Supporting physicians and patients through recommendation: Guidelines and beyond

Anselma, Luca; Bottrighi, Alessio; Hommersom, Arjen; Terenziani, Paolo; Hunter, Anthony. Supporting physicians and patients through recommendation: Guidelines and beyond. Springer Verlag. 2015. pp: 281-286.

in

Foundations of Biomedical Knowledge Representation

Chapter 18
Supporting Physicians through Recommendation: Guidelines and Beyond

Luca Anselma, Alessio Bottrighi, Arjen Hommersom, Paolo Terenziani, and Anthony Hunter

18.1 Clinical Practice Guidelines

The recommendation task, intended as the task of supporting physicians in their activity (and, in particular, in decision making) by providing them indications of the most appropriate way of treating patients, has a long story in Medical Informatics that dates back, for instance, to the first medical expert systems (MYCIN [19]). Many different tools and techniques have been devised, within the Medical Informatics area, in order to provide physicians with recommendations about the most appropriate treatment of patients. Recently, Clinical Practice Guidelines (CPG) have gained a major role in this context. CPGs are, in the definition of the USA Institute of Medicine, 'systematically developed statements to assist practitioner and patient decisions about appropriate health care in specific clinical circumstances' (Institute of Medicine, 2001, p. 151). They are conceived as a way of putting Evidence-Based Medicine into practice, as well as a mean to grant both the quality and the standardization of healthcare services, and the minimization of costs. Thousands of CPGs have been devised in the last years. For instance, the Guideline International Network (http://www.g-i-n.net) groups 77 organizations of 4 continents, and provides a library of more than 5000 CPGs. CPGs aim to reduce errors, unjustified practice variation and wasteful commitment of resources, and encourage best practices and accountability in medicine. Clinical guidelines are typically created by medical experts or panels convened by specialty organizations, who review the relevant studies, perform meta-analysis by contrasting and combining results from different studies and, using a consensus-based process, compile a set of evidence-based recommendations. Their focus may be on screening, diagnosis, management, treatment, or referral of patients with specific clinical conditions. The recommendations are typically written as narrative text and tables, which point back to background material and evidence, ranking the strength of clinical validity, and the strength with which recommendations should be followed according to the guideline authors.

The adoption of computerized approaches to acquire, represent, execute and reason with CPGs can further increase the advantages of CPGs, providing crucial advantages to:

- patients, granting them that they will receive the best quality medical treatments (since CPGs are actually a way of putting EBM into practice);
- physicians, providing them with a standard reference which they may consult, with a way of certifying the quality of their activity (e.g., for insurance or legal purposes), as well as with advanced support to their decision-making activity;
- hospitals and health-care centers, providing them with tools to grant the quality and the standardization of their services, as well as with a means to evaluate quality, and to optimize costs and resources.

However, the main purpose of CPGs is to support physicians in their everyday knowledge-based decision making when treating patients, providing them evidence-based recommendations at the point of care.

Unfortunately, there are several obstacles for a full exploitation of CPGs in the clinical practice. For instance, since CPGs are usually written as standard text in natural language, they tend to be quite long, so that it is difficult for the physician at the point of care to find out the specific part of the guideline that is relevant for the specific patient at hand. Additionally, natural language is inherently ambiguous, so that textually written CPGs are usually not "rigorous" and "formal" enough, possibly leading to un-correct interpretations of physicians using them. Last, but not least, one of the main goals of CPGs is to capture medical evidence. However, from one side, evidence is essentially a form of statistical knowledge, capturing the generalities of classes of patients, rather than the peculiarities of a specific patient. From the other side, demanding to expert committees to characterize all possible executions of a CPG on any possible specific patient in any possible clinical condition is an unfeasible task. Thus, CPGs assume to deal with ideal patients, i.e., patients that have just the single disease considered in the CPG (thus excluding the concurrent application of more than one CPG), and are "statistically relevant" (they model the typical patient affected by the given disease), not presenting rare peculiarities/side-effects. Also CPGs assume to operate in ideal context of execution, so that all necessary resources are available. Unfortunately, however, not all patients and execution contexts are "ideal" (in the above sense). As a consequence, there is always a gap between the generality of CPGs and the specificities of their execution on a specific patient in a specific context. Fulfilling such a gap is a difficult and challenging task, which is usually completely demanded to user physicians.

18.2 Computer Interpretable Guidelines

In the last two decades, Computer Interpretable Guidelines (CIGs) have been introduced in order to overcome some the above problems, and different formalisms and software systems have been developed to support them. CIG formalisms are usually

based on a Task-Network Model (TNM): a (hierarchical) model of the guideline control flow as a network of specific tasks. Such formalisms are "formal" and allow one to unambiguously represent guideline procedures and recommendations. Besides supporting formal languages to acquire and represent CPGs, CIG systems usally also provide execution engines that allow user physicians to "instantiate" general guidelines on specific patients: by accessing the patient clinical data, the execution engine shows to the user physicians only those paths of actions that are applicable to the patient at hand. In such a way, they provide patient-oriented recommendations to physicians, allowing them to fulfill the gap between the generality of the CPG and the specificity of the patient at hand. Given such advantages, many CIG formalisms and systems have been designed/built in the last two decades. Some of them (the list is in alphabetic order, and is far from being exhaustive) are: Asbru [15], EON [9], GEM [18], GLARE [22], GLIF [12], GUIDE [13], PRODIGY [6], PROforma [2], SAGE [23].

A survey and/or a comparative analysis of these systems is outside the goals of this chapter. A comparison of Asbru, EON, GLIF, Guide, PROforma, PRODIGY can be found in [10]. The recent book by Ten Teije [21] represents a consensus of a large part of the computer-oriented CPG community. It presents an assessment of the state of the art, as well as a collection of several recent approaches. Comprehensive surveys of the state-of-the-art about CIG have been already published [1, 3, 4, 5, 7, 24].

The surveys show that a relative consensus has been achieved concerning the representation formalisms. Although there are notable differences among the different approaches, partly due to the different goals they pursue, some important commonalities have been reached. For example, most approaches model guidelines in terms of a Task-Network Model (TNM): a (hierarchical) model of the guideline control flow as a network of specific tasks. Although the terminology may differ, all approaches support a basic set of core guideline tasks, such as decisions, actions and entry criteria. Decisions for example are represented by means of *logic* slots in the Arden Syntax, *Decision steps* in GLIF, *Decision tasks* in PRO*forma* and GLARE, *conditions* in Asbru, and *Decisions* in EON. The TNMs of most approaches define a fixed set of guideline tasks (one remarkable exception is EON, in which new types of tasks may be introduced). Most approaches also provide explicit support for controlled nesting of guidelines in order to model complex guidelines in terms of subguidelines (e.g., GLIF and EON) or subplans (e.g., PRO*forma*, Asbru, GLARE). GLIF also supports the representation of common guideline structures through Macros, which facilitates the reuse of guidelines that are employed often (e.g., 'if-then' rules). EON, PRO*forma* and Asbru also support the use of goals and intentions to formally specify a guideline on a higher level of abstraction.

From the architectural point of view, most CIG approaches provide specific support for at least two subtasks: (i) CPG acquisition and representation and (ii) CPG execution. Concerning acquisition, different issues have been addressed, ranging from the definition of suitable graphical interfaces to enhance the physician-system interaction, to the definition of set of tools supporting the progressive transformation from a textual CPG to its formal representation ([8, 11, 14, 16, 17, 20]).

With respect to execution, most approaches have developed execution engines that support the execution of an acquired CPG on a specific patient. Execution engines access the patient clinical data and use them to discriminate between alternative diagnostic/therapeutic paths, providing user physicians with recommendations about the next actions to be executed on the specific patient at hand.

18.3 Verification of Computer Interpretable Guidelines

While the representation and the execution of CIG seem nowadays to be at least partly consolidated, a very important open issue regards reasoning on CPGs. Indeed, CPGs are, first of all, knowledge sources and, as such, the Artificial Intelligence tradition demonstrates that they may be object of different forms of reasoning. Indeed, Artificial Intelligence widely demonstrates that representation and reasoning are strictly related tasks, complementing each other. In many Artificial Intelligence contexts, knowledge representation is useless without proper reasoning mechanisms operating on it. Indeed, reasoning mechanisms are the tool to "qualify" the represented knowledge, determining its implicit implications and, at the very end, showing its intrinsic underlying semantics.

In the last years, some reasoning tasks concerning CIGs have started to attract increasing attention. CIG *verification* and *conformance* are two of them. Roughly speaking, conformance analysis concerns the execution of a CIG on a specific patient, and is used in order to check whether the CIG recommendations have been followed in the treatment of the patient. A technical description of conformance, and an advanced treatment to it, are proposed in Chapter 5 of this book. On the other hand, in Chapter 19 we focus on CIG verification.

As regards verification, it is worth remembering that, in general, CPGs are a very extensive body of knowledge, which, as long as no formal language is used to represent it, is expressed in an "imprecise" (or partially ambiguous) way. The acquisition and formal representation of a CPG is thus a complex process, so that there is no guarantee that the final formal representation exactly achieves all the desired objectives in terms of correctness and completeness of the specified therapeutic and/or diagnostic treatments. Indeed, there are at least two potential sources of errors. On the one hand, given the large amount of knowledge it contains, there is no guarantee that even the original (textual) guideline correctly covers all the desired cases. On the other hand, the formalization of original (textual) guidelines into some CIG formalism is a complex process that may introduce errors. As a consequence of these problems (and, in general, of the complexity of CPGs), automatic or semi-automatic supports to verification are important, to check, e.g., whether an acquired CIG allows to cope in the desired way with its eligible patients. Only after the check that a CIG verifies the desired properties, physicians can fully trust it and the recommendation it provides. However, CIG verification is a complex task, also in consideration of the fact that CPGs contain heterogeneous forms of knowledge. As a consequence, the adoption of different methodologies (each one appropriate for

a specific type of knowledge/verification) seems to be the best option. In particular, the GLARE system emerges in the literature for the attention devoted to different forms of verification, through the adoption of different Artificial Intelligence formal techniques. In particular, three different forms of verification are considered:

1. verification that the temporal constraints in a CIG are consistent, through constraint-based temporal reasoning techniques;
2. verification of different medical properties of a CIG (e.g., its capability of coping with a given type of patients, or to support specific types of treatments), through model checking;
3. verification of probabilistic properties of a CIG in the context of a probabilistic knowledge base, through probabilistic modelling.

18.4 Aggregation of Evidence using Argumentation

As valuable as guidelines are for drawing the best available evidence into decision making in healthcare, there are also some important limitations.

1. Constructing guidelines can involve **assimilating massive amounts of evidence**. For instance, medical guidelines are based on a rapidly growing body of biomedical evidence, such as clinical trials and other scientific studies (for example, PubMed, the online repository of biomedical abstracts run by the US National Institute of Health has over 20 million articles). Production of evidence-based guidelines therefore requires **considerable human effort and expenditure** since the evidence needs to be systematically reviewed and aggregated.
2. Guidelines can become **out-of-date** quite quickly. For example, in medicine, even when major trials are published on topics, it may take years before the guidelines are rewritten to take account of the large amounts of newly available evidence (for example, PubMed is growing at the rate of 2 articles per minute). Decision makers are thus faced with the problem of assimilating and processing guidelines in combination with large amounts of newly available evidence which may warrant recommendations that conflict with, and so suggest revisions to, those recommendations provided by the guidelines.
3. Often there are **overlapping guidelines** to consider (from different agencies or bodies, and international, national, and local sources), and when there are multiple problems to be resolved (e.g. a patient with both cancer and liver problems). Thus, different guidelines may offer conflicting guidance.
4. Guideline recommendations are often written keeping in mind a **general population** so they need to be interpreted for individual cases with specific features. For example, given a patient with some particular symptoms and test results, the clinician needs to decide if the patient falls into any of the classes of patients for which the guideline offers guidance (e.g. if the patient is from a particular ethnic group, or if they are very young, or if their symptoms do not exactly correspond). If the clinician has doubts, then turning to the primary literature

for fuller descriptions of the relevant clinical trials may be useful. However, the clinician may then need to assimilate and aggregate the results from a number of articles which can be challenging. So after what may be an incomplete study of the evidence, the clinician decides whether or not to accept the recommendation from the guideline for the specific case.

5. Guidelines are **not sensitive to local needs** or circumstances. This may also result in non-compliance by the decision maker in using a guideline. For example, an international guideline may recommend a particular kind of scan for patients with a particular combination of symptoms, but a particular hospital using the guideline might not be able to provide such a scan, and would deviate from the recommendations by the guideline.
6. Use of guidelines can **decouple a decision maker from the evidence** which can be problematical since the decision maker may have valuable knowledge and experience for use in interpreting the evidence.

These shortcomings suggest that there is a need for knowledge aggregation technologies for making evidence-based recommendations based on large repositories of complex, rapidly expanding, incomplete and inconsistent evidence. These technologies should aim to overcome the limitations of guidelines listed above, and offer tools for users who need to make evidence-based decisions, as well as users who need to draft systematic reviews and guidelines, and users who need to undertake research in order to fill gaps or resolve conflicts in the available evidence. In Chapter 20, such knowledge aggregation methodology based on argumentation is presented as a tutorial.

References

[1] Paul A de Clercq et al. “Approaches for creating computer-interpretable guidelines that facilitate decision support”. In: *Artificial Intelligence in Medicine* 31.1 (2004), pp. 1 –27.

[2] John Fox, Nicky Johns, and Ali Rahmanzadeh. “Disseminating medical knowledge: the PRO*forma* approach”. In: *Artificial intelligence in medicine* 14.1 (1998), pp. 157–182.

[3] M. Adela Grando, David Glasspool, and John Fox. “A formal approach to the analysis of clinical computer-interpretable guideline modeling languages”. In: *Artificial Intelligence in Medicine* 54.1 (2012), pp. 1 –13.

[4] B Heller et al., eds. *Computer-Based Support for Clinical Guidelines and Protocols*. Vol. 83. Studies in health technology and informatics. IOS Press, 2001.

[5] David Isern and Antonio Moreno. “Computer-based execution of clinical guidelines: A review”. In: *International Journal of Medical Informatics* 77.12 (2008), pp. 787 –808.

[6] PD Johnso et al. “Using scenarios in chronic disease management guidelines for primary care.” In: *Proceedings of the AMIA Symposium*. American Medical Informatics Association. 2000, pp. 389–93.

[7] K Kaiser, S Miksch, and SW Tu, eds. *Computer-Based Support for Clinical Guidelines and Protocols*. Vol. 101. Studies in health technology and informatics. IOS Press, 2004.

[8] Katharina Kaiser and Silvia Miksch. “Versioning computer-interpretable guidelines: Semi-automatic modeling of Living Guidelines using an information extraction method”. In: *Artificial Intelligence in Medicine* 46.1 (2009), pp. 55 –66.

[9] Mark A Musen et al. “EON: A component-based approach to automation of protocol-directed therapy”. In: *Journal of the American Medical Informatics Association* 3.6 (1996), pp. 367–388.

[10] M. Peleg et al. “Comparing computer-interpretable guideline models: a case-study approach”. In: *Journal of the American Medical Informatics Association* 10.1 (2003), pp. 52–68.

[11] Mor Peleg and Samson W. Tu. “Design patterns for clinical guidelines”. In: *Artificial Intelligence in Medicine* 47.1 (2009), pp. 1 –24.

[12] Mor Peleg et al. “GLIF3: the evolution of a guideline representation format.” In: *Proceedings of the AMIA Symposium*. American Medical Informatics Association. 2000, p. 645.

[13] Silvana Quaglini et al. “Guideline-based careflow systems”. In: *Artificial intelligence in medicine* 20.1 (2000), pp. 5–22.

[14] R. Serban and A. Ten Teije. “Exploiting thesauri knowledge in medical guideline formalization”. In: *Methods of Information in Medicine* 48.5 (2009), pp. 468–474.

[15] Yuval Shahar, Silvia Miksch, and Peter Johnson. “The Asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines”. In: *Artificial intelligence in medicine* 14.1 (1998), pp. 29–51.

[16] Erez Shalom et al. “A quantitative assessment of a methodology for collaborative specification and evaluation of clinical guidelines”. In: *Journal of Biomedical Informatics* 41.6 (2008), pp. 889–903.

[17] Richard N. Shiffman et al. “Building better guidelines with BRIDGE-Wiz: development and evaluation of a software assistant to promote clarity, transparency, and implementability”. In: *JAMIA* 19.1 (2012), pp. 94–101.

[18] Richard N Shiffman et al. “GEM: a proposal for a more comprehensive guideline document model using XML”. In: *Journal of the American Medical Informatics Association* 7.5 (2000), pp. 488–498.

[19] E. Shortliffe. *Computer-based medical consultations: MYCIN*. Elsevier, 1976.

[20] Vojtech Svátek and Marek Ruzicka. “Step-by-step mark-up of medical guideline documents”. In: *International Journal of Medical Informatics* 70.23 (2003), pp. 329–335.

[21] Annette ten Teije, Silvia Miksch, and Peter Lucas. *Computer-based medical guidelines and protocols: a primer and current trends*. Vol. 139. Studies in health technology and informatics. IOS Press, 2008.

[22] Paolo Terenziani, Gianpaolo Molino, and Mauro Torchio. "A modular approach for representing and executing clinical guidelines". In: *Artificial intelligence in medicine* 23.3 (2001), pp. 249–276.

[23] Samson W Tu et al. "The SAGE Guideline Model: achievements and overview". In: *Journal of the American Medical Informatics Association* 14.5 (2007), pp. 589–598.

[24] Dongwen Wang et al. "Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines:: A literature review of guideline representation models". In: *International Journal of Medical Informatics* 68.13 (2002), pp. 59 –70.