Computerized Clinical Decision Support: Contributions from 2015

V. Koutkias[1,2], **J. Bouaud**[3,4], **Section Editors for the IMIA Yearbook Section on Decision Support**
[1] Institute of Applied Biosciences, Centre for Research & Technology Hellas, Thermi, Thessaloniki, Greece
[2] Lab of Computing and Medical Informatics, Department of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
[3] AP-HP, Department of Clinical Research and Development, Paris, France
[4] INSERM, Sorbonne Université, UPMC Univ Paris 06, Université Paris 13, Sorbonne Paris Cité, UMRS 1142, LIMICS, Paris, France

Summary

Objective: To summarize recent research and select the best papers published in 2015 in the field of computerized clinical decision support for the Decision Support section of the IMIA yearbook.

Method: A literature review was performed by searching two bibliographic databases for papers related to clinical decision support systems (CDSSs) and computerized provider order entry (CPOE) systems. The aim was to identify a list of candidate best papers from the retrieved papers that were then peer-reviewed by external reviewers. A consensus meeting between the two section editors and the IMIA editorial team was finally conducted to conclude in the best paper selection.

Results: Among the 974 retrieved papers, the entire review process resulted in the selection of four best papers. One paper reports on a CDSS routinely applied in pediatrics for more than 10 years, relying on adaptations of the Arden Syntax. Another paper assessed the acceptability and feasibility of an important CPOE evaluation tool in hospitals outside the US where it was developed. The third paper is a systematic, qualitative review, concerning usability flaws of medication-related alerting functions, providing an important evidence-based, methodological contribution in the domain of CDSS design and development in general. Lastly, the fourth paper describes a study quantifying the effect of a complex, continuous-care, guideline-based CDSS on the correctness and completeness of clinicians' decisions.

Conclusions: While there are notable examples of routinely used decision support systems, this 2015 review on CDSSs and CPOE systems still shows that, despite methodological contributions, theoretical frameworks, and prototype developments, these technologies are not yet widely spread (at least with their full functionalities) in routine clinical practice. Further research, testing, evaluation, and training are still needed for these tools to be adopted in clinical practice and, ultimately, illustrate the benefits that they promise.

Keywords

Medical informatics, International Medical Informatics Association, yearbook, clinical decision support systems

Introduction

Decision support constitutes a mainstream topic in Medical Informatics as illustrated by the high number and the variety of contributions published in the field every year. In the tradition of the Yearbook of the International Medical Informatics Association (IMIA), this literature review performed for the Decision Support section was targeted to papers published in 2015 related to clinical decision support systems (CDSSs) and computerized provider order entry (CPOE) systems. The goal of this review was to identify a handful of high-quality papers considered as the best papers of the year in the decision support field. No particular focus was given in our review, unlike the survey paper of the Decision Support section of the current IMIA yearbook by Coiera et al. [1], in which the unintended consequences of health information technologies (HIT) and decision support systems in particular are discussed.

The next section of the synopsis briefly presents the best paper selection process with a special note on the modifications to the search strategy that were applied this year as compared to the previous years, in order to increase the specificity of the obtained citations. The following section provides further details of the review process as well as our results in quantitative terms. The last section concludes this synopsis and discusses noticeable characteristics of the works presented in the four selected best papers, emphasizing their contribution to the field of decision support, while also reporting other interesting publications spotted during the review process.

Paper Selection Method

We performed a comprehensive literature search by following an established protocol, which has been applied for the last three years [2]. The search targeted topics related to computerized clinical decision support and CPOE. Queries were developed for two bibliographic databases, namely, PubMed/Medline (from NCBI, National Center for Biotechnology Information) and Web of Science® (WoS, from Thomson Reuters). Besides terms describing the domain of interest, our search criteria only included journal papers published in 2015 (even only electronically, for PubMed in particular) that contained an abstract written in English. The retrieved references were then reviewed by the section editors who concluded with fifteen candidate best papers. Then, these candidate best papers were externally reviewed and rated. Finally, among these rated papers, the Yearbook editorial committee decided to keep some of them as best papers.

In order to improve specificity and thus reduce the number of references to review this year, we slightly modified the search strategy used formerly and adapted the queries accordingly. From our past experience, we observed that WoS query results contained many references that were also indexed in Medline but which had not been retrieved by the PubMed query. Considering first that the Medline database is focused on biomedical literature and includes most of it, and second, that our PubMed query had been fine-tuned in the past years, we assumed that all Medline citations not retrieved by the PubMed query were true negatives, not to be included in the review. In contrast, the WoS database has a broader scope and is not limited to biomedical literature though including it. Searching WoS for our review is interesting for targeting references in the literature outside Medline. As a consequence, we excluded all the Medline-indexed citations from our new WoS search, because they would have been either true-positives, already returned by the PubMed query, or false-positives, only returned by the WoS query and considered as "noise". The final citation set was then made out of the union of the Medline hits, obtained with our PubMed search, and of the hits from the non-Medline literature, but WoS-indexed, obtained by the new WoS strategy.

Review Results

The abovementioned databases were searched on January 8, 2016. 869 references originated from PubMed and 112 from WoS, with none in common. After removing 7 references which had already been considered in the selection procedure of the previous year, a total of 974 references remained. Notably, the new WoS strategy (explained in the previous section) eliminated more than 80% of Medline-indexed references and, in comparison with the previous year (1,254 references [3]), we obtained 282 less references to review. All articles were separately reviewed by the two section editors. During the review process, described in [2], articles were evaluated according to their contribution to diverse DSS topics, e.g., applications and tools, methodology

and design, evaluation studies, experiments as well as methodological reviews. Merging the two reviews identified 28 articles kept aside by at least one section editor. These 28 references were then jointly reviewed by the two section editors to select a consensual list of 15 candidate best papers. Following the IMIA Yearbook best paper selection process, these 15 papers were then peer-reviewed by external reviewers and the Yearbook editors. Four papers were finally selected as best papers. They are discussed in the next section in the order they appear in Table 1. A content summary of these best papers can be found in the Appendix of this synopsis.

Discussion and Outlook

The first paper by Anand et al. [4] reports on the CHICA system, a pediatric decision support system, ten years after its implementation in the practices of the Eskenazi Health System in Indianapolis, Indiana, US. CHICA stands for Child Health Improvement through Computer Automation system, and implements pediatric preventive guidelines. There have been many publications on this system since the first one in 2004 [5]. There are, at least, two characteristics that make CHICA a remarkable CDSS. Firstly, this CDSS has been in routine use for more than a decade serving over 44,000 patients and 755 healthcare professionals in 7 hospital sites. Few guideline-based CDSSs, CPOEs set aside, may claim such routine usage. Secondly, CHICA knowledge bases have been

encoded using the Arden Syntax. The Arden Syntax is a computer-interpretable formalism for representing pieces of medical knowledge as rules, or medical logical modules (MLMs), and for executing these rules in the context of a hospital information system to provide automated decision support. Developed since 1991, it became a standard promoted by the HL7 organization. The paper by Anand et al. [4] is part of a special issue of the Artificial Intelligence in Medicine journal, dedicated to the Arden Syntax for its 25th anniversary. While many papers have been published mentioning the Arden Syntax (119 references in PubMed by June 23th, 2016), the majority of them are about prototypes, feasibility studies, extensions to the language, but CHICA adds the value of effective routine use. The authors provide here an overall description of the CHICA system and the adaptations they performed to the Arden Syntax for their purpose. CHICA is connected to the hospital electronic medical record and is also used to feed it. First, its use is directed toward the patient for collecting pre-screening data directly from the patient's family with a personalized form depending on the child's characteristics. Second, a physician-directed form is issued, synthesizing patient-specific information to prepare the encounter with the patient and providing guideline recommendations. At present time, 41 pediatric clinical issues are covered by CHICA and a total of 429 MLMs have been developed, since the implementation of the system in 2003, where it started with a library of about 200 MLMs. Authors report that during this period, MLMs were fired more than ten

Table 1 Best paper selection of articles for the IMIA Yearbook of Medical Informatics 2016 in the section 'Decision Support'. The articles are listed in alphabetical order of the first author's surname.

Section
Decision Support
▪ Anand V, Carroll AE, Biondich PG, Dugan TM, Downs SM. Pediatric decision support using adapted Arden Syntax. Artif Intell Med 2015 Oct 1.
▪ Cho I, Lee JH, Choi SK, Choi JW, Hwang H, Bates DW. Acceptability and feasibility of the Leapfrog computerized physician order entry evaluation tool for hospitals outside the United States. Int J Med Inform 2015 Sep;84(9):694-701.
▪ Marcilly R, Ammenwerth E, Vasseur F, Roehrer E, Beuscart-Zéphir MC. Usability flaws of medication-related alerting functions: A systematic qualitative review. J Biomed Inform 2015 Jun;55:260-71.
▪ Shalom E, Shahar Y, Parmet Y, Lunenfeld E. A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines. Int J Med Inform 2015 Apr;84(4):248-62.

million times. Some adaptations to the Arden Syntax were performed to successfully implement CHICA's functionalities. These modifications are considered minor, and concern the syntax, the parser, and its parallel execution framework. As one of the most studied CDSS for child health care, the CHICA system demonstrated that the Arden Syntax standard is effective for representing some preventive pediatric guideline knowledge and for delivering decision support in real clinical settings. Moreover, as the authors report, the system's effectiveness has already been assessed in several published randomized controlled trials with positive outcomes.

The second paper, by Cho et al. [6], assessed the acceptability and feasibility of a Leapfrog evaluation in four tertiary and academic hospitals of South Korea. Leapfrog is a systematically-developed tool for evaluating medication-safety related decision support through CPOE systems. In particular, Leapfrog relies its evaluation on testing how the CPOE handles a variety of dangerous medication-ordering scenarios. It includes 11 categories of erroneous medication orders (drug-allergy, drug-route, drug-lab, etc.), as well as various monitoring checks (how well the CPOE generates other information about orders), nuisance orders (whether the CPOE generates warnings or information for slight or inconsequential interactions that may be unnecessary or unwanted), and, finally, a "deception analysis" test category (checking for false positives). Although Leapfrog-based evaluation relies on various simulated clinical scenarios, a study illustrated that its results were close to actual rates of preventable adverse drug events (ADEs).

As the tool has only been available and used in the US, Cho et al. [6] aimed to shed light on unconfirmed CPOE-related issues and effects on patient outcomes in other countries, with potential different healthcare organization and settings. In each of the sites participating to the study, a locally developed CPOE was available and being used for more than 10 years. Interestingly, all sites were regulated by the national drug utilization review process, separating the prescribing and dispensing functions between physicians and pharmacists, respectively.

The study by Cho et al. [6] follows a cross-sectional design, i.e., evaluation was conducted in each hospital and the results were compared among hospitals by measuring: (a) the system response rate, (b) the category completion rate, and (c) the time to complete the evaluation. The scoring system interprets automatically the raw test results reported by the hospitals based on the relative importance of each type of decision support for preventing patient harm, while the score reflects both the severity of a potential ADE (not intercepted by the system) as well as its probability of occurrence. The score indicates test performance across order-checking categories within the following stages: "fully implemented", "good progress in implementing", "good early-stage effort", "completed the evaluation", and "incomplete evaluation". Acceptability was defined as whether the evaluation tool is acceptable to hospitals outside the US, while feasibility was defined as whether the tool is easy to administer and process. The evaluation was conducted in one-week period, requiring approximately eight hours per site. Measurements were calculated anonymously and descriptive rates and proportions were compared among the four sites and then with those of five US community hospitals reported in another study. The overall category completion rates ranged from 67.9% to 75.5%, with a varying degree according to the evaluation category, while the required time to perform the tests was within the allowed test timeframe (between 3.1 to 4 hours). The total evaluation score ranged from 21.6% to 36.5%. Three sites were characterized as "completed the evaluation" and one as "incomplete evaluation stage". In addition, according to the test, three systems could cause severe harm in the "Therapeutic Duplication" category, and one of the three in the "Drug-Allergy" category as well. In comparison with US hospitals, the overall scores of the South Korean systems were lower than the average of the US systems considered in the study, with the two highest scores of the South Korean systems being slightly higher than the lowest score in the US. While the completion rates were above 67% for each system, many differences in terms of error category and system were identified. The authors discussed that this

might be due to the different scope and coverage of CPOE safety performance in the hospitals, despite their similar organizational characteristics. Regarding feasibility, evaluation was tolerable at all four sites. Thus, the study concluded that there is a potential for Leapfrog to be used for hospitals outside the US. While physician acceptability of CPOE is getting higher (80% in the US) and there is a clear need for CPOE systems with alerting functions reflected internationally [7], frameworks like Leapfrog are essential in improving such tools and optimizing the impact they can have in clinical practice and medication safety.

The third paper authored by Marcilly et al. [8] is a systematic review of usability flaws in medication-related alerting functions. It has been widely argued that improving the usability of CDSS functions is necessary, given that usability flaws hamper in multiple ways their optimal use and impact in clinical practice. Nevertheless, existing lists of usability design principles regarding medication-related alerting functions per se, originate mostly from expert consensus or targeted reviews, rather than evidence-based approaches. Thus, Marcilly et al. pursued this study, in order to identify in an evidence-based fashion the different types of usability flaws that can be found in medication-related alerting functions, aiming to complement, foster, and systematize knowledge in the domain. By comprehending the characteristics of these usability flaws, the study aimed to provide insights for better CDSS design and development, in order to overcome such shortcomings. The literature corpus of the study was obtained by searching PubMed, Scopus, and Ergonomics Abstracts databases, while the review was conducted by experts in human factors engineering.

The main extracted data of the study by Marcilly et al. [8] concerned meaningful semantic units representing instances of usability flaws. Data were analyzed through qualitative methods, i.e., a categorization following general usability heuristics, and through an inductive process for the flaws specific to medication-related alerting functions. The main characteristics of the selected papers concerned the mode of alerting (e.g., interruptive, non-interrup-

tive, mixed, etc.), potential integration of the alerting function(s) (e.g., in electronic health records (EHRs), CPOEs, stand-alone), the class of supported functions (e.g., drug-drug interactions, drug allergy, duplicate therapy, dosing guidance, etc.), the implementation stage of the respective CDSS modules (e.g., in use, under development, etc.), and the evaluation means employed (e.g., observations, interviews, retrospective analysis, etc.). Interestingly, the study concluded by identifying 168 instances of usability flaws, which were in turn classified into 13 categories, among which seven concerned general usability principles and six concerned medication-related alerting functions in particular. Categories of general usability flaws included guidance, workload, consistency, control, and adaptability, while categories of usability flows concerned medication-related alerting functions per se, accounted for low signal-to-noise-ratio, alert content, presentation issues, and so forth. Apparently, such flaws do not only concern the user interface of a medication-related alerting function, but rather all the underlying components such as the knowledge model, the triggering model, and the behavior of the function. Nevertheless, given that such usability flaws may negatively affect the quality of the physicians' interaction with the CDSS, the study elaborated on identifying potential links between the categories of flaws that have been observed in medication-related alerting functions and the steps of the interaction. To this end, Marcilly et al. employed Norman's 7-stage model of action as a structured model of user interaction, which discriminates the interaction in two phases, namely, the action and the evaluation phases. In this context, two action/evaluation loops were proposed, the core one dealing with the "display/reading" of the alert, and a second one, called "acknowledgment", depending on the alerting function model (sometimes, no acknowledgment was required). For each of the seven stages, the authors comprehensively discussed potential usability flows. The study also pinpointed the need for establishing reporting guidelines for usability-related studies. Even though it might not cover the entire set of usability flaws that can be found in medication-related

alerting functions, it provided an important evidence-base originated from an exhaustive literature search of the topic.

Interestingly, both the paper of Cho et al. [6] and Marcilly et al. [8] are particularly relevant with the special theme of the current Yearbook, "Unintended Consequences of HIT". Usability flaws in medication-related alerting functions of CDSS may result in unintended consequences and medication errors in particular, thus, systematic evaluation of such systems using frameworks like Leapfrog is necessary to accommodate safety and efficacy concerns. As Cho et al. pinpointed [6], introducing a CPOE is a complex intervention and the implementation of such a tool does not always reduce medical errors but occasionally augments them. We refer the reader to the survey paper by Coiera et al. provided in this Yearbook [1], which summarizes a number of such aspects in the domain of DSS and HIT in general, under a contemporary perspective.

In the fourth paper, Shalom et al. [9] report on the assessment of a guideline-based CDSS for the diagnosis and the management of pre-eclampsia, eclampsia, and toxemia (PET). The CDSS has been developed using the Picard DSS engine developed by the authors and the 2012 PET guidelines of the American College of Obstetricians and Gynecologists. The study has been conducted in the department of Obstetrics and Gynecology of a University hospital in Israel with 36 participating clinicians. The objectives of the study were to assess the effect of the CDSS on the compliance of clinician decisions with the guidelines and to collect clinicians' perceptions of the CDSS use. The study was not based on the implementation of the system as an intervention in actual clinical practices, but it was designed as an "in vitro" experiment with simulated cases. Four remarkable points deserve to be highlighted in the work reported in this paper. Firstly, the developed guideline-based CDSS addresses the full complexity of the management of the disease in a longitudinal manner, from diagnostic to drug therapy management, including the physical examination and the ordering of imaging and lab tests. When the authors built the simulated patient cases, they anticipated several decision points in the scenario, where a clinician

would make decisions, and where the CDSS should be able to issue guideline-based recommendations of various types. This differs from many existing automated CDSSs which are "vertical", in the sense that they are often dedicated to support one kind of decision (typically, diagnosis, therapy, drug management, etc.), and then provide support at only one decision point. Moreover, the six simulated cases were elaborated on the basis of actual disease management situations, while controlling that all the decision points identified were different, mobilizing distinct guideline knowledge. The multiple scenarios included a total of 60 decision points covered by the PET guidelines. Secondly, the Picard DSS engine, which executes the PET CDSS, builds on several research prototypes and tools developed by the authors' team and others, resulting from a long series of research works for decades. In the domain of guideline-based DSS it is noticeable that long lasting medical informatics works, with a significant publication record, end-up to an applicative prototype close to implementation in a clinical setting. Thirdly, the primary outcome of the study was the compliance of clinician decisions with guidelines. Guideline compliance is a classical performance measure for decision support interventions, but how it is calculated is not always well-described, whereas this has an impact on the compliance rates provided. In their paper, the authors provide a good description on how they elaborated their compliance measures. Particularly, they distinguish two dimensions of guideline compliance to assess clinicians' decisions, completeness and correctness. Completeness is characterized by the percentage of guideline-recommended actions that were decided, and decision correctness corresponds to the percentage of decided actions that were guideline-consistent, still making the difference between necessary and redundant actions. These measures provide efficient means to figure out the effect of CDSSs. Fourthly, the assessment protocol of the CDSS in the context of an "in vitro" study with limited clinician resources has been sufficiently well-designed to control many potential biases. With six simulated scenarios, including 60 distinct decision points, and 36 clinicians, they performed

a cross-over experiment where all decision points were presented to all clinicians, and where all decision points have been equally considered in DSS mode and in non-DSS mode. With this design, a total number of 2,160 (36 × 60) decision occurrences could be assessed. For the other aspects of the paper, the authors fairly discuss the limitations of their study, which are usual for case simulations with one CDSS, considering one clinical problem, and in one sole clinical site. As for the results of the study, the authors showed that the correctness of decisions was high (around 94%) and not changed by the CDSS, but that the rate of correct, but redundant, actions was impacted decreasing from 68% in non-DSS mode to 3% in DSS mode. However, the completeness of guideline coverage in decisions was positively impacted by the CDSS, increasing from 41% to 93%. The proportion of non guideline-compliant decisions, qualified as "errors", remained the same in both modes, around 6%. The authors conclude that their CDSS might mostly reduce errors of omissions with respect to completeness, and to a lesser extent errors of commission (redundant actions), while not eliminating incorrect (with respect to the guidelines) decisions for which alternative methods should be sought.

Among the 974 reviewed papers for the Decision Support section of the 2015 edition of the IMIA Yearbook, several contributions brought to light some interesting results and developments and deserve to be cited in this synopsis. From a technical viewpoint, Wilk et al. [10] elaborated on representing the dynamics (i.e., team formation, management, and task-practitioner allocation) of an interdisciplinary healthcare team through semantic components (i.e., an ontology describing concepts and relations in the domain, behavioral rules describing the team dynamics, and the corresponding knowledge base). These semantic components are part of a multi-agent system supporting the selection and maintenance of the most responsible physician as well as complex rules to select practitioners for the workflow tasks. As a proof-of-concept, the approach is illustrated through a clinical scenario of a healthcare team managing a patient with chronic kidney disease. In the domain of big data in healthcare, Khazaei et al. [11] presented a cloud-based framework for big data analytics, supporting both real-time and retrospective analyses. The study introduces an architectural paradigm to implement health-analytics-as-a-service, which is currently a central theme in the era of big data in healthcare (e.g., a relevant effort is conducted in the H2020 AEGLE project, http://www.aegle-uhealth.eu/). Khazaei et al. demonstrated the application of the proposed framework in monitoring a neonatal Intensive Care Unit through a case study on sepsis. Two editions of the platform were proposed, the clinical and the research one, while an interesting part of the study concerned privacy and security issues. In addition, Bettencourt-Silva et al. [12] explored how routinely collected hospital data can be used to develop data-driven pathways that describe the "journeys" that patients take through care and, accordingly, their exploitation in biomedical research. Elaborating on prostate cancer with data obtained from eight different hospital information systems, complemented with information from a local cancer registry, as the case study, Bettencourt-Silva et al. proposed a framework for the construction, quality assessment, and visualization of patient pathways for clinical studies and decision support. As a methodological contribution, Sarker et al. [13] presented a fully automatic method for predicting the quality of medical evidence obtained from literature sources. Given the wide variety of medical literature sources currently available through the Internet, the manual appraisal of the quality of evidence is a time-consuming process. The approach relies on a sequence of high precision classifiers, applied on medical article abstracts, utilizing data from a specialized corpus. The experiments presented in the study suggest that the approach achieves evaluation results comparable to those of human performance.

In terms of CDSS applications, Li et al. [14] presented an algorithmic approach for improving neonatal patient safety through automated detection of medication administration errors in EHRs. High alert medications were considered in the study, e.g., narcotics, vasoactive medication, intravenous fluids, etc., for which an appropriate algorithm was specified (based on standard care practices) and implemented. The rate and the types of the identified medication administration errors were compared to incidence reporting through physician chart review. Interestingly, the study demonstrated the identification of many previously un-identified errors, demonstrating significantly better sensitivity (82% vs. 5%) and precision (70% vs. 50%) than incident reporting for error recognition. In addition, Shoshi et al. [15] presented GraphSAW, a web-based system for graphical analysis of drug interactions and side effects using pharmaceutical and molecular data. GraphSAW is able to analyze single and combined drug-drug interactions, drug-molecule interactions, as well as single and cumulative side effects, by exploiting data from two commercial and two freely available molecular databases.

As regards CDSS adoption and use, Sukums et al. [16] presented a study on the promising implementation of a CDSS for antenatal and intrapartum care in rural primary healthcare facilities in Sub-Saharan Africa (particularly in Ghana and Tanzania regions). The CDSS aimed to facilitate adherence to clinical practice guidelines and to support decision-making during client encounter to bridge the know-do gap of health workers. The implementation and use of the CDSS for over 20 months were overall successful, demonstrating high acceptance and usage rates among its users. Facilitators for this positive outcome included sufficient training and regular support, while unreliable power supply and perceived high workload were considered as major challenges for a sustainable use. Cho et al. [17] conducted a retrospective observational study in order to investigate the relationship between provider characteristics and their response to medication alerts in an outpatient setting. The study revealed that six physician characteristics (physician type, age, number of encounters, medical school ranking, residency hospital ranking, and acceptance of Medicaid) were significantly related to the override rate. Furthermore, house staff were more likely to override than staff physicians, physicians with fewer than 13 average daily encounters were more likely to override than physicians with more than 13 encounters, and

graduates of the top five medical schools were more likely to override than the others. Such indications can be particularly useful when designing CDSS systems with more targeted alerts.

In connection with the special theme of the 2016 IMIA Yearbook, various interesting studies were also published coping with unintended consequences of HIT. For example, Slight et al. [18] studied the vulnerabilities of a wide range of CPOE systems to different types of medication errors, aiming to foster a comprehensive qualitative understanding of how CPOE design could be improved. Thirteen commercial and homegrown CPOE systems across 16 different sites in the US and Canada were assessed based on a random sample of medication error reports, in which CPOE systems were considered a "contributing factor" to errors. Beyond CPOE failure to detect and prevent important medication errors in some cases, alerts/warnings varied widely among systems, with some of them being confusing. The findings are particularly useful for CPOE designers and developers towards the construction of safer prescription systems. Dekarske et al. [19] illustrated in a prospective, randomized crossover study, the increased appropriateness of customized alert acknowledgement explanations compared to non-customized configurations when overriding medication alerts in a CPOE system. The study pinpoints that poor application design or configuration can negatively affect provider behavior, when responding to important medication alerts. Via a prospective observational study, Czock et al. [20] illustrated how tailoring alerts substantially reduces the alert burden in CDSS for drugs that should be avoided in patients with renal disease. The study analyzed critical drug prescriptions in a university-based nephrology clinic and evaluated the effect of four different alerting strategies on the alert burden, demonstrating that strategies considering patient and drug-specific information have the potential to reduce the alert burden by more than 90%.

In conclusion, while we were able to identify some notable examples of routinely used CDSSs, this literature review still shows that, despite a lot of methodological contributions, theoretical frameworks,

prototype developments and evaluations in simulated environments or in specific trials, CDSSs and CPOEs are not yet widely spread (at least with their full functionalities) in routine clinical practice. Further research, testing, evaluation, and training are still needed to incorporate such tools in clinical practice and illustrate the benefits that they promise.

References

1. Coiera E, Ash J, Berg M. The unintended consequences of health information technology revisited. Yearb Med Inform 2016:163-9.
2. Lamy JB, Séroussi B, Griffon N, Kerdelhué G, Jaulent MC, Bouaud J. Toward a formalization of the process to select IMIA Yearbook best papers. Methods Inf Med 2015;54(2):135-44.
3. Bouaud J, Koutkias V. Computerized Clinical Decision Support: Contributions from 2014. Yearb Med Inform 2015:119-24.
4. Anand V, Carroll AE, Biondich PG, Dugan TM, Downs SM. Pediatric decision support using adapted Arden Syntax. Artif Intell Med 2015 Oct 1.
5. Anand V, Biondich PG, Liu G, Rosenman M, Downs SM. Child health improvement through computer automation: the CHICA system. Stud Health Technol Inform 2004;107(Pt 1):187–91.
6. Cho I, Lee JH, Choi SK, Choi JW, Hwang H, Bates DW. Acceptability and feasibility of the Leapfrog computerized physician order entry evaluation tool for hospitals outside the United States. Int J Med Inform 2015 Sep;84(9):694-701.
7. Jung M, Hoerbst A, Hackl WO, Kirrane F, Borbolla D, Jaspers MW, et al. Attitude of physicians towards automatic alerting in computerized physician order entry systems. A comparative international survey. Methods Inf Med 2013;52(2):99-108.
8. Marcilly R, Ammenwerth E, Vasseur F, Roehrer E, Beuscart-Zéphir MC. Usability flaws of medication-related alerting functions: A systematic qualitative review. J Biomed Inform 2015 Jun;55:260-71.
9. Shalom E, Shahar Y, Parmet Y, Lunenfeld E. A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines. Int J Med Inform 2015 Apr;84(4):248-62.
10. Wilk S, Kezadri-Hamiaz M, Rosu D, Kuziemsky C, Michalowski W, Amyot D, Carrier M. Using Semantic Components to Represent Dynamics of an Interdisciplinary Healthcare Team in a Multi-

Agent Decision Support System. J Med Syst 2016 Feb;40(2):42.
11. Khazaei H, McGregor C, Eklund JM, El-Khatib K. Real-Time and Retrospective Health-Analytics-as-a-Service: A Novel Framework. JMIR Med Inform 2015 Nov 18;3(4):e36.
12. Bettencourt-Silva JH, Clark J, Cooper CS, Mills R, Rayward-Smith VJ, de la Iglesia B. Building Data-Driven Pathways From Routinely Collected Hospital Data: A Case Study on Prostate Cancer. JMIR Med Inform. 2015 Jul 10;3(3):e26.
13. Sarker A, Mollá D, Paris C. Automatic evidence quality prediction to support evidence-based decision making. Artif Intell Med 2015 Jun;64(2):89-103.
14. Li Q, Kirkendall ES, Hall ES, Ni Y, Lingren T, Kaiser M et al., Automated detection of medication administration errors in neonatal intensive care. J Biomed Inform 2015 Jul 17.
15. Shoshi A, Hoppe T, Kormeier B, Ogultarhan V, Hofestädt R. GraphSAW: a web-based system for graphical analysis of drug interactions and side effects using pharmaceutical and molecular data. BMC Med Inform Decis Mak 2015 Feb 28;15:15.
16. Sukums F, Mensah N, Mpembeni R, Massawe S, Duysburgh E, Williams A, et al. Promising adoption of an electronic clinical decision support system for antenatal and intrapartum care in rural primary healthcare facilities in sub-Saharan Africa: The QUALMAT experience. Int J Med Inform 2015 Sep;84(9):647-57.
17. Cho I, Slight SP, Nanji KC, Seger DL, Maniam N, Fiskio JM, et al. The effect of provider characteristics on the responses to medication-related decision support alerts. Int J Med Inform 2015 Sep;84(9):630-9.
18. Slight SP, Eguale T, Amato MG, Seger AC, Whitney DL, Bates DW, Schiff GD. The vulnerabilities of computerized physician order entry systems: a qualitative study. J Am Med Inform Assoc 2016 Mar;23(2):311-6.
19. Dekarske BM, Zimmerman CR, Chang R, Grant PJ, Chaffee BW. Increased appropriateness of customized alert acknowledgement reasons for overridden medication alerts in a computerized provider order entry system. Int J Med Inform 2015 Dec;84(12):1085-93.
20. Czock D, Konias M, Seidling HM, Kaltschmidt J, Schwenger V, Zeier M, et al. Tailoring of alerts substantially reduces the alert burden in computerized clinical decision support for drugs that should be avoided in patients with renal disease. J Am Med Inform Assoc 2015 Jul;22(4):881-7.

Correspondence to:
Dr Vassilis Koutkias
Institute of Applied Biosciences
Centre for Research & Technology Hellas
6th Km. Charilaou – Thermi Road
P.O. BOX 60361
GR – 57001 Thermi, Thessaloniki
Greece
Tel. +30 2311 25 76 15
E-mail: vkoutkias@certh.gr

Appendix: Content Summaries of Selected Best Papers for the 2016 IMIA Yearbook, Section Decision Support

Anand V, Carroll AE, Biondich PG, Dugan TM, Downs SM

Pediatric decision support using adapted Arden Syntax

Artif Intell Med 2015 Oct 1

Prevention represents an important part of pediatric care. Though preventive guidelines exist, they are insufficiently followed. To address this problem, a decision support system, the Child Health Improvement through Computer Automation system (CHICA) was developed by the authors more than ten years ago in 2004. Guideline knowledge is encoded in Arden Syntax, a standard computer-interpretable formalism for representing knowledge rules as medical logical modules (MLMs). CHICA is connected to the electronic medical record (EMR) and is also used to feed the EMR. Notably, it is used to collect pre-screening data directly from the patient's family with a personalized form. Then, a patient-specific, physician-directed form is issued, synthesizing information to prepare the encounter with the patient and providing guideline recommendations. CHICA has been running for a decade and is implemented in routine use in seven pediatric clinics of a healthcare system in Indiana, USA. Forty-one pediatric clinical issues are covered by 429 MLMs. More than 44,000 patients and 755 healthcare professionals were served by CHICA. MLMs were fired more than ten million times. In this context, the system effectiveness has been assessed in several published randomized controlled trials with positive outcomes. Some adaptations of the Arden Syntax were performed to successfully implement CHICA's functionalities. These modifications are considered minor, being relevant with the syntax, the parser, and the underlying parallel execution framework. As one of the most studied CDSS for child health care, the CHICA system demonstrated that the Arden Syntax standard is effective for representing preventive pediatric guideline knowledge and for delivering decision support in actual clinical settings.

Cho I, Lee JH, Choi SK, Choi JW, Hwang H, Bates DW

Acceptability and feasibility of the Leapfrog computerized physician order entry evaluation tool for hospitals outside the United States

Int J Med Inform 2015 Sep;84(9):694-701

This study assessed the acceptability and feasibility of Leapfrog, a tool for CPOE evaluation developed in the US, in four tertiary and academic hospitals of South Korea. Leapfrog has been extensively used in the US based on various unsafe dangerous medication-ordering scenarios, covering 11 categories of erroneous medication orders, as well as monitoring checks, nuisance orders and, finally, a "deception analysis" test (checking for false positive indications). Given that the tool has only been available and used in the US, the study tried to shed light on unconfirmed CPOE-related issues and effects on patient outcomes in another country. In each participating hospital site of the study, a self-developed CPOE system was available, which has been used for more than a decade. Interestingly, all sites were regulated by a national drug utilization review process, which enforced them to implement decision support mechanisms. A cross-sectional design was implemented and the system response rate, the category completion rate, as well as the time to complete the evaluation, were measured. In addition, the study compared the evaluation results of the four systems with the scores obtained for five US systems as they have been reported in the literature. Interpretation of raw test results was based on the relative importance of each type of decision support for patient harm prevention, while the assigned score reflected both the severity of a potential ADE which is not intercepted by the system as well as its probability of occurrence. The categorization indicating test performance across order-checking included "fully implemented", "good progress in implementing", "good early-stage effort", "completed the evaluation", and "incomplete evaluation". The evaluation was conducted in a one-week period, requiring approximately eight hours at each site. The measurements were calculated anonymously, while descriptive rates and proportions were compared among the four sites and then with those of five US community hospitals reported in another study. The overall category completion rates ranged from 67.9% to 75.5%, with a varying degree according to the evaluation category, while the required time to finish the tests was within the allowed test timeframe (varying from 3.1 to 4 hours). In the total evaluation score, the hospitals ranged from 21.6% to 36.5%. Three hospitals were assigned the "completed the evaluation" stage and one the "incomplete evaluation" stage. Three systems could cause severe harm in the "Therapeutic Duplication" category, and one of them in the "Drug-Allergy" category as well. The overall scores of the South Korean systems were lower than the average of the five US systems, with the two highest scores of the South Korean systems being slightly higher than the lowest score in the US. The various identified differences in terms of error category and system might be due to the different scope and coverage of CPOE safety performance in hospitals. Regarding feasibility, evaluation was tolerable at the four sites. The study concluded that there is a potential for Leapfrog to be used for hospitals outside the US.

Marcilly R, Ammenwerth E, Vasseur F, Roehrer E, Beuscart-Zéphir MC

Usability flaws of medication-related alerting functions: A systematic qualitative review

J Biomed Inform 2015 Jun;55:260-71

The current paper constitutes a systematic review of usability flaws in medication-related alerting functions. The study aimed to identify the different types of usability flaws in the targeted domain in an evidence-based fashion. Ultimately, the goal was to complement, foster, and systematize knowledge in the domain, which currently originates from expert consensus or targeted reviews. Three bibliographic databases were searched,

providing initially 6,380 references, from which only 26 met the eligibility criteria of the study. The main characteristics of the selected papers concerned the mode of alerting (e.g., interruptive, non-interruptive, mixed, etc.), potential integration of the alerting function(s) (e.g., in EHR, CPOE, standalone), the class of supported functions (e.g., drug-drug interactions, drug allergy, duplicate therapy, dosing guidance, etc.), the implementation stage of the respective CDS modules (e.g., in use, under development, etc.), and the evaluation means employed (e.g., observations, interviews, retrospective analysis, etc.). The study identified 168 instances of usability flaws, which were in turn classified into 13 categories, among which seven concerned general usability principles (i.e., issues related with guidance, workload, significance of codes, consistency, explicit control, adaptability, and error management) and six concerned medication-related alerting functions in particular (i.e., issues concerning low signal-to-noise-ratio, alert content, transparency of functions for the user, alert appearance, task and control distribution, as well as alert features). These flaws are not only relevant for user interfacing aspects but also concern the underlying knowledge model, the triggering model, and the behavior of the alerting functions. Given that such usability flaws may negatively affect the quality of the physicians' interaction with the CDSS, the study elaborated on identifying potential links between the categories of flaws that have been observed in medication-related alerting functions and the steps of the interaction. Using Norman's 7-stage model of action and its two phases (i.e., the action

and the evaluation), two action-evaluation loops were proposed, the core one dealing with the "display/reading" of the alert, and a second one, called "acknowledgment", depending on the alerting function model. For each one of the seven stages, the authors comprehensively discussed potential usability flows. They also raised completeness of usability flaws in medication-related alerting functions as a potential limitation of the study. Nevertheless, the study provided an important evidence originated from this comprehensive literature search and qualitative review of the topic.

Shalom E, Shahar Y, Parmet Y, Lunenfeld E
A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines

In this work, authors' objective was to assess a continuous-care, guideline application engine, named Picard, which they developed. The primary measures, related to clinician's compliance with guidelines, were the completeness and correctness of clinicians' decisions with respect to the American 2012 guidelines for the diagnosis and management of pre-eclampsia and eclampsia, which they implemented as a CDSS with Picard. The secondary measure was the subjective assessment of the system perception by clinicians. They designed a cross-over protocol to quantify the effect of the CDSS using multiple simulated patient cases, or longitudinal scenarios, involving 60 decision points covered by the guidelines.

The study was conducted by 36 clinicians of an academic department of Obstetrics and Gynecology. According to the design, all six scenarios and decision points, were presented to clinicians either in DSS mode or in non-DSS mode, an allocation assuring that all decision points were equally assessed in each modes by half of the participants. Based on the analysis of 2,712 actions decided by the clinicians in the execution of the protocol, authors showed that the correctness of decisions, i.e., the fact that the decided actions were guideline-consistent, was always high (around 94%) and then not changed by the CDSS. In both modes, the rate of errors, i.e., of non-compliant decisions, remained stable. However, the rate of correct, but redundant, actions was impacted, decreasing from a proportion of 68% in non-DSS mode to 3% in DSS mode. This demonstrated the CDSS effect in reducing errors of commissions. As for completeness, i.e., the ratio of recommended actions decided by decision points, it was positively impacted by the CDSS, increasing from 41% to 93%. This demonstrated that the CDSS reduced errors of omissions and fostered guideline adherence. No effect of the level of clinician training was observed. The CDSS was considered potentially useful by the clinicians. Such a system, even in the complex longitudinal management of patient cases, confirms that CDSSs can enhance performance, through guideline adherence, and efficiency by reducing redundant actions. Nevertheless, in this experiment, the CDSS did not affect the rate of incorrect actions, suggesting that other approaches should be considered.