

Article

A Natural Language Processing Approach to Automated Highlighting of New Information in Clinical Notes

Yu-Hsiang Su ^{1,2}, Ching-Ping Chao ³, Ling-Chien Hung ¹, Sheng-Feng Sung ^{1,3,*} 
and Pei-Ju Lee ^{3,4,*}

¹ Division of Neurology, Department of Internal Medicine, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi City 60002, Taiwan; 06500@cych.org (Y.-H.S.); 07177@cych.org (L.-C.H.)

² Department of Cosmetology and Health Care, Min-Hwei Junior College of Health Care Management, Tainan 73658, Taiwan

³ Department of Information Management and Institute of Healthcare Information Management, National Chung Cheng University, Chiayi County 62102, Taiwan; s6951435@gmail.com

⁴ Center for Innovative Research on Aging Society (CIRAS), National Chung Cheng University, Chiayi County 62102, Taiwan

* Correspondence: sfsung@cych.org.tw (S.-F.S.); pjlee@mis.ccu.edu.tw (P.-J.L.)

Received: 26 March 2020; Accepted: 17 April 2020; Published: 19 April 2020



Abstract: Electronic medical records (EMRs) have been used extensively in most medical institutions for more than a decade in Taiwan. However, information overload associated with rapid accumulation of large amounts of clinical narratives has threatened the effective use of EMRs. This situation is further worsened by the use of “copying and pasting”, leading to lots of redundant information in clinical notes. This study aimed to apply natural language processing techniques to address this problem. New information in longitudinal clinical notes was identified based on a bigram language model. The accuracy of automated identification of new information was evaluated using expert annotations as the reference standard. A two-stage cross-over user experiment was conducted to evaluate the impact of highlighting of new information on task demands, task performance, and perceived workload. The automated method identified new information with an F1 score of 0.833. The user experiment found a significant decrease in perceived workload associated with a significantly higher task performance. In conclusion, automated identification of new information in clinical notes is feasible and practical. Highlighting of new information enables healthcare professionals to grasp key information from clinical notes with less perceived workload.

Keywords: bigram language model; electronic medical records; information overload; natural language processing

1. Introduction

Electronic medical records (EMRs) have been developed in Taiwan for more than a decade [1]. They have been implemented in most of the larger medical institutions to store information about encounters and events between patients and healthcare systems [2]. The implementation of EMRs not only enables large-scale storage and collection of patient data but also makes the exchange of healthcare information between healthcare facilities possible through the Electronic Medical Record Exchange Center [2]. Almost all the medical complaints, diagnoses, processes of clinical care, laboratory results, and medication use from various departments of most domestic hospitals are readily available from EMRs. The immediate accessibility of patient information is likely to help healthcare professionals improve patient care delivery and enhance the quality of medical decision-making [3,4]. Moreover,

the development of integrated clinical decision support systems and EMRs has achieved substantial success in reducing medical errors and improving patient safety [5].

Despite the positive changes brought by EMRs, patient information explosion associated with the rapid accumulation of a large amount of unstructured data has become a threat to the effective use of EMRs. The immense amount of information stored in EMRs can lead to information overload for healthcare professionals [6]. In addition, the use of EMRs may interfere with the interaction between patients and practitioners and cause dissatisfaction and burnout among healthcare professionals [7]. Furthermore, because current EMR systems generally provide several time-saving features such as the “copy-and-paste” function, physicians frequently use copying and pasting to reduce the omission of what they consider important information. Clinical notes have thus become full of redundant information and barely able to convey useful information [8]. These redundant contents also cause difficulties in reading and prolong reading time.

As more longitudinal clinical narratives are produced with the increasing number of healthcare encounters, the practice of copying and pasting inevitably generates more redundant information, which is actually noise and masks new and clinically relevant information within notes [9]. A systematic review found that 66% to 90% of clinicians routinely use copy-and-paste and approximately 80% of physicians use copy-and-paste regularly for inpatient documentation [10]. Such practices are similarly common among residents and attendings [11]. Despite the many deficits in notes written using copy-and-paste, approximately 80% of physicians agreed that copy-and-paste behaviors should continue [12] considering that almost half of their work time is spent on EMR-related work [13,14].

Now that the adoption of healthcare information technology has brought about lots of redundant information in EMRs, healthcare information technology should be able to help reduce the interference from redundant information. According to a survey on the usability of EMR systems, “ease of finding the required information on the screen” is the most desired requirement [15]. Hence, many studies have focused on the redesign of EMR interfaces, hopefully facilitating clinicians to keep track of relevant patient information. Well-designed data visualization in EMR systems not only allows healthcare professionals to communicate information efficiently and effectively but also improves data interpretation and clinical reasoning [16,17]. Besides, it is recommended that copied material should be displayed in a different font or color, so that they can be easily identified [10].

With the advances in natural language processing, investigators have tried to build applications around NLP technologies for summarization or extraction of needed information from longitudinal patient records within EMRs [17,18]. Among them, Zhang et al. have developed algorithms based on statistical language models to identify relevant new information in longitudinal clinical narratives [9,19]. Experimentation with a visualization tool for the presentation of new information also found some positive influences on the synthesis of patient information from EMRs [20]. Motivated by these works, this preliminary study aimed to investigate (1) the amount of new versus redundant information in inpatient clinical notes; (2) the accuracy of automated identification of new information in clinical notes; and (3) whether highlighting of new information affects the performance, task demands, and perceived workload of healthcare professionals in reviewing clinical notes.

2. Materials and Methods

2.1. Study Setting

This study was conducted in Ditmanson Medical Foundation Chia-Yi Christian Hospital, a 1000-bed teaching hospital located in southern Taiwan. It employs 3000 staff, with approximately 47,000 admissions, 1,110,000 outpatient visits, and 89,000 emergency visits per year. The study protocol was approved by the Ditmanson Medical Foundation Chia-Yi Christian Hospital Institutional Review Board (CYCH-IRB No.2018085).

2.2. Clinical Notes and Manual Annotation

A purposive sample of ten patients was selected for review of clinical notes from the inpatient population of medical wards and intensive care units. Patients selected for this study had to be hospitalized for more than 10 days, with complex conditions and multiple comorbidities. All clinical notes were checked to ensure that clinicians participating in this study had not taken care of the patient at any time previously. Patient identifiers were replaced by a unique study identification number to ensure confidentiality; the informed consent was thus exempted.

Two experienced attending physicians (LCH and SFS) independently annotated the clinical notes of each patient. After reviewing the admission note, they evaluated the subsequent 9 days of progress notes to identify new information based on all preceding notes chronologically using their clinical judgment. Inter-rater agreement was assessed at the line level using the Kappa statistic. Discrepancies between the two annotators were arbitrated by consensus. The final set of annotations was used as the reference standard for automated highlighting of new information.

2.3. Automated Highlighting Using the Bigram Language Model

A statistical language model is a probability distribution over word sequences. It is useful in many natural language processing applications, such as speech recognition, text categorization, and information retrieval. An n -gram model is a type of language model that approximates the probability of observing a word based on the preceding $n-1$ words in a word sequence [21]. For example, a bigram model estimates the occurrence of a word in the context of the preceding one word.

The automated highlighting of clinical notes largely followed the method developed by Zhang et al. using the bigram language model [19,22]. All the clinical notes from the same patient were ordered chronologically. The text was preprocessed through sentence splitting, stop-word removal, spell checking, and stemming. Then a bigram language model was built based on preceding notes to identify new information in the target note. If a bigram had never appeared in any preceding notes, the sentence containing the bigram was considered new information and was thus highlighted [19].

2.4. Evaluation of Automated Highlighting

Precision (positive predictive value), recall (sensitivity), and F1 score were used to evaluate the performance of automated highlighting of new information against the reference standard at the line level. A true positive means that a line containing new information was identified by both automated and expert annotation. A false positive indicates that a line was highlighted as having new information by automated but not by expert annotation, while a false negative indicates that a line was highlighted as having new information by an expert but not by automated annotation. Precision was calculated as true positives divided by the sum of true positives and false positives, recall as true positives divided by the sum of true positives and false negatives, and F1 score as 2 times precision times recall, divided by the sum of precision and recall.

To determine the optimal number (N) of preceding notes for building the bigram language model, we varied N from 1 to 8. In other words, each target note was highlighted based on the preceding 1 to at most N notes of the target note. The N value that achieved the highest F1 score was used in the user experiment.

2.5. User Experiment

A convenience sample of twelve clinicians from the staff of medical wards and intensive care units was recruited. The participants were contacted in person and were asked if they were willing to participate. Participation was voluntary and compensated. Age, gender, and years of experience in clinical practice were collected for each participant.

Four of the ten patients used in the first experiment were selected for the user experiment. The selection was made to balance the number of clinical notes and the amount of text as best as

possible. The number of lines per progress note was similar among the four patients, ranging from 21 to 23. All of them had multiple underlying comorbidities and a complicated hospitalization course. Each patient had several active problems that needed further investigations to determine the etiology and repeated evaluations of the response to treatment. These intricate clinical scenarios may better represent the daily practice of the study participants.

The user experiment used a 2-period crossover design (Figure 1). The total number of participants was set at 12 so that each patient would appear six times in testing scenarios with notes in the original condition and those with notes in the highlighted condition. Six participants first reviewed the clinical notes of two of the four patients in the original condition (period 1) and then reviewed the notes of the other two patients in the highlighted condition (period 2). After each study period, participants had to fill a National Aeronautics and Space Administration task load index (NASA-TLX) questionnaire. The other six participants reviewed the clinical notes in the reverse order. That is, they first reviewed the notes of two patients in the highlighted condition and then the notes of the other patients in the original condition.

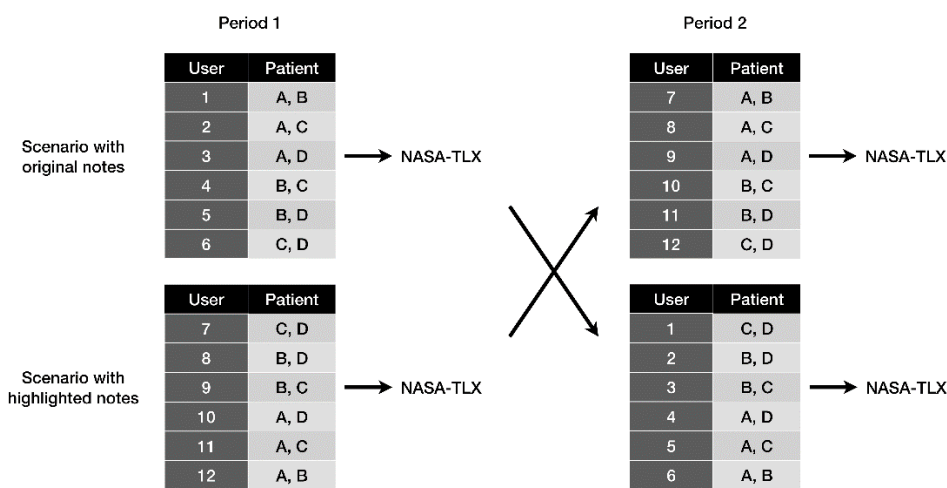


Figure 1. Workflow of the user experiment.

2.6. Measurement of Task Demands

This study used Morae Recorder, version 3.3.4, (TechSmith Corporation, Okemos, MI) to record screen captures and track mouse actions. All participants used the same computer and had access only to the clinical notes displayed in a standard web browser. Immediately before conducting the experiment, participants were instructed on how to browse the clinical notes. They were asked to review the notes at the same pace as they used to, and no time limit was set. The total numbers of mouse clicks, wheels, and mouse moves (in pixels) and the total time for participants to complete each testing scenario were used to measure task demands.

2.7. Measurement of Task Performance

During the note review process, participants had to complete a 20-item task questionnaire for each patient. Task items mainly focused on clinical fact or event finding (e.g., “Does this patient have a history of hypertension?”), date finding (e.g., “When did this symptom start?”), and clinical comparisons (e.g., “Was the condition getting better?”). Half of the answers to these task items came from the admission note and the other half from the progress notes. Reference answers were also obtained from the two expert annotators. Answers of participants were scored as correct or incorrect against the reference answers. Each correct answer was given one point. In addition to the total scores, the scores were subtotaled separately for questions regarding the admission note and those regarding the progress notes. The scores were normalized from 0 to 100 to represent task performance.

2.8. Measurement of Perceived Workload

The NASA-TLX, a widely used tool to assess workload and effectiveness in humans [23], was applied to evaluate perceived workload. The NASA-TLX consists of six dimensions including mental demand, physical demand, temporal demand, overall performance, effort, and frustration level [24]. It has been used to quantify the perceived workload associated with the use of EMRs [25,26]. The workload score ranges from 0 to 100 for each dimension, with a higher score indicating a greater workload. To obtain the weight of each dimension of the NASA-TLX, each participant performed 15 separate pairwise comparisons of the 6 dimensions to determine the relative relevance of each dimension in the task of reviewing clinical notes. Next, an overall NASA-TLX score was obtained by multiplying the dimension score with the corresponding dimension weight, summing across all dimensions, and dividing by 15.

2.9. Statistical Analysis

Given the small sample size in the user experiment, non-parametric statistical analyses were performed. Continuous variables were reported with medians and interquartile ranges. The Wilcoxon signed-rank test was performed for comparison between testing scenarios with original notes and those with highlighted notes because users were measured repeatedly [17]. Two-tailed p values < 0.05 were considered statistically significant. Statistical analyses were performed using Stata 15.1 (StataCorp, College Station, TX, USA).

3. Results

Table 1 lists the characteristics of the clinical notes selected for review. The number of progress notes within 9 days following admission ranged from 7 to 17. The average number of lines per progress note varied from 15 to 35 and increased with the number of progress notes (Figure 2A). The inter-rater agreement (kappa value) for manual annotation between the two attending physicians was 0.767, indicating substantial agreement. Based on the results of manual annotation, 34% to 78% of the lines of the progress notes were determined to contain new information. The proportion of lines with new information was negatively correlated with the average number of lines per progress note (Figure 2B).

Table 1. Characteristics of the clinical notes selected for review.

| | Total Lines in Admission Note | Number of Progress Notes | Total Lines in Progress Notes | Lines per Progress Note, Mean | Lines with New Information |
|----------------------|-------------------------------|--------------------------|-------------------------------|-------------------------------|----------------------------|
| Case 1 ¹ | 74 | 11 | 248 | 23 | 60% |
| Case 2 | 47 | 8 | 183 | 23 | 58% |
| Case 3 | 36 | 8 | 120 | 15 | 78% |
| Case 4 | 117 | 9 | 142 | 16 | 66% |
| Case 5 ¹ | 57 | 9 | 191 | 21 | 61% |
| Case 6 ¹ | 31 | 13 | 286 | 22 | 40% |
| Case 7 | 83 | 17 | 593 | 35 | 48% |
| Case 8 | 69 | 12 | 379 | 32 | 34% |
| Case 9 | 55 | 7 | 178 | 25 | 49% |
| Case 10 ¹ | 58 | 9 | 187 | 21 | 39% |

¹ Used in the user experiment.

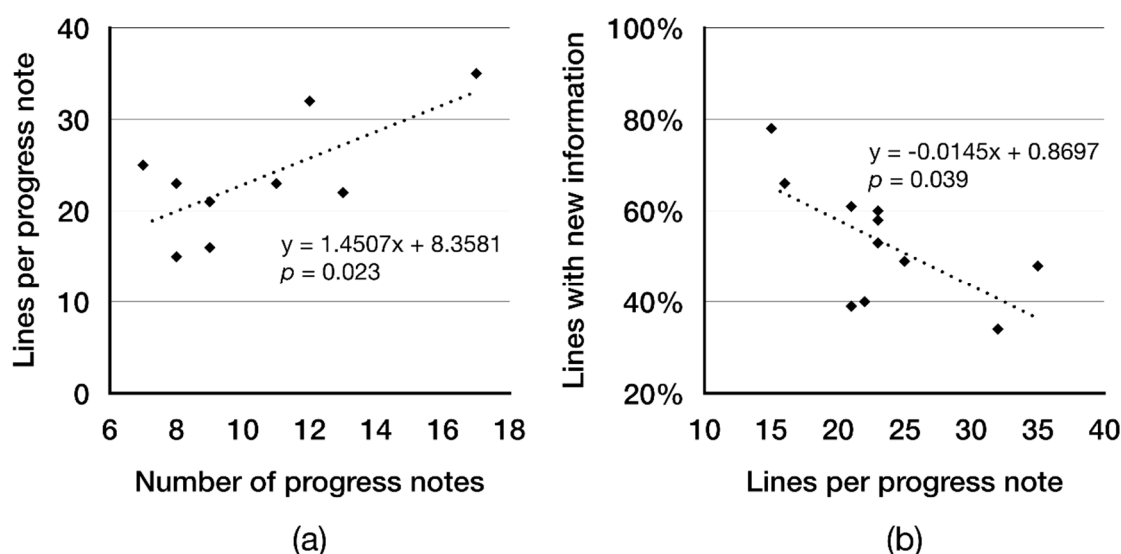


Figure 2. The relationship between lines per progress note and number of progress notes (a) and the relationship between lines with new information and lines per progress note (b).

Table 2 gives the performance of the automated identification of new information across different numbers of preceding notes used in the bigram language model. The highest F1 score (0.833) and accuracy rate (0.814) were achieved when at most four preceding notes were employed to build the bigram language model. Therefore, the optimal number (*N*) of preceding notes was set to 4.

Table 2. Performance of automated identification of new information using manual annotation as the reference standard by increasing the number of previous notes.

| Previous Notes, <i>N</i> | Accuracy | Precision | Recall | F1 Score |
|--------------------------|----------|-----------|--------|----------|
| 1 | 0.764 | 0.690 | 0.953 | 0.801 |
| 2 | 0.792 | 0.723 | 0.945 | 0.819 |
| 3 | 0.811 | 0.747 | 0.938 | 0.832 |
| 4 | 0.814 | 0.751 | 0.936 | 0.833 |
| 5 | 0.814 | 0.754 | 0.929 | 0.832 |
| 6 | 0.810 | 0.752 | 0.921 | 0.828 |
| 7 | 0.809 | 0.753 | 0.918 | 0.827 |
| 8 | 0.808 | 0.752 | 0.916 | 0.826 |

Clinical notes from case 1, 5, 6, and 10 (Table 1) were selected for the user experiment. The new information in each note was highlighted using the bigram language model mentioned above. Four physicians and eight nurse practitioners participated in the experiment. Table S1 lists the characteristics of the participants. Table 3 gives descriptive statistics of task demands, performance, and perceived workload for each testing scenario. No significant differences in task demands were observed between scenarios as quantified by the time to completion as well as the total number of mouse clicks, mouse wheels, or mouse movements.

Table 3. Comparison of task demands, task performance, and perceived workload between scenarios using original notes and highlighted notes.

| | Testing Scenario with Original Notes | Testing Scenario with Highlighted Notes | <i>P</i> |
|--|--------------------------------------|---|----------|
| Task demands | | | |
| Task time, min | 30.7 (29.0–36.0) | 28.7 (25.1–34.1) | 0.060 |
| Mouse click | 132.0 (91.5–169.5) | 116.0 (76.5–176.0) | 0.195 |
| Mouse wheel | 1060.5 (726.0–1147.0) | 797.5 (665.0–1124.5) | 0.308 |
| Mouse movement, 10 ³ pixels | 195.4 (133.8–286.9) | 166.6 (124.5–275.6) | 0.209 |
| Task Performance | | | |
| Admission note | 92.5 (87.5–100.0) | 95.0 (92.5–100.0) | 0.374 |
| Progress notes | 85.0 (80.0–95.0) | 95.0 (90.0–100.0) | 0.003 |
| Overall | 90.0 (82.5–93.8) | 95.0 (91.3–98.8) | 0.003 |
| Perceived Workload | | | |
| Mental | 70.0 (60.0–82.5) | 62.5 (45.0–70.0) | 0.003 |
| Physical | 62.5 (30.0–75.0) | 37.5 (22.5–52.5) | 0.007 |
| Temporal | 70.0 (57.5–80.0) | 50.0 (42.5–62.5) | 0.003 |
| Performance | 30.0 (20.0–52.5) | 30.0 (20.0–50.0) | 0.471 |
| Effort | 72.5 (67.5–85.0) | 60.0 (50.0–75.0) | 0.003 |
| Frustration | 50.0 (20.0–70.0) | 30.0 (20.0–50.0) | 0.038 |
| Overall NASA-TLX | 62.0 (55.5–71.3) | 50.2 (43.3–61.0) | 0.003 |

As for task performance, there was no difference between scenarios in the sub-scores for questions regarding the admission note. In contrast, the overall scores and the sub-scores for questions regarding the progress notes were significantly higher in the testing scenario with notes in the highlighted condition. The overall perceived workload of reviewing highlighted notes was significantly lower than that of reviewing original notes. The workload in different dimensions of the NASA-TLX decreased significantly except for the dimension of perceived overall performance.

4. Discussion

4.1. Effects of Information Redundancy

This study found that a substantial proportion of clinical notes contained redundant information instead of new information. The proportion of redundant information increased with the size of notes. However, the bigram language model could effectively identify new information based on preceding notes. By highlighting new information in clinical notes, healthcare professionals could more accurately extract relevant information from clinical notes. In the meantime, the perceived workload associated with reviewing clinical notes was significantly reduced even though the task demand did not change.

The user experiment showed that participants performed well in extracting information from admission notes and highlighted progress notes. In contrast, participants were less likely to accurately collect relevant information from original non-highlighted progress notes. It may be because admission notes contained only brand-new information while progress notes contained a lot of redundant information. With such abundant redundancy of information, users might be cognitively overloaded and unable to retrieve useful information, thus compromising their performance. This problem is likely to get worse with the increasing number of progress notes. A previous study revealed that the uniqueness of progress notes over the course of hospitalization dramatically decreased with time and contained only 27.7% unique information at the end of hospitalization [27].

4.2. Merits of Highlighting New Information

The performance in extracting relevant information was significantly improved when new information in progress notes was highlighted. Text highlighting enhances not only searching but also reading performance [28]. In addition, the highlighting of new information may help users relieve

information overload and concentrate on new information. The effect of highlighting can probably be explained by the psychological theory of human information processing. As proposed by Schneider and Shiffrin [29], “controlled processing” of information allows humans to read and understand information but requires attention and thus has limited capacity. Their experiments showed that when targets and distractors become more similar, the search tasks become more difficult, especially when the number of distractors increases. Similarly, the excessive redundant information in clinical notes can cause information overload and attention deficit. The highlighting of new information increases the contrast between targets and distractors, leading to a decrease in information overload and improvement in clinical reasoning.

Although a nonsignificant decrease in the time to completion was observed when highlighted notes were reviewed, there was no difference in mouse usage between testing scenarios. The possible reasons are as follows: First, clinical notes, whether highlighted or not, were presented to users in full rather than summarized form. Second, all the participants are trained healthcare professionals. They are accustomed to browsing the whole content of clinical notes to extract relevant information. Third, participants might not be confident in the accuracy of highlighting. Therefore, they would rather thoroughly review the clinical notes than skip non-highlighted redundant information.

It is worth noting that, despite similar task demands between testing scenarios, the overall perceived workload as assessed using the NASA-TLX was significantly reduced when new information was highlighted. Interestingly, the perceived overall performance, one of the six dimensions of the NASA-TLX, did not change between testing scenarios (Table 3). This dimension measures how successful the subject is in performing the task and how satisfied the subject is with their own performance. In the user experiment, even though the participants believed they performed equally well in both scenarios, they actually did worse according to the scores of task performance (Table 3). In other words, the participants were totally unaware of how the increased workload had affected their performance.

4.3. Clinical Implications

The study findings hint that the interference of redundant information to clinical practice is a real but under-recognized problem for healthcare professionals. This problem may become even worse in real-world settings where interruptions and multitasking are common [30], leading to medical errors and patient safety issues. The widespread use of EMRs has brought about several advantages, such as flexibility in storage and retrieval of data, easy access across different locations, and simultaneous use by multiple users. On the other hand, the use of EMRs is inevitably associated with some drawbacks. For example, it takes longer to read text on a computer screen than to read printed text [31]. Readers may need extra cognitive processes to gain the corresponding knowledge from a computer screen than from paper [32]. Furthermore, because of the increase in time spent on completing notes [7,13,14], routine use of copy-and-paste has become highly prevalent among physicians [10,11], thus generating a lot of redundant information in EMRs.

Moreover, the layout and structure of EMRs strongly impact the retrieval of information and sometimes influence clinical decision-making in fundamental ways. Poorly designed interactions with information technology can mislead decision-making and create medical errors, ending in patient harm [33]. An analysis found that 73% of EMR-associated patient safety issues were related to human-computer interaction [34]. Being overwhelmed by less important redundant information and failing to identify relevant new information may further interfere with decision-making [22]. Since it is unrealistic to get rid of all redundant information from EMRs, some measures should be taken to facilitate capturing relevant information more easily. In this regard, highlighting new information can effectively reduce the interference from redundant information and help healthcare professionals grasp key information from EMRs and ameliorate their perceived workload.

4.4. Limitations

This study has several limitations. First, this is a preliminary study with a small sample size from a single institution. Further studies that recruit larger samples are warranted to verify the impact on clinical practice. Second, the study took place in a controlled environment with minimum distracting elements. In real-world settings, healthcare professionals generally need to manage multiple patients in a short period of time, where a higher cognitive demand is required. Therefore, whether highlighting of new information is similarly effective or even more effective under real-world clinical practice conditions is open to question. Third, highlighted clinical notes were entirely new to the study participants, who might not have confidence in the accuracy of highlighting. They might even spend more time reading non-highlighted text. This might increase the time to completion and interfere with the accuracy of task performance. Fourth, because four consecutive preceding clinical notes are required to optimally determine whether the information is redundant, highlighting may not be worthwhile for patients with a short hospital stay. Finally, the majority of the study participants and the writers of clinical notes are not native English speakers even though clinical notes in the study hospital are documented in English. Language barriers may result in different behaviors in copying and pasting clinical text and incur extra cognitive load in extracting information.

5. Conclusions

The use of EMRs brings lots of redundant information, which is potentially harmful to patient safety. Nevertheless, this study shows that automated identification of new information in clinical notes is feasible and practical. Highlighting new information enables users of EMRs to have better understanding of patients' conditions and complete their daily work with less perceived workload. In particular, it may help healthcare professionals grasp key information from clinical notes of unfamiliar patients in situations such as consultations and shifting of services, hopefully improving the quality of medical decision-making.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-3417/10/8/2824/s1>, Table S1: Characteristics of participants in the user experiment.

Author Contributions: Conceptualization, Y.-H.S. and S.-F.S.; data curation, C.-P.C., L.-C.H., and S.-F.S.; formal analysis, Y.-H.S., C.-P.C., and S.-F.S.; funding acquisition, S.-F.S. and P.-J.L.; investigation, Y.-H.S., C.-P.C., L.-C.H., S.-F.S., and P.-J.L.; methodology, S.-F.S. and P.-J.L.; project administration, S.-F.S. and P.-J.L.; resources, P.-J.L.; software, C.-P.C.; validation, L.-C.H. and S.-F.S.; writing—original draft, Y.-H.S., C.-P.C., and L.-C.H.; writing—review and editing, S.-F.S. and P.-J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ditmanson Medical Foundation Chia-Yi Christian Hospital and Center for Innovative Research on Aging Society (CIRAS) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by Ministry of Education (MOE) in Taiwan, grant number RCN004.

Acknowledgments: We would like to thank Li-Ying Sung for English language editing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rau, H.H.; Hsu, C.Y.; Lee, Y.L.; Chen, W. Developing electronic health records in Taiwan. *IT Prof.* **2010**, *12*, 17–25. [[CrossRef](#)]
2. Li, Y.C.; Yen, J.C.; Chiu, W.T.; Jian, W.S.; Syed-Abdul, S.; Hsu, M.H. Building a national electronic medical record exchange system - experiences in Taiwan. *Comput. Methods Programs Biomed.* **2015**, *121*, 14–20. [[CrossRef](#)] [[PubMed](#)]
3. Ben-Assuli, O.; Leshno, M.; Shabtai, I. Using electronic medical record systems for admission decisions in emergency departments: Examining the crowdedness effect. *J. Med. Syst.* **2012**, *36*, 3795–3803. [[CrossRef](#)] [[PubMed](#)]
4. Ben-Assuli, O.; Sagi, D.; Leshno, M.; Ironi, A.; Ziv, A. Improving diagnostic accuracy using EHR in emergency departments: A simulation-based study. *J. Biomed. Inform.* **2015**, *55*, 31–40. [[CrossRef](#)]

5. Hydari, M.Z.; Telang, R.; Marella, W.M. Saving Patient Ryan—Can advanced electronic medical records make patient care safer? *Manag. Sci.* **2018**, *65*, 2041–2059. [[CrossRef](#)]
6. Farri, O.; Pieckiewicz, D.S.; Rahman, A.S.; Adam, T.J.; Pakhomov, S.V.; Melton, G.B. A qualitative analysis of EHR clinical document synthesis by clinicians. *AMIA Annu. Symp. Proc.* **2012**, *2012*, 1211–1220.
7. Shanafelt, T.D.; Dyrbye, L.N.; Sinsky, C.; Hasan, O.; Satele, D.; Sloan, J.; West, C.P. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clin. Proc.* **2016**, *91*, 836–848. [[CrossRef](#)]
8. Hirschtick, R.E. A piece of my mind. Copy-and-paste. *JAMA* **2006**, *295*, 2335–2336. [[CrossRef](#)]
9. Zhang, R.; Pakhomov, S.V.; Lee, J.T.; Melton, G.B. Using language models to identify relevant new information in inpatient clinical notes. *AMIA Annu. Symp. Proc.* **2014**, *2014*, 1268–1276.
10. Tsou, A.Y.; Lehmann, C.U.; Michel, J.; Solomon, R.; Possanza, L.; Gandhi, T. Safe Practices for Copy and Paste in the EHR. Systematic Review, Recommendations, and Novel Model for Health IT Collaboration. *Appl. Clin. Inform.* **2017**, *8*, 12–34.
11. Thornton, J.D.; Schold, J.D.; Venkateshaiah, L.; Lander, B. Prevalence of copied information by attendings and residents in critical care progress notes. *Crit. Care Med.* **2013**, *41*, 382–388. [[CrossRef](#)]
12. O'Donnell, H.C.; Kaushal, R.; Barrón, Y.; Callahan, M.A.; Adelman, R.D.; Siegler, E.L. Physicians' attitudes towards copy and pasting in electronic note writing. *J. Gen. Intern. Med.* **2009**, *24*, 63–68. [[CrossRef](#)]
13. Sinsky, C.; Colligan, L.; Li, L.; Prgomet, M.; Reynolds, S.; Goeders, L.; Westbrook, J.; Tutty, M.; Blike, G. Allocation of physician time in ambulatory practice: A time and motion study in 4 specialties. *Ann. Intern. Med.* **2016**, *165*, 753–760. [[CrossRef](#)]
14. Young, R.A.; Burge, S.K.; Kumar, K.A.; Wilson, J.M.; Ortiz, D.F. A time-motion study of primary care physicians' work in the electronic health record era. *Fam. Med.* **2018**, *50*, 91–99. [[CrossRef](#)]
15. Farzandipour, M.; Meidani, Z.; Riazi, H.; Sadeqi Jabali, M. Task-specific usability requirements of electronic medical records systems: Lessons learned from a national survey of end-users. *Inform. Health Soc. Care* **2017**, *13*, 1–20. [[CrossRef](#)]
16. Lee, S.; Kim, E.; Monsen, K.A. Public health nurse perceptions of Omaha System data visualization. *Int. J. Med. Inform.* **2015**, *84*, 826–834. [[CrossRef](#)]
17. Hirsch, J.S.; Tanenbaum, J.S.; Lipsky Gorman, S.; Liu, C.; Schmitz, E.; Hashorva, D.; Ervits, A.; Vawdrey, D.; Sturm, M.; Elhadad, N. Harvest, a longitudinal patient record summarizer. *J. Am. Med. Inform. Assoc.* **2015**, *22*, 263–274. [[CrossRef](#)]
18. Pivovarov, R.; Elhadad, N. Automated methods for the summarization of electronic health records. *J. Am. Med. Inform. Assoc.* **2015**, *22*, 938–947. [[CrossRef](#)]
19. Zhang, R.; Pakhomov, S.; Melton, G.B. Automated identification of relevant new information in clinical narrative. In Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, Miami, FL, USA, 28–30 January 2012; Association for Computing Machinery: New York, NY, USA, 2012; pp. 837–842.
20. Farri, O.; Rahman, A.; Monsen, K.A.; Zhang, R.; Pakhomov, S.V.; Pieczkiewicz, D.S.; Speedie, S.M.; Melton, G.B. Impact of a prototype visualization tool for new information in EHR clinical documents. *Appl. Clin. Inform.* **2012**, *3*, 404–418. [[CrossRef](#)]
21. Liu, X.Y.; Croft, W.B. Statistical language modeling for information retrieval. *Annu. Rev. Inf. Sci. Technol.* **2005**, *39*, 3–31.
22. Zhang, R.; Pakhomov, S.; Lee, J.T.; Melton, G.B. Navigating longitudinal clinical notes with an automated method for detecting new information. *Stud. Health Technol. Inform.* **2013**, *192*, 754–758.
23. Hart, S.G. NASA-task load index (NASA-TLX); 20 years later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908. [[CrossRef](#)]
24. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv. Psychol.* **1988**, *52*, 139–183.
25. Belden, J.L.; Koopman, R.J.; Patil, S.J.; Lowrance, N.J.; Petroski, G.F.; Smith, J.B. Dynamic electronic health record note prototype: Seeing more by showing less. *J. Am. Board. Fam. Med.* **2017**, *30*, 691–700. [[CrossRef](#)]
26. Hosseini, M.; Faiola, A.; Jones, J.; Vreeman, D.J.; Wu, H.; Dixon, B.E. Impact of document consolidation on healthcare providers' perceived workload and information reconciliation tasks: A mixed methods study. *J. Am. Med. Inform. Assoc.* **2019**, *26*, 134–142. [[CrossRef](#)]
27. Wrenn, J.O.; Stein, D.M.; Bakken, S.; Stetson, P.D. Quantifying clinical narrative redundancy in an electronic health record. *J. Am. Med. Inform. Assoc.* **2010**, *17*, 49–53. [[CrossRef](#)]

28. Wu, J.H.; Yuan, Y.F. Improving searching and reading performance: The effect of highlighting and text color coding. *Inf. Manag.* **2003**, *40*, 617–637. [[CrossRef](#)]
29. Schneider, W.; Shiffrin, R.M. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* **1977**, *84*, 1–66. [[CrossRef](#)]
30. Laxmisan, A.; Hakimzada, F.; Sayan, O.R.; Green, R.A.; Zhang, J.; Patel, V.L. The multitasking clinician: Decision-making and cognitive demand during and after team handoffs in emergency care. *Int. J. Med. Inform.* **2007**, *76*, 801–811. [[CrossRef](#)]
31. Walsh, S.H. The clinician's perspective on electronic health records and how they can affect patient care. *BMJ* **2004**, *328*, 1184–1187. [[CrossRef](#)]
32. Nygren, E.; Lind, M.; Johnson, M.; Sandblad, B. The art of the obvious. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Monterey, CA, USA, 3–7 May 1992; Association for Computing Machinery: New York, NY, USA, 1992; pp. 235–239.
33. Coiera, E. Technology, cognition and error. *BMJ Qual. Saf.* **2015**, *24*, 417–422. [[CrossRef](#)]
34. Palojoki, S.; Mäkelä, M.; Lehtonen, L.; Saranto, K. An analysis of electronic health record-related patient safety incidents. *Health Inform. J.* **2017**, *23*, 134–145. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).