Formative Usability Testing Reduces Severe Blood Product Ordering Errors

Evan W. Orenstein[1,2] Jeanne Boudreaux[1,3] Margo Rollins[3,4] Jennifer Jones[3] Christy Bryant[5] Dean Karavite[6] Naveen Muthu[6] Jessica Hike[5] Herb Williams[5] Tania Kilgore[5] Alexis B. Carter[7] Cassandra D. Josephson[1,3,4,7]

[1] Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia, United States
[2] Division of Hospital Medicine, Children's Healthcare of Atlanta, Atlanta, Georgia, United States
[3] Aflac Cancer and Blood Disorders Program, Children's Healthcare of Atlanta, Atlanta, Georgia, United States
[4] Department of Pathology and Laboratory Medicine, Center for Transfusion and Cellular Therapies, Emory University School of Medicine, Atlanta, Georgia, United States
[5] Information Services and Technology, Children's Healthcare of Atlanta, Atlanta, Georgia, United States
[6] Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States
[7] Department of Pathology and Laboratory Medicine, Children's Healthcare of Atlanta, Atlanta, Georgia, United States

Address for correspondence Evan W. Orenstein, MD, 1405 Clifton Road, Atlanta, GA 30322, United States (e-mail: Evan.Orenstein@emory.edu).

Appl Clin Inform 2019;10:981–990.

Abstract

Background Medical errors in blood product orders and administration are common, especially for pediatric patients. A failure modes and effects analysis in our health care system indicated high risk from the electronic blood ordering process.

Objectives There are two objectives of this study as follows:
(1) To describe differences in the design of the original blood product orders and order sets in the system (original design), new orders and order sets designed by expert committee (DEC), and a third-version developed through user-centered design (UCD).
(2) To compare the number and type of ordering errors, task completion rates, time on task, and user preferences between the original design and that developed via UCD.

Methods A multidisciplinary expert committee proposed adjustments to existing blood product order sets resulting in the DEC order set. When that order set was tested with front-line users, persistent failure modes were detected, so orders and order sets were redesigned again via formative usability testing. Front-line users in their native clinical workspaces were observed ordering blood in realistic simulated scenarios using a think-aloud protocol. Iterative adjustments were made between participants. In summative testing, participants were randomized to use the original design or UCD for five simulated scenarios. We evaluated differences in ordering errors, time on task, and users' design preference with two-sample *t*-tests.

Results Formative usability testing with 27 providers from seven specialties led to 18 changes made to the DEC to produce the UCD. In summative testing, error-free task

Keywords
- electronic health records
- blood transfusion
- clinical decision support
- usability
- human–computer interaction

received
August 18, 2019
accepted after revision
November 9, 2019

© 2019 Georg Thieme Verlag KG
Stuttgart · New York

DOI https://doi.org/10.1055/s-0039-3402714.
ISSN 1869-0327.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

completion for the original design was 36%, which increased to 66% in UCD (30%, 95% confidence interval [CI]: 3.9–57%; $p = 0.03$). Time on task did not vary significantly.

Conclusion UCD led to substantially different blood product orders and order sets than DEC. Users made fewer errors when ordering blood products for pediatric patients in simulated scenarios when using the UCD orders and order sets compared with the original design.

Background and Significance

Ordering blood products for transfusion is a complicated task. Medical errors, related to human factors and information technology, have been frequently reported in transfusion safety surveillance systems.[1] Providers must consider multiple factors when ordering blood products. These include the volume and transfusion rate of the blood product,[2,3] whether especially prepared products (e.g., irradiated and washed) are required based on the patient's medical condition(s),[4] and workflow (e.g., preparing blood for immediate use vs. future transfusion during a procedure).[5] Transfusion-related errors can be benign or can result in long-term adverse sequelae or death. Clinical decision support (CDS) focused on the transfusion ordering process has reduced inappropriate transfusions in adults.[6,7] Blood product orders for pediatric patients are more complex than for adults, primarily due to the small size of children (increasing the risk of under- or over transfusion), and the relative immunocompromised status of infants less than 6 months of age. As a result, children are at higher risk of adverse outcomes because they may not tolerate errors that would be less impactful to adults.[8] During a failure modes and effects analysis (FMEA)[9] at Children's Healthcare of Atlanta, the five potential failure points with highest risk predictive number were all associated with the ordering process, prompting an organization-wide effort to standardize blood product orders, and order sets.

Electronic order sets can improve adherence to evidence-based practices, reduce variation in care, and minimize cognitive load.[10–12] However, poorly designed order sets can unintentionally and negatively impact practice patterns[13] and/or can lead to medical errors.[14] The "five rights" framework for CDS from Osheroff et al notes that good decision support must provide the right information to the right person in the right format through the right channel at the right time in workflow.[15,16] In the absence of formal usability testing, individuals' designing orders and order sets may not sufficiently understand how users will interact with the system to complete tasks for realistic scenarios, leading to an inadequate application of the "five rights" and misalignment between order sets and workflow.[17] User-centered design informed by simulated testing can lead to substantial design adjustments that ultimately improve adherence to evidence-based practices while reducing the user's cognitive workload.[7,18]

In this study, the original electronic blood ordering process (original design, see ►Supplementary Figs. S1A–C and S2A and B; available in the online version) was initially redesigned by an internal multidisciplinary expert committee (design by expert committee; DEC, see ►Supplementary Fig. S1D, available in the online verion). After simulated testing of these redesigned order sets identified persistent user errors in the ordering process, iterative adjustments were made resulting in a second, user-centered redesign of orders, and order sets (user-centered design; UCD, see ►Supplementary Figs. S1E and F and S2C and D, available in the online version). The aims of this study were to (1) compare the types of changes made to orders and order sets between the original design, DEC, and UCD; (2) compare the number and severity of blood ordering errors for the original design versus UCD in summative usability testing; and (3) determine if the UCD improved users' clinical knowledge of appropriate special processing for blood products compared with the original design.

Methods

We began with an original design of blood product orders and order sets, which was initially redesigned through an expert committee and subsequently via user-centered design (►Fig. 1).

Design by Expert Committee

We performed an FMEA with a multidisciplinary group of stakeholders that included representatives from surgery, anesthesia, hematology–oncology, hospital medicine, clinical informatics, information technology, and laboratory including transfusion medicine. Blood product order sets in the Electronic health record (EHR; Epic Systems, Verona, Wisconsin, United States) were identified as high-risk sources of errors. All blood product orders and order sets were considered in scope for this project with the exception of emergency release orders and massive transfusion protocols. The expert committee reviewed existing blood product orders and order sets and proposed adjustments, which were built into orders and order sets in an EHR test environment, and comprised the DEC.

User-Centered Design

Formative usability testing was performed starting with the DEC orders and order sets using a think-aloud protocol.[19] Front-line providers from a purposive set of specialties who were on service in their usual patient care areas were asked to participate in a 10-minute usability test. Providers who agreed to participate (participants) were instructed to interrupt the test if they had to attend to patient care and to

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

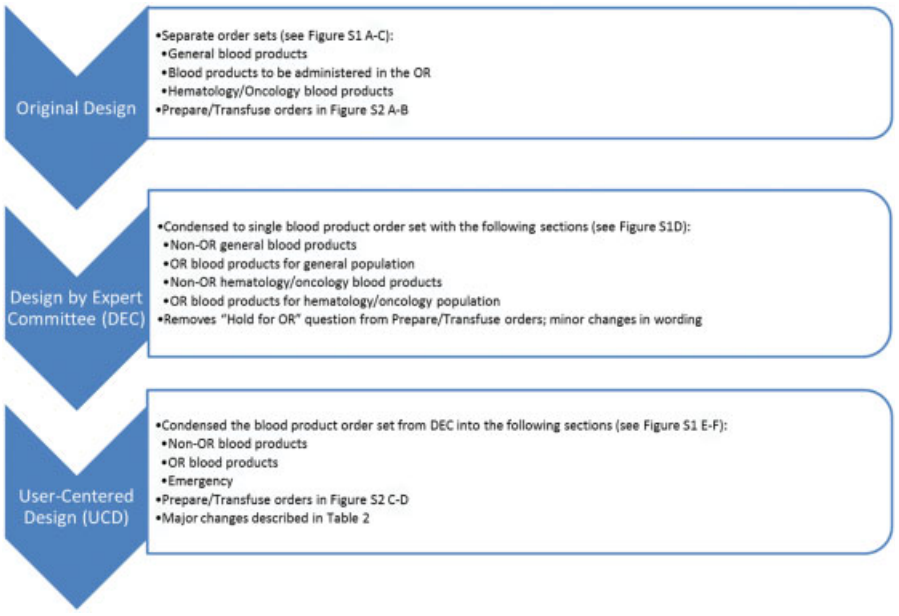

Fig. 1 Progression of order and order set design. We began with the original blood product orders and order sets, in which blood product orders were distributed across three main order sets (see ►**Supplementary Figs. S1 A–C** and **S2 A** and **B**, available in the online version). In design by expert committee (DEC), we condensed these into a single order set and made minor adjustments in the "prepare" and "transfuse" orders (see ►**Supplementary Fig. S1 D**, available in the online version). In user-centered design (UCD), we condensed the order set even further into fewer sections and made major changes to both the order set structure and the "prepare" and "transfuse" orders (see ►**Supplementary Figs. S1 E** and **F** and **S2C** and **D**, available in the online version). OR, operating room.

resume the test when they were again available. Participants were provided a scenario appropriate to their clinical specialty and asked to order blood products in an EHR test environment which, except for the redesigned blood orders being tested, had identical infrastructure and function to the production EHR environment. While the participants performed the usability test, they were instructed to verbally describe their goals, thoughts, and actions out loud. Ordering activities were observed by a transfusion safety specialist (J.J.) and a clinical informaticist (E.W.O.) who took notes on users' actions within the EHR test environment and identified any errors that occurred, with special attention to the volume, rate, and special processing of blood product orders. After completing a set of one to four scenarios, each participant was shown his/her ordering errors and asked to suggest alternative designs that might have prevented errors and/or reduced confusion. The observers (J.J. and E.W.O.) wrote down key comments which were reviewed with the user at the end of each session (member checking).[20] Sessions were not audio recorded. Iterative adjustments were made in the EHR test environment to the blood orders and order sets between participants, generally after two or three providers had made the same or similar error. Formative testing was stopped after five unique providers did not make any errors across 10 scenarios and had no new suggestions for design improvement to the orders or order sets which comprised the UCD.

Summative Usability Test

Usability of the finalized UCD orders and order sets was evaluated and compared with the original blood product ordering process through summative usability testing.[19] Summative usability testing participants were recruited in a similar manner to formative usability testing participants, while they were in their usual clinical setting and given the same instructions regarding necessary interruptions for patient care duties. However, in summative usability testing participants were told that participation was expected to take 20 to 30 minutes and were not explicitly asked to think aloud. No participants from formative usability testing participated in summative usability testing. Summative usability testing was completed in five steps (►**Fig. 2**).

Step 1: participants answered five multiple choice quiz questions asking which special processing requests would be appropriate in each of five clinical scenarios (►**Table 1**). The clinical scenarios for the summative usability test were created by a pediatric hematologist (J.B.) and two transfusion medicine specialists (M.R. and C.J.) based on a combination of well-known and common ordering errors, as well as errors, directly observed during formative usability testing.[21]

Step 2: participants were randomized to use either the original orders and order sets or the UCD orders and order sets and instructed to place the appropriate blood product orders in the EHR test environment for the first four of the five scenarios listed in ►**Table 1**. As with the formative usability test, the EHR test environment had the same structure and functions as the production environment with the noted exception of the blood orders and order sets being tested. Participants completed all four of these scenarios with the same set of orders and order sets. The details of each clinical scenario were read to the participant; these were not built into the records of the test patients in the EHR

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

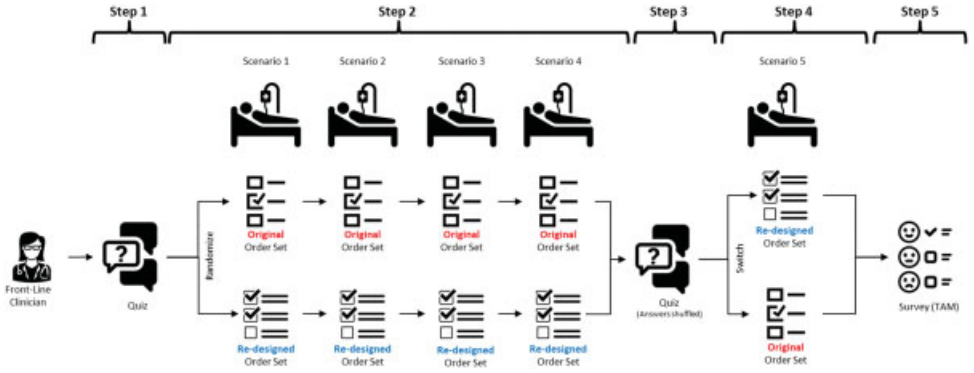

Fig. 2 Summative testing design. Participants (1) completed a quiz assessing their knowledge of appropriate special processing requests for blood products, (2) completed blood ordering tasks for 4 scenarios with either the original design or UCD orders and order set, (3) retook the same quiz with questions and answers shuffled, (4) completed one additional blood ordering scenario with the opposite design, and (5) answered survey questions on their preference between the designs and the perceived usability, ease of use, efficiency, and overall rating of the UCD. UCD, user-centered design.

test environment. Time on task was recorded from the end of the verbal scenario prompt until the participant signed the orders in the EHR test environment, with pauses during interruptions for patient care needs. Ordering errors were also recorded by a physician observer and were classified as follows:

- *Moderate ordering error*: error in a signed order that, if followed exactly by both the blood bank technologists preparing the product(s) and the nurse administering the transfusion, would lead to an unnecessary special processing request (e.g., order for cytomegalovirus-negative blood in a patient who does not require it) or a delay in nonurgent transfusion (e.g., failed to place a transfuse order for the nurse when the provider intended to transfuse the patient). These errors can result in unnecessary work or mild delays but do not generally result in patient harm.
- *Severe ordering error*: error in a signed order that, if followed exactly by both the blood bank technologists preparing the product(s) and the nurse administering the transfusion, would lead to an over transfusion (higher volume of blood or faster rate of infusion delivered to the patient than was intended), under transfusion, or missed special processing request (e.g., nonirradiated blood transfused to a patient with a clinical condition, such as T-cell dysfunction, in which administering nonirradiated blood is contraindicated). These errors can cause significant patient harm or death.

Task completion with no errors was defined as signing blood product orders with no moderate or severe errors as described above.

Step 3: all participants (original design and UCD groups) retook the same five-question quiz with the sequence of questions and answers shuffled. We compared the difference in quiz grades between steps 1 and 3 for participants randomized to the UCD orders and order sets versus those

Table 1 Clinical scenarios for summative usability testing

Scenario	Correct special processing requests
A 15-year-old girl with menorrhagia has a history of hives that appeared with a transfusion of pRBCs that she received two months ago. The transfusion was completed with administration of Benadryl, and there were no signs of anaphylaxis (GI distress, hypotension and respiratory distress). She is presenting with symptomatic anemia and needs another transfusion of pRBCs.	None
You are caring for a premature infant born at 27 weeks of gestation (birth weight 1,100 g) who is now at 35 weeks adjusted gestational age (weight 2,500 g). This patient was recently diagnosed with necrotizing enterocolitis, and you would like to treat this patient with a blood transfusion.	Irradiated
An 18 year-old with acute myeloid leukemia has received multiple platelet transfusions for platelet counts of less than 10,000/μL (<10 k/μL). Two days ago, his 1-hour postplatelet transfusion count remained <10 k/μL. Transfusion medicine was consulted and confirmed that the patient has antibody mediated platelet refractoriness. He needs another platelet transfusion today for a platelet count of 6 k/μL.	Irradiated HLA-matched
A 12-year-old with hemoglobin SC disease presents to the Emergency Department for management of splenic sequestration. You would like to transfuse the patient an aliquot of 5 mL/kg pRBCs. Historically, this patient has no autoantibodies or alloantibodies.	Phenotypically similar
A 7-month-old boy diagnosed with DiGeorge's syndrome and severe combined immunodeficiency presents in aplastic crisis and needs a transfusion of pRBCs.	Irradiated

Abbreviations: GI: gastrointestinal; HLA: human leukocyte antigen; pRBCs: packed red blood cells.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

randomized to the original design. We hypothesized that participants randomized to use the UCD orders and order sets would improve their knowledge of appropriate special processing requests, leading to a greater improvement in quiz grades than those randomized to use the original design.

Step 4: participants used the opposite design to order blood products in the EHR test environment for the fifth scenario (row 5 of ►Table 1). For example, if a participant was randomized to use the original orders and order sets for the first four scenarios in step 2, they used the UCD orders and order set for the fifth scenario in step 4, and vice versa. This scenario was not graded for task completion; rather it was intended to expose participants to the opposite design so they could rate their design preference in the subsequent step.

Step 5: the authors developed a short survey focused on the themes of perceived ease of use and usefulness from the technology acceptance model.[22] Participants (1) rated their preference between the original design versus that developed via UCD on a visual analog scale of −1 (completely prefer original design) to 1 (completely prefer UCD) and (2) rated their agreement on a 9-point Likert's scale with statements that the UCD orders and order sets were (1) easy to use, (2) useful, (3) allowed the user to perform tasks efficiently, and (4) overall satisfied (►Supplementary Fig. S3, available in the online version).

The rationale behind this design which included four scenarios of one design and then only one scenario of the opposite design was (1) the need to complete multiple scenarios to capture the variety of error types we were seeing, (2) the goal of testing the effect of each design on participants' knowledge of special processing requests, where we wanted participants to have only seen one design between the pre- and postquizzes, and (3) the goal of limiting the total number of scenarios to minimize participant time away from patient care. We therefore did not perform a balanced or incomplete block design.

Differences in task completion rates, ordering error rates, and change in quiz grades were compared using two-sample *t*-tests. Differences in specialty and clinical experience between the two groups were assessed with Fisher's exact test. All statistical tests were completed in R version 3.5.2.[23] This work was felt to be primarily focused on quality improvement and therefore deemed nonhuman subjects research by the Institutional Review Board of Children's Healthcare of Atlanta.

Results

Design by Expert Committee

Twenty-nine stakeholders were identified across clinical, laboratory, and information technology departments. During committee review, it was felt that users frequently used the incorrect order set for the intended patient population (e.g., using the general population order set for hematology/oncology patients) which could lead to errors. Therefore, multiple blood product order sets for different clinical settings (e.g., general inpatient floors, surgery, and oncology) were condensed into a single unified blood administration order set (►Supplementary Fig. S1, available in the online version) with sections organized similarly to the original order sets. Minor changes in wording were also made to improve clarity of the order which instructs the transfusion service (blood bank) to crossmatch, prepare and allocate blood products to a specific patient (the "prepare" order), as well as the order to the nurse to transfuse the patient with the prepared blood product(s) (the "transfuse" order).

User-Centered Design

Formative usability testing was completed in eight half-day sessions (∼30 hours) by a clinical informaticist (E.W.O.), transfusion safety specialist (J.J.), and EHR training expert (H.W.) with 27 providers from seven specialties using 30 unique clinical scenarios and a total of 70 scenarios administered (range: 1–4 scenarios per user). Key error types identified during formative testing included transfusion volume ordering errors, errors in selection of special processing requests, and errors when preparing and transfusing two separate aliquots of blood product from the same donor unit (common in pediatric transfusion). During formative usability testing, a total of 18 design changes were made to avoid observed errors. Six of these changes were considered major (►Table 2) and included major adjustments to the order set structure (►Supplementary Fig. S1; available in the online version), as well as changes to the "prepare" and "transfuse" orders (►Supplementary Fig. S2; available in the online version) which required downstream changes in the transfusion laboratory information system. Another 12 changes were considered minor (e.g., changing wording, allowing multiselect instead of single select on order questions).

Summative Usability Testing

Fifteen front-line ordering clinicians participated in summative usability testing. Each participant completed four scenarios in *step 2* of summative usability testing for a total of 60 scenarios primarily focused on task completion without errors. Seven participants were randomized to the original design (28 scenarios), while eight participants were randomized to the UCD orders and order sets (32 scenarios). Participants included front-line ordering providers from general pediatrics, cardiology, hematology/oncology, neonatology, critical care, and surgery. Four participants (two in each group) had <2 years of postgraduate clinical experience, five had 3 to 4 years of experience, three had 5 to 9 years of experience, and one participant had ≥10 years of clinical experience. There were no significant differences in specialty distribution or experience between the groups by Fisher's exact test.

Compared with the original design, task completion without any errors significantly improved with the UCD orders and order set compared with the original design (UCD: 65.6% vs. original design: 35.7%; $p = 0.03$), and the rate of severe order errors significantly decreased (UCD: 0.125 severe errors per scenario vs. original design: 0.5 severe errors

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Table 2 Major design changes made to the original design and DEC orders and order sets based on formative usability testing, resulting in the development of UCD orders and order sets

Design change	Rationale
Original design and DEC: when user selects blood product, both "prepare" and "transfuse" orders are default checked. *UCD:* When user selects blood product, "prepare" order is default checked but not "transfuse" order. If user signs "prepare" order without pre-op indication and no "transfuse" order present, sees alert to confirm.	In formative testing, 2 of 3 providers signaled nurse to transfuse when scenario did not call for it with original design and DEC. With UCD, no transfuse orders placed when scenario did not call for it, but 1 of 16 providers did not signal nurse to transfuse when scenario did call for it.
Original design and DEC: when preparing blood products in mL, always ask volume of 1st aliquot and volume of 2nd aliquot[a]. *UCD:* ask for 1st aliquot volume, then ask "2nd aliquot needed?" If yes, cascade to 2nd aliquot volume.	In scenario requiring only 1 aliquot, 2 users entered the volume for both 1st and 2nd aliquot with the original design and DEC, which would lead to over transfusion.
Original design: separate order sets for general, OR, and Hem/Onc populations. *DEC:* collapsed into 1 order set with sections for "non-OR, general," "non-OR, Hem/Onc," "OR general," "OR Hem/Onc." *UCD:* single order set with "non-OR," "OR," and "emergency O-negative" sections; no special Hem/Onc sections.	In scenarios of Hem/Onc patients, 2 participants made order selections in the wrong subsection for the clinical scenario ("OR" vs. "non-OR"), leading to inaccurate use of "transfuse" order.
Original design and DEC: User can order split aliquots[a] when ordering by volume, not by units. *UCD:* allow option to order 1/2 unit packed red blood cells or platelets; when 1/2 unit selection, cascade option of "2nd aliquot needed?"	Anesthesia and Hem/Onc providers described use cases for ordering 1/2 unit packed red blood cells. In original design and DEC, these providers wrote instructions to split a single unit in the order comments, which were sometimes missed by blood bank technologists and risks over-transfusion.
Original design and DEC: ask providers to choose each special processing request (i.e., irradiated, washed, phenotypically similar, CMV-negative, etc.). *UCD:* ask providers if indications present for each special processing request and provide buttons to select indication.	Multiple providers did not know when to select special processing requests because they didn't know the indications requiring special processing and made frequent errors. When participants were presented with a list of indications, accuracy and reported satisfaction increased.
Original design and DEC: when ordering 2 aliquots, order 2 occurrences of "transfuse" order so nurse can document administration of each aliquot. *UCD:* each aliquot requires separate "Transfuse" order; require user to specify nursing tasks before 2nd aliquot (e.g., "notify MD," "Draw Labs").	Hem/Onc participants noted that nurse would often transfuse both aliquots without seeing instructions to notify provider or draw laboratories. Original design and DEC provided affordance to administer and document 2nd aliquot. UCD removed this affordance and made instructions for tasks between aliquots more prominent for nurse.

Abbreviations: CMV, cytomegalovirus; DEC, design by expert committee; Hem/Onc, hematology/oncology; OR, operating room; UCD, user-centered design.

[a]Smaller pediatric patients cannot receive full units of blood product and are therefore transfused in mL aliquots. The blood product is allocated to the individual patient and reserved for future aliquots as needed for the same patient to minimize antigen exposure and the risk of developing alloantibodies, especially for patients requiring frequent transfusions.

per scenario; $p = 0.002$; ►Table 3). The frequency of moderate order errors (UCD: 0.28 moderate errors per scenario vs. original design 0.29 moderate errors per scenario; $p = 0.97$) and mean time per scenario (UCD: 95.4 seconds vs. original design 87.4 seconds; $p = 0.57$) were not significantly different between the two designs. On a visual analog scale of −1 (completely prefer original design) to 1 (completely prefer UCD), participants preferred the UCD with a mean score of 0.56 (95% CI: 0.23–0.89). Overall, 71% of participants agreed or strongly agreed that they were satisfied with the UCD system, 79% agreed that it was easy to use, 71% agreed that it allowed efficient task performance, and 79% agreed that it was useful.

There was a trend toward greater improvement in quiz grade for those using the UCD orders and order set, but this was not significant. Of note, with eight participants in the UCD group, seven in the original design group, and a standard deviation in quiz grade across the whole sample of 17%, we would have been 80% powered to detect a quiz grade difference of 27% or greater.[24]

Discussion

User-centered design of the blood product ordering process for pediatric patients through scenario-based, formative usability testing significantly reduced severe ordering errors

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Table 3 Summative testing results

	Original design (7 participants, 28 scenarios)	UCD (8 participants, 32 scenarios)	Difference (95% CI, *p*-value)
Task completion with no errors	35.7%	65.6%	29.9% (3.9–56.5%, $p = 0.03$)
Severe errors (per scenario)	0.5	0.125	−0.375 (−0.58 to −0.17, $p = 0.002$)
Moderate errors	0.29	0.28	−0.005 (−0.29 to 0.29, $p = 0.97$)
Mean time per scenario (sec)	87.4	95.4	8.0 (−37.3 to 21.4, $p = 0.57$)
Mean change in quiz grade	−2.9%	10%	12.9% (−5.4 to 31%, $p = 0.15$)

Abbreviations: CI, confidence interval; UCD, user-centered design.

in simulation-based testing without increasing time on task. Key design changes from the UCD process included (1) asking providers for the patient's clinical indications for special processing of blood products rather than the name of the special processing requests themselves, (2) removing fields that facilitated ordering and administering two aliquots of blood products when two aliquots were unnecessary, and (3) simplifying the structure of the order set to make it easier to know when a "transfuse" order was appropriate. To our knowledge, no standards exist for the appropriate design of electronic blood product orders for pediatric patients or adults. These results demonstrate that some designs may predispose to errors, while adjustments targeting those errors can more safely accommodate pediatric transfusion needs.[7,8]

Even after UCD, we continued to see an error rate in summative usability testing of 34%, with 0.125 severe ordering errors per scenario. This high-error rate is partially because the scenarios created by a pediatric hematologist and transfusion specialist were intentionally difficult and representative of known errors from prior safety incidents, as well as errors, directly observed during formative usability testing. The scenarios were not representative of the most common blood ordering scenarios encountered in general practice. Other studies of complex workflows have found similar error rates even after interface redesign. For example, Horsky and Ramelson demonstrated that a side-by-side medication reconciliation interface reduced errors during a complex task from 1.29 errors per participant to 0.37 errors per participant.[25] Medical errors are more common in complex patients, emphasizing the need for multiple layers of safety in addition to better design of each step in the ordering process; for example, the blood bank at our institution maintains independent lists of patients requiring special processing to double check orders. Nonetheless, based on the Swiss cheese model of patient safety, reducing error rates in each stage of a care process is likely to reduce overall errors.[26]

To reduce resource needs to perform formative usability testing for an operational optimization project, participants were approached in their usual patient care areas. This method of in situ testing and immediate adjustments based on observed errors accelerated recruitment and led to more rapid iterations in the design. Additionally, participants experienced real interruptions due to patient care needs. While this approach may reduce the internal validity of conclusions, since we did not control the ordering environment with laboratory precision and participants may have felt more hurried or distracted by their patient care duties, it may in fact improve the external validity (i.e., likelihood that our results generalize to nonsimulated settings) by increasing the fidelity of the sessions to actual clinical care at low cost.

We hypothesized based on participant feedback during formative testing that the redesign itself might improve users' knowledge of when to order special processing requests without additional educational interventions. We saw a trend in improved quiz grades but it was not significant. This may be due to small sample size as the difference in quiz grades detected was below the level and we had 80% power to detect with our sample. It may also be due to lack of true effect. In a study of resident physician's knowledge of cystic fibrosis and chronic obstructive pulmonary disease before and after implementation of disease-specific order sets, Yu et al[27] found a similar trend of improved test scores in those exposed to the order sets. While some argue that order sets and decision support in general prevents medical trainees from learning,[28] these findings suggest that order sets may actually improve resident education.

Limitations

This study is limited by a small sample size at a single institution with unique considerations and constraints. Providers were used to order blood products in a certain way, and this may have impacted their reactions to novel design approaches. Therefore, the benefits of the specific design changes implemented at our institution may not generalize to other settings where providers have different expectations. Nonetheless, formative usability testing identified many ordering errors which led to a very different final design in our institution than DEC. The benefits of formative testing to inform user-centered design likely do generalize across health care settings as they have across multiple industries.[29,30]

In the summative usability testing section of this study, we compared the UCD orders and order sets to the original design, not the DEC. This choice was based on operational goals to determine if the final design reduced ordering errors when compared with the design to which providers were already accustomed. The frequent error rates with the DEC seen in formative usability, testing limited operational

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

interest in formal evaluation of its performance. However, since the summative comparison was not directly between the UCD and the DEC, it is possible that the DEC approach would have outperformed the UCD orders and order sets. Similarly an expert committee with a different structure may have yielded a different and potentially more effective design.

Additional limitations included having a single clinical informaticist and transfusion specialist observing the test sessions, defining errors, and making order set adjustments. While the clinical informaticist (E.W.O.) had completed a cognitive informatics and EHR usability elective during clinical informatics fellowship, no formally trained human factors expert was involved. This could lead to bias and lack of generalizability based on the individual characteristics of the observers without a degree in human factors engineering. Similarly, we did not evaluate differences in providers who agreed to participate compared with those who did not agree. Additionally, the survey administered to participants to assess their design preference, perceived ease of use, and usability of the UCD was developed by the authors without validation. This may have led to misclassification which in a small sample size could bias the results.

While the test EHR environment mirrored the true EHR in structure and function, the clinical scenarios were not fully detailed in the test EHR environment. Participants were read details of the scenario but did not have the ability to read along making comprehension more difficult for participants who process information visually. These approaches may have led to participant confusion which could explain higher error rates. However, the approach was not different between the two groups, and observers noted that participants spent nearly all of their time on ordering screens rather than seeking information in the chart, decreasing the chance that errors were due to inaccurate information in test patient charts.

Finally, this study only demonstrated benefits in specific simulated scenarios that were developed by clinical experts based on prior errors but never formally validated. Our results may be due to unintended errors in the scenarios leading to participant confusion. Furthermore, the design may have been overfit to the scenarios created by the authors and the targeted error types, but may lead to unanticipated effects when applied more generally to patient care in untested clinical situations.

Conclusion

User-centered design through scenario-based usability testing of a pediatric blood transfusion ordering process reduced severe ordering errors in simulated settings without any impact on ordering efficiency. The design produced after formative usability testing differed substantially from that developed by an expert committee. In situ formative testing can be performed by a small, skilled team within a short period of time (2 weeks) and limited effort (~100 person-hours) and therefore does not create an unreasonable operational burden for organizations with trained staff members. In summative testing, participants randomized to the original design versus UCD orders and order sets had similar changes in scores for quizzes that tested knowledge of when special processing requests were clinically appropriate. Future work will examine the effectiveness of these UCD orders and order sets on reducing ordering errors and their influence on clinical practice patterns and patient outcomes in real clinical settings.

Clinical Relevance Statement

In many organizations, clinical decision support processes and tools, such as order sets, are designed by committees of stakeholders in the absence of usability testing performed by appropriate clinical end users. We developed clinical decision support for pediatric blood product order sets through a stakeholder committee and subsequently through scenario-based formative usability testing. The final order set product after formative usability testing was substantially different from that created by the stakeholder committee alone. Additionally, in summative testing, users made significantly fewer severe ordering errors with the user-centered design of orders and order sets than the original design. These results demonstrate that (1) ordering errors for pediatric blood products are common and (2) order set design through formative usability testing adds additional value to design by stakeholder committee alone.

Multiple Choice Questions

1. A health care organization notices ordering errors in blood product transfusions leading to patient harm. Which of the following approaches to redesigning the blood product ordering process is most likely to induce optimal ordering behavior from front-line clinicians?
   a. A stakeholder committee including representatives from information technology, blood bank, and clinical heads of service.
   b. Scenario-based formative usability testing using a "think aloud" protocol.
   c. Design by the medical director of the blood bank
   d. Allow front-line clinicians to customize their own personal order sets.

   Correct Answer: The correct answer is option b. in this study, design changes based on formative usability testing in which typical users would "think aloud" as they attempt to order blood products for specific clinical scenarios led to a substantially different order set design than stakeholder committee alone. This new design led to significantly fewer severe ordering errors in subsequent simulated scenarios. While stakeholder committees likely have the knowledge about which blood products are appropriate for which clinical scenarios, they may not have a full understanding of front-line clinician workflow or thought processes and may not be able to predict front-line user behavior in response to a specific order set representations. By contrast, while front-line clinicians may have the best understanding of their own workflow,

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

which they could use to customize their own personal order sets, front-line clinicians may have insufficient medical knowledge regarding transfusion practice to ensure that customized personal order sets are appropriate and lead to correct ordering behaviors.

2. At an academic health system, you are placed in charge of developing order sets to improve adherence to certain published guidelines. One of your physician colleagues complains that the trainees (residents and fellows) at your organization may do a better job of implementing the guidelines but will suffer in terms of their education because they will rely on the order set instead of going to look up the guidelines themselves.
   Which of the following is the most appropriate response to your colleague's concerns?
   a. Guideline-based order sets likely do not detract from and may even improve guideline awareness.
   b. It does not matter if physician trainees learn about guidelines so long as they implement their recommendations.
   c. Order sets do not have any effect on prescriber behavior, so physician trainees will have to look up the guideline anyway.
   d. Your colleague is correct and so you should not develop guideline-based order sets at organizations responsible for training new physicians.

Correct Answer: The correct answer is option a. in the summative testing portion of our study, we administered a quiz for participants before and after exposure to the order set, assessing their knowledge of appropriate special processing requests for pediatric blood transfusions. Participants randomized to the UCD order set with an explicit algorithm for choosing special processing requests did not have worse quiz grades compared with participants using the original design with no such guidance. In fact, there was a trend (though not statistically significant) toward improved quiz grades in the group using the UCD orders and order sets. Similar trends have been seen with the introduction of cystic fibrosis and chronic obstructive pulmonary disease order sets.[27]

Protection of Human and Animal Subjects

This work was felt to be primarily focused on quality improvement and therefore deemed nonhuman subjects research by the Institutional Review Board of Children's Healthcare of Atlanta.

Funding

None.

Conflict of Interest

A.B.C. reports personal fees from American Medical Informatics Association, outside the submitted work. N.M. reports equity in Phrase Health, a clinical decision support analytics company. E.W.O. reports other from Phrase Health, outside the submitted work. J.B. has nothing to disclose. M.R. has nothing to disclose.

Acknowledgments

We thank the many clinicians who contributed their time both in formative and summative usability testing to optimize the design of the pediatric blood product order process. We also thank the Information Systems and Technology Department at Children's Healthcare of Atlanta for their extensive efforts to create and test novel blood product ordering designs.

References

1 Rowley M. Errors related to information technology (IT). In: Bolton-Maggs P, ed. Annual SHOT Report: 2017. Manchester, England: Serious Hazards of Transfusion (SHOT). 2018:104–105
2 New HV, Berryman J, Bolton-Maggs PH, et al; British Committee for Standards in Haematology. Guidelines on transfusion for fetuses, neonates and older children. Br J Haematol 2016;175 (05):784–828
3 Goel R, Cushing MM, Tobian AAR. Pediatric patient blood management programs: not just transfusing little adults. Transfus Med Rev 2016;30(04):235–241
4 Zantek ND, Parker RI, van de Watering LM, et al; Pediatric Critical Care Transfusion and Anemia Expertise Initiative (TAXI); Pediatric Critical Care Blood Research Network (BloodNet), and the Pediatric Acute Lung Injury and Sepsis Investigators (PALISI) Network. Recommendations on selection and processing of RBC components for pediatric patients from the pediatric critical care transfusion and anemia expertise initiative. Pediatr Crit Care Med 2018;19(9S, Suppl 1):S163–S169
5 Jones H, Reeve K. Transfusion guidelines in children: II. Anaesth Intensive Care Med 2017;18:546–550
6 Rothschild JM, McGurk S, Honour M, et al. Assessment of education and computerized decision support interventions for improving transfusion practice. Transfusion 2007;47(02):228–239
7 McGreevey JD III. Order sets in electronic health records: principles of good practice. Chest 2013;143(01):228–235
8 Yarahuan JW, Billet A, Hron JD. A quality improvement initiative to decrease platelet ordering errors and a proposed model for evaluating clinical decision support effectiveness. Appl Clin Inform 2019;10(03):505–512
9 Institute for Healthcare Improvement. Quality improvement essentials toolkit. Available at: http://www.ihi.org/resources/Pages/Tools/Quality-Improvement-Essentials-Toolkit.aspx. Accessed December 4, 2019
10 Ahmadian L, Khajouei R. Impact of computerized order sets on practitioner performance. Stud Health Technol Inform 2012; 180:1129–1131
11 Chan AJ, Chan J, Cafazzo JA, et al. Order sets in health care: a systematic review of their effects. Int J Technol Assess Health Care 2012;28(03):235–240
12 Gartner D, Zhang Y, Padman R, Gartner gartnerd D. Cognitive workload reduction in hospital information systems decision support for order set optimization. Health Care Manage Sci 2018;21(02):224–243
13 Kara A, Isaacs AN, Nisly SA. Prescriptions for bedtime sedatives after the introduction of a general admission order set at an academic health center: The potential and pitfalls of order sets. J Patient Saf 2017;13(04):232–236
14 Leu MG, Morelli SA, Chung O-Y, Radford S. Systematic update of computerized physician order entry order sets to improve quality of care: a case study. Pediatrics 2013;131(Suppl 1):S60–S67
15 Osheroff JA, Teich J, Levick D, et al. Improving Outcomes with Clinical Decision Support: An Implementer's Guide. 2nd ed. Chicago, IL: HIMSS Publishing; 2012
16 Greenes RA, Bates DW, Kawamoto K, Middleton B, Osheroff J, Shahar Y. Clinical decision support models and frameworks:

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Seeking to address research issues underlying implementation successes and failures. J Biomed Inform 2018;78:134–143
17 Li RC, Wang JK, Sharp C, Chen JH. When order sets do not align with clinician workflow: assessing practice patterns in the electronic health record. BMJ Qual Saf 2019;28(12):987–996
18 Melton BL, Zillich AJ, Russell SA, et al. Reducing prescribing errors through creatinine clearance alert redesign. Am J Med 2015;128 (10):1117–1125
19 Schumacher RM, Lowry SZ, Locke G, Gallagher PD. NIST guide to the processes approach for improving the usability of electronic health records. Available at: https://www.nist.gov/system/files/documents/itl/hit/Guide_Final_Publication_Version.pdf. Accessed December 4, 2019
20 Ash JS, Smith AC III, Stavri PZ. Performing subjectivist studies in the qualitative traditions responsive to users. In: Friedman CP, Wyatt JC, eds. Evaluation Methods in Medical Informatics. New York, NY: Springer-Verlag; 2006:267–300
21 Russ AL, Saleem JJ. Ten factors to consider when developing usability scenarios and tasks for health information technology. J Biomed Inform 2018;78:123–133
22 Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000;46:186–204
23 R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/
24 Champely S, Ekstrom C, Dalgaard P, et al. Package 'pwr': Basic Functions for Power Analysis. Available at: https://cran.r-project.org/web/packages/pwr/pwr.pdf. Accessed December 4, 2019
25 Horsky J, Ramelson HZ. Cognitive errors in reconciling complex medication lists. AMIA Annu Symp Proc 2017;2016:638–646
26 Reason J. Human error: models and management. BMJ 2000;320 (7237):768–770
27 Yu CH, Stephenson AL, Gupta S. The effect of patient care order sets on medical resident education: a prospective before-after study. BMC Med Educ 2013;13:146
28 Lewis JB Jr, Ryder K. Medical education and decision-support systems. Virtual Mentor 2011;13(03):156–160
29 Miller CA. Lessons from another industry: aviation, usability, and medical device design. Biomed Instrum Technol 2013;47 (Suppl):40–44
30 Kushniruk A. Evaluation in the design of health information systems: application of approaches emerging from usability engineering. Comput Biol Med 2002;32(03):141–149

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.