Detecting professional interpreter use among patients with limited English proficiency: Derivation and validation study

Jalal Soleimani1, Alberto Marquez2, Sawsan Fathma1, Timothy J Weister2 and Amelia K Barwise3

Abstract
Objectives: The objective of this derivation and validation study was to develop and validate a search strategy algorithm to detect patients who used professional interpreter services.

Methods: We identified all adults who had at least one intensive care unit admission during their hospital stay across the Mayo Clinic Enterprise between 1 January 2015 and 30 June 2020. Three random subsets of 100 patients were extracted from 60,268 patients to develop the search strategy algorithm. Two physician reviewers conducted gold standard manual chart review and any discrepancies were resolved by a third reviewer. These results were compared with the search strategy algorithm each time it was refined. Sensitivity and specificity were calculated during each phase by comparing the search strategy results to the reference gold standard for both derivation cohorts and the final validation cohort.

Results: The first search strategy resulted in a sensitivity of 100% and a specificity of 89%. The second revised search strategy achieved a sensitivity of 100% and a specificity of 87%. The final version of the search strategy was applied to the validation subset and sensitivity and specificity were 100% and 89%, respectively.

Conclusion: We derived and validated a search strategy algorithm to assess interpreter use among hospitalized patients. Using a search strategy algorithm with high sensitivity and specificity can reduce the time required to abstract data from the electronic medical records compared with manual data abstraction.

Keywords
Implementation and deployment, electronic health record and systems, intensive care, data validation and derivation

Date received: 16 July 2021; accepted: 13 April 2022

Introduction
The number of residents in the United States with Limited English Proficiency (LEP) has grown in recent decades.1 According to the 2017 US Census, more than 64 million people aged 5 years and older speak a language other than English at home, and more than 25 million of the US population are classified as “speaking English less than very well” or having LEP.2,3 Furthermore, approximately 60,000 to 85,000 patients travel to prestigious medical centers in the United States every year for treatments not available in their native countries and frequently these patients also have LEP.4,5

Although there is a federal mandate for language services deployment when patients with LEP navigate the healthcare system, interpreters are underused. Negative health outcomes and suboptimal healthcare quality have been documented for patients with LEP.6–14 However, integrating medical interpreting standards into clinical practice has been challenging.15–17 Research demonstrates the benefits of using interpreter services in clinical practice. These include improved communication, improved patient and family satisfaction, and improved adherence to treatment plans.

1 Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, MN, USA
2 Anesthesia Clinical Research Unit (ACRU), Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, MN, USA
3 Division of Pulmonary and Critical Care Medicine, Mayo Clinic, Rochester, MN, USA

Corresponding author:
Amelia Barwise, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
Email: barwise.amelia@mayo.edu
Other beneficial healthcare outcomes include the reduction of complications and length of hospital stay, as well as increased use of preventive healthcare services.\textsuperscript{11–13,18–21}

Despite evidence of the advantages of using interpreters, physicians still try to “get-by” with the use of their own limited language skills or those of family members as interpreters.\textsuperscript{15,22–24} Physician behaviors may be influenced by the lack of access and availability of professional interpreter services based on organizational structures and resources.\textsuperscript{25} Evidence suggests that professional interpreters provide higher quality interpretation than family members and therefore should be used.\textsuperscript{26,27}

It is important to understand the use of interpreters in the clinical setting to be able to optimize systems to increase access to interpreters for patients who need one. This knowledge will be a useful step toward sustained change to address disparities among those with LEP.

The adoption of electronic health records (EHRs) in place of traditional paper charts has provided unprecedented amounts of information (“Big data”) that allows researchers to evaluate larger cohorts of patients than traditional research approaches.\textsuperscript{28,29} Harnessing the potential of the EHR and the vast amounts of data within is challenging but developing and applying automated search strategies is a helpful approach. By improving validated algorithms to identify when an interpreter is utilized, we can support quality improvement initiatives, enhance clinical practice, and improve our research accuracy and efficiency.

**Methods**

**Setting and study design**

This is a derivation and validation cohort study which was approved by the Mayo Clinic Institutional Review Board (IRB). The IRB reviewed the study and the procedures and deemed the study exempt. This study did not include patient contact. We only reviewed and included the EHR of patients who had provided research authorization in accordance with Minnesota state statutes if applicable.\textsuperscript{30}

**Study participants**

We included (1) consecutive adult patients (≥18 years), (2) with research authorization if the state required it, (3) admitted to the intensive care units (ICU) across the Mayo Clinic enterprise (Minnesota, Wisconsin, Arizona, and Florida), (4) between 1 January 2015 and 30 June 2020 (5½ years). The search strategies to identify when an interpreter was used incorporated the interpreter flag indicating an interpreter was needed and a combination of a preferred language other than English and interpreter flow sheets. Patients who did not need an interpreter or whose primary language was listed as English were randomly selected for conducting the derivation and validation. If a patient was admitted several times during this period, only the first admission including an ICU stay was included.

Our search strategy was designed to detect those who used professional interpreters including phone, video, or in-person interpreters. Since this study focused on the use of professional interpreters, we excluded encounters in which interpretation by family, friends, and healthcare team members occurred. Three random subsets of 100 patients were used for the derivation and validation phases (see Figure 1). No sample size calculation was performed. This size cohort has been used in previous similar derivation and validation studies we have conducted and is considered the accepted and standard size for this type of work.\textsuperscript{28,29}

**Manual data extraction strategies (reference gold standard)**

Before developing the automated search strategy and to formulate the gold standard for interpreter use, members of the study team (J.S., A.B., and S.F.) manually reviewed the EHRs of random patients classified as having a preferred language other than English and those patients whose EHR indicated that an interpreter was required. The study team examined the different ways that the “interpreter used” was documented by healthcare providers in the EHR including within flow sheets, progress, and encounter notes.

During the derivation and validation process, the reference gold standard comparison involved EHR manual review by two physician researcher reviewers. They examined the electronic medical record between the specified admission dates to assess whether an interpreter was used. They did this by reviewing the medical note documentation and the flow sheets as well as the patient-provided information. Two reviewers (J.S. and S.F.) conducted gold standard electronic medical record reviews and any disagreements were resolved by a third reviewer (A.B.). We have used this approach in other studies in which manual chart review is the gold standard.\textsuperscript{31}

**Automated electronic search strategy**

Although “interpreter flags” appear in the patient-provided information of the EHR, this does not specify whether an interpreter was actually used, simply that an interpreter was needed or that the primary language was not English. The search strategy algorithm was developed in several stages. The first derivation used patient-provided information in the EHR specifically “Interpreter Indicator” (which indicates if a patient needs Interpreter Services) and “Preferred Language.” We also used applicable interpreter flow sheets from the EHR. The second derivation used modifications to the flowsheet search and interpreter needed indicator. The preferred language was not included in the second search strategy. We used the interpreter indicator search in the ICU DataMart. DataMart is an extensive data warehouse...
containing a near-real-time normalized replica of Mayo Clinic’s EHR. We accessed the DataMart warehouse and searched the data by using JMP software. DataMart contains patient demographic characteristics, diagnoses, laboratory results, and clinical flow sheets, gathered from various sources within the institution. The data within DataMart has been validated and is reliable.

We also used an interpreter indicator search in Mayo Clinic’s Advanced Cohort Explorer (ACE), an electronic retrieval query database within Mayo Clinic’s Unified Data Platform (UDP). ACE is a powerful web-based software toolset that enables the search of the EHR by specific text phrases or terms in specific parts of the clinical notes. All data extracted by ACE can be exported to Excel to enable further statistical analysis. Each subset sample contained 100 patients who were randomly selected from DataMart. (Figure 1) These cohorts consisted of both those likely to use interpreter services and those not likely to use interpreter services (see Figure 1). In order for patients to be categorized as “Interpreter used” using the automated search strategy, the patients needed to have both an interpreter indicator “Yes” within their demographic information as well as an applicable flowsheet articulating that an interpreter was used at the time of the encounter we were examining.

Following the manual review of each of the derivation subsets, A.M. refined the electronic search strategy through several iterations of evaluation and refinement of the electronic search algorithm in the derivation cohorts. Once the search strategy was assessed and modified in the derivation cohorts, it was validated in the third cohort.

**Statistical analysis**

An overall percent agreement between the electronic search algorithm and the manual EHR review was calculated. Sensitivity and specificity were calculated by comparing the results to the reference gold standard for each derivation and validation subsets. JMP statistical software (version 10.0.0; SAS Institute Inc., Cary, NC) was used for all analyses. Furthermore, we calculated a kappa statistic to assess the
agreement between reviewers 1 and 2 conducting gold standard manual EHR reviews.

**Results**

The cohort of those with hospital admissions and at least one ICU admission across the Mayo Clinic Enterprise between 1 January 2015 and 30 June 2020 was 60,268 admissions.

**Primary outcome**

The primary outcome of the study was to derive and validate a search strategy to identify when an interpreter was used by patients with LEP by assessing sensitivity and specificity with classification performance.

In the first derivation cohort, our manual review resulted in a sensitivity of 100% (95% confidence interval [CI] 93.5–100) and a specificity of 89% (95% CI 75.9–96.3). The supervised algorithm was used for the second derivation subset. The second derivation achieved a sensitivity of 100% (95% CI 92.6–100) and a specificity of 87% (95% CI 74.2–94.4). The final version of the search strategy for interpreter use was applied to the validation subset and this achieved a sensitivity of 100% and a specificity of 89% (95% CI 75.9–96.3) (Table 1). Kappa agreement between reviewers with the first derivation was 0.90, and with the second derivation and validation were 0.86, and 0.88, respectively. These can be construed as near perfect agreement. Despite the high sensitivity, which remained at 100% in the final cohort, the specificity of the validation subset did not increase beyond 89%, meaning there is a small possibility of false-positive results. The differences between the first, second, and third cohorts are not statistically significant.

**Strengths and limitations**

Strengths of the study include the following. We used robust approaches during the derivation and validation processes with an experienced team. We did gold standard EHR review by having two reviewers and a third reviewer to resolve discrepancies. Our kappa statistics are 0.86–0.90 and therefore demonstrate near perfect agreement. Despite the high sensitivity, which remained at 100% in the final cohort, the specificity of the validation subset did not increase beyond 89%, meaning there is a small possibility of false-positive results.

Based on estimated times required to conduct manual EHR review during this study, we believe our search strategy will prove very useful for future identification of patients who used an interpreter across our healthcare enterprise. Furthermore, it will provide useful foundational knowledge to build more real-time algorithms to identify patients who would benefit from an interpreter. Other institutions that use EPIC may also be able to leverage this approach to identify patients who used or need an interpreter. Based on 2019 data, EPIC has a market share of almost one-third of acute care multi-specialty hospital EHRs. It is becoming increasingly dominant and being adopted by healthcare systems nationwide.

The study has some limitations. It is worth noting that we were able to leverage institutional software and specific data sets (ACE and DataMart) to conduct our study and these may not be available in all institutions. Therefore, external validation of this search strategy is needed especially if different electronic infrastructures exist in those institutions. Other institutions with diverse computational infrastructure may need to modify our search strategy within their patient data sets and EHRs.

We wanted to focus on professional interpreter use in this study as there is evidence that unless absolutely necessary, or in an emergency situation, interpretation by family and friends should be avoided. Using family and friends as interpreters threatens the accuracy and completeness of interpretation. This can lead to vital information being inadvertently or deliberately omitted or misinterpreted. Using clinical team members who do not have sufficient language skills can also lead to information being inaccurately interpreted compromising communication.

In comparison to other papers describing validation and derivation of automated digital algorithms in which sensitivity varied from 77% to 100%, and specificity ranged from 91% to 99.7%, the results of this study are favorable.

**Table 1. Sensitivity and specificity of subset groups.**

| Subset                  | Sensitivity (%) (95% CI) | Specificity (%) (95% CI) |
|-------------------------|--------------------------|--------------------------|
| 1st derivation cohort   | 100% (93.5–100)          | 89% (75.9–96.3)          |
| 2nd derivation cohort   | 100% (92.6–100)          | 87% (74.2–94.4)          |
| Validation cohort       | 100% (93.5–100)          | 89% (75.9–96.3)          |

CI: confidence interval.
Amra et al.’s \cite{38} automated electronic search algorithm sensitivity ranged from 94\% to 97\% and specificity ranged from 93\% to 99\%, while Singh et al. achieved a sensitivity between 91\% to 100\% and specificity from 98\% to 100\%.\cite{33} Others automated electronic queries demonstrated sensitivity ranging from 77\% to 100\% and specificity $\geq 96\%$.\cite{34}

The purpose of this study was to develop an electronic automated search algorithm to accurately and reliably detect patients who used professional interpreter services. This could then be used to reduce the time and effort needed to identify those who used interpreter services in a large data set. It is important to better understand professional interpreter use to foster strategies to increase timely and appropriate interpreter use in near real-time. Interpreter use can mitigate disparities experienced by patients with LEP.\cite{31-36} Interpreters can reduce cultural, language, and literacy barriers by improving communication between patients and clinicians.\cite{48}

**Conclusion**

We have successfully derived and validated an “Interpreter used” search strategy that identifies if a patient used a professional in-person interpreter, video-linked interpreter, or telephone interpreter. It can be deployed in our enterprise EHR demonstrating a sensitivity of 100\% and a specificity of 89\%. This method of electronic data extraction by an automated algorithm through institutional software connected to EHR is accurate, time-saving, and cost-effective.

**Authors’ note**

This work was performed at Mayo Clinic, Rochester, MN, USA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Mayo Clinic, Rochester.

**Author contributions**

A.B. as a principal investigator designed the study, and did data interpretation. J.S. helped design the study, performed data acquisition, data interpretation, and manuscript writing. S.F. acted as the second reviewer during gold standard EHR manual review doing data interpretation. A.M. designed the search strategy algorithm and provided statistical analysis. T.W. helped design the study, designed the search strategy algorithm, and provided statistical analysis. All authors drafted the manuscript and/or revised it critically for important intellectual content and gave final approval of the manuscript with all the accountability herein. All authors read and approved the final manuscript.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethical approval**

Ethical approval for this study was obtained from the Mayo Clinic Institutional Review Board (IRB#: 19-009625).

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Division of Critical Care Medicine with no direct financial support.

**Informed consent**

Written consent was waived by the Institutional Review Board/Ethics Committee.

**ORCID iDs**

Jalal Soleimani https://orcid.org/0000-0002-1066-800X

Amelia K. Barwise https://orcid.org/0000-0001-9267-1298

**References**

1. Pandya C, McHugh M and Batalova J. Limited English proficient individuals in the United States: number, share, growth, and linguistic diversity. LEP data brief. Migration Policy Institute, Washington, DC, 2011.
2. Bureau UC. Selected social characteristics in the United States: 2011-2015 American Community Survey 5-year estimates, https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2015/5-year.html (2015, accessed 2 September 2020).
3. Patel AT, Lee BR, Donegan R, et al. Length of stay for patients with limited English proficiency in pediatric urgent care. Clin Pediatr 2020; 59(4-5): 421–428.
4. Ehrbeck T, Guevara C and Mango PD. Mapping the market for medical travel. The Mckinsey Quarterly, 2008; 11, https://www.lindsayresnick.com/Resource_Links/MedicalTravel.pdf
5. Van Dusen A. U.S. hospitals worth the trip. Forbes.com, 29 May 2008, https://www.forbes.com/2008/05/25/health-hospitals-care-forbeslife-cx_avd_outsourcing08_0529healthoutsourcing.html?sh=47d426ecf152e1
6. Lu T and Myerson R. Disparities in health insurance coverage and access to care by English language proficiency in the USA, 2006–2016. J Gen Intern Med 2020; 35(5): 1490–1497.
7. Ridgeway JL, Njeru JW, Breitkopf CR, et al. Closing the gap: participatory formative evaluation to reduce cancer screening disparities among patients with limited English proficiency. J Cancer Educ 2020; 36: 795–803.
8. John-Baptiste A, Naglie G, Tomlinson G, et al. The effect of English language proficiency on length of stay and in-hospital mortality. J Gen Intern Med 2004; 19(3): 221–228.
9. Karliner LS, Kim SE, Meltzer DO, et al. Influence of language barriers on outcomes of hospital care for general medicine inpatients. J Hosp Med 2010; 5(5): 276–282.
10. Betancourt JR, Renfrew MR, Green AR, et al. Improving patient safety systems for patients with limited English proficiency: a guide for hospitals. Rockville, MD: Agency for Healthcare Research and Quality, 2012.
11. Flores G. The impact of medical interpreter services on the quality of health care: a systematic review. Med Care Res Rev 2005; 62(3): 255–299.
12. Jacobs EA, Shepard DS, Suaya JA, et al. Overcoming language barriers in health care: costs and benefits of interpreter services. Am J Public Health 2004; 94(5): 866–869.
13. Bagchi AD, Dale S, Verbitsky-Savitz N, et al. Examining effectiveness of medical interpreters in emergency departments for Spanish-speaking patients with limited English proficiency: results of a randomized controlled trial. *Ann Emerg Med* 2011; 57(3): 248–256.

14. Flores G, Abreu M, Barone CP, et al. Errors of medical interpretation and their potential clinical consequences: a comparison of professional versus ad hoc versus no interpreters. *Ann Emerg Med* 2012; 60(5): 545–553.

15. Diamond LC, Schenker Y, Curry L, et al. Getting by: underuse of interpreters by resident physicians. *J Gen Intern Med* 2009; 24(2): 256–262.

16. Jacobs EA, Diamond LC and Stevak L. The importance of teaching clinicians when and how to work with interpreters. *Patient Educ Couns* 2010; 78(2): 149–153.

17. Lor M, Bowers BJ and Jacobs EA. Navigating challenges of medical interpreting standards and expectations of patients and health care professionals: the interpreter perspective. *Qual Health Res* 2019; 29(6): 820–832.

18. Barwise AK, Nyquist CA, Espinoza Suarez NR, et al. End of life decision making for ICU patients with limited English proficiency: a qualitative study of healthcare team insights. *Crit Care Med* 2019; 47(10): 1380–1387.

19. Jacobs EA, Lauderdale DS, Meltzer D, et al. Impact of interpreter services on delivery of health care to limited-English-proficient patients. *J Gen Intern Med* 2001; 16(7): 468–474.

20. Chen Y, Criss SD, Watson TR, et al. Cost and utilization of lung cancer end-of-life care among racial-ethnic minority groups in the United States. *Oncoologist* 2020; 25(1): e120–e129.

21. Espinoza Suarez NR, Urtecho M, Nyquist CA, et al. Consequences of suboptimal communication for patients with limited English proficiency in the intensive care unit and suggestions for a way forward: a qualitative study of healthcare team perceptions. *J Crit Care* 2021; 61: 247–251.

22. Parsons JA, Baker NA, Smith-Gorvie T, et al. To “get by” or “get help”? A qualitative study of physicians’ challenges and dilemmas when patients have limited English proficiency. *BMJ Open* 2014; 4(6): e004613.

23. Kale E and Syed HR. Language barriers and the use of interpreters in the public health services. *Patient Educ Couns* 2010; 81(2): 187–191.

24. López L, Rodriguez F, Huerta D, et al. 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.

25. Hsieh E. Not just “getting by”: factors influencing providers’ choice of interpreters. *J Gen Intern Med* 2015; 30(1): 75–82.

26. Gray B, Hilder J and Stubbe M. How to use interpreters in general practice: the development of a New Zealand toolkit. *Journal of Primary Health Care* 2012; 4(1): 52–61.

27. Gray B, Hilder J and Donaldson H. Why do we not use trained interpreters for all patients with limited English proficiency? Is there a place for using family members? *Australian Journal of Primary Health* 2011; 17(3): 240–249.

28. Amra S, O’Horo JC, Singh TD, et al. Derivation and validation of the automated search algorithms to identify cognitive impairment and dementia in electronic health records. *J Crit Care* 2017; 37: 202–205.

29. Guru PK, Singh TD, Passe M, et al. Derivation and validation of a search algorithm to retrospectively identify CRRT initiation in the ECMO patients. *Appl Clin Inform* 2016; 7(2): 596–603.

30. Statutes 2006 section Minnesota, 144.335. https://www.revisor.mn.gov/statutes/2006/cite/144335 (2006, accessed 13 July 2020).

31. Soleimani J, Pinevich Y, Barwise AK, et al. Feasibility and reliability testing of manual electronic health record reviews as a tool for timely identification of diagnostic error in patients at risk. *Appl Clin Inform* 2020; 11(3): 474–482.

32. Herasevich V, Pickering BW, Dong Y, et al. Informatics infrastructure for syndrome surveillance, decision support, reporting, and modeling of critical illness. *Mayo Clin Proc* 2010; 85(3): 247–254.

33. Singh B, Singh A, Ahmed A, et al. Derivation and validation of automated electronic search strategies to extract Charlson comorbidities from electronic medical records. *Mayo Clin Proc* 2012; 87(9): 817–824.

34. Alsara A, Warner DO, Li G, et al. Derivation and validation of automated electronic search strategies to identify pertinent risk factors for postoperative acute lung injury. *Mayo Clin Proc* 2011; 86(5): 382–388.

35. Viera AJ and Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; 37(5): 360–363.

36. Centers for Medicare Medicaid Services (CMS) HHS. Medicare and Medicaid programs; electronic health record incentive program—final rule. *Fed Reg* 2010; 75(144): 44313–44588.

37. Sittig DF and Singh H. Electronic health records and national patient-safety goals. *N Engl J Med* 2012; 367(19): 1854.

38. Kruse CS, Mileski M, Alaysiev V, et al. Adoption factors associated with electronic health record among long-term care facilities: a systematic review. *BMJ Open* 2015; 5(1): e006615.

39. Milinovich A and Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med* 2018; 6(3): 42.

40. Johnson RJ. A comprehensive review of an electronic health record system soon to assume market ascendency: EPIC. *J Healthc Commun* 2016; 1(4): 36.

41. Smischney NJ, Velagapudi VM, Onigkje JA, et al. Derivation and validation of a search algorithm to retrospectively identify mechanical ventilation initiation in the intensive care unit. *BMC Med Inform Decim Making* 2014; 14(1): 1–6.

42. US Hospital EMR and market share, 2020, https://klasresearch.com/report/us-hospital-emr-market-share-2020/1616 (accessed 22 March 2021).

43. Rosenberg E, Seller R and Leanza Y. Through interpreters’ eyes: comparing roles of professional and family interpreters. *Patient Educ Couns* 2008; 70(1): 87–93.

44. Bernstein J, Bernstein E, Dave A, et al. Trained medical interpreters in the emergency department: effects on services, subsequent charges, and follow-up. *J Immigr Health* 2002; 4(4): 171–176.

45. Juckett G and Unger K. Appropriate use of medical interpreters. *Am Fam Phys* 2014; 90(7): 476–480.

46. Blake C. Ethical considerations in working with culturally diverse populations: the essential role of professional interpreters. *Bull Can Psychiat Assoc* 2003; 34: 21–23.

47. Rollins G. Translation, por favor. *Hosp Health Netw* 2002; 76(12): 46–501.

48. Cooper LA, Hill MN and Powe NR. Designing and evaluating interventions to eliminate racial and ethnic disparities in health care. *J Gen Intern Med* 2002; 17(6): 477–486.