory. This allowed the team to better know the size of the user population of the software.

Three days before the survey deadline, each laboratory supervisor received a reminder email to please ask his/her staff to complete the survey. The survey was available after the deadline passed in an effort to include as many people as possible in the survey. As a thank you for helping with the evaluation, a token of appreciation, individual bags of candy, were mailed to the laboratory for each staff member that was requested to participate in the survey.

9.4.3 Survey Analysis

STATPACK™ has approximately 150 trained users. We received 64 responses from STATPACK™ users, yielding a response rate of 42 %. However, several respondents marked one or more items as "Not Applicable/Could Not Answer" due to their infrequent use of the system, which was expected given the culture of clinical laboratories. Laboratorians are used to dealing with results that are absolute as they identify pathogens and offer diagnoses. This group may hesitate to offer an opinion if they did not have a very high confidence level in their knowledge about the system.

Descriptive statistics were used to examine the success of STATPACK™ based on the quantitative survey responses. We examined the averages and standard deviations for both individual items as well as for each dimension of IS success (see Appendix B). To analyze the qualitative responses, similar responses were grouped together and counted. A report containing a list of all responses as well as the frequency of the issue was created and sent to the supervisors following the completion of the evaluation.

9.5 Results

Overall, each of these dimensions of success was rated highly on a 7-point scale with 1 being low and 7 being high. All averages were above the midpoint of 3.5. This suggests that overall, those that responded to the survey are generally happy with the system. No individual items received any scores below 4.0 on the 7-point scale. Table 9.2 is a summary of the average scores for STATPACK™ for each of the IS success variables.

The averages for each variable suggest that users are pleased with the STATPACK™ system, particularly with the quality of support from STATPACK™ IT staff and User Satisfaction. The survey also shows that users are very likely to use STATPACK™ should an emergency arise which further indicates the success of the system. The lowest rated dimension of IS success for STATPACK™ is Use. This result was expected because STATPACK™ is an ERMIS. Use of the system is primarily only in the case of an emergency, which means that STATPACK™ use is highly varied based on the circumstances and the types of situations that may arise. As discussed earlier, one of our challenges was to measure the Use for a system that is used generally only for emergency situations. Use is a complex measure (Burton-Jones and Straub 2006; Hsieh and Wang 2007); therefore, we adapted the survey to measure Use based on how often laboratorians used different functionalities of STATPACK™ as a measure of depth of use. We learned that storing microscopic images and reviewing images are the most often used functions of STATPACK™.

Our evaluation also included open-ended questions. The qualitative feedback provided additional insight on features liked by the users and areas needing improvement. In the open-ended responses, many survey respondents referred to the value of capturing images for education, training, and consultation. Others expressed they appreciated the ability to connect to state public health laboratories for consultation if need be. Several respondents commented that the system was very easy to use.

Our interpretation is that STATPACK™ has been a success on all dimensions, but like all systems, it has room for improvement. The most commonly stated request was the ability to print within the STATPACK™ software. Other requests included the ability to organize images, delete images, and email images to a person of choice, among others. This feedback had not been received through other means,

Table 9.2 Evaluation of STATPACK™

Construct	# Items	Mean	Std. Dev.
System quality	9	5.6	1.1
Information quality	6	5.8	1.1
Service quality	8	6.2	0.9
User satisfaction	2	6.0	1.0
Use (Type)	5	4.2	1.6
Intention to use	3	6.3	0.9
Net benefits	9	5.4	1.2

such as calls to the help desk, comments to trainers, or complaints via email. This information about potential new functionality or improvements to functionality would not have been realized except through this evaluation process.

9.6 Discussion

Given the widespread use of DeLone and McLean IS success model to evaluate different types of information systems (Petter et al. 2008), we expected that it would be useful in the ERMIS context. We have applied the model to the evaluation of ERMIS by illustrating that DeLone and McLean's IS success model can be adapted to emergency response systems that are used on a non-routine basis. Using this model and our analytical process, we were able to create a concise survey for seven variables that measure information systems success. Furthermore, we were able to document a process that can be used by others interested in the evaluation of ERMIS.

The results demonstrate the STATPACK™ information system was indeed beneficial to users and accomplishing the system goals as intended. Users identified that the system offers benefits beyond their own work processes. We learned that, overall, users could see the value of this system for their organization and patients. The evaluation of STATPACK™ success confirmed that the information systems is performing as intended, but also yielded helpful feedback on the aspects of the information system that are performing well and new suggestions for improvement.

Perhaps the more insightful component of the STATPACK™ success evaluation was the qualitative responses. For example, we learned of additional functionality to consider for future development, such as deleting messages and sorting images. We also learned about the need to educate the users about how this system is designed to comply with government regulations.

In creating this method to evaluate the success of ERMIS, we learned many lessons about designing an evaluation questionnaire, data gathering, and analyzing the data. For example, the rich insights we obtained about potential new features for the system were only provided because we included open-ended questions in the survey. We learned from an evaluation process perspective that qualitative data is essential, particularly to identify areas of strength and potential improvements.

Next, administering the survey online may have limited the number of responses that we received. It also may have limited our respondents to those that are most technically savvy. Those that are less comfortable with technology may not have been comfortable using an online survey. To mitigate this risk, we provided names, phone numbers, and email addresses of the researchers if the respondents had questions and two people contacted us during the survey period with questions. We also allowed laboratory supervisors to request paper copies of the survey. We did receive one survey via fax and four surveys via mail. The remaining responses were received online.

Emergency response systems may not be used very often which limits the ability to evaluate them. There are segments of the user population that may not have used the system in quite some time, limiting their ability to assess the system. We attempted to mitigate this risk by timing the survey after a recent training exercise, but not every possible participant completed the exercise because only one medical technologist was required to respond from each laboratory during an exercise.

We were conscientious about creating a survey that was comprehensive, but short enough to encourage as many users as possible to respond. We knew the medical technologist community would not tolerate a lengthy survey given the demands and time pressures within their environment. This may also be a challenge when evaluating the success of other ERMIS or medical information systems. However, we were able to develop a questionnaire that was concise, but addressed all aspects of the system required to properly evaluate the success of STATPACK™ system.

The evaluation approach was extremely cost effective. There was a nominal fee required for the online survey tool, mailing a few paper surveys, and mailing tokens of appreciation to all sites. However, for this very small fee, we gleaned very important information about STATPACK™ in terms of its current success and how to continue to improve it for the future.

9.7 Conclusion

The purpose of this paper is to demonstrate how to evaluate the success of emergency response medical information systems. This paper does so by documenting a method to develop questions and consider the process for administering the survey. We demonstrated the feasibility of the approach by evaluating the success of an ERMIS, STATPACK™. The results of the evaluation suggest that overall the STATPACK™ system has been a success and the evaluation process was useful.

For those interested in evaluating the success of an ERMIS currently in use in an organization, one can use and adapt the approach presented in this paper. To use this method of evaluating the success of an ERMIS, one would begin with the items listed in Appendix A as a starting point for Phase 1 of survey development (i.e., creating the question pool). If needed, additional items can be added based on the features of the ERMIS under study. Then, considering the features and context of the ERMIS, the list of question can be reduced to eliminate redundancy and to ensure the items are applicable (Phase 2—Narrowing the Question Pool). The next step would be to ask subject matter experts to review the items for clarity and completeness. Wording can be refined and any additional redundant items can be eliminated when completing Phase 3 of survey development (i.e., Reviewing the Question Pool).

Once the survey items are identified, the various considerations for administering the survey may be examined. Decisions should be made regarding the survey scale and whether or not to include open-ended questions. It is also important to identify how the survey will be shared and communicated with ERMIS managers and users. Finally, a pilot to verify the survey administration procedures and the final survey questions should be executed before the full scale success evaluation is performed.

When analyzing results, particularly with smaller sample sizes, descriptive statistics (i.e., mean and standard deviation) can examine which aspects of the ERMIS are successful and which areas may need improvement. By examining specific items within each success dimension (e.g., the specific items in System Quality), more detailed insights about the success of the ERMIS can be identified. If the sample size is sufficient, more sophisticated statistical analyses can also be performed if necessary. For example, one could examine if there are differences in perceptions across different groups of users of the ERMIS.

This method did not require extensive resources; with a relatively small degree of effort, we were able to evaluate the success and ERMIS and obtain useful feedback. The method offers a flexible approach to evaluate an ERMIS to ascertain if the system is a success across multiple dimensions or variables of information systems success.

Acknowledgments The authors would like to thank Sandra Vlasnik and Gregory Hoff for their assistance in data collection. We would also like to thank Guy Pare and others that shared their insights into our work at the SIG eHealth workshop at the Americas Conference on Information Systems. In addition, we want to acknowledge the Nebraska State Public Health, Anthony Sambol, Assistant Director, and Karen Stiles for the ongoing support. This research was funded through a Nebraska Foundation grant.

Appendix A: Measures Identified in Survey Development Process

This appendix identifies the measures that were identified and considered as part of the survey development process. Most items were identified from validated instruments in the literature, while other items were developed by the authors specific to the ERMIS being evaluated.

In each table, the item and source of the item is identified. The “X” denotes which phases the items were considered during the survey development process. When evaluating an ERMIS system using the method described in this paper, these items in the following tables could be considered in Phase 1 (Creating the Question Pool). Unnecessary items to evaluate a specific ERMIS can then be deleted in Phase 2 (Narrowing the Question Pool). Finally, the measures can be piloted and further refined as part of Phase 3 (Reviewing the Question Pool).

System Quality Measures

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS is user friendly	Doll et al. (2004)	X		
ERMIS is easy to use		X	X	X
ERMIS is easy to learn	Gable et al. (2008)		X	
It is easy to get *ERMIS* to do what I want it to do	Wixom and Todd (2005)	X		
ERMIS is easy to operate		X		
Learning to use *ERMIS* is easy for me	Davis (1989)	X		
It is easy for me to become skillful at using *ERMIS*		X		
I am knowledgeable on how to use *ERMIS*	Authors			X
ERMIS requires only the minimum number of fields and screens to achieve a task	Gable et al. (2008)	X	X	
ERMIS operates reliably	Wixom and Todd (2005)	X		
ERMIS performs reliably		X		
ERMIS is available when I need it	Authors			X
The operation of *ERMIS* is dependable	Wixom and Todd (2005)	X		
The downtime of *ERMIS* is minimal	Chang and King (2005)	X		
The *ERMIS* is always up-and-running as necessary	Gable et al. (2008)	X	X	
The *ERMIS* responds quickly enough to commands.		X	X	X
ERMIS allows information to be readily accessible to me	Wixom and Todd (2005)	X	X	X
ERMIS makes information very accessible		X		
ERMIS makes information easy to access		X		
It is often difficult to access information that is in *ERMIS*	Gable et al. (2008)		X	
ERMIS can be adapted to meet a variety of needs	Wixom and Todd (2005)	X		
ERMIS can flexibly adjust to new demands or conditions		X		

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS is versatile in addressing needs as they arise		X		
The *ERMIS* user interface can be easily adapted to one's personal approach	Gable et al. (2008)	X	X	
ERMIS can be easily modified, corrected or improved.		X		
ERMIS effectively integrates data from different areas of the company	Wixom and Todd (2005)	X		
ERMIS pulls together information that used to come from different places in the company		X		
ERMIS effectively combines data from different areas of the company		X		
ERMIS provides me appropriate information about hardware availability	Authors			X
All data within *ERMIS* is fully integrated and consistent	Gable et al. (2008)	X		
It takes too long for *ERMIS* to respond to my requests.	Wixom and Todd (2005)	X		
ERMIS provides information in a timely fashion		X		
ERMIS returns answers to my requests quickly		X		
ERMIS has all of the features that I need to do my job	Sedera and Gable (2004)	X	X	X
The system functionality of *ERMIS* is complete		X		
ERMIS is sufficiently sophisticated to meet my needs		X		
ERMIS meets the requirements of my organization	Gable et al. (2008)	X	X	
ERMIS includes necessary features and functions		X	X	
ERMIS always does what it should		X	X	X

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
In terms of system quality, I would rate *ERMIS* highly	Wixom and Todd (2005)	X	X	
Overall, *ERMIS* is of high quality		X		
Overall, I would give the quality of *ERMIS* a high rating		X	X	X
ERMIS meets your expectations	Chang and King (2005)	X		

Information Quality Measures

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS provides me with a complete set of information	Wixom and Todd (2005)	X	X	
ERMIS produces comprehensive information		X		
ERMIS provides me with all of the information that I need		X		
Information from *ERMIS* is unavailable elsewhere	Gable et al. (2008)	X	X	
The information provided by *ERMIS* is well formatted	Wixom and Todd (2005)	X		
The information provided by *ERMIS* is well laid out		X		
The information provided by *ERMIS* is clearly presented on the screen		X		
Information from *ERMIS* is in a form that is readily usable	Gable et al. (2008)	X	X	X
Information from *ERMIS* appears readable, clear and well formatted		X	X	
Information from *ERMIS* is concise			X	
ERMIS produces correct information	Wixom and Todd (2005)	X		
There are few errors in the information I obtain from *ERMIS*		X		
The information provided by *ERMIS* is accurate		X	X	

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS provides the precise information you need	Doll et al. (2004)	X		
ERMIS provides output that seems to be exactly what is needed	Gable et al. (2008)		X	
The resolution of the images captured and stored in *ERMIS* meets my standards	Authors			X
The images captured and stored in *ERMIS* accurately reflect the actual sample	Authors			X
It is easy to identify errors in information from *ERMIS*		X		
Though data from *ERMIS* may be accurate, outputs sometimes are not	Gable et al. (2008)	X		
ERMIS provides me with the most recent information	Wixom and Todd (2005)	X		
ERMIS produces the most current information		X		
The information from *ERMIS* is always up to date.		X		
Data from *ERMIS* is current enough	Gable et al. (2008)	X		
Information from *ERMIS* is always timely		X	X	
ERMIS provides reliable information	Ives et al. (1983)	X		
The information provided by *ERMIS* is consistent	Authors	X	X	
ERMIS produces information that I can depend on	Authors	X		
ERMIS has information that can be easily maintained	Chang and King (2005)	X		
The information in *ERMIS* can be easily changed		X		
I can easily update the information from *ERMIS*		X	X	
The information from *ERMIS* can be easily compared to past information		X		
The information from *ERMIS* can be easily integrated		X		
The information from *ERMIS* can be used for multiple purposes		X		

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS contains information that is important		X	X	X
The information from *ERMIS* is clear	Doll et al. (2004)	X		
ERMIS provides information that is interpretable		X		
I can understand the information from *ERMIS*	Chang and King (2005)	X		
Information from *ERMIS* is easy to understand	Gable et al. (2008)	X	X	X
The information from *ERMIS* is relevant	Ives et al. (1983)	X		
ERMIS provides information that meets my needs	Authors	X		
ERMIS produces information that is related to my task	Authors	X		
Information available from *ERMIS* is important	Gable et al. (2008)	X		
Overall, I would give the information from *ERMIS* high marks	Wixom and Todd (2005)	X		
Overall, I would give the information provided by *ERMIS* a high rating in terms of quality		X	X	
In general, *ERMIS* provides me with high-quality information		X	X	X

Service Quality Measures

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS support staff has up-to-date hardware and software	Jiang et al. (2002)	X		
ERMIS support staff physical facilities are visually appealing		X		
ERMIS support staff employees are well-dressed and neat in appearance		X	X	

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
The appearance of the physical facilities of the *ERMIS support staff* is in keeping with the kind of services provided		X		
When *ERMIS support staff* promises to do something by a certain time, it does so		X	X	
When users have a problem, *ERMIS support staff* shows a sincere interest in solving it		X	X	X
ERMIS support staff is dependable		X	X	X
ERMIS support staff provides its services at the times it promises to do so		X	X	
ERMIS support staff insists on error-free records		X	X	
ERMIS support staff tells users exactly when services will be performed		X	X	
ERMIS support staff employees give prompt service to users		X	X	X
ERMIS support staff employees are always willing to help others		X	X	
ERMIS support staff employees are never be too busy to respond to users' requests		X	X	
ERMIS support staff employees are available when I need them	Authors		X	X
The behavior of *ERMIS support staff* employees instills confidence in users	Jiang et al. (2002)	X	X	
Users feel safe in their transactions with *ERMIS support staff*		X	X	
ERMIS support staff employees are consistently courteous with users		X	X	X
ERMIS support staff employees have the knowledge to do their job well		X	X	X

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
ERMIS support staff gives users individual attention		X	X	
ERMIS support staff has operation hours convenient to all their users		X	X	
ERMIS support staff has employees who give users personal attention		X	X	
ERMIS support staff has the users' best interest at heart		X	X	
Support staff of *ERMIS* understand the specific needs of their users		X	X	X
Overall, I would rate *ERMIS support staff* highly in terms of their ability to provide quality service	Authors	X	X	X
In general, *ERMIS support staff* provides me with high-quality service	Authors	X	X	

User Satisfaction Measures

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
How adequate do you feel *ERMIS* meets the information processing needs of your area of responsibility?	Seddon and Yip (1992)	X		
How efficient do you feel *ERMIS* is?		X		
How effective do you feel *ERMIS* is?		X		
Overall, I am satisfied with *ERMIS*	Seddon and Yip (1992) and Briggs et al. (2008)	X	X	X
I like having *ERMIS* available	Briggs et al. (2008)		X	X

Use Measures[1]

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
How many times have you used *ERMIS* (for both training and for other purposes)?	Authors		X	X
How many times have you used *ERMIS* for purposes other than training?	Authors		X	X
When were you first introduced to *ERMIS*?	Authors		X	X
How often do you believe your organization uses *ERMIS* (weekly, monthly, quarterly, yearly, never)?	Authors		X	X
I often use *ERMIS* to capture microscopic images for consultation	Authors		X	X
I often use *ERMIS* to capture gross images for consultation	Authors		X	X
I often use *ERMIS* to store microscopic and gross images for consultation	Authors		X	X
I often use *ERMIS* to review images from the Image History for consultation	Authors		X	X
I often use *ERMIS* the Electronic Textbook capability	Authors		X	X
When *ERMIS* is used for consultation, I am the person that usually uses *ERMIS*	Authors			X

[1] These items were developed to relate to features that are specific to STATPack™. Authors wanting to use these measures to evaluate other information systems should adapt these measures based on the functionality of the information system.

Intention to Use Measures

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
I am likely to use *ERMIS* in an emergency	Davis (1989)			X
I intend to use *ERMIS* in the future				X
Should a situation arise, I plan to use *ERMIS*				X

Net Benefits Measures

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
I have learnt much through the presence of *ERMIS*	Gable et al. (2008)	X	X	
ERMIS enhances my awareness and recall of job related information		X		
ERMIS enhances my effectiveness in the job		X	X	X
ERMIS has resulted in overall productivity improvement for consultations		X	X	
Using *ERMIS* in my job enables me to accomplish tasks more quickly	Davis (1989)	X	X	
Using *ERMIS* improves my job performance		X		
Using *ERMIS* in my job increases my productivity	X			
Using *ERMIS* enhances my effectiveness on the job		X	X	
Using *ERMIS* makes it easier to do my job		X		
I find *ERMIS* useful for my job		X	X	X
Using *ERMIS* improves my decisions	Chang and King (2005)	X	X	X
Using *ERMIS* gives me confidence to accomplish my job		X	X	X
Using *ERMIS* increases my participation in decisions		X	X	

Item	Source	Creating question pool	Narrowing question pool	Reviewing question pool
Using *ERMIS* increases my awareness of job-related information		X		
Using *ERMIS* improves the quality of my work product		X		
Using *ERMIS* enhances my problem-solving ability		X	X	
ERMIS is useful	Authors	X		
Using *ERMIS* speeds up service delivery	Authors	X		
Using *ERMIS* streamlines work processes	Authors	X	X	X
Using *ERMIS* reduces cycle times	Authors	X	X	
Overall, I believe *ERMIS* is useful to me	Authors		X	
In general, *ERMIS* is a positive impact on my work	Authors		X	X
Using *ERMIS* helps me to address patient needs	Authors		X	
Using *ERMIS* helps me better work with physicians	Authors		X	
Using *ERMIS* improves our organization's care to patients	Authors		X	X
ERMIS has resulted in improved outcomes or outputs	Gable et al. (2008)		X	
ERMIS has resulted in an increased capacity to manage a growing volume of activity (e.g. population growth, epidemics, bioterrorism attack, etc.)			X	
ERMIS has resulted in overall quality improvement for consultations	Authors			X
Overall, *ERMIS* provides value to our organization	Authors		X	X
In general, *ERMIS* is a positive addition to our organization	Authors		X	

Appendix B: STATPACK™ Evaluation Results

The following table contains the measures used to evaluate STATPack™. In addition, the number of responses, the mean, and standard deviation of each item is also included.

System quality	N	Mean	Std Dev
The STATPack™ system is easy to use	62	5.69	0.97
I am knowledgeable on how to use the STATPack™ system	63	5.44	1.30
The STATPack™ system has all of the features that I need for remote public health microbiology laboratory consultations and interactions	62	5.39	1.28
The STATPack™ system provides me appropriate information about camera, microscope and network availability	57	5.35	1.25
The STATPack™ system is available when I need it	61	6.20	0.75
The STATPack™ software always does what I expect it to do	61	5.48	1.15
The STATPack™ software performs quickly enough to commands	62	5.68	1.10
The STATPack™ system makes information readily accessible to me	60	5.38	1.14
In terms of overall system quality, I would rate the STATPack™ system highly	62	5.95	1.00

Information quality	N	Mean	Std Dev
Information from STATPack™ is in a form that is readily usable	61	5.61	1.16
The information presented by STATPack™ is easy to understand	61	5.80	0.98
The resolution of the images captured and stored in STATPack™ meets my standards	62	5.60	1.36
Information available from STATPack™ is important	60	5.78	1.04
The images captured and stored in STATPack™ accurately reflect the actual sample	62	6.06	0.81
In general, STATPack™ provides me with high-quality information (text and images)	61	5.72	1.05

Service quality	N	Mean	Std Dev
The technical support staff for the STATPack™ system is consistently courteous with users	57	6.39	0.88
The technical support staff for the STATPack™ system has the knowledge to do their job well	53	6.32	0.92
When users have a problem, the technical support staff for the STATPack™ system shows a sincere interest in solving it	54	6.33	0.87

Service quality	N	Mean	Std Dev
The technical support staff for the STATPack™ system is dependable	58	6.33	0.89
The technical support staff for the STATPack™ system gives prompt service to users	55	6.20	0.89
The technical support staff for the STATPack™ system is available when I need them	54	6.13	0.89
The technical support staff for the STATPack™ system understands the specific needs of their users	57	5.95	1.17
Overall, I would rate the technical support staff for the STATPack™ system highly in terms of their ability to provide quality service	57	6.28	0.84

User satisfaction	N	Mean	Std Dev
Overall, I am satisfied with STATPack™	62	5.82	1.02
I like having the STATPack™ system available	63	6.17	0.89

Use	N	Mean	Std Dev
I often use the STATPack™ capability to capture microscopic images for consultation	56	4.14	1.60
I often use the STATPack™ capability to capture gross images for consultation	55	3.93	1.56
I often use the STATPack™ capability to store microscopic and gross images locally	56	4.70	1.80
I often use the STATPack™ capability to review images from the Image History.	60	4.45	1.48
I often use the STATPack™ Electronic Textbook capability	59	3.69	1.50
When the STATPack™ system is used for consultation, I am the person in the lab who usually does this	58	5.07	1.50
When were you first introduced to/trained on STATPack™ (i.e., how long have you been using STATPack™)?	59	2.88	1.76
How many times have you used the STATPack™ system (considering both the times when you were trained on the system and for other purposes)?	62	32.94	128.00
How many times have you used STATPack™ for purposes other than being trained?	62	27.42	127.39
How often do you believe your organization uses the STATPack™ system? (1 = weekly, 2 = monthly, 3 = quarterly, 4 = yearly, 5 = never)	62	2.74	0.83

Intention to use	N	Mean	Std Dev
I am likely to use the STATPack™ system in an emergency	63	6.21	0.85
I intend to use the STATPack™ system in the future	62	6.21	0.89
Should a situation arise, I plan to use the STATPack™ system	63	6.37	0.81

Net benefits	N	Mean	Std Dev
The STATPack™ system enhances my effectiveness as a medical technologist	56	5.36	1.05
I find the STATPack™ system useful for my job	61	5.31	1.15
Using the STATPack™ system improves my decisions	54	5.17	1.22
Using the STATPack™ system gives me confidence to accomplish my job	60	5.22	1.17
In general, STATPack™ is a positive impact on my work	60	5.40	1.20
The STATPack™ system streamlines consultation and other work processes	52	5.60	1.21
Using the STATPack™ system improves our organization's care to patients	56	5.21	1.25
The STATPack™ system has resulted in overall quality improvement for consultations	53	5.36	1.33
Overall, the STATPack™ system provides value to our organization	63	5.86	1.05

Demographic variables	N	Mean	Std Dev
I believe I am qualified to evaluate the STATPack™ system. (7 point Likert scale)	62	4.97	1.64
Age	61	3.82	1.09
Gender	60	1.77	0.43
Organization Type (1 = State Public Health Laboratory; 2 = Other)	61	1.90	0.30
Length of time in Organization (in years)	60	18.2	12.1
Have you encountered the Technical Support Staff for STATPack™? (1 = yes, 2 = no)	60	1.40	0.64
How many times have you encountered the technical support staff?	44	8.86	20.6
How would you rate your experience with STATPack™? (1 = novice; 2 = intermediate; 3 = advanced)	60	1.73	0.61

Open ended questions
Are there features or functionalities that the STATPack™ system is missing? If so, please explain
What feature do you find the most valuable? Why?
Does STATPack™ provide value to you, the organization, or public health? If so, how? If not, please explain your thoughts
What do you like the most about the STATPack™ system?
What do you like the least about the STATPack™ system?

References

Ammenwerth E, Mansmann U, Iller C, Eichstadter R (2003) Factors affecting and affected by user acceptance of computer-based nursing documentation: results of a two-year study. J Am Med Inform Assoc 10(1):69–84

Briggs RO, Reinig BA, deVreede G-J (2008) The yield shift theory of satisfaction and its application to the IS/IT domain. J Assoc Inf Syst 9(5):267–293

Burton-Jones A, Straub DW (2006) Reconceptualizing system usage: an approach and empirical test. MIS Quarterly 17(3):228–246

Chang C-J, King J (2005) Measuring the performance of information systems: a functional scorecard. J Manag Inf Syst 22(1):85–115

Davis FD (1989) Perceived usefulness, perceived ease of Use, and user acceptance of information technology. MIS Quarterly 13(3):319–340

DeLone WH, McLean ER (1992) Information systems success: the quest for the dependent variable. Inf Syst Res 3(1):60–95

DeLone WH, McLean ER (2003) The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst 19(4):9–30

Doll WJ, Deng XD, Raghunathan TS, Torkzadeh G, Xia WD (2004) The meaning and measurement of user satisfaction: a multigroup invariance analysis of the end-user computing satisfaction instrument. J Manag Inf Syst 21(1):227–262

Fruhling A (2006) Examining the critical requirements, design approaches and evaluation methods for a public health emergency response system. Commun Assoc Inf Syst 18:431–450

Fruhling A (2010) STATPack—an emergency response system for microbiology laboratory diagnostics and consultation. Adv Manag Inf Syst 16:123–149

Gable GG, Chan T, Sedera D (2008) Re-conceptualizing information system success: the IS-impact measurement model. J Assoc Inf Syst 9(7):377–408

Hsieh JJPA, Wang W (2007) Explaining employees' extended Use of complex information systems. Eur J Inf Syst 16(3):216–217

Ives B, Olson M, Baroudi JJ (1983) The measurement of user information satisfaction. Commun ACM 26(10):785–793

Jiang JJ, Klein G, Carr CL (2002) Measuring information system service quality: SERVQUAL from the other side. MIS Quarterly 26(2):145–166

Jick TD (1979) Mixing qualitative and quantitative methods: triangulation in action. Adm Sci Q 24(4):602–611

Pare G, Lepanto L, Aubry D, Sicotte C (2005) Toward a multidimensional assessment of picture archiving and communication system success. Int J Technol Assess Health Care 21(4):1–9

Petter S, Fruhling A (2011) Evaluating the success of an emergency response medical information system. Int J Med Inform 80(7):480–489

Petter S, DeLone WH, McLean ER (2008) Measuring information systems success: models, dimensions, measures, and interrelationships. Eur J Inf Syst 17(3):236–263

Seddon P, Yip S-K (1992) An empirical evaluation of user information satisfaction (UIS) measures for use with general ledger accounting software. J Inf Syst 6(1):75–98

Sedera W, Gable GG (2004) A factor and structural equation analysis of the enterprise systems success measurement model. Twenty-Fifth International Conference on Information Systems, 2004

Sicotte C, Pare G, Kra Bini K, Moreault M-P, Laverdure G (2009) Virtual organization of hospital medical imaging: a user satisfaction survey. J Digit Imaging 23(6):689–700

Straub DW (1989) Validating instruments in MIS research. MIS Quarterly 13(2):146–169

Turoff M (2002) Past and future emergency response information systems. Commun Assoc Comput Mach 45(4):29–33

Turoff M, Van de Walle B (2004) Preface to the special issue on emergency preparedness and response information systems. J Inf Technol Theory Appl 6(3):3–5 van der Meijden MJ, Tange HJ, Troost J, Hasman A (2003) Determinants of success of inpatient clinical information systems: a literature review. J Am Med Inform Assoc 10(3):235–243

Wilson EV, Lankton NK (2004) Modeling patients' acceptance of provider-delivered E-health. J Am Med Inform Assoc 11(4):241–248

Wixom BH, Todd PA (2005) A theoretical integration of user satisfaction and technology acceptance. Inf Syst Res 16(1):85–102

Chapter 10
Effective Use of Clinical Decision Support in Critical Care: Using Risk Assessment Framework for Evaluation of a Computerized Weaning Protocol

Sahiti Myneni[1], Debra McGinnis[2], Khalid Almoosa[3], Trevor Cohen[1], Bela Patel[3], and Vimla L. Patel[4]

[1]Center for Cognitive Informatics and Decision Making, School of Biomedical Informatics, UT Health, 7000 Fannin St, Suite 600, Houston, TX 77030, USA

[2]Memorial Hermann Hospital, Houston, TX, USA

[3]Divisions of Critical Care Medicine, UT Health, Houston, TX, USA

[4]Center for Cognitive Studies in Medicine and Public Health, New York Academy of Medicine, New York, NY, USA

Abstract **Background**: Clinical decision support aids such as computerized weaning protocols (CWPs) aim to reduce medical errors and improve patient safety. However, the dynamic nature of critical care environments demands context-specific and complexity -inclusive assessment of these support tools for optimal results.

Objective: To apply and validate the use of a risk assessment method called Functional Resonance Accident Method (FRAM), which is originally proposed for adverse event analysis in the aviation industry, to evaluate effective use of a CWP in a medical intensive care unit.

Study Design and Methods: Multiple data collection methods including (1) ethnographic observations, (2) semi-structured interviews, and (3) review of hospital documents related to workflow, procedures, and training were used to simulate a FRAM based model of the CWP and identify factors affecting its use. Subsequently, we validated our findings by shadowing clinicians during 65 weaning attempts (120 h of in vivo data).

Results: The factors posing risk to effective use of CWP included misinterpretation of CWP's sedation assessment scale, communication and collaboration breakdowns, problems with on-time support delivery, and negative perception of

S. Myneni (✉)
Center for Cognitive Informatics and Decision Making, School of Biomedical Informatics, UT Health, 7000 Fannin St, Suite 600, Houston, TX 77030, USA
e-mail: Sahiti.Myneni@uth.tmc.edu

© Springer International Publishing Switzerland 2016
A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*, Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_10

the protocol among clinicians. During the in-situ validation, we found that 45 of the 65 attempts were favorable, 16 fell under near-miss category, while the remaining four were unfavorable.

Conclusions: Non-linear risk assessment method based on resilience engineering concepts is an effective approach for identification of factors for safe use of decision support aids in the real- world health care environment.

Keywords Complexity • Medical error • Risk assessment • Computerized weaning protocol • Clinical decision support

10.1 Introduction

Clinical Decision Support (CDS) provides clinicians with knowledge and patient-specific information, intelligently filtered at appropriate times (Osheroff et al. 2005). It encompasses a variety of interventions such as computerized alerts, reminders, clinical guidelines, order sets, patient data reports, documentation templates, diagnostic support, and clinical workflow tools (Osheroff et al. 2007). CDS has been shown to be a successful strategy to reduce medical errors and improve patient safety (Kawamoto et al. 2005). However, outcomes of several evidence-based CDS interventions are often suboptimal because compliance with these guidelines in clinical settings is poor (Cabana et al. 1999). Therefore, systematic evaluation of CDS workflow, implementation, user satisfaction, and other systemic factors are recommended to exploit CDS advantages to their full extent (Kaplan 2001; Friedman 2005). It has been suggested that such assessments should take the context and complexity of CDS environment into account for high yield in quality improvement (Sinuff et al. 2007; Weiss and Amaral 2010), and that failure to assess the environment prior to implementation of an intervention can have harmful unintended consequences (Patel and Cohen 2008; Ash et al. 2007). In this paper, we provide the findings from one such context-aware complexity-inclusive assessment of a CDS, a computerized weaning protocol (CWP).

10.2 Background

While mechanical ventilation (MV) is a lifesaving procedure, prolonged ventilation can expose patients to unnecessary risks including increased mortality, ventilator-associated pneumonia, and airway trauma (Burns 1999). On the other hand, premature discontinuation of MV can result in unsuccessful extubation, requiring re-intubation (MacIntyre et al. 2001). In order to avoid the aforementioned risks, critical care units have adopted a variety of weaning protocols (WPs). Comparison and description of these strategies can be found in (Brochard et al. 1994; Esteban et al. 1995; Esen et al. 1992; Lellouche et al. 2004; Rose et al. 2008; Burns et al. 2009).

It is reported that protocols employing every day weaning can result in a reduction of ventilator days, lower ICU costs, and fewer related complications (Ely et al. 2001). Randolph et al. identified implementation issues including software errors, underlying logic errors, and user misunderstandings of the protocol instructions (Randolph et al. 1998). Ely and colleagues found that healthcare professionals perceived protocols as removing clinical judgment without considering all facets of the patients involved (Ely et al. 2001). Physician unfamiliarity with the protocol, respiratory therapists' inconsistent performance in seeking extubation orders, and resource constraints were shown as the reasons behind protocol incompliance (Ely et al. 1999). The model for accelerating improvement was used to improve implementation outcomes of CWP (McLean et al. 2006). Educational interventions to foster an understanding of the advantages and mechanisms of protocol use, and inculcation of collaborative practices were identified as high-priority aspects for successful protocol implementation (Ely et al. 1999; McLean et al. 2006; Vitacca et al. 2000). However, there is limited work documenting the application of systems engineering principles and complexity theory for evaluation of CWPs in critical care.

Implementation of a CWP or any other CDS in a critical care unit involves more than just a procedural change. Given the socio-technical nature of critical care environment, other factors are likely to play significant role in effective use of the protocols. Detailed understanding of complex dynamics and interdependencies commonly observed in such socio-technical systems is essential in order to mitigate adverse events and unintended consequences of any new intervention including CDS (Amalberti 2001; Hollnagel 2008; Myneni et al. 2010).

The objective of our study is to identify risks inherent in the use of the CWP, which could result in sub-optimal performance in a Medical Intensive Care Unit (MICU). To account for the complexity of MICU environment, we chose a dynamic risk assessment methodology called Functional Resonance Accident Method (FRAM) motivated by complex systems research. FRAM is a systemic method originally developed for the analysis and prediction of adverse events in the aviation industry, and has been successfully applied to adverse event and risk analysis in other complex domains such as manufacturing plants and financial markets (Hollnagel et al. 2008; Sundström et al. 2008).

10.3 Theoretical Rationale

FRAM provides a way to describe how multiple functions and conditions can combine to produce an adverse outcome. It illustrates the dynamic interactions within the socio-technical system and lets us understand the how and why of a particular event chain (Herrera and Woltjer 2010). FRAM is based on the following four major principles (Hollnagel et al. 2008):

1. **The principle of equivalence of successes and failures**: FRAM adheres to the resilience engineering view that failures represent the flip side of the adaptations necessary to cope with real-world complexity (Hollnagel et al. 2005). Success depends on the ability of teams and individuals to anticipate risks and critical situations, to recognize them in time, and to take appropriate action.
   - Input (I): that which the function transforms or that which starts the function
   - Output (O): that which is the result of the function, either an entity or a state change
   - Preconditions (P): conditions that must exist before a function can be executed
   - Resources (R): that which the function needs or consumes to produce the output
   - Time (T): temporal constraints affecting the function (with regard to starting time, finishing time, or duration)
   - Control (C): how the function is monitored or controlled
2. **The principle of approximate adjustments**: Since the conditions of work never completely match what has been specified, individuals must adjust their performance so that they can succeed under the existing conditions.
3. **The principle of emergence**: The variability of normal performance is rarely large enough to be the cause of an ineffective activity in itself or even to constitute a risk. But the variability from multiple functions may combine in unexpected ways, leading to consequences that are disproportionally large producing a non-linear effect.
4. **The principle of functional resonance**. The variability of a number of functions may resonate, i.e., reinforce each other and thereby cause the variability of one function to exceed normal limits. The consequences may spread through tight couplings rather than via identifiable and enumerable cause-effect links.

The steps to apply FRAM for evaluation of the effective use of CDS ((in this context, a CWP) are given below:

1. **Identify and characterize essential CDS functions**: All functions required to complete a CDS activity are specified in this step. Each function is separately identified, but not arranged in any way. A function may, for instance, be to update the medication list of a patient. Each function is modeled using six parameters: Input, Output, Time, Resource, Precondition, and Control (see Fig. 10.1).
2. **Describe potential variability of functions**: The variability of each of the parameters in each function is characterized using a checklist, which was based on common performance conditions (CPCs) listed in (1) Hollnagel's cognitive reliability and error analysis method (CREAM) (Hollnagel et al. 2008), and (2) Ten Commandments for effective use of CDS (Bates et al. 2003). A list of CPCs from both the above sources was presented to an expert physician, who chose the final list with 12 CPCs (Table 10.1) that best define the working conditions in MICU.
3. **Identify functional resonance and potential variability**: The functions identified in Step 1 may be coupled via their parameters. For example, the output

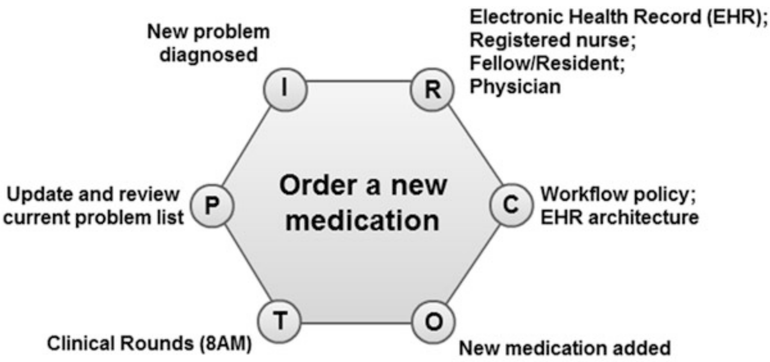


Fig. 10.1 A FRAM module describing a function. (I-Input, O-Output, T-Time, R-Resource, C-Constraint, P-Precondition)

Table 10.1 Variability checklist for context-dependent evaluation of clinical decision support

Conditions for effective clinical decision support	Rating scale
On-time support delivery	
Fit into user's workflow	
Usability	
Positive perception of clinicians	
Collaboration quality	Adequate
Communication quality	
Training and experience	Inadequate
Monitoring impact and feedback	
Time needed/available	Unpredictable
Knowledge management and update	
Quality and support of organization	
Operational support	

of one function may be an input to another function, or provide a resource, fulfill a pre-condition, or enforce a control. Couplings between functions can be identified by analyzing commonly related parameters. These couplings may then be combined with the results of Step 2, the characterization of variability, to specify how the variability of one function may have an impact on the variability of another. To understand which CPCs influence a given function, the functions need to be classified into one of the following three categories: (1) Human, (2) Technology, and (3) Organization. Hollnagel, 2004 provides a detailed overview of the functional classification (Hollnagel 2004). Functional dependencies can spread variability across the activity and result in an adverse or unfavorable event.

4. **Propose variability monitoring interventions**: Once the causes for, and nature of variability in CDS are understood, interventions are developed to contain the variability and improve CDS performance.

10.4 Methods

The Respiratory Therapist (RT) led CWP under study has been implemented in MICU in 2008. The protocol itself acts as a CDS tool which provides clinicians with patient-specific, relevant information as needed for making informed decisions regarding extubation (Mack et al. 2009). Patient-related data collected by RTs (night shift and day shift) as part of the protocol's requirements were recorded in the Electronic Health Record (EHR). All mechanically ventilated patients were screened daily for a Spontaneous Breathing Trial (SBT) by the night shift RT starting at 4 AM every day. The elements of the screen included physiological parameters such as hemodynamic stability, respiratory rate, positive end-expiratory pressure, and fractional concentration of inspired oxygen. Once the results of the screening assessment were entered into EHR CWP module, the system automatically produced the results (Pass/Fail). For patients who passed the above screen, the day shift RT conducted sedation assessment and SBT. The CWP guided the RT through every step (cuff leak checks, ventilator mode selection) and the SBT results were manually entered into EHR. All data related to weaning mechanics (e.g. tidal volume, rapid shallow breathing index) were collected using CWP module and the aggregated data were presented to the physician for the final decision.

The study protocol was approved by our Institutional Review Board. In the first phase of the study, Phase I, we adopted our clinical version of FRAM to identify the risk factors creating barriers to effective use of the CWP in the MICU. A FRAM-based normative model of the CWP was created using multiple data collection methods: (1) Discussions with expert physicians, RTs, Registered Nurses (RNs), (2) 15 h of ethnographic observations guided by the RT manager, (3) semi-structured interviews with eight MICU RTs and four Attending Physicians (APs), and (4) a review of hospital documents related to workflow, procedures, and training manuals. Once the potential risk factors were identified using FRAM, we validated our findings by shadowing four MICU RTs for 65 weaning sessions in Phase II. The objective of Phase II was to understand the criticality and urgency to address the issues identified in Phase I. In Phase II, a trained researcher unobtrusively observed RTs and APs as they conducted weaning sessions using the CWP (Van Maanen 1996). A total of 120 h of observational data were collected and detailed notes were taken. The purpose of in-vivo observations was to understand the real-time operating conditions with respect to CWP in MICU and gain insight into protocol related work flow, information sharing, and interaction strategies among MICU clinicians. By the end of Phase II, we developed possible intervention strategies to mitigate the effects of the risk factors on effective and efficient performance of the CWP in MICU. In summary, multiple data collection and analysis methods were used to analyze CWP use in a critical care unit.

10.5 Results

10.5.1 Phase I: Identification of Risk Factors Affecting Performance of CWP

A list of possible factors that could affect the performance and effective use of the CWP in the MICU was populated using FRAM methodology.

Step 1: The first step in FRAM was to identify essential functions of an activity. Five essential steps in the CWP were identified using multiple methods (observations, review of hospital manuals, semi-structured interviews) as follows: (1) patient inclusion, (2) SBT screening assessment, (3) sedation assessment using RASS score (see Table 10.2), (4) SBT, and (5) decision making: extubation. As shown in Table 10.3, each of these functions was modeled using six parameters -input (I), output (O), resource (R), time (T), pre-condition (P), and control (C).

Step 2: The variability of the CWP parameters was characterized (see Table 10.4). The ratings were given based on input from domain experts- physician and RT. Cohen's Kappa measure was used to determine inter-rater reliability. The raters had an outstanding reliability of 0.855 ($p < 0.001$) with only one disagreement in "positive perception of clinicians" category. The disagreement has been resolved by asking two additional raters to assign a rating for that particular condition, and the final rating was the one that has most agreement. It was clear from Table 10.4 that there were many factors that were inadequate and unpredictable, which might ultimately affect the effective use of the CWP.

Table 10.2 Richmond Agitation Sedation Scale (RASS) used for sedation assessment as part of weaning protocol in Medical Intensive Care Unit

RASS score	Description
+4	Combative, violent, danger to staff
+3	Pulls or removes tube(s) or catheters; aggressive
+2	Frequent nonpurposeful movement, fights ventilator
+1	Anxious, apprehensive, but not aggressive
0	Alert and calm
−1	Awakens to voice(eye opening/contact) >10 s
−2	Light sedation, briefly awakens to voice(eye opening/contact) <10 s
−3	Moderate sedation, movement or eye opening. No eye contact
−4	Deep sedation, no response to voice, but movement or eye opening to physical stimulation
−5	Unarousable, no response to voice or physical stimulation

Table 10.3 FRAM based characterization of essential functions in a computerized weaning protocol

Function	Input	Output	Resource	Time	Control	Pre-condition
Function 1: *Patient inclusion*	Ventilator settings; Protocol order	SBT screening assessment Order;	EHR; Day RT; Fellow		MICU weaning policy	Patient exclusion criteria
Function 2: *SBT screening assessment*	SBT screening assessment order	Eligible/ineligible for SBT; Respiratory mechanics	EHR; Night RT	4 AM	MICU weaning policy	Patient exclusion criteria; Protocol order
Function 3: *Sedation assessment*(*RASS Score*)		Arousability score	EHR; Day RT; RN	7:30 AM	MICU weaning policy; RASS	Sedation Holiday; Eligible for SBT
Function 4: *Spontaneous Breathing Trial* (*SBT*)	Ventilator settings	Pass/Fail SBT	Day RT; EHR		MICU weaning policy	Eligible/Ineligible for SBT; Arousability score > -2
Function 5: *Decision Making*: *Extubation*	Ventilator settings	Patient extubation	RN; Day RT; Physician; Fellow; Residents; EHR	8 AM	Clinical Objectives; MICU Weaning policy	Pass/Fail SBT; Arousability > -2

Table 10.4 FRAM based variability checklist for weaning protocol

Conditions for effective clinical decision support	Rating
On-time delivery of decision support	***Unpredictable***
Fit into user's workflow	Adequate
Usability and understanding	*Inadequate*
Positive perception of clinicians	*Inadequate*
Collaboration quality	***Unpredictable***
Communication quality	***Unpredictable***
Training and experience	*Inadequate*
Monitoring impact and feedback	*Inadequate*
Time needed/available	***Unpredictable***
Knowledge management and update	*Inadequate*
Quality and support of organization	Adequate
Operational support	Adequate

Step 3: Identify functional resonance and potential variability

The local and global interactivity of the CWP components were considered during this analysis. It is interesting to note the degree of dependencies among functions in terms of the six defining parameters. We identified and analyzed possible ways in which these variability sources might resonate and affect the performance of the protocol. For the purpose of illustration, we will focus on how FRAM enabled the identification and understanding of one variability source, the "Sedation assessment (RASS Score)", i.e. function 3.

Based on the results from Step 2, the use of RASS for assessment of sedation level of a patient was the control for function 3 (see Fig. 10.2). Appropriate use of RASS is essential if the output of function 3 is to be correct. In turn, the output of function 3 is a precondition for functions 4 and 5. Therefore, variability in the RASS usage could potentially affect the implementation of functions 4 and 5.

At this point, we used Table 10.4 in which we characterized variability checklist using FRAM Step 2. To understand which CPC affects a particular function, the functions were classified as Human/Technology/Organization. Function 2 mainly depends on the person carrying out the task and therefore it is a human function. Based on the nature of function 2 and expert input, the variability in the CPCs was traced to (1) usability and understanding, (2) training and experience. Both the CPCs were rated inadequate and if an RT assigns a wrong score to a patient because of misinterpretation of the RASS, the CWP would skip function 4, which is essential for timely extubation. This variability in function 2 could be curbed at function 5, when the RT discusses the RASS score assignment with the rest of the clinical team. Given that checklist in Table 10.4, communication and collaboration quality among clinicians was quite unpredictable, thus making human resources for all functions variable. For example, clinicians (such as RTs) might be attending the case of another critically ill patient, or the unit may be crowded. Therefore, the misinterpreted sedation scale might not be corrected because of reinforced variability

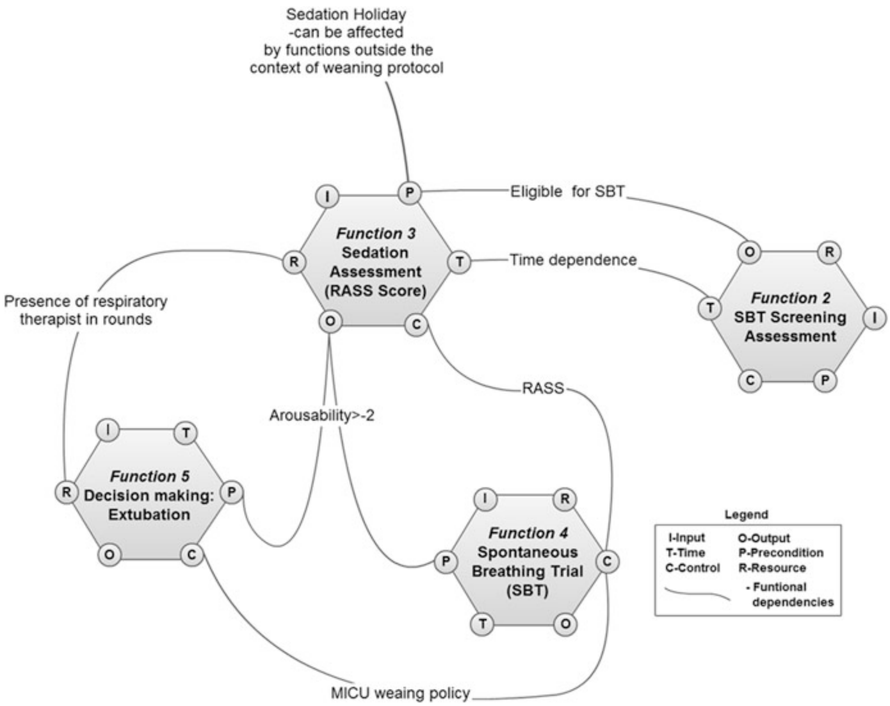


Fig. 10.2 Functional dependencies of Function 2- Sedation Assessment (RASS Score)

induced by a situation in which a RT did not communicate with the clinical team. This particular case of functional resonance can lead to an unfavorable event- delay in extubation. From this example, it was evident that inadequate understanding of protocol mechanisms (such as RASS) and unpredictable communication among clinicians were the two major problems that might pose risks to the effective use of CWP.

Similarly, using FRAM as explained above, the variability sources in the protocol posing risk were identified to be: (a) misinterpretation of the sedation scale, (b) lack of RTs presence in the daily rounds, (c) communication breakdown among clinicians, (d) problems of on-time support delivery, (e) clinicians' negative perception of the protocol. Variability in one or more of these sources might reinforce one another and affect the performance of the CWP. Having identified major factors that might limit the effectiveness of the CWP, we validated our results from Phase I by collecting in-vivo data on the CWP operations in MICU. Phase II data also allowed us to understand the urgency and criticality of each of the risk factors. Step 4 of the FRAM analysis is to propose intervention strategies to limit variability and its spread among functions. Details about the proposed interventions for effective CWP use is presented in the following sections of the paper (Sect. 10.5.4).

10.5.2 *Phase II: In Vivo Findings of Variability Sources Posing Risk to Effective Use of CWP in MICU*

The following results were based on data collected by shadowing MICU clinicians for 65 weaning sessions.

- *Misinterpretation of sedation scale*: In seven of these sessions, the sedation scale was misinterpreted, which in one case resulted in a patient being erroneously extubated based on the incorrect sedation score, and subsequently re-intubated. Aside from this being a waste of human and material resources, the patient was unnecessarily placed at risk of harm from a period of inadequate respiratory support, and the potential complications of re-intubation. The sedation scores were incorrectly assessed because the RTs misinterpreted the word “sedation” in the RASS scale as referring to the prescription sedative, instead of an assessment of the physical state of the patient. As patients on ventilation were usually sedated, the presence of a prescribed sedative does not help to discern whether or not a patient would tolerate extubation.
- *Lack of cognitive support and/or knowledge issues with the protocol mechanisms*: There were five instances during which the SBT was prolonged for more than 150 min, where the protocol-based time limit was 30–120 min. The RTs placed the patients on minimal ventilator support and did not return the ventilator settings back to their original value. In addition to placing a patient at higher levels of discomfort and anxiety, this was potentially life-threatening for patients with airway management issues.
- *Problems with on-time data delivery*: During two sessions, the AP had to make a decision without aid from the CWP because the RTs were not able to complete SBTs on time to present the protocol data during rounds. The reason behind this delay was attributed to ICU crowding and resource allocation to another critically ill patient.
- *Lack of adherence to CWP*: During one instance, although a patient failed the night RT assessment, the physician ordered an extubation. The reason was that the patient was on the ventilator “for a long time”. The patient was re-intubated. There were another two instances where the patient was extubated, despite not meeting the protocol pass criteria. The protocol was not followed by physicians for one of the two reasons below:

    1. Some physicians did not trust the protocol, although the protocol is evidence-based, and henceforth, disregarded RT’s data.
    2. Physicians disregarded protocol data for “bottom-line patients” (patients who were slightly below the passing criteria were considered bottom-line).

    However, in these instances RTs were not told the reasons as to why the protocol was not followed by the physicians. During follow-up interviews, a few RTs attributed this as the reason for them being insufficiently rigorous in their implementation of the protocol on account of the perception that their data were not acknowledged during final decision point. It is important to note that some such communication lapses between APs and RTs stemmed due to lack of RTs presence and participation during daily morning clinical rounds. In addition, although the protocol was RT-driven, it was observed that the there was lack of clarity in defining the roles of RTs, RNs, and APs. The RTs assumed that their role in the protocol was to collect ventilator data and that the final decision was made by the AP. However, the protocol was intended to support interdisciplinary collaborative decision-making (Meade and Ely 2002).

10.5.3 Evaluation of CWP's Effectiveness

We coded the data using the following outcome definitions for the CWP. The performance of the CWP was defined in terms of three categories- "favorable outcome", "unfavorable outcome", and "near-miss".

Favorable outcome criteria:

- If a mechanically-ventilated (MV) patient passes night-RT assessment AND sedation assessment AND spontaneous breathing trial, and subsequently he/she is extubated.
- If a MV patient fails night RT assessment OR sedation assessment OR spontaneous breathing trial), and subsequently he/she is not extubated.
- If a MV patient passes night RT assessment AND sedation assessment AND spontaneous breathing trial, then he/she is not extubated because of airway management issues or other clinical objectives.

Unfavorable outcome criteria (*Caused by functional coupling among components*)

- If a MV patient passes night RT assessment AND sedation assessment AND spontaneous breathing trial, but he/she is not extubated.
- If a MV patient fails night RT assessment OR sedation assessment OR spontaneous breathing trial, but is extubated and is again re-intubated
- A physician makes a decision on extubation with no or erroneous data from the protocol

Near- Miss criteria (*No functional coupling among components*)

- Uncoupled Errors during execution of protocol which did not result in an unfavorable outcome. But, may cause an adverse event in some other context.

Data from 65 weaning sessions were coded based on the above definitions. 45 (69 %) of the 65 sessions were favorable, 16 (25 %) fell under near-miss category, while the remaining four (6 %) were unfavorable (Myneni et al. in press).

In summary, the problems identified with the CWP were misinterpretation of sedation scores, issues with on-time delivery support, inadequate communication and collaboration among clinicians, and insufficient feedback of protocol's impact on quality of care delivery in MICU. Our findings from Phase II were in agreement with the results from Phase I. The FRAM based analysis to identify risk factors acting as barriers to effective use of the CWP identified in Phase I positively predicted 81 % of the variability sources that resonated to cause near-misses and unfavorable outcomes observed during Phase II.

10.5.4 Intervention Strategies

We believe that the majority of problems identified with the CWP can be addressed through clinician education, improved communication, and impact demonstration. Multi-disciplinary collaborative input from clinicians involved in the daily use of the CWP needs to be considered in view to identify and address confusing aspects in the protocol's procedures.

1. *Training and Education*: Design a training module for clinicians to fill existing knowledge gaps and conceptual misunderstandings.
2. *Feedback and Impact Monitoring System*: A platform to solicit frequent feedback from clinicians regarding the operation of the CWP can help improve its effectiveness. In addition, the system should be able to disseminate relevant quality data reflecting the impact of CWP among clinician community. Such a system may enable better information sharing among clinicians and leadership. It may also serve as a resource to demonstrate the efficacy of the CWP in MICU setting, thus bolstering clinicians' trust and thereby improving their compliance with it.

10.6 Conclusions and Discussion

Adoption of a non-linear risk assessment methodology based on resilience engineering concepts is a valuable approach to address dynamic, non-deterministic nature of critical care environment. Functional Resonance Accident Method when complemented with common performance conditions representing critical care context can help us determine variability risk sources leading to sub-optimal use of CDS. Our study provides insight into factors affecting the effective and safe use of a CWP. These factors are socio-technical in nature: inherent to the protocol and externalized in the environment, in addition to trust and understanding. Some of the potential risks, such as clinicians' negative perception, protocol misinterpretation, and inadequate collaborative practices identified using FRAM are consistent with the results from previous research (Burns et al. 2009; Ely et al. 1999, 2001; Randolph et al. 1998). However, other risks including untimely delivery of support and communication breakdowns are not reported elsewhere.

In-vivo analysis revealed that near-misses and adverse events constituted almost 34 % of protocol outcomes. While CDS such as CWP are essential for improvement of patient safety and to reduce medical errors, these “safety nets” require continuous assessment and refinement in order for them to reach optimal working conditions in a complex environment like critical care. Ways to improve the performance of such CDS are context-specific and can range from education and motivation to workflow re-engineering (Weiss and Amaral 2010). Current method based on resilience engineering concepts shows strong potential for assessment of critical care safety interventions. It should be pointed out that this method is still evolving and needs to be formalized for use in health care settings. Given the extensive analysis required by FRAM, the method is time and labor intensive. In addition, it is important to note that not all variability of a system is risky in nature. Deviations from normal working conditions might sometimes be an act of resilience and positive adaption to an unanticipated or emerging event (Hollnagel et al. 2005). It is essential to understand the aspects of cognitive risk management employed by clinicians during error detection and recovery (Morel et al. 2008). This understanding can inform the design of better support systems in critical care.

Acknowledgements This research was supported by an award from the James S McDonnell Foundation (Grant 220020152) to Vimla L. Patel. We thank all of the clinicians for their time and involvement in the study. *Disclosures*: The authors have nothing to disclose and have no Conflict of Interest.

References

Amalberti R (2001) The paradoxes of almost totally safe transportation systems. Saf Sci 37:109–126

Ash JS, Sittig DF, Poon EG, Guappone K, Campbell E, Dykstra RH (2007) The extent and importance of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 14(4):415–423

Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L et al (2003) Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inf Assoc 10:523–530

Brochard L, Rauss A, Benito S et al (1994) Comparison of three methods of gradual withdrawal from ventilatory support during weaning from mechanical ventilation. Am J Respir Crit Care Med 150:896–903

Burns SM (1999) Making weaning easier. Crit Care Nurs Clin North Am 11:465–479

Burns KE, Meade MO, Lessard MR, Keenan SP, Lellouche F (2009) Wean earlier and automatically with New technology (the WEAN study): a protocol of a multicentre, pilot randomized controlled trial. Trials 10:81

Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA et al (1999) Why don’t physicians follow clinical practice guidelines? A framework for improvement. JAMA 282:1458–1465

Ely EW, Bennett PA, Bowton DL, Murphy SM, Florance AM, Haponik EF (1999) Large scale implementation of a respiratory therapist driven protocol for ventilator weaning. Am J Respir Crit Care Med 159(2):439–446

Ely EW, Meade MO, Haponik EF, Kollef MH, Cook DJ, Guyatt GH et al (2001) Mechanical ventilator weaning protocols driven by non-physician healthcare professionals: evidence-based clinical practice guidelines. Chest 20:454S–463S

Esen F, Denkel T, Telci L et al (1992) Comparison of pressure support ventilation and intermittent mandatory ventilation during weaning in patients with acute respiratory failure. Adv Exp Med Biol 317:371–376

Esteban A, Frutos F, Tobin MJ et al (1995) A comparison of four methods of weaning patients from mechanical ventilation. N Engl J Med 332:345–350

Friedman CP (2005) "Smallball" evaluation: a prescription for studying community-based information interventions. J Med Libr Assoc 93(4 Suppl):S43–S48

Herrera IA, Woltjer R (2010) Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis. Reliab Eng Syst Saf 95(12):1269–1275

Hollnagel E (2004) Barriers and accident prevention. Ashgate, Aldershot

Hollnagel E (2008) The changing nature of risks. Ergonomics Austr J 22(1–2):33–46

Hollnagel E, Woods DD, Leveson N (eds) (2005) Resilience engineering: concepts and precepts. Ashgate Publishing Company, Brookfield

Hollnagel E, Pruchnicki S, Woltjer R, Etcher S. Analysis of Comair flight 5191 with the Functional Resonance Accident Model. Proceedings of the 8th International Symposium of the Australian Aviation Psychology Association, Sydney, 2008

Kaplan B (2001) Evaluating informatics applications-clinical decision support systems literature review. Int J Med Inform 64(1):15–37

Kawamoto K, Houlihan CA, Balas EA, Lobach DF (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 330:765

Lellouche F, Mancebo J, Jolliet RJP et al (2004) Computer-driven ventilation reduces duration of weaning: a multicenter randomized controlled study. Intensive Care Med 30:S69

MacIntyre NR, Cook DJ, Ely EW et al (2001) Evidence-based guidelines for weaning and discontinuing ventilatory support: a collective task force facilitated by the American College of Chest Physicians; the American Association for Respiratory Care; and the American College of Critical Care Medicine. Chest 120(6 Suppl):375S–395S

Mack EH, Wheeler DS, Embi PJ (2009) Clinical decision support systems in the pediatric intensive care unit. Pediatr Crit Care Med 10:23–28

McLean SE, Jensen LA, Schroeder DG, Gibney NR, Skjodt NM (2006) Improving adherence to a mechanical ventilation weaning protocol for critically ill adults: outcomes after an implementation program. Am J Crit Care 15:299–309

Meade MO, Ely EW (2002) Protocols to improve the care of critically ill pediatric and adult patients. JAMA 288:2601–2603

Morel G, Amalberti R, Chauvin C (2008) Articulating the differences between safety and resilience: the decision-making process of professional sea-fishing skippers. Hum Factors 50:1–16

Myneni S, Cohen T, Almoosa KF, Patel VL (2014) Standard solutions for complex settings: the idiosyncrasies of a weaning protocol use in practice. In Cognitive Informatics in Health and Biomedicine (pp. 183–202). Springer London

Myneni S, McGinnis D, Almoosa K., Cohen T, Patel B, Patel V (2011) Socio-technical barriers to effective use of a weaning protocol in a medical intensive care unit. Critical Care Medicine, 39(12):153

Osheroff JA, Pifer EA, Teich JM, Sittig DF, Jenders RA (2005) Improving outcomes with clinical decision support: an Implementer's guide. Healthcare Information and Management Systems Society, Chicago

Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE (2007) A roadmap for national action on clinical decision support. J Am Med Inf Assoc 14:141–145

Patel VL, Cohen T (2008) New perspectives on error in critical care. Curr Opin Crit Care 14:456–459

Randolph AG, Clemmer TP, East TD et al (1998) Evaluation of compliance with a computerized protocol: weaning from mechanical ventilator support using pressure support. Comput Methods Programs Biomed 57:201–215

Rose L, Presneill JJ, Johnston L, Cade JF (2008) A randomized, controlled trial of conventional versus automated weaning from mechanical ventilation using SmartCare/PS. Intensive Care Med 34(10):1788–1795, Epub 2008 Jun 25

Sinuff T, Cook D, Giacomini M, Heyland D, Dodek P (2007) Facilitating clinician adherence to guidelines in the intensive care unit: a multicenter, qualitative study. Crit Care Med 35:2083–2089

Sundström GA, Hollnagel E. Modelling risk in financial services systems: a functional risk modelling perspective. Third resilience engineering symposium, Antibes—Juan-les-Pins, France, 2008

Van Maanen J (1996) Ethnography. In: Kuper A, Kuper J (eds) The social science encyclopedia, 2nd edn. Routledge, London, pp 263–265

Vitacca M, Clini E, Porta R, Ambrosino N (2000) Preliminary results on nursing workload in a dedicated weaning center. Intensive Care Med 26(6):796–799

Weiss CH, Amaral LA (2010) Moving the science of quality improvement in critical care medicine forward. Am J Respir Crit Care Med 182(12):1461–1462

Chapter 11
Virtual Worlds in Healthcare

Prabal Khanal[1], Ashish Gupta[2], and Marshall Smith[3]

[1]3D Systems Corporation, Andover, MA, USA
[2]College of Business, University of Tennessee Chattanooga, Chattanooga, TN, USA
[3]Banner Health, Phoenix, AZ, USA

Abstract A well-structured medical education system focusing on proper cognitive as well as psychomotor skills is the key to better health care. However, medical errors still remain one of the leading causes of death in the United States. The recent advancement of healthcare technologies along with various tools and techniques help with the information sharing and provide a better mechanism for learning. One such technology that is gaining popularity as a tool for delivering medical education and health care is virtual worlds. Virtual worlds (VWs) are based on web 2.0 technologies that enable users with internet connectivity to access the respective systems any time. This review paper presents various functions and applications of VWs in medical education and healthcare. This paper surveys various architectures of VWs that focus on medical education and identifies related training tasks that can be achieved using VWs.

Keywords Virtual worlds • Medical training • Healthcare • Literature review

11.1 Background

Medical error is one of the leading causes of death in the United States (Committee on Quality of Health Care in A and I. Medicine 2000). It has also been proposed that periodic trainings and examinations should be implemented to reduce the medical errors (Committee on Quality of Health Care in A and I. Medicine 2000) in order to

P. Khanal (✉)
3D Systems Corporation, Andover, MA, USA
e-mail: prabalkhanal@gmail.com

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_11

improve healthcare (and patient-care) (Committee on Quality of Health Care in A and I. Medicine 2000; Koerner 2003). For better patient-care, a medical professional (doctor, nurse) must have high level of expertise in cognitive as well as clinical skills. Cognitive skills are generally associated with one's abilities to know and correctly use theoretical concepts for making healthcare decisions, whereas clinical skills are tied to psychomotor abilities required for performing various clinical procedures. In the traditional method of learning, cognitive skills are acquired from text books, in-class lectures, and ward-based medical education and clinical skills are generally acquired by performing various clinical procedures on cadavers and manikins (Rhienmora 2007). The main disadvantage of the traditional cognitive learning methods is that students generally do not retain the skills after 10 months (Schwid et al. 1999). In case of clinical skills learning, the necessary reliance on cadavers and manikins for training purposes is expensive and prone to damages leading to their restricted use. Furthermore, some clinical procedures require team-based training involving interaction and communication among the team members. This becomes an issue when the members are located at disparate places.

With the advancement of technology, various tools and techniques have evolved to improve training for both cognitive and clinical skills (Koerner 2003). Medical training focused on delivering clinical skills is increasingly adopting Virtual Reality (VR) based simulation techniques that are now proving to increase the retention period (Schwid et al. 1999). Additionally this technique provides two major benefits over the traditional techniques: first, such techniques allow medical students and care providers to practice using computer-based training models that can be used multiple times without any risk to the models, unlike the traditional physical models; second, such approaches provide more number of practice sessions for the students to improve their procedural skills. A majority of state-of-art VR simulators focus on single user (Center 2010). Existing research alludes to the necessity for such tools to not only focus on providing individual procedural training but go beyond to offer group oriented training. However, existing VR simulators haven't yet addressed the issue of team training in medical education. To fill in this gap, web 2.0 tools such as Social Networks (SN) and Virtual Worlds (VW) have been instrumental in facilitating theoretical learning, exchanging ideas and information, while providing a platform for discourse and interactions.

Given the implications of medical education on improving healthcare delivery and reducing avoidable errors, this review focuses on providing an overview of different virtual worlds and their applications that are currently used for medical training and education purposes in an effort to encourage research in this area.

11.1.1 Virtual Worlds: History and a Brief Background

Virtual world (VW) is a kind of social network that provides 3D virtual environments for participants to connect and share a common virtual space through a computer network (Kawamoto et al. 2006; Steve et al. 2001). Since several participants can simultaneously join and perform collaborative tasks, VWs are also referred to as collaborative virtual environment (CVE) (Steve et al. 2001). Kawamoto et al. describes VW as combination of computer supported cooperative work (CSCW) and virtual reality (VR) (Kawamoto et al. 2006). CSCW can be described as inherently being a computer supported activity that is performed by a team of collaborating individuals, often at different locations, in a computer-based shared environment (Steve et al. 2001; Baecker 1995).

VW represent users by their 'avatars' that can communicate with each other using text or voice. The VWs provide two added functionalities to the users: a) content and information sharing, and b) social interaction. VWs ability to share information in multiple different formats such as auditory, visual, and text makes them increasingly conducive for conducting remote training and education. VW also provide animated gestures for users to support multi-modal communication platform, which makes the interaction more immersive.

VW evolved from gaming environment. The Maze War was the first multi-user system that introduced users and developers to a multi-user 3D environments connected through a network (ARPANET) (digibarn.com 2004). However, the system had restricted access due to 24/7 un-availability. The prevalence of Internet and the evolution of GUI based browsers propelled the development of 'always on' 3D social spaces. Several 3D VW such as AlphaWorld, WorldsAway, Traveler evolved during the mid-90s that allowed for social interaction among the users (Terdiman 2006). Since then, a large number of 3D VWs have evolved that focus on different domains such as gaming (e.g. World of Warcrafts), education, business, and entertainment (e.g. Second Life, Tixeo, I-maginer, Moove). The major application areas are briefly explained in the following section.

11.1.2 General Applications of VW

VW have applications in a wide spectrum of fields and can be designed to have various functionalities (discussed later in Sect. 11.2.1). The major fields where VW can be used are: (1) social interaction, (2) education and training, (3) games and entertainment, (4) business and e-commerce, and (5) healthcare.

VW can provide for more realistic social interaction by incorporating features such as 3D immersive environment, avatars, and voice communication. Such design features can facilitate interaction among users through virtual conferences where real participants from the real world can attend (Terdiman 2006). A majority of traffic within any VW environment is typically for **gaming and entertainment purposes**. For example, World of Warcraft, a 3D VW game, has more than 12 million subscribers as of October, 2010 (Warcraft 2010).

Several applications of VW have now been emerging for **business use as well**. Various e-commerce tasks such as currency exchanges, transactions, etc. and development tasks such as virtual islands, renting virtual houses etc. are performed within VW. In addition to buying and selling activities, various companies such as IBM, Coke, Dell, Nissan have now build their virtual islands in the VW (Messinger et al. 2009).

Although VWs are primarily used for social interaction, entertainment and business, they have a wide range of applications and utilities in higher **education**. The advantage of using such tools for education purposes is to enable participants (students, trainees, mentors) to attend remote training from anywhere and anytime. VWs are capable of incorporating visual, interactive and dynamic simulations of various educational procedures across different majors (Wiecha et al. 2010; Schmidt and Stewart 2010). Such systems allow users to interact with the system and receive real time feedback thus facilitating our understanding of the dynamic behavior of a particular system or procedure (Callaghan et al. 2009). Features such as synchronous interactions, voice communication and 3D models further add to their abilities for training and educational purposes (Messinger et al. 2009).

The following section discusses various research activities performed in the field of medical education as well as healthcare using VW.

11.2 Related work: VW in Medical Education and Healthcare

As mentioned in previous sections, virtual reality (VR) based medical training systems for medical education and healthcare have already been available in medical education. The VR-based training simulators are designed and developed so a trainee becomes more comfortable with scenarios when presented with similar real life medical situations thereby resulting in less error in healthcare. However, most of these training simulators focus on improving psychomotor and/or cognitive skills. In addition, they lack the team-based training components as well as supporting communication aspects that are integral to group medical training. Since the architecture of VWs allow us to integrate various parties at one common space, VWs have greater potential to provide team-based training for medical students as well as healthcare professionals. Several more recent studies have focused on training, education, awareness and healthcare delivery. The applications of VWs in the field of healthcare and education can be classified into four major categories that are explained below with examples.

11.2.1 Education for Medical/Nursing Students

VWs provide a platform to replicate real life environments, scenarios, and cases in a virtual environment, which allows for the creation of various real life scenarios in the virtual world focusing on areas such as diseases, disability and different medical procedures. In addition to cognitive and psychomotor skills, VWs provide training on procedural as well as communication skills required for team-based or collaborative training, which makes the VWs best suited for case-based or problem-based learning (PBL). Such kind of collaborative training is applicable in emergency scenarios within or outside a hospital facility.

Several research studies report on the potential of VWs in the field of medical education. For example, Chodos et al. (2010) described two case studies on Emergency Medical Training/Emergency Room (EMT/ER), training simulation and InterD-410 course. The EMT/ER training simulation focuses on providing training to EMT/ER personnel about basic procedures for assessing and stabilizing accident victims before transferring them to the hospital (for EMR personnel). For ER personnel, this simulator focuses on training for receiving, assessing and starting different treatments for the victims. This simulator intends to provide training on handoff process that involves information exchange by the EMT personnel to the ER personnel. The Inter-D410 course simulation explains the idea that health delivery is a team activity and intends to teach this concept to the medical students. Conradi et al. (2009) designed a VW training simulator for training paramedic students using five different scenarios (or cases). Schmidt and Stewart (2010) explains how VW can be used to familiarize nursing students with various public health services as well as public health nursing interventions. Different scenarios that are included in the game include nutrition activity, virtual support groups, disaster scenario, public health resources and libraries. The students had to work in a team on each scenario during the training sessions. Play2Train (Boulos et al. 2008) is another example of an environments that train the participants on how to assign different roles to complete a set of tasks in emergency conditions.

VWs can be helpful in other aspects of medical education such as organizing didactic sessions. This is similar to tele-conferencing but VWs provide an additional advantage by enabling better direct interaction among the participants through virtual objects and animation procedures unlike tele-conference where participants can only see and/or hear others using different media. Ann Myers Medical Centre (2009) and nursing program at Duke University (Johnson et al. 2009) use VW to create a meeting place in order to present virtual lectures and educational materials to the students, and interact with each other.

11.2.2 Training for Healthcare Professionals

Another application of VWs is to provide training to healthcare professionals. Since all the VWs are using Internet based technology, multiple healthcare professionals can simultaneously login to the same virtual space for interacting with each other. This enables VWs to hold virtual seminars on current medical affairs, training methods, and latest medical technologies. According to Orizio et al. (2010), the first International Virtual Association of Surgeons meeting was held in Second Life (SL) in 2008 and was attended by 47 delegates from five nations. The major objective of this meeting was to exchange ideas on using surgical robots and latest surgery technology. In addition to interaction, VWs have been in use to provide various kinds of training to healthcare professionals. Melús-Palazón et al.'s (2012) work designed a Continuing Professional Development (CPD) program for primary healthcare professionals in SL. In their study, two training workshops for healthcare professionals from nine health centers were organized.

The virtual training was divided into several sessions based on different fields such as preventive medicine, family and community medicine, prescription drugs, and new technologies. However, the authors mentioned that less than half of the participants considered VW to be equal or superior that of face-to-face methods. In a similar study by Wiecha et al. (2010) VW was used to organize a virtual seminar for medical education on insulin therapy for type 2-diabetes to evaluate the effectiveness of the virtual world training with a pre-post questionnaire distribution.

11.2.3 Care Delivery Management

While VWs have prominent applications in entertainment and education, researchers are continuously trying to find various avenues to provide better healthcare using VWs. Although VWs have potential for effective healthcare, the real physical examination of patients is not possible using the existing technology and architecture of VWs. However, that does not limit the possibilities of VWs to provide care for patients with psychological disorders. Since VWs provide an environment to involve both patients and their therapists, they have been already in use for providing behavioral healthcare to the patients. In World was the original online tool to deliver behavioral healthcare by enhancing cognitive therapy, counseling, training, education, and supervision (Solutions 2010). SECTER (Simulated Environment for Counseling, Training, Evaluation and Rehabilitation) is another customized VW where patients are assigned roles and can communicate with the avatars of their therapists (Frenkel 2009). These environments are being used for treatment of troubled teens, patient's with Asperger's syndrome (Phillips 2008), anorexia and bulimia, anxiety disorders, post-traumatic stress syndrome, alcoholism, and disabilities in stroke victims (Frenkel 2009). Some VW provide various ongoing activities in SL where patients can login to various virtual spaces and meet people similar to them (Parsons 2008). This feature offers the patients a virtual platform to socialize and get information about their physical and cognitive disabilities.

11.2.4 Awareness

Awareness implies having the knowledge of something. The process of people being 'aware' of something varies with different domains. Since VWs do not disclose the identities of users, it encourages users to share medical information with other users. There are numerous virtual spaces—'islands' in SL dedicated to health awareness. These islands allow participants to learn symptoms, prevention, and care for particular health related problems. For example, Center of Disease Control has its own island in SL. The major objective of CDC island is to create awareness in public about health related problems and remediation Similarly, an island in SL is available for Association for International Cancer Research to provide awareness related to cancer. Boulos et al. (2008) created an island for

HIV-AIDS awareness. Virtual worlds have tremendous potential for use in other health problems such as smoking and alcoholism.

The four major categories discussed in this study are not necessarily disjoint. The applications in medical education for students also apply for healthcare professionals. The research study by Wiecha et al. (2010) on insulin therapy can be applied to medical/nursing students as well. Likewise, EMT training can be used to train healthcare professionals while renewing their certification.

11.3 Architectures of VWs for Medical Education and Healthcare

This section explains various VWs architectures. (Thrasyvoulos et al. 2009) provides information regarding various VWs that are currently available and compare various VWs primarily for educational purposes. This study concludes that the selection of VW for a training system depends on the purpose of the system (Thrasyvoulos et al. 2009). What we can infer from their conclusion (Thrasyvoulos et al. 2009) is that the selection of a VW should be based on the complexity of a training system, which can be gauged once the system-architecture is known. For this purpose, we have categorized the architectures that have been used in the development of medical education as well as healthcare systems using VW. This classification will be helpful for researchers/developers while designing a VW-based system for medical education and healthcare.

We first define the terminologies used in this section and its sub-sections. The *architecture* of a VW is based on a platform created using various objects and component as modules. *Objects* in this context are the virtual elements that represent real entities or things such as people (avatars), rooms, tables, mannequins, etc. *Components or modules* are parts of a design that represent non-physical elements of the space. The examples of such modules are learning module, device handling module, etc.

The most basic architecture follows client–server model where a VW client sends requests through the Internet to a remote VW server to access the objects that are available. The server responds to the client's requests and the objects or modules are then displayed at the client's computer. Most of the communities built in VW, including virtual environments designed for cognitive behavior therapy (Frenkel 2009; Phillips 2008; Nation 2010; Mollman 2008) are based on this architecture. Although the basic architecture supports easier development and allows for basic interaction among objects and avatars in the VW, it does not allow users to create customized interaction with the objects in the VW or record the activities (or actions) performed inside the VW.

Several advanced objects and components/modules, typically created using third party software, can be added to the existing basic architecture. They then need to be converted into native-virtual-world format and uploaded to the VW server from where they can be downloaded back to the VW client to be displayed. VWs provide a list of interaction scripts, which are responsible for different actions performed by 'avatars' as well as other animation sequences performed by objects inside the VW.

11.3.1 Web-Based Learning Component

An important aspect of the VW is to be able to assess the trainees on their learning. A web-based learning component can be added to the basic architecture so that the trainees have to answer the questions relevant to their VW sessions. These are made available on separate websites other than the VW. There are two types of web-based learning components for VW:

1. Separate training and assessment components
2. Integrated training and assessment components

The architecture that includes separate training and assessment components (Fig. 11.1a) works similar to e-learning methods where students/trainees get tested after completing the learning session in the VW. The evaluation procedure starts with the users following the tasks of a VW training session. They then would login into the website containing the test at the client's computer, separated from the VW system. The scores then get stored in a database on the remote server for performance assessment. One VW system that follows this architecture was designed by Weicha et al. (2010).

In the integrated training and assessment modules, the training module (in VW) and testing module (web-based) are both connected in the client computer (see Fig. 11.1b). The procedure of evaluation starts with users logging into the website where the tasks are listed. Then they have to visit the VW client and perform the tasks while they are concurrently being monitored and evaluated. Evaluations are stored in the database as the tasks are performed inside the VW. The advantage of this type of component is that the users do not have to re-login to the website to take the test. The limitation of this architecture is that the developers can only have basic default interactions provided by the VW. Callaghan et al. designed a similar education system by integrating VW (Second Life) and virtual learning environment (SLOODLE) (Callaghan et al. 2009). There are not other known implementations that integrate VW to website component directly.

11.3.2 Interaction Control Component

The VW architecture that includes an interaction control component provides users with the ability to apply tailored interactions between the avatars and the objects (Fig. 11.2). In this architecture, the VW client can be controlled by user-defined script. With user-defined script, custom interactions as well as assessment criteria can be designed in the VW. This architecture is better suited for medical education

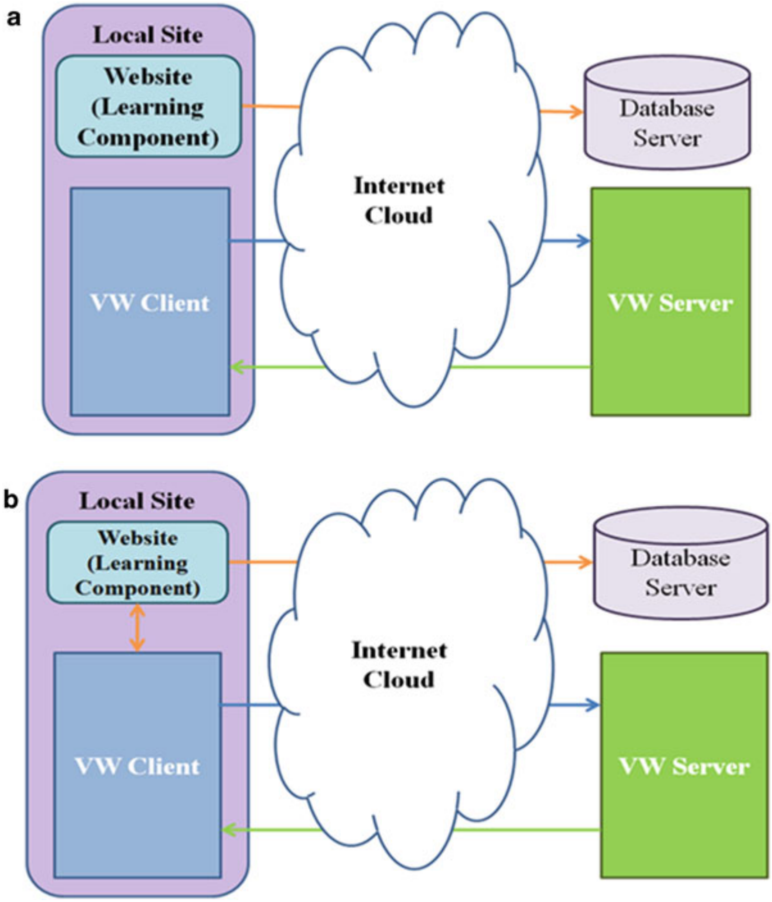


Fig. 11.1 (**a**) Architecture with website learning component (Separate training and assessment modules). (**b**) Architecture with website learning component (integrated training and assessment modules)

later Section (11.3.1) because custom interaction can be built to train users more effectively. Similarly, custom scripts also facilitate complex interaction among various stakeholders. Examples include interactions between patient and patient, patient and clinician, and clinician and clinician. Database server acts as the central repository of all the actions/interactions that are performed in the VW. Depending on the requirements, variables such as time taken to complete a task, task completion level, and scores can be stored in the same database server. Instead of database server to record the values of different variables, plain text files can also be used for limited capabilities. Both storage techniques (database or plain text files) enable data to be used for providing feedback to the users and performance evaluation purposes.

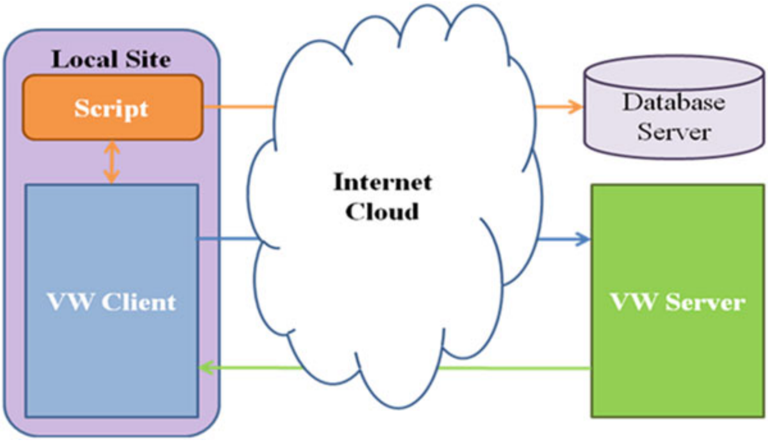


Fig. 11.2 Architecture with interaction control script component

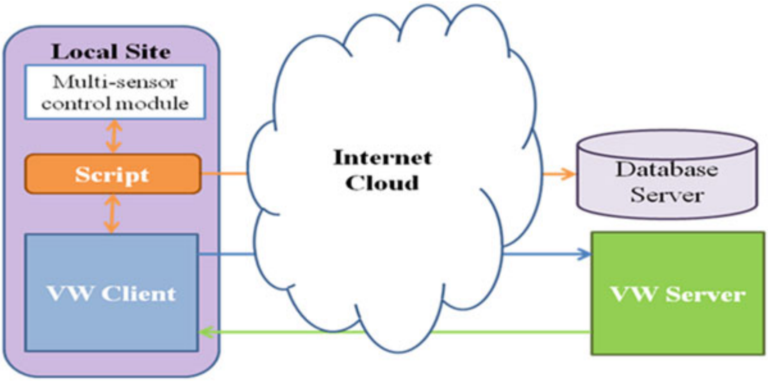


Fig. 11.3 Architecture integrating multi-sensory devices with VW

11.3.3 Multi-Sensory Devices Control Modules for VW

This architecture introduces a mechanism to integrate multi-sensory devices like RFID tags, accelerometer, and haptic devices with VW. In this architecture, the sensor devices can be connected to the local (client) computer and can be controlled by user-controlled scripts that are integrated with interaction control scripts and the VW system (Fig. 11.3). Smart Condo Project (Eleni et al. 2009) discusses about the integration of RFID tags to SL. Khanal and Kahol (2010) presents similar method to integrate haptic device to a VW (Active Worlds) and implement the architecture to train CPR skills.

11.4 Interaction

Various types of interactions need to be supported within the VW due to varying user roles. For example, a user could be playing the role of a teacher or a student within the VW. The original role of the user could also use other specifications such buyer-seller, trainer-trainee, doctor-patient, etc. In the field of medical education, a predominant set of roles is teacher-student (or master-apprentice, or trainer-trainee) where teachers share training resources and students learn theoretical and clinical skills. Virtual worlds are also able to support other forms of faculty-student interactions described earlier. As VW provide more immersion within the 3D virtual environment, and users login to the VW with their own avatar, the interactions are more life-like as the avatars will be interacting with other avatars. They can also interact with text-based chat or voice-chat.

Healthcare is another critical domain where SN and VW can be helpful tools for outcome improvements. A large body of research focuses on patient-provider communication. The communication, or interaction, required for better healthcare system could be broadly categorized into three groups:

1. Patient-patient interaction
2. Patient-care provider interaction
3. Provider-provider interaction

11.4.1 Patient-Patient Interaction

The main objective of providing patient-patient interaction in SN as well as VW is for patient to quickly receive some information about diseases, risk factors, symptoms, and solutions from one source rather than searching from multiple sources. Healing Island is a virtual space in SL with several hundred members who can interact with each other to recover from depression, mental illness and emotional trauma by engaging in virtual friendship (Parsons 2008). Boulos et al. (2008), and Press kit (Luminos 2008) are different islands in SL that are dedicated to provide HIV/AIDS education and outreach. Likewise, Fearless Nation PTSD Retreat (Nation F 2010) is a virtual space in SL that offers information and support for the patients with PTSD or combat-related stress that are not ready for in-person contact (Anonymous 2010).

11.4.2 Patient-Care Provider interaction

This category of interaction is often the most complex to be integrated with VW. There is always a thin margin between the openness of doctor to the patient and the level of secrecy to be maintained by the doctor (Guseh et al. 2009). Virtual worlds, with the aid as well as limitation of the current technology, are used to deliver behavioral healthcare (Solutions 2010). Cognitive behavior therapy typically has a higher degree of patient and therapist engagement and often involves role-play while mimicking different situations that a patient may have to handle. However, patient with attention deficit hyperactivity disorder (ADHD) affects the patient's ability to engage and focus in role-play situation (Frenkel 2009). Virtual worlds provide an opportunity to interact with therapists in an immersive environment. Patients often find themselves familiar with the virtual environment and are able to easily interact with the therapists using voice chat (Frenkel 2009). Similarly, patients with Asperger's syndrome (Phillips 2008), anxiety disorders, post-traumatic stress disorder (home C 2009), and alcoholism (Mollman 2008) has been treated with VW solutions. Future VWs could focus on improved patient-therapist interaction with sensor-enabled features such as 3D depth sensing and data analytic support.

11.4.3 Provider-Provider Interaction

An example of such interaction is communication between medical resident and physician. Various educational virtual spaces (Chodos et al. 2010; Johnson et al. 2009; Boulos et al. 2007) are developed in recent years where medical students and physicians (mentors) interact with each other. Johnson et al. (2009), present a way to organize a virtual meeting place in VW where physicians can interact with other physicians and students on different topics. A good example of such VW where multiple providerr-provider interactions take place are described in Khanal et al. 2014; Vankipuram et al. 2014.

11.5 Future Research

VWs are still in early stage of development in healthcare domain. There are several areas where the use of VW could be improved and expanded to different conditions. For example, the next generation of VW could focus on healthcare stakeholders such as care providers, patients, trainers/educators, trainees, etc. VWs could be designed to integrate other components of healthcare delivery such as medical education, awareness, etc. We describe below several such components

11.5.1 Use of Sensory Devices

Various haptic devices could be integrated with VW for conducting intensive group trainings such as cardiopulmonary resuscitation (CPR) skills training. The trainees logged in to a virtual world, choose their "avatar" and learn how to perform CPR on a haptic joystick. The avatar can visualize the CPR gesture in the virtual world when the CPR was being performed in the real world hat is integrated with VW. Similar design principles could be used for implementation of advanced cardiac life support

(ACLS) training system, which opens several novel opportunities for researchers to consider using VWs for time intensive critical training in healthcare. One of the examples would be a *remote patient monitoring system* in VWs. Different kinds of sensors can be attached to patients or people with disabilities. These sensors can transmit location information as well physical information such as temperature, oxygen saturation level, and heart rate to the VW server. This information can be remotely monitored, allowing for patients to stay at home and be taken care of from a remote care provider site. Additionally, VWs can maintain anonymity so patients do not have to disclose their identity while they are being monitored. This may have a potential impact on the number of readmissions to hospitals significantly.

11.5.2 Team-Based Learning and Patient Education

Team-based learning focuses on extracting knowledge from an activity where several individuals are involved in situations such as in emergency. Although individual medical training is the common form of training, it does not train the individual in important skills such as communication between the team members and coordination of the tasks in the emergency scenarios. Furthermore, in conventional face-to-face training, the team-based training requires individuals to be at the same physical location, which is difficult to organize frequently due to conflicting schedules. VWs have tremendous potential for use in such team-based learning where they have the capabilities of providing both cognitive as well as psychomotor skills.

11.5.3 Frequent Training Sessions and Improved Patient Care

The team-based learning in VWs can also be implemented for reducing errors in patient care processes or delivery. VWs provide a platform where participants can meet from remote places and organize multiple training sessions unlike infrequent training sessions that are typically organized once or twice a year. Since it requires more practice to be proficient in any medical task, VWs can prove to be of tremendous value and serve as a platform where participants can learn about adherence to the guidelines of special team-based medical activities.

11.5.4 Patient Education and Awareness

One of the major application areas of VWs is to raise awareness against diseases and patient literacy. VWs provide an environment where patients can login anonymously and interact with each other. This provides a very important opportunity to create a system to educate patients about diseases, especially chronic diseases.

11.5.5 Avatars in Teaching

"Avatar" is the most important component within a virtual world. An avatar is a character in a virtual world that represents an individual in a real world. They can be customized as per individual requirements and can perform various set of actions or animations such as walking, dancing to name a few. Unlike in traditional web-based learning, avatars provide more interactive and immersive environment for learning along with better communication. Future work could focus on developing teaching components where training on different medical tasks, diseases or any other activities can be provided from remote locations

11.6 Conclusion

In this paper, we performed an extensive review of extant literature related to key elements of VW related to its development and implementation. For example, various architectures used for designing VWs for medical education, important persuasive components that can help users to continue using VWs, interaction features among different groups of users are also discussed. The categorization of the architecture, supporting persuasive components and interactions provide researchers and developers some guidelines on choosing the appropriate design for their VW based on the need assessment.

Various medical educational institutions are seeing a surge in SNs and VWs use as supplementary materials (Hewett (IAMSE) 2010). However, current technological limitations, the complexity of team based clinical procedures and trainings are keeping the adoption of VWs still in embryonic stage. Current research works focusing on integration of various sensors with VWs (Eleni et al. 2009; Khanal and Kahol 2010) are good indicators of the possibilities of practicing and enhancing users' procedural skills in a virtual team environment. In the field of healthcare, VWs are predominantly used in cognitive behavior therapies (Frenkel 2009; Phillips 2008).

There is, however, an issue of content reliability with VWs. Content validation is the only way to verify the reliability of the SNs (Orizio et al. 2010). To the best of our knowledge, no validation mechanisms are yet available in the current healthcare VWs. Some of the VWs (Messinger et al. 2009; Wiecha et al. 2010; Schmidt and Stewart 2010) in medical education undergo usability and learnability validation of the system based on survey questionnaires. There is a significant paucity of validation studies that focus on outcomes of VW usage in healthcare. This is a critical research problem in the field of medical education and healthcare.

References

Anonymous P (2010) Option: fearless nation. [cited 2010], http://www.ptsdanonymous.org/option-fearless-nation

Baecker RM (1995) Readings in human-computer interaction: toward the year 2000. Morgan Kaufmann Publishers, San Francisco

Boulos MN, Hetherington L, Wheeler S (2007) Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. Health Info Libr J 24(4):233–245

Boulos K et al (2008) Web 3D for public, environmental and occupational health: early examples from second life. Int J Environ Res Public Health 5(4):290–317

Callaghan, MJ et al (2009) Integrating virtual worlds: virtual learning environments for online education. Games Innovations Conference. ICE-GIC 2009. International IEEE Consumer Electronics Society's, 2009

Center AMM (2009) Ann Myers Medical Center. [cited 2010 September 1]. http://ammc.wordpress.com/

Center PHS (2010) Virtual reality simulators. [cited 2010 November 15, 2010] http://pennstatehershey.org/web/simlab/home/available/vrs

Chodos D et al (2010) Healthcare education with virtual-world simulations. Proceedings of the 2010 ICSE Workshop on Software Engineering in Health Care, ACM: Cape Town

Committee on Quality of Health Care in A. and I. Medicine. (2000) To Err Is Human: Building a Safer Health System. National Academies Press.

Conradi E et al (2009) Virtual patients in a virtual world: training paramedic students for practice. Med Teach 31(8):713–720

digibarn.com (2004) The DigiBarn's Maze War 30 Year Retrospective "The First First Person Shooter", [cited 2010 October, 2010]. http://www.digibarn.com/history/04-VCF7-MazeWar/index.html

Eleni S et al (2009) Software engineering for health education and care delivery systems: The Smart Condo project. Proceedings of the 2009 ICSE Workshop on Software Engineering in Health Care, IEEE Computer Society, 2009

Frenkel KA (2009) Therapists use virtual worlds to address real problems. [cited 2010]. http://www.scientificamerican.com/article.cfm?id=therapists-use-virtual-worlds

Guseh JS 2nd, Brendel RW, Brendel DH (2009) Medical professionalism in the age of online social networking. J Med Ethics 35(9):584–586

Hewett JK (2010) The use of Social Media in Medical Education. International Association of Medical Science Educators (IAMSE) IAMSE web Seminar, http://www.slideshare.net/jkhewett/the-use-of-social-media-in-medical-education

home C (2009) Coming Home. [cited 2010], http://projects.ict.usc.edu/force/cominghome/.

Johnson C, Vorderstrasse A, Shaw R (2009) Virtual worlds in health care higher education. J Virtual Worlds Res 2(2):1–12

Kawamoto ALS, Kirner TG, Kirner C (2006) Experience on the implementation of a collaborative virtual environment for educational applications. Systems, Man and Cybernetics. SMC'06. IEEE International Conference on 2006

Khanal P, Kahol K (2010) Interactive haptic virtual collaborative training simulator to retain CPR skills, Arizona State University

Khanal, P., Vankipuram, A., Ashby, A., , Vankipuram, M., Gupta, A. Drumm-Gurnee, D., Josey, K., Tinker, L., Smith, M. (2014) Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles, Journal of Biomedical Informatics , Volume 51 , 49–59

Koerner JG (2003) The virtues of the virtual world. Enhancing the technology/knowledge professional interface for life-long learning. Nurs Adm Q 27(1):9–17

Luminos N (2008) Second Life Celebrates World AIDS Day with New Island. [cited 2010 December 2, 2010], http://www.associatedcontent.com/article/1257442/second_life_celebrates_world_aids_day.html?cat=3

Melús-Palazón E, Bartolomé-Moreno C, Palacín-Arbués JC, Lafuente-Lafuente A, García IG, Guillen S, Esteban AB, Clemente S, Marco AM, Gargallo PM, López C, Magallón-Botaya R (2012) Experience with using second life for medical education in a family and community medicine education unit. BMC Med Educ 12:30

Mesko B (2007) Top 10: virtual medical sites in Second Life! [cited 2010]. http://scienceroll.com/2007/06/17/top-10-virtual-medical-sites-in-second-life/. Accessed 17 June 2007

Messinger PR et al (2009) Virtual worlds—past, present, and future: new directions in social computing. Dec Support Syst 47(3):204–228

Mollman S (2008) Avatars in rehab: Getting therapy in virtual worlds. [cited 2010]. http://articles.cnn.com/2008-07-16/tech/db.secondlifetherapy_1_second-life-avatar-virtual-world?_s=PM:TECH. Accessed 16 July 2008

Nation F (2010) Fearless nation PTSD retreat. [cited 2010], http://secondlife.com/destination/fearless-nation-ptsd-retreat

Orizio G et al (2010) The World of e-Patients: a content analysis of online social networks focusing on diseases. Telemed J E Health 16(10):1060–1066

Parsons C (2008) Second Life offers healing, therapeutic options for users. [cited 2010]. http://articles.sfgate.com/2008-07-13/living/17173237_1_second-life-virtual-world-virtual-spaces. Accessed 13 July 2008

Phillips A (2008) Asperger's therapy hits second life: experts express concern about applying online actions to real life. abcNEWS

Rhienmora P (2007) Visualization and virtual reality in clinical skills training. Asian Institute of Technology, Bangkok

Schmidt B, Stewart S (2010) Implementing the virtual world of Second Life into community nursing theory and clinical courses. Nurse Educ 35(2):74–78

Schwid HA et al (1999) Use of a computerized advanced cardiac life support simulator improves retention of advanced cardiac life support guidelines better than a textbook review. Crit Care Med 27(4):821–824

Solutions I (2010) [cited 2010]. http://www.inworldsolutions.com/

Steve B et al (2001) Collaborative virtual environments. Commun ACM 44(7):79–85

Terdiman D (2006) A brief history of the virtual world. [cited 2010]. http://news.cnet.com/A-brief-history-of-the-virtual-world/2008-1043_3-6134110.html

Thrasyvoulos T et al (2009) Implementing collaborative e-learning techniques in collaborative virtual environments: the case of Second Life. Proceedings of the 2009 Fourth Balkan Conference in Informatics, IEEE Computer Society, 2009

Vankipuram, A., Khanal P, Ashby A, Vankipuram M, Gupta A, DrummGurnee D, Josey K, Smith M, Design and Development of a Virtual Reality Simulator for Advanced Cardiac Life Support Training, IEEE Transactions of Biomedical & Health INformatics, 18 (4)

Warcraft WO (2010) World of Warcraft® Subscriber Base Reaches 12 million worldwide. [cited 2010]. http://us.blizzard.com/en-us/company/press/pressreleases.html?101007. Accessed 7 Oct 2010

Wiecha J et al (2010) Learning in a virtual world: experience with using second life for medical education. J Med Internet Res 12(1):233–245

Chapter 12
Natural Language Processing for Understanding Contraceptive Use at the VA

Matthew Scotch[1,2], Cynthia Brandt[1,3], Sylvia Leung[4], and Julie Womack[1]

[1]VA Connecticut Healthcare System, West Haven, CT, USA

[2]Department of Biomedical Informatics, Arizona State University, Samuel C. Johnson Research Bldg, 13212 East Shea Boulevard, Scottsdale, AZ 85259, USA

[3]Yale Center for Medical Informatics, New Haven, CT, USA

[4]VA Palo Alto Health Care System, Palo Alto, CA, USA

Abstract Objective: To evaluate the potential of Natural Language Processing (NLP) for understanding contraceptive use among female Veterans seeking care at Veterans Administration (VA) healthcare facilities.

Design: Retrospective chart review of a subset of female Veterans enrolled in the Women Veterans Cohort Study (WVCS) who sought care at the VA Connecticut Healthcare facility (in West Haven, CT) in 2009 and completed a survey that included self-reported contraceptive use. In addition, only notes that were annotated for contraceptive use from a prior study that included 227 patients WVCS participants were selected.

Methods: A biomedical ontology of contraceptive terms and concepts was created that included both permanent methods (e.g. hysterectomy) as well as non-permanent methods (e.g. oral contraceptives). The new ontology, along with a section of the VA's National Drug File was used as the knowledge base for information extraction from the free-text medical records. Included were 208 annotated notes across 39 patients. The General Architecture for Text Engineering (GATE), an open-source application for development of NLP pipelines was used. The ontology was added to GATE along with a processing resource that was developed in order to create an ontology-aware information extraction plugin for the pipeline. In addition, prior resources developed for negation of concepts (*e.g.* The patient *denies* using a emergency contraceptive) were utilized.

The NLP pipeline extracted contraceptives currently used by the patient, ones not currently used (prior use or recommended use by the clinician), or whose use

M. Scotch (✉)
Department of Biomedical Informatics, Arizona State University, Samuel C. Johnson Research Bldg, 13212 East Shea Boulevard, Scottsdale, AZ 85259, USA
e-mail: matthew.scotch@asu.edu

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2_12

was negated. A Boolean matrix of concepts by each patient was produced for input into a decision tree classifier. Tenfold cross validation created iterations of training and testing sets to estimate active versus inactive contraceptive. Responses to self-reported contraceptive use on the prior survey were used as the gold standard.

Results: The use of manual annotation, development of a biomedical ontology, and creation of a natural language processing pipeline achieved high precision (0.83) and recall (0.84). The weighted F-measure was 0.83.

Conclusion: Our combined approach utilized annotation of concepts, a biomedical ontology of contraceptives, and a natural language processing pipeline for information extraction. Our results highlight the potential for biomedical informatics to support research of contraceptive use among female Veterans at the VA. Additional research needs to be done that evaluates the accuracy of contraceptive information in the VA's Electronic Health Record (EHR) with the consideration of both free text and semi-structured data such as pharmacy records.

Keywords Natural language processing • Contraceptive agents • Veterans • Medical informatics

12.1 Introduction and Background

Recent efforts by the Veterans Health Administration (VHA) highlight the agency's recognition of the central importance of reproductive health (Yano et al. 2006). Little attention, however, has been paid to contraceptive use. One explanation for this could be that information about contraception and its use is difficult to extract from the electronic health record (EHR) because structured data (e.g. ICD-9-CM, CPT, or V codes and pharmacy data) are often inadequate as women may receive their reproductive health care, including contraception, outside of the VHA. Unstructured data in free-text clinical notes may better capture contraceptive use. Natural language processing (NLP) is an automated approach to identifying relevant information from free-text and has been widely used in the clinical domain for health services research and decision support including adverse drug events (Warrer et al. 2012), cancer screening (Denny et al. 2012), Alzheimer's Disease (Lee and Gonzalez 2011), and radiology reports (Rubin et al. 2010). This approach presents challenges for capturing contraceptive use because of the wide variety of drug/device names including brand and generic names, as well as idiosyncratic abbreviations, synonyms, and misspellings. The use of biomedical ontologies may provide one approach for addressing this problem. Biomedical ontologies create semantic relationships between concepts and use hierarchies to establish formal definitions of concepts. The emerging role of biomedical ontologies is highlighted in the existence of the National Center for Biomedical Ontologies (NCBO) Biportal (2010; Whetzel et al. 2011), a resource of biomedical ontologies, many of which have been used in collaboration with NLP research in the clinical domain. *Ontology-*

aware NLP applications use ontologies as a knowledge base for identifying and extracting concepts from the free text clinical record.

Our goal was to evaluate the use of a contraceptive ontology in the context of an integrated biomedical informatics approach to information extraction to identify contraceptive use among female Veterans at VHA healthcare facilities. In addition to a contraceptive ontology, our approach includes the use of an ontology-aware NLP pipeline to extract contraceptive information from clinical progress notes. The output of our NLP pipeline is then loaded into a machine-learning model to assess active versus inactive (*i.e.* not currently using) contraceptive use. We compute precision, recall, and F-measure of these classifications using the results from a self-report survey on contraceptive use of participants enrolled in the Women Veterans Cohort Study (WVCS) (Womack et al. 2012). WVCS is a study of Veterans of Operations Enduring (Afghanistan) and /Iraqi Freedom (OEF/OIF) who are enrolled in VHA services (Haskell et al. 2011).

12.2 Methods

12.2.1 Ontology Development

We developed an ontology of contraceptives to represent a knowledge base for the NLP pipeline. The ontology contained concepts and synonyms (variations of names) that the authors determined to be relevant to contraceptives including medications and devices. We first considered existing ontologies such as SNOMED (IHTSDO 2012; Cote and Robboy 1980), the VHA's National Drug File (NDF) (2012), (Brown et al. 2004), and RxNorm (NLM 2012; Coonan 2004). However it was the opinion of the authors that none of the existing ontologies (with wide clinical and biomedical scope) adequately covered the multitude of contraceptive concepts and their variations of names. We thus developed a new ontology focused exclusively on contraceptives but also included the NDF as a supplement to increase our knowledge base.

One of the authors (JW), a certified nurse midwife, created the initial hierarchy of the ontology with assistance from another clinician (CB). To develop the ontology, we used Protégé version 4.1 (2011) (Noy et al. 2003). High-level categories included *female_permanent* (e.g. hysterectomy) and *female_reversible* contraceptives. Reversible contraceptives included *combined estrogen and progestin* medications as well as *progestin-only*. Once the initial structure of the ontology was developed, a third author (MS) added specific contraceptive names at the leaf-level using online sources including About.com (2012) and Drugs.com (2012). In addition, variants (such as brand names) were added as synonyms. The final version of the ontology (a portion which is shown in Fig. 12.1) was approved by all the authors.

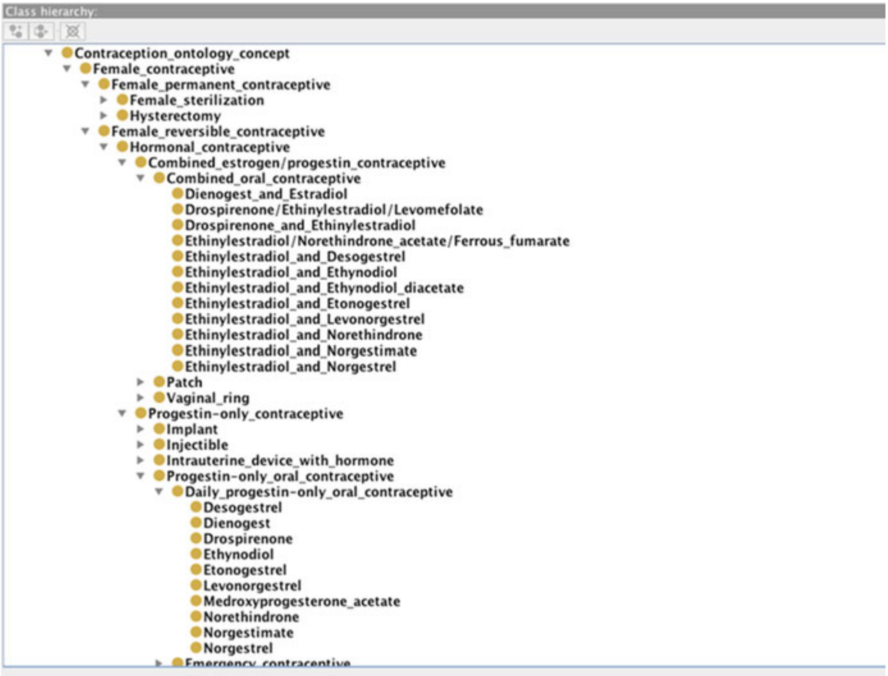


Fig. 12.1 A portion of the ontology of contraceptives

12.2.2 Annotation of Clinical Notes

In order to facilitate information extraction and to train the machine learning classifier, we used annotated corpora from our prior research (Womack et al. 2012). The notes originated from women enrolled in WVCS who participated in the baseline health survey (conducted in 2009/2010) and who had entries in the VHA electronic health record (EHR) in the year prior to survey completion. Fifty-nine women with 1739 notes met these criteria. Of these, 39 women representing 208 notes addressed contraception.

One of the authors (SL) led the development of an annotation schema, in collaboration with the other authors and the Consortium for Healthcare Informatics Research (CHIR) (Protege 4.1 2011), a VHA-based research group that focuses on the development and use of informatics tools to facilitate health services research at the VA. The schema was created within Protégé and Knowtator (Ogren 2006, 2007) to direct the annotation of the free-text notes. The annotation schema identified the contraceptive mention in the note (intrauterine device, hormonal method, barrier method, fertility awareness, permanent method, unspecified method and no method), the *status* of the contraceptive (*e.g.* current, past, future and potential use) and its *temporality* (*e.g.* duration, frequency or date of onset). *Temporality* and *status* could be annotated either as slots within the contraceptive class or as separate classes. For example, the phrase "Used DMPA in the past", would include a *contraceptive* annotation for 'DMPA', and as a separate class, a *status* annotation for 'Used'. However, if the provider had written "Contraception: DMPA" then 'DMPA' would be annotated as a *progestin-only* contraceptive, and within that annotation, the slot *status_implied* would be used. Here, in the absence of a specific status term (such as 'used') the assumption is made that the patient is currently using this form of contraception.

12.2.3 Knowtator Annotations

The annotations from Knowtator were exported as XML files which were then converted, using python script, into a slightly different XML format that was acceptable to GATE. Both the annotations and their corresponding text were then loaded as a corpus of documents into GATE.

12.2.4 NLP Pipeline

The NLP pipeline was developed using the General Architecture for Text Engineering (GATE) (2011), a free open-source platform created in the Java programming language. GATE supports a pipeline consisting of language and processing resources that are added to the pipeline as *plugins* and that perform information extraction. The GATE platform supports the use of existing resources as well as those developed by the user, including language and processing resources in Java. These resources can be *ontology-aware,* connecting to a specific ontology as a knowledge base for annotating corpora. Our pipeline included existing GATE resources (such as a tokenizer and a sentence splitter), as well as those we developed or were developed by our colleagues. For the latter, we used a package of processing resources created by colleagues at the Tampa VA and designed for negation of concepts. This plugin was based on the NegEx algorithm (Chapman et al. 2001) for negating clinical concepts (i.e. *decided against* emergency contraception).

We developed four processing resources to support our NLP pipeline. The first was a processing resource to map words from the clinical notes to concepts in our ontology. This resource enabled our pipeline to be *ontology-aware* and to utilize the knowledge and the structure of our ontology to better identify the concepts of interest. For example, *Dienogest and Estradiol*, two components of oral contraceptive pills, could be grouped into their higher-level class, *Oral contraceptive*, which in turn could be grouped into the higher class, *Hormonal Contraceptive*. In addition to our contraceptive-specific ontology, we also included the *Contraceptives*, *Systemic* branch from the NDF and loaded it into GATE as a gazetteer.

Our second resource processed the annotations from Knowtator and considered the presence of any slots or status classes to define the contraceptive as *active* or *inactive*. Here, any concept annotated as *Active* or *Sporadic Use* was labeled *Active*. On the contrary, any concept annotated *Retired* (*e.g.* "Patient had used Desogestrel for 3 years"), or *Prescribed or Recommended*, (*e.g.* "Talked with patient about the benefits of the patch") was labeled *Inactive*. In addition, anything annotated as *No_Contraception* was labeled *negated*. This negation information was used in addition to the NegEx plugin.

Our pipeline extracted both the Knowtator annotations (done manually by the two annotators) as well as any concept mapped to our ontology. However, ontology concepts within the span of a Knowtator annotation were not considered 'novel'. Only ontology concepts found outside the span of a Knowtator annotation were considered novel. We thus created a third processing resource that identified the span of each annotation and assigned an ontology concept as 'novel' if it met our criterion.

Our final processing resource was a writer that wrote the annotations to a flat file. For each patient, we considered only the presence or absence of the concept and did not consider how many times the concept was annotated in the text. The final version of the pipeline is shown in Fig. 12.2.

12.2.5 Machine Learning

We sought to evaluate the accuracy of classifying a patient's contraceptive use (of any kind) as *active* or *inactive* based on the concepts that were extracted by the NLP pipeline. We first wrote a preprocessing script in python to create a Boolean matrix (of presence/absence) of each concept per patient. We used a patient study identifier to link the results back to the survey (our gold standard). On the survey, we considered two questions. One asked *yes/no* if they used contraceptives in the last 12 months. The other asked if they used condoms. Since we considered condoms as a form of contraception, a participant indicating "no" on the first question but "yes" to condoms was classified as actively using contraception. We added an additional *class* column in our matrix that represented active or inactive use of contraceptives. We used Weka (2012), a publically available software tool, for the machine learning classifier. In Weka, we selected the J48 decision tree classifier and used tenfold cross validation for training and testing of our data. Weka reports statistics such as precision and recall in order to evaluate the classifier.

12.3 Results

Our final result set included 37 patients. Two of the original 39 did not complete at least one of the two questions on the survey and were thus excluded from the analysis. We first explored how well our machine-learning classifier identified active

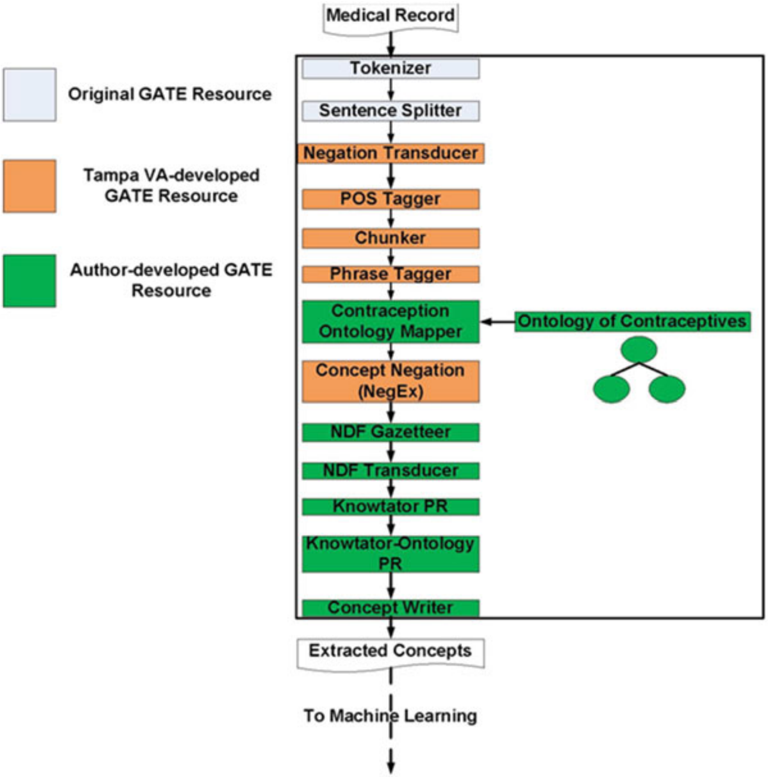


Fig. 12.2 GATE pipeline for information extraction of contraceptives. The processes are color-coded to denote their source. *Light blue boxes* represent processes that are available in the standard GATE environment. *Orange boxes* represent those provided by the Tampa VA. *Green boxes* represent those developed by the authors

and inactive contraceptives. The results after tenfold cross validation included a true positive rate of 0.84, a false positive rate of 0.28, precision of 0.83, recall of 0.84, and an F-measure of 0.83.

The machine learning classifier included 17 contraceptive attributes that are shown in Table 12.1 and organized by their source (whether from the ontology, manual annotation from Knowtator, or the NDF).

12.4 Discussion

Prior work has shown that NLP can be useful for unlocking free-text information in the medical record. Friedman's MedLEE system was originally applied to free-text radiology reports (Friedman 1997) but has demonstrated usefulness for other

Table 12.1 Annotations in the machine learning model by their source: ontology or Knowtator

Ontology of contraceptives	Knowtator annotation (schema)	VA National drug file
• Diaphragm	• Barrier method	• NDF
• Hormonal contraceptive—negated	• Barrier method—inactive	
• Hormonal contraceptive	• Contraception unspecified	
• Withdrawal	• Contraception unspecified—inactive	
• Withdrawal—negated	• Hormonal method	
	• Hormonal method—inactive	
	• Intrauterine device	
	• Intrauterine device—inactive	
	• No contraception—inactive	
	• No contraception—negated	
	• Permanent method	

In addition, any annotation from the VHA NDF gazetteer was given the type *NDF*

health services research in cardiology (Chiang et al. 2010), nursing (Hyun et al. 2009), disease-symptom associations (Wang et al. 2008) and de-identification of clinical progress notes (Morrison et al. 2009). Other NLP systems have been used in the clinical arena including Mayo's cTAKES (Savova et al. 2010), IBM's UIMA (Garvin et al. 2012), and the National Library of Medicine's (NLM) MetaMap (Meystre and Haug 2005). Minimal research however has focused on the clinical domain of contraceptive use among women. Our ontology-enriched NLP pipeline performed well in comparison to a manual review of VHA progress notes. In comparison, our prior work that did not include an ontology but included only our annotation schema resulted in a far more limited identification of contraceptive use (Womack et al. 2012).

Chapman et al. have identified barriers to using NLP for clinical research and decision support including: lack of access to shared data, lack of annotated datasets, insufficient conventions for annotation, reproducibility, limited collaboration, and user-centered development and scalability (Chapman et al. 2011). Our work addresses many of these barriers. In terms of shared data, through the VHA and VHA-funded research efforts such as WVCS and CHIR (Veterans Health 2007), we have access to data from multiple clinical sites. To address the lack of annotated datasets, we used an annotated corpora specifically targeted to contraceptive use to train and test our NLP-based classification model. The corpora are available to all VHA researchers interested in exploring contraceptive use through CHIR and WVCS. Also, the development of our contraceptive annotation schema addresses the barrier of *insufficient conventions for annotation* as the annotators systematically marked up the corpora. Our work relied on extensive collaboration with colleagues across the country (addressing *limited collaboration*) and used GATE plugins developed by our colleagues (*reproducibility*). Our work will also contribute to NLP communities worldwide via the GATE plugins we developed.

Our work demonstrates that information about contraceptive use can be extracted using a combination of biomedical informatics approaches including biomedical ontologies, annotation of relevant concepts from clinical progress notes and NLP for automated information extraction. This work is not only useful for researchers but also for clinicians and those interested in markers for quality of care in women's health.

12.4.1 Limitations

The authors recognize several limitations to this study. Since we only considered a subset of patients at the VHA, we cannot conclude that our results are generalizable to other healthcare institutions including other VHA facilities. In addition, we did not consider all aspects of the medical record including pharmacy notes that contain semi-structured data that might be useful for identifying contraceptives. Finally, we only considered two outcomes: *active* versus *inactive* use. Thus, we cannot address how useful our approach would be on more complex issues such as duration of use and co-occurrence of certain contraceptives.

12.5 Conclusion

Our combined approach utilized annotation of concepts, a biomedical ontology, and a natural language processing pipeline for information extraction. Our results achieved both high precision and recall and highlight the potential for biomedical informatics to support research of contraceptive use among female Veterans at the VHA. Additional research needs to be done that evaluates the accuracy of contraceptive information in the VHA's EHR with the consideration of both free-text and semi-structured data such as pharmacy records.

Acknowledgements This work was supported in part by a Veterans Affairs Health Services Research & Development (HSR&D) grant HIR 09-007 and is a translational use case project within the VA-funded Consortium for Healthcare Informatics Research (CHIR). In addition, this work is supported in part by VA grant DHI 07-065-1 to CB. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs. The authors would like to thank colleagues from the Tampa VA, specifically Dr. James McCart, Mr. Jay Jarman, and Dr. Stephen Luther for providing GATE plugins. The authors would also like to thank Ms. Harini Bathulapalli for her database work and Mr. Brett South for providing code to export Knowtator annotations. Finally, the authors would like to thank Dr. Jyotishman Pathak for his feedback on the project.

References

About.com. (2012). Contraception terms. [cited 2011 Dec 1]. http://contraception.about.com

BioPortal. 2010 [cited 2010 Dec 5]. http://bioportal.bioontology.org/

Brown SH et al (2004) VA National Drug File Reference Terminology: a cross-institutional content coverage study. Stud Health Technol Inform 107(Pt 1):477–481

Chapman WW et al (2001) A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 34(5):301–310

Chapman WW et al (2011) Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions. J Am Med Inform Assoc 18(5):540–543

Chiang JH, Lin JW, Yang CW (2010) Automated evaluation of electronic discharge notes to assess quality of care for cardiovascular diseases using Medical Language Extraction and Encoding System (MedLEE). J Am Med Inform Assoc 17(3):245–252

Coonan KM (2004) Medical informatics standards applicable to emergency department information systems: making sense of the jumble. Acad Emerg Med 11(11):1198–1205

Cote RA, Robboy S (1980) Progress in medical information management. Systematized nomenclature of medicine (SNOMED). JAMA 243(8):756–762

Denny JC et al (2012) Natural language processing improves identification of colorectal cancer testing in the electronic medical record. Med Decis Making 32(1):188–197

Drugs.com. (2012) Contraception drugs. [cited 2011 Dec 2]. http://www.drugs.com/condition/contraception.html

Friedman C (1997) Towards a comprehensive medical language processing system: methods and issues. Proc AMIA Annu Fall Symp 1997:595–599

Garvin JH et al (2012) Automated extraction of ejection fraction for quality measurement using regular expressions in Unstructured Information Management Architecture (UIMA) for heart failure. J Am Med Inform Assoc 19(5):859–866

GATE (2011) [cited 2012 Apr 16]. http://gate.ac.uk/

Haskell SG et al (2011) The burden of illness in the first year home: do male and female VA users differ in health conditions and healthcare utilization. Womens Health Issues 21(1):92–97

Hyun S, Johnson SB, Bakken S (2009) Exploring the ability of natural language processing to extract data from nursing narratives. Comput Inform Nurs 27(4):215–223; quiz 224–5

IHTSDO (2012) SNOMED. [cited 2012 May 10]. http://www.ihtsdo.org/snomed-ct/

Lee JH, Gonzalez GH (2011) Towards integrative gene prioritization in Alzheimer's disease. Pac Symp Biocomput 4:13

Meystre SM, Haug PJ (2005) Comparing natural language processing tools to extract medical problems from narrative text. AMIA Annu Symp Proc 2005:525–529

Morrison FP, Sengupta S, Hripcsak G (2009) Using a pipeline to improve de-identification performance. AMIA Annu Symp Proc 2009:447–451

NDF (2012). [cited 2012 May 10]. http://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/VANDF/attributes.html

NLM (2012) RxNorm. [cited 2012 May 10]. http://www.nlm.nih.gov/research/umls/rxnorm/

Noy NF et al (2003) Protege-2000: an open-source ontology-development and knowledge-acquisition environment. AMIA Annu Symp Proc 2003:953

Ogren PV (2006) Knowtator: a protégé plug-in for annotated corpus construction. Proceedings of the 2006 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, 2006

Ogren P (2007) Knowtator . [cited 2012 May 12]. http://knowtator.sourceforge.net/

Protege 4.1. 2011. [cited 2011 Dec 1]. http://protege.stanford.edu/download/protege/4.1/installanywhere/Web_Installers/

Rubin D et al (2010) Natural language processing for lines and devices in portable chest X-rays. AMIA Annu Symp Proc 2010:692–696

Savova GK et al (2010) Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. J Am Med Inform Assoc 17(5):507–513

Veterans Health Administration (2007) Program Announcement for Request for Concept Paper for Service Directed Research: Consortium for Healthcare Informatics Research (CHIR). [cited 2012 May 12]. http://www.research.va.gov/funding/solicitations/docs/Consortium-Healthcare-Informatics.pdf

Veterans Health Administration (2012) VINCI. [cited 2012 May 12]. http://www.hsrd.research.va.gov/for_researchers/vinci/

Wang X et al (2008) Automated knowledge acquisition from clinical narrative reports. AMIA Annu Symp Proc 2008:783–787

Warrer P et al (2012) Using text-mining techniques in electronic patient records to identify ADRs from medicine use. Br J Clin Pharmacol 73(5):674–684

Weka (2012). http://www.cs.waikato.ac.nz/ml/weka/

Whetzel PL et al (2011) BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic Acids Res 39:541–545

Womack J et al (2012) Analysis of contraceptive use among female veterans at the VA. AMIA Summit on Clinical Research Informatics, San Francisco

Yano EM et al (2006) Toward a VA Women's Health Research Agenda: setting evidence-based priorities to improve the health and health care of women veterans. J Gen Intern Med 21(Suppl 3):S93–S101

Index

© Springer International Publishing Switzerland 2016

A. Gupta et al. (eds.), *Advances in Healthcare Informatics And Analytics*,
Annals of Information Systems 19, DOI 10.1007/978-3-319-23294-2