THIEME
OPEN
ACCESS

Usability Evaluation of Visual Representation Formats for Emergency Department Records

Nathaniel Brown[1,2]   Aboozar Eghdam[1]   Sabine Koch[1]

[1] Department of Learning, Informatics, Management and Ethics, Health Informatics Centre, Karolinska Institutet, Stockholm, Sweden

[2] Department of Emergency Medicine, Medical College of Wisconsin, Milwaukee, Wisconsin, United States

Address for correspondence Nathaniel Brown, MD, MSc, 700 South Park Street, Suite 104, Madison, WI 53715, United States (e-mail: nathanpatrickbrown@gmail.com).

Abstract

Background Integration of electronic information is a challenge for multitasking emergency providers, with implications for patient safety. Visual representations can assist sense-making of complex data sets; however, benefit and acceptability in emergency care is unproven.

Objectives This article evaluates visually focused alternatives to lists or tabular formats, to better understand possible usability in Emergency Department Information System (EDIS).

Methods A counterbalanced, repeated-measures experiment, satisfaction surveys, and narrative content analysis was conducted remotely by Web platform. Participants were 37 American emergency physicians; they completed 16 clinical cases comparing 4 visual designs to the control formats from a commercially available EDIS. They then evaluated two additional chart overview representations without controls.

Results Visual designs provided benefit in several areas compared to controls. Task correctness (90% to 76%; $p = 0.003$) and completion time (median: 49–74 seconds; $p < 0.001$) were superior for a medication history timeline with class and schedule highlighting. Completion time (median: 45–60 seconds; $p = 0.03$) was superior for a past medical history design, using pertinent diagnosis codes in highlighting rules. Less mental effort was reported for visual allergy ($p = 0.04$), past medical history ($p < 0.001$), and medication timeline ($p < 0.001$) designs. Most of the participants agreed with statements of likeability, preference, and benefit for visual designs; nonetheless, contrary opinions were seen, and more complex designs were viewed less favorably.

Conclusion Physician performance with visual representations of clinical data can in some cases exceed standard formats, even in absence of training. Highlighting of priority clinical categories was rated easier-to-use on average than unhighlighted controls. Perceived complexity of timeline representations can limit desirability for a subset of users, despite potential benefit.

Keywords
- user–computer interface
- emergency service
- hospital
- computer graphics
- clinical decision-making

Background and Significance

The effect of cognitive load on interpretation of data is of elevated importance to patient safety in the emergency department (ED) setting.[1,2] Interruptions, time constraints, workflow variation, and information overload are cognitive threats that increase risk of errors.[3–6] With digitization, Emergency Department Information System (EDIS) have a larger role in ED error reduction, quality, and efficiency efforts.[1,7–10] Yet, high cognitive load, poor information presentation, lack of customizability, and efficiency problems are observed with existing record systems.[1,11–15]

Can visual features be part of the solution? In other domains, "big data" processing and graphics software advances help to create visual meaning from large data sets.[16–19] Likewise, visual approaches might also assist prioritization and sense-making by ED providers, who review complex records amid distractions, thus accelerating their clinical efforts.[20]

Most prior publications for visualization in EDIS address multiple-patient views, such as dashboards or ED tracking lists ("whiteboards").[21–25] On the other hand, evaluations of single-patient formats are infrequently seen. In 2014, Ozturk et al, described a visual medication list using a timeline format; however, the authors did not show comparative usability evaluation, in the sense of the standards used in human–computer interaction, including measures of effectiveness, efficiency, and satisfaction.[26,27]

More visual approaches are seen in other settings of care. Examples reported include laboratory data sparklines, metaphorical icons, integrated visual displays, and visual timelines for past medical records.[28–32] In 2018, a medication history timeline, evaluated by Belden et al, showed performance benefit among outpatient providers.[33] Yet, despite the research prototypes, in a 2011 review, Rind et al reported that most available electronic medical records (EMRs) do not offer advanced visualization features, and nearly all lack usability testing among emergency medicine specialist physicians.[34–36]

Visual Formats for Data Recognition and Prioritization

Risk recognition and short-term recall are of high value for an emergency provider's multitasking workflow.[3,37] However, sense-making of common text formats, such as lists or tables, requires reading list items sequentially and mentally managing priority and context using adjacent content.[11,38,39] As per Patel et al, clinical cognition can be viewed as a hierarchy —so the effort expended to gathering and prioritizing lower-level observations and excluding distractors could delay higher complexity findings and actions that affect patient management.[40,41]

Alternatively, "highlighting" is a visual design strategy utilizing preattentive styles of size, color, orientation, and shape (perceived in under 250 ms in early vision) to guide users to priority items in the visual field.[42–45] Huang et al describe use of visual recommendations as a means of "accelerating decision-making performance."[45] Additionally, use of object shape and pattern from expert domains ("objects of expertise") have been shown to improve recall capacity, which may reduce cognitive effort.[46–49]

Visual Formats for Time and Numerical Comparisons

Visual approaches may also improve interpretability of comparisons among numerical data and timestamps in health records, reducing cognitive load. Graphical depiction is likely cognitively beneficial (per cognitive science author David Kirsh) because external representations (graphs or drawings) can "coordinate," or "lower the cost of controlling thought" better than mental images (called "internal representations").[50–52] In a series of experiments in 1994, Zhang and Norman showed that external representations are more than just memory aids, but provide information that is usable without explicit mental formulation.[51] For example, nearby graph values are perceived as equivalent, without mentally reciting the underlying numbers.

Connecting Clinical Data and Visual Design

As visual representations of large data sets gain technical feasibility, the usefulness to emergency medicine remains uncertain. Making clinically important features in lengthy patient medical records more accessible during emergency decision-making could improve care. Cognition-based strategies for evaluating large data sets, described above, including, visual highlighting and content recommendation, visual objects for recognition and recall, and use of external representations such as timelines, should be better evaluated to identify potentially effective uses for emergency care. Usability evaluation by emergency physicians, including hypothesis tests comparing available visual concepts to more standard tabular formats, would provide a level of evidence not currently available in the literature.

Objectives

We aim to evaluate several visual representation formats for single ED patient records, either available from current literature, or supported by published cognition theories, and then more clearly describe possible performance benefits of use observed, such as improved emergency physician effectiveness or efficiency, that would justify added costs and EDIS development effort. We also aim to assess emergency physicians' subjective experiences completing clinical tasks using new designs and describe the acceptability and possible usability challenges of EDIS visual representations, varied in purpose, EDIS area, and complexity, if introduced among users in this practice setting.

Research Questions

How do graphical or visual approaches for representation of clinical data affect performance and satisfaction of emergency physicians, in clinical problem-solving tasks? What benefit or weaknesses of visual designs are seen among examples evaluated by participants?

Methods

Setting

Common practice situations for the medical specialty of emergency medicine were represented using simulated

clinical cases, conducted remotely among emergency specialist physicians in the United States, using an online learning-management system. The project was conducted from January to May 2018, and data collection occurred between March 4 and April 25, 2018.

Participants

Recruited participants were American Board of Emergency Medicine (ABEM)-certified physicians, currently active in clinical practice. Fifty-four ABEM-certified physicians were contacted with direct email as a purposive, heterogeneous sample of age, gender, and background with EDIS software.[53] From those, 37 completed the study.

Emergency medicine experience of participants ranged from 5 to 38 years, including residency (median: 10.5, mean: 13.9, and standard deviation: 9.5). Participants were primarily based or trained in the Midwest region of the United States, practicing at 15 clinical sites in 5 U.S. states, in 1 of 2 academic practices or 6 community hospital practices. Current work ranged from 32 to 180 clinical hours average per month, excluding administrative hours (median: 130, mean 123.4, standard deviation: 32.7). All participants had current use of EDIS in their EDs; 35 of 37 used EDIS for greater than 5 years, and over half reported using EDIS for 10 years or longer. Of systems in current use, 20 reported using Epic ASAP, 14 used Cerner FirstNet, and 3 reported using other systems. Nine participants were over 50 years old. Fourteen were women. Three reported color vision impairment, and two with difficulty reading small text.

Design

The study was a mixed methods, comparative usability evaluation, including a hypothesis test with a counterbalanced, repeated-measures design.[54] The methodology is deductive from theories of visual cognition, with a null hypothesis of no observed differences as physicians when used visual formats or control formats to complete study tasks.

The main hypothesis test consisted of 16 written "board exam"-style simulated cases, focused on comparing the use of visual formats to control formats in four EDIS areas: allergy list, past medical history (PMH), vital signs, and medication list (Meds). This was followed by satisfaction surveys. The test comprised four cases for each EDIS area, with two independently timed questions (tasks) per case. An example case evaluating tasks related to using a Meds area is provided in ►Table 1. The example utilizes ►Figs. 5 and 9 as visual and control images, respectively. The case order, and whether the participant received visual format or control format data, was determined by randomization group (►Fig. 1). Participants were randomized by exact order of response to simultaneous direct email, using sequential assignment. Physicians over age 50 were assigned separately, so stratifying the groups with respect to age.

Both the visual and control image for each case contained equivalent clinical information, but differences in formatting of colors, shape, position, temporal data, and clinical logic. Two case sequence groups accounted for learning effects—a simplified Latin square.[54]

Study Material

The study was developed on the e-learning platform Canvas (Instructure; www.canvaslms.com; Salt Lake City, Utah, United States). Clinical content areas were identified from among key topics in the *2016 Model of the Clinical Practice of Emergency Medicine*, by Counselman et al.[55] Data integration and decision-making tasks encountered during care of typical ED patients were incorporated, drawing on experience of nonparticipating physicians for details of clinical management.

Thirty-five images were used as case materials in the simulated cases. They include 16 experimental images for

Table 1 Case "H" for medication list (using ►Figs. 5 or 9)

Case prompt (Participant reads prompt before revealing tasks and timing)
A 16–year-old male patient with bipolar disorder and autism spectrum disorder is in ED with parents. He texted his sister 2–3 hours ago that he took a "handful" of his medications after his parents took his Xbox away. He's awake in the ED, but seems a little tired. With his frequent medication changes, there's difficulty confirming the med quantities missing. His mom reports throwing out all his old bipolar medications two weeks ago after he threatened to take pills. His EKG has a sinus rhythm of 75 showing normal QRS and QTc intervals.
Task 1 (Independently timed)
When was the patient's most recent prescription for quetiapine (Seroquel) written? 1. 1 week ago 2. Around 2 months ago 3. Around 6 months ago 4. Over 12 months ago
Task 2 (Independently timed)
In addition to ordering typical labs, which additional order is likely most useful? 1. Tegretol (carbamazepine) level 2. Lithium level 3. Serial EKGs q1 hour 4. Depakote (valproate) level

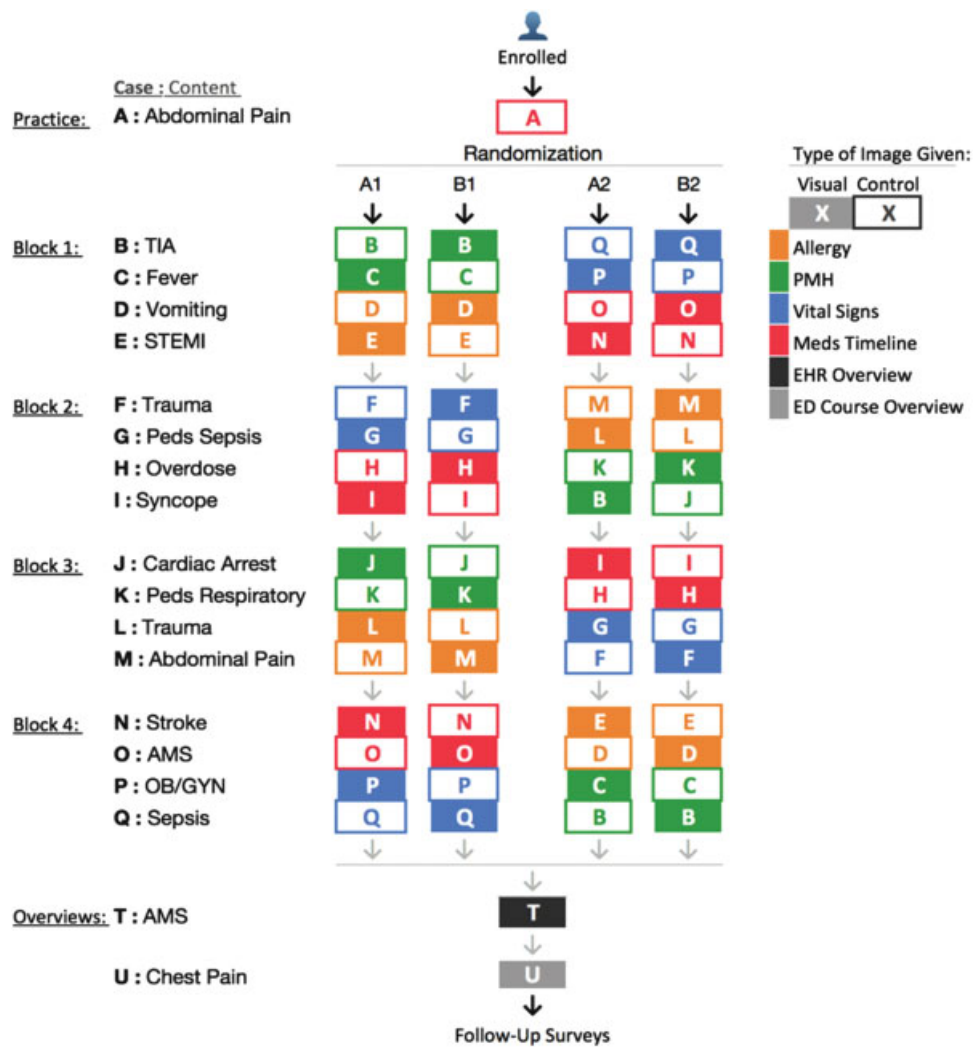Abbreviations: ED, emergency department; EKG, electrocardiogram.

Fig. 1 Diagram showing counterbalancing, randomization, and case content.

Fig. 2 Visual allergy severity and type categories highlighted by position and color.
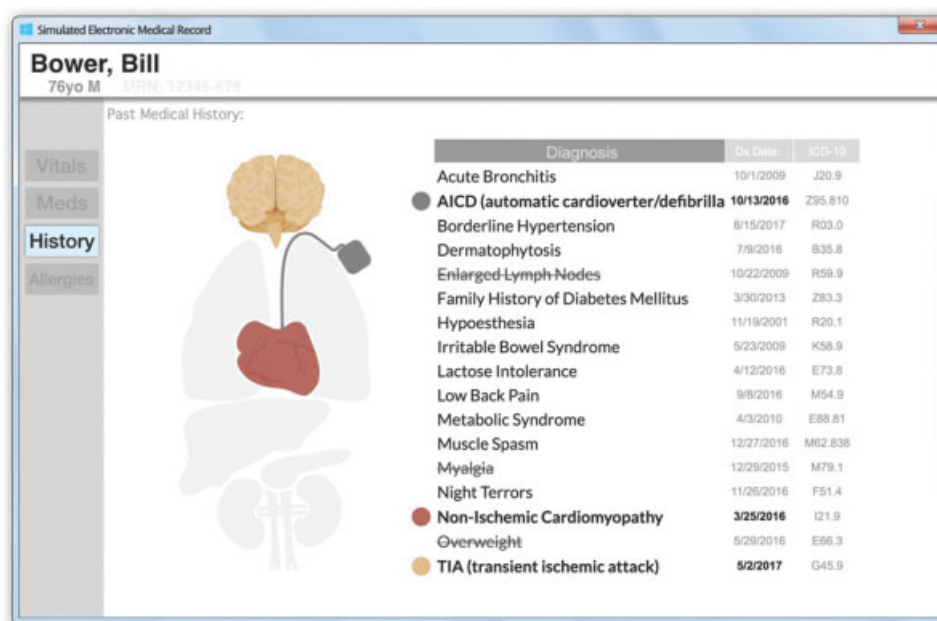
Fig. 3 Visual past medical history highlighted with anatomic objects and color according to emergency department pertinent International Classification of Diseases, Tenth Revision diagnoses.
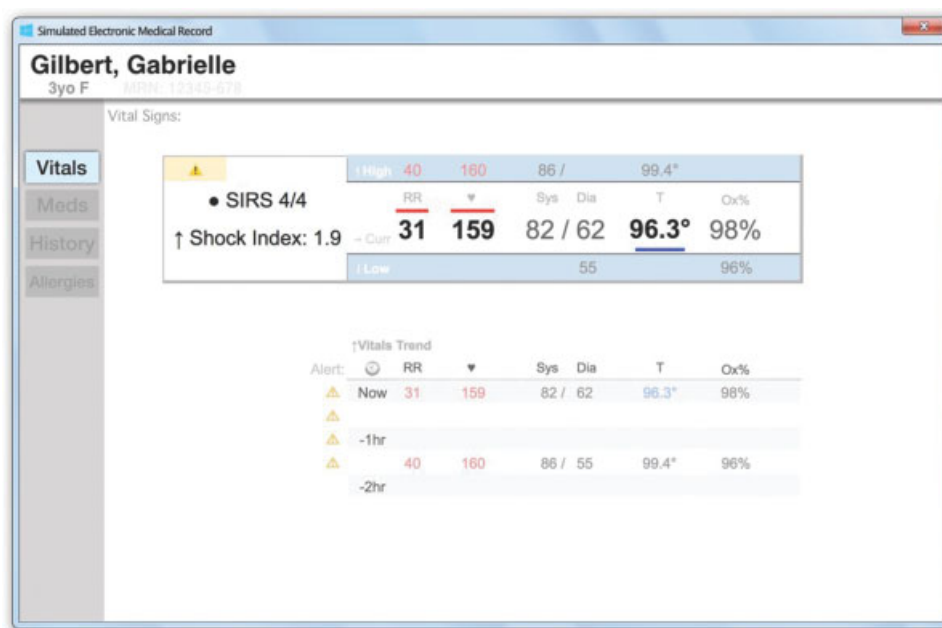

Fig. 4 Visual vital signs with high and low range colors and prominent critical illness calculations.

visual formats, in the 4 common EDIS areas (examples in ►Figs. 2–5), and were compared to 16 matched control images containing equivalent clinical information in the standard commercial format (examples in ►Figs. 6–9).

Images were developed using Google Sheets (Google LLC, Mountain View, California, United States), Adobe Illustrator CS6 (Adobe Systems, San Jose, California, United States), and Keynote 7.3.1(Apple Inc., Cupertino, California, United States). For design of the visual images, we adapted specific concepts from visualization or cognition publications. The visual allergy list (►Fig. 2) uses spatial and color highlighting

approaches for data visualization described in Huang et al, applied to high- and low-risk allergy categories.[45] PMH (►Fig. 3) incorporates highlighting of ED priority International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes with an external representation of anatomic objects of expertise, inspired by the "Studio TACK" entry in a 2013 Veterans Affairs Healthcare IT design challenge.[45,46,56] In addition to spatial and color highlighting of abnormal values, vital signs (►Fig. 4) calculates alerts for higher order findings (including systemic inflammatory response syndrome) from the individual vital

Fig. 5 Visual medication history (Meds) as a timeline, with highlighting for drug schedule and class.

Fig. 6 Control allergy list as a plain table.

signs, a concept based in the hierarchical decision-making process described by Patel et al.[40,45] Meds (►**Fig. 5**) incorporates an external timeline representation inspired by Ozturk et al in 2014, and others such as Alonso et al for representing temporal clinical histories.[26,30] Control images (►**Figs. 6–9**) were inspired by a production implementation of Cerner Millennium FirstNet used by many participants (Cerner Corporation, Kansas City, Missouri, United States).

Cases, questions, and case images were evaluated and revised by four nonparticipating physicians before the experiment. Afterward when surveyed, the participants unanimously reported agreement that scenarios in the study had real practice relevance.

Data Collection

Participants completed the study at their own pace during the 8-week data collection period, using personal macOS ($n = 19$) or Windows ($n = 18$) desktop or laptop computers. Chrome ($n = 18$), Safari ($n = 11$), and Internet Explorer ($n = 6$) were the most frequent browsers used. Cases and surveys required between 60 and 90 minutes for most participants to complete. Participants could take breaks between cases, and then come back for additional sections. Data collection incorporated quantitative comparative measures, and both quantitative and qualitative satisfaction and opinion as is typical of usability assessments.[57,58]
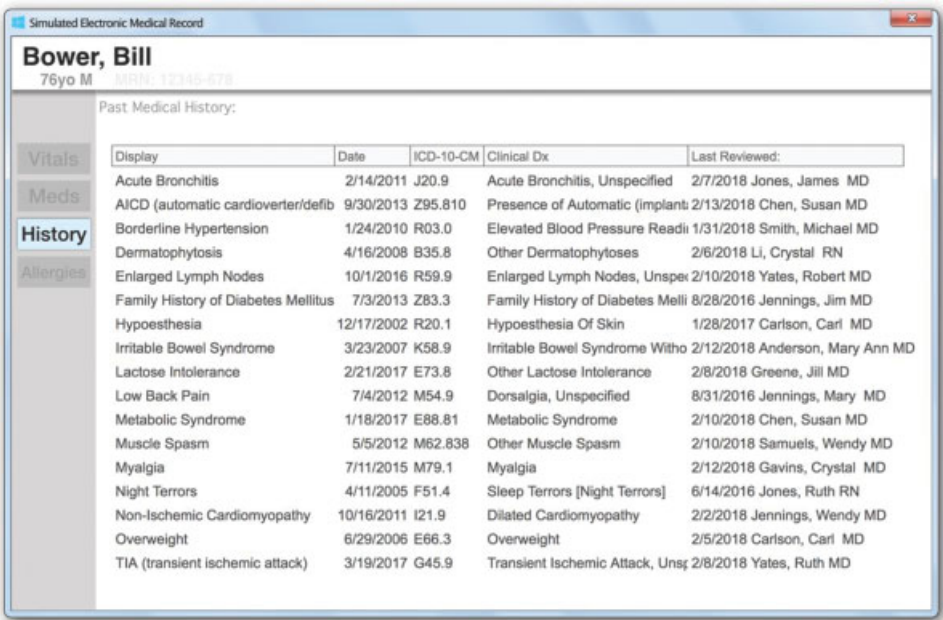
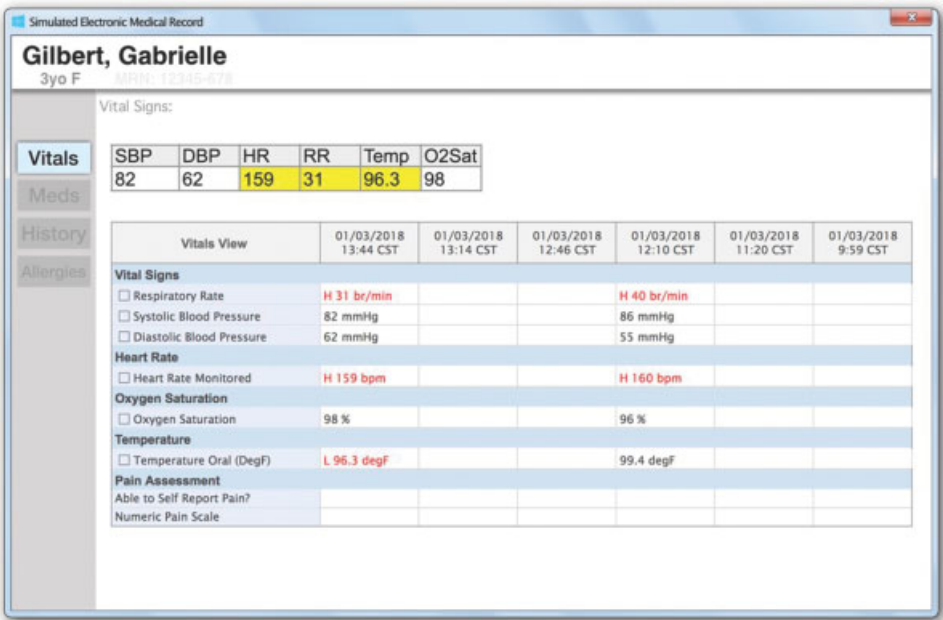**Fig. 7** Control past medical history as a plain table.

Fig. 8 Control vital signs in a plain table with single color highlighting of abnormal values.

To simulate an urgent pace of emergency practice, time for each case was limited, to maximum 4 minutes, with a single attempt. In the allotted time, the physician reviews a 2- to 3-sentence case prompt, marks completion; reviews an experimental or a control image; and then reads, and answers two different, and separately timed, multiple-choice questions, and then finally rates, "how hard was the case while using the image?," using a slider widget for numerical scale of difficulty (adapted from "Subjective Mental Effort Questionnaire," Sauro and Dumas), and embedded in the study platform.[59] Responses and elapsed times of responses were collected by the online platform for later analysis.

After the main hypothesis test of 16 cases, 2 additional cases were presented and evaluated formats for timeline overview of temporal data, inspired by timelines of Alonso et al[30] (►**Figs. 10** and **11**). An equivalent process in real life would be step-by-step information gathering, navigating between tables in different EMR modules, but navigation introduced a complexity level infeasible to recreate in our project controls.

Following the cases, the participants subjectively assessed each format in a Google Form survey (www.google.com/forms), using simple 5-point Likert scale questions adapted for image formats from System Usability Scale (SUS), as well

Fig. 9 Control medication list (Meds) as a plain table.
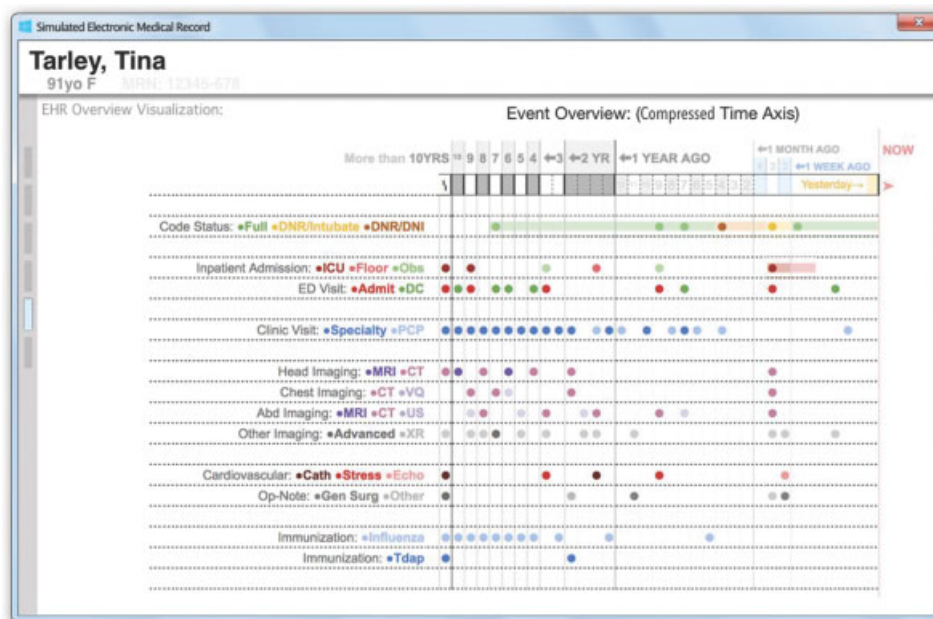
Fig. 10 Overview timeline of electronic health record (EHR) types, showing records pertinent to emergency department (ED) care with starfield plot.

as two further questions regarding perceived benefit and preference between visual and standard format images, and free-text response.[60]

At least one potential participant withdrew from the study after a recurrent delay with Safari auto-refresh of the study platform. Another withdrew after being blocked by a workplace firewall. Four participants completed the entire comparative test, but then omitted one or more of the following sections, including overview format cases or follow-up surveys.

Data Analysis
Data from collected responses, completion times, and after-surveys were aggregated and anonymized using an SQLite

relational database and Python 2.7 (PyCharm 2017.3.4), and utilizing the Canvas LMS REST API.

For correct task completion, results were pooled in 2×2 contingency tables (|Visual: Control| $\times$ |Correct: Incorrect|) and Fisher's exact tests are used for statistical significance.[61] Each of the four designs was assessed separately.

For task completion time, a survival analysis model was used—assessed for significance with the Cox proportional hazards model (*lifelines* Python package).[62–64] Event survival (event: correct completion) was selected over other approaches given tradeoffs in addressing in-process time for incorrect or omitted responders. Survival models retain individuals with incorrect and omitted responses in the "in-

Fig. 11 Overview timeline of emergency department (ED) course and results, showing event flow with tree graph.

process" population, as long as they have not yet answered or concluded the attempt.[64] This better reflects the in-process task delay experienced by all users, regardless of incorrect, or omitted responses.[65] This approach was demonstrated in a similarly designed study assessing intensive care unit displays by Koch et al and also in related methods used in general usability research.[31,65,66]

For subjective effort ("Subjective Mental Effort Questionnaire"), median and quartiles were calculated. Mann–Whitney U tests are used for significance, noting long-tailed response data with a lower bound.[59,61]

For the two final overview concepts and subjective opinion surveys, descriptive statistics are used, and distributions of the result data are provided, given infeasibility of additional hypothesis tests in these areas of the project. We did not report a calculated SUS score, due to added questions and question context that no longer referenced a complete system.

Narrative comments written by participants were assessed qualitatively. We used content analysis to elaborate and clarify possible explanations for main, deductive results (obtain "complementarity" per Johnson and Onwuegbuzie).[58,67] Ideas within participant comments were isolated as distinct meaning units, then separated into categories of favorable, unfavorable, or suggestion, and then coded, and themes derived.[67]

Results

Effectiveness: Correct Completion of Tasks
The correct responses for questions corresponded to standard clinical management, after all available image information was considered in each case. Use of the medication timeline visual (red) showed superior performance among
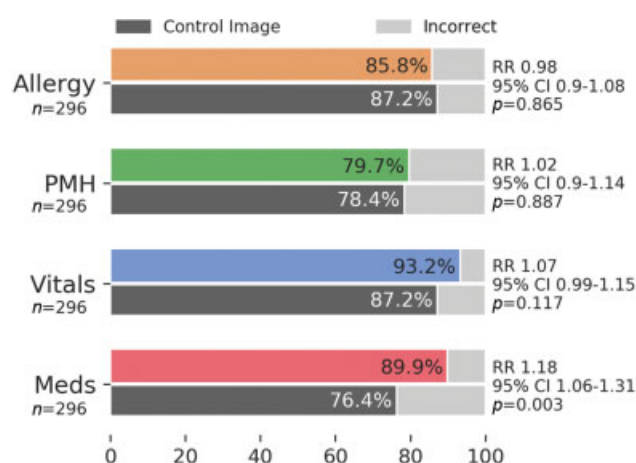
Fig. 12 Correct task completion using each visual format versus controls.

participants compared using the control Meds (corresponding black), and this difference was statistically significant ($p = 0.003$) (►**Fig. 12**). Allergy and PMH image comparison were equivalent (orange and green), and a trend visible for the vital signs visual (blue) was nonsignificant.

Efficiency: Time to Correct Completion
The rate that participants gave correct responses for two case questions were modeled using Kaplan–Meier survival curves (examples in ►**Fig. 13**). Trend lines show proportion of remaining unsolved questions (vertical axis), over time (horizontal axis). A more aggressively downsloping trend indicates faster task completion rates (i.e., increased hazard).[63] The medication timeline visual (red) and PMH visual (green) showed a faster completion rate compared to controls (matched gray), and the difference (in hazard ratio) is significant, with a Cox proportional hazards model.[63]
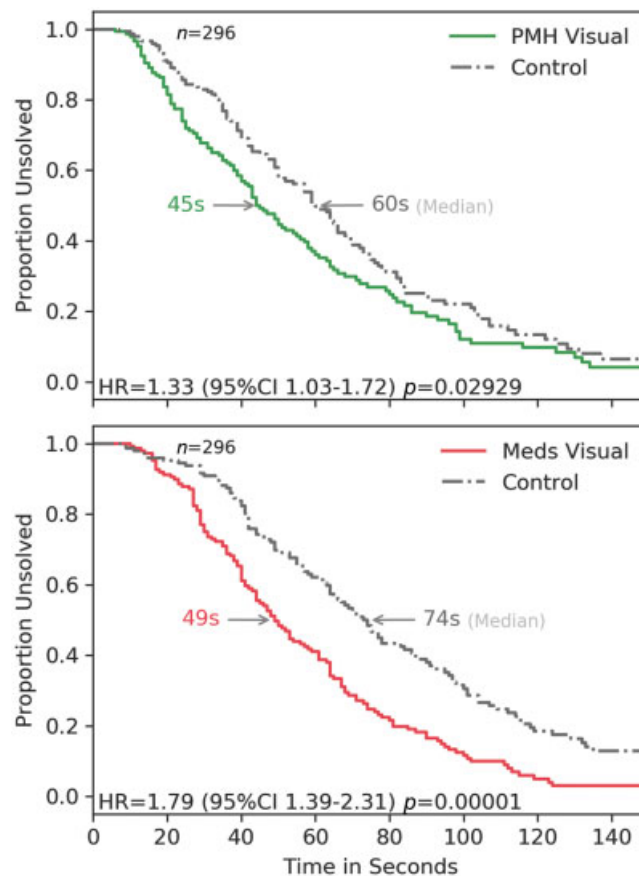
Fig. 13 Kaplan–Meier curves of time interval to each correct task completion for past medical history (PMH) and medication list (Meds) visuals.

Median time to completion was 25 seconds faster, and 15 seconds faster than controls, respectively. Other visual designs were nonsignificant with respect to time to completion.

Subjective Effort: Subjective Mental Effort Questionnaire

In each case, participants rated mental effort using a visual sliding scale.[59] Three of four visuals were associated with easier difficulty ratings than the corresponding standard controls; this was significant with Mann–Whitney U tests (►**Fig. 14**). For the allergy area, the median rating with the visual image was "Not very hard" compared to "A bit hard" using the control ($p = 0.037$). PMH ($p < 0.001$), and medication timeline ($p < 0.001$), both showed significant effects.

Results for the Two Additional Designs: Overview Formats

Two additional cases with the visual concepts for electronic health record (EHR) Overview and ED Course Overview were evaluated for purpose of broadening discussion regarding users' satisfaction using new visual concepts, particularly with more complex formats in the literature.[30] Correct completion of overview tasks did not significantly vary from the four experimentally tested designs. Task completion rate showed a slower rate and median time-of-comple-

tion for overview formats, but this was not statistically significant ($p = 0.101$). In the Subjective Mental Effort Questionnaire after each case, both overview designs, EHR Overview and ED Course Overview, were rated as more difficult than other visuals. Their median corresponded to "A bit hard," compared to "Not very hard" in other designs.

Satisfaction: Opinion Surveys

We observed favorable trends with statements of preference and benefit for all six designs used by the participants, compared to the standard formats (►**Table 2**). The primary designs, including allergy, PMH, vital signs, and medication history—performance compared to controls—had mostly favorable responses to the statement "would like to use frequently," statements of learnability, and statements of difficulty (►**Figs. 15**, **16** and ►**Table 2**). However, opinion was more mixed regarding the more visually complex overview designs, with both unfavorable and favorable opinions. The mean and distribution among Likert scale responses are presented visually, where each dot represents the rating of a single participant (►**Figs. 15** and **16**).

Satisfaction: Narrative Comment Content Analysis

Optional comments were provided for the "best or worst" aspects of the formats, or improvement suggestions. Themes and code areas from content analysis are presented in ►**Table 3**.
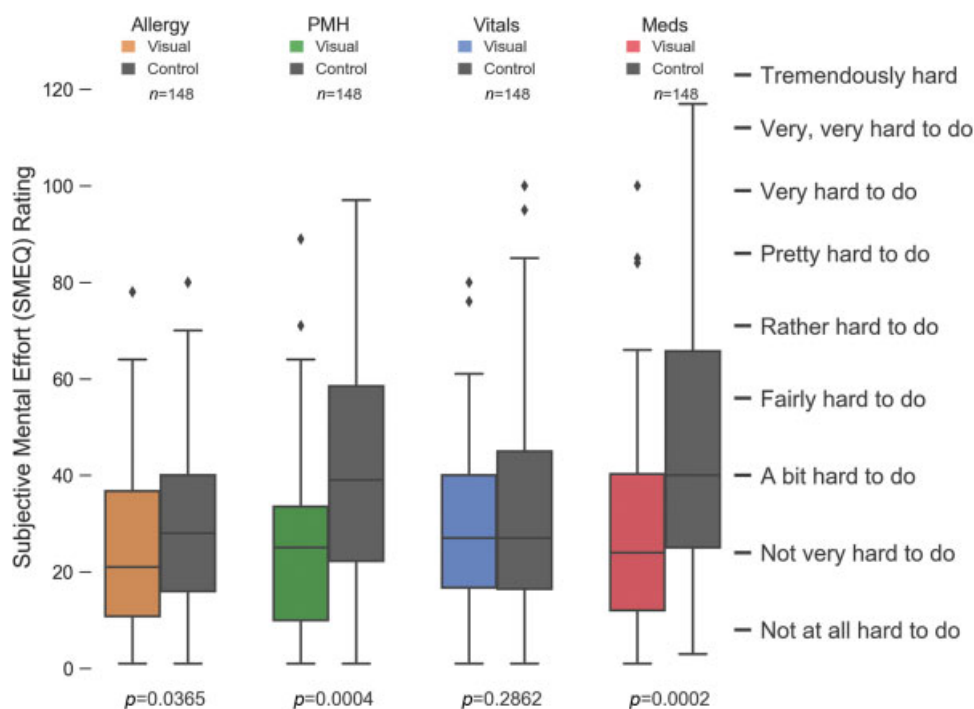
Fig. 14 Immediate mental effort reported following each case.

Table 2 Mean satisfaction survey responses (1–5 scale) by visual format

Allergy	PMH	Vital signs	Meds	EHR overview	ED course overview	Rate agreement with the statement:
Likeability and learnability (higher score is better): 1 = strongly disagree; 3 = neutral; 5 = strongly agree						
4.1	3.6	3.8	3.5	3.0	2.9	I would like to use the design frequently
4.1	4.0	4.3	4.2	3.8	4.0	There are benefits over regular approaches
3.8	3.4	3.6	3.5	3.6	3.5	Prefer compared to regular approaches
4.0	3.5	4.0	3.5	2.8	2.8	Easy to use
3.9	3.8	3.9	3.8	3.5	3.5	Well-integrated
4.0	3.5	4.0	3.7	2.9	2.8	Most would learn format quickly
3.9	3.7	4.1	3.7	3.2	3.3	Felt confident using design
Difficulty (lower score is better): 1 = strongly disagree 1; 3 = neutral; 5 = strongly agree						
2.4	2.8	2.3	2.6	3.3	3.4	Format unnecessarily complex
1.9	2.6	1.9	2.2	2.9	3.2	I would need additional assistance to use
2.3	2.0	2.2	2.1	2.2	2.6	Too much inconsistency
2.0	2.3	2.0	2.6	3.3	3.0	Very cumbersome to use
2.0	2.3	2.0	2.3	3.0	3.2	Needed to learn a lot before I could get started

Abbreviations: ED, emergency department; EHR, electronic health record; PMH, past medical history.

Discussion

In a high-pace and high-consequence clinical environment, incremental performance gains using EDIS may mean a cumulative benefit to the care process. Our evaluation illustrates several formats that visually represent EDIS data, and shows how the designs may affect effectiveness, efficiency, and user satisfaction of board-certified emergency physicians completing clinical decision-making tasks.

While results varied among designs, the advantages observed in performance support further attention to visual and cognitive approaches as external representations applied in EDIS interfaces, as has been described generally in other content areas.[50,51,68] Our medication timeline (►**Fig. 5**, inspired by Ozturk et al) had not previously been quantitatively evaluated among emergency physicians, yet showed 13% greater correct completion and 25 seconds
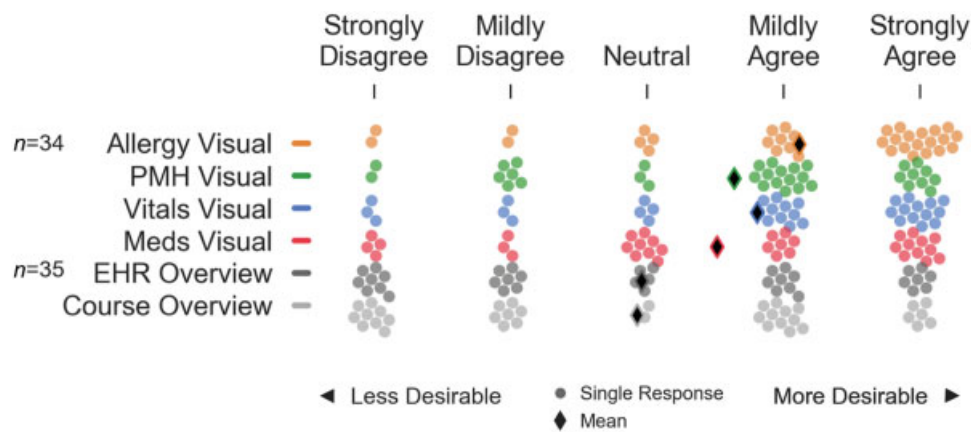
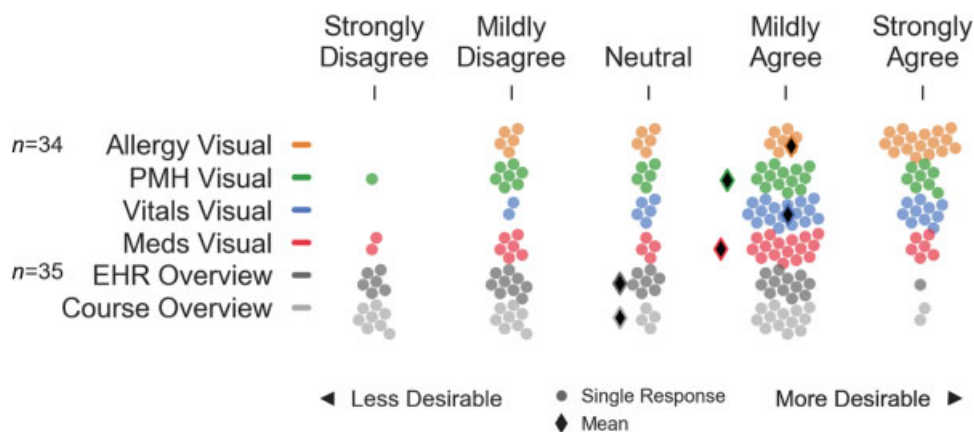**Fig. 15** Distribution and mean of likeability for all visual designs.

Fig. 16 Distribution and mean of ease-of-use for all visual designs.

faster median completion time for medical history tasks compared to tabular formats among participants.[26,33] Our highlighting and graphical anatomic representation of ICD-10-CM diagnosis codes most pertinent to emergency care in PMH data (►**Fig. 3** and ►**Table 4**) also showed a 15-second faster median completion time.

The above graphical representations, known in cognition science literature as "external representations," are described more generally to affect the "cost structure" of tasks, freeing cognitive resources, and potentially improving cognitive performance.[50,52] Indeed, our participant comments report better clarity using timelines to distinguish old, or discontinued Meds from current medications; and note better recognition of pertinent past history items from among organ systems (►**Table 3**). Graphical external representations did not feature as strongly in allergy list or vital signs formats (►**Figs. 2** and **4**), and these formats did not have significant differences in performance compared to the tabular controls.

Nevertheless, a lower level of effort was subjectively reported in all the designs when visual highlighting with preattentive features were used, when in comparison to a control format that did not use highlighting (control images for vital signs did highlight abnormal values, ►**Fig. 8**). Improved perception of effort may result from a known effect of high-

lighting to improve the ease of "information foraging," a behavior by users searching for trails of available visual clues to find valued information.[52,69] The highlighting used in our visual designs was based on generic, reproducible rules of clinical priority—for example, highlighting the most severe drug allergies, most ED-relevant ICD-10 codes (►**Table 4**), and most current or complaint-relevant drug classes. Visual recommendation influences enhanced discovery—in our examples, an emphasis for the most generally pertinent information to experienced emergency providers.[45,68]

Notably, narrative content analysis identified a concern that highlighting could mean overlooking or distracting from something important (that is not highlighted) (►**Table 3**). Such premature anchoring is a cognitive bias discussed in clinical cognition; and the impact of highlighting is to help jump to conclusions about which EDIS information is most relevant.[70,71] However, highlighting is still a tradeoff between improved visibility for *priority* EDIS data, or for *nonpriority* data—begging the question: which is more problematic? For many participants this critique was reversed.

While majorities of emergency physician users agreed with likeability- and learnability-related statements, while disagreed with difficulty-related statements for allergy, PMH, vital sign, and medication history designs, the opinions for the more complex, overview visuals (seen without

Table 3 Summary of content analysis of narrative comments by participants

Theme for design area	Most frequent code (positive vs. negative)	Example comment	No.
Visual allergy highlighting, de-emphasis, and design bring tradeoffs for recognition	Categorization by allergy class	"Separating out into classes, visually more accessible"	7
	Missed information	"Gray colors essentially make me ignore those items"	3
Visual PMH highlighting, expectations, and context affect sense-making	Benefit to visual highlighting	"Good to highlight organ systems that might get glossed over in a standard chart review"	8
	Doubt real world benefit	"I wonder if it could distract in real life"	4
Visual vital signs formatting assists interpretation but brings a new challenge	Visible critical illness calculations	"Important information (i.e., SIRS info) was at top right–the first place my eye went"	7
	Inconsistency from expectations	"Flow of data is not consistent with the way we usually see vitals and may lead to confusion"	3
Meds timeline helps comparing information but brings a new challenge	Timeline visual representation	"Clearly showed which meds were current and which had been discontinued long ago"	5
	Difficult to learn	"At first difficult but did adapt to it"	3
EHR Overview complexity affects ease of use, offsetting efficiency of information gathering	Ease of finding information	"Can help simplify our daily work by reducing clicks and searching through filters by integrating important info on one graph/slide/page"	4
	Difficult to learn	"A lot of the initial difficulty with it is that it's so foreign, but for good reasons"	7
ED Course complexity affects ease of use, offsetting beneficial operations comparisons	Useful in quality review and handoffs	"The ED timeline with interventions is a great mapping tool for quality improvement"	6
	Visual complexity	"Interesting way to visualize the ED course but might be too complex to use in real time"	8

Abbreviations: ED, emergency department; EHR, electronic health record; PMH, past medical history; SIRS, systemic inflammatory response syndrome.

control) were mixed. Majorities still agreed to benefit and preference, but complexity of some external visual representation approaches (such as timelines, star-field plots, and tree graphs) remained an important limitation to more widespread acceptability, making the formats comparatively more difficult to learn.[72–74] Our Meds timeline, which showed the most significant overall performance benefit, was also affected by the lower perceived learnability and higher difficulty compared to simpler formats. In comments, many remarked that more complex designs would be additionally useful after a longer learning period, or that the ideas were a helpful new paradigm, but would take getting used to (►**Table 3**). Variation in observed performance benefits, and a range of subjective opinions regarding the six explored format concepts continue to support a user-centered, and iterative process, and one allowing features such as highlighting to be configured based on preference.

Limitations

The most prominent issues that reduce the validity and generalizability of our findings are a small sample size, possibility of sampling bias, and other methodological challenges encountered conducting physician usability testing, and with conducting remote data collection. Comparatively, prior literature of EMR visualization that we reviewed has

reported evaluations of designs only among graduate students with no clinical training, evaluations in absence of any emergency physicians, or in absence of objective, comparative measures of performance.[26,34,35]

An ideal study size in usability research is a debated area, with diminishing practical benefits of problem discovery or comparative discrimination seen after achieving 10, 20, or 30 or more study subjects, yet adding further study cost.[75] The reported typical practice of 10 to 25 subjects in usability studies varies from ideal sample sizes calculated to achieve power and statistical measures in other (or solely) quantitative research, more often in hundreds of individuals.[75] In our planning, we calculated 140 to 150 as a group size for ideal statistical power, when using a clinically meaningful difference in task completion of 1 of 8 available tasks, and a 10-second difference in completion time. We measured that many observations using four unique iterations of each visual concept, 35 individuals, and multiple tasks, acknowledging the limitation that any dependence of the repeat measures would reduce calculated sample power, increasing the likelihood of type 2 error, but that obtaining the participation of over 30 individuals did exceed typical group sizes in usability test procedures and met our feasibility constraints.[75]

A high level of standardization of emergency medicine training and practice in the United States suggests a basis for

Table 4 Important ICD-10 codes selected for PMH highlighting

System	Problem	ICD-10-CM
Cardiac	Coronary diseases	I20-25
	Heart failure	I50,42,43
	Pacemaker/Bypass	Z95
Pulmonary	Restrictive	J60-70, J80-84
	Obstructive/Asthma	J41-47
Renal	End-stage disease	N18, Z49,Z99
Hepatic	Liver failure	I85,K72,K74, K70.3-4,K71.7
Neurologic	Cognitive Impairment	G30-31, G20-21
	Primary neurologic	G80-83, G35, G10-14
	Stroke/Ischemia	I63-69, G45.8-9
	Cerebral hemorrhage	I60-62
	Seizure	G40
Psychologic	Psychosis	F20-29
Transplant	Major organ	Z94.1-4,.83,.81
Immune	Deficiencies	D80-84
Neoplasm	Cancer	C00-96
Vascular	Thrombosis, sclerosis	I26, I71-72
Endocrine	Diabetes	E08-13
	Insulin pump	Z96.41
DNR		Z66
Cardiac arrest		I46

Abbreviations: DNR, do not resuscitate; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification; PMH, past medical history.

applicability to other U.S. emergency physicians. However, despite purposive methods for demographic and EDIS groups, known limitations of nonrandom sampling, including sampling bias, require caution.

Using simulated, online cases improved study feasibility by reducing initial barriers, such as access to the physicians, access to patient data, obtaining hospital system consent for clinical disruptions, and cooperation of EMR vendors. The approach enabled a counterbalanced design, intended to offset effects of learning, and differences between subjects. However, this reductionist approach creates less generalizability and validity to real-life systems. For example, a single control format for each visual image simplified production of our study design and materials, yet makes it more difficult to transfer the conclusions to other current and future EDIS interface formats in the market. We also did not introduce intentional confounders, such as duplicated data or workflow interruptions, which can create usability problems in real EDIS use.[3]

The remote, Web-based format also did limit full control of the testing environment, meaning at least two participants left the study due to technical issues on their systems, and four participants did not complete one or more follow-up survey. A counterbalanced design does address this attrition, but a negative confounder in follow-up survey results could occur due to technical frustrations with the platform, unrelated to the visual designs themselves.

Related U.S. practice settings with a high volume of new, general complaints—such as urgent care or hospitalist medicine, or among advanced practice providers—may have overlapping scenarios of use. However, varied expectations of brevity, visual complexity, and efficiency based on training and experience suggest potential for important differences.

Future Research

Many providers expressed interest in new EDIS visualization tools. While we were limited to evaluating six designs, other instances of visually focused formats are possible, with similar theoretical premises, including refinement and iteration of the ideas demonstrated, or incorporation and evaluation in production systems.

Suggestions from participant comments include dimensionality reduction to improve visual simplicity. Also, requirements for physicians with limitations of color perception or spatial resolution may also be better described targeting more affected physicians with other color combinations and sizes. Lastly, visual formats could also be used with more sophisticated statistical analysis or data mining of patient records.

Conclusion

Visual design factors affect emergency physicians in clinical decision-making scenarios; not only in terms of user satisfaction, but in some cases, by achieving relevant goals with better performance. Our experiment suggests timelines and highlighting can offer more effective and more efficient interfaces for reviewing medication histories, compared to standard tables. More generally, highlighting priority information using clinical logic rules can increase problem-solving speed and decrease mental effort. In a workplace where cognitive overload can be a threat to patient safety, the advantages are not trivial.

Greater complexity of some timeline formats was dissatisfying among many physicians, despite performance benefits seen (and in opinions of other participants). This is likely due to perception (by some) that the effort of new ways of visual abstraction offsets its advantages. To account for tradeoffs for different users, formats showing benefit (such as the medication timeline) could be user-centered and optional—available for physicians who see benefit, while allowing others to select plain representations.

Clinical Relevance Statement

This research describes the acceptability and usability of visual designs for Emergency Department Information System, using clinical, logic-based highlighting and external

representations. The findings may provide rationale for innovation of commercial electronic health record system user interface design.

Multiple Choice Questions

1. Which of the following experimental designs showed the greatest difference in correct responses by physicians, compared to control formats?
   a. Allergy category highlighting.
   b. Past medical history by systems with pertinent diagnoses.
   c. Medical history as a timeline with drug classes.
   d. Vital signs with severity calculations visible.

   Correct Answer: The correct answer is option c.

2. Which of the following was the most common concern reported by emergency physicians after using timeline overview formats to view pertinent case data?
   a. Direction of the timeline.
   b. Learning a complex format.
   c. Choice of colors.
   d. Wrong data types available.

   Correct Answer: The correct answer is option b.

Protection of Human and Animal Subjects
The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. A formal ethics declaration was evaluated in institutional review and approved prior to the project. No real patient data was used. All patient cases are fictional. All physician participants completed a standard written informed consent. Benefits to participants exceeded risks of harm; they are not a vulnerable group and freely chose to participate.

Conflict of Interest
None declared.

References

1 Farley HL, Baumlin KM, Hamedani AG, et al. Quality and safety implications of Emergency Department Information Systems. Ann Emerg Med 2013;62(04):399–407

2 Laxmisan A, Hakimzada F, Sayan OR, Green RA, Zhang J, Patel VL. The multitasking clinician: decision-making and cognitive demand during and after team handoffs in emergency care. Int J Med Inform 2007;76(11-12):801–811

3 Chisholm CD, Collison EK, Nelson DR, Cordell WH. Emergency department workplace interruptions: are emergency physicians "interrupt-driven" and "multitasking"? Acad Emerg Med 2000;7(11):1239–1243

4 Morrison JB, Rudolph JW. Learning from accident and error: avoiding the hazards of workload, stress, and routine interruptions in the emergency department. Acad Emerg Med 2011;18(12):1246–1254

5 Källberg A-S, Göransson KE, Florin J, Östergren J, Brixey JJ, Ehrenberg A. Contributing factors to errors in Swedish emergency departments. Int Emerg Nurs 2015;23(02):156–161

6 Franklin A, Liu Y, Li Z, et al. Opportunistic decision making and complexity in emergency care. J Biomed Inform 2011;44(03):469–476

7 Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004;11(02):104–112

8 Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006;144(10):742–752

9 Hersh WR. Medical informatics: improving health care through information. JAMA 2002;288(16):1955–1958

10 Handel DA, Wears RL, Nathanson LA, Pines JM. Using information technology to improve the quality and safety of emergency care. Acad Emerg Med 2011;18(06):e45–e51

11 Zahabi M, Kaber DB, Swangnetr M. Usability and safety in electronic medical records interface design: a review of recent literature and guideline formulation. Hum Factors 2015;57(05):805–834

12 Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. J Am Med Inform Assoc 2017;24(01):218–226

13 Kim MO, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. J Am Med Inform Assoc 2017;24(02):246–250

14 Caudill-Slosberg M, Weeks WB. Case study: identifying potential problems at the human/technical interface in complex clinical systems. Am J Med Qual 2005;20(06):353–357

15 Mcdonnell C, Werner K, Wendel L. Electronic Health Record Usability: Vendor Practices and Perspectives. Rockville, MD: Agency for Healthcare Research and Quality; 2010

16 Jung H, Kim T, Yang Y, et al. Aesthetics in data visualization: case studies and design issues. In: Huang M, Huang W, eds. Innovative Approaches of Data Visualization and Visual Analytics. Hershey, PA: IGI Global: Information Service Reference; 2014:1–24

17 Ola O, Sedig K. Beyond simple charts: design of visualizations for big health data. Online J Public Health Inform 2016;8(03):e195

18 Meyer J. Performance with tables and graphs: effects of training and a Visual Search Model. Ergonomics 2000;43(11):1840–1865

19 Zoss A. Cognitive processes and traits related to graphic comprehension. In: Huang M, Huang W, eds. Innovative Approaches of Data Visualization and Visual Analytics. Hershey, PA: IGI Global: Information Service Reference; 2014:94–105

20 Stead WW, Miller RA, Musen MA, Hersh WR. Integration and beyond: linking information from disparate sources and into workflow. J Am Med Inform Assoc 2000;7(02):135–145

21 Hertzum M, Simonsen J. Effects of electronic emergency-department whiteboards on clinicians' time distribution and mental workload. Health Informatics J 2016;22(01):3–20

22 Dexheimer JW, Kennebeck S. Modifications and integration of the electronic tracking board in a pediatric emergency department. Pediatr Emerg Care 2013;29(07):852–857

23 Sopan A, Plaisant C, Powsner S, Shneiderman B. Reducing wrong patient selection errors: exploring the design space of user interface techniques. In: AMIA Annual Symposium Proceedings. American Medical Informatics Association; 2014:1056–1065. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25954415. Accessed October 31, 2017

24 Franklin A, Gantela S, Shifarraw S, et al. Dashboard visualizations: supporting real-time throughput decision-making. J Biomed Inform 2017;71:211–221

25 Rasmussen R. Electronic whiteboards in emergency medicine: a systematic review. In: Luo G, Liu J, eds. Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium Publication. Vol. 1. New York, NY: Association for Computing

Machinery; 2012:483–492. Available at: https://rucforsk.ruc.dk/ws/files/37541649/Rasmussen_2012.pdf. Accessed May 18, 2018

26 Ozturk S, Kayaalp M, McDonald CJ. Visualization of patient prescription history data in emergency care. In: AMIA Annual Symposium Proceedings. American Medical Informatics Association 2014:963–968. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25954404. Accessed October 31, 2017

27 International Organization for Standardization. ISO 9241–11: Ergonomics of Human-System Interaction — Part 11: Usability: Definitions and Concepts; 2018. Available at: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en. Accessed January 20, 2019

28 Bauer DT, Guerlain S, Brown PJ. The design and evaluation of a graphical display for laboratory data. J Am Med Inform Assoc 2010;17(04):416–424

29 Tasa UB, Ozcan O, Yantac AE, Unluer A. A case study on better iconographic design in electronic medical records' user interface. Inform Health Soc Care 2008;33(02):125–138

30 Alonso DL, Rose A, Plaisant C, Norman KL. Viewing personal history records: a comparison of tabular format and graphical presentation using lifelines. Behav Inf Technol 1998;17(05): 249–262

31 Koch SH, Staggers N, Weir C, Agutter J, Liu D, Westenskow DR. Integrated information displays for ICU nurses: field observations, display design, and display evaluation. Proc Hum Factors Ergon Soc Annu Meet 2010;54(12):932–936

32 Effken JA, Loeb RG, Kang Y, Lin ZC. Clinical information displays to improve ICU outcomes. Int J Med Inform 2008;77(11):765–777

33 Belden JL, Wegier P, Patel J, et al. Designing a medication timeline for patients and physicians. J Am Med Inform Assoc 2019;26(02): 95–105

34 Rind A, Wang TD, Aigner W, et al. Interactive information visualization to explore and query electronic health records. Found Trends Human–Computer Interact 2013;5(03):207–298

35 Wongsuphasawat K, John Alexis Guerra G, Catherine P, Taowei David W, Ben S, Taieb-Maimon M. LifeFlow: visualizing an overview of event sequences. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Vancouver, BC, Canada 2011:1747–1756

36 Cousins SB, Kahn MG. The visual display of temporal information. Artif Intell Med 1991;3(06):341–357

37 Berg LM, Källberg A-S, Göransson KE, Östergren J, Florin J, Ehrenberg A. Interruptions in emergency department work: an observational and interview study. BMJ Qual Saf 2013;22(08):656–663

38 Berlin A, Sorani M, Sim I. A taxonomic description of computer-based clinical decision support systems. J Biomed Inform 2006;39 (06):656–667

39 Starren J, Johnson SB. An object-oriented taxonomy of medical data presentations. J Am Med Inform Assoc 2000;7(01):1–20

40 Patel VL, Arocha JF, Kaufman DR. A primer on aspects of cognition for medical informatics. J Am Med Inform Assoc 2001;8(04): 324–343

41 Treisman A, Gormican S, Warren HC, et al. Feature analysis in early vision: evidence from search asymmetries. Psychol Rev 1988;95 (01):15–48

42 Wolfe JM. Guided Search 2.0: a revised model of visual search. Psychon Bull Rev 1994;1(02):202–238

43 Hesse PN, Schmitt C, Klingenhoefer S, Bremmer F. Preattentive processing of numerical visual information. Front Hum Neurosci 2017;11:70

44 Grabowecky M, Robertson LC, Treisman A. Preattentive processes guide visual search: evidence from patients with unilateral visual neglect. J Cogn Neurosci 1993;5(03):288–302

45 Huang ML, Liang J, Huang W. Highlighting in visual data analytics. In: Huang M, Huang W, eds. Innovative Approaches of Data Visualization and Visual Analytics. Hershey, PA: IGI Global: Information Service Reference; 2014:178–190

46 Curby KM, Glazek K, Gauthier I. A visual short-term memory advantage for objects of expertise. J Exp Psychol Hum Percept Perform 2009;35(01):94–107

47 Ericsson KA, Kintsch W. Long-term working memory. Psychol Rev 1995;102(02):211–245

48 Few S. Data presentation: tapping the power of visual perception. Intell Enterp 2004;13(07):33

49 Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. Nature 1997;390(6657):279–281

50 Kirsh D. Interaction, external representation and sense making. AI Soc 2010;25:441–454

51 Zhang J, Norman DA. Representations in distributed cognitive tasks. Cogn Sci 1994;18(01):87–122

52 Russell DM, Stefik MJ, Pirolli P, Card SK. The cost structure of sensemaking. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 1993:269–276

53 Etikan I, Abubakar Musa S, Sunusi Alkassim R. Comparison of convenience sampling and purposive sampling. Am J Theor Appl Stat 2016;5(01):1–4

54 Pollatsek A, Well AD. On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. J Exp Psychol Learn Mem Cogn 1995;21(03):785–794

55 Counselman FL, Babu K, Edens MA, et al; 2016 EM Model Review Task Force; American Board of Emergency Medicine. The 2016 model of the clinical practice of emergency medicine. J Emerg Med 2017;52(06):846–849

56 Caldwell R, Salem L, Smith WB. Health Design Challenge: Studio TACK. Health Design Challenge. 2013. Available at: http://health-designchallenge.com/showcase/studiotack/studiotack.pdf. Accessed March 17, 2018

57 Simes-Marques M. L. I. Usability of interfaces. In: Nunes IL, ed. Ergonomics - A Systems Approach. Rijeka, Croatia: InTech; 2012: 155–170

58 Johnson R, Onwuegbuzie A. Mixed methods research: a research paradigm whose time has come. Educ Res 2004;33(07):14–26

59 Sauro J, Dumas JS. Comparison of three one-question, post-task usability questionnaires. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI '09. Boston, MA; 2009:1599–1608

60 U.S. Department of Health & Human Services Washington DC. System Usability Scale [SUS]. Usability.gov - Improving the User Experience. Published September 6, 2013. Available at: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html. Accessed January 17, 2018

61 J de Smith M. Statistical Analysis Handbook. 2018. Available at: http://www.statsref.com/HTML/index.html. Accessed April 21, 2017

62 Altman DG, Bland JM. Time to event (survival) data. BMJ 1998;317 (7156):468–469

63 Abadi A, Saadat S, Yavari P, Bajdik C, Jalili P. Comparison of Aalen's additive and Cox proportional hazards models for breast cancer survival: analysis of population- based data from British Columbia, Canada. Asian Pac J Cancer Prev 2011;12(11):3113–3116

64 Machin D, Cheung YB, Parmar MKB. Survival Analysis: A Practical Approach. 2nd ed. West Sussex, England: John Wiley & Sons, Ltd; 2006

65 Rummel B. Beyond average: Weibull analysis of task completion times. J Usability Stud 2017;12(02):56–72

66 Liu C, White RW, Dumais S. Understanding web browsing behaviors through Weibull analysis of dwell time. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '10; 2010:379–387

67 Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004;24(02):105–112

68 Fisher DL, Coury BG, Tengs TO, Duffy SA. Minimizing the time to search visual displays: the role of highlighting. Hum Factors 1989; 31(02):167–182

69 Pirolli P, Card S. Information foraging. Psychol Rev 1999;106(04): 643–675

70 Croskerry P. From mindless to mindful practice–cognitive bias and clinical decision making. N Engl J Med 2013;368(26): 2445–2448

71 Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med 2003;78(08): 775–780

72 Aigner W, Miksch S, Müller W, Schumann H, Tominski C. Visualizing time-oriented data—a systematic view. Comput Graph 2007; 31(03):401–409

73 Sanchez A. Understanding collections and their implicit structures through information visualization. In: Huang ML, Huang W, eds. Innovative Approaches of Data Visualization and Visual Analytics. Hershey, PA: IGI Global: Information Service Reference; 2014:151–171

74 Zhang S, Wu K. Feature-based uncertainty visualization. In: Huang M, Huang W, eds. Innovative Approaches of Data Visualization and Visual Analytics. Hershey, PA: IGI Global: Information Service Reference; 2014:68–93

75 Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. J Usability Stud 2009;5 (01):34–45