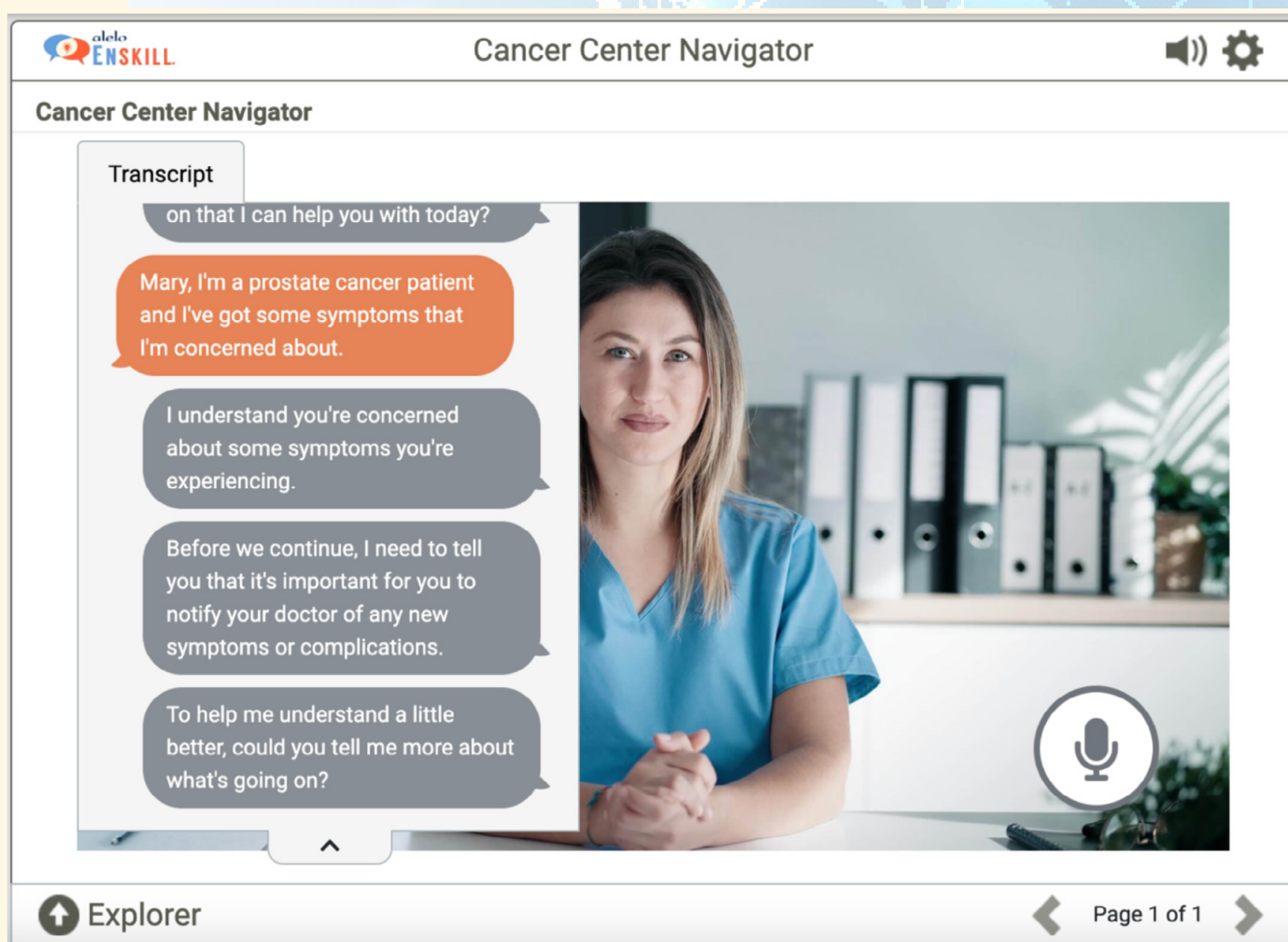


BUILDING TRUSTWORTHINESS INTO AI HEALTH NAVIGATORS

Lewis Johnson, Ph.D.
Alelo Inc.



The screenshot displays the 'Cancer Center Navigator' interface. At the top left is the 'alelo ENSKILL' logo, and at the top right are speaker and settings icons. The main content area features a 'Transcript' section on the left with a list of messages: 'on that I can help you with today?', 'Mary, I'm a prostate cancer patient and I've got some symptoms that I'm concerned about.', 'I understand you're concerned about some symptoms you're experiencing.', 'Before we continue, I need to tell you that it's important for you to notify your doctor of any new symptoms or complications.', and 'To help me understand a little better, could you tell me more about what's going on?'. To the right of the transcript is a video feed of a woman in blue scrubs sitting at a desk with binders in the background. A microphone icon is overlaid on the bottom right of the video. At the bottom left is an 'Explorer' button with an upward arrow, and at the bottom right are navigation arrows and the text 'Page 1 of 1'.

AI-powered tools offer great promise to relieve burdens on health teams and improve patient experiences, as noted in [a recent report by Harvard Business Review Analytic Services](#). They can help address what is sometimes referred to as the [Electronic Health Record Inbox Problem](#). Melanie Dixon, director of IS tech solutions at Ochsner Health Systems, is quoted in the HBR report as saying: “The way I see AI having an impact is in creating capacity, particularly for our providers. They spend so much time after hours answering emails, and it is overloading them. How could we use generative AI to help them answer some questions?”

For generative AI to be beneficial in interacting with patients and answering their questions, it must be reliable and provide trustworthy information, without time-consuming human supervision. [Health misinformation is a growing problem](#), and [generative AI solutions that hallucinate only add to the problem](#). Popular large language models (LLMs) offer medical advice, and there are [increasing concerns about the quality of that advice](#). Generative AI solutions that require constant expert review and editing negate the potential productivity benefits of generative AI. [A recent study of AI-generated draft email replies integrated into electronic health records](#), which require physician review and editing, showed that generated replies correlated with increased message read time, no change in reply time, and significantly longer replies. Experts who have studied the use of generative AI in healthcare have argued that ["it may be some time-or perhaps it will never be time-for practices to reply directly to patients without a human in the loop."](#)

At Alelo we have developed a new approach to applying generative AI to the provision of health information, called health navigators. Instead of generating unreliable and error-prone health information across a wide range of topics, health navigators focus on a constrained range of health-related topics, and provide trustworthy responses within that domain. Health navigators are inspired by the work of human patient navigators, who assist patients within the limits of their role and do not diagnose conditions or make treatment decisions. The work is informed by Alelo’s work in using AI-powered tools to [train community health workers](#) and patient navigators, [now employing generative AI](#). Alelo’s health navigators currently provide education and support to patients and their families. They also can provide trustworthy information to health professionals; for example, they can provide information about how to administer drugs safely and effectively, and provide references to supporting research publications. They can

recommend educational resources that patient educators can share with their patients. We envision expanding their role to cover a wide range of navigation services.

This white paper describes our process for building trustworthiness into health navigators, using current solutions as examples. Our approach ensures that health navigators meet the requirements of applicable medical, legal, and regulatory (MLR) review so that they can safely interact with patients, family members, and health professionals.

Types of AI Health Navigators

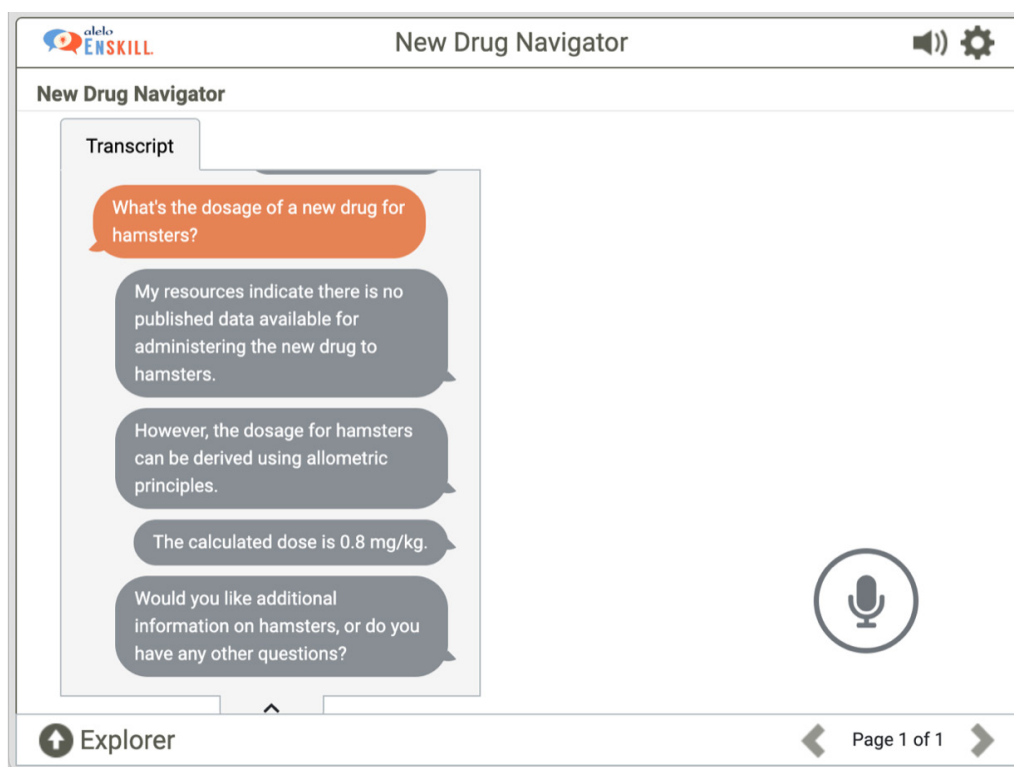
Alelo currently develops three main types of health navigators. They have related capabilities but address the needs of different types of users in different situations.

- **Virtual patient navigators** answer patients' questions, connect them with resources, and generally help patients through their care journeys. Virtual patient navigators operate within patient portals and supplement other services and resources that health systems provide through such portals.
- **Content navigators** are similar to virtual patient navigators, but are open to the public. They are intended to interact with prospective patients, caregivers and family members, and others seeking reliable health-related information. They also can support patient educators by suggesting educational content that would be relevant for their patients. They refer visitors to medical online content relevant to their concerns, answer basic questions, and provide contact information to schedule an appointment with a medical professional when appropriate.
- **Drug navigators** provide information about specific drugs and their use. They answer common questions and refer visitors to relevant documents such as package inserts, medical brochures, etc. They can provide references to supporting research publications. They provide guidance on how to report drug-related problems, such as adverse events. They can be designed to interact with health professionals or the general public.

Alelo's Cancer Center Navigator, shown in the figure at the top of this article, is an example of a health navigator that can be deployed either as a virtual patient navigator or as a content navigator. It asks users about the problem

they are seeking help with, and refers them to approved educational resources most relevant to their situation. For example, if the user is a patient who has just been diagnosed with lung cancer it can refer them to web pages that describe common lung cancer treatments, and also give them information about supportive services. If the user is a caregiver for a cancer patient, it can provide information specifically aimed at caregivers. If the user is a patient educator it suggests educational resources that they can share with their patients. We are integrating it with health system portals so that it can identify each visiting patient and recognize patients that are on a return visit. This will allow it to follow up on previous visits to see if the patient has any questions about educational materials that it recommended during the previous visit.

The screenshot below is an example of a drug navigator we developed for a veterinary drug company, Fidelis Animal Health. This navigator is able to answer a range of common questions about the product, relieving drug company staff of the burden of responding to emails with routine questions. If the visitor is reporting an important problem such as an adverse event it provides contact information to seek help from a staff member. This process can be automated so that the navigator collects information about the event from the visitor and sends it directly to the appropriate staff member for attention. This is advantageous from a regulatory perspective because it ensures that all adverse events are properly responded to.



Use Responsible AI as a Foundation

The first step in building trustworthy health navigators is to utilize foundation models that adhere to principles of [trustworthy and responsible AI](#), as articulated by the National Institute of Standards and Technology and other standards bodies. This includes making sure that the model is Reliable, Accountable, Fair, and Transparent in its use and of data. The foundation model must further support HIPAA compliance, to ensure safety and privacy of patient data. These factors restrict the choices of foundations model that we use in health navigators. Many model developers do not offer HIPAA-compliant services or do not provide specifics about how they adhere to responsible AI standards.

Yet even when foundation models are developed in accordance with trustworthy AI principles, they can still make mistakes if they are not used properly. [A recent independent evaluation of large language models](#) showed that state-of-the-art LLMs do not accurately diagnose patients across all pathologies, follow neither diagnostic nor treatment guidelines, cannot interpret laboratory results, and cannot be easily integrated into existing workflows because they often fail to follow instructions and are sensitive to both the quantity and order of information. Health AI applications that use large language models must therefore be designed so as to avoid such problems.

Design with MLR Review in Mind

Alelo health navigators are explicitly designed and developed so that their responses will pass MLR (medical, legal, and regulatory) review. We regard such reviews as the preferred standard for trustworthiness. Keeping MLR review in mind through the development process makes the task of MLR review teams easier, and helps to ensure that health navigators meet users' expectations.

When we design a navigator for a client, we determine the intended audiences for the navigator (patients, caregivers, oncology navigators, veterinarians, etc.) and the MLR review processes that the client employs for content aimed at those audiences. We reach agreement with the client as to which tasks and questions the navigator should handle on its own and what it should refer to a doctor or licensed medical, dental, or veterinary professional. We define clear boundaries for the navigator so that operates within its expected role. It may engage in dialogue with the user in order to

clarify its understanding of the user's problem, but it usually should not try to diagnose the problem or offer medical advice. We develop the navigator in a transparent way so that MLR teams can understand why the navigator generates the responses that it does and make sure that the responses meet expectations.

One consequence of this review-focused approach is that we typically build navigator capabilities incrementally. Generative AI systems can behave in complex, unpredictable ways, which poses a challenge for MLR review teams. Our preferred approach is therefore to start with a more limited set of navigator capabilities, get it approved, and then develop more advanced versions and submit them for approval. This simplifies and streamlines the approval process, as lessons learned from MLR review of early versions are applied to subsequent versions.

Build from an Inventory of Approved Resources

We always build health navigators from an agreed-upon inventory of approved resources. Resources may include any type of informative content: brochures, videos, web pages, any materials that are created to communicate to the target audience. "Approved" means that the resources have been approved through the organization's MLR review process, or by the MLR review process of another organization that the client trusts. Reliance on trustworthy resources helps to ensure that the resulting navigator is also trustworthy.

Some of the resources that a navigator uses may not require MLR review. For example, a virtual patient navigator for a cancer center might use resources that explain to new patients where to park their car, how to arrive via public transportation, and where to go when they arrive at the center. Nevertheless, we document all resources that are used to create the navigator, so that the client can determine what review processes are necessary for those resources.

The New-Drug Navigator for Fidelis Animal Health currently uses a small inventory of resources: the drug's package insert, the Frequently Asked Questions page on the product website, the laboratory veterinarian brochure, and the veterinary practitioner brochure. There are approximately forty research publications that relate to the use of the drug, and we may add these to the resource inventory for the navigator. The Cancer Center Navigator currently uses an inventory of 42 resources. These include web

pages and documents found on a major cancer center's website, and videos on the center's YouTube channel. Some were developed by the cancer center itself, and some were developed by other cancer resource organizations such as the American Cancer Society. This inventory could be greatly expanded. The cancer center's YouTube channel has 750 videos, and the center's website includes a large number of webpages describing a wide variety of cancer conditions and treatments. The sheer number of resources makes it difficult for users to find relevant and appropriate materials, which is one reason why navigators can help.

Alelo's navigators rely exclusively on the approved resources when generating responses to give to users. They explicitly provide links to the resources that they used to generate the responses. They do not search the Internet for additional information. This helps MLR teams to make sure that responses can be relied upon. In contrast, chatbots that search the Internet for information sometimes draw on resources whose information is not always reliable, such as WebMD. ChatGPT asserts that it draws on reputable resources, but it does not cite any specific resources that an MLR team can check for reliability. This makes it difficult to detect possible hallucinations.

Process the Approved Resources

Once the resources have been selected, they must be processed so that the navigator can make use of them. First, a description of each resource is needed so that the navigator can determine its relevance to the user's situation, explain its relevance, and answer basic questions about it.

Some navigators are designed to answer user questions using information drawn from the contents of the resources. Although we generally recommend this, we check with the MLR review team to understand their expectations regarding responses to questions. Is the navigator free to generate free-form responses using the approved resources, or does the MLR team have specific expectations about how the navigator should answer specific questions? In either case we can provide documentation of navigator responses to facilitate MLR review. It should be understood however that the wording of generative AI's responses can vary and one should not expect the navigator to generate verbatim responses to a fixed set of questions.

Ingest and Employ the Approved Resources

Once the processing is done, we ingest the resources into a database that the navigator can employ using RAG (retrieval augmented generation) techniques. We provide the navigator with specific instructions as to how it should ingest the resources if there are issues that require particular attention. For example, in the case of the drug navigators must be avoid discussion of off-label use, and recognize any user questions that involve unapproved use or imply plans to use the product outside of its intended and approved parameters.

We provide the navigator with guidelines for how to interact with the user. The navigator is instructed to engage in coherent dialogue with the user instead of simply responding to questions. Follow-up questions and answers should take into account and connect with the previous discussion with the user. Navigators are instructed to adhere to core principles and guidelines in communicating with users. In the case of patient-facing navigators they should be patient-centered, focusing on the patient's needs and concerns. They should be empathetic and compassionate, and provide information that is informative and actionable. They should engage in active listening to acknowledge the patient's feelings and concerns. They should communicate in an ethical and professional manner. Drug navigators should prioritize safety and compliance, as well as conciseness and clarity. This all contrasts with typical chatbots that tend to generate long-winded responses without taking the time to understand the patient's needs and concerns.

We also provide a list of types of questions that the navigator should not answer, or that require special action. Navigators should not answer questions that require medical judgment, and should instead refer them to a medical professional. We also provide instructions for what to do if available information is insufficient to fully answer the question.

The MLR Approval Process

To support the MLR approval process, we document clearly what resources were used, how they were processed, and how the navigator employs them. This provides transparency so that MLR review teams can make sure that the resources and intermediate work products are trustworthy. We can also provide test results to show how the navigator responds to various types of questions. We provide review teams with beta versions of the navigator for their independent testing and review.

Alelo's Enskill platform anonymizes and saves transcripts of each interaction with the navigator.

MLR teams can review the transcripts as needed to ensure that compliance is maintained. We can also generate summary reports to facilitate the review process.

Conclusion

This document describes how Alelo works to ensure that health navigators are trustworthy. We apply care throughout to ensure that Alelo health navigators rely only on trusted information sources, and use them to generate trustworthy responses. We design navigators with MLR review in mind, and document the development process so that MLR teams can easily review and approve as necessary, and audit performance as needed.

Generative AI is a highly dynamic field. We likely will revise our development process as new technologies emerge, as health providers develop their MLR review methods for AI-powered tools, or as new regulations are established governing generative AI in healthcare.

We believe that the development of AI for health requires deep attention to safety and transparency. These characteristics are currently sadly deficient in the field of generative AI. Providers focus on delivering advanced new capabilities that are prone to error, and leave it to the user to check results themselves. It can be difficult to determine how a bot developed its recommendations and whether they are based on reliable sources. Although generative AI by its nature can be variable in its responses, we find that by carefully delineating the role that AI should perform and ensuring that it works in a transparent way using trustworthy sources, it is possible to create AI that is medically accurate and fully meets legal and regulatory requirements.

References

Alelo (2024). KDHRC and Alelo Develop Avatar-Based CHW Training Simulations to Promote Postpartum Maternal Health. Retrieved Dec. 16, 2024 from

<https://www.alelo.com/2024/11/kdhrc-and-alelo-develop-avatar-based-chw-training-simulations-to-promote-postpartum-maternal-health/>.

Hager, P., Jungmann, F., et al. (2024). Evaluation and mitigation of the limitations of large language models in clinical decision-making. *Nat Med.* 2024 Sep;30(9):2613-2622. doi: 10.1038/s41591-024-03097-1. Epub 2024 Jul 4.

Harvard Business Review Analytics Services (2024). *AI in the Health Care Industry: Driving Efficiency and Better Patient Experiences*. Harvard Business School Publishing.

Hatem, R., Simmons, B., & Thornton, J.E. (2023). A Call to Address “Hallucinations” and How Healthcare Professionals Can Mitigate Their Risks. *Cureus* 2023 Sep 5;15(9):e44720. doi: [10.7759/cureus.44720](https://doi.org/10.7759/cureus.44720).

Johnson, W.L. (2021). IELA Award Winner-Case Study: Rapid Training of Community Health Workers using Enskill®. *International Journal of Advanced Corporate Learning* 14(2), <https://doi.org/10.3991/ijac.v14i2.25929>.

National Institute of Standards and Technology (2024). *AI Risks and Trustworthiness*. Retrieved Dec. 16, 2024 from https://airc.nist.gov/AI_RMF_Knowledge_Base/AI_RMF/Foundational_Information/3-sec-characteristics.

Nyce, C.M. (2024). Google is Playing a Dangerous Game with AI Search. *The Atlantic*, May 24, 2024.

Office of the Surgeon General (2024). *Confronting Health Misinformation: The U.S. Surgeon General’s Advisory on Building a Healthy Information Environment*. Retrieved Dec. 16, 2024 from <https://www.hhs.gov/surgeongeneral/priorities/health-misinformation/index.html>.

Rotenstein, L.S. & Wachter, R. M. (2024). Are Artificial Intelligence-Generated Replies the Answer to the Electronic Record Inbox Problem? *JAMA Netw Open.* 2024;7(10):e2438528. doi:[10.1001/jamanetworkopen.2024.38528](https://doi.org/10.1001/jamanetworkopen.2024.38528).

Sinsky, C.A., Shanafelt, T.D., & Ripp, J.A. (2022). The Electronic Health Record Inbox: Recommendations for Relief. *J Gen Intern Med.* 2022 Aug 29;37(15):4002-4003. doi: 10.1007/s11606-022-07766-0

Tai-Seale, M., Baxter, S.L. et al. (2024). AI-Generated Draft Replies Integrated Into Health Records and Physicians' Electronic Communication. *JAMA Netw Open.* 2024 Apr 15;7(4):e246565. doi: [10.1001/jamanetworkopen.2024.6565](https://doi.org/10.1001/jamanetworkopen.2024.6565).