AI in the Family: Care Collaboration in Pediatrics as a Testbed for Challenges Facing AI in Healthcare

Sarah Nikkhah and Andrew D. Miller

Indiana University-Purdue University Indianapolis (IUPUI), USA

In this position paper, we argue that family caregiving collaboration is a promising context for studying the role of HCI in the implementation of AI in real-world healthcare. We draw on our own fieldwork in pediatric cancer care to explain how some of the grand challenges facing AI could be meaningfully studied through the context of family caregiving collaboration. We identify some of the current challenges in the implementation of AI, such as 'explainability,' actionability,' and performance,' and describe how family caregiving coordination can be used as a testbed for implementing AI in healthcare.

CCS CONCEPTS • Information systems → Information systems applications.

Additional Keywords and Phrases: artificial intelligence; human computer interaction; family; caregiving;

1 INTRODUCTION: FROM AI IN PERSONAL HEALTHCARE TO AI IN FAMILY HEALTHCARE

AI is usually thought of as a means to help an individual understand a recommendation or understand the individual's mindset to provide justifiable recommendations or information. Challenges in implementing personalized health recommendation and analysis algorithms in the real-world setting have been referred to as the 'last mile [2,3]' for AI in healthcare. Many AI methods and applications in healthcare are being developed to provide services to patients and providers. These applications can range from AI for diagnosis assistance (such as IBM Watson, ADA, Your.MD, DeepMind Health, Face2Gene) to personal informatics (such as ADA, Your.MD, Cardio, Gymfitty, and Get in Shape) [7].

However, individuals exist within and interact with a diverse ecology of environments, stakeholders, and social contexts. If AI is to truly transfer healthcare, it must be designed to account for the varied interpersonal effects[11] on real-world use. While some AI applications have also been designed to support patient-provider interaction (such as ADA and Your.MD [7]), there appear to be no AI applications designed explicitly for caregivers' collaboration with each other or even caregiver collaboration with the patient and medical team. In the pediatric context,

In our research, we have been studying the caregiving collaboration needs of parents and other family members of children diagnosed with cancer. These children often need extended hospitalizations, careful monitoring, and multiple rounds of treatment over many months. Their families reshape their lives in order to care for their ill child while maintaining existing personal and professional responsibilities. Pediatric cancer care also increasingly requires more nuanced decisions about a variety of promising treatments, including precision medicine treatments. Families could benefit not only from AI that supports clinical decision-making but also AI that supports family collaboration and symptom tracking. However, in order to unlock the potential of these technologies, AI tools face numerous challenges beyond those already identified. We are currently working with families to co-design technologies that could allow them to keep track of symptoms, come to consensus decisions around treatment and care options, and reduce isolation and information disparities within the caregiving group.

In this position paper, we describe several potential opportunities for AI in healthcare, drawn from our fieldwork with families with a child with cancer. We suggest that current AI technologies in healthcare need to expand their focus from individuals and include interpersonal relationships and caregiving groups. Some of the opportunities for family-centered AI represent 'low hanging fruit.' For example, while dictation, transcription, and summarization technologies for clinical practitioners are in active development [14], these tools are not currently available for caregivers. One of the most common challenges experienced by families in our research is family-centered daily rounds during hospitalization. Parents and other caregivers struggle to attend to the rapid-fire discussion while also keeping notes to share with each other. Sometimes no caregiver is able to be present during rounds, and patients alone may not be in a position to keep track of this vital information (due to the age of the child or treatment-related fatigue). Caregiver-focused transcription and summarization technologies could meaningfully improve this daily caregiving task, allowing caregivers and patients to accurately recall the content of clinical encounters and keeping the caregiver network more informed. But such technologies could *also* support unanchored contexts [9]: for example, transcribing a call between caregivers helps them manage caregiving responsibilities and be connected and informed about the patient's health status and treatment plan.

However, the road from personal AI to family-centered AI will not be an easy one. Even the examples that seem 'obvious' to us as family-centered HCI researchers will not be achievable by simply repurposing existing algorithms. Moreover, significant sociotechnical implementation barriers exist as patients, providers, and caregivers negotiate the proper role and responsibility for AI. In the sections that follow, we describe three key challenges to family-centered AI in healthcare: explainability, actionability, and performance. For each challenge, we propose some new ways to reframe the discussion, grounded in our experience in the field.

2 FROM 'EXPLAINABILITY' TO 'RE-EXPLAINABILITY'

From an end-user perspective, current AI recommendations or syntheses work like 'black boxes,' transforming inputs into outputs without justifications [8,15]. This makes it hard for humans to assess the reliability and trustworthiness of AI-generated information. The 'explainability' challenge is particularly critical for AI in healthcare contexts [6,10,12,13], where recommendations can have a meaningful impact on health outcomes. Current research in this area typically focuses on delivering explanations to individuals, such as "the doctor" or "the patient."

However, different individuals will need different kinds of explanations and will also often need to explain AI-generated data to *each other*. Our ongoing fieldwork with the parents of hospitalized children shows that each family caregiver might have different levels of understanding and knowledge about the diagnosis and treatment. Additionally, real-world caregiver networks include many individuals who currently only hear medical explanations' second hand.' For example, 'primary' caregivers usually spend more time with the child in the hospital and know more about the diagnosis and treatments. These primary caregivers then share information with other family members. In our research with families of children hospitalized for cancer care, mothers tend to be the primary caregivers. Moms stayed at the hospital and provided care, and shared required updates and information with dad and other family members in the caregiving circle. Secondary caregivers, usually fathers, mentioned how they do not have the opportunity to be with their hospitalized child due to their work and that they try to stay connected with mom in the hospital to be updated about their child's health.

Explainable AI systems, therefore, may need to not only provide *an* account of their conclusions but *multiple* accounts, suitable for different levels of familiarity with the underlying data, health literacy, and level of decision-making authority. If AI systems are to facilitate explanations across the caregiver network, they will need to address these differences and answer follow-up questions much like a doctor and be able to do so to different caregivers.

The family context also provides a further opportunity for explainable AI systems: the ability to support caregivers and patients as they *re-explain* information within the caregiver network. Such re-explanation systems might take the form of a chatbot feature or a messaging group with different members and channels of information. Where not only the bot tries to explain a recommendation or decision, but family members can help each other understand the explanation and decide collectively.

3 FROM 'ACTIONABILITY' TO 'INTERACTIONABILITY'

One grand challenge for AI in healthcare is the 'implementation gap': few machine learning algorithms ever make it to the bedside [6,15]. AI health experts argue that a new standard is needed: clinical 'actionability [15].' That is, recommendations or assessments need not merely be *accurate* but also *useful* in supporting human decision-making and improving health outcomes. Even the most sophisticated learning model is useless 'in the wild' unless it clearly indicates a follow-up action, such as ordering a test or prescribing medication.

However, in many cases where AI could be transformative, mere 'actionability' may not be enough. Often, the next course of action is not up to the sole discretion of a single individual at a single moment. To truly realize AI's potential in clinical outcomes, learning algorithms will need to support *interaction*, both responding to an individual's queries and context and fitting into existing social ecologies. In the transcription and summary example we mentioned in the introduction, providers we work with disagree about whether such tools would even be allowable during a clinical encounter, and additional review processes may be required in order to gain consent from all affected parties.

Moreover, that was just for a *descriptive* algorithm. These challenges multiply when AI tools begin offering prescriptions. For example, what if the patient or caregivers disagree with the action suggested or taken by the AI? What if caregivers disagree with *each other?* We have seen many instances of this kind of issue even without AI tools. For example, one family in an interview study recounted serious disagreements with clinical staff about a planned discharge. The clinicians who came in on the weekend interpreted their child's levels differently from the weekday clinical team and recommended discharge. The situation was ultimately resolved, and the child was not discharged. In this scenario, the final decision was based on interactions between family caregivers and providers coming together to interpret the data collaboratively. An algorithm that supports such a collaborative process will therefore not only need to provide an *actionable* recommendation (stay inpatient vs. discharge) but to offer its explanations in a way that supports interaction. That is, AI for family care situations must be prepared to participate as social actors in the family system.

Enabling *interactionability* remains a key challenge and one that family-centered Human-Computer Interaction work is primed to meet. We are already exploring ways to enable interactionability as we design prototype family collaboration technologies for pediatric cancer. For example, many families we encounter primarily share information through group chats or texts. They find it difficult to keep track of what information they have shared, who has agreed to do what, and when is the right time for a live chat. AI could support all of these tasks, but it is not yet clear to what degree the AI should present itself as an active collaborator (proactive agent) vs. 'smart' process (performing analysis on demand) vs. invisible infrastructure (setting defaults, performing routine tasks such as transcription, photo segmentation, 'feed' cleanup, etc.). In many ways, this is a similar challenge to the 'ghost work[4]' or 'articulation work[1]' concepts from CSCW research. AIs can and will perform and create much hidden work. As Park et al. noted in their call for a previous workshop on this topic [13], accounting for AI in the wild also necessitates accounting for invisible work. When AI-enabled health tools are implemented in the family context, the family member who already takes more responsibility in providing care of the child will end up being the one who has to add data to the system and interpret the results because they are more aware and informed about the child symptoms and medical information. Therefore, well-functioning AI systems may appear to add *extra work*. However, it may be that these systems reduce the overall burden if indeed the 'extra work' is merely ghost work that had been previously unacknowledged. A focus on the ability for AIs to *act* as social actors, and *interact* with caregiving groups, both performing and creating hidden work, will allow us to truly assess whether a given AI is actually worth implementing.

4 RETHINKING 'PERFORMANCE'

How do we know if an AI tool works well in the real world? Currently, AI tools are commonly judged in terms of 'model performance' metrics. At the very least, an AI tool needs to be accurate and reliable. Being accurate means that the AI system should predict systematic errors, and being reliable means it should predict random errors and being consistent. There are different performance measures such as precision, sensitivity, accuracy, and F1 scores that help assess an AI system's performance [16]. However, these metrics are not enough to assess the successful implementation of AI applications in the real world. We suggest that performance metrics of AI tools need to be revised to include metrics relevant to the successful implementation in the healthcare setting and that interpersonal and caregiver-focused metrics are essential to real-world performance.

In the family caregiving context, an AI tool can only be considered 'performant' if it enables understanding, fosters collaboration, and supports conflict resolution. Rethinking explainability and actionability will go a long way to a more honest appraisal of performance, but there are many additional factors waiting in the wings. The family context affects the entire data pipeline, from provenance to cleaning. For example, one of us (Andrew Miller) is collaborating with health researchers on a pediatric oncology symptom reporting tool that allows parents and children to report a wide variety of experiences and conditions in a systematic way [5]. The data from this tool will hopefully be used to identify symptom trends and develop precision medicine treatments. However, any models developed on top of this data must take into account family-centered aspects of the data entry itself. Different caregivers of the same patient may assess symptom severity in different ways, report at a differing frequency, and have different levels of symptom recall. Current AI tools do not take such differences into account as they do not accept multiple inputs from different participants for the same data.

Another challenge facing AI in healthcare is related to information safety and privacy [11,12]. Large datasets, including data related to medical information of the individuals, need to be fed to an AI system to train the system so that the system can be trained. This data can be accessed at different levels by different stakeholders such as medical providers, healthcare professionals, hospitals, software companies, software developers, and data collectors. Therefore, in addition to proper methods to reliably deidentify patient's information, there is a need for robust data and algorithm protection techniques and encryption systems to avoid access to third parties and ensure cybersecurity and information safety. In addition to such challenges, when implementing AI-enabled healthcare systems in real-world settings, even at the level of the family system, different family members, including patients and their caregivers, may require or prefer different privacy settings for specific information and AI tools need to account for these needs and preferences.

At this time, model performance metrics, including precision, sensitivity, accuracy, and F1 scores, are the main factors to assess the abilities of an AI tool. Nevertheless, in the healthcare setting, we argue that mere model performance metrics are not enough. Proper 'performance' requires attention to interpersonal metrics such as privacy, re-explainability, and interactionability. Defining and integrating these additional performance measures will be essential to the successful implementation of AI.

5 CONCLUSION

In this position paper, we showed that family caregiving collaboration is a promising context for studying the implementation of AI in real-world healthcare. Drawing on our own fieldwork in pediatric cancer care, we identified some of the current challenges in the implementation of AI, such as 'explainability,' actionability,' and performance.' There are no easy answers yet. But we are optimistic that a sociotechnical framing and a true accounting of context of use will allow for the development and implementation of truly transformative AI technologies.

6 ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1850273.

7 REFERENCES

[1] Geoffrey C. Bowker and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. MIT Press.

[2] Federico Cabitza, Andrea Campagner, and Clara Balsano. 2020. Bridging the "last mile" gap between AI implementation and operation: "data awareness" that matters. *Ann Transl Med* 8, 7 (April 2020). DOI:https://doi.org/10.21037/atm.2020.03.63

[3] Enrico Coiera. 2019. The Last Mile: Where Artificial Intelligence Meets Reality. *Journal of Medical Internet Research* 21, 11 (2019), e16323. DOI:https://doi.org/10.2196/16323

[4] Mary L. Gray and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt.

[5] by Sherri Deatherage Green. Crane receives grant to improve symptom reporting for kids with cancer. Retrieved February 28, 2021 from https://www.uth.edu/news/story.htm?id=23240f34-e998-4005-be84-ab61c6a2cfb7

[6] David Higgins and Vince I. Madai. 2020. From Bit to Bedside: A Practical Framework for Artificial Intelligence Product Development in Healthcare. *Advanced Intelligent Systems* 2, 10 (2020), 2000052. DOI:https://doi.org/10.1002/aisy.202000052

[7] Oksana Iliashenko, Zilia Bikkulova, and Alissa Dubgorn. 2019. Opportunities and challenges of artificial intelligence in healthcare. *E3S Web Conf.* 110, (2019), 02028. DOI:https://doi.org/10.1051/e3sconf/201911002028

[8] T. J. Kasperbauer. Conflicting roles for humans in learning health systems and AI-enabled healthcare. *Journal of Evaluation in Clinical Practice* n/a, n/a . DOI:https://doi.org/10.1111/jep.13510

[9] Predrag Klasnja, Andrea Civan Hartzler, Kent T. Unruh, and Wanda Pratt. 2010. Blowing in the wind: unanchored patient information work during cancer care. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10), Association for Computing Machinery, New York, NY, USA, 193–202. DOI:https://doi.org/10.1145/1753326.1753355

[10] Tamra Lysaght, Hannah Yeefen Lim, Vicki Xafis, and Kee Yuan Ngiam. 2019. AI-Assisted Decision-making in Healthcare. *ABR* 11, 3 (September 2019), 299–314. DOI:https://doi.org/10.1007/s41649-019-00096-0

[11] Jessica Morley and Luciano Floridi. 2019. *NHS AI Lab: Why We Need to Be Ethically Mindful About AI for Healthcare*. Social Science Research Network, Rochester, NY. DOI:https://doi.org/10.2139/ssrn.3445421

[12] K. Paranjape, M. Schinkel, and P. Nanayakkara. 2020. Short Keynote Paper: Mainstreaming Personalized Healthcare–Transforming Healthcare Through New Era of Artificial Intelligence. *IEEE Journal of Biomedical and Health Informatics* 24, 7 (July 2020), 1860–1863. DOI:https://doi.org/10.1109/JBHI.2020.2970807

[13] Sun Young Park, Pei-Yi Kuo, Andrea Barbarin, Elizabeth Kaziunas, Astrid Chow, Karandeep Singh, Lauren Wilcox, and Walter S. Lasecki. 2019. Identifying Challenges and Opportunities in Human-AI Collaboration in Healthcare. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, ACM, Austin TX USA, 506–510. DOI:https://doi.org/10.1145/3311957.3359433

[14] Kshitij Saxena, Robert Diamond, Reid F. Conant, Terri H. Mitchell, ir. Guido Gallopyn, and Kristin E. Yakimow. 2018. Provider Adoption of Speech Recognition and its Impact on Satisfaction, Documentation Quality, Efficiency, and Cost in an Inpatient EHR. *AMIA Jt Summits Transl Sci Proc* 2018, (May 2018), 186–195.

[15] Martin G. Seneviratne, Nigam H. Shah, and Larry Chu. 2020. Bridging the implementation gap of machine learning in healthcare. *BMJ Innovations* 6, 2 (April 2020). DOI:https://doi.org/10.1136/bmjinnov-2019-000359
[16] 2019. 4 things you need to know about AI: accuracy, precision, recall and F1 scores. *lawtomated*. Retrieved February 28, 2021 from https://lawtomated.com/accuracy-precision-recall-and-f1-scores-for-lawyers/