Chatting Beyond ChatGPT: Advancing Equity Through AI-Driven Language Interpretation

Leen Bakdash, BA[1], *Areeba Abid, BS*[1], *Amritha Gourisankar, BS, MPH*[1], *and Tracey L. Henry, MD, MPH, MS*[2]

[1]Emory University School of Medicine, Atlanta, GA 30322, USA; [2]Division of General Medicine, Emory University School of Medicine, Atlanta, GA, USA

ABSTRACT

Medical interpretation is an underutilized resource, despite its legal mandate and proven efficacy in improving health outcomes for populations with low English proficiency. This disconnect can often be attributed to the costs and wait-times associated with traditional means of interpretation, making the service inaccessible and burdensome. Technology has improved access to translation through phone and video interpretation; with the acceleration of artificial intelligence (AI) large language models, we have an opportunity to further improve interpreter access through real-time, automated translation. The impetus to utilize this burgeoning tool for improved health equity must be combined with a critical view of the safety, privacy, and clinical decision-making risks involved. Physicians must be active participants and collaborators in both the mobilization of AI tools to improve clinical care and the development of regulations to mitigate harm.

J Gen Intern Med 39(3):492–5
DOI: 10.1007/s11606-023-08497-6

INTRODUCTION

We are living amidst an artificial intelligence (AI) revolution. Over the last few years, it has become clear this technology is an inevitable force, changing much of our lives, whether we recognize it or not. As society reckons with the implications of consumer-facing AI, like ChatGPT, we have the opportunity to design solutions to one of healthcare's greatest challenges—immigrant health equity. Twenty-five million individuals in the USA have low English proficiency (LEP).[1] Despite the growing number of patients with LEP, our healthcare system continually fails this population, particularly in terms of interpretation. As patients with LEP receive inferior care due to language barriers, we must examine whether emerging technologies can address this enduring problem. Here, we will explore the nature of language barriers for patients with LEP, the potential for AI-based language models to address this issue, and the work ahead to make this solution a safe reality.

HISTORY

Addressing the communication needs of patients with LEP is a legal and clinical necessity for healthcare systems. The right to an interpreter was cemented in US legislation on August 11, 2000, when President Clinton issued Executive Order 13,166, "Improving Access to Services for Persons with Limited English Proficiency." This major step in improving care for LEP populations required organizations that receive federal funding, including hospitals serving Medicare or Medicaid patients, to provide interpreter services to all patients who need them. The Affordable Care Act expanded these laws, increasing the standards for language access, including assessment of interpreters for quality and prohibition of hospitals from relying on bilingual family members (especially minors) except in emergencies.[1] These are vital standards for the 25 million individuals with LEP in the USA. *Professional* medical interpretation is associated with improved clinical outcomes and patient satisfaction, and decreased medical costs.[2]

THE PROBLEM

Despite the medical, legal, ethical, and financial imperatives, interpreters remain underutilized. Estimates of interpreter usage are limited and varied: at an academic tertiary care center, 65% of LEP patients had no documented interpreter use and only 4.8% of patients had interpreter use when seeing a hospitalist.[3] An investigation of orthopedic clinics in California found that a secret shopper "patient" was asked to rely on a non-qualified interpreter at 80% of clinics and asked to bring a friend or family member for interpretation at 28% of clinics.[4] A 2023 study of pediatric hospitalists found only 65% of respondents always used interpreters for admissions, 57% for discharges, and 40% on rounds.[5] These studies suggest that professional interpreters are not used in a majority of clinical encounters.

Received July 17, 2023
Accepted October 16, 2023
Published online October 30, 2023

COMMON BARRIERS

Given how critical interpretation is to quality care, why is it used so inconsistently? Cost is a common barrier, especially for smaller, community-based clinics and free clinics that care for marginalized populations. Interpreter costs range from $1.25 to 3/min for telephone interpreters, $1.95–3.49/min for video interpreters, and $45–150/h for in-person interpreters. Federally funded insurance, like Medicaid and the Children's Health Insurance Program (CHIP), only provides reimbursement for interpreter costs in 14 states.[1] Furthermore, major interpreter shortages make these resources difficult to access, especially for less common languages. [7] Even when interpreters are available, providers often forgo their use, citing delays in wait-time for interpreters and concerns about increased visit lengths. Many settle for subpar communication, relying on informal interpreters like family members or no interpreter at all.[5,6]

TURNING THE TIDE

The right to medical interpretation has been legally enshrined for over two decades but current interpretation options do not meet patient or provider needs. Rather than accepting existing options, we must consider novel ideas for the longstanding problems with medical interpretation. Since 2000, we have witnessed a technological revolution: two decades ago, the internet was a budding novelty and now most Americans carry portable supercomputers in their pockets; electronic medical records are widespread and apps exist for almost every imaginable purpose. In the twenty-first century, there is an enormous untapped opportunity for technology to fill gaps in medical interpreter use, particularly with the emergence of AI capabilities.

Emerging AI capabilities can expand the impact technology has made on interpreter availability. Phone and video interpreters for hundreds of languages on demand provide faster and cheaper access, even for rare languages. Online trainings make translator certification easier, increasing the supply of interpreters. In extremely low-resource settings, crowd-sourced interpretation helps connect refugees and migrants with volunteer interpreters across the world. Despite these improvements, time and cost continue to be barriers.

Recent advancements in AI language models, the ubiquity of smartphones among providers, and the resulting potential for automated bedside translations could reduce these barriers. While in-person interpreters have been established as the gold standard, followed by video and then phone interpreters, we cannot leave our patients without interpretation while we wait for access to traditional interpretation to improve.[8] Long term, we must address the root causes of these barriers by improving reimbursements for interpreter use, increasing the supply of in-person and remote interpreters, and educating physicians on best practices for efficient interpreter use. However, AI language models bring us solutions that can improve health equity while these systemic changes are implemented. Having automated translation on providers' smartphones, translating in real-time conversation between provider and patient, provides a solution, without the need for patients to have smartphone access, to bridge the gap in interpreter access and use.

Automated translation is not new. As early as the 1980s, automated translation models for text were being developed; hundreds of apps are now available for medical translation. However, these publicly available apps have varying quality and limited usefulness in a clinical setting.[9] While large language models like ChatGPT dramatically improved the ability of computers to accomplish a wide range of language tasks, they are currently inadequate for medical interpretation. For example, these models are trained on online text datasets and perform better with languages that are well represented online; they perform very well in English and other high resource languages, but poorly in low resource languages.[10] Though automated translation cannot yet be used reliably in clinical settings, the rate of development of AI language models makes it a reasonable future possibility. While we wait for structural improvements to improve access to human interpreters, we can utilize AI, with its low cost, lack of wait time, and fast interpretation, to reduce barriers and improve the accessibility of interpretation.

THE TASK AT HAND

Research

First, software engineers, clinical researchers, and research foundations must collaborate to advance and validate this technology. The most immediate need is to improve large language models for low-resource languages. Techniques are already being developed to improve performance for underrepresented languages, and we hope to soon see machine translation available for a wide range of languages.[10] Early studies have demonstrated that the major flaw of AI interpretation is accuracy, which has obvious clinical implications.[9] Before utilization of these technologies, clinician-researchers must evaluate their accuracy, acceptability across languages, and safety for use with patient data in clinical contexts. The National Institute of Health and Agency for Healthcare Research and Quality can catalyze this progress by creating financial incentives for those conducting research in AI interpretation and special emphasis notices.

Privacy

Ethicists, software engineers, and policymakers must consider the privacy risks of AI interpretation. AI language

interpretation could be developed without any risks for patient privacy by utilizing publicly available data instead of patient data and downloading the technology onto devices as an app. In this way, no patient data would be utilized. This model, not trained on patient data, may be less useful in clinical settings and less easily improved upon if unable to learn from patient conversations. In contrast, an adaptive model could be developed, which would store and learn from patient data. Protecting patient privacy when training AI models on patient data is a major area of debate. If data is to be collected, then patient consent would be a reasonable requirement, to be obtained before the technology is used to collect any identifying data. This may delay medical care and would be subject to the same barriers to consent that are generally experienced in research for patients with LEP.[11] However, the data could be de-identified using automated de-identification tools to prevent release of patient data in the case of security breaches and mitigate privacy risk. [12]

Policy

Once the technology has been clinically and legally validated, it must be regulated. The policies governing the privacy and security of Electronic Health Records (EHR) can guide the regulation of AI interpretation. The Department of Health and Human Services sets the standard for EHRs, with the HITECH Act governing how providers and hospitals can access patient information in EHRs.[13] A similar policy should be implemented to govern patient information accessed and stored by companies building language learning models for interpretation. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) currently provides certification for health centers meeting various EHR quality metrics.[14] Their accreditation requires a fee, which may be a barrier for clinics that primarily serve underinsured patients. However, these metrics for AI interpretation may still be useful to evaluate the technology's use in hospitals already undergoing JCAHO accreditation. The Office of the National Coordinator for Health Information Technology (ONC) is ideally positioned to evaluate the efficacy and security of AI interpretation programs. ONC recently released a security risk assessment (SRA) tool to help healthcare systems evaluate potential breaches of electronic protected health information (ePHI).[15] A similar risk assessment tool developed by ONC for the proposed technology would allow a standardized national evaluation of privacy risks.

Funding

We must consider the incentivization of these technologies. Venture capitalists and the private sector can invest in advancing this technology. While likely cheaper than current methods, hospitals/clinics would likely need to subscribe to these services. Mirroring the funding model for EHR implementation, one method to reduce the cost of AI interpretation services is to provide CMS-funded incentives for hospitals implementing these services.[16] These incentives were provided to institutions that utilized EHRs meeting certain standards of e-prescribing, patient-provider communication, and others.[17] Replicating this program for AI interpreter services would reiterate the quality standards set by policy, decreasing the cost of utilization only for hospitals that meet quality metrics.

Education

Lastly, providers must be trained in using the resources. Physicians and professional societies like the Society of General Internal Medicine and American College of Physicians can promote education and awareness through workshops and toolkits on these technologies and their use. Physicians can also promote education about these resources at their institutions.

CONCLUSION

AI advancements in medical technology are inevitable. AI applications are being developed at a breakneck pace, and automated translation is a timely solution to address the widespread underutilization of interpreter services. Inadequate interpreter use negatively affects healthcare outcomes in our most marginalized populations. Clinicians should be more than consumers, collaborating with medical ethicists, policymakers, and software engineers to ensure the safe and effective use of this rapidly evolving technology and advance health equity for millions.

Corresponding Author: *Leen Bakdash, BA; , Emory University School of Medicine, Atlanta, GA, 30322, USA (e-mail: lbakdas@emory.edu).*

Declarations

Conflict of Interest *There are no conflicts of interest for the authors.*

REFERENCES

1. **Jacobs B, Ryan AM, Henrichs KS, Weiss BD.** Medical Interpreters in Outpatient Practice. Ann Fam Med. 2018;16(1):70-76. doi: https://doi.org/10.1370/afm.2154.
2. **Karliner LS, Jacobs EA, Chen AH, Mutha S.** Do professional interpreters improve clinical care for patients with limited English proficiency? A systematic review of the literature. Health Serv Res. 2007;42(2):727-54. doi: https://doi.org/10.1111/j.1475-6773.2006.00629.x.
3. **López L, Rodriguez F, Huerta D, Soukup J, Hicks L.** Use of interpreters by physicians for hospitalized limited English proficient patients and its impact on patient outcomes. J Gen Intern Med. 2015;30(6):783-789. doi:https://doi.org/10.1007/s11606-015-3213-x
4. **Greene NE, Fuentes-Juárez BN, Sabatini CS.** Access to Orthopaedic Care for Spanish-Speaking Patients in California. J Bone Joint Surg Am. 2019;101(18):e95. doi:https://doi.org/10.2106/JBJS.18.01080

5. **Rajbhandari P, Glick AF, Brown MF, VanGeest J.** Linguistic Services for Hospitalized Children With Non-English Language Preference: A PRIS Network Survey. Hosp Pediatr. 2023;13(3):191-203. doi:https://doi.org/10.1542/hpeds.2022-006906
6. **Diamond LC, Schenker Y, Curry L, Bradley EH, Fernandez A.** Getting by: underuse of interpreters by resident physicians. J Gen Intern Med. 2009;24(2):256-262. doi:https://doi.org/10.1007/s11606-008-0875-7
7. Modern Healthcare. Medical Interpreters Face High Demand, Low Reimbursement. 2023. https://www.modernhealthcare.com/safety-quality/medical-interpreters-demand-covid-19-pandemic-recruiting-training-amn. Accessed 15 Jul 2023.
8. **Heath M, Hvass AMF, Wejse CM.** Interpreter services and effect on healthcare - a systematic review of the impact of different types of interpreters on patient outcome. J Migr Health. 2023;7:100162. doi: https://doi.org/10.1016/j.jmh.2023.100162.
9. **Panayiotou A, Gardner A, Williams S, et al.** Language Translation Apps in Health Care Settings: Expert Opinion. JMIR Mhealth Uhealth. 2019;7(4):e11316. Published 2019 Apr 9. doi:https://doi.org/10.2196/11316
10. **Jiao W, Wang W, Huang JT, Wang, X, Tu Z** (2023) Is ChatGPT a good translator? A preliminary study. arXiv preprint arXiv:2301.08745.
11. **Zeidan AJ, Smith M, Leff R, Cordone A, Moran TP, Brackett A, Agrawal P** (2023). Limited English Proficiency as a Barrier to Inclusion in Emergency Medicine-Based Clinical Stroke Research. J Immigr Minor Health, 25(1), 181–189. https://doi.org/10.1007/s10903-022-01368-y
12. **Cohn I, Laish I, Beryozkin G, Li G, Shafran I, Szpektor I, ... Matias Y** (2019). Audio de-identification: A new entity recognition task. *arXiv preprint* arXiv:1903.07037
13. TITLE XIII—HEALTH INFORMATION TECHNOLOGY, Pub. L. No. 201, § 13001, 42 USC (2009).
14. Joint Commission Accreditation | The Joint Commission. Accessed October 4, 2023. https://www.jointcommission.org/what-we-offer/accreditation/. Accessed 4 Oct 2023.
15. HealthITSecurity. ONC, OCR Release Security Risk Assessment Tool Version 3.4. HealthITSecurity, 2023. https://healthitsecurity.com/news/onc-ocr-release-security-risk-assessment-tool-version-3.4. Accessed 4 Oct 2023.
16. Centers for Medicare & Medicaid Services (CMS), HHS. Medicare and Medicaid Programs; Electronic Health Record Incentive Program. Final Rule. Federal Register 75, no. 144 (July 28, 2010): 44313–588.
17. Federal Register. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program, 2020. https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification. Accessed 4 Oct 2023.