

# Contributions from the 2017 Literature on Clinical Decision Support

V. Koutkias<sup>1</sup>, J. Bouaud<sup>2,3</sup>, Section Editors for the IMIA Yearbook Section on Decision Support

<sup>1</sup> Institute of Applied Biosciences, Centre for Research & Technology Hellas, Thermi, Thessaloniki, Greece

<sup>2</sup> Assistance Publique—Hôpitaux de Paris, Delegation for Clinical Research and Innovation, Paris, France

<sup>3</sup> Sorbonne Université, Université Paris 13, Sorbonne Paris Cité, INSERM, UMR\_S 1142, LIMICS, Paris, France

## Summary

**Objectives:** To summarize recent research and select the best papers published in 2017 in the field of computerized clinical decision support for the Decision Support section of the International Medical Informatics Association (IMIA) yearbook.

**Methods:** A literature review was performed by searching two bibliographic databases for papers referring to clinical decision support systems (CDSSs). The aim was to identify a list of candidate best papers from the retrieved bibliographic records, which were then peer-reviewed by external reviewers. A consensus meeting of the IMIA editorial team finally selected the best papers on the basis of all reviews and the section editors' evaluation.

**Results:** Among the 1,194 retrieved papers, the entire review process resulted in the selection of four best papers. The first paper studies the impact of recency and of longitudinal extent of electronic health record (EHR) datasets used to train a data-driven predictive model of inpatient admission orders. The second paper presents a decision support tool for surgical team selection, relying on the history of surgical team members and the specific characteristics of the patient. The third paper compares three commercial drug-drug interaction knowledge bases, particularly against a reference list of highly-significant known interactions. The fourth paper focuses on supporting the diagnosis of postoperative delirium using an adaptation of the "anchor and learn" framework, which was applied in unstructured texts contained in EHRs.

**Conclusions:** The conducted review illustrated also this year that research in the field of CDSS is very active. Of note is the increase in publications concerning data-driven CDSSs, as revealed by the review process and also reflected by the four papers that have been selected. This trend is in line with the current attention that "Big Data" and data-driven artificial intelligence have gained in the domain of health and CDSSs in particular.

## Keywords

Medical informatics; International Medical Informatics Association; Yearbook; Decision Support Systems

Yearb Med Inform 2018;122-8

<http://dx.doi.org/10.1055/s-0038-1641222>

## Introduction

Decision support is a landmark topic in medical informatics. Since the inception of the Yearbook of the International Medical Informatics Association (IMIA) in 1992, a section has been dedicated to this topic. The goal of this synopsis is to summarize recent research in the domain of decision support and to select the best papers published in this field during 2017, based on a comprehensive literature review. Our review targeted clinical decision support systems (CDSSs) and computerized provider order entry (CPOE) systems. Of note is this year's survey paper of the decision support section by Cho and Bates [1], which elaborates on the novel perspective of behavioral economics interventions for clinical decision support.

The synopsis is organized as follows: the next section briefly describes the review protocol and the methods employed for selecting the best papers on the topic; the following section presents the results of this year's selection process, and the last section discusses the main contributions of the four best papers as well as noticeable research works in the domain of decision support, which were identified during the selection process.

## Paper Selection Method

A comprehensive literature search on topics related to CDSSs and CPOE systems was performed to identify candidate best papers following the established protocol applied in the past years [2]. We used two bibliographic databases, primarily the PubMed/MEDLINE

database (from NCBI, National Center for Biotechnology Information) as it is dedicated to biomedical literature and, secondarily, Web of Science® (WoS, from Thomson Reuters) to retrieve publications which are not referenced in PubMed, since WoS has a broader scope. Both databases were searched with similar queries targeting journal papers published in 2017, written in English, and on the aforementioned topics. The adopted strategy, which was first implemented last year [3] and replicated this year, was based on four exclusive queries yielding four disjoint citation subsets:  $Q_{Pub\_plain}$ , based on a plain-text search in PubMed titles and abstracts using keywords;  $Q_{Pub\_indexed}$ , based on the PubMed indexing scheme using MeSH terms and exclusive of the previous set;  $Q_{WoS\_restricted}$ , based on a WoS search on non PubMed-indexed papers restricted to the two subject areas "Medical Informatics" and "Health Care Sciences & Services" and, finally,  $Q_{WoS\_filtered}$ , based on other non-PubMed-indexed papers filtered by non-relevant subject areas. It should be noted that, due to the delay in the PubMed population process, some papers published during 2016 were not yet retrievable at the date we queried the bibliographic database last year; thus, PubMed queries were modified accordingly so that papers missed last year could be considered for the 2017 selection.

A first review of the four subsets of retrieved citations was performed by the two section editors to select 15 candidate best papers. These candidate best papers were then individually reviewed and rated by external reviewers from the international Medical Informatics community. Based on reviewers' ratings and comments, the Yearbook editorial committee had to select three to six best papers of the year in the decision support domain.

## Review Results

Database extraction on the 2017 literature with the four queries was performed on January 13, 2018. A total of 1,194 references were obtained, distributed as follows: 859 for  $Q_{Pub\_plain}$ , 163 for  $Q_{Pub\_indexed}$ , 30 for  $Q_{WoS\_restricted}$  and 142 for  $Q_{WoS\_filtered}$ , yielding sub-totals of 1,022 references from PubMed and 172 from WoS. Compared to the previous year, we retrieved 49 papers more in total. The two section editors reviewed the four batches of citations separately. The non-rejected citations were then merged, yielding 57 papers that were reviewed again to select 15 candidate best papers. Following the IMIA Yearbook best paper selection process, these papers were then peer-reviewed by external reviewers and the Yearbook editors. Four papers were finally selected as best papers for 2017 [4–7], all indexed in PubMed. The four papers are listed in Table 1 (in alphabetical order of the first author's surname), and they are discussed in the next section. Summaries of their contents are available in the Appendix of this synopsis.

## Discussion and Outlook

In the first paper, Chen *et al.* [4] adopted the position that the existing static knowledge-based, or guideline-based, approach to clinical decision support is limited in scale due to both the lack of evidence for all interventions and the cost of human authoring processes, which do not allow to account for the perpetually evolving practice of medicine. On the contrary, consistent with the paradigm of learning health systems and taking advantage of data accumulated in electronic health records (EHRs), they assumed that data-driven clinical decision support could be effective to predict clinical practice patterns and reduce practice variability. However, learning from past practices in order to make future decisions is somehow paradoxical and raises concerns about the adaptability of the decision support to the aforementioned continuous evolution of medicine. In their paper, the authors studied how varying longitudinal historical training data can impact the prediction of

**Table 1** Best paper selection of articles for the IMIA Yearbook of Medical Informatics 2018 in the section 'Decision Support'. The articles are listed in alphabetical order of the first author's surname.

Section
Decision Support
▪ Chen JH, Alagappan M, Goldstein MK, Asch SM, Altman RB. Decaying relevance of clinical data towards future decisions in data-driven inpatient clinical order sets. <i>Int J Med Inform</i> 2017 Jun;102:71-9.
▪ Ebadi A, Tighe PJ, Zhang L, Rashidi P. DisTeam: A decision support tool for surgical team selection. <i>Artif Intell Med</i> 2017 Feb;76:16-26.
▪ Fung KW, Kupusnik-Uner J, Cunningham J, Higby-Baker S, Bodenreider O. Comparison of three commercial knowledge bases for detection of drug-drug interactions in clinical decision support. <i>J Am Med Inform Assoc</i> 2017 Jul 1;24(4):806-12.
▪ Mikalsen KØ, Soguero-Ruiz C, Jensen K, Hindberg K, Gran M, Revhaug A, Lindsetmo RO, Skrøvseth SO, Godtliebsen F, Jenssen R. Using anchors from free text in electronic health records to diagnose postoperative delirium. <i>Comput Methods Programs Biomed</i> 2017 Dec;152:105-14.

future clinical decisions. A clinical order recommender system, analogous to Netflix or Amazon's product recommenders based on customers' prior purchase, was used to predict admission orders in a tertiary academic hospital based on patients' diagnoses and recorded orders at admission. The objective of the study was to assess the impact on the accuracy of decision prediction of varying historical datasets used to train the clinical order recommender system and to estimate the decay rate of the relevance of prior data. Nine training sets were built on available EHR data from 2009 to 2012 considering different periods varying in duration, from one month to the whole period (4 years), and in starting year. In parallel, "classical" order sets, expert-based and human-authored, attached to admission diagnoses were elaborated. Predicted orders as well as human-authored order sets were compared to actual 2013 data. Results showed that the accuracy of predicted decisions for the reference period (2013) was significantly better when the system was trained on just one month on recent data (2012) than with one year of old data (2009). Using more data from a longer period, four years in the past, was not better than using the most recent data (2012) except when applying a decaying weighting scheme. In this context, testing several values, an efficient half-life of data relevance was estimated at four months. The authors concluded that data-driven models predict decisions better when trained on small recent datasets than on larger sets augmented with older data. Adding older training data may

lead to less efficient predictions unless a decaying weighting function is used. The authors pointed out that, whatever the training set, predicted decisions using data mining were more accurate than knowledge-based predefined, human-authored, order sets. However, the questions of what is "good" practice in this context and what is the gold standard are of paramount importance; maybe past suboptimal practice was endorsed as a model for future, suboptimal, practice. This issue is addressed in the last two paragraphs of the discussion section of the paper and is worth reading since the position is questionable. The reported study is remarkable, but it would call for a qualitative analysis of the quality of prior admission orders, as they are the foundation of the automated learning process that would inform about the reliability of the approach. Likewise, given that such data-driven approaches are conservative, the question about the consideration of the perpetual evolution of medicine needs to be thought, especially if the training decision sets could themselves be biased by data-driven decision support systems.

The second paper, authored by Ebadi *et al.* [5], introduced DisTeam, a decision support tool aiming to address optimal surgical team selection having as its cornerstone a genetic algorithm. To this end, DisTeam entails a "training" mode and an "operational" mode. The "training" mode consists of the "patient clustering" module (allowing the tailoring of the surgical team selection procedure to patient characteristics) and the "extracting existing teams" module (fetching distinct

surgical service providers that are being used to formulate intermediate solutions), while the optimization procedure is performed in the “operational” mode through the genetic algorithm. More specifically, for the “patient clustering” a retrospective dataset of patients forming groups of similar patients was employed (based on features including age, ethnicity, race, Charlson comorbidity index, and body mass index). This allowed the identification of the most representative cluster for a given patient using K-Prototypes, a variation of the well-known K-means clustering method. The “extracting existing teams” module extracts the surgical team (consisting of all the healthcare professionals who provided care to the patient, e.g. a surgeon, an anesthesiologist, a nurse, etc.) associated with each patient. Whenever the information of a new patient is entered into DisTeam, the best possible team is suggested through the “operational” mode by finding the most similar patient cluster, extracting candidate teams and, ultimately, selecting the best candidate. Through the “patient clustering” module, DisTeam determines which cluster best represents the new patient and then selects the target cluster, which is being used by the “optimization” module to select the best surgical team. The respective fitness function relies on the number of complications that have occurred during the surgery, which are in turn associated with providers at three different levels, i.e. collective level (any past surgical case involving “all” current team members), pairwise level (any past surgical case involving any “two” current team members), and individual level (any past surgical case involving any current team member). First, past collective surgical cases of the given candidate team are considered. If there are no collective surgical cases, then pairwise surgical cases of any members of the candidate team are checked and, if no pairwise cooperation is found, individual provider’s performance is considered. DisTeam demonstrated high effectiveness in its evaluation based on intra-operative data from 6,065 unique orthopedic surgery cases, involving a total of 440 surgeons, anesthesiologists, and circulators. Overall, DisTeam introduced a complementary perspective for decision-making for surgical team selection, exceeding criteria such as the healthcare pro-

viders’ availability and preferences, which are typically employed by existing tools.

The third paper, authored by Fung *et al.* [6], compares three commercial drug-drug interaction (DDI) knowledge bases (KBs) used for automated decision support. Medication safety has always been a central concern to healthcare delivery since drug dispensing is one of the major causes of iatrogenesis. Adverse events due to known DDIs may be among the most preventable, and CPOEs that generate DDI alerts are the most disseminated kind of decision support systems. Previous studies, *e.g.*, McEvoy *et al.* [8] cited in the last year’s Decision Support Synopsis of the IMIA Yearbook [3], have highlighted variability in DDI resources, alerts, and implementation. The study reported by Fung *et al.* [6] aimed at conducting a comprehensive comparison of the commercial KBs widely used in US hospitals, clinics, and pharmacies. First, a normalization process was performed on all drug resources, in which listed drugs were mapped to RxNorm (<https://www.nlm.nih.gov/research/umls/rxnorm/>) to allow for comparisons. Then, the contents of the KBs were statically compared to assess how they overlapped. It was also determined whether each KBs covered a reference list of highly significant DDIs from the Office of the National Coordinator for Health Information Technology, referred to as the ONC list. Finally, all KBs as well as the ONC list were applied to an actual dataset of 14 million prescriptions to trigger DDI alerts and simulate their effect for clinical decision support. Five drug KB vendors were contacted and three accepted to participate to the study. It must be noted that vendor representatives are co-authors of the published paper. Results showed that the number of drug-drug pairs listed in each KB varied from a factor of three. A total of 8.6 million unique drug pairs were identified in the three KBs, among which 79% were present in only one KB and 5% in all three KBs. This low number supports the finding that DDI resources are highly variable. Further content analysis showed however that within the subset of common pairs, there was more agreement than disagreement in the severity ranking of the DDIs, especially for contraindications, which represent the most important category

of DDIs. When considering the high priority DDIs of the ONC list, each KB covered them at least in 99%. This result showed that DDIs identified as among the most important were correctly handled despite quantitative variations in size and contents. Applied to the prescription dataset, the total number of alerts varied according to the KB. However, the ONC list alerts were again all covered by each KB, though differences in the severity ranking were observed. Notably, two statins and QT-prolonging agents were responsible for more than 97% of all ONC alerts. While KBs significantly cover the reference ONC DDIs, the authors suggested that other contraindicated DDIs shared by all KBs might complement the current ONC list. They concluded that observed variations in size and contents call for better standardization of drug KBs supported by better evidence, preferably obtained from EHR-derived patient outcomes rather than from expert panel consensus.

In the fourth paper, Mikalsen *et al.* [7] presented an adaptation of the “anchor and learn” framework, which was applied in the exploitation of free-text EHRs for addressing the challenging problem of diagnosing post-operative delirium. In particular, through this data-driven CDSS approach, Mikalsen *et al.* introduced a new method for anchor specification based on domain knowledge and exploratory data analysis. This analysis relied on clustering and visualization techniques and provides the opportunity to obtain a labeled training set without manual label annotation. In addition, compared to the original anchor and learn framework, which relies on L2-regularized logistic regression for classification [9], Mikalsen *et al.* employed instead “elastic net” as a robust solution in settings where the dimension is higher than the sample size. The paper provides a comprehensive description of all the steps entailed in the application of the proposed framework, as well as the limitations of the study. The proposed framework was applied in a quite large number of patient EHRs, which were extracted from the Department of Gastrointestinal Surgery (DGS) at the University Hospital of North Norway from 2004 to 2012. The dataset included both structured information, such as ICD-10 diagnosis codes, age, sex, length of surgery,

blood test results, etc., as well as free-text from documents such as doctor notes, radiology reports, and semi-structured nurse notes. For testing, a clinician created a list of major abdominal surgeries requiring general anesthesia, defining a cohort of patients who could potentially suffer from postoperative delirium. The clinician manually read the EHRs for a subset of the cohort, in order to find patients that experienced postoperative delirium; thus, a training set was obtained with the remaining “unlabeled” patients. The study illustrated a significant increase in the performance through the proposed approach, i.e. the area under the precision-recall curve (AUC-PR) value was 0.51, when creating the labels in a naive way, and 0.96 when defining them through the adapted learn and anchor framework. The study concluded that the proposed method could be quite successful when applied in problems where no obvious anchors exist as well as in other application domains, such as the preoperative identification of malnourished patients and the prediction of patients at risk for postoperative complications.

Besides the best papers selected for the Decision Support section of the 2018 edition of the IMIA Yearbook, which are discussed in this synopsis, several contributions obtained from our literature search brought to light some interesting results and developments and, thus, deserve to be presented. For example, with respect to technical contributions, Merone *et al.* [10] presented a decision support system for tele-monitoring chronic obstructive pulmonary disease (COPD)-related worrisome events. The system comprises a binary finite state machine, the training stage of which allows for subject-specific personalization of its underlying predictive model, which triggers warnings and alarms as the health status evolves over time. Yet *et al.* [11] introduced a framework for representing the evidence-base of a Bayesian network (BN) decision support model, aiming to present all the clinical evidence alongside the BN itself (i.e. supporting and conflicting evidence, as well as evidence associated with relevant but excluded factors). The framework is applied on a BN for predicting acute traumatic coagulopathy. Oliveira *et al.* [12] presented a temporally-oriented healthcare assistant,

the underlying model of which provides a comprehensive representation of temporal constraints in Clinical Practice Guidelines (CPGs). The expressiveness of the model is illustrated via a case study featuring CPGs for the diagnosis and management of colon cancer. Mohammadhassanzadeh *et al.* [13] elaborated on semantics-based plausible reasoning, aiming to extend the coverage of medical KBs for improved clinical decision support. The work relied on Semantic Web technology to solve complex clinical decision support queries and it was evaluated using a real-world medical dataset of patients with hepatitis, from which different percentages of data were randomly removed to reflect scenarios with increasing amounts of incomplete medical knowledge. Gräßer *et al.* [14] presented a system for data-driven therapy decision support based on techniques from the field of recommender systems, while Danaley *et al.* [15] introduced the Genomic Prescribing System, an online, secure, electronic custom interface aiming to simplify the use of pharmacogenomics in clinical practice.

In terms of CDSS evaluation and impact assessment, Peleg *et al.* [16] presented a comprehensive study assessing a patient-centered, mobile decision support system for patients and their care providers, which relies on clinical guidelines and semantically integrated EHRs. The assessment concerned two domains, i.e. atrial fibrillation and gestational diabetes mellitus, focusing particularly on both patient and care provider compliance to guideline recommendations and overall satisfaction with the CDSS, as well as on patient quality of life. Focusing on medication safety per se, Kannampallil *et al.* [17] analyzed medication orders voiding in CPOE systems by exploiting data from an academic medical center for a 6-year period, while Baysari *et al.* [18] conducted a longitudinal study to obtain user experiences concerning the implementation of a CPOE system in a pediatric hospital. Ip *et al.* [19] elaborated on identifying CDS factors contributing to imaging order cancellation or modification, a study performed across four institutions participating in the Medicare Imaging Demonstration with findings that may have implications for the future design of such CDSSs. Finally, Liberati *et al.* [20]

conducted a qualitative study and introduced an implementation framework for CDSSs in hospitals by assessing what hinders CDSS uptake in the hospital environment. The study concluded that the respective barriers and facilitators are dynamic in nature and may exist prior to the CDSS introduction in the clinical context. In addition, factors such as clinicians' attitude towards scientific evidence and guidelines, the quality of inter-disciplinary relationships, and an organizational ethos of transparency and accountability need to be considered when exploring the readiness of a hospital to adopt CDSSs.

As also remarked in the synopsis of the Decision Support Section of the 2017 IMIA Yearbook [3], the review conducted this year illustrates that the research in the field of CDS remains very active. We should note that we witnessed a significant increase in publications concerning data-driven CDSSs, an observation that is reflected by the four papers that have been selected as best papers. This trend is to some extent expected, given the extraordinary attention that “Big Data” and data-driven artificial intelligence are currently experiencing in the health domain overall and in decision support systems in particular.

## Acknowledgement

We would like to thank Martina Hutter and Adrien Ugon for their support, and the reviewers for their participation to the selection process of the Decision Support section of the IMIA Yearbook.

## References

- Cho I, Bates DW. Behavioral economics interventions in clinical decision support systems. *Yearb Med Inform* 2018;114-21.
- 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.
- Koutkias V, Bouaud J. Contributions from the 2016 Literature on Clinical Decision Support. *Yearb Med Inform* 2017 Aug;26(1):133-8.
- Chen JH, Alagappan M, Goldstein MK, Asch SM, Altman RB. Decaying relevance of clinical data towards future decisions in data-driven inpatient clinical order sets. *Int J Med Inform* 2017 Jun;102:71-9.
- Ebadli A, Tighe PJ, Zhang L, Rashidi P. DisTeam:

- A decision support tool for surgical team selection. *Artif Intell Med* 2017 Feb;76:16-26.
6. Fung KW, Kapusnik-Uner J, Cunningham J, Higby-Baker S, Bodenreider O. Comparison of three commercial knowledge bases for detection of drug-drug interactions in clinical decision support. *J Am Med Inform Assoc* 2017 Jul 1;24(4):806-12.
  7. Mikalsen KØ, Soguero-Ruiz C, Jensen K, Hindberg K, Gran M, Revhaug A, Lindsetmo RO, Skrovseth SO, Godtliebsen F, Jenssen R. Using anchors from free text in electronic health records to diagnose postoperative delirium. *Comput Methods Programs Biomed* 2017 Dec;152:105-14.
  8. McEvoy DS, Sittig DF, Hickman TT, Aaron S, Ai A, Amato M, et al. Variation in high-priority drug-drug interaction alerts across institutions and electronic health records. *J Am Med Inform Assoc* 2017 Mar 1;24(2):331-8.
  9. Halpern Y, Horng S, Choi Y, Sontag D. Electronic medical record phenotyping using the anchor and learn framework. *J Am Med Inform Assoc* 2016 Jul;23(4):731-40.
  10. Merone M, Pedone C, Capasso G, Incalzi RA, Soda P. A Decision Support System for Tele-Monitoring COPD-Related Worrisome Events. *IEEE J Biomed Health Inform* 2017 Mar;21(2):296-302.
  11. Yet B, Perkins ZB, Tai NRM, Marsh WR. Clinical evidence framework for Bayesian networks. *Knowl Inf Syst* 2017;50(1):117-43.
  12. Oliveira T, Silva A, Neves J, Novais P. Decision support provided by a temporally oriented health care assistant: an implementation of computer-interpretable guidelines. *J Med Syst* 2017 Jan;41(1):13.
  13. Mohammadhassanzadeh H, Van Woensel W, Abidi SR, Abidi SSR. Semantics-based plausible reasoning to extend the knowledge coverage of medical knowledge bases for improved clinical decision support. *BioData Min* 2017 Feb 10;10:7.
  14. Gräßer F, Beckert S, Küster D, Schmitt J, Abraham S, Malberg H, Zaunseder S. Therapy decision support based on recommender system methods. *J Healthc Eng* 2017;2017:8659460.
  15. Danahey K, Borden BA, Furner B, Yukman P, Hussain S, Saner D, Volchenboum SL, Ratain MJ, O'Donnell PH. Simplifying the use of pharmacogenomics in clinical practice: Building the genomic prescribing system. *J Biomed Inform* 2017 Nov;75:110-21.
  16. Peleg M, Shahar Y, Quaglini S, Broens T, Budasu R, Fung N, et al. Assessment of a personalized and distributed patient guidance system. *Int J Med Inform* 2017 May;101:108-30.
  17. Kannampallil TG, Abraham J, Solotskaya A, Philip SG, Lambert BL, Schiff GD, et al. Learning from errors: analysis of medication order voiding in CPOE systems. *J Am Med Inform Assoc* 2017 Jul 1;24(4):762-8.
  18. Baysari MT, Hardie RA, Lake R, Richardson L, McCullagh C, Gardo A, et al. Longitudinal study of user experiences of a CPOE system in a pediatric hospital. *Int J Med Inform* 2018 Jan;109:5-14.
  19. Ip IK, Lacson R, Hentel K, Malhotra S, Darer J, Langlotz C, et al. JOURNAL CLUB: Predictors of provider response to clinical decision support: Lessons learned from the Medicare Imaging Demonstration. *AJR Am J Roentgenol* 2017 Feb;208(2):351-7.
  20. Liberati EG, Ruggiero F, Galuppo L, Gorli M, González-Lorenzo M, Maraldi M, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci* 2017 Sep 15;12(1):113.

#### 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 2018 IMIA Yearbook, Section Decision Support

Chen JH, Alagappan M, Goldstein MK, Asch SM, Altman RB

**Decaying relevance of clinical data towards future decisions in data-driven inpatient clinical order sets**

**Int J Med Inform 2017 Jun;102:71-9**

Taking note that the current knowledge-based decision support approaches to promote best practice are limited by classical trial-based clinical research and human authoring, Chen *et al.* advocate for data-driven approaches that would take advantage of data accumulated in EHRs to predict clinical practice patterns and then offer automated decision support. A clinical order recommender system, analogous to Netflix or Amazon's product recommenders, was used to predict admission orders in a tertiary academic hospital based on admission diagnoses. The objective of the study was to assess the impact of this approach on the accuracy of decision prediction for varying historical datasets used to train the clinical order recommender system and to estimate the decay rate of the relevance of prior data. The training sets were built on available data from 2009 to 2012 with different periods varying in duration, from one to 48 months, and varying in their starting year. Predicted orders and human-authored order sets were compared to actual 2013 data. Results showed that the accuracy of predicted decisions for the reference period (2013) was significantly better when the system was trained on just one month and on recent data (2012) than when it was trained on one year and on older data (2009). Interestingly, using more data from a longer period (2009-2012) was not better than using the most recent data (2012), except when applying a decaying weighting scheme. In this context, the half-life of data relevance was estimated at four months. Whatever the training set, the predicted decisions using data mining were more accurate than the predefined, human-authored, order

sets. The authors conclude that data-driven models predict decisions better when trained on small recent datasets than on larger sets augmented with older data. Adding older training data may lead to less efficient predictions, unless a decaying weighting function is used.

Ebadi A, Tighe PJ, Zhang L, Rashidi P

**DisTeam: A decision support tool for surgical team selection**

**Artif Intell Med 2017 Feb;76:16-26**

DisTeam is a novel decision support tool for surgical team selection aspiring to contribute in reduced conflicts, better coordination, and better patient outcomes. While many studies have elaborated on how to organize optimal human resources allocation in the hospital environment, DisTeam relies on a matchmaking framework to accommodate optimal surgical team selection among individual healthcare professionals (surgeons, anesthesiologists, and nurse circulators). In particular, having as input the historical data about surgical teams (i.e. surgical complications associated with teams and their members as well as their teamwork history) and the individual characteristics of the patient (e.g. age, body mass index, and Charlson comorbidity index score), Ebadi *et al.* introduced via DisTeam a metaheuristic framework for objective evaluation of surgical teams, in order to identify the optimal team for a given patient. Relying on a genetic algorithm, DisTeam generates a ranked list of possible surgical teams, the evaluation of which suggested high effectiveness. Specifically, DisTeam converges quickly to the optimal solution, providing the best surgical team as well as additional teams that could be employed alternatively. DisTeam was evaluated using intra-operative data from 6,065 unique orthopedic surgery cases, involving 60 distinct surgeons, 157 anesthesiologists, and 223 circulators. Compared to the current state-of-the-art, DisTeam extends existing scheduling software by considering team structure and history as well as patient's specific characteristics, beyond personnel preferences (about days, shifts, units, etc.) as well as regulatory and union requirements.

Fung KW, Kapusnik-Uner J, Cunningham J, Higby-Baker S, Bodenreider O

**Comparison of three commercial knowledge bases for detection of drug-drug interactions in clinical decision support**

**J Am Med Inform Assoc 2017 Jul 1;24(4):806-12**

Medication safety is of primary importance to healthcare delivery. Known DDIs may be among the most preventable adverse events. The authors compared three commercial drug KBs largely used in the US for automated decision support in CPOE systems generating DDI alerts. In a normalization phase, all drug resources were mapped to RxNorm to allow for comparisons. The contents of KBs were statically compared to assess how they overlapped. It was also checked whether the KBs covered a reference list of highly significant DDIs from the Office of the National Coordinator for Health Information Technology (ONC). Finally, each KB and the ONC list were applied to an actual dataset of 14 million prescriptions to simulate their effect as clinical decision support. Results showed that the number of drug-drug pairs listed in each KB varied from a factor of three. Among the 8.6 million unique pairs, 79% were present in only one KB and 5% in all three KBs. Content analysis showed that there was more agreement than disagreement in the severity ranking of the DDIs in common pairs, especially for contraindications. The selected DDIs of the ONC list were covered at 99% or more in the three KBs. Applied to the prescription dataset, the number of alerts varied according to the KB. However, ONC alerts were all covered by the KBs, though differences in the severity ranking were observed. Notably, two statins and QT-prolonging agents were responsible of more than 97% of all ONC alerts. Observed variations in size and contents call for better standardization of drug KBs. Despite this, KBs significantly cover the reference ONC DDIs. The authors suggest that other contraindicated DDIs shared by all KBs might complement the current ONC list.

Mikalsen KØ, Soguero-Ruiz C, Jensen K, Hindberg K, Gran M, Revhaug A,

Lindsetmo RO, Skrøvseth SO, Godtliebsen F, Jenssen R

**Using anchors from free text in electronic health records to diagnose postoperative delirium**

**Comput Methods Programs Biomed 2017 Dec;152:105-14**

Although being a common complication after major surgery with serious consequences, especially in the elderly population, postoperative delirium remains often undetected. To this end, Mikalsen *et al.* exploited free-text parts of EHRs (a rich information resource, given that nurses monitor the patient's health status after the surgery and report them three times a day) to construct

data-driven CDSSs for addressing this problem. However, this task typically relies on the labeling of training data, which is both time-consuming and expensive to perform as a process. This shortcoming was addressed by Mikalsen *et al.* by adopting an anchor-based learning framework which transforms the key observations contained in the free-text (i.e. the anchors) into labels. Learning with anchors presents a method of efficiently learning statistically-driven phenotypes with minimal manual intervention, under the assumption that the presence of an anchor variable implies the presence of the latent label of interest. In order to eliminate the problem of specifying reliable anchors, Mikalsen *et al.* developed a problem-specific method (based on domain knowledge and exploratory data analysis using clustering

and visualization techniques) and employed the elastic-net based classification, which forces sparsity and has been shown to provide robustness in settings where the dimensionality is higher than the sample size. Aiming to assess the proposed framework in the detection of postoperative delirium, Mikalsen *et al.* exploited EHR data (corresponding to 7,741 patients) from a Norwegian university hospital and observed an increase in the AUC-PR from 0.51 to 0.96 compared to baselines. Overall, the study concluded that the proposed method could be successfully applied in problems where no obvious anchors exist as well as in other application domains, such as the preoperative identification of malnourished patients and the prediction of patients at risk for postoperative complications.