Health Document Presentation in Patient-Centered Recommender Systems with Carousel Interfaces

Behnam Rahdari[1], Peter Brusilovsky[1], Daqing He[1], Khushboo Thaker[1], Mohammad Hassany[1], Youjia Wang[2], Young Ji Lee[2] and Heidi Donovan[2]

[1]*School of Computing and Information, University of Pittsburgh, Pittsburgh, USA*
[2]*School of Nursing, University of Pittsburgh, Pittsburgh, USA*

Abstract

Despite the increasing availability of health information, many users still find it difficult to navigate and comprehend this content effectively. Addressing these challenges requires innovative approaches, including personalized recommendations and more efficient methods of information delivery. In this paper, we explore the use of generative AI to improve access to health article recommendations within a carousel-based interface, utilizing our system, HELPeR. Our focus is on both generating and evaluating these summaries through a three-stage online experiment with domain experts. The results reveal the potential and complexities of employing generative AI for summarizing recommended health articles for ovarian cancer patients and their caregivers.

Keywords

Health Recommender Systems, Generative AI, Personalized Health Information, User Interface Design

1. Introduction

The increased availability of health information online has transformed the way patients and caregivers access knowledge about diseases, treatments, and health management. However, navigating this vast amount of information can be overwhelming, especially for non-experts. The complexity of medical terminology, coupled with the sheer volume of available content, often leaves users struggling to find and understand the information most relevant to their needs. This challenge underscores the need for more intuitive methods of information delivery that cater to varying levels of information needs and health literacy.

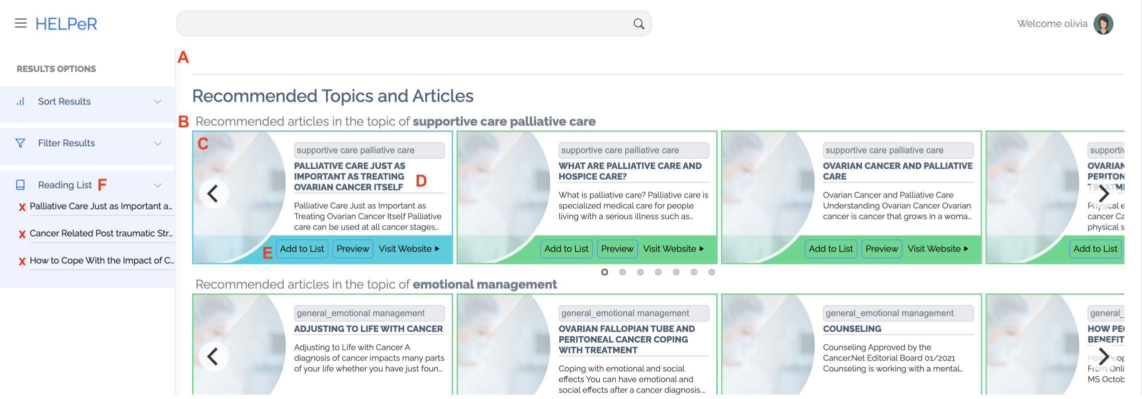
In response to these challenges, personalized recommendations and advanced methods of information delivery have emerged as key strategies to bridge the gap between users and the information they seek. These approaches aim to adapt the content to individual users, considering factors such as their treatment history, disease trajectory, and cognitive abilities. By offering personalized recommendations, these systems can guide users through relevant information more efficiently, potentially improving their understanding and engagement with health content.

Despite the progress made in this area, important questions remain unanswered. One of the key issues is understanding how users decide whether to explore a recommendation further, especially when only minimal information is provided. While visual cues such as images are effective in capturing users' attention in domains like entertainment and e-commerce, the health domain primarily relies on textual content, with images often serving a decorative role in most online health articles. This raises the question of how to present health information in a way that encourages deeper exploration without the visual allure that is typical in other fields.

This challenge became particularly evident in our work with the HELPeR system [1] depicted in Figure 1, which used a carousel-based recommendation interface to present relevant health documents to cancer patients. The use of a carousel-based interface was important in our recommendation context, since the system has to data to decide which information need brought the user to HELPeR for each particular session. While the system maintains a fine-grain model of user interests and knowledge, the

HealthRecSys'24: The 6th Workshop on Health Recommender Systems co-located with ACM RecSys 2024

ber58@pitt.edu (B. Rahdari); peterb@pitt.edu (P. Brusilovsky); dah44@pitt.edu (D. He); k.thaker@pitt.edu (K. Thaker); moh70@pitt.edu (M. Hassany); yow14@pitt.edu (Y. Wang); leeyoung@pitt.edu (Y. J. Lee); donovanh@pitt.edu (H. Donovan)



Recommended Topics and Articles

Recommended articles in the topic of **supportive care palliative care**

supportive care palliative care
PALLIATIVE CARE JUST AS IMPORTANT AS TREATING OVARIAN CANCER ITSELF
Palliative Care Just as Important as Treating Ovarian Cancer Itself Palliative care can be used at all cancer stages...

supportive care palliative care
WHAT ARE PALLIATIVE CARE AND HOSPICE CARE?
What is palliative care? Palliative care is specialized medical care for people living with a serious illness such as...

supportive care palliative care
OVARIAN CANCER AND PALLIATIVE CARE
Ovarian Cancer and Palliative Care Understanding Ovarian Cancer Ovarian cancer is cancer that grows in a woma...

supporti
OVARIAN PERITON TREATMI
Physical e cancer Ca physical s

Recommended articles in the topic of **emotional management**

general_emotional management
ADJUSTING TO LIFE WITH CANCER
Adjusting to Life with Cancer A diagnosis of cancer impacts many parts of your life whether you have just foun...

general_emotional management
OVARIAN FALLOPIAN TUBE AND PERITONEAL CANCER COPING WITH TREATMENT
Coping with emotional and social effects You can have emotional and social effects after a cancer diagnosis...

general_emotional management
COUNSELING
Counseling Approved by the Cancer.Net Editorial Board 01/2021 Counseling is working with a mental...

general_
HOW PE BENEFIT
How Peop From Onli MS Octob

Reading List
X Palliative Care Just as Important a...
X Cancer Related Post traumatic Str...
X How to Cope With the Impact of C...

Figure 1: The HELPeR system interface design includes the following main components: A: Recommended Articles, B: A carousel containing recommended articles within the same topic, C: A specific recommended article, D: Details of the selected item, E: Options for user interaction with the recommendation, and F: Reading list (Personal Library)

majority of the users maintain their interests in several health topics, and choosing user chief concern reliably at each session is as hard as guessing which movie genre a user would like to watch on Netflix at each login. Just like Netflix uses a carousel-based interface to support human-AI collaboration and allow the user to choose the genre she preferred today, HELPeR's carousel-based interface allows the user to choose the priority topics from several most likely topics shown by the carousels.

The carousel format is familiar and effective in presenting multiple options in parallel, but in the context of health information, it presented unique difficulties. While carousels are useful for organizing and displaying visual content (such as movie posters or book covers), it is a challenge to choose textual content to present on a carousel card to ensure that this information is sufficient for the user to make an informed decision. This challenge was compounded by the fact that the recommended content was often not available in a cohesive, summarized format. The headers of health documents are usually too long to be readable or even fit on a card (Figure 1). Document summaries are either too complex or not available at all. To address this challenge, we turned to generative AI to create concise document overviews and summaries that could effectively fit within the carousel cards, helping users make informed decisions with the limited available space.

In this paper, we present our attempts to use Large Language Models (LLM) to generate brief but informative overviews and summaries of health articles related to ovarian cancer, which can fit to the cards of HELPeR's carousel-based recommendation interface. To assess the relevance, clarity, and informativeness of these AI-generated overviews and summaries, we performed a three-phase expert evaluation study. Our findings highlight the challenges of generating concise yet informative content that meets the needs of diverse users, offering valuable insights into the future of AI-driven health information delivery.

Our study contributes to this evolving landscape by evaluating the use of AI-generated summaries within a patient-centered health recommender system, specifically designed for ovarian cancer patients. Although previous research has established the potential of AI in personalizing health information, our focus is on understanding how these AI-generated summaries can be integrated into carousel interfaces to enhance user engagement and decision-making. Through a three-phase expert evaluation, we aim to provide insights that will inform the design of future recommender systems, making them more responsive to the diverse and changing needs of patients.

The paper is organized as follows. In the next section, we provide a review of related work in personalized health information systems. The methodology section explains the approach to generating summaries and a balanced document selection approach that we applied to minimize bias in our evaluation process. We then describe the evaluation study and present our findings. Finally, we discuss the implications of our results, outline the limitations of our study, and suggest directions for future work.

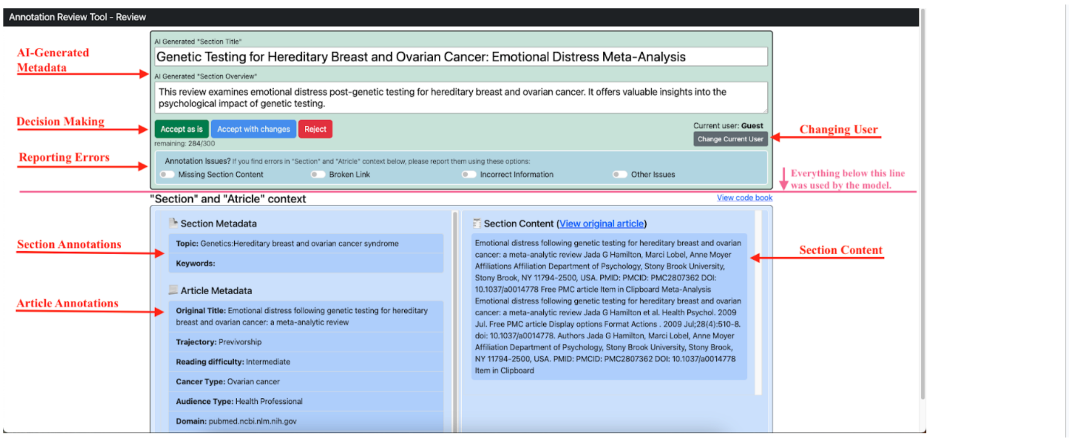
2. Related Work

The increasing availability of health information online has greatly influenced how patients and caregivers manage care and make informed decisions. Early recommender systems in this domain focused on providing personalized health information by tailoring content to often static user profiles. These systems reported in [2, 3] were important in demonstrating the benefits of personalized information delivery. However, as the complexity and volume of health information increased, so did the need for systems that could adapt to the evolving needs of users more effectively.

In response to these needs, the field has seen a shift towards more interactive and dynamic recommender systems. These systems known as conversational and critique-based recommenders [4, 5] allow users to actively engage with the recommendations, refining the information needs based on real-time feedback. Although these interactive models enhanced personalization, they also introduced new challenges, such as the cognitive load associated with continuous interaction, which could be particularly burdensome for users with lower health literacy.

To address these challenges, recent research explored visual interfaces for recommender systems [6], which offered a higher expressive power and a better opportunity to add transparency [7] and user control [8] to the recommendation process. Among other visual recommender interfaces, carousel-based interfaces became especially popular for their ability to display multiple pieces of content in a compact and navigable format [9, 10] and offer users a simple control over the recommendation process. These interfaces combined power and simplicity enabling users to quickly scan through recommendations and make informed decisions without feeling overwhelmed.

Despite these advancements, the challenge of effectively delivering personalized health information remains unsolved. As noted by Chi et al. [11] and Thaker et al. [12], users' information needs are not static; they could change freqiently following the change in health status, the progress of treatment and personal circumstances. This variability requires a more dynamic approach to content presentation, one that can adapt to the immediate context of the user. The integration of generative AI into health recommender systems, as explored in our work, offers a promising solution by creating concise, contextually relevant summaries that fit within the limited space of carousel interfaces [1].



AI Generated "Section Title"

Genetic Testing for Hereditary Breast and Ovarian Cancer: Emotional Distress Meta-Analysis

AI Generated "Section Overview"

This review examines emotional distress post-genetic testing for hereditary breast and ovarian cancer. It offers valuable insights into the psychological impact of genetic testing.

remaining: 284/300

Current user: **Guest**

Annotation Issues? If you find errors in "Section" and "Atricle" context below, please report them using these options:

○ Missing Section Content ○ Broken Link ○ Incorrect Information ○ Other Issues

"Section" and "Atricle" context

Section Metadata

Topic: Genetics:Hereditary breast and ovarian cancer syndrome

Keywords:

Article Metadata

Original Title: Emotional distress following genetic testing for hereditary breast and ovarian cancer: a meta-analytic review

Trajectory: Previvorship

Reading difficulty: Intermediate

Cancer Type: Ovarian cancer

Audience Type: Health Professional

Domain: pubmed.ncbi.nlm.nih.gov

Section Content (View original article)

Emotional distress following genetic testing for hereditary breast and ovarian cancer: a meta-analytic review Jada G Hamilton, Marci Lobel, Anne Moyer Affiliations Affiliation Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, USA. PMID: PMCID: PMC2807362 DOI: 10.1037/a0014778 Free PMC article Item in Clipboard Meta-Analysis Emotional distress following genetic testing for hereditary breast and ovarian cancer: a meta-analytic review Jada G Hamilton et al. Health Psychol. 2009 Jul. Free PMC article Display options Format Actions . 2009 Jul;28(4):510-8. doi: 10.1037/a0014778. Authors Jada G Hamilton, Marci Lobel, Anne Moyer Affiliation Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, USA. PMID: PMCID: PMC2807362 DOI: 10.1037/a0014778 Item in Clipboard

Figure 2: Evaluation Interface for AI-Generated Metadata

3. Method

3.1. Generating the title and overview

The HELPeR system [1, 13], is an interactive recommender system designed for ovarian cancer patients and their caregivers. It uses a knowledge base built from a curated collection of documents, which includes public health information, research articles, clinical trial results, and other relevant resources.

These documents undergo a rigorous curation process that involves sectioning, topic modeling, and key-phrase extraction to ensure that the information is reliable, up-to-date, and relevant to the needs of the users. For each document, HELPeR's knowledge base includes a range of textual and metadata information, such as the article title, topics, difficulty, relevant key-phrase, and full text of each document section.

As mentioned above, neither document titles nor section content were immediately suitable to be displayed on carousel card. The titles were frequently long and confusing and summaries either very long on not available. To address this problem, the most recent version of our system explored the use of LLM [1] to generate a document representation (title and summary) that can fit on a carousel card providing concise information about the document behind the card.

The exact prompt used in our study to generate a title and summary for each document is shown below. This prompt was selected through a prompt-tuning process to maximize clarity and relevance within the constraints of a limited format. Note that we passed all information about each document to LLM as part of the prompt:

```
prompt = f""" Given the following section of an
article titled '{article_title}', with the
topic of {topic}, covering keywords: {keywords}
and containing this text: "{section_text}"

1- Generate a 4-7 word title reflecting the
section's essence and aligning with the
article's theme.

2- Write a 20-25 words, two-sentence summary
capturing key points, serving as an informative
overview of the article for readers.

Respond ONLY and PRECISELY in this format:
[{{ "title": "the generated title" }},
{{"summary": "the generated summary" }}]"""
```

3.2. Selecting a Diverse Subset of Annotations

Given the large number of documents in our collection related to ovarian cancer, selecting a manageable yet representative subset for our human-centered evaluation was essential. A subset that accurately represents the diversity of the full dataset is crucial to avoid biases that could skew our results, particularly in our collection, that certain topics or audience types are over-represented. To address this, we used a modified version of the Maximal Marginal Relevance (MMR) algorithm [14], traditionally used to balance relevance and novelty in information retrieval tasks.

The primary challenge in selecting this subset was to ensure that it not only reflected the diversity of topics, domains, and audience types present in the dataset, but also maintained the overall distribution of the documents and topics in our collection. We adapted the MMR algorithm to emphasize the intrinsic properties of the document, allowing it to select a subset that balanced these two critical factors.

Given a set of all documents D, where $n = |D|$ denotes the total number of documents, we represented the document set as a matrix X, with each document encoded as a one-hot vector across m categories. The subset S, initially empty, was populated by the execution of the MMR algorithm. The key parameter λ, which ranged from 0 to 1, modulated the trade-off between relevance and diversity.

The process began with the random selection of an initial document d_q from D, which served as a pseudo-query, establishing the initial subset $S = \{d_q\}$. Then each document d_i within D was assigned a relevance score, $sim(d_i)$, assumed for simplicity to follow a uniform distribution $sim(d_i) \sim U(0, 1)$.

[1] OpenAI APIs - gpt-3.5-turbo-1106

During each iteration, the algorithm evaluated each candidate document d_c not yet included in S. The diversity score $div(d_c, S)$ was calculated based on the average cosine dissimilarity between d_c and all documents currently in S:

$$div(d_c, S) = 1 - \frac{1}{|S|} \sum_{d_s \in S} \text{cosine\_similarity}(\mathbf{x}_{d_c}, \mathbf{x}_{d_s})$$

The MMR score for each candidate document was then calculated by combining the relevance and diversity scores:

$$MMR(d_c) = \lambda \cdot sim(d_c) + (1 - \lambda) \cdot div(d_c, S)$$

The document d_c with the highest MMR score was added to S, thereby progressively refining the document subset to be both *relevant* and *diverse*.

Our modified approach departs from the traditional reliance on a fixed query in the MMR algorithm by dynamically calculating diversity relative to the evolving subset S. This modification is particularly advantageous in environments where queries are undefined or fluid, such as unsupervised document clustering or information retrieval systems where query independence is crucial. By structuring the selection process around these principles, we ensured that the resulting subset was both representative of the dataset's diversity and suitable for our human-centered evaluation of the AI-generated summaries.

As illustrated in Figure 3, the subset selected using the Maximal Marginal Relevance (MMR) algorithm shows a more balanced distribution of features across various categories, such as article knowledge level, audience type, and domain, compared to a randomly selected subset. The MMR-based selection achieves a broader coverage of diverse topics, ensuring that underrepresented areas are included, thus mitigating the biases that are evident in the random sample where certain categories, such as "Survivorship" and "Patient and Caregiver," dominate.

4. Experiment

4.1. Experimental Design

To evaluate the quality and consistency of AI-generated *titles and summaries* for documents in HELPeR knowledge base we designed a multi-phase study. We alternate the articles and their order in which they are presented in each phase. This study design allows us to gain insights about both the quality of AI-generated content and the evaluation process itself.

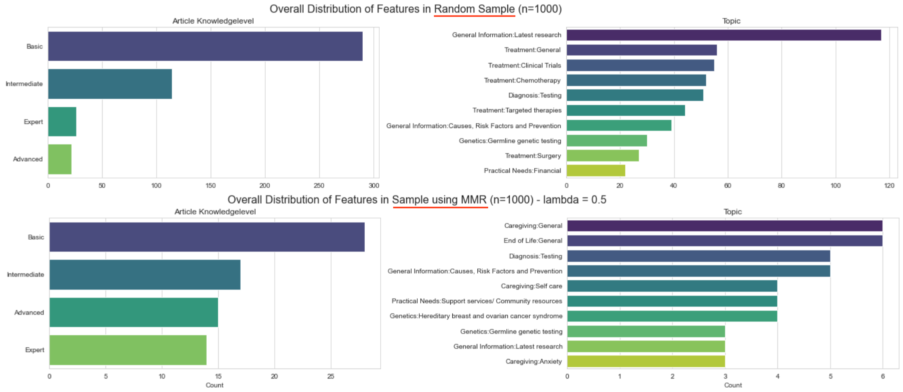


Figure 3: Comparison of Overall Distribution of Features in the Random Sample (top) vs. MMR-Selected Sample (bottom) - $n = 1000$, $\lambda = 0.5$ - Color intensity corresponds to feature frequency.

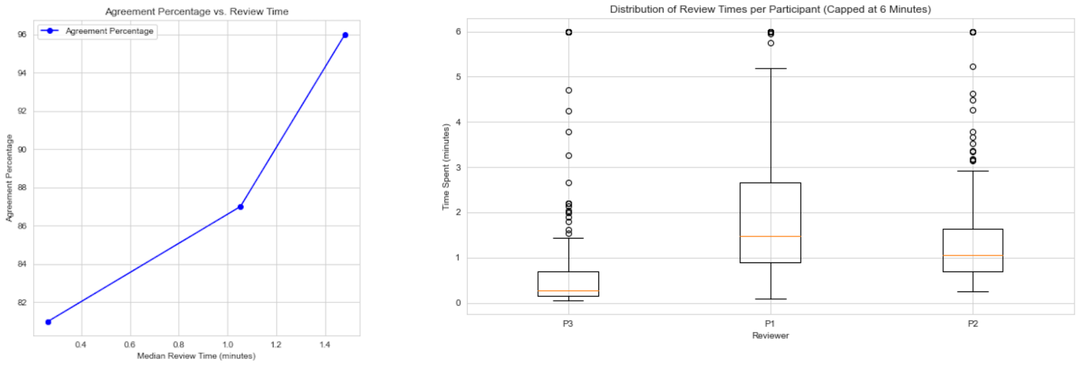


Figure 4: Left: Agreement percentage goes up with more time spend on reviewing. Right: Time spent by each reviewer.

4.2. Participants

The study involved two groups of participants: domain experts and nursing students with specialized knowledge in ovarian cancer. In the first phase, we have 100 unique documents to each of three domain experts in our research team. In the next phases, three nursing students (not belonging to the research team) were recruited to review the same sample in two rounds. Each student reviewed 300 instances, with 100 documents that overlapped between the two rounds. The use of nursing students in addition to domain experts was important to better represent the prospect of caregivers and reduce evaluation bias.

4.3. Procedure

4.3.1. Phase 1

In the first phase, AI-generated titles and overviews were created for a sub-sample of 300 documents selected by the modified MMR algorithm. Then these summaries were presented to three domain experts via a custom web interface (Figure 2). Each expert was tasked with evaluating 100 documents. To our surprise, only 55% of the summaries were accepted without modification, a number that raises many questions about the abilities of Generative-AI in summarizing health articles.

4.3.2. Phase 2

Following the insights gained from the first phase, we conducted unstructured interviews with domain experts on our team to understand the reason behind their high rejection rate. As it turns out, the reason for rejecting in many cases was related to a problem with missing or inaccurate annotations and uncertainty about the decision-making criteria. To address these issues, several enhancements were made in the second phase. A detailed code-book was developed to provide clear guidelines and decision-making criteria for participants, ensuring consistency across evaluations. Additionally, the web interface was improved to include an interactive tutorial and more user-friendly features including the ability to report inconsistent annotations and other issues that are not directly related to the quality of AI-Generated contents. In this phase, three external nursing students, each with specialized knowledge in ovarian cancer, were recruited to evaluate the same 300 documents. Each student reviewed 100 documents in this phase.

4.3.3. Phase 3

In the final phase, the consistency of the evaluations was tested. The external nursing students were asked to review an additional set of 200 documents. This set included 100 new documents and 100 documents that they had reviewed in the previous round. These documents were randomly distributed.

This phase aimed to measure the stability and reliability of their evaluations over time. At the end of this phase, each document had been reviewed by multiple experts, providing a comprehensive dataset for our analysis.

5. Results

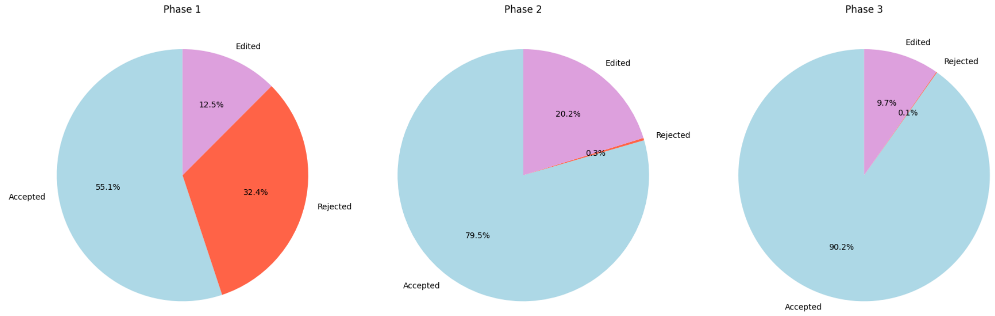


Figure 5: Document quality ratings before and after improvements (left to right)

As explained above, the initial evaluation phase, conducted by domain experts, indicated that only 55% of the AI-generated summaries were accepted without modifications. This lower-than-expected acceptance rate highlighted several issues in the evaluation process, which we attempted to address by introducing several enhancements reviewed above. Following these enhancements, the second phase of our study, in which nursing students with specialized knowledge in ovarian cancer were recruited as evaluators, showed a significant change, with 98.7% of the content being accepted either outright or with minor edits. Specifically, 78.7% of the summaries were accepted as they were, while 20% required slight modifications. Only a small fraction, 1.3%, of the content was rejected. These results demonstrated that the enhancements implemented after the first phase, such as the code-book and UI improvements, had a substantial positive impact on the perceived quality of the AI-generated summaries. Notably, our results suggest that the main factor contributing to this improvement was the added ability for users to report issues with the generated content separately from judging its overall quality. (Figure 5)

The final phase of the study focused on assessing the quality of evaluation itself by analyzing the consistency between our raters. The results from this ,depicted in Table 1,phase showed an 88% agreement in the reviews, indicating the reliability of the evaluation process. However, Cohen's Kappa scores of 0.37 suggested moderate to low consistency among reviewers, pointing to individual differences in judgment and the influence of subjective factors on the evaluation process. We are aware that our small sample size could have major influence of this results. Additionally, pairwise agreement percentages among reviewers varied, with the highest consistency observed between reviewers P2 and P3 at 92%.

To better understand the individual differences between our reviewers, we analyzed the ratio between the duration of the review session and the consistency with peers. Our analysis, illustrated in Figure 3-left, revealed a positive correlation between longer review times and higher consistency in evaluations as can be seen in Figure 4. This suggests that a more deliberate review process leads to more stable and reliable judgments. Reviewers who spent less time on evaluations exhibited lower consistency, emphasizing the importance of thoroughness in assessing AI-generated content.

Overall, the results of this study, although preliminary in nature, underscore the challenges inherent in evaluating AI-generated summaries, particularly in a complex domain like ovarian cancer information. The improvements made between phases significantly improved reviewer's judgment of acceptability of the AI outputs, yet individual differences among reviewers remain a critical factor in the consistency of evaluations. These findings highlight the need for clear guidelines and rigorous training to ensure reliable and consistent assessments in future AI evaluation tasks.

Table 1
Inter-Rating Consistency

Metric	**Value**
Total reviews analyzed	300
Unchanged reviews	264 (88.00%)
Changed reviews	36 (12.00%)
Changes frequency	
Edited → Accepted	32
Accepted → Edited	4
Inter-rater consistency	
Total sections	100
Some agreement	99%
No agreement	1% (actual AI mistake)
Full agreement	63 (63.00%)
Partial agreement	36 (36.00%)

6. Discussion, Limitations, and Future Work

The results of this study highlight both the potential and the challenges associated with using AI-generated summaries in the domain of ovarian cancer information. The improvements observed in the second phase, following the implementation of a comprehensive code-book [2] and user interface enhancements, underscore the importance of clear guidelines and a user-friendly evaluation environment. These adjustments significantly increased the acceptability of the AI outputs, demonstrating that careful attention to the evaluation process can substantially enhance the performance of AI systems in generating accurate and relevant content.

However, the study also revealed certain limitations. One of the primary concerns is the variability in reviewer judgments, as indicated by the moderate to low Cohen's Kappa scores. This suggests that individual differences in interpretation and the inherent subjectivity in content evaluation can impact the consistency of the results. Additionally, the correlation between review time and consistency points to the importance of thorough evaluations, but it also raises questions about the feasibility of scaling such processes for larger datasets.

Another limitation is related to the specific domain of ovarian cancer, which, while critical, may not encompass the full spectrum of challenges that AI-generated content could face in other medical or health-related domains. Therefore, while the findings of this study are valuable, they may not be fully generalizable to other areas without further validation.

For future work, several avenues can be explored. First, expanding the evaluation to include a more diverse set of medical topics could help to understand how AI performs in different domains. Furthermore, exploring automated methods to assess the consistency of evaluations, while keeping the human in the loop, could reduce the reliance on manual review processes, making the evaluation of AI-generated content more scalable.

Finally, addressing the subjective nature of content evaluation by developing more objective criteria or incorporating a larger pool of evaluators could improve the robustness of the evaluation process. Future studies could also examine the impact of these AI-generated summaries on end-users, such as patients and caregivers.

Acknowledgments

This work was supported by awards from the National Library of Medicine (NLM) of the National Institutes of Health (NIH) (Award Number: R01-LM013038).

[2] https://bit.ly/helper-codebook

References

[1] B. Rahdari, P. Brusilovsky, D. He, K. M. Thaker, Z. Luo, Y. J. Lee, Helper: An interactive recommender system for ovarian cancer patients and caregivers, in: Proceedings of the 16th ACM Conference on Recommender Systems, RecSys '22, Association for Computing Machinery, New York, NY, USA, 2022, p. 644–647.

[2] K. Lee, K. Hoti, J. D. Hughes, L. M. Emmerton, Consumer use of "dr google": a survey on health information-seeking behaviors and navigational needs, Journal of medical Internet research 17 (2015) e4345.

[3] S. Kanthawala, A. Vermeesch, B. Given, J. Huh, Answers to health questions: Internet search results versus online health community responses, Journal of medical Internet research 18 (2016) e5369.

[4] B. Smyth, L. McGinty, J. Reilly, K. McCarthy, Compound critiques for conversational recommender systems, in: Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence, WI '04, IEEE Computer Society, USA, 2004, p. 145–151.

[5] L. Chen, P. Pu, Critiquing-based recommenders: survey and emerging trends, User Modeling and User-Adapted Interaction 22 (2012) 125–150.

[6] C. He, D. Parra, K. Verbert, Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities, Expert Systems with Applications 56 (2016) 9–27.

[7] K. Verbert, D. Parra-Santander, P. Brusilovsky, E. Duval, Visualizing recommendations to support exploration, transparency and controllability, in: the 2013 International Conference on Intelligent User Interfaces, IUI '2013, ACM Press, 2013, pp. 351–362.

[8] D. Jannach, S. Naveed, M. Jugovac, User control in recommender systems: Overview and interaction challenges, in: E-Commerce and Web Technologies: 17th International Conference, EC-Web 2016, Porto, Portugal, September 5-8, 2016, Revised Selected Papers 17, Springer, 2017, pp. 21–33.

[9] B. Rahdari, B. Kveton, P. Brusilovsky, The magic of carousels: Single vs. multi-list recommender systems, in: Proceedings of the 33rd ACM Conference on Hypertext and Social Media, 2022, pp. 166–174.

[10] N. Felicioni, M. Ferrari Dacrema, P. Cremonesi, A methodology for the offline evaluation of recommender systems in a user interface with multiple carousels, in: Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, 2021, pp. 10–15.

[11] Y. Chi, D. He, W. Jeng, Laypeople's source selection in online health information-seeking process, Journal of the Association for Information Science and Technology 71 (2020) 1484–1499.

[12] K. Thaker, Y. Chi, S. Birkhoff, D. He, H. Donovan, L. Rosenblum, P. Brusilovsky, V. Hui, Y. J. Lee, Exploring resource-sharing behaviors for finding relevant health resources: analysis of an online ovarian cancer community, JMIR cancer 8 (2022) e33110.

[13] K. Thaker, B. Rahadari, V. Hui, Z. Luo, Y. Wang, P. Brusilovsky, D. He, H. Donovan, Y. J. Lee, Helper: Interface design decision and evaluation, in: Innovation in Applied Nursing Informatics, IOS Press, 2024, pp. 750–751.

[14] J. Carbonell, J. Goldstein, The use of mmr, diversity-based reranking for reordering documents and producing summaries, in: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, 1998, pp. 335–336.