Original Research

Evaluation of clinical practice guideline-derived clinical decision support systems using a novel quality model

Julián Alberto García-García [a,*], Manuel Carrero [a], María José Escalona [a], David Lizcano [b]

[a] ES3 Group (Engineering and Science for Software Systems group). Computer Languages and Systems Department. Escuela Técnica Superior de Ingeniería Informática., Avda. Reina Mercedes s/n. 41012 Seville, Spain
[b] School of Computer Science, Madrid Open University, UDIMA., Campus Collado Villalba, Madrid, Spain

ARTICLE INFO

Keywords:
Comparative studies
Clinical practice guideline
Clinical decision support system
Quality model

ABSTRACT

Over the last decade, clinical practice guidelines (CPGs) have become an important asset for daily life in healthcare organizations. Efficient management and digitization of CPGs help achieve organizational objectives and improve patient care and healthcare quality by reducing variability. However, digitizing CPGs is a difficult, complex task because they are usually expressed as text, and this often leads to the development of partial software solutions. At present, different research proposals and CPG-derived CDSS (clinical decision support system) do exist for managing CPG digitalization lifecycles (from modeling to deployment and execution), but they do not all provide full lifecycle support, making it more difficult to choose solutions or proposals that fully meet the needs of a healthcare organization. This paper proposes a method based on quality models to uniformly compare and evaluate technological tools, providing a rigorous method that uses qualitative and quantitative analysis of technological aspects. In addition, this paper also presents how this method has been instantiated to evaluate and compare CPG-derived CDSS by highlighting each phase of the CPG digitization lifecycle. Finally, discussion and analysis of currently available tools are presented, identifying gaps and limitations.

1. Introduction

Today, any organization requires innovative, flexible solutions to digitize and automate their processes [2] in conjunction with new technologies [3] for increasing their competitiveness and productivity [4,5]. In healthcare environments, digitization has greater disruptive potential because it affects aspects like patient care, spiraling costs, quality and rewarding value [6]. However, before the technological boom of the last decades, patient care was usually based on the manual application of clinical practice guidelines[1] (CPG); i.e., it was usually based on paper-based medical reporting without automatic support (such as computerized systems for clinical decision support).

Today, many authors have studied the benefits of establishing well-defined CPG digitization processes[2] to both healthcare professionals and patients [7–9]. CPGs help reduce variability in clinical practice and improve the quality of clinicians' performance and decision-making. CPGs pool existing knowledge to facilitate the use of effective, reliable interventions based on empirical evidence and clinical experience [10], but the CPG digitization processes make this effort more efficient and reliable by using computerized clinical decision support systems.

Over the last decade, many scientific initiatives and technological proposals have been published to facilitate the automation and digitization of clinical guidelines. These technological initiatives are referred to as CPG-derived CDSS (clinical decision support system) in this paper. Most of them, however, have been limited in their functional scope, or their practical application has focused on treating specific pathologies in controlled environments. After an initial rise of formalisms and languages [11], it is necessary to assume and address the fact that the actual applications of CPG-derived CDSS are limited. This situation has been analyzed by several authors from different perspectives and the reasons for this are heterogeneous [12,13], but the reasons are usually

* Corresponding author.
E-mail addresses: juliangg@us.es (J.A. García-García), mcarrero@us.es (M. Carrero), mjescalona@us.es (M.J. Escalona), david.lizcano@udima.es (D. Lizcano).

[1] The definition of CPG (clinical practice guidelines) assumed in this paper is the one offered by Steinberg et al. [1], who define CPG as *≪statements that include healthcare processes, clinical rules and 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≫*.

[2] The concept of CPG digitization process is understood in this paper as the process of translating the content of CPG content to digital form for machine-interpretable understanding.

https://doi.org/10.1016/j.jbi.2023.104573

related to clinicians' limited understanding and trust in the underlying models, thus creating poor engagement in their usage [14]. In addition, other barriers have been identified in the scientific literature [15]: limitations for the integration of these systems into EHR (Electronic Health Record) systems and clinical workflows; lack of sufficient patient-specificity; mismatch to the cognitive tasks and processes of the end user (healthcare professional); lack of change management mechanisms in clinical recommendations and clinical processes in runtime; and effective interoperability mechanisms.

In this context, digital health innovations related to CPG have not been adopted on a large scale and are usually abandoned when they are not upscaled or kept in use over time at organizational or system level [16,17]. As mentioned above, the factors influencing non-adoption and abandonment are complex and include health conditions, technology, value propositions, adopters' systems (professional staff, patients, and lay caregivers), organization(s), institutional contexts, as well as interaction and mutual adaptation between these factors over time [16,18]. Consequently, the success of the CPG-derived CDSS implementation will clearly depend on the analysis of previous critical success factors, and modeling efforts should allow for the broadest and most effective use of these systems based on models of technology adoption, evidence-based practices, and conceptual models in clinical practice [15].

The contributions of this paper. The final purpose of our work is to support decision makers with a method based on quality models that provides stakeholders with information regarding the eventual adoption of a new technology into their organizations according to their objectives. For this purpose, this paper describes a method based on quality models to uniformly compare and evaluate technological tools, offering a rigorous method that uses qualitative and quantitative analysis of technological aspects. Later, this method is instantiated to evaluate and compare five currently available and specific CPG-derived CDSS (GLEE, ArdenSuite, GLARE, DeGeL, and KnowWe) by highlighting each phase of the CPG digitization lifecycle. The opinion of technology consultants who are experts in the application of information and communications technology (ICT) in the healthcare environment were considered for this purpose.[3] This evaluation is carried out objectively and uniformly testing each technological tool on each technological aspect included in our quality model. After carrying out this evaluation, the final support score of each CPG-derived CDSS is obtained applying the method described in this paper.

Table 1 summarizes the contribution of our paper to the existing literature considering the problem, what is already known and what this paper adds.

The rest of the paper is organized as follows. Section 2 presents the methods used in this paper; specifically, it describes the phases that make up this method, as well as the quality model (c.f., Section 2.2.2) and scoring method (c.f., Section 2.2.3) that was used to evaluate each tool under study. Subsequently, Section 3 presents our results and how our quality model on CPG-derived CDSS has been applied and what results of its evaluation have been, respectively. Later the discussion and analysis is presented in Section 4. Finally, Sections 5 and 6 describe some related works and conclusions as well as future work, respectively.

[3] Regarding the partner entities and experts who were consulted, our research group (the ES3 group—Engineering and Science for Software Systems), has extensive experience in university collaboration with private companies (including FujitsuTS, Everis, Wellness Telecom, and Soltel, among others) and Spanish Health organizations (such as, private andalusian assisted reproduction clinic Inebir, and the Andalusian Regional Government's Department of Health and Social Welfare, among others), where we have applied computer science techniques in the health field.

2. Methods

In this section, the method applied to compare and evaluate CPG-derived CDSS is explained; this method allows to support decision makers based on quality models providing the quantitative values necessary to justify a strategic decision. For this purpose, the technical and research questions (TRQ) and the search protocol are described in Section 2.1. Later, Section 2.2 describes our quality model, which defines each characteristic that has to be valued for each CPG-derived CDSS, as well as our rating methods (which establish a quantitative scoring procedure to homogenously assess each technological solution).

2.1. Technical and research questions, and search protocol

The objective of this study was to answer the following TRQ: (1) ≪*What are the main published CPG-derived CDSS currently available for managing CPG digitalization lifecycles?*≫ and (2) ≪*What is the functional scope and limitations of these systems considering features on modeling, design, deployment, implementation and operation, monitoring and control, analysis and other criteria?*≫.

Many search keywords could be used to answer these questions. Some of them were: ≪careflow≫, ≪care workflow≫, and ≪clinical guideline tool≫, among other. These keywords were used to carry out exhaustive searches in different digital libraries. The libraries were chosen in accordance with the recommendations of Brereton [19], who identified IEEExplore, ACM Digital library, GoogleScholar, Citeseer library, Inspec, ScienceDirect, and EI Compendex as the most relevant digital libraries in the field of software engineering. We also considered other digital libraries relevant to the use of computer science in the health context (such as PubMed). In addition, other general purpose information sources were considered in our study (such as Google, technical YouTube channels, whitepapers, among others) because usually technical information is included in these forums by technicians.

2.2. Defining the quality model

2.2.1. The CPG digitization lifecycle: previous context

Before describing the quality model used to uniformly evaluate CPG-derived CDSS, a expert group was consulted to determine the phases of a minimum lifecycle to strategically digitize clinical guidelines. This group of experts was composed of consultants, practitioners and technologists in healthcare ICT. On the one hand, software engineering managers from several private companies (including FujitsuTS, Everis, Wellness Telecom, and Soltel) were consulted considering their experience in the design and development of process-oriented HIS (Hospital Information Systems) capable of supporting the implementation of CPG . On the other hand, managers of public health organizations (such as the Andalusian Regional Government's Department of Health and Social Welfare) were also consulted considering their end-user experience in the use of CPG-derived CDSS.

After considering the experience of this expert group and the results of our interviews, we concluded that it was possible to establish a minimum lifecycle to strategically manage clinical guidelines (c.f., Fig. 1). This lifecycle was defined drawing inspiration from Hill's proposal [20], which also addressed the definition of a management model for the continuous, incremental improvement of CPGs. In this regard, we decided to concentrate our quality model based on this lifecycle with the following phases: Modeling; Design; Deployment; Execution and Operation; Monitoring and Control; and Analysis. Section 2.2.2 describes these phases and their features in detail.

2.2.2. Characterization scheme of the quality model

As mentioned above, our quality model is based on a characterization scheme that categorizes different features in the stages of our proposal of CPG digitalization lifecycle (c.f., Fig. 1). These stages and features are detailed in the next subsection.

Table 1
Statement of significance.

Problem	Many technological proposals have been published to facilitate the automation and digitization of clinical practical guidelines (CPG), which are referred to as CPG-derived CDSS (clinical decision support system). Most of them have limits in their functional scope, or their practical applications.
What is Already Known	Although there have been similar comparative studies on CPG-derived CDSS, a prominent limitation in these studies is that their methods propose isolated functional characteristics with no connection to the CPG digitization lifecycle, which limits their evaluation considering the needs and requirements of each healthcare entity.
What This Paper Adds	This paper describes a method based on quality models to uniformly compare and evaluate technological tools, offering a rigorous method that uses qualitative and quantitative analysis of technological aspects. Also, this paper instantiate our quality model to evaluate and compare five currently available and specific CPG-derived CDSS (GLEE, ArdenSuite, GLARE, DeGeL, and KnowWe) by highlighting each phase of the CPG digitization lifecycle.

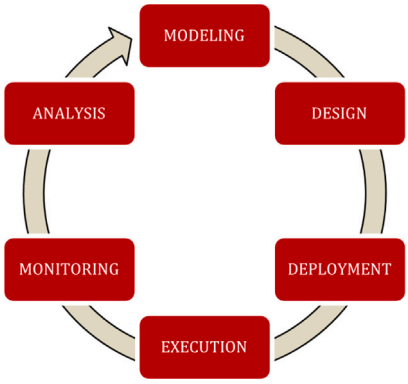

Fig. 1. Proposal of CPG digitalization lifecycle.

2.2.2.1. Modeling criteria. Over the last decade, different modeling languages have been proposed in which to define CPGs, but each one has its own specific peculiarities, advantages and disadvantages. The objective of the modeling stage was to represent and express any CPG both structurally and formally, allowing it to be unequivocally interpreted and automatically processed by a software system. This representation is usually achieved using formal languages to define things like information flows, clinical decision rules, roles and workflows. The modeling of the views or perspectives of stakeholders involved in the healthcare process could also be a relevant feature because it makes it possible to focus on responsibilities by user profile. The modeling stage therefore included the following modeling (MO) criteria and features:

- **MO-01. Formal modeling language**. The system must allow the use of a standardized language to represent the elements in the CPG.
- **MO-02. Expressivity**. The modeling language used had to have sufficient elements (e.g., complete representation of clinical actions and automatic actions, flow control mechanisms, definition of parallel paths, description of patient data and medical knowledge, etc.) to adequately express or represent the different concepts and relationships in the problem domain.
- **MO-03. Multilanguage support**. The system had to accept more than one standardized representation language, making it possible to choose which language to use to represent the CPG.
- **MO-04. Modeling tool**. The system had to allow clinical processes to be modeled by means of process modeling tools or incorporate some type of process editor for this purpose.
- **MO-05. Visual editor based on flowcharts**. The system had to incorporate or support the use of visual editors for modeling clinical processes by means of flowcharts.
- **MO-06. Import/export support**. The system had to present a way to use already represented clinical guidelines by importing them. It also had to offer the possibility of exporting a clinical guide previously modeled in the system to a file.
- **MO-07. Granularity**. The system had to allow several tasks in the general clinical process to be joined together to form a "major" task or "major tasks" (composite tasks, also called subprocesses in the clinical process). The sub-processes created had to be reusable in other CPG, in order to allow the reuse of common factors.
- **MO-08. Multiview support**. The system had to have different ways of displaying the clinical guideline modeling information, such as by showing a flowchart of the process in diagram form, showing a hierarchical list of the activities or tasks in the clinical process, or showing the activities to be performed by a specific professional, etc.
- **MO-09. Documentary references**. The system had to provide tool or option that would facilitate access to the medical and/or scientific reference documentation on which the CPG had been modeled.
- **MO-10. Clinical rule modeling**. The system had to have mechanisms to establish process flow control by means of clinical rules: for example, by means of logic expressions, decision tables, decision trees, etc.
- **MO-11. Formal rule language modeling**. The system had to use the syntax of a formal expression language to model decision criteria.
- **MO-12. CPI modeling**. The system had to provide mechanisms with which it would be possible to define different types of Clinical Performance Indicators (CPI).
- **MO-13. Data protection modeling**. The system had to provide mechanisms to enforce or assist in the enforcement or modeling of data protection laws relating to the use of patient data, and of any other legislation pertaining to privacy.
- **MO-14. Standard terminology**. The system had to allow the use of standard medical terminologies suite such as SNOMED-CT (Systematized Nomenclature of Medicine-Clinical Terms) [21], LOINC (Logical Observation Identifiers Names and Codes) [22], CPT (Current Procedural Terminology) [23], ICD-9 (International Classification of Diseases. Ninth Revision) [24] or UMLS (Unified Medical Language System) [25], among other.
- **MO-15. Automatic consistency check**. The system had to be able to automatically verify the consistency of clinical process flows, warning of possible logical inconsistencies before they are executed.

2.2.2.2. Design criteria. This design phase addressed the criteria that had to be met, if possible, to make the system capable of efficiently implementing the proposed model. The aim here was to define a series of characteristics that, forming part of the system structure itself, would allow or facilitate the subsequent execution of the CPG. To this end, the

system had to include the possibility of making the model defined in the modeling phase executable, facilitate communications with other systems to make use of essential external services and, for example, provide tools for end users to interact with the system by means of graphical user interfaces (GUI).

- **DE-01. Interoperability**. The system had to have APIs or other interconnection methods that would allow it to connect with other medical systems or services to obtain the data necessary.
- **DE-02. GUI design tools**. The system had to enable the manual or automatic creation of GUI with which a healthcare professional could easily interact with the system (for example, by creating pages with web forms, dialog boxes, etc.).
- **DE-04. Data protection mechanisms**. The system had to offer security mechanisms to protect both data and the system itself, such as data encryption, firewalls or tools against cyber-attacks.
- **DE-04. User access control**. The system had to allow control of user access by means of identification and subject the necessary access permissions.
- **DE-05. Import organizational info**. Support to enable the system to import information about the healthcare organization, including, for example, information about the management of staff schedules, available healthcare resources (operating rooms, magnetic resonance imaging machines, etc.), and the organizational structure (names of professionals, availability, specialty areas, etc.). This was important for the subsequent assignment of clinical process tasks to the members of the organization and the time scheduling of tasks according to available resources.
- **DE-06. Assigning roles to tasks**. Support for assigning healthcare organization roles to the different tasks or activities in the CPG. Tasks or activities to be carried out by nursing staff, for example, would be assigned the nursing staff role.
- **DE-07. SLA support**. The system had to allow the pre-establishment of a given level of service quality or SLA (Service Level Agreement) and its relationship with key points in the CPG where that level could be measured.
- **DE-08. Handling of technical errors**. The system had to provide tool to control or manage possible technical errors caused during the execution of clinical processes, such as system crashes, connection errors with other services that block execution, freezing of the tool, etc.
- **DE-09. Automatic compilation of models**. The system had to allow automatic translation of the modeled CPG to an executable model interpretable by a computer system.
- **DE-10. Backup support**. The system had to incorporate a method for making backup copies of the data used in the clinical process.

2.2.2.3. Deployment criteria. It was important for the system to be able to deploy flexibly and be used more extensively within the organization's IT infrastructure. The factors that would influence the use of the system included the size of the organization, the number of patients to be attended or, for example, the workload of the healthcare workers, so it was necessary for the tool to have mechanisms to be able to adapt to these situations.

- **DP-01. Distributed environments**. The system had to be executable on several machines so that it could be used in distributed environments to guarantee availability in case of error or to balance loads and avoid possible saturation.
- **DP-02. Integration with external systems**. The system had to offer facilities for integration with other systems in the healthcare institution.
- **DP-04. External access**. The system had to be accessible from other machines either via a local network or through Internet.
- **DP-04. Web interface**. The system had to have a web-based GUI that could be accessed by users from a web browser.

2.2.2.4. Execution and operational criteria. This section covers those features that would make it possible to successfully execute the CPG. It is in this execution phase where the task flows established in the guideline model would be implemented, depending on the different events, the clinical data of the patient, and/or interactions with the healthcare professionals. Relevant system criteria here would include the next features:

- **EX-01. Personal task management**. The system had to allow users to manage information about their own tasks or activities (task queries, completion deadlines, information about themselves provided in the tasks already performed, etc.).
- **EX-02. Notifications**. The system had to allow notifications or warnings to be sent directly to users via e-mail, SMS or by other means.
- **EX-04. Version control**. The system had to offer the possibility of executing multiple versions of the same CPG model, with each version presenting different changes that could be fully monitored (versioning and version control).
- **EX-04. Flexibility at runtime**. The system had to allow changes to be made in the task flow or activity flow of a CPG currently being instantiated. For example, the user should be able to choose an alternative flow to the one proposed, alter the order of the tasks to be performed, or even add a task to the flow once the clinical process has started.
- **EX-05. Optimizing the process at runtime**. The system had to allow the execution of clinical activity flows to be optimized in line with specific criteria, for example by changing the process or activity flow structure of a running instance to respond to special situations, such as emergency room saturation.
- **EX-06. Changing task manager at runtime**. The system had to offer the possibility of changing the role assigned to a task or activity corresponding to a clinical process in execution. This way, a task assigned to the "physician" role, for example, could be reassigned to the "nurse" role during the execution of the process.
- **EX-07. Execution breadcrumb**. The system had to be able to display the path followed by a patient in a CPG model, so that it would be possible to track the clinical decisions taken at each stage and corroborate the corresponding clinical data.
- **EX-08. Simultaneous execution**. The system had to allow several CPG models to be simultaneously executed and accessed during their runtime.

2.2.2.5. Monitoring and control criteria. For the continuous improvement of clinical processes, it was necessary to monitor how the model behaved during the stage in which it was being executed and to keep in mind different aspects at both clinical and technical levels. Ideally, the system should have features that would make it possible to know what resources were available at the health institution, so that they could be appropriately distributed during the process. In this phase, attention was also paid to the system's capacity to deal with clinical and resource-related setbacks.

- **MC-01. Technical monitoring of infrastructure at runtime**. The system had to provide mechanisms for monitoring the IT infrastructure, such as resource consumption, availability, failures, etc.
- **MC-02. Technical monitoring at runtime**. The system had to allow the monitoring of clinical processes during execution, for example by displaying general technical information about the running processes (e.g., the time they have been running, the resources consumed, KPI, etc.).
- **MC-04. Changing rules at runtime**. The system had to allow the modification of clinical rules (the conditions for performing one activity or another) during clinical processes in execution.

- **MC-04. Recovery mechanisms**. The system had to provide mechanisms, tools, or options for recovering from any technical errors that may occur. For example, errors in the CPG engine, errors during the execution of an activity, errors in communications with other systems, etc.
- **MC-05. Real-time performance data**. The system had to offer tools for viewing execution information in real time by means of control panels or informative graphic elements with sections dedicated to relevant data.
- **MC-06. View-based monitoring data**. The system had to allow information on the technical monitoring of the CPG engine and clinical data to be viewed from different data perspectives.
- **MC-07. Manual workload distribution**. The system had to offer mechanisms for sharing the workload (tasks or activities) between the organization's staff. This would avoid, for example, activity overload for some professionals and undertaking for others.
- **MC-08. Automatic workload distribution**. The system had to offer mechanisms for automatically sharing the workload (tasks or activities) between users according to given criteria, for example by altering professionals' task assignments during vacation periods.

2.2.2.6. Analysis criteria. In the analysis phase, desirable features were identified that would make it possible to view and evaluate the data collected during the execution of the process. To raise the quality of clinical processes and improve the system's technical performance, the following criteria were established.

- **AN-01. Clinical reports**. The system had to allow the generation of medical reports related to the clinical processes followed by patients.
- **AN-02. Historical execution technical data**. Execution records should be kept of things like the time spent on a process, calculated performance indicators, application data, and the status of the IT infrastructure.
- **AN-04. Reuse of clinical data**. The system had to allow the reuse of data that have been used in previous versions of a CPG model, even though this one has not been fully completed.
- **AN-04. Historical clinical data**. The system had to keep a record of historical clinical data used during the execution of CPG models, making it possible to retrieve all the information about a process followed by a patient.
- **AN-05. Optimization recommendations**. The system had to offer suggestions for improving the flow of clinical processes, for example by identifying possible tasking bottlenecks during execution, proposing the use of more resources, or altering the order of tasks to mitigate the situation.
- **AN-06. BI & process mining Support**. The system had to offer facilities for connecting data with business intelligence and/or process mining analysis tools.

2.2.2.7. Other features. This section includes criteria for system features that were not included in the other categories, such as commercial support and the provision of sufficient didactic material for users.

- **OF-01. Commercial support**. The developer or supplier of the tool had to offer commercial support, assistance with new features and advice on its installation within the IT infrastructure.
- **OF-02. Manuals and training**. The system had to have documentation for learning how to use it or offer other options for doing so. These might include video training, case studies, examples, webinars or community events on the tool itself.
- **OF-04. Documentation generation**. The system had to have tools or options to create documentation about the use of the tool itself, for example in PDF or HTML format.
- **OF-04. Maturity**. Time since the first version of the system was published, and date of latest version. It had to be possible to chart the development of the product from its release.

2.2.3. Rating method

After establishing the above-described quality criteria for the model, it was necessary to establish a scoring method with which to homogenously assess each CPG-derived CDSS. For this purpose, a quantitative evaluation method was used for each criterion and category of criteria, allowing homogeneous comparison on three levels:

- **Basic Scoring (BS)**. This was related to the basic feature score and was obtained by assigning an integer score based on a numerical scale of zero to four [0..4], where: 4 points meant that the CPG-derived CDSS provided full native support; 3 points meant partial native support; 2 points mean that the CPG-derived CDSS included programming interfaces that supported the evaluated feature; 1 point meant that a third party component was necessary to support the feature; and 0 points meant that the CPG-derived CDSS did not support the feature.
- **Partial Score (PS)**. This score is associated with each category of features and it is calculated by adding the basic scores of each feature belonging to a specific category, dividing the result by the maximum score and, later, multiplying the result by 10. Eq. (1) represents this calculation; where *k* means a specific tool under evaluation, *j* means a specific category belonging to our quality model, *n* represents the number of features in the category *j*, *BS(k,j,i)* represents the score of the feature *i* (belonging to category *j* for tool *k*), and *max(BS(k,j)* represents the maximum basic score of the category *j* for tool *k*.

$$PS(k,j) = \frac{\sum_{i=1}^{n} BS(k,j,i)}{max(BS(k,j))} \cdot 10; \quad k \in \{CPG\ suites\}; \quad j \in \{quality\ model\ categories\} \tag{1}$$

- **Final Score (FS)**. This score represents the final rating of each CPG-derived CDSS and it is calculated by adding partial scores of all the categories. Eq. (2) represents this calculation; where *k* means a specific CPG-derived CDSS under evaluation, *j* means a specific category belonging to our quality model, *m* means the number of categories, *PS (k,j)* represents the partial score of the category *j* for the suite *k*, and *max(PS(k)* represents the total maximum score of all categories for the suite *k*.

$$PS(k) = \frac{\sum_{j=1}^{m} PS(k,j)}{max(PS(k))} \cdot 10; \quad k \in \{CPG\ suites\} \tag{2}$$

3. Results

This section analyzes the most representative CPG-derived CDSS (i.e., GLEE, ArdenSuite, GLARE, DeGel and KnowWE) to illustrate the application of our method and quality model. For this purpose, each technological tool was evaluated and tested on each technological aspect included in Section 2.2.2. The selection of these systems was established after consulting the opinion and experience of technology consultants (who are experts in the application of IT techniques in healthcare) and healthcare professionals who are responsible of Spanish healthcare organizations (e.g., the private Andalusian assisted reproduction clinic Inebir, and Ministry of Health and Social Welfare of the Andalusian Regional Government, among others).

3.1. Evaluation of the GLEE system

GLIF3 Guideline Execution Engine (GLEE) [26] is a clinical guideline execution system capable of interpreting and executing guidelines that are encoded in the representation format known as GLIF3 [27]. GLEE has interfaces that allow it to connect to a medical institution's HIS and access patient data in order to properly implement a clinical guideline [28]. It even has the ability to integrate with clinical event

monitors [26]. GLEE-compatible guidelines can be created with modeling tools like the GLIF Editor or the Protégé-2000 tool [27,28]. During guideline execution, the healthcare staff will receive recommendations based on scientific evidence for treating an illness, together with relevant information about the patient's condition. The GLEE system also allows a record to be made of the actual guideline execution, thus making it possible to later review and analyze the data generated during the process [28].

Finally, Table 1 (included as a table in the supplementary file) presents a summary of the quantitative evaluation carried out in this paper for GLEE using our quality model.

3.2. Evaluation of the ArdenSuite system

ArdenSuite is a platform distributed by Medexter Healthcare [29] that implements Clinical Decision Support (CDS) software solutions [30]. It uses the well-known Arden Syntax [31] to facilitate the representation of medical knowledge in units known as Medical Logic Modules (MLMs). Each of these modules encodes enough knowledge to make at least one clinical decision [32]. The ArdenSuite platform has two main tools: ArdenSuite IDE and ArdenSuite Server [33].

ArdenSuite IDE, a software based on the well-known Eclipse IDE, is used to write and subsequently compile MLMs. It also allows the use of plugins for the visual modeling of the processes [34,35]. ArdenSuite Server runs the compiled MLMs and manages interoperability with other services. Thanks to its comprehensive API [36], it can even interact with Activiti process management software [34]. The developer also offers different demos to test various functionalities of the system [37].

Finally, Table 2 (included as a table in the supplementary file) presents a summary of the quantitative evaluation carried out in this paper for ArdenSuite using our quality model.

3.3. Evaluation of the GLARE system

GLARE (Guideline Acquisition, Representation and Execution) is a clinical guidelines manager [38] developed in collaboration with one of the largest hospitals in Italy [38]. The system has an acquisition module and a clinical guidelines execution module. It thus distinguishes between, on the one hand, the acquisition and formalization of medical knowledge and, on the other, the implementation of this knowledge on a patient. Thanks to the acquisition module, GLARE provides a simple graphical environment for the visual modeling of clinical guidelines [39] while employing mechanisms to automatically check and ensure the consistency of the process being represented.

The GLARE run module executes a guideline for a given patient and includes clinical decision support functionality to help healthcare professionals choose between different therapeutic or diagnostic alternatives for their patients [38].

Finally, Table 3 (included as a table in the supplementary file) presents a summary of the quantitative evaluation carried out in this paper for GLARE using our quality model.

3.4. Evaluation of the DeGeL system

The name Digital Electronic Guidelines Library (DeGeL) may at first sight seem to refer only to a library for storing executable clinical guidelines. However, the DeGel system (or framework) goes beyond that insofar that it comprises not only a guideline storage repository but also a very complete set of tools encompassing the entire guideline development cycle, from the initial guideline in free text through to its conversion into a computer-interpretable guideline and its practical implementation. The DeGel system tools include: (1) Uruz, an application for the semantic tagging of free-text clinical guidelines; (2) Gesher, an application for the graphical specification of clinical guidelines; (3) Vaidurya and DegeLook, applications for searching for and consulting clinical guidelines; and (4) Spock, an application with a clinical guideline execution engine.

Fig. 2. Summary of the global ranking.

Finally, Table 4 (included as a table in the supplementary file) presents a summary of the quantitative evaluation carried out in this paper for DeGeL using our quality model.

3.5. Evaluation of the KnowWe system

KnowWe is a semantic wiki integrated in d3web [40], a platform that facilitates the creation of expert diagnostic systems and provides solutions in different fields of application, including medical and therapeutic diagnosis, diagnosis of technical failures and the monitoring of technical devices. To represent clinical guidelines on this platform, a graphical modeling language called DiaFlux [41] has been developed, which allows guidelines to be expressed as flowcharts. KnowWe functions as a collaborative workspace [42] where each user can write data or information that he/she considers relevant about a topic or domain, and has mechanisms to make those data easily interpretable by a computer. KnowWe is therefore conceived as a tool for creating knowledge bases for any domain. It is also a wiki-based process modeling environment. KnowWe acts as a user interface for interaction with the main module of d3web, the d3web-core engine, which implements reasoning and knowledge persistence components like decision trees, rules (heuristics), diagnostic flowcharts, etc.

Finally, Table 5 (included as a table in the supplementary file) presents a summary of the quantitative evaluation carried out in this paper for KnowWe using our quality model.

4. Discussion and analysis

This section aims to discuss and analyze the advantages and disadvantages of each CPG-derived CDSS, considering the rating method defined in Section 2.2.3. For this purpose, Section 4.1 to Section 4.7 discusses all CPG-derived CDSSs against each other per category of our quality model (specifically, modeling, design, deployment, execution and operation, monitoring and control, analysis and other criteria, respectively). This analyses is presented in Tables 2–3, which summarizes the assessment, scope and degree of support for each CPG-derived CDSS by characteristic. Fig. 2 summarizes the overall ranking of each tool compared to the rest of them. Although our quality model and strategy systematically and methodologically allow the comparison of CPG-derived CDSS, it is important to note that it is not possible to determine the best proposal. This choice depends on the requirements and needs of each organization. Finally, Section 4.8 describes some limitations of this study.

Table 2
Summary of evaluations for each CPG-derived CDSS (cont.).

FEATURE	GLEE	Arden Suite	GLARE	DeGel	Know-WE
MODELING CRITERIA	**7,5**	**7**	**7,17**	**7,83**	**6,67**
MO-01. Formal modeling language	4	4	4	4	4
MO-02. Expressivity of the modeling language	4	4	4	4	4
MO-03. Multilanguage support	0	0	0	4	0
MO-04. Modeling tool	4	4	4	4	4
MO-05. Visual editor based on flowcharts	4	4	4	4	4
MO-06. Import/export support	4	4	4	4	4
MO-07. Granularity	4	3	4	4	4
MO-08. Multiview support	4	4	4	4	4
MO-09. Documentary references	2	4	4	3	4
MO-10. Clinical rule modeling	4	4	4	4	4
MO-11. Formal rule language modeling	4	4	0	4	4
MO-12. CPI modeling	0	0	0	0	0
MO-14. Data protection modeling	0	0	0	0	0
MO-14. Standard terminology	3	0	3	4	0
MO-15. Automatic consistency check	4	3	4	0	0
DESIGN CRITERIA	2,5	3	3,75	3	3,5
DE-01. Interoperability	4	4	3	4	1
DE-02. GUI design tools	0	1	0	0	4
DE-04. Data protection mechanisms	0	0	0	0	0
DE-04. User access control	1	4	0	4	4
DE-05. Import organizational info	0	0	4	0	1
DE-06. Assigning roles to tasks	0	0	4	0	0
DE-07. SLA support	0	0	0	0	0
DE-08. Handling of technical errors	1	0	0	0	0
DE-09. Automatic compilation of models	4	3	4	4	4
DE-10. Backup support	0	0	0	0	0
DEPLOYMENT CRITERIA	6,25	8,75	5,63	5,63	6,88
DP-01. Distributed environments	2	2	4	2	2
DP-02. Integration with external systems	4	4	1	0	1
DP-04. External access	4	4	4	4	4
DP-04. Web interface for users	0	4	0	3	4
EXECUTION AND OPERATION CRITERIA	5,63	4,69	6,25	3,44	3,75
EX-01. Personal task management	0	0	3	0	0
EX-02. Notifications	3	4	0	0	0
EX-04. Version control	4	3	3	0	4
EX-04. Flexibility at runtime	3	0	3	3	0
EX-05. Optimizing the process at runtime	0	0	3	0	0
EX-06. Changing task manager at runtime	0	0	4	0	0
EX-07. Execution breadcrumb	4	4	0	4	4
EX-08. Simultaneous execution	4	4	4	4	4

Table 3
Summary of evaluations for each CPG-derived CDSS (final).

FEATURE	GLEE	Arden Suite	GLARE	DeGel	Know-WE
MONITORING AND CONTROL CRITERIA	1,88	1,88	2,81	1,25	1,56
MC-01. Technical monitoring of infrastructure at runtime	1	1	1	1	1
MC-02. Technical monitoring at runtime	3	0	0	0	0
MC-04. Changing rules at runtime	0	4	2	0	0
MC-04. Recovery mechanisms	1	0	0	3	0
MC-05. Real-time performance data	1	1	0	0	4
MC-06. View-based monitoring data	0	0	0	0	0
MC-07. Manual workload distribution	0	0	3	0	0
MC-08. Automatic workload distribution	0	0	3	0	0
ANALYSIS CRITERIA	5,83	3,33	6,67	5	4,58
AN-01. Clinical reports	1	1	1	2	2
AN-02. Historical execution technical data	4	0	4	0	4
AN-04. Reuse of clinical data	4	2	2	2	0
AN-04. Historical clinical data	4	4	4	4	4
AN-05. Optimization recommendations	0	0	4	0	0
AN-06. BI & process mining support	1	1	1	4	1
OTHER CATEGORIES	1,25	5,63	2,5	1,25	5,63
OF-01. Commercial support	0	4	0	0	3
OF-02. Manuals and training	0	3	0	0	1
OF-04. Documentation generation	0	0	0	0	3
OF-04. Maturity	2	2	4	2	2
FINAL SCORE	4,4	4,9	4,97	3,91	4,65

4.1. Modeling criteria

Fig. 3 summarizes the modeling score of each tool compared to the rest of them.

In this case, all the systems analyzed support a formal modeling language, thus allowing a standardized language to be used to represent the elements of a CPG. The modeling languages present in the different tools also have sufficient elements to properly express or represent the different concepts and relationships of the problem domain. However, except for the DeGeL system, none of the other tools are compatible with more than one clinical process modeling language, so they do not allow users to choose which language to use to represent the guidelines.

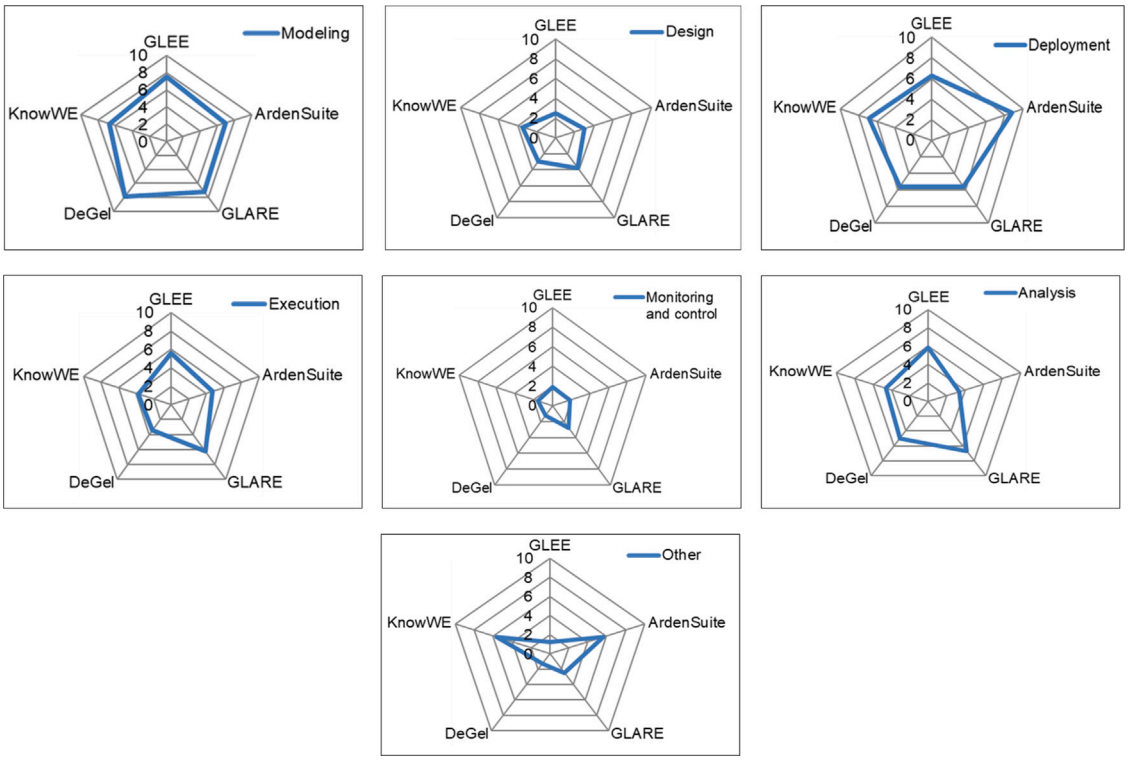

Fig. 3. Summary of the evaluation of each CPG-derived CDSS.

The sole exception, DeGeL, is compatible with languages such as Asbru, GLIF, Arden and GEM.

All the systems evaluated offer tools for modeling clinical processes or incorporate some type of process editor. They all support the use of visual editors for modeling clinical processes by means of flow diagrams and offer mechanisms for importing or exporting clinical guidelines previously modeled in files or other media such as databases, in order to be able to share guidelines with other systems. The systems also support the creation of composite tasks or subprocesses in a clinical process (i.e., the integration of several tasks or elements into one single task), although each system has its own particular name for this type of task: in GLEE they are called "subguides", in GLARE "composite actions", in DeGeL "subplans" and in KnowWe "composite nodes". In ArdenSuite, however, there is no specific term for composite tasks because, thanks to the fact that the system's MLMs can invoke other modules, this functionality is already partially supported.

The systems can all display clinical guideline modeling information. The visual tool contained in GLARE, for example, shows not only a flowchart but also a hierarchical list with the actions to be performed by the user, whereas in KnowWe modeling information can be presented in as many ways as the wiki page format allows.

All the systems evaluated have some kind of mechanism or option for referencing the clinical guideline from which a clinical process was modeled. Some, such as GLARE, with its GLARE-Edu tool, constitute a complete system capable of referencing extracts of the medical and scientific documentation that motivated the system's own clinical decisions. KnowWe, thanks to its own wiki format in which pages containing more information can be added, and ArdenSuite, with a dedicated section in its GUI, offer options for adding medical references. Other systems, however, do not offer such versatility. GLEE, for example, offers only the name and reference of the textual guide that has been modeled, while the only clinical description in DeGeL would be that of the elements in the process.

All the systems allow the modeling of clinical rules, and therefore have mechanisms for controlling process flows. In all tools except GLARE, modeling is done using formal expressions. GLARE differs from other systems in that it uses "punctuated" diagnostic decisions. This means that, unlike the other systems, GLARE does not have a formal expression language with a syntax that allows the modeling of decision criteria. Instead, GLARE models clinical decisions using tabular representation.

No possibility was found in any of the tools analyzed of defining Clinical Process Indicators (CPI). Neither did we find any mechanisms for applying or assisting in the application of patient data protection laws or privacy protocols. The only system that proposes the use of standard medical terminologies and vocabularies is DeGeL, which is able to use LOINC, ICD-9 and CPT. The developers of GLEE and GLARE do not propose any one specific terminology or vocabulary, but the architecture of both systems could be enabled to use any terminology suite. This possibility is not, however, offered in the architecture of ArdenSuite or KnowWe.

Regarding the automatic checking of logical consistency in clinical processes, GLEE and GLARE support such checking during modeling. In ArdenSuite, this functionality is partially supported, as it is possible to logically check the MLMs individually. In DeGeL, this feature is not mentioned and KnowWe does not support it.

4.2. Design criteria

Fig. 3 summarizes the modeling score of each tool compared to the rest of them.

GLEE, ArdenSuite and DeGeL natively support general connectivity with medical systems or services to obtain the data needed in a clinical process. GLARE partially supports this feature because it can only connect to the specific EHR (Electronic Health Record) system for which it was programmed, although, according to its developers, its layered design and its use of XML could also allow it to connect to other, different EHR systems. KnowWe is not intended for interconnection with other systems, although its d3web API leaves this possibility open. What KnowWe does fully allow is the customized creation of graphical interfaces for the user to interact with the system itself, something that GLEE, GLARE and DeGeL do not allow. ArdenSuite at least has

several APIs that can facilitate access to the system by external client applications with modifiable graphical interfaces.

None of the systems offer security mechanisms for protecting both the data and the systems themselves. ArdenSuite, DeGeL and KnowWe allow control of user access to the system. GLEE does not allow user access control directly, but its architecture does allow such control to be implemented by the medical applications through which this system is accessed. User access control is not specified in the GLARE system.

GLARE does, however, support the importing of organizational information from a healthcare institution—something that is not offered by GLEE, ArdenSuite or DeGeL. In KnowWe, information about the institution can only be loaded into the system manually on a wiki page. GLARE is also the only system that supports the assignment of roles to clinical process tasks or activities. None of the systems offer support for adding SLAs and linking them to CPIs, or for handling technical errors during the execution of a clinical process. In this regard, only GLEE can provide a minimum of information about possible technical failures because it offers the possibility of viewing process execution traces.

All of the systems allow the automatic translation of the modeled guide into an executable model, but none of them include methods for backing up the data used in a clinical process.

4.3. Deployment criteria

Fig. 3 summarizes the modeling score of each tool compared to the rest of them.

GLARE and DeGeL provide native support for execution in distributed environments. This functionality is not implemented in any of the other systems but could be developed in GLEE and KnowWe, thanks to their client–server architecture, and in ArdenSuite, thanks to its use of web services. Regarding support for integration with other systems, GLEE and ArdenSuite offer defined interfaces that allow them to cover this functionality. GLARE and KnowWe do not natively support integration but some kind of integration could be implemented based on their own internal operation. In DeGeL, however, this option is not even contemplated. All the systems can be accessed from other machines via a network.

ArdenSuite and KnowWe have graphical user interfaces that can be accessed from a web browser. In DeGeL this feature is partially supported because this suite contains tools with web-based versions. GLEE and GLARE have no web version of their graphical user interfaces.

4.4. Execution and operation criteria

Fig. 3 summarizes the modeling score of each tool compared to the rest of them.

GLARE is the only system that, at least partially, allows users to manage information about their own tasks. The others do not offer this option.

ArdenSuite fully supports the sending of alerts or warnings directly to users telematically. GLEE partially supports this, as it requires a clinical event monitor to connect to, but this option is not available in the other systems. With the sole exception of DeGeL, all of the tools offer some type of version control for clinical processes. GLEE, GLARE and DeGeL partially support making changes to the task flow of a running clinical process, but ArdenSuite and KnowWe have no such capability.

GLARE is the only system analyzed that includes support for optimizing the running of a clinical process during its execution according to certain criteria. It is also the only system that allows the role assigned to a task to be changed in a running process. GLEE, ArdenSuite, DeGeL and KnowWe allow the user to see the path followed by a patient in a running clinical process, while in GLARE this feature is not specified. All the tools evaluated allow several clinical processes to be executed and accessed simultaneously.

4.5. Monitoring and control criteria

Fig. 3 summarizes the modeling score of each tool compared to the rest of them.

None of the systems support technical monitoring of the IT infrastructure, although in all of them it is possible to use the resource control mechanisms of the machine on which the system is installed for this purpose. Only GLEE offers partial support for the technical monitoring of clinical processes in execution, showing some basic technical values. The other systems do not contemplate this feature. Regarding the modification of clinical rules in clinical processes in execution, only ArdenSuite presents a comprehensive method for doing this, although it may also be feasible in GLARE thanks to this system's mechanisms for dealing with exceptions. In GLEE, DeGeL and KnowWe it is not possible. On the other hand, DeGeL is the only suite that partially supports recovery from technical errors that occur in the system. ArdenSuite, GLARE and KnowWe all lack this feature. In GLEE, the error recovery mechanism is only informative because it simply tracks the execution trace where it is possible to see in which part of the execution the system has failed.

KnowWe includes a page with data relevant to the real-time execution of the processes, something that is not offered by the other systems. The interfaces defined by GLEE and ArdenSuite suggest that real time monitoring could be implemented in these tools. However, in none of the systems analyzed is it possible to display monitoring information from different views or perspectives.

Regarding mechanisms for distributing the workload among the organization's personnel, the only tool that partially allows this is GLARE, which is also the only suite that allows the automatic distribution of workloads according to pre-specified criteria.

4.6. Analysis criteria

Fig. 3 summarizes the modeling score of each tool compared to the rest of them.

GLEE, ArdenSuite and GLARE do not offer the option of directly generating automatic medical reports on the clinical process followed by a patient, but it would be possible to develop an external application capable of accessing the data stored in these systems in order to create such automatic reports. DeGeL and KnowWe, on the other hand, do provide options for the healthcare professional to include medical comments on the clinical process, although this must be done manually.

GLEE, GLARE and KnowWe keep records of technical data concerning the execution of clinical processes. KnowWe is the only system analyzed that does not provide a way to reuse the clinical data collected in previous versions of a process, while GLEE is the most complete system in this regard because it is able to store the patient's status as an entry point for data reuse. All the tools store the historical clinical data used during the execution of processes. Only GLARE is able to propose suggestions for improving activity flows in clinical processes. Regarding support for business intelligence and/or process mining tools, DeGel is the only one that has its own tool for this task; the other systems offer interfaces through which developers can explore this possibility.

4.7. Other criteria

Fig. 3 summarizes the other criteria score of each tool compared to the rest of them.

The only system that offers full commercial support is ArdenSuite, while KnowWe offers d3web, a part of the system itself. The other systems have no commercial support, or at least it is not specified by their developers. As to whether there is sufficient documentation or options for training in the use of the systems, ArdenSuite offers a series of training manuals on how to use some parts of the tool. KnowWe also presents some information about its use on its developer's website, but this is very limited. Documentation about the GLEE, GLARE and DeGeL

systems is limited to academic articles describing features and advances in the tools that make up the systems, but there is no explanation of how to use them.

Thanks to its wiki format, KnowWe very rudimentarily facilitates the incorporation of user documentation to aid use of the tool. The other systems, however, have no mechanisms or options for creating relevant documentation of this type.

With respect to the maturity of the systems analyzed, understood as the time elapsed between the publication of the first version of the system and the date of its latest version, GLARE can be said to be the tool with the most stable, most continuous development over time in comparison with the other systems.

4.8. Threats to validation

Below, some threats to the validity of this research paper are presented. On the one hand, the quality model presented in this paper was instantiated to evaluate specific tools. The selection of these tools may constitute another threat to validity because it was influenced by the keywords used in the different digital libraries (c.f., Section 2.1). These keywords may have returned an insufficient set of results. This was mitigated by carrying out two iterations of the search protocol. The first iteration produced preliminary results whereas the second made it possible to refine the keywords used in the first iteration. On the other hand, although the quality model and its characterization scheme were considered by the authors to be a valid contribution of this research paper, some readers may see this contribution as being subjective rather than objective. This was mitigated by unifying three quality models (each researcher in the study created his/her own quality model), which were then shared and merged. The quality model was also reviewed by ICT experts within the healthcare environment. This strategy made it possible to jointly identify as many features as possible.

5. Related works

When identifying and analyzing related works, no research publications were found which specifically featured CPG-derived CDSS quality and characterization models. However, a few studies and discussions include more limited comparisons. The most relevant works in this field are briefly described below.

Greenes et al. [15] discuss critical aspects that affect the application of CDSS approaches to support clinical guidelines. Specifically, authors identify and quantify success and failure factors to develop and adopt new approaches to addressing theoretical models and frameworks for CDSS. Although these aspects are superficially discussed, authors provide good wishes for CDSS and valuable future paths for new related research: (a) integration/adaptation to workflow; (b) construction of CDSS artifacts; (c) knowledge management, interoperability, and sharing; (d) cognitive tasks/reasoning processes to be supported; (e) health system priorities and implementation (or adoption) paradigms; (f) quality improvement impacts; and (g) evaluation of effectiveness of a CDSS intervention. The challenges of each aspect are discussed individually, but authors mention that it is possible to intuit a sort of loose lifecycle connecting these considerations, in that a developer may find appropriate models or frameworks addressing these aspects to be useful on stages of CDSS conceptualization, design, modeling, formalizing, integrating into workflow, deploying, and evaluating. However, authors acknowledge that this lifecycle is only an intuition that is not validated and does not cover other aspects such as formal approaches to requirements elicitation, information presentation to support visualization and cognition, data quality issues, explicit consideration of confidentiality and security especially in integrating multiple data sources as inputs to a CDSS, among others.

Dadam et al. [43] carried out a state-of-the-art study to identify desirable features of process-oriented HIS supporting healthcare workflows. Here, the authors discussed and identified six features related to the expressiveness (control flow, temporal constraints and data flow), verification and consistency of models, dynamic flexibility at runtime, integration with HISs, and control of time dependencies between tasks. They also described how these features were applied in a real project that they executed, although no comparisons were made between different systems to evaluate them.

Sim et al. [44] discussed CDSS and their practical application, and later proposed a taxonomy to evaluate functional and design aspects of those systems. In this sense, the authors defined 24 characteristics, grouped into 5 categories related to the clinical management context, data and knowledge sources, decision support, information delivery and workflow. However, the proposed taxonomy was presented theoretically, without quantitative evaluation mechanisms, and was not instantiated.

Isern et al. [7] carried out a systematic review to identify and compare CPG-derived CDSS. They analyzed 8 systems providing execution capabilities. The comparison focused on 11 aspects related to modeling aspects, technological architecture, integration and coordination of the system with external elements, type of execution engine, security aspects and the use of standard terminology.

Gooch et al. [45] carried out a systematic review to identify challenges in designing and developing process-oriented HIS supported by and integrated with formal models of clinical guidelines and care workflows. Having completed the review, the authors identified 25 common features in the studies analyzed. These features were closely related to features such as the integration of models into individual and collaborative clinical workflow systems, to the possibility of checking a model with reasonable run-time behavior, the mapping of EHR data to procedural tasks in the guideline or pathway, the reporting of features, flexibility, pathway adaptability at run-time, and usability.

What is common in these works, and in many similar studies is that the features to be evaluated were presented in isolation and were not associated with the clinical guideline management lifecycle. This definition is not a wrong strategy in itself, however, we think any set of characteristics should be contextualized within the CPG management lifecycle to facilitate its evaluation considering the needs and requirements of each healthcare entity. This hypothesis was corroborated by IT experts in healthcare who were consulted. This fact prompted that our own paper establishes a grouping by phase in the CPG digitalization lifecycle (as described in Section 4.2).

Finally, one important difference between the studies mentioned above and our own work is that we propose a quality module based on a methodological framework with which we propose to evaluate systems in a systematic way, combining quantitative method and qualitative analysis. The application of this quality model in practice, and feedback obtained from IT experts in healthcare, provides an objective and quantitative indicator of the functional and technical benefits that each CPG-derived CDSS offers.

6. Conclusions and future works

Clinical guidelines are useful tools for standardizing existing clinical knowledge in a specific context. The main handicap of CPGs is their representation, which is usually textual. This causes ambiguity and variability when CPGs are applied in clinical practice by health professionals. There currently exist software systems that could help clinicians to improve CPG automation, but the use of these systems is a just a first step towards their widespread use in medical organizations.

In this context, due to the variability of functionalities supported by different CPG-derived CDSS, this paper proposes a method based on quality models to uniformly compare these systems by highlighting each phase of CPG digitalization lifecycle. For this purpose, our method integrates qualitative and quantitative analysis of technological aspects to evaluate and compare CPG-derived CDSS. The ultimate goal of this proposal is to provide healthcare stakeholders information regarding

the strategic adoption of a CPG-derived CDSS into their healthcare organizations according to their objectives. This paper also presents how our proposal has been instantiated to compare five currently available tool (GLEE, ArdenSuite, GLARE, DeGeL, and KnowWe). This evaluation was carried out objectively and uniformly testing each technological tool on each technological aspect included in our quality model.

After carrying out this study, it is possible to observe that, at present, CPG support systems pose challenges in daily practice in healthcare institutions because of these systems have a series of shortcomings related to their integration into EHR systems. Specifically, we have observed that each CPG-derived CDSS offers its own non-standardized integration mechanisms with clinical flows, which causes ad hoc integration solutions with complex maintenance and with poor (or null) flexibility when a change on clinical recommendations or clinical process is required by healthcare professionals to improve healthcare. To address this limitation, it might be necessary to offer highly granular definition mechanisms, which allow to define unambiguously and exactly the correct clinical information offered to the right healthcare staff at the right time (to improve decision-making).

Additionally, after evaluating each CPG-derived CDSS, we have observed great limitations to align the healthcare process models with the actual care and attention to patients. Each tool provides its own modeling mechanisms, which are often insufficient due to difficulty in determining where and when the CPGs are triggered in an actual patient. For example, it is difficult to model how the current state of the patient is determined, how healthcare professionals interact with the system at the right time of patient care, and how the clinical recommendations are integrated into the healthcare process.

Moreover, other conclusions were obtained. In summary, the final scores obtained for the different systems are diminished by a lack of mechanisms or options for system monitoring. This not only affected the scoring of the monitoring phase criteria, but also had indirect consequences for other criteria in different phases (for example, system design criteria). The general lack of support from system developers also drastically reduced the scores. On the other hand, the highest scoring criteria were those associated with modeling. If we associate the different scores obtained by the systems with the particular characteristics that they present, it can be seen that GLEE scored high in criteria related to the process engine itself, standing out from other systems in both the modeling and execution of clinical processes. ArdenSuite, which scored well in modeling and deployment, stands out from the other systems due to its commercial support, which basically means that there is a company behind its development. The GLARE system, with the highest overall rating, is the only one that allows role assignment. It also gained points thanks to its support for optimizing processes during execution. DeGeL, on the other hand, differs from the other systems in that it offers a series of tools that allow clinical guidelines to be modeled progressively, from modeling to analysis. This makes it the only system to include tools for analyzing clinical data after a process has been executed. Finally, KnowWe stands out for its ability to add information, either textual or in other formats, to the system itself thanks to its implementation as a wiki system. It is also the only system to feature a unique modeling language developed from scratch to take advantage of its reasoning engine.

Moreover, in future works, we will continue to study the main features of CPG-based execution systems to improve our quality model. We will also present new comparative studies evaluating other possible CPG support systems. This paper is a starting point from which to study how these systems can reduce ambiguity and variability when clinical guidelines are automated in real health contexts. In addition, we will take our paper as initial point to study what CPG-derived CDSS could be the best one for a kind of healthcare organism (large or small entity). In this sense, we also plan to adapt our scoring method, allowing to set weights to each characteristic and functional category. In this way, our method can be adapted according to the needs and requirements of the healthcare organization in which the CPG-based execution system is to be implemented. Finally, we plan to develop a new model, which allows to express the characteristics needed to improve these systems or to build new ones.

CRediT authorship contribution statement

Julián Alberto García-García: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Manuel Carrero:** Writing – original draft. **María José Escalona:** Methodology, Supervision. **David Lizcano:** Methodology, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research paper has been carried out by the first author during his international postdoctoral stay at School of Computer Science, Madrid Open University (UDIMA) Madrid, Spain. In addition, it is a pleasure to mention that this research is partially supported by: (1) the GIMO-PD (RTC2019-007150-1) project of the Spanish Ministry of Economy and Competitiveness, financed by European funds; (2) the NICO project (PID2019-105455GB-C31) of the Spanish Ministry of Economy and Competitiveness; and (3) the EQUAVEL project (PID2022-137646OB-C31), funded by MCIN/AEI/10.13039/501100011033/ and by the ‘‘European Union’’.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jbi.2023.104573.

References

[1] E. Steinberg, S. Greenfield, D.M. Wolman, M. Mancher, Clinical Practice Guidelines We Can Trust, National Academies Press, 2011,
[2] J. Van Den Bergh, S. Thijs, S. Viaene, Transforming Through Processes: Leading Voices on BPM, People and Technology, Springer, 2014.
[3] Y. Alotaibi, Business process modelling challenges and solutions: A literature review, J. Intell. Manuf. 27 (4) (2016) 701–723.
[4] A. Van Looy, A quantitative and qualitative study of the link between business process management and digital innovation, Inf. Manag. 58 (2) (2021) 103413.
[5] T. Schmiedel, J. vom Brocke, Business process management: Potentials and challenges of driving innovation, in: BPM - Driving Innovation in a Digital World, Springer, Cham, 2015, pp. 3–15.
[6] G.C. Kane, Digital innovation lights the fuse for better health care outcomes, MIT Sloan Manag. Rev. 59 (1) (2017).
[7] D. Isern, A. Moreno, Computer-based execution of clinical guidelines: A review, Int. J. Med. Inf. 77 (12) (2008) 787–808.
[8] R. Lenz, R. Blaser, M. Beyer, O. Heger, C. Biber, M. Bäumlein, M. Schnabel, IT support for clinical pathways-lessons learned, Int. J. Med. Inf. 76 (2007) S397–S402.
[9] F. Rutten, W. Brouwer, L. Niessen, Practice guidelines based on clinical and economic evidence. Indispensable tools in future market oriented Health care, Eur. J. Health Econ. 6 (2) (2005) 91–93.
[10] Evidence-Based Medicine Working Group, et al., Evidence-based medicine. A new approach to teaching the practice of medicine, JAMA 268 (17) (1992) 2420.
[11] M. Peleg, Computer-interpretable clinical guidelines: A methodological review, J. Biomed. Inf. 46 (4) (2013) 744–763.
[12] S. Van de Velde, A. Heselmans, N. Delvaux, L. Brandt, L. Marco-Ruiz, D. Spitaels, H. Cloetens, T. Kortteisto, P. Roshanov, I. Kunnamo, et al., A systematic review of trials evaluating success factors of interventions with computerised clinical decision support, Implement. Sci. 13 (1) (2018) 1–11.
[13] T. Rawson, L. Moore, B. Hernandez, E. Charani, E. Castro-Sanchez, P. Herrero, B. Hayhoe, W. Hope, P. Georgiou, A. Holmes, A systematic review of clinical decision support systems for antimicrobial management: Are we failing to investigate these interventions appropriately? Clin. Microbiol. Infect. 23 (8) (2017) 524–532.
[14] Y. Shi, A. Amill-Rosario, R.S. Rudin, S.H. Fischer, P. Shekelle, D.P. Scanlon, C.L. Damberg, Barriers to using clinical decision support in ambulatory care: Do clinics in health systems fare better? J. Am. Med. Inf. Assoc. 28 (8) (2021) 1667–1675.

[15] R.A. Greenes, D.W. Bates, K. Kawamoto, B. Middleton, J. Osheroff, Y. Shahar, Clinical decision support models and frameworks: Seeking to address research issues underlying implementation successes and failures, J. Biomed. Inf. 78 (2018) 134–143.
[16] T. Greenhalgh, J. Wherton, C. Papoutsi, J. Lynch, G. Hughes, C. A'Court, S. Hinder, N. Fahy, R. Procter, S. Shaw, Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies, J. Med. Internet Res. 19 (11) (2017).
[17] J. Cho, The impact of post-adoption beliefs on the continued use of health apps, Int. J. Med. Inf. 87 (2016) 75–83.
[18] M. Zhang, M. Luo, R. Nie, Y. Zhang, Technical attributes, health attribute, consumer attributes and their roles in adoption intention of healthcare wearable technology, Int. J. Med. Inf. 108 (2017) 97–109.
[19] P. Brereton, B.A. Kitchenham, D. Budgen, M. Turner, M. Khalil, Lessons from applying the systematic literature review process within the software engineering domain, J. Syst. Softw. 80 (4) (2007) 571–583.
[20] J.B. Hill, J. Sinur, D. Flint, M.J. Melenovsky, Gartner's position on business process management, Gartner Res. 136533 (2006).
[21] SNOMED Internacional, Systematized nomenclature of medicine - clinical term (SNOMED CT), 2022, URL https://www.snomed.org/. (Last access: 25 July 2022).
[22] Regenstrief Institute l, Logical observation identifiers names and codes (LOINC), 2022, URL https://loinc.org/. (Last access: 25 July 2022).
[23] J.A. Hirsch, T.M. Leslie-Mazwi, G.N. Nicola, R.M. Barr, J.A. Bello, W.D. Donovan, R. Tu, M.D. Alson, L. Manchikanti, Current procedural terminology; A primer, J. Neurointervent. Surg. 7 (4) (2015) 309–312.
[24] D.J. Cartwright, ICD-9-CM to ICD-10-CM codes: what? why? how?, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, 2013.
[25] Regenstrief Institute, Unified medical language system (UMLS), 2022, URL https://loinc.org/. (Last access: 25 July 2022).
[26] D. Wang, E.H. Shortliffe, GLEE–A model-driven execution system for computer-based implementation of clinical practice guidelines, in: Proceedings of the AMIA Symposium, American Medical Informatics Association, 2002, pp. 855–859.
[27] A.A. Boxwala, M. Peleg, S. Tu, O. Ogunyemi, Q.T. Zeng, D. Wang, V.L. Patel, R.A. Greenes, E.H. Shortliffe, GLIF3: A representation format for sharable computer-interpretable clinical practice guidelines, J. Biomed. Inform. 37 (3) (2004) 147–161.
[28] D. Wang, M. Peleg, S.W. Tu, A.A. Boxwala, O. Ogunyemi, Q. Zeng, R.A. Greenes, V.L. Patel, E.H. Shortliffe, Design and implementation of the GLIF3 guideline execution engine, J. Biomed. Inform. 37 (5) (2004) 305–318.
[29] Medexter Healthcare, ArdenSuite support, 2020b, URL https://www.medexter.com/ardensuite-support/. (Last Access: 9 July 2022).
[30] Medexter Healthcare, Learning center, 2022f, URL https://www.medexter.com/products-and-services/learning-center. (Last Access: 8 July 2022).
[31] J.S. de Bruin, K.-P. Adlassnig, H. Leitich, A. Rappelsberger, Separating business logic from medical knowledge in digital clinical workflows using business process model and notation and arden syntax, in: Health Informatics Meets EHealth, IOS Press, 2018, pp. 17–24.
[32] M. Samwald, K. Fehre, J. De Bruin, K.-P. Adlassnig, The Arden Syntax standard for clinical decision support: Experiences and directions, J. Biomed. Inf. 45 (4) (2012) 711–718.
[33] Medexter Healthcare, How to compile, test and deploy MLMs with the ArdenSuite, 2019, URL https://www.medexter.com/ardensuite/How_To_Compile_Test_Deploy_MLMs.pdf. (Last Access: 10 July 2022).
[34] Medexter Healthcare, Achieving system connectivity between activiti BPMN platform and ArdenSuite, 2018, URL https://www.medexter.com/ardensuite/ArdenSuite_Extension_for_Activiti_How_To.pdf. (Last Access: 9 July 2022).
[35] M. Spineth, A. Rappelsberger, K.-P. Adlassnig, Achieving interoperability between arden-syntax-based clinical decision support and openEHR-based data systems, Stud. Health Technol. Inf. 248 (2018) 338–344.
[36] Medexter Healthcare, ArdenSuite support. APIs, 2021a, URL https://www.medexter.com/ardensuite-support/server/apis/. (Last Access: 9 July 2022).
[37] Medexter Healthcare, Demos, 2022a, URL https://www.medexter.com/demos. (Last Access: 15 July 2022).
[38] P. Terenziani, S. Montani, A. Bottrighi, M. Torchio, G. Molino, G. Correndo, The GLARE approach to clinical guidelines: main features, in: Computer-Based Support for Clinical Guidelines and Protocols, IOS Press, 2004, pp. 162–166.
[39] P. Terenziani, S. Montani, A. Bottrighi, M. Torchio, G. Molino, L. Anselma, G. Correndo, Applying artificial intelligence to clinical guidelines: The GLARE approach, in: AI*IA 2003: Advances in Artificial Intelligence, Springer, 2003, pp. 536–547.
[40] D3web, DOC FaQ, 2022b, URL https://www.d3web.de/Wiki.jsp?page=Doc%20FAQ#top. [Last Access: 16 July 2022].
[41] R. Hatko, J. Baumeister, V. Belli, F. Puppe, DiaFlux: A graphical language for computer-interpretable guidelines, in: International Workshop on Knowledge Representation for Health Care, Springer, 2011, pp. 94–107.
[42] J. Baumeister, J. Reutelshoefer, V. Belli, A. Striffler, R. Hatko, M. Friedrich, KnowWE - A Wiki for knowledge base development, in: The 8th Workshop on Knowledge Engineering and Software Engineering, KESE2012, 2012.
[43] P. Dadam, M. Reichert, K. Kuhn, Clinical Workflows-the killer application for process-oriented information systems, in: BIS 2000, Springer, 2000, pp. 36–59.
[44] I. Sim, A. Berlin, A framework for classifying decision support systems, in: AMIA Annual Symposium Proceedings, 2003, pp. 599–603.
[45] P. Gooch, A. Roudsari, Computerization of workflows, guidelines, and care pathways: A review of implementation challenges for process-oriented health information systems, J. Am. Med. Inf. Assoc. 18 (6) (2011) 738–748.