

# Chapter 3

## Towards a Knowledge and Data-Driven Perspective in Medical Processes



Carlos Fernandez-Llatas and Mar Marcos

### 3.1 Introduction

Healthcare is one of the most challenging problems that our society is facing currently. The population of the world is growing around 1% per year [33]. In addition, the expectancy of life is increasing thanks to the new advances and the better quality of health solutions. For this reason, people are reaching older ages, which entails more chronic illnesses, with more co-morbidities. This supposes a great increase in the complexity of the illnesses. In addition, thanks to the new age of internet patients are more aware of their illnesses, having higher expectations of the health system. Altogether, this causes a great impact in the sustainability of healthcare, which should cover this scenario with the same budget. This juncture is demanding a new paradigm that will be able to deal with the complexity and continuous changes in the health domain in the coming years, in order to guarantee the sustainability of the system.

From the 1990s, when the Evidence-Based Medicine paradigm emerged [35], there has been an increasing interest in providing tools for empowering health professionals in the application of new methodologies and paradigms that could solve this problem. Sacket defined Evidence-Based Medicine as the “*conscientious, explicit, and judicious use of current best evidence in making decisions about*

---

C. Fernandez-Llatas (✉)

Process Mining 4 Health Lab – SABIEN – ITACA Institute, Universitat Politècnica de València, Valencia, Spain

CLINTEC – Karolinska Institutet, Sweden

e-mail: [cfllatas@itaca.upv.es](mailto:cfllatas@itaca.upv.es)

M. Marcos

Department of Computer Engineering and Science, Universitat Jaume I, Castelló de la Plana, Castelló, Spain

e-mail: [mar.marcos@uji.es](mailto:mar.marcos@uji.es)

*the care of individual patients”* [35]. This paradigm tried to unify knowledge gathered from the best research evidence (what the literature says), with the clinical experience (what the clinician knows), centered in providing the best experience to the patient (what the patient wants). In this scenario, the idea of creating formalized processes that support the daily clinical practice with the best evidence available arises. In this scenario, the idea of creating formalized processes (or protocols) that support the daily clinical practice with the best evidence available arises. It promotes the formalization of clinical research results so that they can be applied in daily practice by clinical professionals. This is done through the specification of protocols that are thought to be well-defined standards of care. In this line, these protocols can serve to improve the clinical effectiveness, provide solutions for risk management, and trace the actual care process to reduce the variability of the treatments in healthcare.

In the literature there are different approaches for the definition of such formalized processes. The aim of this chapter is to analyse the most prominent approaches for supporting clinical experts in the representation of medical processes. First, the two process-related perspectives in healthcare (which we have named as patient & process centered and clinician & knowledge centered, respectively) are presented and compared. Moreover, the different instruments developed by the medical profession related to this concept are described. After that, the two main approaches available in the literature for building medical processes are reviewed: on one hand, knowledge-driven Clinical Decision-Making technologies, and, on the other hand, Clinical Process Management technologies, which rely on a data-driven approach. Finally, the concluding section discusses new challenges towards the formalization of medical processes leveraging the advantages of these two technologies.

### **3.2 Process-Related Perspectives in Healthcare**

Health systems are struggling to meet the growing demand for healthcare services from ageing population while maintaining consistent quality standards. As Peleg and González-Ferrer point out, two strategies are being used for this purpose, both sharing a process-based perspective [30]. One strategy focuses on improving the management of the processes (e.g. interventions, interactions) that the patient goes through in relation to a clinical encounter. The other strategy concentrates on supporting decision making by the clinician at the point of care using specific-purpose tools (i.e. dedicated to a specific medical condition) that incorporate knowledge about clinical processes. This knowledge is mostly based on the best evidence available that can be found in documents such as clinical practice guidelines, but can also refer to medical background knowledge contained in textbooks and manuals. The former strategy takes the perspective of the patient journey and considers organizational issues of healthcare processes, including the coordination of multidisciplinary teams and the allocation of resources, and thus can be described as *patient & process centered*. In contrast, the latter strategy focuses on the perspective of the clinician when managing an individual patient, with an

**Table 3.1** Summary of the features of the *clinician & knowledge centered* (CKC) and *patient & process centered* (PPC) strategies in healthcare

Features \ Strategies	CKC	PPC
<i>Main subject</i>	Health professional	Patient
<i>Main process</i>	Health professional's actions and decisions	Patient journey
<i>Patients</i>	Individual patient	Multiple patients
<i>Health professionals</i>	Individual health professional	Multidisciplinary health team
<i>Diseases</i>	Single disease	Single or multi-disease
<i>Main usage</i>	Prescriptive	Analytical
<i>Orientation</i>	Knowledge-driven	Knowledge or data-driven

emphasis on knowledge-intensive decision tasks, therefore it might be considered as *clinician & knowledge centered*.

There exist significant differences in how the previous strategies can be exploited in the healthcare context. Most notably, the clinician & knowledge centered (CKC) strategy, relying on the recommendations issued by medical experts, can be used to determine *what should be performed* (or *what is prescribed*) given the specific clinical circumstances of a patient, typically in the context of a single disease. On the other hand, the patient & process centered (PPC) strategy can be applied to inspect *what has been performed* and makes it possible e.g. to monitor the itinerary (or itineraries) actually followed by patients with a particular clinical profile, possibly involving multiple diseases. In other words, the usage of the CKC strategy would be primarily prescriptive, whereas that of the PPC one would be analytical. Lastly, the two strategies may differ in their positioning with respect to knowledge and data. Although a knowledge-driven orientation can be taken in both cases, in the case of the PPC strategy the use of process models obtained from clinical data in the Electronic Health Record, i.e. a data-driven orientation, is a common practice. Table 3.1 summarises the main characteristics (and differences) of these strategies.

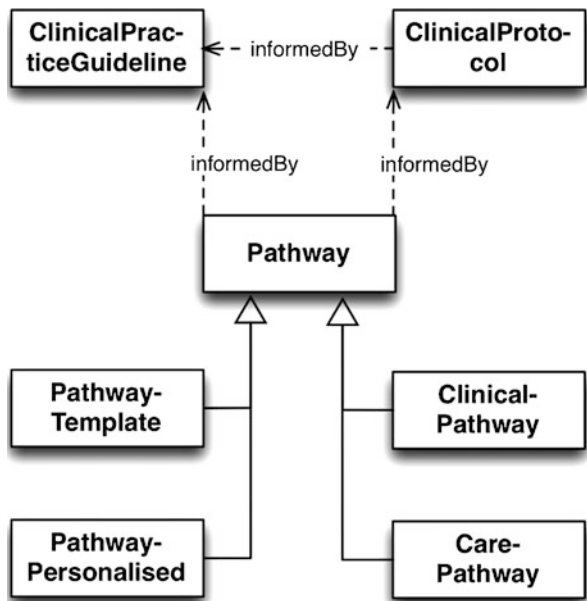
The instruments developed by the medical profession to support the concept of consistent and high-quality healthcare are very much related to what has been exposed. *Clinical Practice Guidelines* (CPGs) are the core instrument. According to the most recent definition, CPGs are defined as “*statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options*” [15]. In line with the view of Evidence-Based Medicine, the development of CPGs is usually commissioned to a group of experts who are responsible for collecting and analysing the best and most up-to-date evidence about a particular clinical condition, and for agreeing a set of general recommendations regarding the main management aspects thereof.

*Clinical Protocols* are related yet distinct from CPGs. A clinical protocol is a locally agreed statement about a specific clinical issue with steps based on CPGs and/or organizational consensus [2]. Usually, clinical protocols are specific to a

health organization. *Care Pathways* likewise adapt CPG recommendations to the needs and particularities of a health organization. Thus, both clinical protocols and care pathways can be regarded as instruments for the implementation at local level of the evidence base from CPGs. However, care pathways differ in that they describe many more aspects (and in more detail), including: an explicit statement of the goals and key elements of care when solving one or several clinical issues, the description of the communication among the care team members and with patients and their families, and the specification of the coordination aspects of the care process (with roles, sequencing of decisions and actions, etc.) [36]. Care pathways also define the information to be recorded so that it is possible to monitor deviations of the actual care with respect to the recommended procedure. *Clinical Pathways (CPs)* in turn differ from care pathways in that they are confined to the paths within a hospital, i.e. excluding outpatient clinic and follow-up activities. Common to most of the concepts, it is possible to distinguish the general instrument (template) from the versions adapted to the values and preferences of the patient, giving rise e.g. to “personalised care pathways”. Figure 3.1 depicts the relationships among these concepts.

As explained before, the CKC strategy strongly relies on knowledge about clinical processes and decisions. Most typically, CPGs are used as source for such knowledge. For their part, CPs (and pathways in general) are very well suited for the purposes of the PPC strategy, due to their focus on the monitoring of care processes. Naturally, CPGs have a knowledge-driven orientation, whereas either a knowledge-driven or a data-driven orientation can be adopted for CPs.

**Fig. 3.1** Relationships between clinical practice guidelines, clinical protocols, and pathways, including pathway variations. (Adapted from Figure 1 in Benson’s article [2])



### 3.3 Technologies for Clinical Decision-Making

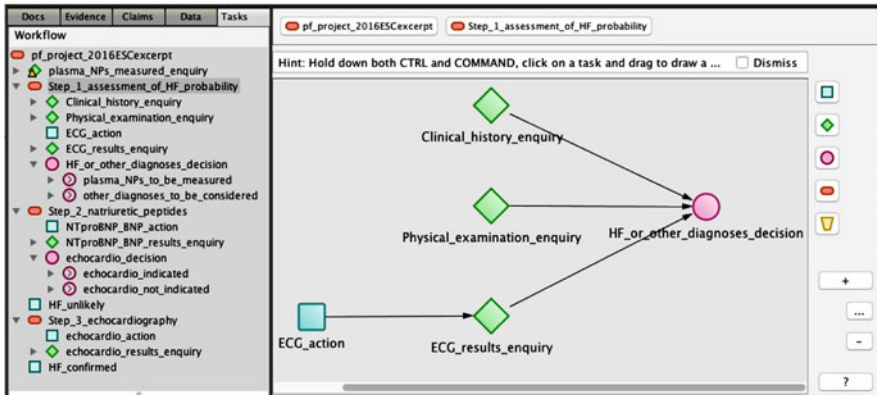
#### 3.3.1 Computer-Interpretable Guidelines

CPGs have shown the potential to foster the translation of clinical research results into practice, and to improve the quality and outcomes of healthcare. However, the practical utility of CPGs is often hindered by the text-based format in which they are predominantly disseminated. Another problem is their emphasis on the general principles of care, rather than on the actual processes along the patient journey [30]. In this context, *Computer-Interpretable Guidelines (CIGs)* emerge as a tool to make patient-customized CPG recommendations available to clinicians in an easier and more immediate way, compared to text-based CPGs. Thus, CIGs can be defined as formalized versions of CPG contents intended to be used as decision-support systems. The beneficial effects of the use of CIGs in the clinical setting have been documented in the literature, and include improved CPG adherence and increased efficiency of the healthcare processes (e.g. thanks to the reduction of unnecessary test requests) [20].

Several CIG representation languages have been proposed in the fields of Artificial Intelligence in Medicine and Medical Informatics, the most prominent of which are Arden Syntax, PROforma, Asbru, EON, GLIF, and GUIDE [5, 31]. CPGs contain a wealth of knowledge of diverse types. To accommodate this variety, CIG languages provide a wide range of modelling constructs. Peleg et al. recognize two main representational categories, namely structuring in plans of decisions and actions, and linking to patient data and medical concepts, and identify a total of eight dimensions within them [31]. These dimensions are: (1) organization of plans, (2) goals, (3) action model, (4) decision model, (5) expression language, (6) data interpretation/abstraction, (7) medical concept model, and (8) patient model.

Many of the CIG languages take an approach to the description of plans (above dimensions (1) through (4)) which has been named *Task-Network Model (TNM)*. The TNM approach consists in describing guidelines in terms of a hierarchical decomposition of networks of component tasks. The task types, as well as the types of control-flow constructs (sequence, in parallel, etc.), vary in the different TNM approaches. Still, all of them provide support for actions, decisions and nested tasks. A highly distinctive feature of CIG languages lies in the decision model. In this regard, PROforma's decision model, which was subsequently adopted by other CIG languages, deserves a special mention. In PROforma, decisions are described in terms of the alternative options (or candidates) considered, each one with an associated set of arguments. These arguments are logical conditions that, when they are met, provide different kinds of support for the candidate, namely for, against, confirming or excluding the candidate.

As an illustration, Fig. 3.2 shows a PROforma excerpt corresponding to the algorithm for the diagnosis of heart failure in the non-acute setting [23], based on the 2016 guidelines of the European Society of Cardiology. It comes as no surprise that, although CIG languages were specifically geared for CPGs, they have also



**Fig. 3.2** PROforma model for the diagnosis of heart failure in the non-acute setting [23]. To the left, task tree representing the hierarchical decomposition of tasks; to the right, task network corresponding to the plan *Step\_1\_assessment\_of\_HF\_probability* (first step of diagnosis), including action, enquiry and decision tasks

proven to be useful for modelling and supporting complex clinical processes from a broad spectrum. This includes care pathways for different purposes, e.g. for the management of triple assessment in breast cancer [28] or for the monitoring of patients with multiple comorbidities [21].

### 3.3.2 *Development and Maintenance Issues with Computer-Interpretable Guidelines*

The representational richness of CIG languages makes them difficult to use (not to mention mastering them) for non-technical users like clinicians. Furthermore, it is well recognized that CPG knowledge is intrinsically complex and hence difficult to comprehend and formalize [19]. As a consequence of these factors, the encoding of CPG knowledge in a CIG language is a difficult and labour-intensive task which requires the joint collaboration of both clinical and IT professionals. On the one hand, clinical expertise is essential for a complete and adequate understanding of CPG recommendations. On the other hand, IT skills are required to analyse the clinical processes they include, as well as to shape them in terms of the constructs of the CIG language chosen [25]. This explains why the topic of CIG knowledge acquisition and specification has been the focus of a large number of research works in the literature. Concretely, in relation with the life-cycle of CIG development, knowledge acquisition and specification is the topic to which more efforts have been devoted, after CIG languages [29]. Noteworthy among these approaches are the application of cognitive methodologies to guide the encoding of CPGs into CIGs, and the use of pattern-based information extraction methods to support the

translation of CPG texts into a semi-structured format. Despite these efforts, CIG development tasks remain largely manual. This may lead to a significant delay between the time when a CPG is issued and the time when a fully functional and validated CIG is ready for its implementation, which could be unacceptable from a clinical perspective.

Once implemented, CIGs necessitate some kind of quality control to determine whether the impact they have on the healthcare processes is as expected. The aim of clinical decision support systems in general, and of CIGs in particular, is to improve the quality, safety and cost-effectiveness of care processes. A monitoring of evidence-grounded quality metrics, together with an appropriate feedback to the health organization, can serve as a stimulus for process improvement [8]. Quality metrics provide a framework for comparison, e.g. to detect outlier cases in which CPG recommendations have not been followed. Such cases may point to procedure parts where modifications should be considered. CIG compliance analysis has been the topic of a number of works in the literature [29]. There are two types of approaches for evaluating compliance with CIGs: approaches directly comparing the concrete actions performed by the physician, and those comparing the actual processes discovered from clinical activity logs using Process Mining methods. The former range from informal (manual) methods to more formal methods based e.g. on model checking. The approaches based on Process Mining methods have recently attracted growing interest because of their potential to recognize variations with respect to the prescriptive process embodied in CIGs.

## 3.4 Technologies for Clinical Process Management

### 3.4.1 *Process Discovery and Continuous Improvement*

Due to the difficulties of the manual development process of CPGs, data-driven approaches have emerged in the literature for supporting health experts in the definition of guidelines. Data-driven models use the data available in healthcare databases to infer the underlying processes and thereby provide Decision Support Systems without the need for a purely manual development by clinical experts. The idea is to develop algorithms that discover automatically such underlying processes.

Data-driven solutions have been used successfully for the automatic learning of models that can support experts in different fields. With this aim, different approaches within this paradigm have been applied to the medicine domain. One of the most common approach is the creation of classifiers for supporting in the daily decisions [40]. However, these tools do not provide a process view. These systems only provide a statistical probability of the current status of the patient at a certain moment in time. Other options, like Temporal Abstractions [4], offer a vision about the trends in the biomedical signals that enable a dynamic measure of the patient status. However, although these techniques can be incorporated in the CKC or PPC

strategies, they do not allow to discover the rules behind the progression of the disease in the patient.

Other works have tried to discover the behaviour of the medical processes by inferring their inherent rules using information routinely collected in healthcare databases [6]. However, from a process management perspective, these rules do not provide the natural view of the process as it is provided by workflows. This fact has a negative impact on the understandability of these systems, which as a consequence usually appear as *black boxes* in the eyes of medical practitioners.

Process Mining appears in the middle of this juncture [38]. This paradigm uses time-stamped events existing in healthcare databases to offer a workflow-based view. Using these techniques, there are works providing tools to infer the underlying medical process, offering partial patterns [18], general patterns avoiding infrequent behavior [17], or complete views of CP [11, 39]. But Process Mining is not only about the discovery of processes. Process Mining aims to provide a complete set of technologies for supporting medical professionals not only in the process design phase but also in the traceability, analysis and optimization of the process deployed. Health systems produce continuous data flow that can be used for analyzing how the processes behave in actual scenarios [9]. With that, it is possible not only to show a snapshot of the pathway, but also to make a comparison over time to discover any variation of the medical procedures, e.g. due to the application of new protocols [3, 32].

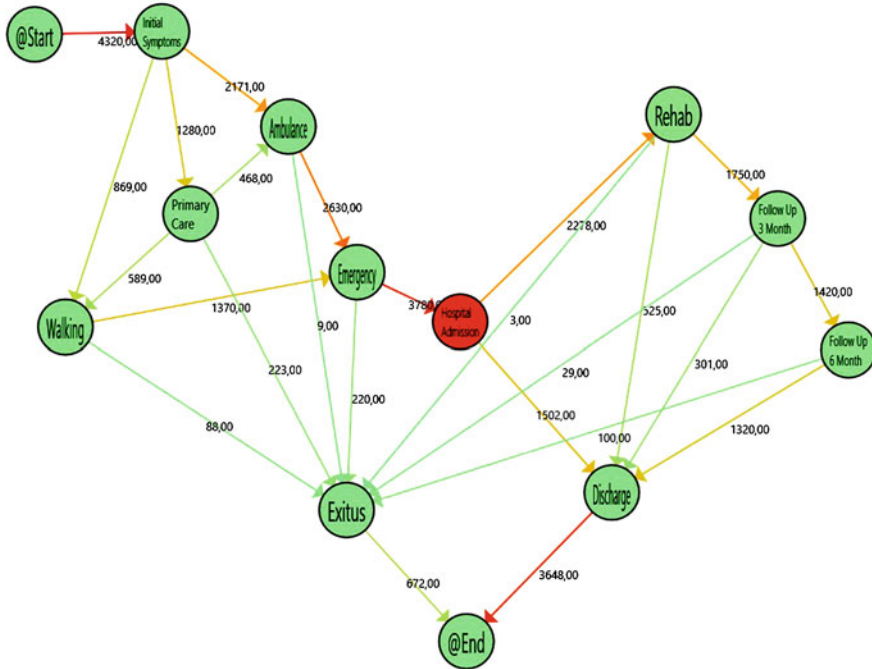
Figure 3.3 shows an example of how Process Mining can represent the processes inferred from data available in medical databases. This process represents the flow of patients in a surgery area, and was automatically inferred from real data in existing databases [9]. The model not only represents the flow of the process but, also, colours represent the performance in their execution. This information can be crucial for a better understanding of how processes are deployed in a real scenario.

These technologies can be applied in an iterative way allowing for a continuous optimization of the process [12]. This allows the user not only an easier design of the process, but also an iterative adaption of the process that converges to the best optimized solution.

### 3.4.2 Workflow Inference Models

For achieving an adequate process standardization, algorithms should provide formal models that can be used for standardize the care. In this line, The Business Process Management (BPM) field [7] aims to offer solutions for supporting the creation of those processes in a general-purpose way. In this way, the concept of *Workflow* is proposed. The Workflow Management Coalition defines it as “*the computerised facilitation or automation of a business process, in whole or part*” [16]. In other words, a Workflow is a formal specification designed to automate a process. Process Mining provides tools for building Workflows from events existing in medical databases.





**Fig. 3.3** Process inferred using a Process Mining Discovery algorithm. The colours of the nodes represent the average Length of Stay in each one of the stages, and the colours in the arcs represent the percentage of patients that follows each path

Workflows are designed to deal with process standardization via the definition of graphical structures, without ambiguities and focused on their automation, with a view towards automatic guidance by computer systems or replication by human experts. Workflows are devised for supporting the design of a process that: (1) needs high-level legibility, intended to be understood by human experts, not only for its creation but also for its optimization; (2) requires traceability, to make possible a continuous analysis of the current status of the process flow; and (3) guides users over a set of steps, allowing them to know the trace of the process flow to the current status and showing the possibilities after it.

BPM techniques have been tested in the medical field for representing clinical workflows. In this line, there are some works in the literature dealing with different Workflow models. Some works use well known mathematical representation languages like: Petri Nets [22, 32]; Deterministic Finite Automaton (DFA) or graphs [26]; other more specific formal mathematical models like Timed Parallel Automaton (TPA) [10]; and other models specifically created for increasing the understandability like the Business Process Management Notation (BPMN) [27].

To select an adequate Workflow Language for each problem it is necessary to analyze its characteristics. In healthcare, three characteristics should be taken into account [10]:

- **Expressivity.** It is the capability for representing all the dynamic flow behaviours of the processes in the field. This characteristic can be measured thanks to the so-called *Workflow Patterns* [34]. Workflow patterns are different situations that are possible in a process (sequences, parallelism, milestones, etc.). The more workflow patterns a language can express, the best expressivity it has. The objective is to have the best language able to express all the possible patterns in each medical field. Otherwise, the lack of expressivity can result in inaccurate and ambiguous models.
- **Understandability.** It represents how easy to read and understand is a language. Clinicians are not IT engineers and not all the languages are suitable for them to comprehend. The lack of understandability increases the risk of rejection by the clinician. Even worst, it can result in inaccuracies and errors in the models' design that could not only make the system to fail but also lead to an inappropriate recommendation for the patient.
- **Complexity.** The complexity of a language is related to the quantity of information that it conveys and how computers can process it. This complexity depends on the grammar on which the language is based. The complexity of a grammar has an influence on its interaction with computers. It is well known in compilation theory that the more complex a language of a specific grammar is, the more difficult it is to process and interpret it [1]. A complex language is harder to execute and to infer using data-driven techniques. So, the more complex a language is, the more difficult it is to create applications to understand it and to create data-driven accurate models without the use of heuristics.

Selecting the best language for a specific field requires a trade-off between having the desired expressivity, while maximizing the understandability and keeping the least possible complexity. For example, Petri Nets are probably the most expressive language for representing any kind of process, but the difficulty for clinicians to understand it and its complexity makes it necessary to use heuristics for the inference algorithm. On the other hand, DFA has a very low complexity and this allows an easy interpretation and very powerful techniques for inference. It is also easy to understand, however its expressivity is very limited. There are mixed solutions like TPAs, which are expressive as Safe Petri Nets and low complexity as DFAs. Other specific languages, like BPMN, have been specifically created so that the users can understand them, using graphical metaphors for adding semantics to workflows in a human-understandable way. However, BPMN has a higher complexity to be executed. In this line, all the characteristics of the problem to be solved should be evaluated, for selecting an adequate solution in each case.

### 3.5 Challenges of Clinical Decision-Making and Process Management Technologies

Despite the hard work in research to leverage the advantages of clinical decision-making and process management technologies in the medical domain, there is still a big gap between the possibilities of these technologies and their joint application in real scenarios.

On one hand, from the perspective of clinicians, there is a need for creating safe, non-intrusive, adaptive and trusted tools that offer the confidence required for their implementation in real scenarios. CPGs and CIGs can offer solutions in this direction, but the problems associated to their development process, requiring a high consensus of experts and the need for a continuous revision, make the judicious and successful use of these instruments a challenge. On the other hand, from a cognitive and perceptual computing perspective [37], there is a need for more data as well as for better self-adaptive algorithms to provide a holistic approach from a data-driven point of view [39]. However, this would imply reducing to a minimum human intervention during the process of model creation and adaption. But, is there a real need for completely self-adaptive tools? Why exclude the human in the process of automatic learning? In the Interactive Pattern Recognition paradigm, it has been demonstrated that the involvement of the expert in the loop not only provides better and quicker results than classical Data-Driven approaches, but also ensures a better understanding and improved confidence in each iteration [12].

The research community of the fields of CIGs and process mining for health is claiming for the combination of these technologies, taking advantage of the best of two worlds [14]. The importance of combining clinical decision-support and workflow technologies to provide realistic support for complex processes, like extended care pathways and multidisciplinary care, was identified more than a decade ago [13]. Beyond that, the need for human expert participation is key for ensuring the adequacy of the models inferred by data-driven approaches. With this assistance experts could incorporate their background knowledge in the model, e.g. to correct possible inference errors. This would allow the creation of models, Workflow or CIG ones, of a better quality and error-free. Furthermore, in process mining, the implementation of solutions to the problem of data denoising would be more effective in an interactive way [24].

The involvement of the expert in the loop using an interactive paradigm mixing data and knowledge-driven solutions opens a set of new perspectives with huge potential. On one hand, process mining approaches could greatly benefit from knowledge intensive models such as CIGs, e.g. using them as a layer for the purpose of improving the explainability of their models to clinicians. On the other hand, the application of interactive process mining methods could play a role of paramount importance in the development and continuous adaption of CPGs (and CIGs), e.g. enabling the integration of tried-and-tested procedures inferred from healthcare data as a complementary source of knowledge in addition to evidence-based and expert knowledge. The main challenge will be how to articulate the design of tools so that a

perfect integration of these two technologies can be achieved and, at the same time, their respective benefits can be leveraged to improve both the quality standards and the management aspects of healthcare processes.

## References

1. Aho AV, Sethi R, Ullman JD. *Compilers, principles, techniques*. Addison Wesley. 1986;7(8):9.
2. Benson T. *Care pathways*. Technical report, NHS National Programme for Information Technology (NPfIT), 2005.
3. Conca T, Saint-Pierre C, Herskovic V, Sepúlveda M, Capurro D, Prieto F, Fernandez-Llatas C. Multidisciplinary collaboration in the treatment of patients with type 2 diabetes in primary care: analysis using process mining. *J Med Internet Res*. 2018;20(4).
4. Concaro S, Sacchi L, Cerra C, Fratino P, Bellazzi R. Mining healthcare data with temporal association rules: improvements and assessment for a practical use. In: *Conference on artificial intelligence in medicine in Europe*. Springer; 2009. p. 16–25.
5. de Clercq PA, Blom JA, Korsten HHM, Hasman A. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artif Intell Med*. 2004;31(1):1–27.
6. Du G, Jiang Z, Diao X, Yao Y. Knowledge extraction algorithm for variances handling of cp using integrated hybrid genetic double multi-group cooperative pso and dpso. *J Med Syst*. 2012;36(2):979–94.
7. Dumas M, La Rosa M, Mendling J, Reijers HA, et al. *Fundamentals of business process management*, vol. 1. Springer; 2013.
8. Eisenberg F. Chapter 4 – the role of quality measurement and reporting feedback as a driver for care improvement. In: Greenes RA, editor. *Clinical decision support. The road to broad adoption*. 2nd ed. Oxford: Academic; 2014. p. 145–64.
9. Fernandez-Llatas C, Lizondo A, Monton E, Benedi J-M, Traver V. Process mining methodology for health process tracking using real-time indoor location systems. *Sensors*. 2015;15(12):29821–40.
10. Fernandez-Llatas C, Pileggi SF, Traver V, Benedi JM. Timed parallel automaton: a mathematical tool for defining highly expressive formal workflows. In: *2011 Fifth Asia modelling symposium*. IEEE; 2011. p. 56–61.
11. Fernandez-Llatas C, Valdivieso B, Traver V, Benedi JM. Using process mining for automatic support of clinical pathways design. In: *Data mining in clinical medicine*. Springer; 2015. p. 79–88.
12. Fernández-Llatas C, Meneu T, Traver V, Benedi J-M. Applying evidence-based medicine in telehealth: an interactive pattern recognition approximation. *Int J Environ Res Public Health*. 2013;10(11):5671–82.
13. Fox J, Black E, Chronakis I, Dunlop R, Patkar V, South M, Thomson R. From guidelines to careflows: modelling and supporting complex clinical processes. *Stud Health Technol Inform*. 2008;139:44–62.
14. Gatta R, Vallati M, Fernandez-Llatas C, Martinez-Millana A, Orini S, Sacchi L, Lenkowicz J, Marcos M, Munoz-Gama J, Cuendet M, et al. Clinical guidelines: a crossroad of many research areas. Challenges and opportunities in process mining for healthcare. In: *International conference on business process management*. Springer; 2019. p. 545–56.
15. Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E. *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press; 2011.
16. Hollingsworth D, Hampshire UK. *Workflow management coalition: the workflow reference model*. Workflow Management Coalition. Document Number TC00-1003. 1995;19:16.
17. Huang Z, Lu X, Duan H. On mining clinical pathway patterns from medical behaviors. *Artif Intell Med*. 2012;56(1):35–50.

18. Huang Z, Lu X, Duan H, Fan W. Summarizing clinical pathways from event logs. *J Biomed Inform.* 2013;46(1):111–27.
19. Kaiser K, Marcos M. Leveraging workflow control patterns in the domain of clinical practice guidelines. *BMC Med Inform Decis Mak.* 2016;16:20.
20. Latoszek-Berendsen A, Tange H, van den Herik HJ, Hasman A. From clinical practice guidelines to computer-interpretable guidelines. A literature overview. *Methods Inf Med.* 2010;49(6):550–70.
21. Lozano E, Marcos M, Martínez-Salvador B, Alonso A, Alonso JR. Experiences in the development of electronic care plans for the management of comorbidities. In: Riaño D, Teije A, Miksch S, Peleg M, editors. *Knowledge representation for health-care. Data, processes and guidelines.* Berlin/Heidelberg: Springer; 2010. p. 113–23.
22. Mahulea C, Mahulea L, García-Soriano J-M, Colom J-M. Petri nets with resources for modeling primary healthcare systems. In: 2014 18th International conference on system theory, control and computing (ICSTCC). IEEE; 2014. p. 639–44.
23. Marcos M, Campos C, Martínez-Salvador B. A practical exercise on re-engineering clinical guideline models using different representation languages. In: Marcos M, Juarez JM, Lenz R, Nalepa GJ, Nowaczyk S, Peleg M, Stefanowski J, Stiglic G, editors. *Artificial intelligence in medicine: knowledge representation and transparent and explainable systems.* Springer International Publishing; 2019.
24. Martin N, Martinez-Millana A, Valdivieso B, Fernández-Llatas C. Interactive data cleaning for process mining: a case study of an outpatient clinic’s appointment system. In: International conference on business process management. Springer; 2019. p. 532–44.
25. Martínez-Salvador B, Marcos M. Supporting the refinement of clinical process models to computer-interpretable guideline models. *Bus Inform Syst Eng.* 2016;58(5):355–66.
26. Mesner O, Davis A, Casman E, Simhan H, Shalizi C, Keenan-Devlin L, Borders A, Krishnamurti T. Using graph learning to understand adverse pregnancy outcomes and stress pathways. *PLoS One.* 2019;14(9):e0223319.
27. Müller R, Rogge-Solti A. Bpmn for healthcare processes. In: *Proceedings of the 3rd central-European workshop on services and their composition (ZEUS 2011), Karlsruhe, vol. 1, 2011.*
28. Patkar V, Fox J. Clinical guidelines and care pathways: a case study applying proforma decision support technology to the breast cancer care pathway. *Stud Health Technol Inform.* 2008;139:233–42.
29. Peleg M. Computer-interpretable clinical guidelines: a methodological review. *J Biomed Inform.* 2013;46(4):744–63.
30. Peleg M, González-Ferrer A. Chapter 16 – guidelines and workflow models. In: Greenes RA, editor. *Clinical decision support. The road to broad adoption.* 2nd ed. Oxford: Academic; 2014. p. 435–64.
31. Peleg M, Tu S, Bury J, Ciccarese P, Fox J, Greenes RA, Hall R, Johnson PD, Jones N, Kumar A, Silvia Miksch, Quaglini S, Seyfang A, Shortliffe EH, Stefanelli M. Comparing computer-interpretable guideline models: a case-study approach. *J Am Med Inform Assoc.* 2003;10(1):52–68.
32. Rebuge Á, Ferreira DR. Business process analysis in healthcare environments: a methodology based on process mining. *Inf Syst.* 2012;37(2):99–116.
33. Roser M, Ritchie H, Ortiz-Ospina E. *World population growth.* Our world in data, 2013.
34. Russell N, Van Der Aalst W, Ter Hofstede A. *Workflow patterns: the definitive guide.* MIT Press, 2016. <https://ieeexplore.ieee.org/book/7453725>.
35. Sackett DL, Rosenberg WMC, Gray MJA, Haynes BR, Richardson SW. Evidence based medicine: what it is and what it isn’t. *BMJ.* 1996;312(7023):71–2.
36. Schrijvers G, van Hoorn A, Huiskes N. The care pathway: concepts and theories: an introduction. *Int J Integr Care.* 2012;12(Spec Ed Integrated Care Pathways):e192.
37. Sheth A. Internet of things to smart iot through semantic, cognitive, and perceptual computing. *IEEE Intell Syst.* 2016;31(2):108–12.
38. Van Der Aalst W. *Process mining. Data science in action.* Springer; 2016.

39. Yang W, Su Q. Process mining for clinical pathway: literature review and future directions. In: 2014 11th international conference on service systems and service management (ICSSSM). IEEE; 2014. p. 1–5.
40. Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang J-F, Hua L. Data mining in healthcare and biomedicine: a survey of the literature. *J Med Syst*. 2012;36(4):2431–48.