Towards an “Operational” Educational Model in Healthcare: Exploiting Computer-Interpretable Guidelines

Alessio Bottrighi\textsuperscript{1}, Gianpaolo Molino\textsuperscript{2}, Luca Piovesan\textsuperscript{1} and Paolo Terenziani\textsuperscript{1}

\textsuperscript{1}DISIT, Università del Piemonte Orientale, Viale Teresa Michel 11, Alessandria, Italy
\textsuperscript{2}Az. Osp. San Giovanni Battista, C.so Bramante 88, Torino, Italy

Keywords: Clinical Guidelines, Practitioners, Healthcare-agents, Medical Students Education.

Abstract: Clinical guidelines (GLs) encode the best medical practices. GLs have been widely exploited to enhance the quality of patient care, and to optimize it, and several computer-based approaches to manage computer-interpretable guidelines (CIGs) have been proposed in the literature. Quite surprisingly, however, the potentialities of CIG systems in medical education have not been considered yet. In this position paper we argue that, since CIG systems support the “simulation” of the application of GLs on specific patients, they can be used to show students how to apply medical knowledge and best practices on specific patients. Therefore, using CIG systems, students may learn an “operational methodology” that, otherwise, they could only learn from the medical practice. In this paper, we have taken GLARE (and its extension, META-GLARE) as an example of CIG system, and we have addressed the roadmap we intend to follow to fully exploit its potentialities in medical education.

1 INTRODUCTION

Clinical guidelines (GLs) have been defined as “systematically developed statements to assist practitioner and patient decisions about appropriate healthcare under specific clinical circumstances” (Field and Lohr, 1990). They encode best medical practices, promoting the adoption of evidence-based medicine and supporting the quality and the standardization of healthcare services, and the optimization of costs. Thousands of GLs have been devised. For example, the Guideline International Network (http://www.g-i-n.net) groups 106 organisations in 54 countries, and provides a library of more than 6500 GLs. In this paper, we argue that such a valuable body of knowledge may be paired with software tools supporting its application and simulation (the so-called Computer-Interpretable Guideline Systems, see below), and fruitfully used to complement “traditional” education in medicine.

Despite the large number of GLs, and their diffusion, GLs have not fully achieved all their goals, in terms of quality and optimization of the healthcare services. The discussion of the reasons for GL not-full success is outside the goals of this paper. Here we just highlight two of such reasons, which have largely motivated the introduction of ICT support tools to GL management. Non-computerized GLs are large bodies of knowledge (even hundreds of pages), mostly expressed as free text, describing the “best practice” recommendations for the treatment of a given disease. Using such large bodies of (textual) knowledge to diagnose and treat a specific patient is a difficult issue for physicians, who are left alone

(i) to interpret the textual description (which, as any natural language text, contains imprecisions and ambiguities)

(ii) to identify the “mapping” between the general recommendations in the GL and the specific patient (and disease or clinical condition) at hand; indeed, when diagnosing or treating a patient, physicians should quickly identify, among pages and pages of free text, the few parts that are relevant for the specific patient at hand.

Considering (among the others) issues (i) and (ii) above, since the 90’s, the medical community has started to develop many different systems and projects to support physicians in the management of GLs. In particular, in Computer-Interpretable Guidelines (CIG) systems, GLs

(1) are represented in a formal and unambiguous way, and
(2) most CIG systems provide execution tools that are automatically connected to the patients’ clinical records, so that they are able to automatically detect the GL recommendations that are more appropriate for the given patient. Besides disease-based support tools (e.g., software tools considering only a specific disease), several disease-independent tools for CIGs have been proposed in the literature (see e.g. (Fridsma, 2001; Gordon and Christensen, 1995; Peleg, 2013)). Such systems usually provide facilities to acquire, represent and/or execute GLs in different (medical) domains. Such tools have been mainly developed to support physicians in patient care, while, quite surprisingly, their impact in medical education has been quite neglected until now.

2 CIG SYSTEMS AND EDUCATION

Since GLs encode “best medical practices”, their role in the medical practice has been widely investigated. Indeed, several studies have shown that the quality of patient treatments is higher in case physicians have been educated to adopt GLs (see, e.g., (Corriere et al., 2014) as regards the case of diabetes).

CIG systems have been quite widely used in order to acquire GLs into a computer format, and to support practitioners in the diagnosis and treatment of patients.

However, to the best of our knowledge, the possible impact of the adoption of CIG systems in education has not been investigated yet. Indeed, the main claim of this position paper is that although CIG systems until now have been conceived as a support to practitioners to deal with patients affected by specific diseases, they can evolve to support also medical education. Indeed, in this paper we argue that CIGs have the potentiality of drastically improving the education of healthcare agents (medical students/practitioners/nurses), proposing them an “operational” methodology showing how to (reason in order to) apply best practice recommendations to specific patients.

First of all, it is important to point out that the approach we propose may support the education of both practitioners (e.g., their continuous education) and medical students. Indeed, though the former have responsibilities that the latter do not have, both of them have to learn how to apply best medical practices to patients. Indeed, for the same reason, also other types of healthcare agents (e.g., nurses) can take advantage of our approach.

Second, we want to highlight that the approach we propose is not intended to substitute “traditional” education in medicine (erogated through courses and textbooks), but to complement it.

As a matter of facts, “traditional” education provided by medical texts and courses covers a wide range of knowledge, ranging from human anatomy to the description of diseases, and of their treatments. While such aspects are certainly very important (and necessarily needed), another aspect should also deserve a specific attention, the “operational” (our terminology) aspect: how to “operate” on a specific patient? How to proceed to diagnose and treat her/him through the best medical practices? Such a kind of “operational” knowledge is usually not considered in textbooks, so that it can be learned by medical students (and healthcare agents in general) only “by practicing”.

On the other hand, CIG systems have the potentiality of enriching education through an operational methodology, showing how to proceed to diagnose and to treat specific patients. Indeed, as discussed in the introduction, any CIG supports healthcare agents in the treatment of a specific patient, by automatically “focusing” on the part of the GL which is relevant to the (status of) the patient at hand (see problem (ii) and the solution (2) in the introduction). This is, indeed, the process that students have to learn to be able to cope with real patients: how to focus on the more appropriate parts of their general medical knowledge and of the best practices and to apply them when considering the patient.

Until now, CIG system have been developed in order to support practitioners in the diagnose and treatment of patients. To do so, most CIG systems provide execution modules, which support physicians in the application of a general GL to a specific patient. Execution modules take in input the patient’s clinical record and, on the basis of the patient’s data, suggest the proper actions and, above all, they help physicians to take the decision appropriate to the patient. Intuitively speaking, one can say that CIG execution tools support the focusing on the specific parts of the general GL that are appropriate for the patient at hand, and help to take the best decisions on the basis of the patient’s data. In other words, they make best medical practices “operational”, by supporting the application of general GLs to the patients.

In this paper, we propose a new use of CIG systems. Instead of being used to support physicians in the execution (application) of a GL to a real patient, they could also be used in education, to simulate such an execution (application). The learning healthcare
agent (e.g., student, practitioner) is given a patient (in the form of the history of the evolution of the patient’s data), and the CIG system can be used to show how the GL would recommend to operate on her/him (while, of course, no action is “physically” executed: it is just a simulation).

Notably, since it is a simulation, the presence of “physical” patients is not required: all that is needed is the evolution of patients’ clinical data. As a remarkable consequence, teachers may propose to students (the data of) “significant” patients, and/or invent (the data of) patients in such a way to force students to explore the diagnosis/treatment of most important (and/or problematic) cases.

This contribution may provide, in our opinion, crucial benefits in the area of education, to complement textbooks and traditional courses with “practical” examples (through simulation) of how to apply medical knowledge on specific patients. Indeed, different modalities of education through simulations can be provided:

(i) **“standard” simulation**, in which the CIG system is used to show students how the (real or invented) patient should be treated, step-by-step, given the CIG recommendations

(ii) **“second-opinion” simulation**, in which a student has to indicate how s/he would treat a (real or invented) patient, and the CIG system is used to indicate to the student where s/he has followed the recommendations of the CIG, and where s/he has violated them (with additional explanations).

Notably, the educational approach we propose can be strongly based on the facilities provided by currently existing CIG systems. In particular, the acquisition of CIGs can be done through the acquisition tools already provided by such systems, with usually automatically interact with patient clinical records. As a consequence, the starting point of each simulation could simply be the loading of

- an already acquired CIG
- the evolution of the data of a (real or invented) patient.

As regards the simulation, “standard” simulation can be performed by taking advantage of the execution module already provided by most CIG systems. On the other hand, the “second-opinion” simulation require a modification of current execution tools, as discussed in the next Section.

While in Section 4 we further elaborate such issues, here we highlight that, for education, the disease-independent CIG systems have a major advantage with respect to the disease-based ones. In fact, with such systems, one can acquire a library of different CIGs (expressed using the same formalism, the one provided by the CIG system), and provide a unique software tool for simulating each of them. Therefore, they allow one to develop a uniform educational environment, in which all the CIGs are homogeneously represented through the same formalism, and the same simulation mechanism is applied to all of them. On the contrary, the adoption of disease-dependent CIG systems would force the adoption of multiple different representation formalisms and simulation mechanisms, one for each different CIG. This move would force students to learn different formalisms and to practice different software tools for simulation, which would only be a loss of time for medical students.

Though our position and the discussion until now is fully general, in our future research we aim at making it fully operative taking advantage of GLARE (Guideline Acquisition, Representation and Execution; (Terenziani et al., 2008)), and its extension, META-GLARE (Bottrighi and Terenziani, 2016). Before highlighting how we plan to cope with education through CIGs, in Section 3 we quickly mention the current status of GLARE and META-GLARE.
3 STATUS OF GLARE PROJECT

GLARE is a disease-independent system for the acquisition and execution of CIGs, which we are developing since 1997, in collaboration with the physicians of Azienda Ospedaliera San Giovanni Battista in Torino, Italy (one of the major hospitals in Italy). The core of GLARE (see the reference and the box on the left of Fig. 1) is a modular architecture. CG_KRM (Clinical Guidelines Knowledge Representation Manager) is the main module of the system: it manages the internal representation of GLs, and operates as a domain-independent and task-independent knowledge server for the other modules; moreover, it permanently stores the acquired GLs in a dedicated Clinical Guidelines Database (CG-DB). The Clinical Guidelines Acquisition Manager (CG_AM) provides expert-physicians with a user-friendly graphical interface to introduce the GL into the CG_KRM and to describe them. It may interact with four databases: the Pharmacological DB, storing a structured list of drugs and their costs; the Resources DB, listing the resources that are available in a given hospital; the ICD DB, containing an international coding system of diseases; the Clinical DB, providing a “standard” terminology to be used when building a new GL, and storing the descriptions and the set of possible values of clinical findings. The execution module (CG-EM) executes a GL for a specific patient, considering the patient’s data (retrieved from the Patient DB). CG-EM stores the execution status in another DB (CG Instances) and interacts with the user-physician via a graphical interface (CG-IM).

GLARE’s architecture is open. In the latest years, several new modules and/or methodologies have been added to cope with automatic resource-based contextualization (ADAPT module, (Terenziani et al., 2004)), temporal reasoning (TR, (Anselma et al., 2006)), decision making support (DECIDE_HELP, (Montani et al., 2005)), model-based verification (VERIFY, (Bottrighi et al., 2010)), and comorbidities (COMORBID, (Piovesan et al., 2014) (Piovesan et al., 2015).

Representation Formalism. The core of GLARE is the definition of the representation formalism used to model GLs. Notably, a unique (disease-independent) formalism is provided, and is used by all the modules of GLARE. In GLARE, a GL is represented through the set of actions composing it. GLARE distinguishes between atomic and composite actions. Atomic actions can be regarded as elementary steps in a GL, in the sense that they do not need a further decomposition into sub-actions to be executed. Composite actions are defined in terms of their components (atomic or composite actions), via the “has-part” relation. GLARE adopts four different types of atomic actions. Work actions represent basic atomic actions which must be executed on the patient, and can be described in terms of a set of attributes, such as name, (textual) description, cost, time, resources, goals. Query actions are requests of information, which can be obtained from the outside world (physicians, Databases, knowledge bases). Conclusions represent the output of decision actions. Decision actions are specific types of actions embodying the criteria which can be used to select alternative paths in a GL. They are crucial also to the education task, so that they are detailed in the following subsection.

Control relations establish which actions might be executed next and in what order. We distinguish among four different control relations: sequence, constrained, alternative and repetition. Temporal constraints between actions (e.g., overlaps, during) are also supported.

Testing. GLARE has been applied to different medical domains, including bladder cancer, reflux esophagitis, heart failure, and ischemic stroke.

META-GLARE

In the last years, we have defined a new GL system, META-GLARE (Bottrighi and Terenziani, 2016), on top of GLARE. META-GLARE has been designed to make extensions of CIG formalisms easy to implement-manage, so that its availability is very important, to support the extensions of GLARE with the constructs needed to support educations. However, META-GLARE is a support for system designers, while users (e.g., physicians, or students) never have to directly interact with META-GLARE.

3.1 Decisions in GLARE/META-GLARE

Decisions are probably the most crucial aspect of GLs, since they allow user-physicians to discriminate among alternative actions. GLARE supports two different types of decisions: diagnostic and therapeutic decisions.

Diagnostic decisions consider a set of parameters (to be evaluated on the basis of the status of the patient) to discriminate among different diagnoses. Of course, different parameters have to be considered, depending on the specific diagnostic decision. In GLARE, we support “scored” diagnostic decisions: for each one of the relevant parameter (e.g.
fever) and considering the alternatives to be discriminated, expert-physicians define a priori a set of values/ranges (e.g., fever < 37, between 37 and 38, between 38 and 39, > 39; or, qualitatively: no, low, medium, high). For each one of the values/ranges of the findings in each one of the alternatives, the expert-physician defines a score. Moreover, a threshold is fixed to separate recommended/disliked actions. GLARE execution engine considers the patient’s data to evaluate each parameter, and, for each alternative diagnosis, sums up the scores. Only the alternatives whose additive score is greater than the threshold are recommended for selection to the user.

To summarize, a decision among a set of diagnoses can be represented as an open set of triples <diagnosis, parameter, score> (where, in turn, a parameter is a triple <data, attribute, value>), plus a threshold which is relative to the sum of the scores. For example, Fig. 2 shows a diagnostic decision within the Gastro-Esophageal Reflux Disease (GERD) guideline, namely the “differential diagnosis”, which allows the physician to discriminate between “possible GERD”, “probable GERD” and “no GERD” according to the values of several parameters: heartburn absent (abbreviated by “no-hb” in Fig. 2), or lasted not more than 3 months (“hb≤3”), or lasted more than 3 months (“hb>3”); dysphagia absent (“no-dys”) or present (“dys”); occurrence of weight loss (“wl”) or non-occurrence (“no-wl”); absence (“no-hem”) or presence (“hem”) of hematemesis; postural reflux absent (“no-ref”), or lasted not more than 3 months (“ref≤3”) or more than three months (“ref>3”). The threshold for such a decision (not shown in Fig. 2) is >9. E.g., given the scores in Fig. 2, one should conclude “no GERD” only if heartburn, dysphagia, weight loss, hematemesis and postural reflux are all absent.

Notably, also Boolean diagnostic decisions are supported by META-GLARE, and will be used in the educational task.

On the other hand, in the therapeutic decisions, physicians have to choose from different therapies (treatments) according to a given set of parameters: effectiveness, cost, side-effects, compliance, duration; which have to be specified for each one of the alternative therapies to be discriminated. Thus, a decision action can be represented by the set of the parameters above, for each one of the alternatives.

E.g., in Fig. 3 we show a possible way of modelling the decision among four different treatments in the case of “possible GERD”, in the tabular fashion which is used by GLARE acquisition module.

### 3.2 GLARE in Education: First Steps

GLARE has been already tested in the training of physicians in the context of emergency medicine, to cope with poly-trauma, within the ROPHS (Report on the Piedmont Health System) project (Leonardi et al., 2012). In such a project, GLARE has been used “as it is” for educational purposes following the lines suggested in Section 2, point (1). Specifically:

1. together with expert physicians, we have acquired a GL for poly-trauma in GLARE
2. expert physicians have defined for us a set of “typical” patients affected by poly-trauma. We
have stored the evolution of the data of such patients in the DB of clinical records.

(iii) Expert physicians have used GLARE execution module to provide a “standard simulation” (see Section 2) of the treatement of the patients.

The adoption of GLARE has supported the physicians attending the course while learning the “process”: we showed the simulation of the application of the polytrauma CIG to different patients, showing step-by-step the results of applying the CIG’s recommendations. However, we did not have the possibility to split the physicians attending the course into two groups (one taking advantage of GLARE, and the other not using it), so that we could not have a practical analysis of the educational advantages provided by the adoption of GLARE. However, this has just been our first experience in adopting GLARE (and META-GLARE) for educational purposes. Our roadmap is much more ambitious, and it is briefly presented in the following section.

4 ROADMAP TO A FULL EXPLOITATION OF META-GLARE FOR EDUCATION

Besides the “conservative” application of GLARE for education in the ROPHS project, in which GLARE has been used “as it is” to show students practical examples about how to operate on specific (polytrauma) patients, our roadmap plans much more ambitious applications, and consequently, a lot of extensions. Notably, to achieve a fast prototyping of such extensions, we will take advantage of META-GLARE, which has been explicitly designed in order to make extensions of CIG formalisms easy to implement-manage.

4.1 Explanatory Annotations

For the sake of education, it is very important that each action in the CIGs is carefully annotated with detailed explanations and motivations. Indeed, GLs are based on best practices and medical evidence, and such knowledge has to be explicitly stored in order to support education. Of course, such an extension does not require specific efforts from the computer science point of view, but, indeed, the acquisition of such annotation is a long and time-consuming task, to be performed mostly by domain experts.

4.2 Propose & Check Execution Modality

As discussed in Section 2, issue (ii), also “second-opinion” simulation can play an important role in medical education. However, to support “second-opinion” simulation, GLARE execution engine must be enriched with a new modality, in which GLARE, instead of suggesting the most suitable choice for diagnostic and therapeutic decisions, let the student choose without any suggestion, and then compares the student’s choice with the one that would have been suggested by the CIG. Especially in case of disagreement (but also in case the choice of the student is the same that would have been recommended by the GL) the motivations for the choice recommended by the GL should be provided to the students (e.g., by showing the logic underlying the decision – see the discussion about therapeutic and diagnostic decisions in Section 3.1).

4.3 Fake Decisions and Paths

By definition, GLs contain the “best practices” to cope with specific diseases. However, during the educational processes, diagnostic and therapeutic problems might be made more complex for students by adding wrong decisions leading to non-existing (in the real GL) alternative paths into the GL (called fake decisions and paths in the following). Fake decisions and path can be used in the “second-opinion” simulation, to increase the complexity of the problem, especially in case such fake alternatives represent cases of frequent-plausible medical errors. Of course, the explanatory facility is fundamental in this context: each fake path must be paired with exhaustive annotations, detailing why such a path should not be applied.

4.4 What-if Analysis

The “what-if” analysis in an important cognitive tool in human reasoning in general, and in the medical context in particular, since it allows one to analyse the consequences of a given action or choice. GLARE already supports a quite reach mechanism to support what-if analysis (Terenziani et al., 2002). GLARE’s “what-if” facility is the implementation of a form of hypothetical reasoning. In particular, users are helped in gathering the various types of information necessary for discriminating among the alternative paths of actions at any stage of the GL. Relevant decision parameters (e.g., costs, resources, times) are gathered from the alternative paths in an automatic
way. At the end of this process, the tool displays the values of the chosen parameters for each of the alternative paths.

Such a mechanism can be used for education in two different ways:

(1) In the “standard” simulation, to show to the students the different consequences of decision actions in the GLs.

(2) In the “second-opinion” simulation, to support students’ choices. Before taking a decision, students could use such a facility to explore the consequences of choosing one alternative or another.

4.5 Patients’ Data Generation

A central issue is the definition of patients. Indeed, (the data of) a specific patient will be the starting point of each educational session, in which students will have to simulate the application of the GL “best practices” on her/him. Since we are speaking of education, we do not “physically” operate on real patient. Therefore, while in a real application (of a GL on a patient) the whole evolution of the data of the patient cannot be known “a-priori” (but it is discovered “step-by-step”, looking at the evolution of the status of the patient), in the educational context we only simulate an application, so that we can hypothesize to know a-priori the whole evolution of the patient. However, different forms of simulations can be provided, depending on the model one adopts to represent the patient data. At least two different approaches are possible, that we call (i) deterministic patient and (ii) probabilistic patient.

In the “deterministic patient” model, we assume that the clinical record of the patients contains, “a-priori” (i.e., already at the beginning of the simulation) the whole evolution of the patient’s data. There is only one specific evolution, corresponding to a specific path in the CIG, the one recommended for such a patient. A temporal (relational) database is used to manage such temporal data (as regards GLARE, consider, e.g., (Anselma et al., 2018; Stantic et al., 2012)). The deterministic patient model can be used in the “standard” simulation (since the data evolution corresponds to the path of actions that one wants to show to the students). It may also be used for the “second-opinion” simulation. However, in such a case, whenever students deviate from the recommended path, they have to receive a warning (and exhaustive explanations), and then be forced to continue the simulation considering the recommended path (since only the data corresponding to the recommended path are available to the system).

On the other hand, the probabilistic patient model represents the initial status of a patient, and the probabilistic history of her/his evolutions, depending on the GL actions performed on her/him. Obviously, the definition of (“probabilistic”) patients is complex, and requires the availability of a lot of knowledge, e.g. (Dagan et al., 2007), (Real et al., 2015). However, it supports also flexible “second-opinion” simulations, in which students may also follow non-recommended paths in the CIGs.

Finally, it is important to emphasize that, while the whole history of the evolution of the (data of the) patients must be provided as input to the simulation, the students will operate “step-by-step” in the application of the CIGs to the patients, so that, at each step, they will only see the data holding at that point of the execution, plus past data.

4.6 Experimental Evaluation

Needless to say, the experimental evaluation is necessarily a cue issue in the educational context. One of the major goals we have in this context is to be able to establish partnerships or projects with educational institutions, in order to be able to provide extended and accurate experimental evaluations of the impact of adopting CIG systems in medical education.

5 CONCLUSIONS

GLs and CIGs have a quite consolidated role in the standardization and optimization of the healthcare services. On the other hand, their potentiality in the education of medical students and, more generally, of healthcare agents has not been adequately explored. In this position paper, we claim that CIG systems have great potentiality in educations. Since they support the “simulation” of GLs on specific patients, they can be used to show students how to apply medical knowledge and best practices on specific patients, providing them an “operational methodology” that, otherwise, they could only learn from the medical practice.

In this paper, we have taken GLARE as an example of CIG system, and we have addressed the roadmap we intend to follow in order to fully exploit its potentialities in medical education.