ORIGINAL RESEARCH

Dynamic Electronic Health Record Note Prototype: Seeing More by Showing Less

Jeffery L. Belden, MD, Richelle J. Koopman, MD, MS, Sonal J. Patil, MD, MSPH, Nathan J. Lowrance, MS, Gregory F. Petroski, PhD, and Jamie B. Smith, MA

Introduction: Cluttered documentation may contribute adversely to physician readers' cognitive load, inadvertently obscuring high-value information with less valuable information. We test the hypothesis that a novel, collapsible assessment, plan, subjective, objective (APSO) note design would be faster, more accurate, and more satisfying to use than a conventional electronic health record (EHR) subjective, objective, assessment, plan (SOAP) note for finding information needed for ambulatory chronic disease care.

Methods: We iteratively developed physician clinic note prototypes with features designed to emphasize more important information and de-emphasize less clinically relevant information. Sixteen primary care physicians reviewed comparable clinic notes with the 4 note styles presented in random order to find key information in the notes during timed tasks. The 4 note styles were denoted A (traditional SOAP note), B (2-column APSO note), C (collapsible APSO note), and D (2-column collapsible APSO note). The 4 unique note styles were designed to have equal amounts of information in each section. We simulated their utility for clinical practice by imposing time limits and by interrupting 1 of the tasks with a typical clinical interruption. For each session, we recorded audio, computer-screen activity, eye tracking, and made field notes. We obtained usability ratings (System Usability Scale), new feature preference ratings, and performed semistructured post-task interviews with subsequent content analysis. We compared the effectiveness of the 4 note styles by measuring time on task, task success (accuracy), and effort as measured by NASA Task Load Index.

Results: Note styles C and D were significantly faster than A and B for the Review of Systems and Physical Examination tasks, as we expected. Notes B and C had the best success (finding requested data) scores. Users strongly endorsed all the new note features incorporated into the new note prototypes. Previously expressed concerns about temporarily hiding parts of the note (using the accordion display design pattern) were allayed. Usability ratings for note A were worst but comparably better for note styles B, C, and D.

Discussion: The new APSO note prototypes performed better than the traditional SOAP note format for speed, task success (accuracy), and usability for physician users acquiring information needed for a typical chronic disease visit in primary care. Moving Assessment and Plan to the top is 1 easily accomplished feature change. Innovative documentation displays of EHR data can safely improve information display without eliminating data from the record of the visit. (J Am Board Fam Med 2017;30:691–700.)

Keywords: Data Display, Documentation, Electronic Health Records, Information Systems, Medical Informatics

Physician clinic visit notes can be long and challenging to read. One reason for this is that currently these notes are designed to meet more than just the needs of the clinical care team. Additional stakeholders in the design of office notes are audi-

tors, attorneys, billing staff, and insurance companies. Their tasks and information needs are often

This article was externally peer reviewed. Submitted 6 February 2017; revised 25 May 2017; accepted 28 May 2017.

From the Department of Family and Community Medicine (JLB, RJK, SJP, JBS), the School of Information Science and Learning Technologies, College of Education

(NJL), the Department of Health Management & Informatics (GFP), and the Informatics Institute (JLB), University of Missouri–Columbia, Columbia, MO.

Funding: This work was supported by a faculty development grant from the Department of Family & Community Medicine, University of Missouri.

Conflict of interest: none declared.

Corresponding author: Jeffery L. Belden, MD, Department of Family & Community Medicine, University of Missouri School of Medicine, M239 Medical Science, Columbia, MO 65212 (E-mail: beldenj@health.missouri.edu).

tangential to physicians' patient-centered information needs.¹

Previous work evaluating physician note content has focused on a variety of issues, including copypaste of material from previous notes, automatic import of electronic health record (EHR) data, use of structured versus unstructured narrative data, and information overload from verbiage added to reduce risk exposure or to enhance quality measure reporting.²⁻⁶ Beasley et al⁷ report that information chaos, composed of information overload, information underload, information scatter, information conflict, and erroneous information, is common in primary care. Using a human factors approach, they propose a framework of information chaos and its effect on physician mental workload and situational awareness, which affects physicians' problem-solving and decision-making capacity. "Information chaos is more than inconvenient, annoying, and frustrating; there are operational implications that can impair physician performance, increase workload, and reduce the safety and quality of care delivered."⁷ They call for improved display techniques to present the data needed at the time of the patient visit. Note quality has been evaluated by a variety of measures, including completeness and correctness; 22- and 9-item validated scales (Physician Documentation Quality Instruments [PDQI and PDQI-9]); and a multi-stakeholder evaluation of quality characteristics, desired elements, and system supports to improve note quality.8-11 A physician satisfaction study sponsored by the American Medical Association found a number of dissatisfactions with EHR use, among them "degraded clinical documentation (as a consequence of templatebased notes)."12

Lin et al¹³ compared user satisfaction among physicians at an academic health center that implemented assessment, plan, subjective, objective (APSO) notes institution-wide for authoring and creating APSO versus standard subjective, objective, assessment, plan (SOAP) notes and compared their ability to find data in both note types. They found physicians favored the change to APSO both as authors and as readers. Physicians reported the APSO notes were faster and easier to use, although objective speed measurements did not detect a difference.¹³ Various editorials and blog posts have endorsed moving the Assessment and Plan to the top of physician notes.^{14,15} Brown et al¹⁶ used an eye-tracking device to assess the visual attention

patterns of 10 hospitalists as they read 3 electronic notes. They found that of all the sections of inpatient hospitalist progress notes, physicians read the Assessment and Plan preferentially and most closely; indeed, over 90% of information conveyed during a verbal handoff came from the Assessment and Plan.¹⁶

Our earlier examination of information needs of physicians, patients, and billing staff led us to develop a new model of physician office notes. 17,18 We held focus groups showing physicians a hide/reveal feature that temporarily hides information needed only by nonphysician stakeholders. This feature provoked a dichotomy of reactions to the relative value and risks of such an approach. Some physicians endorsed the feature, whereas others feared that hiding some information would create a safety hazard, provoking the question, "Does hiding less-relevant information help or hinder physicians in their information-gathering task for patient care?"

To answer this important question, we built interactive prototypes that would allow us to usertest the hypothesis that the new note design would be faster, more accurate, and more satisfying to use than a standard note format for finding information needed for ambulatory chronic disease care. We based the designs on human factors as well as iterative user feedback from previous design-review sessions.

We aimed to make the most desired information more visually prominent, whereas subduing information considered less clinically relevant to reduce cognitive load, improve perceived ease of use, and improve speed and accuracy of information retrieval. Features that make desired information more prominent include the following:

- 1. Adding visual emphasis to abnormal data values with bold type and color to take advantage of preattentive visual processing¹⁹,
- 2. Adding abnormal data elements to the collapsed accordion header to reduce scrolling and remove visual distraction from normal text, and
- 3. Using a collapsible accordion header to hide some information considered less relevant (but retrievable with a single click) to reduce necessary reading and scrolling.

The interactive prototype incorporated the features in our model, moving the most sought-after sec-

Table 1. Description of Features for Four Model Electronic Health Record Note Visual Displays

Letter	Name of Note Visual Display Model	Features Distinguishing This Note Model			
A	SOAP note	Usual electronic health record format, serves as control			
В	APSO two-column note	Moves Assessment & Plan to top, adds second column for static data from the Past Family, Medical, and Social History			
С	Collapsible APSO one column	Same features as B, but only one column of text. Adds: Interactive, collapsible accordion display; headers for Review of Systems, Physical Exam, & Results displaying the abnormal text and showing abnormal item count & total organ system count; visual emphasis (bold & color type) for abnormal values in the body text of Review of Sytems, Physical Exam, & Results.			
D	Collapsible APSO two column	Same as C, but two columns instead of one			

APSO, assessment, plan, subjective, objective; SOAP, subjective, objective, assessment, plan.

tions of the note (Assessment and Plan; History of Present Illness) to the top of the note, while allowing the dynamic hiding/revealing of unwanted or less desired sections of the note. Those hidden sections have headers that still display key summary information while hiding more verbose and cluttered portions of each section.

Methods

We designed 4 options for presentation of information in physician clinic notes. We then used comparative usability testing to examine time to acquire information and accuracy of information acquired with the use of different formats of physician clinic notes. We also assessed cognitive load and perceived usability and usefulness of the different note formats.

Note Design

We iteratively developed physician clinic note prototypes with features designed to emphasize vital information and de-emphasize less important information, with level of importance identified by clinicians in a previous study. 17 We utilized 4 different visual display models (listed in Table 1 and pictured in Figure 1) for formatting the progress notes. We developed clinical content for the 4 simulated ambulatory physician progress notes for a patient with chronic disease. We designed the notes to be of similar complexity but with different clinical content to test 4 different visual display models. The notes were written by 2 authors (JB, SP) and were reviewed by 3 of the authors (JB, SP, RK). Each note included 4 problems, 1 of which was diabetes, 1 of which was an acute illness, and 2 of which chronic problems.

Figure 1. Diagram of content for 4 model notes. APSO, assessment, plan, subjective, objective; SOAP, subjective, objective, assessment, plan.

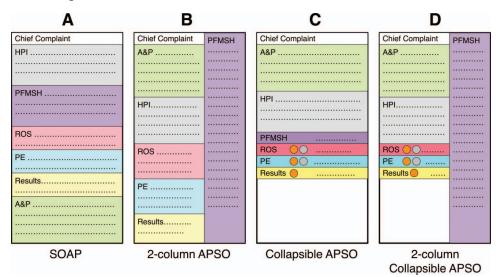
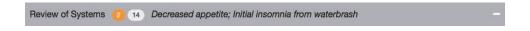


Figure 2. Collapsible section header with counts (abnormal values, total organ systems) and abnormal text.



Before testing, the researchers conducted design reviews with practicing physicians to ensure that the 4 clinical designs used in the study were free from coding mistakes or design oversights. Designs were typical of an EHR note without being specific to any particular EHR.

For each collapsible note design (notes C and D), any abnormal elements from Review of Systems, Physical Examination, and Results are listed by numeric count of abnormal terms, along with the total count of organ systems (billing bullet points) for the Review of Systems and Physical Examination (see Figure 2). Abnormal terms have added visual emphasis (eg, bold type and color) in the expanded text. Abnormal text elements are also listed in the section header adjacent to the counts and truncated by an ellipsis when the characters exceed the capacity of the header to display that entire content in 1 line. Collapsed content can then be expanded to reveal the full text content of that section (see Figure 3).

Sample

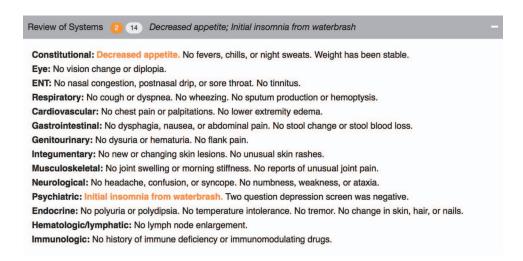
We recruited sixteen physicians who practice in the ambulatory clinics of the University of Missouri. We sampled for heterogeneity in sex, years since residency graduation, and years using EHR. Family medicine attending physicians formed the sample majority; smaller numbers of internal medicine attending physicians also participated.

Testing

Each session was performed on a Windows laptop using Morae software (TechSmith Corporation, Okemos, MI) to record audio, video, onscreen activity, and keyboard/mouse input, and a Tobii X2-60 eye-tracker (Danderyd, Sweden) to record eye movement patterns. Participants evaluated 4 different model notes from the role of a physician preparing to see a patient for the first time using a colleague's notes. They completed a series of tasks with each note sequentially, and after reviewing each note, provided an assessment of that note.

Each of the 4 visual display models was evaluated with 4 separate simulated physician progress notes with similar complexity but different clinical content. The order of exposure to the 4 notes changed with each user. To simulate the stress of clinical practice, researchers gave participants a limited amount of time to complete each task. Researchers also interrupted participants during the first task with questions common in actual clinical practice (eg, "The next patient has back pain and thinks she

Figure 3. Section header expanded to reveal content detail with abnormal values visually emphasized.



has a UTI. Should I do a urinalysis?") to simulate the increased cognitive load of typical clinical-setting interruptions.

Participating physicians viewed 1 note at a time and engaged in a sequence of tasks 1 at a time, addressing the following 4 tasks for each note:

- 1. What were the diagnoses, the medications ordered or changed, and the labs ordered at the last visit? (30 seconds)
- 2. Does the patient have any abnormal Review of Systems findings from the last visit? If so, what were they? (20 seconds)
- 3. Does the patient have any abnormal Physical Examination findings from the last visit? If so, what were they? (25 seconds)
- 4. Does the patient have any abnormal Results from the last visit? If so, what were they? (15 seconds)

The eye-tracking data and audio-recorded information was used to determine time to first fixation for sections relevant to each question, total time in each section, and the accuracy of note review relating to each task question.

After the participants completed each note review, they answered 3 additional sets of questions before proceeding to the next note. First, they were asked to complete the NASA Task Load Index (NASA-TLX 7) consisting of 7 Likert-scale questions to determine task difficulty for each design.²⁰ The second set of questions was composed of the System Usability Scale (SUS).²¹ These 10 questions are often used in evaluating system usability. The final set of questions asked participants to mark which sections of the physician's note they would want to be open by default for every patient if collapsible sections were available and which features of the design they thought were the most useful. Participant sessions typically lasted 40 to 60 minutes. The University of Missouri Institutional Review Board reviewed and approved this study.

Analysis

We compared time on task for notes A, B, C, and D using a Cox proportional hazards model to test whether task completion times differed by note model, task, and a note model-task interaction term.

Because each task was time-limited to mimic the sense of the urgency in actual clinical practice, there were a few physicians who did not complete a task. This then created a need to deal statistically with both the skewed distribution common with timed data and the time for noncompleters; we decided to use median as the summary statistic to account for skew and to censor noncompleters for the time-to-completion analyses.

Task success for items detected is expressed as percent correct. Task success for the different note models was compared using a logistic regression model with note A as the reference category. In the logistic model, we used "participant" as a random effect to account for the repeated measures of tasks taken from the same participant. We used paired *t*-tests to compare the means of notes B, C, and D to note A for scores from the NASA-TLX and the raw scores of the SUS. We also conducted a 1-way Analysis of Variance (ANOVA) comparing note A NASA-TLX and SUS scores to those for notes B, C, and D with "participant" as a random effect.

Results

Sample

Of the sixteen physicians who participated, 56% were female. Most of the participants were faculty (81%), 13% of whom were internal medicine physicians. There was a range of time in practice since graduation, but 94% reported over 5 years' experience using the EHR.

Task Time

Task time comparisons are influenced largely by the design features of each note section. The A&P (Assessment and Plan) design is the same in all 4 note models, and the median task times were not statistically different between any of the note models. The Results section is by far the simplest, containing the least content and the fewest visual distractions. However, the Review of Systems and Physical Examination include significant redesign of note models C (collapsible APSO) and D (2column collapsible APSO), with several features that make the target content more visible to the subject (Figures 2 and 3). Note models C and D containing the collapsible accordion elements have substantially faster median task times of 13.4 and 12.9 seconds for the Review of Systems, versus 20.0 and 18.7 seconds for note models A (traditional SOAP) and B (2-column APSO), respectively. Note models C and D have much faster median times for the Physical Examination section as well, 12.2 and 12.6 seconds versus 23.0 and 22.0 seconds for note models A and B (Table 2), respectively. These comparisons are significant in the Cox proportional hazards regression model as shown in Table 3.

Task Success

By design, the task success (accuracy) rates for all note sections were less than perfect in our time-limited information retrieval tasks. The lowest task success rate (63% to 77%) was observed in the Assessment & Plan section (the most complex task, one that included a user interruption), while the greatest task success rate was observed in the Results section (85% to 97%), the simplest task containing the fewest visual distractions.

Regarding note model comparisons, we unfortunately discovered a content error in the Physical Examination section of note D (2-column collapsible APSO) that caused note D to perform disproportionately worse in task success rates, both in comparison with the other note models and across note sections within note D (See Table 4). To adjust for the impact of the erroneous item, we recalculated the note task success rates after removing the Physical Examination section from each of the 4 note models. After adjustment, compared with note A, physicians had increased odds of task success with notes B (odds ratio [OR], 1.66; 95% CI, 1.05 to 2.63) and C (OR, 1.73; 95% CI, 1.10 to 2.72); odds of task success were not statistically different with note D (OR, 1.37; 95% CI, 0.87 to 2.15) (see Table 5).

Task Load

We measured task load by employing the NASA Task Load Index. As expected, none of the notes posed a significant physical demand (Table 6, Physical). Compared with note A, note B (2-column APSO) differed on the frustration subscale (3.31 vs 2.69; P=.01), note C (collapsible APSO) differed on the performance subscale (3.81 vs 2.69; P=.048), and note D (2-column collapsible APSO) differed on the effort subscale (3.94 vs 2.69; P=.043). However, in ANOVA with participant as a random effect, none of the subscales differed significantly by note type.

Usability (SUS)

Our physician test subjects found note A the least usable compared with the other note models, based

on the SUS (Table 6). In paired t-tests, SUS scores for note A were significantly lower than for note B (58.50 vs 74.83; P = .007), note C (58.50 vs 81.83; P = .005), and note D (58.50 vs 77.50; P = .009). In 1-way ANOVA with "participant" as a random effect, there was a significant mean difference between notes A and C (difference in means = 23.3; P = .01).

Feature Preferences

Users strongly endorsed all the new note features incorporated into the new note prototypes. Previously expressed concerns¹⁶ about temporarily hiding parts of the note (using the accordion display design pattern) were allayed. A large majority of subjects preferred to display these sections open by default in the expanded position in the collapsible note models (C & D): Chief Complaint, Assessment & Plan, Present Illness, Problem List, Medication List.

Discussion

Our study shows that collapsible accordion design notes may reduce physician time spent reviewing Review of Systems and Physical Examination sections of the notes. Physicians in our study completed increased proportion of tasks correctly with 2-column APSO and the collapsible APSO note designs. The SOAP note was perceived to be the least usable and the collapsible accordion design the most usable by our study participants. There were no differences in task difficulty between note designs. Lin et al¹³ reported comparable note reading speed and accuracy with APSO notes compared with SOAP notes.

Physicians face a dilemma in balancing competing values in their documentation work, both from a consumption and from a production perspective. Finding that balance has not been easy. Methods that make production easier, such as copying forward text from a previous note, automatically importing lists and labs, and using template text designed to satisfy nonclinician stakeholder demands, all add to the information overload and visual clutter when the time comes to consume that same note. Our 3 new model note styles address the needs of primary care physicians identified in our earlier studies^{17,18} and incorporate a number of human factors missing from current note designs. By far, the technologically simplest change is to

40 0 Table 2. Median Task Time (Interquartile Range) for Completion of All Assigned Tasks for Each Electronic Health Record (EHR) Note Section (n = 16) 30 Task Time (seconds) 25 20 15 10 Median (IQR) 24.8 (11.2) 12.6 (10.4) 25.7 (8.0) 20.0 (3.7) 18.7 (2.4) 13.4 (7.5) 12.9 (8.5) 12.2 (7.6) 13.3 (3.7) 13.1 (3.0) 10.3 (5.3) 23.0 (5.7) 26.2 (9.3) 27.4 (5.8) 22.0 (5.3) 12.2 (3.7) Note Model Д О Assessment & Plan (30 seconds): Find all diagnoses, Physical Exam (25 seconds): Find abnormal values in PE Note Section Name (Time Limit) Task Category Results (15 seconds): Find abnormal result values Review of Systems (20 seconds): Find abnormal values in ROS medication ordered/changed, & lab ordered

Physicians sequentially viewed each of 4 note EHR note models, completing each task sequence one note at a time.

Table 3. Comparing Task Completion Time Differences (P < .05) among Different Electronic Health Record Note Models for Each Note Section

Task	Note Versu Mo	P	
Physical exam	A	С	.0125
	A	D	<.0001
	В	С	.0023
	В	D	.0051
Results	A	С	.0012
	A	D	.0412
	В	С	.0476
Review of systems	A	С	.0255
	A	D	.0048
	В	С	.0065
	В	D	.0065

Note models: A (traditional SOAP), B (two-column APSO), C (collapsible APSO), D (two-column collapsible APSO). Only significant differences (P < .05) are displayed in the table. APSO, assessment, plan, subjective, objective; SOAP, subjective, objective, assessment, plan.

move Assessment & Plan to the top of the note. This can readily be done with almost any existing EHR, either by the vendor or client information technology team or even by individual clinician users.

We demonstrated alternative methods of adding visual emphasis to abnormal list items in the Review of Systems, Physical Examination, and Results sections. The simplest is to add emphasis employing colored text and bold font to abnormal items. A second method was to separately list the abnormal item count and total organ system count, and to display the abnormal text in the section header bar itself (Figure 2).

Table 5. Pairwise Comparisons of Accurate Retrieval of All Information for Each Electronic Health Record Note Type

Note v	s. Note	t Value	Odds Ratio (95% CI)
В	С	-0.16	0.96 (0.59 to 1.56)
В	D	0.81	1.22 (0.75 to 1.99)
В	A	2.22	1.66 (1.05 to 2.63)
C	D	0.98	1.27 (0.78 to 2.05)
C	A	2.43	1.73 (1.10 to 2.72)
D	A	1.37	1.37 (0.87 to 2.15)

Note models: A (traditional SOAP), B (two-column APSO), C (collapsible APSO), D (two-column collapsible APSO). Scores for each note pair were calculated without the faulty Physical Exam component.

APSO, assessment, plan, subjective, objective; SOAP, subjective, objective, assessment, plan.

In addition, we used the collapse/expand feature from the accordion-display design pattern. Clicking a section header alternately expands or collapses the content associated with the header. This feature is technically feasible with modern text display methods such as XHTML but much less feasible with older rich text format (RTF) or PDF displays.

There are significant technical and cultural challenges to enabling the consistent display of abnormal items within the Review of Systems, Physical Examination, and Results. Abnormal values may be identified by several different mechanisms. When words are selected as discrete data elements when physicians create the note (eg, clicking on the word "fever"), it is easy and measurable if each item carries a designation as abnormal or not. Typed or transcribed text is more problematic to identify as abnormal. Natural language processing can be em-

Table 4. Percent of Items Accurately Retrieved from All Possible Information for Each Electronic Health Record Note Type

Note Model	Note Sections Included	Mean (SD)	Median (min-max)	
A (traditional SOAP)	All	68.8 (11.9)	64.7 (52.9 to 88.2)	
	Excluding Physical Exam	69.6 (11.8)	64.3 (57.1 to 62.9)	
B (2-column APSO)	All	79.2 (14.6)	83.3 (46.7 to 93.3)	
	Excluding Physical Exam	78.9 (13.9)	84.6 (46.2 to 92.3)	
C (collapsible APSO)	All	80.5 (20.8)	90.6 (37.5 to 100.0)	
	Excluding Physical Exam	79.5 (21.0)	89.3 (42.9 to 100.0)	
D (2-column collapsible APSO)	All	69.1 (19.2)	75.0 (25.0 to 93.8)	
•	Excluding Physical Exam	75.5 (21.4)	75.0 (25.0 to 100.0)	

Note D contained a content error in the Physical Exam section that caused note D to perform disproportionately worse, thus we recalculated the results after removing the Physical Exam section from each of the four note models.

APSO, assessment, plan, subjective, objective; SOAP, subjective, objective, assessment, plan; SD, standard deviation.

Table 6. Comparison (Paired t-test) of Note Types for Perceived Workload (NASA-TLX) and Usability (SUS)

	Note						Paired t-Test Results				
	A		В		С		D		A Versus B	A Versus C	A Versus D
Note	Mean	SD	Mean	SD	Mean	SD	Mean	SD	P-Value	P-Value	P-Value
TLX—Mental	3.25	(1.69)	3.00	(1.32)	2.50	(1.83)	2.56	(1.36)	.52	.21	.21
TLX—Physical	1.75	(1.13)	1.56	(0.81)	1.44	(0.73)	1.38	(0.72)	.38	.33	.23
TLX—Timing	4.25	(1.57)	4.31	(1.49)	3.25	(1.95)	3.50	(1.71)	.85	.12	.18
TLX—Performance	3.81	(1.47)	3.63	(1.45)	2.69	(1.70)	3.31	(1.54)	.66	.048	.43
TLX—Effort	3.94	(1.77)	3.50	(1.32)	2.88	(2.00)	2.69	(1.74)	.30	.13	.043
TLX—Frustration	3.31	(1.78)	2.69	(1.54)	2.75	(1.98)	2.63	(1.75)	.01	.41	.28
TLX—Overall	3.39	(1.21)	3.11	(0.86)	2.58	(1.51)	2.68	(1.14)	.24	.11	.10
SUS*	58.50	(22.22)	74.83	(15.10)	81.83	(21.9)	77.50	(26.86)	.007	.005	.009

^{*}There was one subject who had missing data for note A. This subject was excluded from analysis specific to SUS. Bold reflects significant mean differences where P < .05.

ployed after the note is completed but adds cost, complexity, and ambiguity. Abnormal lab values can be expected to arrive with a flag denoting abnormality (eg, a serum potassium result will have a flag for high, low, or critical if outside the normal range). If the labels of abnormality are not consistently reliable, users will not trust the information display, and its utility will suffer dramatically. Our erroneous note (note model D the with the faulty Physical Examination values) vividly illustrated that erosion of trust once a test subject discovered the discrepancy between the header summary and the remainder of the section hidden by default.

Several limitations should be acknowledged. Our sample size was small, so some differences between note performance measures may not have been detected. Although we purposefully sampled both family medicine and general internal medicine physicians with different levels of experience as well as both attending and resident physicians, our study was limited to a single academic center. Our model notes were not actual notes from a current EHR, but rather were HTML page displays designed to be representative of the typical EHR note in their clinical content.

Conclusion

Starting with an understanding of physicians' information needs for a primary care chronic disease

visit and using human factors design principles, we developed innovative note models incorporating an array of display improvements. The 3 new note models offered equivalent or improved speed, accuracy, and user satisfaction over the standard note. Assessment-plan-subjective-objective (APSO) notes would be simple and inexpensive to implement for most organizations, as would adding emphasis to abnormal elements with selective use of color and typography. Two of these models hide and reveal note content strategically resulting in reduced information chaos by reducing information overload and information scatter. Using these same human factors principles in other aspects of information transfer, such as in computerized order entry or in clinical decision support, might reduce clinicians' burden further, and deserves systematic exploration. The next step is to incorporate as many features as feasible into the local implementation of our commercial EHR.

The authors thank Clayton Hicklin at the Tiger Institute for Health Innovation for development of the HTML interactive prototypes, Kenny Haggerty, Neeley Current, and Fatih Demir PhD at the Information Experience Lab for assistance with data collection and analysis. We thank Gaia Guirl-Stearley for critical review of the final manuscript.

To see this article online, please go to: http://jabfm.org/content/30/6/691.full.

Note models: A (traditional SOAP), B (two-column APSO), C (collapsible APSO), D (two-column collapsible APSO).

NASA-TLX contains 6 subscale items and an overall mean subscale score, each on a 7-point Likert scale in which lower TLX scores indicate less workload.

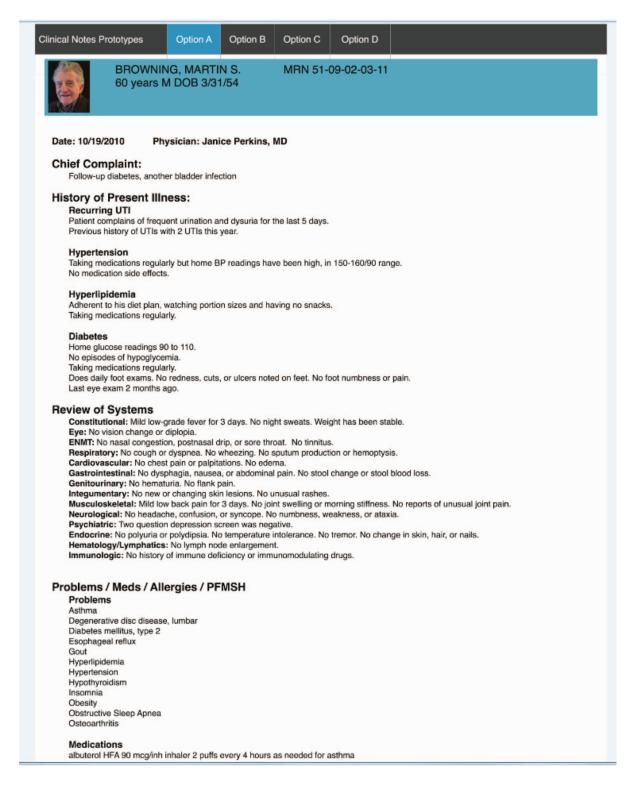
System Usability Scale (SUS) is reported as a raw score (scale of 0 to 100) in which larger values are considered better usability. TLX, Task Load Index; APSO, assessment, plan, subjective, objective; SOAP, subjective, objective, assessment, plan.

References

- 1. Belden JL, Koopman RJ. Making EHR notes more readable. Fam Pract Manag 2013;20:8-9.
- 2. Oxentenko AS, West CP, Popkave C, Weinberger SE, Kolars JC. Time spent on clinical documentation: A survey of internal medicine residents and program directors. Arch Intern Med 2010;170:377-
- 3. Shoolin J, Ozeran L, Hamann C, Bria W 2nd. Association of Medical Directors of Information Systems consensus on inpatient electronic health record documentation. Appl Clin Inform 2013;4:293-303.
- 4. Siegler EL, Adelman R. Copy and paste: A remediable hazard of electronic health records. Am J Med 2009;122:495-496.
- 5. Rosenbloom ST, Denny JC, Xu H, Lorenzi N, Stead WW, Johnson KB. Data from clinical notes: A perspective on the tension between structure and flexible documentation. J Am Med Inform Assoc 2011; 18:181-186.
- 6. Singh H, Wilson L, Reis B, Sawhney MK, Espadas D, Sittig DF. Ten strategies to improve management of abnormal test result alerts in the electronic health record. J Patient Saf 2010;6:121-123.
- 7. Beasley JW, Wetterneck TB, Temte J, et al. Information chaos in primary care: Implications for physician performance and patient safety. J Am Board Fam Med 2011;24:745-751.
- 8. Logan JR, Gorman PN, Middleton B. Measuring the quality of medical records: A method for comparing completeness and correctness of clinical encounter data. Proc AMIA Symp 2001;408-412.
- 9. Stetson PD, Bakken S, Wrenn JO, Siegler EL. Assessing electronic note quality using the Physician Documentation Quality Instrument (PDQI-9). Appl Clin Inform 2012;3:164-174.
- 10. Stetson PD, Morrison FP, Bakken S, Johnson SB. Preliminary development of the physician documen-

- tation quality instrument. J Am Med Inform Assoc 2008;15:534-541.
- 11. Hanson JL, Stephens MB, Pangaro LN, Gimbel RW. Quality of outpatient clinical notes: A stakeholder definition derived through qualitative research. BMC Health Serv Res 2012;12:407.
- 12. Friedberg MW, Chen PG, Van Busum KR, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. The RAND Corporation; 2015.
- 13. Lin CT, McKenzie M, Pell J, Caplan L. Health care provider satisfaction with a new electronic progress note format: SOAP vs APSO format. JAMA Intern Med 2013;173:160-162.
- 14. Pullen E. APSO needs to replace SOAP in EMRs. The Health Care Blog. 2010.
- 15. Versel N. APSO vs. SOAP, continued. Meaningful HIT News. 2014.
- 16. Brown PJ, Marquard JL, Amster B, et al. What do physicians read (and ignore) in electronic progress notes? Appl Clin Inform 2014;5:430-444.
- 17. Koopman RJ, Steege LM, Moore JL, et al. Physician information needs and electronic health records (EHRs): Time to reengineer the clinic note. J Am Board Fam Med 2015;28:316-323.
- 18. Clarke MA, Steege LM, Moore JL, Koopman RJ, Belden JL, Kim MS. Determining primary care physician information needs to inform ambulatory visit note display. Appl Clin Inform 2014;5:169-190.
- 19. Belden J, Patel J, Lowrance N, et al. Inspired EHRs: Designing for clinicians. 2014.
- 20. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advance Psychol 1988;52:139-183.
- 21. Sauro J. Measuring usability with the system usability scale (SUS). 2011. Available from: https:// measuringu.com/sus/. Accessed January 29, 2017.

Appendix A. Electronic Health Record Note Model A: Subjective-Objective-Assessment-Plan (SOAP) Note



allopurinol 300 mg 1 tablet daily aspirin 81 mg 1 tablet daily atorvastatin 40 mg 1 tablet daily furosemide 40 mg 1 tablet daily levothyroxin 100 mcg 1 tablet daily lisinopril 20 mg 1 tablet daily meloxicam 15 mg 1 tablet daily metformin 1000 mg 1 tablet twice daily omeprazole 20 mg 1 capsule twice daily potassium chloride 20 mEq 1 tablet twice daily trazodone 100 mg 2 tablets at bedtime

Allergies

Codeine Penicillin rash

Past Medical History

Viral meningitis, age 20

Procedures

Colonoscopy 2008 Appendectomy Cholecystectomy Arthroscopy R knee 1997

Denies cigarette use. Alcohol 3 drinks weekly. Denies other drug use. Exercises 3 days a week 30 to 60 minutes at the gym. Married, 2

Family History

Negative

Health Maintenance / Immunizations

7-1-12 Tdap 7-1-14

Physical Exam:

Vitals & Measurements: T: 37.5°C HR: 66 BP 165/72 WT: 60.3 kg

General: Awake, alert, oriented x3. No acute distress.

HEENT: PERRLA. Moist mucous membranes. Nostril normal. Thyroid normal.

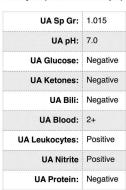
Respiratory: Bilateral lung sounds equal with no wheezing, rales, or rhonchi. No chest wall tenderness. Cardiovascular: Regular rate and rhythm. No ectopics. No JVD.

Abdomen: Mild suprapubic tenderness. Soft. No flank tenderness. Nondistended. No masses or hepatosplenomegaly. Bowel sounds

Extremities: 1+ non-pitting ankle edema. Capillary refill <3 seconds. No ulcers or gangrene noted.

Results:

Urinalysis (Reflex microscopic) (Today)



Assessment and Plan:

Acute UTI

Treat with nitrofurantoin 100 mg BID for 7 days since previous culture sensitive to it.

Urine culture sent.

Repeat UA at follow up in 4 weeks. Consider urology referral if hematuria is persistent, or if culture positive since 3rd UTI.

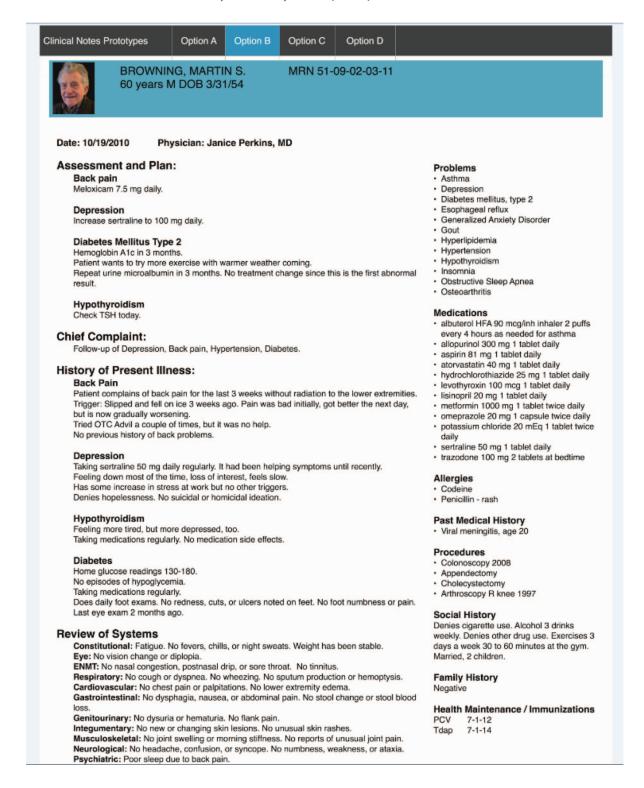
Add lisinopril-HCTZ 10-12.5 mg daily. That should help the mild edema. No symptoms of CHF. Recheck BMP in 2 weeks after starting lisinopril-HCTZ.

Diabetes Mellitus Type 2

Well-controlled.

Hyperlipidemia

Appendix B. Electronic Health Record Note Model B: 2-column Assessment-Plan-Subjective-Objective (APSO) Note



Endocrine: No polyuria or polydipsia. No temperature intolerance. No tremor. No change in

skin, hair, or nails.

Hematology/Lymphatics: No lymph node enlargement.

Immunologic: No history of immune deficiency or immunomodulating drugs.

Physical Exam:

BP 109/70 Vitals & Measurements: T: 37.5°C HR: 66 WT: 50.3 kg

General: awake, alert, oriented x3. No acute distress.

HEENT: PERRLA. Moist mucous membranes. Nostrils normal. Thyroid normal.

Respiratory: Bilateral lung sounds equal with no wheezing, rales, or rhonchi. No chest wall

Cardiovascular: Regular rate and rhythm. No ectopics. No JVD.

Back exam: Lumbar paraspinous tenderness. ROM limited by pain. SLR negative for

radiation to lower extremity. No masses or rash noted.

Lower extremity strength and sensation equal bilaterally. DTR 2+. Heel to toe gait intact.

Extremities: No edema. Capillary refill <3 seconds. No ulcers or gangrene noted.

Mental Status: Affect sad. Normal thought process and content. No paranoia, delusions, or

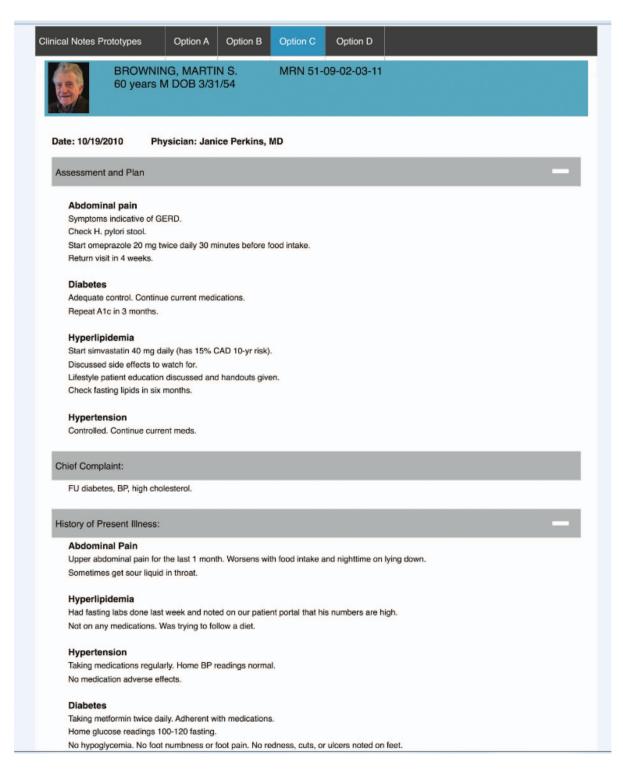
hallucinations. No suicidal ideation or thoughts.

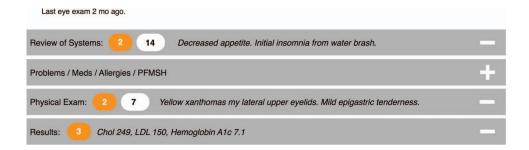
Results:

Hemoglobin A1c: (3 days ago) 8.1

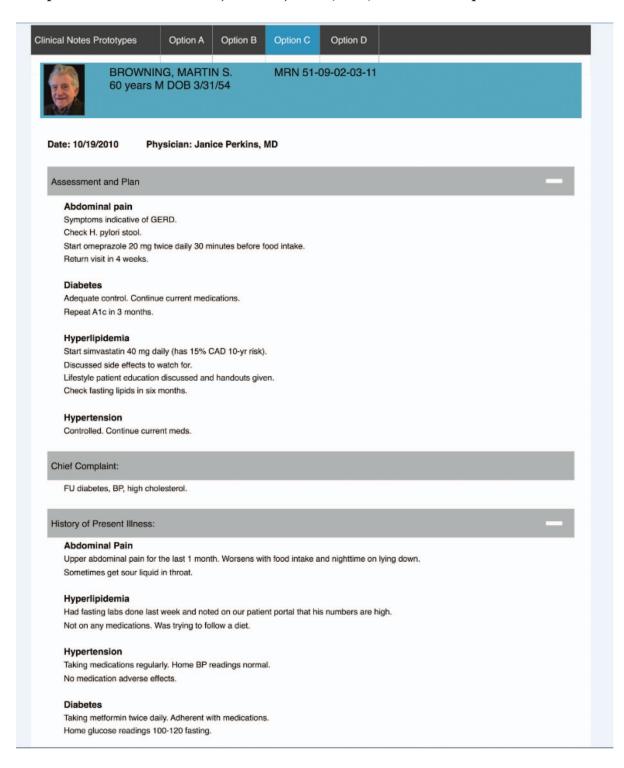
Urine microalbumin/creatinine ratio: (3 days ago)

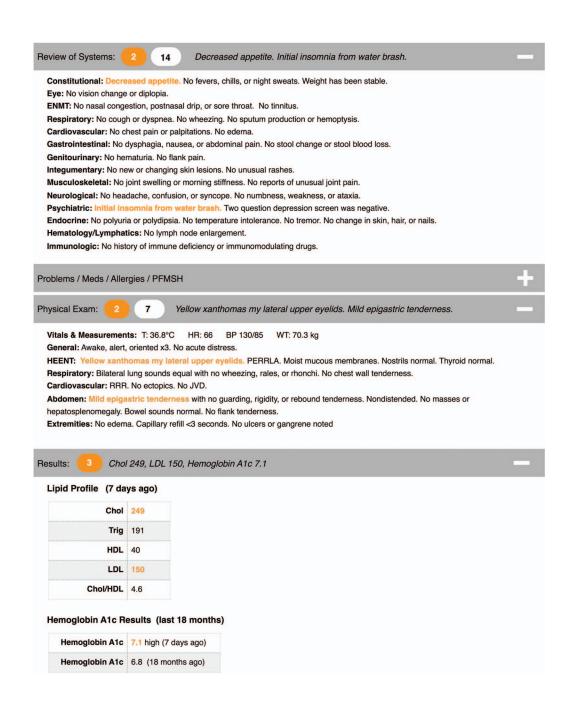
Appendix C-1. Electronic Health Record Note Model C: Collapsible Assessment-Plan-Subjective-Objective (APSO) note, shown collapsed



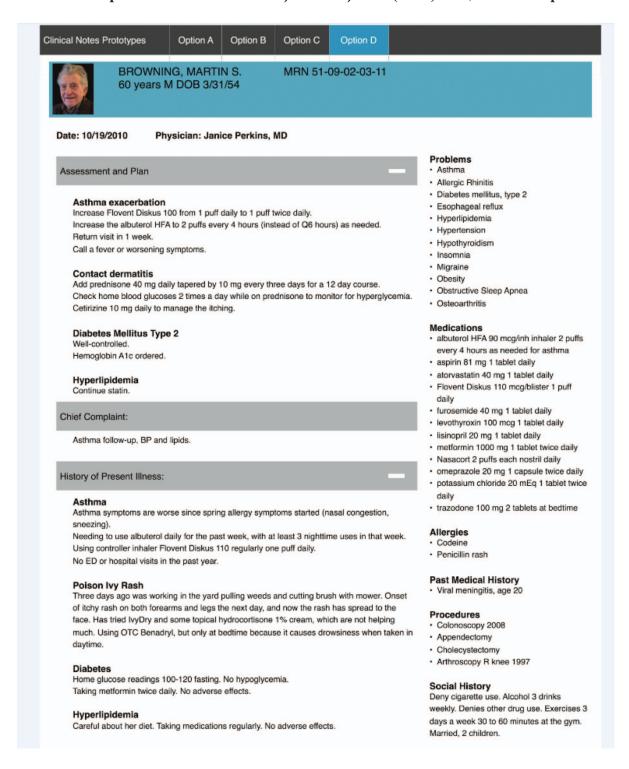


Appendix C-2. Electronic Health Record Note Model C: Collapsible Assessment-Plan-Subjective-Objective (APSO) note, shown expanded





Appendix D-1. Electronic Health Record Note Model D: 2-column collapsible Assessment-Plan-Subjective-Objective (APSO) note, shown collapsed



Review of Systems: 2 14 Occasional nosebleeds, Nocturnal cough	Family History Negative		
Physical Exam: 8 Pale, boggy nasal mucosa	Health Maintenance / Immunizations		
Results: Peak flow = 300	PCV 7-1-12 Tdap 7-1-15		

Appendix D-2. Electronic Health Record Note Model D:

2-column collapsible Assessment-Plan-Subjective-Objective (APSO) note, shown expanded

