BioMedInformatics

Article

Evaluation of Replies to Voice Queries in Gynecologic Oncology by Virtual Assistants Siri, Alexa, Google, and Cortana

Jamie M. Land [1], Edward J. Pavlik [1,*], Elizabeth Ueland [2,†], Sara Ueland [3,†], Nicholas Per [1], Kristen Quick [1], Justin W. Gorski [1], McKayla J. Riggs [1], Megan L. Hutchcraft [1], Josie D. Llanora [1] and Do Hyun Yun [1]

[1] Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Chandler Medical Center-Markey Cancer Center, University of Kentucky College of Medicine, Lexington, KY 40536-0293, USA; jamie.land@uky.edu (J.M.L.)
[2] Scripps College, Claremont, CA 91711, USA
[3] Claremont McKenna College, Claremont, CA 91711, USA
* Correspondence: edward.pavlik@uky.edu; Tel.: +1-(859)-323-3830
† These authors contributed equally to this work.

Abstract: Women that receive news that they have a malignancy of gynecologic origin can have questions about their diagnosis. These questions might be posed as voice queries to the virtual assistants Siri, Alexa, Google, and Cortana. Because our world has increasingly adopted smart phones and standalone voice query devices, this study focused on the accuracy of audible replies by the virtual assistants (VAs) Siri, Alexa, Google, and Cortana to voice queries related to gynecologic oncology. Twenty-one evaluators analyzed VA audible answers to select voice queries related to gynecologic oncology. Questions were posed in three different ways for each voice query in order to maximize the likelihood of acceptability to the VAs in a 24-question panel. For general queries that were not related to gynecologic oncology, Google provided the most correct audible replies (83.3% correct), followed by Alexa (66.7% correct), Siri (45.8% correct), and Cortana (20.8% correct). For gynecologic oncology-related queries, the accuracy of the VAs was considerably lower: Google provided the most correct audible replies (18.1%), followed by Alexa (6.5%), Siri (5.5%), and Cortana (2.3%). There was a considerable drop in the accuracy of audible replies to oral queries on topics in gynecologic oncology relative to general queries that were not related to gynecologic oncology. There is considerable room for improvement in VA performance, so that caution is advised when using VAs for medical queries in gynecologic oncology. Our specific findings related to gynecologic oncology extend the work of others with regard to the low usability of general medical information obtained from VAs, so that reliance on conversational assistants for actionable medical information represents a safety risk for patients and consumers.

Keywords: virtual assistants; gynecologic; oncology; Siri; Alexa; Google; Cortana; validity; accuracy

Citation: Land, J.M.; Pavlik, E.J.; Ueland, E.; Ueland, S.; Per, N.; Quick, K.; Gorski, J.W.; Riggs, M.J.; Hutchcraft, M.L.; Llanora, J.D.; et al. Evaluation of Replies to Voice Queries in Gynecologic Oncology by Virtual Assistants Siri, Alexa, Google, and Cortana. *BioMedInformatics* **2023**, *3*, 553–562. https://doi.org/10.3390/biomedinformatics3030038

Academic Editor: Shuncong Wang

Received: 4 May 2023
Revised: 26 June 2023
Accepted: 3 July 2023
Published: 11 July 2023

1. Introduction

The Internet can be used by individuals to obtain health information [1]. Voice technology allows Internet searches for health information through verbal queries, which are answered audibly by a virtual assistant. Virtual assistants (VAs), such as Siri (Apple), Alexa (Amazon), Cortana (Microsoft), and the Google Assistant, are used ubiquitously. Google Assistant deploys searches on Google, as does Siri, while Alexa and Cortana use Bing as their search engine [2,3]. In 2020, Google Assistant was available on >1 billion devices and was used by >500 million users monthly [4], providing 27% of all global web searches. More than 500 million Apple customers use Siri as a virtual assistant [5]. The Amazon Echo Home Speaker, which uses Alexa, has >40 million users in the United States [6]. Siri, Alexa, Google, and Cortana are among the top 10 best VAs, hence we have chosen to investigate these VAs regarding the accuracy of healthcare information that they provide [7]. Inaccurate healthcare information can be responsible for misunderstandings by patients that can lead them to refuse or withdraw from treatments that otherwise may have proven beneficial. In this paper, we seek to identify risks for gynecologic oncology patients that can originate from inaccuracies in VA responses.

Since the introduction of Siri as a feature of the iPhone 4S in 2011, VAs have become mainstream. Siri can make phone calls or send text messages when users cannot manually enter information while driving, biking, or walking [8]. VAs have utility for the elderly in situations that are simplified by speaking, especially when there is physical impairment [9]. There have been steady declines in the personal computer market for the last eight years as people opt to go online with smartphones, where voice queries are easier than keying in searches [10,11]. A surge in the use of VAs is underway in healthcare. VA usage increased to 21.0% of U.S. adults in 2021, with 54.4 million people using VAs for questions in healthcare about symptoms, medical information, and treatments [12]. VA responses to voice queries have been utilized for information on postpartum depression [12], healthy lifestyles [13], vaccinations [14], addiction [15], mental health, and interpersonal violence [16]. Several U.S. hospitals have smart speakers installed in patients' rooms, enabling them to make requests that can relieve the clinical staff [17]. Widespread access and utilization can be contrasted by the quality of the results received from VAs, especially in extending simple everyday voice queries to more complex questions related to gynecologic oncology. In the absence of literature for VA queries in oncology or gynecologic oncology, we show in this paper, for the first time, the degree to which voice queries related to gynecologic oncology are accurately addressed by audible replies from Siri, Alexa, Google, and Cortana.

2. Materials and Methods

Power was calculated between the VAs using McNemar's test for comparing the percent correct between two sources of information. McNemar's test was used because it provides the best fit for our data set in which the data were paired and proportions were different. To calculate power, VA performance was estimated based on general queries made by two investigators, as shown in the "Percent correct responses" line of Table 1 (bottom line). In this setting, Google was correct (83.3%) more often than Siri (45.8%: p value = 0.27, power = 0.90) and Cortana (20.8%: p value = 0.0003, power = 0.99), but not Alexa, and Alexa (66.7%) was correct more often than Cortana (20.8%: p value = 0.0023, power = 0.98). By enlarging the data set to additional rounds of 24 queries, assuming the results obtained in Table 1 would continue to hold, then to detect a difference between Google and Alexa, 90 queries would attain 90% power, and this would hold for Alexa vs. Siri, as well as Siri vs. Cortana. To approach this power, 21 evaluators each presented a set of 24 questions specific to gynecologic oncology in querying the VAs. The general questions used for estimating power were selected from QUIZ Daily (1550 Larimer Street, Suite 431, Denver, CO 80202, U.S.A.) and checked against Google and Bing for matching correct answers. Our baseline general question evaluation of VA performance was within 1–3% of the performance reported by others, except for Cortana, for which our estimate was lower (20.8% vs. 45%) [18,19]. This difference does not affect our power calculation.

Evaluators were chosen at random to submit queries in English. Evaluators' speech characteristics varied, but we do not believe this influenced the data that was collected, because each VA is equipped to assess over 100 languages and dialects. Moreover, preliminary use of a VA includes practice voice recognition steps that adjust for pronunciations and accents. Each evaluator accessed the most recent updated versions of Siri (iOS 14), Alexa (version 2.2.427375.0), and Google (version 1.9.28702) on their smartphones. To access Cortana (version 3.2106.14307.0), a Dell Latitude E6430 Windows 10 laptop was provided to each evaluator that did not have access to Windows 10, while evaluators who already had access completed the test on their own device. The queries that were settled on were deemed appropriate by a team of gynecologic oncologists with regard to their expectation that patients fail to know answers to these questions, which could have significant impacts on their health and well-being. The panel of 24 questions related to gynecologic oncology were posed in three different templates to all of the VAs by 21 evaluators. In order to maximize the chances of a correct answer by each VA, and to assess which query presentation could be validated, a 3-tier template was used for each inquiry: "X?" (A), "What is X?" (B), and "Define X?" (C). The questions asked to each VA are listed in Table 2. We weighted the 24-question panel toward ovarian cancer queries (16 questions) with the premise that the lower incidence of ovarian cancer might provide a more robust test of the VAs than more common gynecologic malignancies. In addition, six queries related to more prevalent cervical cancers were included. The remaining two questions were aimed at borderline epithelial tumors of the ovary and endometrial cancer.

Table 1. Responses to General Questions Presented to Virtual Assistants. General questions that were verbally asked to the VAs are identified. Chi-square and Fisher's exact probability tests showed that Siri and Cortana underperformed Google ($p < 0.05$), and that Cortana underperformed Alexa ($p < 0.05$). The performance of Google and Alexa was not significantly different.

Question: YES = Correct Answer, NO = No Audible Answer or Incorrect	Google	Alexa	Siri	Cortana
What is the current temperature?	YES	YES	YES	NO
What is the weather forecast?	YES	YES	YES	YES
What is the weather tomorrow?	YES	YES	YES	YES
What is USD 5 in Euros?	YES	YES	YES	YES
What is a flock of flamingos known as?	YES	YES	NO	NO
Which actress has the most Oscar nominations without a win?	YES	YES	NO	NO
What is the world's largest body of water?	YES	YES	YES	NO
What river feeds the Dead Sea?	YES	YES	YES	NO
What is the minimum age for a U.S. President?	YES	YES	YES	NO
What popular vegetable is poisonous if eaten raw?	NO	NO	NO	NO
Which cathedral has more statues than any other in the world?	YES	NO	NO	NO
What President appointed a former President to the Supreme Court?	YES	YES	NO	NO
Who was awarded the highest-ever rank in the U.S. armed forces?	NO	YES	NO	NO
What is the smallest planet in our solar system?	YES	YES	YES	NO
Who was the drummer for the Beatles?	YES	YES	YES	NO
Which state produces nearly half of all U.S. mushrooms?	YES	YES	NO	NO
Which President owned a haberdashery?	YES	NO	NO	NO
What is the name of the land bridge that once connected Asia and North America?	YES	NO	YES	NO
What is the only state with an official jelly?	YES	NO	NO	NO
Which artist is best known for their work in Cubism?	YES	YES	NO	NO
In which state is it illegal to serve margarine instead of butter in restaurants?	NO	NO	NO	YES
Which state's official dance is the polka?	NO	NO	NO	NO
Where is the longest fence in the world?	YES	YES	NO	YES
What does "Häagen-Dazs" mean?	YES	NO	YES	NO
Total number of correct responses	20	16	11	5
Percent correct responses	83.3%	66.7%	45.8%	20.8%

Table 2. Gynecologic Oncology Questions Asked to Virtual Assistants. This table shows the questions that were asked to each VA in the format "What is X?", where X is underlined. Two additional formats, "X?" and "Define X?" were also used for each query. Sources of correct answers from the Society of Gynecologic Oncology, American Cancer Society, American College of Gynecology, Centers for Disease Control, American Society of Clinical Oncology, the National Cancer Institute and UpToDate are hyperlinked to each query in the far-right column. Links were accessed on 6 July 2023.

Question #	Question	Correct Answer Link
1	What is stage I ovarian cancer?	Answer
2	What is stage II ovarian cancer?	Answer
3	What is stage III ovarian cancer?	Answer
4	What is stage IV ovarian cancer?	Answer
5	What is stage IC1 ovarian cancer?	Answer
6	What is stage IIIA1 ovarian cancer?	Answer
7	What is stage IVB ovarian cancer?	Answer
8	What are the subtypes of epithelial ovarian cancer?	Answer
9	What is screening for ovarian cancer?	Answer
10	What are the screening recommendations for ovarian cancer?	Answer
11	What are ways to prevent ovarian cancer?	Answer
12	What are the symptoms of ovarian cancer?	Answer
13	What is hereditary ovarian cancer?	Answer
14	What is ovarian cancer risk reduction?	Answer
15	What is screening for cervical cancer?	Answer
16	What are the screening recommendations for cervical cancer?	Answer
17	What are the options for a 20-year-old sexually active woman who requests a Pap smear?	Answer
18	What is the HPV vaccine?	Answer
19	What are the ages for HPV vaccination?	Answer
20	What are the three dose HPV vaccine recommendations?	Answer
21	What are borderline epithelial tumors of the ovary?	Answer
22	What is carcinosarcoma of the ovary?	Answer
23	What are high-grade serous tumors of the ovary?	Answer
24	What is stage IB endometrial cancer?	Answer

Individuals making queries were provided answers against which responses returned by VAs were graded. Correct answers were determined from the most recently updated consensus recommendations published by the Society of Gynecologic Oncology, American Cancer Society, American College of Gynecology, Centers for Disease Control, American Society of Clinical Oncology, the National Cancer Institute and UpToDate. Audible answers from the VAs were scored as incorrect = 0, does not understand or know or returns only web-links = 1, <40% correct = 2, 40–50% correct = 3, 50–85% correct = 4, 100% correct = 5. We calculated the overall intraclass correlation coefficient (ICC) using the results from a one-way ANOVA on Winstat (version 2012.1) and calculated individual coefficients of variation for each question, as well as results across VAs and templates. Finally, we quantified the responses across the VAs by aggregating the frequency of each score, and have reported the respective results in percentages. Graded scores were averaged across all graders, and expressed with the standard error of the mean (SEM). An average graded score ± SEM was calculated across individuals that evaluated responses by each VA. An average % score was determined as a percentage of the total possible score. Median scores and 75th percentiles were determined as a function of the total possible score. Minimum and maximum scores were determined as a function of the total possible score, and the difference between maximum and minimum as a function of the total possible score. Counts of total correct answers to VA queries are expressed as a percentage of total queries. Significant differences were determined as $p < 0.05$.

Chi-square and Fisher's exact probability tests were used for nonparametric analyses.

3. Results

Google provided the most correct audible replies with the general questions ($n = 20$; 83.3% correct), followed by Alexa ($n = 16$; 66.7% correct), Siri ($n = 11$; 45.8% correct), and Cortana ($n = 5$; 20.8% correct), Table 1. For queries related to gynecologic oncology, Google's average graded score (2.88 ± 0.04, Table 3, column B) was significantly higher ($p < 0.01$) than the average graded scores for Alexa, Siri, and Cortana (1.60 ± 0.04, 1.52 ± 0.04, 1.28 ± 0.03, Table 3, column B). The graded score for Cortana was statistically lower than the other VAs. When the average graded scores were expressed as a percent of the highest possible score, Google graded almost twice as high as the other VAs (57.6% ± 0.9 vs. 32% ± 0.8, 30.5% ± 0.8, 25.7% ± 0.6%, Table 3, column C). Google's performance was mirrored by comparing scores at the median (Table 3, column D) and at the 75th percentile (Table 3, column E). Although the range displayed between the lowest and highest maximum scores was wide (Table 3, column H), the agreement between graders showed acceptable reliability, with an ICC score of 0.525. The clearest evaluation, less subjective than a graded scoring metric, is summarized by how often each VA provided correct answers to queries in gynecologic oncology (Figure 1). Thus, while a difference between VAs was observed in responses to both general queries and queries in gynecologic oncology, correct responses to queries in gynecologic oncology were considerably reduced for all VAs. Totally correct replies to Google queries were higher ($n = 222$, 18.2%) than the other VAs (2.3–6.5%), Table 3, column I. Examination of the query formats in terms of totally correct responses showed that there was no significant difference in the Google responses to the three formats, while Alexa, Siri, and Cortana had more correct responses to the "What is X" format, $p \leq 0.05$ (Table 4). Consequently, query formats can influence VA responses to queries. These data show that Google provided the most accurate responses for both general queries and gynecologic oncology-related questions. However, all of the VAs provided fewer correct responses to the queries that were related to gynecologic oncology.

Table 3. Response Summary to Questions Related to GYN-ONC. Graded scores were averaged across all graders and expressed together with the standard error of the mean (±SEM). The number of queries (column A) varied due to dissimilar access devices for VA applications by different graders. Average graded score ± standard error (SEM) (column B) was calculated across individuals that evaluated responses by each VA. Average % score was determined as a percentage of the total possible score (column C). Medians (column D) and 75th percentiles (column E) were determined as a function of the total possible score. Minimum (column F) and maximum (column G) scores were determined as a function of the total possible score, and the difference between maximum and minimum as a function of the total possible score (column H).

		% Possible Score = Graded Score/Perfect Score							
	(A) $N_{queries}$ =	(B) Average Graded Score	(C) Average % Score	(D) Median Score	(E) 75th Percentile	(F) Minimum Score	(G) Maximum Score	(H) Difference Max-Min	(I) N Totally Correct (%)
Google	1224	2.88 ± 0.04 [a]	57.6 ± 0.9%	60%	80%	23.6%	75.6%	51.9%	222 (18.2%) [b]
Alexa	1152	1.60 ± 0.04	32.0 ± 0.8%	20%	40%	24.7%	58.3%	33.6%	75 (6.5%)
Siri	1008	1.52 ± 0.04	30.5 ± 0.8%	20%	20%	23.6%	80.3%	57.2%	55 (5.5%)
Cortana	1008	1.28 ± 0.03	25.7 ± 0.6%	20%	20%	16.4%	43.3%	26.9%	23 (2.3%)

Count of total correct answers to VA queries for each VA and percentages are expressed as a percentage of total queries (I). Significantly different $p < 0.05$ ANOVA [a] or Chi-square [b].

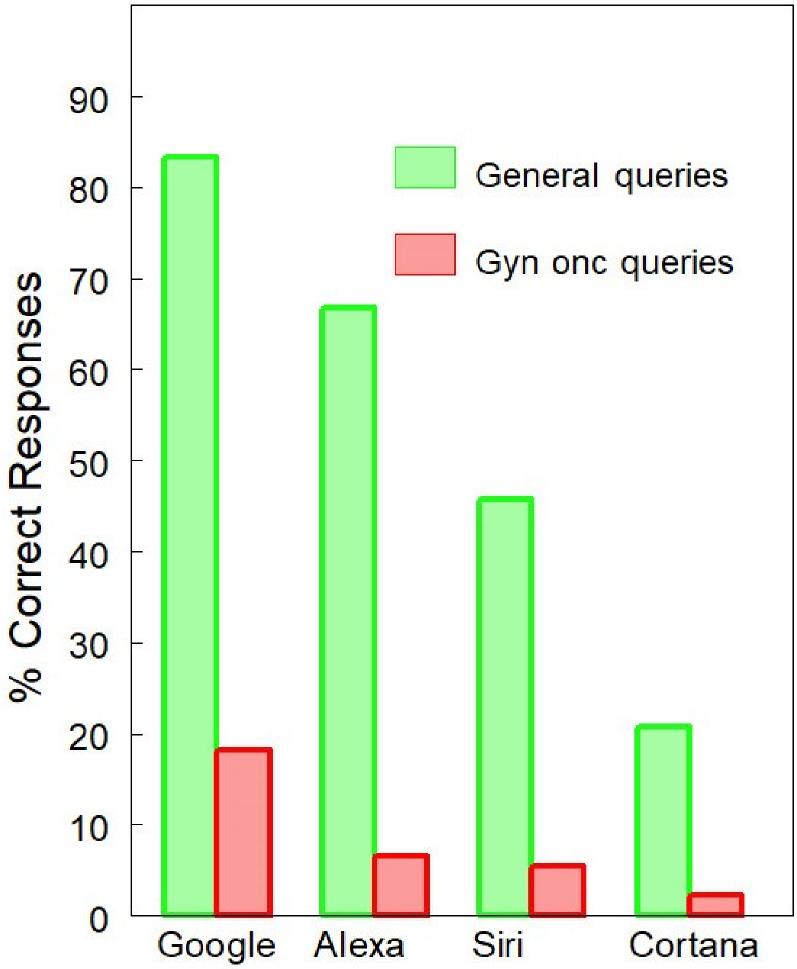

Figure 1. Comparative Performance of Google, Alexa, Siri, and Cortana on General Queries (green bars) and Queries Related to Gynecologic Oncology (red bars).

Table 4. Analyses of Query Format Presented to Google, Alexa, Siri, and Cortana. Number of totally correct responses by each VA. Each percentage in parenthesis is based on the number of totally correct answers by each VA for the template queries: "X?", "What is X?", and "Define X." N = total number of correct responses to gynecologic oncology-related queries, responses with a grade of 5 (meaning 100% correct). Significance was determined both by the Chi-square test and Fisher's exact probability test (* $p > 0.05$, ** $p \leq 0.05$).

	Google	Alexa	Siri	Cortana
"X?"	74 (33.3%) *	20 (26.7%) *	12 (21.8%) *	7 (30.4%) *
"What is X?"	67 (30.2%) *	35 (46.7%) **	30 (54.5%) **	10 (43.5%) **
"Define X"	81 (36.5%) *	20 (26.7%) *	13 (23.6%) *	6 (26.1%) *
N =	222 (100%)	75 (100%)	55 (100%)	23 (100%)

4. Discussion

In summary, audible replies by VAs to queries related to gynecologic oncology have room for improved accuracy. Our findings support those of Bickmore et al., that patients should not rely solely on VAs to answer medical questions [20]. For the queries evaluated, a well-trained gynecologic oncologist would answer the queries with more accuracy than the VAs. Our results are consistent with surveys showing that Google Assistant performed better than Siri or Alexa in addressing nonclinical queries [19]. Against a background of queries related to medicine, VA responses related to family medicine revealed room for improvement [18], and did not provide the addiction help that was intended [15]. We have not found reports on VA queries by gynecologic oncology patients; however, virtual visits to gynecologic oncologists have increased during the pandemic [21] and should continue in the future [22]. A proof-of-concept utilization of a virtual assistant has been reported for creating a molecular tumor board in gynecologic oncology, to integrate automated methods in collaborative treatment decisions [23]. Certain gynecologic oncology providers have already introduced "live chat" [24,25], which can be readily augmented with "voice chat" through specialized VAs. A data-driven approach, substituting virtual visits for in-person visits, has been suggested for the identification of symptoms related to ovarian cancer recurrence [26]. While not specific to gynecologic oncology alone, the Alberta Health Service is the first public healthcare organization in Canada to offer healthcare information via voice queries on Google and Amazon devices [27]. In addition, the United Kingdom's National Health Service is partnering with Amazon Alexa to answer health-related questions [28]. Amazon is also partnering with a telemedicine provider to start a voice-activated virtual program that prompts a call back from a telemedicine physician [28]. Our view is that the availability and utilization of VAs for medical information is increasing. This paper indicates that significant improvements are needed in the accuracy of gynecologic oncology information provided by the VAs considered here. Differences in VA performance may be related to the search engine employed by the VA, as well as the proprietary artificial intelligence employed by each. It is notable that keyboard entries yielded results with accurate answers on the questions posed here, while voice-based queries did not, indicating that information about gynecologic oncology can be discovered through a browser-based search on the Internet. It is possible that voice queries are filtered through different servers that are unable to adequately analyze and interpret the voice queries. The degree to which physical devices can clearly interpret speech without interference from background sounds is also a factor in VA performance. It should be remembered that preliminary use of a VA includes practice voice recognition steps that adjust for pronunciations and accents, so that device-based adjustments for these are made when setting up voice recognition. In the present study, 21 evaluators made 24 queries in three different formats, using different access devices, and hence should mirror performance expected in the real world. The present work provides an empirical standard for evaluating the reliability of information obtained from virtual assistants. Inaccuracy may be reduced by improvements at the search engine and VA artificial intelligence levels. Utilization of VAs internationally is a function of language and dialect [29]. Alexa supports 8 languages and 10 dialects, while Google Assistant supports 12 languages and 13 dialects, and has been working on more than 115 languages capable of speech recognition and natural language understanding [30]. By early 2022, Google became conversant in 30 languages in 80 countries, with Siri supporting 21 languages in 36 countries, and Cortana supporting 8 languages in 16 countries [30]. VA availability in an expanding number of languages attests to the international relevance of VA utilization.

In recent news, Microsoft has announced that it will launch an AI-powered Bing search engine on the Edge browser to "deliver better search, more complete answers, a new chat experience and the ability to generate content" [31]. Google plans a rollout of its AI chatbot named BARD, which is intended to enhance Google Search; however, in a preview demonstration, Google BARD provided erroneous information about discoveries made by the James Webb Space [32,33], while Microsoft's AI-powered Bing has also generated false information [34,35]. At present, neither Microsoft nor Google has associated these AI-technologies with receiving and responding to voice instructions. With new technology such as these, further analysis of such VAs may be required when the technology has been improved. The new AI-powered Bing search Engine and Google BARD will need to be examined for their validity in answering healthcare questions, in particular as related to gynecologic oncology. Specifically, after they are integrated into a VA environment, the first steps must evaluate how well they respond to both general queries and queries in gynecologic oncology. Subsequently, since they are based on deep learning, they will need to be evaluated for a learning style that improves performance and keeps up with changes relevant to gynecologic oncology.

In addition to Microsoft's AI-powered Bing search engine, OpenAI has released an AI system called Generative Pretrained Transformer 4 (GPT-4) that has a chat interface [36]. The chatbot gives a natural-language "response", normally within 1 s, that is relevant to the prompt [36]. This opens up a new conversation for how the chatbot could pertain to the medical field. Just like with the VAs, there is concern about the accuracy of the information from the chatbot. According to an article in the New England Journal of Medicine, a false response given by GPT-4 is referred to as a "hallucination" [36]. These errors can be dangerous in medical scenarios because they can often be subtle and stated in a manner in which the chatbot is very convincing.

5. Conclusions

The accuracy of VA responses to voice queries related to gynecologic oncology was very low and inferior to queries that were of a general nature. Since treatment options for gynecologic malignancies are highly dependent on disease classifications such as stage, inaccurate informative responses by VAs could lead to misunderstandings by patients that cause them to refuse or withdraw from treatments that otherwise might have been beneficial. The significance of the present work is that it identifies risks for patients that originate from inaccurate VA responses. The VA performance that we have reported on is most pertinent to questions that require the VA to reply with a correct definition. For questions that are more complex, it can be expected that VA performance will be even poorer, so that more complex queries to VAs should be avoided. In summary, audible replies by VAs to voice queries related to gynecologic oncology have considerable room for improved accuracy. Overall, we recommend caution when using VAs to obtain information in gynecologic oncology.

Author Contributions: Conceptualization, E.J.P. and J.M.L.; methodology, J.M.L.; software, E.J.P. and J.M.L.; validation, E.J.P. and J.M.L.; formal analysis, E.J.P. and J.M.L.; investigation, J.M.L., E.U. and S.U.; resources, E.J.P.; data curation, J.M.L., E.U., S.U., N.P., K.Q., J.W.G., M.J.R., M.L.H., J.D.L., D.H.Y. and E.J.P.; writing—original draft preparation, E.J.P. and J.M.L.; writing—review and editing, J.M.L., E.U., S.U., N.P., K.Q., J.W.G., M.J.R., M.L.H., J.D.L., D.H.Y. and E.J.P.; visualization, J.M.L.; supervision, E.J.P.; project administration, J.M.L.; funding acquisition, E.J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data will be made available if requested.

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

References

1. Fox, S.; Duggan, M. Health Online 2013. Pew Research Center 2013. Available online: https://www.pewresearch.org/internet/2013/01/15/health-online-2013/ (accessed on 16 June 2023).
2. Shah, P. How to Change Siri's Search Engine (And Other Tricks). Guiding Tech. 2019. Available online: https://www.guidingtech.com/change-siri-search-engine-tricks/ (accessed on 16 June 2023).
3. Snead, A. What Search Engine Does Alexa Use? And Can I Use Google to... Smarter Home Guide 2020. Available online: https://smarterhomeguide.com/alexa-search-engine/ (accessed on 16 June 2023).
4. DBS Interactive. Voice Search Statistics and Emerging Trends- Voice Search Statistics and Emerging Trends. Available online: https://www.dbswebsite.com/blog/trends-in-voice-search/ (accessed on 16 June 2023).
5. Georgiev, D. 2023's Voice Search Statistics- Is Voice Search Growing? Review 42. Available online: https://review42.com/resources/voice-search-stats/ (accessed on 16 June 2023).
6. SafeAtLast. Intriguing Amazon Alexa Statistics You Need to Know in 2022. Available online: https://safeatlast.co/blog/amazon-alexa-statistics/ (accessed on 6 June 2023).

7. McFarland, A. 10 Best AI Assistants (November 2022). Unite. AI. Available online: https://www.unite.ai/10-best-ai-assistants/ (accessed on 16 June 2023).
8. CNBC. Here's How Siri Made It Onto Your iPhone. Available online: https://www.cnbc.com/2017/06/29/how-siri-got-on-the-iphone.html (accessed on 16 June 2023).
9. Yaghoubzadeh, R.; Kramer, M.; Pitsch, K.; Kopp, S. Virtual agents as daily assistants for elderly or cognitively impaired people. In Proceedings of the 13th International Conference on Intelligent Virtual Agents, Edinburgh, UK, 29–31 August 2013; pp. 79–91. [CrossRef]
10. Phys Org. Personal Computer Sales Fall for Fifth Year in a Row. Available online: https://phys.org/news/2017-01-personal-sales-fall-year-row.html (accessed on 16 June 2023).
11. Roopinder, T. The Desktop Computer Was in Decline. The Pandemic Made It Worse. Engineering. Available online: https://www.engineering.com/story/the-desktop-computer-was-in-decline-the-pandemic-made-it-worse (accessed on 16 June 2023).
12. Yang, S.; Lee, J.; Sezgin, E.; Bridge, J.; Lin, S. Clinical advice by voice assistants on postpartum depression: Cross-sectional investigation using Apple Siri, Amazon Alexa, Google Assistant, and Microsoft Cortana. *JMIR Mhealth Uhealth* **2021**, *9*, e24045. [CrossRef] [PubMed]
13. Kocaballi, A.B.; Quiroz, J.C.; Rezazadegan, D.; Berkovsky, S.; Magrabi, F.; Coiera, E.; Laranjo, L. Responses of conversational agents to health and lifestyle prompts: Investigation of appropriateness and presentation structures. *J. Med. Internet Res.* **2020**, *22*, e15823. [CrossRef] [PubMed]
14. Alagha, E.C.; Helbing, R.R. Evaluating the quality of voice assistants' responses to consumer health questions about vaccines: An exploratory comparison of Alexa, Google Assistant and Siri. *BMJ Health Care Inform.* **2019**, *26*, e100075. [CrossRef] [PubMed]
15. Nobles, A.L.; Leas, E.C.; Caputi, T.L.; Zju, S.-H.; Strathdee, S.A.; Ayers, J.W. Responses to addiction help-seeking from Alexa, Siri, Google Assistant, Cortana, and Bixby intelligent virtual assistants. *NPJ Digit. Med.* **2020**, *3*, 11. [CrossRef] [PubMed]
16. Miner, A.S.; Milstein, A.; Schueller, S.; Hedge, R.; Mangurian, C.; Linos, E. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Intern. Med.* **2016**, *176*, 619–625. [CrossRef] [PubMed]
17. Wired. Amazon's Creep Into Health Care Has Some Experts Spooked. Available online: https://www.wired.com/story/amazons-creep-into-health-care-has-some-experts-spooked/ (accessed on 6 June 2023).
18. Hong, G.; Folcarelli, A.; Less, J.; Wang, C.; Erbasi, N.; Lin, S. Voice Assistants and Cancer Screening: A Comparison of Alexa, Siri, Google Assistant, and Cortana. *Ann. Fam. Med.* **2021**, *19*, 447–449. [CrossRef] [PubMed]
19. Laricchia, L. Share of Questions Answered Correctly by Selected Digital Assistants as of 2019, by Category. Statista. Available online: https://www.statista.com/statistics/1040539/digital-assistant-performance-comparison/ (accessed on 6 June 2023).
20. Bickmore, T.W.; Trinh, H.; Olafsson, S.; O'Leary, T.K.; Asadi, R.; Rickles, N.M.; Cruz, R. Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. *J. Med. Internet Res.* **2018**, *20*, e11510. [CrossRef] [PubMed]
21. McAlarnen, A.; Tsaih, S.-W.; Aliani, R.; Simske, N.M.; Hopp, E.E. Virtual visits among gynecologic oncology patients during the COVID-19 pandemic are accessible across the social vulnerability spectrum. *Gynecol. Oncol.* **2021**, *162*, 4–11. [CrossRef] [PubMed]
22. Mancebo, G.; Solé-Sedeño, J.; Membrive, I.; Taus, A.; Castells, M.; Serrano, L.; Serrano, L.; Carreras, R.; Mirapeix, E. Gynecologic cancer surveillance in the era of SARS-CoV-2 (COVID-19). *Int. J. Gynecol. Cancer* **2021**, *31*, 914–919. [CrossRef] [PubMed]
23. Macchia, G.; Ferrandina, G.; Patarnello, S.; Autorino, R.; Masciocchi, C.; Pisapia, V.; Calvani, C.; Lacomini, C.; Cesario, A.; Boldrini, L.; et al. Multidisciplinary Tumor Board Smart Virtual Assistant in Locally Advanced Cervical Cancer: A Proof of Concept. *Front. Oncol.* **2022**, *11*, 797454. [CrossRef] [PubMed]
24. Virtua Health. Gynecologic Oncology. Available online: https://www.virtua.org/services/cancer-treatment/gynecologic-oncology (accessed on 6 June 2023).
25. Dignity Health. Treating Gynecologic Cancers. Available online: https://www.dignityhealth.org/campaign-landers/gyn-oncology-surgery (accessed on 16 June 2023).
26. Feinberg, J.; Carthew, K.; Webster, E.; Chang, K.; McNeil, N.; Chi, S.; Roche, K.L.; Gardnerm, G.; Zivanovic, O.; Sonodo, Y. Ovarian cancer recurrence detection may not require in-person physical examination: An MSK team ovary study. *Int. J. Gynecol. Cancer* **2022**, *32*, 159–164. [CrossRef] [PubMed]
27. Brown, S. Partnerships between health authorities and Amazon Alexa raise many possibilities—And just as many questions. *CMAJ* **2019**, *191*, E1141–E1142. [CrossRef]
28. You Can Now Ask Amazon's Alexa to Call You a Doctor. Associated Press. New York Post. Available online: https://nypost.com/2022/02/28/amazons-voice-assistant-alexa-to-start-seeking-doctor-help/ (accessed on 6 June 2023).
29. Summa Linguae. Language Support in Voice Assistants Compared. Available online: https://summalinguae.com/language-technology/language-support-voice-assistants-compared/ (accessed on 6 June 2023).
30. Wiggers, K. Which Voice Assistant Speaks the Most Languages, and Why? The Machine. Available online: https://venturebeat.com/ai/which-voice-assistant-speaks-the-most-languages-and-why/ (accessed on 6 June 2023).
31. Mehdi, Y. Reinventing Search with a New AI-Powered Microsoft Bing and Edge, Your Copilot for the Web. Official Microsoft Blog. Available online: https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/ (accessed on 6 June 2023).

32. Miao, H. Alphabet Stock Drops 8% After Google Rollout of AI Search Features. The Wall Street Journal. Available online: https://www.wsj.com/livecoverage/stock-market-news-today-02-08-2023/card/alphabet-stock-drops-after-google-parent-introduces-ai-search-features-wgCJG3IDoSbfL3SgyrNI (accessed on 6 June 2023).
33. Martindale, J. How to Use Google Bard, the Latest AI Chatbot Service. DigitalTrends. Available online: https://www.digitaltrends.com/computing/how-to-use-google-bard/ (accessed on 6 June 2023).
34. Hao, K. What Is ChatGPT? What to Know About the AI Chatbot That Will Power Microsoft Bing. The Wall Street Journal. 10 February 2023. Available online: https://www.wsj.com/articles/chatgpt-ai-chatbot-app-explained-11675865177?st=q4wbp2ercfh1zo3&reflink=share_mobilewebshare (accessed on 6 June 2023).
35. Quach, K. Microsoft's AI Bing Also Factually Wrong, Fabricated Text During Launch Demo. The Register. Available online: https://www.theregister.com/2023/02/14/microsoft_ai_bing_error/ (accessed on 16 February 2023).
36. Lee, P.; Bubeck, S.; Petro, J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *N. Engl. J. Med.* **2023**, *388*, 1233–1239. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.