

Paving the COWpath: data-driven design of pediatric order sets

Yiye Zhang,¹ Rema Padman,² James E Levin³

¹School of Information Systems Management, H John Heinz III College, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

²H John Heinz III College, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

³(Late) CMIO, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, Pennsylvania, USA

Correspondence to

Yiye Zhang, School of Information Systems Management, H John Heinz III College, Carnegie Mellon University, 4800 Forbes Ave, Pittsburgh, PA 15213, USA; yiyez@andrew.cmu.edu

Received 3 September 2013

Revised 5 February 2014

Accepted 24 February 2014

Published Online First

27 March 2014

ABSTRACT

Objective Evidence indicates that users incur significant physical and cognitive costs in the use of order sets, a core feature of computerized provider order entry systems. This paper develops data-driven approaches for automating the construction of order sets that match closely with user preferences and workflow while minimizing physical and cognitive workload.

Materials and methods We developed and tested optimization-based models embedded with clustering techniques using physical and cognitive click cost criteria. By judiciously learning from users' actual actions, our methods identify items for constituting order sets that are relevant according to historical ordering data and grouped on the basis of order similarity and ordering time. We evaluated performance of the methods using 47 099 orders from the year 2011 for asthma, appendectomy and pneumonia management in a pediatric inpatient setting.

Results In comparison with existing order sets, those developed using the new approach significantly reduce the physical and cognitive workload associated with usage by 14–52%. This approach is also capable of accommodating variations in clinical conditions that affect order set usage and development.

Discussion There is a critical need to investigate the cognitive complexity imposed on users by complex clinical information systems, and to design their features according to 'human factors' best practices. Optimizing order set generation using cognitive cost criteria introduces a new approach that can potentially improve ordering efficiency, reduce unintended variations in order placement, and enhance patient safety.

Conclusions We demonstrate that data-driven methods offer a promising approach for designing order sets that are generalizable, data-driven, condition-based, and up to date with current best practices.

BACKGROUND AND SIGNIFICANCE

As technology assumes an increasingly critical role in healthcare delivery, recent literature indicates that the interaction of technology and humans in healthcare settings warrants a thorough investigation.¹ Poor usability of health information technology (IT) causes new types of medical errors that are unique to the technology era, particularly through insufficient customization of the systems to workflow in the care delivery setting, end user information overload, and lack of adequate knowledge about user behaviors.^{1–4} In this paper, we examine an unintended, adverse consequence of implementing computerized physician order entry (CPOE) systems—the excessive physical and cognitive workload imposed on users.⁵ This is a significant contributor to poor judgment, inaccuracies, and

erroneous actions that may lead to adverse outcomes in the clinical care setting.⁶ Poor usability stems from the lack of coordination between technology and human practices.^{6–9} Previous research indicates that knowledge of user behaviors can have a profound impact on achieving health IT use, and incorporating this knowledge into the design of health IT systems is vital for enhancing their usability.¹

To address this challenge, we introduce a 'paving the COWpath' approach into the design of health IT system components,^{10–11} where 'COWpath' indicates the ordering 'paths' taken by busy providers, typically using a 'computer on wheels'.¹² This approach is a design rationale where the system components function by learning from and embracing the best practices performed repeatedly by users, such that they are able to discover users' 'desire paths' and facilitate actual workflow without requiring excessive human intervention.¹³ We suggest this as a mechanism to combine explicit knowledge¹⁴—available from evidence-based guidelines—with the tacit, non-codified knowledge embedded in the user experiences of domain experts resulting from the interactions between processes, culture, and routines.^{15–17} As medication ordering and management is a critical patient-safety-sensitive electronic health record function,¹ we focus on the development of 'order sets', a CPOE component that has the potential to make order entry faster, easier, and safer.⁶

Order sets have the potential to increase the efficiency of order entry.⁶ There are two ways users can place orders via CPOE. One is a la carte order placement, where users search for an order using a keyword and place a single order found from the search results. When multiple orders are needed, users repeat this process as many times as the number of orders needed. However, by using order sets, users can enter multiple system-suggested orders with as few as one mouse click. Instead of searching by order name, users can search for an order set by disease condition, and be presented with a set of expert-suggested/evidence-based orders, all related to the condition of interest. Hence, order sets also serve as a decision support tool to simplify complex order placement and reduce prescribing variation and errors.¹⁸ Traditionally developed by experts in each clinical area, order sets include multiple orders related by clinical purposes that less-experienced/time-constrained users may refer to. Since a majority of potential adverse drug events are the result of errors during prescribing,¹⁹ well-designed order sets can increase compliance on order placement based on best practices and improve patient safety by lowering the chances of prescribing variations and errors.



CrossMark

To cite: Zhang Y, Padman R, Levin JE. *J Am Med Inform Assoc* 2014;**21**:e304–e311.

Today, the creation of order sets has been widely accepted as a core prerequisite for successful CPOE implementation.^{6 20–23}

Despite these benefits, historical data indicate a low rate of usage of order sets, primarily due to lack of order set content to accommodate diverse patient conditions, physicians' personal preferences and unfamiliarity with order sets, and, most importantly, inconsistency of order set content with current best practices.^{24–27} Instead, we observe heavy use of a la carte orders to compensate for the limitations of order sets, generating a COWpath that provides evidence of evolving care delivery needs and adapted to users' workflows, but may be highly inefficient. Indeed, owing to the time-, manual labor-, and knowledge-intensive order set development process, the growth, modification, and maintenance of order sets are often not kept up to date with changing provider skills and interests, quality and clinical resource management initiatives, new types of services, regulatory requirements, and an evolving standard of care.^{24–27} Consequently, there may be a significant mismatch between CPOE component capability and provider needs that harms treatment quality and exposes patients to potential medical errors.

Order set usage has been examined in a number of previous studies to identify usability issues,^{28–31} and ambulatory and laboratory corollary order sets have been created using data-mining techniques.^{30 32} For example, Hulse *et al*³³ created an automated, knowledge-based order set feedback system that analyzes order set usage relative to the original template and suggests modifications based on actual usage. In a study by Avansino and Leu,³⁴ systematically designed order sets were shown to provide a reduction in cognitive workload, but there was no improvement in terms of mouse clicks or time spent on order sets. In this paper, we demonstrate that potential reductions in both physical and cognitive workload are possible through order set optimization. We summarize our research to automate the entire order set design process using machine learning and optimization techniques while incorporating physical and cognitive workload as major development criteria.^{24–27} The machine learning methods enable us to learn the tacit knowledge embedded in the COWpath and combine them with the evidence built into the current order sets, while the optimization criteria allow us to accommodate human factors in the design process. We further evaluate our methods using CPOE data for inpatient asthma (a chronic medical condition), appendectomy (an acute surgical condition), and pneumonia (an acute medical condition). This approach has the potential to be more broadly generalizable to other conditions and care delivery settings.

OBJECTIVE

We describe the complexity of order set usage, analyze its effect on users and patient safety, and develop order sets that reduce physical and cognitive workload imposed on users. Specifically, we propose an order set development method by leveraging historical ordering data from the inpatient setting. As shown in figure 1, at an abstract level, using only a la carte orders at one extreme or a single, comprehensive order set that includes all possible orders at the other extreme will both incur very high physical and cognitive workload. Between these two, we hypothesize that there must be a 'satisficing' number of order sets, each containing an 'acceptable' number of order items, such that users can place orders safely with much lower physical and cognitive workload.³⁵ Therefore, our methods model the order set development problem as an optimization process

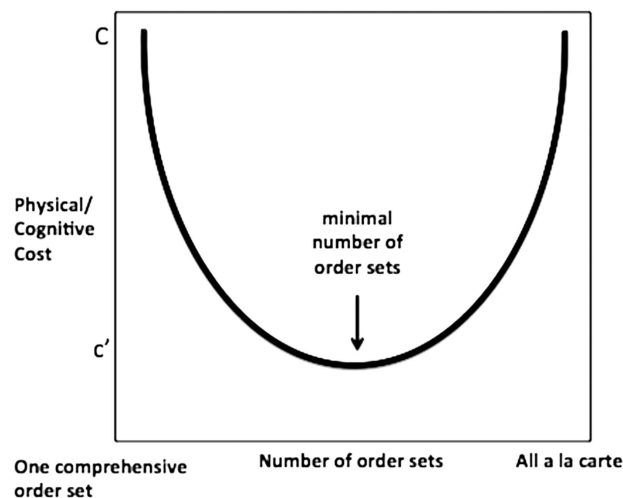


Figure 1 Order sets versus usage cost.

embedded with machine learning techniques to identify orders that are clinically and temporally relevant.

Physical and cognitive workloads are measured in terms of mouse clicks. We define mouse click cost (MCC) as the number of mouse clicks users apply to complete ordering tasks, and cognitive click cost (CCC) as MCC multiplied by weights indicating the complexity of specific ordering tasks.^{24 27} Using historical order data for inpatient management of asthma, appendectomy, and pneumonia, we demonstrate the method's performance on different clinical conditions that may affect the automated order set development process differentially. Rather than merely extracting popularly used orders to develop order sets, we take into consideration each condition's characteristics, as well as usability issues, to create order sets that are data-driven, condition-based, and up to date with the current best practices.

MATERIALS AND METHODS

Data

Data for this study were obtained from the Children's Hospital of Pittsburgh (CHP) of the University of Pittsburgh Medical Center (UPMC). As the first pediatric hospital in the USA to reach Healthcare Information and Management Systems Society (HIMSS) Analytics Stage 7 electronic medical record implementation status, CHP has been all-digital with a single integrated eRecord (Cerner Millennium) since 2009. With over 13 000 inpatient visits in 2011, a patient at CHP was hospitalized, on average, for 5.5 days, during which time 36 unique individuals entered 846 order actions. Over 10 million order actions are stored in the clinical data warehouse, which is an exact copy of the production eRecord.²⁵ There are 1559 departmental order sets in the eRecord system covering 17 978 orders, designed in 2002 by CHP clinicians, mainly for medical admission and surgery. The order set template has not been updated since its creation in 2002, and all analyses performed for the study were based on the same template. All patient-identifiable health information was removed to create a deidentified dataset for this study. This study was designated as exempt by the University of Pittsburgh Institutional Review Board.

Important variables extracted from the data warehouse include the order, order action time, care set flag, deidentified patient ID, diagnosis, severity, age at admission, gender, and length of stay in hours. For this study, order action time includes time from 20 h before admission to 24 h after admission, as

most order set usage is concentrated on the first day of admission. Data from 20 h before admission often come from patients' emergency department visits immediately before they are admitted to the inpatient setting. Orders are from the same encounter as the admission, but the data do not exclude orders that are placed for other conditions when the patient has other comorbidities. Diagnosis and severity are captured by drug description and severity of illness using all patient refined diagnosis-related groups (APR-DRGs).³⁶

For this study, we extracted the top 26 APR-DRGs, including 4331 cases that required a total of 490 049 orders that were placed in calendar year 2011. From this set, 292 asthma cases with 20 076 orders and 40 order sets, 105 appendectomy cases with 13 280 orders and 51 order sets, and 179 pneumonia cases with 13 743 orders and 69 order sets were selected for evaluating the methods developed in this study. Patient age is categorized into age groups following American Academy of Pediatrics (AAP) age groups: 1=neonate (<30 days); 2=infancy (≥30 days and <1 year); 3=early childhood (≥1 and <5 years); 4=late childhood (≥5 and <13 years); 5=adolescence (≥13 and ≤18 years); 6=adult (≥18 years). A care set flag identifies order set items against a la carte items. A care set flag of 0 refers to an a la carte item, 1 an order set parent, 2 an order set, and 3 an order set nested in another order set. In this study, we include only orders that have a care set flag of 0, 2 and 3.

The size of order sets ranges from two to more than 50 unique items. Each item in order sets can be defaulted 'ON' or 'OFF' according to clinical relevance and frequency of use. An order item can be included in multiple order sets when necessary. Items with a checked checkbox are default ON items, and those with blank checkboxes are default OFF items. Figure 2 is a screen shot that shows part of the 'Admission Orders Pneumonia' order set. Unlike a la carte order placement, where users need to apply a mouse click every time to select an individual order, default ON items are automatically selected when an order set is chosen. With additional clicks, users can add default OFF items to the selection or deselect default ON items from the order placement. Ordering through order sets is not mandatory, with a few exceptions.

Method

Click costs

As shown in figure 2, MCC counts the exact number of mouse clicks that users have to apply for each order entry. One unit of MCC is incurred when a user places an a la carte order, selects an order set, deselects a defaulted ON order set item, or selects a defaulted OFF order set item. Furthermore, given that these tasks require different levels of cognitive workload, we multiply MCC by a different cognitive cost coefficient for each task to calculate CCC. Assuming that these coefficients are independent of the patient, order set, time interval, and physician (this assumption can be easily relaxed in the models but incur significant information gathering overhead), we applied two approaches to derive the coefficients for CCC to represent the cognitive workload that accompanies each mouse click in order placement. One was provided by a clinical and medical informatics expert, and the other was estimated from a survey of 15 users including physicians and nurses.²⁷ Cognitive-click-cost-expert-estimate (CCCE) coefficients are made assuming that users typically trust the order set contents. Hence CCCE attributes high cost coefficients to actions that disagree with the default setting, such as deselecting ON items or selecting OFF items. Conversely, cognitive-click-cost-survey-estimate (CCCS) coefficients reveal that, while deselecting ON items and

selecting OFF items requires high costs, almost the same amount of cognitive workload is required of users to ensure that default settings are suitable for patients, leading to small differences in the four coefficients. In this study, MCC and CCC focus on clicks that are related to ordering only and do not include other activities while ordering such as up or down clicks, and do not account for incomplete orders and rework. Table 1 summarizes the cost coefficients derived for MCC and CCC.

Coverage rate

We introduce a term called coverage rate to measure the goodness of order sets. Coverage rate of an order set *c* used by *p*=1, ..., *N* patients is defined as

$$\frac{1}{N} \sum_{p=1}^N \frac{i_{p,c}}{i_c}$$

where *i_{p,c}* is the number of unique items given to patient *p* from order set *c*, and *i_c* is the total number of unique items originally in order set *c*. Generally, orders from multiple order sets are given to a single patient during one order entry *e*. Given *S* order sets and *N* patients, the average coverage rate (ACR) during one order entry is

$$\frac{1}{S} \sum_{c=1}^S \left(\frac{1}{N} \sum_{p=1}^N \frac{i_{p,c}}{i_c} \right)$$

Then, the average overall coverage rate across *S* order sets, *M* order entries, and *N* patients is

$$\frac{1}{S} \sum_{c=1}^S \left(\frac{1}{M} \sum_{e=1}^M \left(\frac{1}{N} \sum_{p=1}^N \frac{i_{p,e,c}}{i_c} \right) \right)$$

where *i_{p,e,c}* is the number of unique items given to patient *p* from order set *c* during order entry *e*. High coverage rate for an order set suggests that users can use the order set with few further modifications on average.

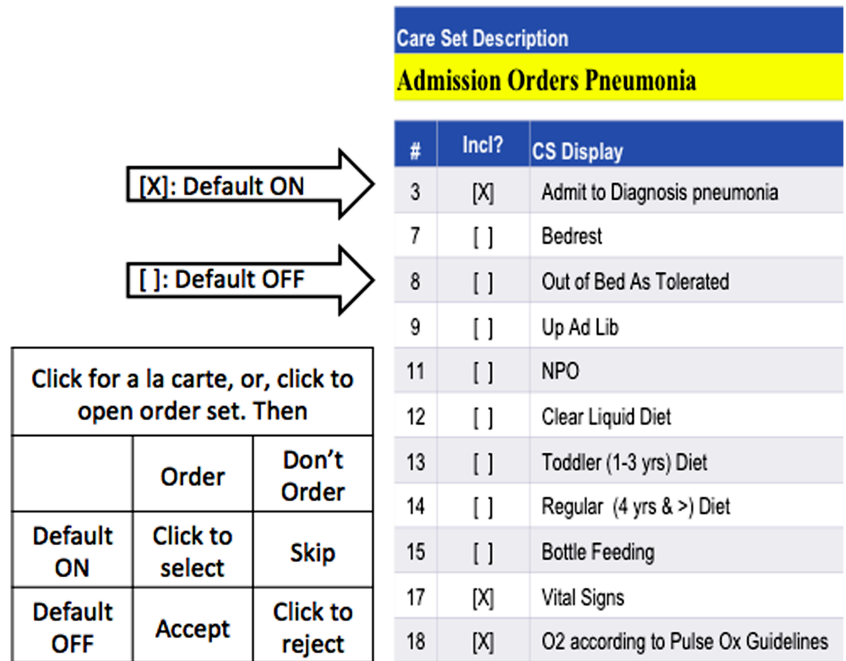
Relative risk

We attempt to place order items that tend to be ordered together for a single patient during an order entry session into one order set. In order to find clinically relevant order items, we apply a similarity measure called relative risk (RR), commonly used in biomedical informatics to look for similarity between diseases.³⁷ Given a pair of orders *x* and *y*, let *U_{xy}* be the number of unique patients who had both orders, *V_x* be the number of times order *x* was placed for patients, *V_y* be the number of times order *y* was placed for patients, and *N* be the total number of patients. Relative risk, *RR_{xy}*, for this order pair is defined as (*U_{xy}N/V_xV_y*). Given *i* number of orders, each order has *i*-1 number of RR measures when paired with each of the rest of the orders. Bisecting K-means clustering is applied to the *i* by *i* matrix of RR to identify similar orders that should belong to the same order sets. Bisecting K-means clustering is known to perform clustering faster than K-means clustering.³⁸

Iterative algorithm

Since order placement is driven by patients' conditions, we develop order sets by extracting order items most suitable for constituting order sets according to different time intervals. This is modeled as a two-stage optimization problem with MCC and CCC as the objective function, respectively, while satisfying

Figure 2 A Children’s Hospital of Pittsburgh order set for pneumonia.



constraints on time and order set content. Before the optimization process, default settings for orders are updated such that an item is defaulted ON if more than 80% of the study patients have used it at least once, and OFF if otherwise, thus judiciously learning from the COWpath. For each clinical condition, the algorithm iteratively extracts orders present in the data in increasing time intervals, starting with a 2 h interval. In each iteration, the bisecting K-means algorithm clusters items that often co-occur in patients’ order lists into the same order sets. Combinations of order sets with the lowest MCC or CCC are kept as the best solutions. Finally, the iteration ends when MCC and CCC are compared across time intervals of different lengths to select the optimal time interval. The end point of the previous time interval becomes the starting point for the next time interval (figure 3). Each item is part of no more than one order set within one time interval. However, an item can be included in multiple order sets across different time intervals. This process is repeated from 20 h before admission until 24 h after admission.

To computationally evaluate the methods and illustrate the optimization results, the data are divided into training and test sets. While training sets include patients who had only the particular APR_DRG diagnosis as their final diagnosis upon discharge, test sets evaluate the performance of the machine-derived order sets using data on patients who exhibited comorbidities so that the robustness of the order sets can be

determined in a realistic manner. For example, the asthma training set includes patients who had a diagnosis of ‘asthma, unspecified’, while its testing set includes patients who had comorbidities such as ‘contact dermatitis and other eczema, unspecified cause’ at some point during their hospital stay. Similarly, the appendectomy training set includes a sample of patients with ‘acute appendicitis’ and ‘appendicitis, unqualified’, and the pneumonia training set includes patients with ‘pneumonia, organism unspecified.’ In addition, the severity of the condition was controlled to reduce the noise in the optimization process. We use the severity group with the largest sample size in each condition for reporting the computational results. Based on the assignment rule discussed above, we have 84 asthma patients for training and 208 for testing, 93 appendectomy patients for training and 57 for testing, and 96 pneumonia patients for training and 83 for testing.

Clustering and optimization were performed using an integrated computational platform that we developed (figure 4). Users can select clinical conditions and severity level and optimize order sets based on MCC, CCCE, and CCCS criteria, and time intervals of interest. Currently available clinical conditions are asthma, appendectomy, and pneumonia, although new conditions can be readily added. Resulting orders and their respective order set assignments, MCC and CCC values, and parameters used for each optimization session on the platform are stored in a secure MySQL database.

RESULTS

Descriptive statistics

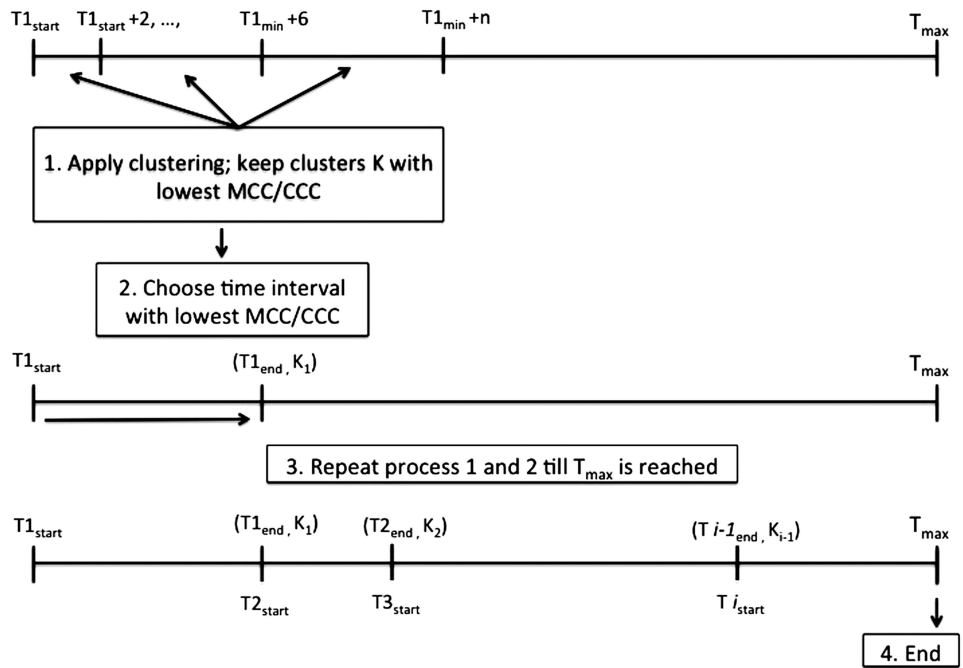
Table 2 shows order set usage associated with asthma, appendectomy, and pneumonia management. ‘Number of unique order sets used’ is the total number of unique order sets used across all the patients with each condition. On the other hand, ‘Number of times order set opened’ is the total number of times users chose order sets to place orders. ‘Number of order set orders’ is the number of orders placed from order sets, as opposed to a la carte. ‘Number of clicks to use order set’ is the number of mouse clicks that users applied to place order set orders. ‘Clicks saved by order sets’ is the difference between the

Table 1 MCC and CCC coefficients

	MCC		CCCE		CCCS	
	Select	Deselect	Select	Deselect	Select	Deselect
Order set/a la carte	1	–	1.2	–	1.1	–
Default ON	0	1	0.2	1.5	1	1.3
Default OFF	1	0	0.5	0.1	1.4	1.1

CCCE, cognitive-click-cost-expert-estimate; CCCS, cognitive-click-cost-survey-estimate; MCC, mouse click cost.

Figure 3 Iterative process for order set timing and content. CCC, cognitive click cost; MCC, mouse click cost.



number of order set orders and the number of clicks to place order set orders. For example, the actual number of order set orders needed for pneumonia patients was 8935, but users had to make 9827 mouse clicks to complete order placement from 2662 order sets, adding 10% higher physical workload compared with a la carte order placement, where users simply needed to make 8935 mouse clicks. Yet, had the 2662 order sets been a perfect fit for patients, users would have completed the order placement process with just 2662 clicks, saving 70% of the physical workload. While it is highly unlikely that order sets will match every patient’s requirements perfectly, and hence will always require some sort of modification, this extreme

example demonstrates the potentially significant impact that improved usability of order sets may have on treatment efficiency.

Optimization results

Table 3 shows a summary of optimization results under three objective functions: MCC, CCCE, and CCCS. Reduction in clicks per patient is the difference in the total number of clicks users have to apply when using current and machine-derived order sets per patient. The new approach potentially reduced both the physical and cognitive workload of order placement by 14–52% in the study sample. For example, pneumonia patients in the training set needed 73 orders on average, which can be placed with 67 clicks on average if machine-derived order sets were used, instead of 79 clicks on average if current order sets were used. The number of new time intervals is the number of optimal time intervals identified by the algorithm, so that the order set content reflects the clinical needs during the particular time interval. There is no time interval for the current order sets. In the redesigned solution, the number of optimal time intervals ranges from 15 to 20, most of which are 2 h intervals. This suggests that ordering patterns tend to change every 2 h.

To make the comparison of the number of order sets and their size easier, figure 5 shows the number of distinct order sets and the total number of items covered by these order sets by time interval, under current and MCC models using appendectomy patients. As the plot shows, the number of machine-derived order sets tends to be smaller, but the number of distinct order sets in each time interval does not differ significantly. The mean percentage increase in ACR across time intervals is shown in the ‘% increase’ column for training and test sets in table 3. The statistical significance of the percentage increase was calculated using the Mann–Whitney test in R Statistical Package V2.15.1.³⁹ A significant increase in ACR, surpassing 100% in training sets, was recorded for asthma and appendectomy data. The percentage increase for pneumonia was not as high, presumably because of diversity in the patient population, which potentially introduced noise in the optimization process. The average number of comorbidities exhibited

Order Set Content Optimization Platform

Step 1: Select Data

First, please tell us the clinical area and severity level associated with the order sets you are interested in.

1. Please select your area:

appendectomy asthma pneumonia

2. Please select severity level:

minor moderate major

Order Set Content Optimization Platform

Step 2: Confirm

you picked area: pneumonia
you picked severity: moderate

Report

your data look like this:
total number of patients: 96
total number of orders: 7144
orders from a la carte: 4758
orders from order sets: 2386
number of unique order sets used: 55

Order Set Content Optimization Platform

Step 3: Optimize

1. Please enter time interval to focus:

2. Please enter minimal time interval:

3. Please select the cost function

CCC MCC

Figure 4 Order set content computational platform. (A) step 1; (B) step 2; (C) step 3.

Table 2 Order set usage of the selected diagnosis groups in 2011

Diagnosis group	Patients (n)	Orders (n)	Unique order sets used (n)	Times order sets were opened (n)	Order set orders (n (% of total))	Clicks to use order sets (n)	Clicks saved by order sets (n (%))
Asthma	415	32 736	77	3895	11 794 (36)	12 254	460 (-4)
Appendectomy	353	37 557	69	1505	9900 (26)	6170	3730 (38)
Pneumonia	357	34 505	114	2662	8935 (26)	9827	892 (-10)

during the inpatient stay supports this explanation: 4.7 for patients with pneumonia compared with 2.4 for asthma and 2.1 for appendectomy. Clinical input combined with optimization may help achieve greater improvement, especially for complex conditions such as pneumonia.

Sample case

A sample case demonstrates the changes made to order sets based on the three objective function criteria. Orders administered for a patient during one ordering action is shown in table 4 with the current and new order set assignment and default settings. Twelve items ordered for this patient currently come from four order sets (C1–C4). After optimization, these 12 items were reassigned to different order sets (M1–M4, E1–E6, S1–S4), or designated as a la carte items (A) based on MCC, CCCE, and CCCS, respectively. The status quo MCC is 15, CCCE is 20.3, and CCCS is 76. After optimization, MCC is 13 (15.4% lower), CCCE is 12.9 (36.4% lower), and CCCS is 23.1 (69.6% lower). Our algorithm also shortened the length of order sets and turned on OFF items that are often used, and vice versa. For example, we found that more than 80% of the study patients had used ‘initial pulse oximetry continuous’—an OFF item currently—at least once. Hence, our algorithm grouped ‘subsequent pulse oximetry continuous’ and ‘initial pulse oximetry continuous’ into one order set, and defaulted both items to be ON. Also, instead of having an order set such as C1, which contains 63 items currently, our algorithm broke it down into smaller order sets to ease the cognitive burden of reviewing a long list of orders,⁸ many of which may not be selected for most patients.

DISCUSSION

Other approaches to order set design

Order sets are intended to assist users in performing their order placement tasks efficiently and effectively. Current approaches rely on continuous and consistent manual updating to accommodate new knowledge, making it a challenge to adapt to changing workflows and patient requirements. We believe a combination of order set development and order set modification methods can better address the evolving care-delivery needs of diverse groups of patients in distinct care-delivery environments. In a prior, preliminary study, we investigated five models of order set usage: (1) status quo; (2) add a la carte orders to current order sets; (3) change default setting of current order sets; (4) combination of (2) and (3); (5) add order sets created from a la carte items using clustering to (4).²⁴ Results indicated that model (5) had the lowest MCC and CCC, suggesting that combining experts’ consensus with evidence from data is a promising method for increasing usability while maintaining clinical validity. This method preserves experts’ consensus by learning ordering patterns from current order sets while incorporating evidence from actual workflows such that rarely used orders are deleted and frequently used a la carte orders are added to order sets. When scientific knowledge and practice-based evidence both increase at a rapid pace, a complete redesign of the order sets may be necessary at periodic intervals, which can utilize the COWpath approaches developed in this paper.

Limitations

Order sets produced by machine-derived techniques suffer from the disadvantage that they will not capture clinically meaningful

Table 3 Computational results for asthma, appendectomy, and pneumonia

Diagnosis group	Reduction in clicks per patient (n (%))	New time intervals (n)	Order sets per time interval (mean±SD)	Order set items per time interval (mean±SD)	Training set ACR† (mean±SD)			Test set ACR‡ (mean±SD)		
					Current	New	% increase	Current	New	% increase
Asthma										
MCC	9 (14)	15	5±1.21	13±3.44	0.27±0.04	0.67±0.08	196*	0.27±0.03	0.54±0.08	108*
CCCE	14 (18)	15	5±1.41	12±3.61	0.26±0.04	0.63±0.08	156*	0.27±0.03	0.49±0.08	90*
CCCS	72 (52)	16	4±1.29	12±3.35	0.26±0.04	0.59±0.10	141*	0.27±0.03	0.47±0.08	84*
Appendectomy										
MCC	19 (19%)	18	5±1.36	27±5.61	0.32±0.07	0.57±0.12	130*	0.32±0.10	0.35±0.10	23
CCCE	26 (21%)	18	6±1.85	32±7.47	0.32±0.07	0.54±0.11	123*	0.32±0.10	0.34±0.09	34
CCCS	36 (20%)	18	5±1.41	25±5.67	0.32±0.07	0.53±0.11	133*	0.32±0.10	0.32±0.09	20
Pneumonia										
MCC	12 (15%)	20	6±2.19	20±5.32	0.30±0.07	0.40±0.04	37*	0.30±0.05	0.35±0.04	30
CCCE	18 (19%)	20	7±2.03	22±5.31	0.30±0.06	0.40±0.03	43*	0.30±0.05	0.38±0.04	32*
CCCS	30 (19%)	18	6±2.31	20±5.63	0.30±0.05	0.30±0.03	26*	0.30±0.04	0.36±0.05	23*

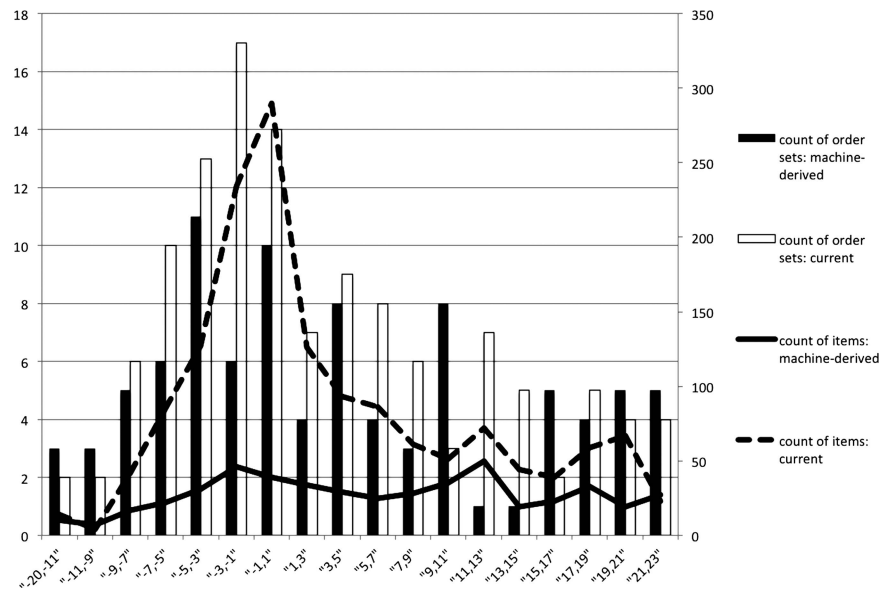
*p<0.05.

†84 patients for asthma; 93 patients for appendectomy; 96 patients for pneumonia.

‡208 patients for asthma; 57 patients for appendectomy; 83 patients for pneumonia.

ACR, average coverage rate; CCCE, cognitive-click-cost-expert-estimate; CCCS, cognitive-click-cost-survey-estimate; MCC, mouse click cost.

Figure 5 Count of order sets and items: current versus mouse click cost model, appendectomy.



but rare events. Therefore, it is important to have experts evaluate order sets and, if necessary, embed known scientific evidence and guidelines into them. In addition, the detailed nature of treatment options and possibility of many other complications for each condition play an important role in the success of order set development. Large variations in order placement patterns associated with varied and multiple complications remains the biggest challenge, as illustrated by pneumonia orders. Changing certain parameters in the optimization method, applying different clustering methods according to condition type, or heuristic approaches for convergence may lead to even better results such as improved coverage rates. Moreover, the retrospective data used to build order sets in this paper are more or less influenced by the current order sets, since users are initially trained to place orders using the current order sets. Users may have learned to avoid some flaws in the current system, and incorporate workarounds that are not the best practices.

Provider variability is another factor that may affect the data. It is not uncommon for orders to be placed by interns or residents who may be biased because of lack of knowledge or experience, or both. Hence, the models and methods proposed in this study are a first step towards automating the redesign of order sets on a periodic basis. The results of our methods must be reviewed for clinical appropriateness based on evidence or by an institutional clinical effectiveness group. In addition, since our study uses only 1 year’s worth of data, from 2011, during which the order set template also remained unchanged, it is likely that biases due to provider variability may have limited impact in this study. A future study may explore variations in usage of order sets across provider groups and across years to classify changes as those based on evidence, poor design, or bias.

Future steps

Testing the machine-derived order sets by actual users in realistic settings is a vital next step, which can be conducted first as a pilot study with interested users and in a simulated environment. Clinical review of the order sets by experts needs to be conducted before actual clinical use of machine-derived order sets. Also, since machine-derived order sets learn from actual practices, regular content update and revision should be followed to keep order sets up to date with changes in the workflow and maintain clinical validity. Updates of the order set template can follow the cycle described in this paper, starting with relearning of the workflow using the COWpath approach, followed by expert evaluation.

CONCLUSION

This paper focuses on a critical, unintended consequence of IT in healthcare delivery, the excessive physical and cognitive workload on end users due to poor usability in the context of using order sets within CPOE. We propose a ‘paving the COWpath’ approach by modeling a two-stage optimization method embedded with a machine learning technique to develop data-driven order sets that used retrospective order data from the pediatric inpatient setting. An evaluation of the new approach provides evidence of a successful reduction in physical and cognitive workloads as well as an increase in order set coverage for diverse patient conditions. The results indicate that automated

Table 4 Sample case

Item	Current		Machine-derived			
	Order set	Default	MCC	CCCE	CCCS	Default
Admit to	C1	ON	M1	E1	S1	ON
Height	C2	ON	M2	E2	S2	OFF
Weight	C1	ON	M2	E2	S2	OFF
Notify MD for oxygen saturations	C1	OFF	A	E3	S3	ON
Notify MD for TPR	C1	OFF	A	E2	S2	OFF
Regular (≥4 years) diet	C1	OFF	A	E3	S3	OFF
Up Ad Lib	C1	OFF	A	E4	S3	OFF
Vital signs	C1	OFF	M1	E1	S1	ON
Subsequent oxygen therapy	C3	ON	M3	E5	S1	OFF
Initial oxygen therapy	C3	OFF	M3	E5	S1	OFF
Subsequent pulse oximetry continuous	C4	ON	M4	E6	S4	ON
Initial pulse oximetry continuous	C4	OFF	M4	E6	S4	ON

A, a la carte; Ad Lib, without restraint or imposed limit; CCCE, cognitive-click-cost-expert-estimate; CCCS, cognitive-click-cost-survey-estimate; MCC, mouse click cost; MD, medical doctor; TPR, temperature pulse respiration.

development is time efficient and able to identify commonly co-placed orders that are also supported by evidence. In this paper, the optimization focused only on reducing physical and cognitive workload, but it is possible to add other developmental criteria as well. Also, while our methods were built on the basis of CHP orders, the data-driven approach is generalizable for different diagnosis groups and workflows, with minor modifications. Hence, we believe that it can be a powerful method to use as a starting point for informed order set review and redesign and, in some instances, to replace the traditional manual development. The easy implementation of our methods can facilitate ongoing update of order set content by healthcare organizations and allow treatment quality to stay up to date with the evolving care delivery standards. We anticipate that data-driven order set design will facilitate order sets to be developed on the basis of current best practices, potentially leading to increased order set acceptance, ordering efficiency, and improved patient safety.

Correction notice This article has been corrected since it was published Online First.

Acknowledgements Dr James E Levin passed away on 11 February 2013. We are greatly indebted to his vision, contributions and support that made this study possible.

Contributors JEL initiated the study and obtained data from the Children's Hospital of Pittsburgh of UPMC. Study design and implementation were conducted jointly by YZ, RP, and JEL. YZ performed the data analysis and drafted the paper, which was revised and confirmed carefully by RP and JEL. RP revision of the paper mainly focused on the analytical components, whereas JEL reviewed the clinical aspects of the paper. Data interpretation was performed jointly by the three authors.

Competing interests None.

Ethics approval This study was designated as exempt by the University of Pittsburgh Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data are the property of the Children's Hospital of Pittsburgh of UPMC.

REFERENCES

- Middleton B, Bloomrosen M, Dente MA, et al. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. *J Am Med Inform Assoc* 2013;20(e1):e2–8.
- Jung M, Riedmann D, Hackl WO, et al. Physicians' perceptions on the usefulness of contextual information for prioritizing and presenting alerts in Computerized Physician Order Entry systems. *BMC Med Inform Decis Mak* 2012;12:111.
- Candy PC. Preventing "information overdose": developing information-literate practitioners. *J Contin Educ Health Prof* 2000;20:228–37.
- Westbrook JL, Baysari MT, Li L, et al. The safety of electronic prescribing: manifestations, mechanisms, and rates of system-related errors associated with two commercial systems in hospitals. *J Am Med Inform Assoc* 2013;20:1159–67.
- Harrison MI, Koppel R, Bar-Lev S. Unintended consequences of information technologies in health care—an interactive sociotechnical analysis. *J Am Med Inform Assoc* 2007;14:542–9.
- Horsky J, Kaufman DR, Oppenheim MI, et al. A framework for analyzing the cognitive complexity of computer-assisted clinical ordering. *J Biomed Inform* 2003;36:4–22.
- Schumacher RM, Lowry SZ. NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records http://www.nist.gov/itl/hit/upload/Guide_Final_Publication_Version.pdf [updated 2010; cited 2013 08/09/2013].
- Patel VL, Kushniruk AW, Yang S, et al. Impact of a computer-based patient record system on data collection, knowledge organization, and reasoning. *J Am Med Inform Assoc* 2000;7:569–85.
- Kushniruk AW, Kaufman DR, Patel VL, et al. Assessment of a computerized patient record system: a cognitive approach to evaluating medical technology. *MD Comput* 1996;13:406–15.
- Levin JE. Paving the COWpath: Order Set Optimization at The Children's Hospital of Pittsburgh of UPMC. Mayo Quality & Systems Engineering Conference; May 16, 2012; Rochester, MN, 2012.
- Maguire D. Neural cowpaths in the NICU. *Neonatal Netw* 2002;21:61–2.
- UPMC TCsHoPo. Barcode Technology/Positive Patient Identification <http://www.chp.edu/CHP/barcode+technology> [cited 2013 08/09/2013].
- Nissen ME. An experiment to assess the performance of a redesign knowledge system. *J Manag Inform Syst* 2000;17:25–43.
- Dienes Z, Perner J. A theory of implicit and explicit knowledge. *Behav Brain Sci* 1999;22:735–808.
- Reber AS. Implicit learning and tacit knowledge. *J Exp Psychol Gen* 1989;118:219–35.
- Eraut M. Non-formal learning and tacit knowledge in professional work. *Br J Educ Psychol* 2000;70:113–36.
- Goh JM, Gao GG, Agarwal R. Evolving work routines: adaptive routinization of information technology in healthcare. *Inform Syst Res* 2011;22:565–85.
- Jacobs BR, Hart KW, Rucker DW. Reduction in clinical variance using targeted design changes in computerized provider order entry (CPOE) order sets: impact on hospitalized children with acute asthma exacerbation. *Appl Clin Inform* 2012;3:52–63.
- Kaushal R, Bates DW, Landrigan C, et al. Medication errors and adverse drug events in pediatric inpatients. *JAMA* 2001;285:2114–20.
- Payne TH, Hoey PJ, Nichol P, et al. Preparation and use of preconstructed orders, order sets, and order menus in a computerized provider order entry system. *J Am Med Inform Assoc* 2003;10:322–9.
- McDonald CJ, Overhage JM, Mamlin BW, et al. Physicians, information technology, and health care systems: a journey, not a destination. *J Am Med Inform Assoc* 2004;11:121–4.
- Wright A, Feblowitz JC, Pang JE, et al. Use of order sets in inpatient computerized provider order entry systems: a comparative analysis of usage patterns at seven sites. *Int J Med Inform* 2012;81:733–45.
- Ahmad A, Teater P, Bentley TD, et al. Key attributes of a successful physician order entry system implementation in a multi-hospital environment. *J Am Med Inform Assoc* 2002;9:16–24.
- Zhang Y, Levin JE, Padman R. Toward order set optimization using click cost criteria in the pediatric environment. 46th Hawaii International Conference on System Sciences (HICSS), 7–10 Jan. 2013:2575–84.
- Zhang Y, Levin JE, Padman R. Data-driven order set generation and evaluation in the pediatric environment. *AMIA Annual Symposium Proceedings* 2012; 1469–78.
- Zhang Y, Padman R, Levin JE. *Clustering Methods for Data-driven Order Set Development in the Pediatric Environment*. Phoenix, AZ: INFORMS 2012 DM-HI Workshop, 2013.
- Zhang Y, Padman R, Levin JE. Reducing Provider Cognitive Workload in CPOE Use: Optimizing Order Sets. the 14th World Congress on Medical Informatics, Copenhagen, Denmark, 2013.
- Chan J, Shojania KG, Easty AC, et al. Usability evaluation of order sets in a computerized provider order entry system. *BMJ Qual Saf* 2011;20:932–40.
- McAlearney AS, Chisolm D, Veneris S, et al. Utilization of evidence-based computerized order sets in pediatrics. *Int J Med Inform* 2006;75:501–12.
- Wright A, Sittig DF. Automated development of order sets and corollary orders by data mining in an ambulatory computerized physician order entry system. *AMIA Annual Symposium Proceedings* 2006;819–23.
- Munasinghe RL, Arsene C, Abraham TK, et al. Improving the utilization of admission order sets in a computerized physician order entry system by integrating modular disease specific order subsets into a general medicine admission order set. *J Am Med Inform Assoc* 2011;18:322–6.
- Klann J, Schadow G, McCoy JM. A recommendation algorithm for automating corollary order generation. *AMIA Annual Symposium Proceedings* 2009;333–7.
- Hulse NC, Del Fiol G, Bradshaw RL, et al. Towards an on-demand peer feedback system for a clinical knowledge base: a case study with order sets. *J Biomed Inform* 2008;41:152–64.
- Avansino J, Leu MG. Effects of CPOE on provider cognitive workload: a randomized crossover trial. *Pediatrics* 2012;130:e547–52.
- Simon HA. Rational choice and the structure of the environment. *Psychol Rev* 1956;63:129–38.
- Averill RF, Goldfield N, Hughes JS, et al. All patient refined diagnosis related groups (APR-DRGs) Version 20.0 Methodology Overview 2003 08/09/2013.
- Hidalgo CA, Blumm N, Barabasi AL, et al. A dynamic network approach for the study of human phenotypes. *PLoS Comput Biol* 2009;5:e1000353.
- Li YCS. Parallel bisecting k-means with prediction clustering algorithm. *J Supercomput* 2007;39:19–37.
- R Core Team. *R: A Language and Environment for Statistical Computing*. 2.15.1 edn, Vienna, Austria: R Foundation for Statistical Computing; 2012..