Title

Inaccurate Language Interpretation and Its Clinical Significance in the Medical Encounters of Spanish-speaking Latinos

Journal

Medical Care, 53(11)

ISSN

0025-7079

Authors

Nápoles, Anna M
Santoyo-Olsson, Jasmine
Karliner, Leah S
et al.

Publication Date

2015-11-01

DOI

10.1097/mlr.0000000000000422

Peer reviewed

HHS Public Access

Author manuscript

Med Care. Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:

Med Care. 2015 November ; 53(11): 940–947. doi:10.1097/MLR.0000000000000422.

Inaccurate Language Interpretation and its Clinical Significance in the Medical Encounters of Spanish-speaking Latinos

Anna M. Nápoles, PhD, MPH[1], **Jasmine Santoyo-Olsson, MS**[1], **Leah S. Karliner, MD, MAS**[1], **Steven E. Gregorich, PhD**[1], and **Eliseo J. Pérez-Stable, MD**[1]

[1]Medical Effectiveness Research Center for Diverse Populations, Division of General Internal Medicine, Department of Medicine, University of California San Francisco (UCSF), 3333 California St, Suite 335, San Francisco, CA 94118

Abstract

Background—Limited English-proficient (LEP) patients suffer poorer quality of care and outcomes. Interpreters can ameliorate these disparities; however, evidence is lacking on the quality of different interpretation modes.

Objective—Compare accuracy of interpretation for in-person professional (IP), professional videoconferencing (VC), and ad hoc interpretation (AH).

Design—Cross-sectional study of transcribed audiotaped primary care visits

Subjects—32 Spanish-speaking Latino patients; 14 clinicians

Measures—Independent coding of transcripts by four coders (two were internists) for accurate and inaccurate interpretation instances. Unit of analysis was a segment of continuous speech or text unit (TU). Two internists independently verified inaccurate interpretation instances and rated their clinical significance as clinically insignificant, mildly, moderately or highly clinically significant.

Results—Accurate interpretation made up 70% of total coded TUs and inaccurate interpretation (errors) made up 30%. Inaccurate interpretation occurred at twice the rate for AH (54% of coded TUs) versus IP (25%) and VC (23%) interpretation, due to more errors of *omission* ($p<0.001$) and *answers for patient or clinician* ($p<0.001$). Mean number of errors per visit was 27, with 7.1% of errors rated as moderately/highly clinically significant. In adjusted models, the odds of inaccurate interpretation were lower for IP (OR = −1.25, 95% CI −1.56, −0.95) and VC (OR = −1.05; 95% CI −1.26, −0.84) than for AH interpreted visits; the odds of a moderately/highly clinically

Please address correspondence to: Eliseo J Pérez-Stable, MD, University of California, San Francisco, 1545 Divisadero Street, Box 0320, San Francisco, CA 94143-0320, Phone: (415) 514-8662, Fax: (415) 514-8666, eliseo.perez-stable@ucsf.edu.

Complete Author Information

Anna M. Nápoles, PhD, MPH[1], University of California San Francisco, 3333 California St, Suite 335, San Francisco, CA 94118, (415) 476-6290 (t), (415) 502-8291 (f), anapoles@ucsf.edu

Jasmine Santoyo-Olsson, MS, University of California San Francisco, 3333 California St, Suite 335, San Francisco, CA 94118, (415) 514-3355 (t), (415) 502-8291 (f), jasmine.santoyo-olsson@ucsf.edu

Leah S. Karliner, MD, MAS, University of California San Francisco, 1545 Divisadero St, San Francisco, CA 94115, (415) 885-7893 (t), (415) 514-8666 (f), leah.karliner@ucsf.edu

Steven E. Gregorich, PhD, University of California San Francisco, 3333 California St, Suite 335, San Francisco, CA 94118, (415) 502-5235 (t), (415) 502-8291 (f), steven.gregorich@ucsf.edu

Eliseo J. Pérez-Stable, MD, University of California, San Francisco, 1545 Divisadero Street, Box 0320, San Francisco, CA 94143-0320, Phone: (415) 514-8662, Fax: (415) 514-8666, eliseo.perez-stable@.ucsf.edu

significant error were lower for IP (OR = −0.06; 95% CI −1.05, 0.92) than for AH interpreted visits.

Conclusions—Inaccurate language interpretation in medical encounters is common and more frequent when untrained interpreters are used compared to professional in-person or via videoconferencing. Professional video conferencing interpretation may increase access to higher quality medical interpretation services.

Keywords

Language interpretation; quality of care; Latino/Hispanic; primary care

Introduction

In 2011, 25.3 million limited English proficient (LEP) individuals resided in the U.S., constituting nine percent of the population and an 81% increase in the LEP population since 1990.[1] Federal law requires that health care organizations receiving federal funds provide interpretation services by bilingual-bicultural staff or professional interpreters for LEP patients at no cost.[2, 3] However, health systems are challenged by the volume of visits, diversity of languages, high costs, and lack of reimbursement mechanisms.[4] When professional interpreters are unavailable, patients and clinicians rely on their limited language abilities or ad hoc interpreters with no interpreter training, potentially compromising the quality of communication.[5–7]

Compared to English-speaking patients, LEP patients suffer poorer quality of care. [8, 9] A study of six hospitals found that of all adverse events, those occurring among LEP patients were more likely than those experienced by English-speaking patients to involve some physical harm, to be more severe, and to be due to communication errors. [10] Interpreter services can ameliorate these disparities in quality and outcomes of care,[11] including increased receipt of preventive services,[12] reduced emergency department visits,[13] better understanding of physician recommendations,[14] a more positive health outlook,[15] greater patient satisfaction[16, 17] and more patient-centered care[8, 18] Although evidence is mixed, interpreter services may lower health care costs among LEP patients.[19, 20]

Ad hoc interpretation by untrained persons continues to be the default mode in many health care settings.[5–7] Use of ad hoc interpreters often results in inferior translation,[21, 22] communication errors with potential clinical consequences,[23–25] worse patient comprehension and clinical outcomes, less patient satisfaction,[8, 11] and reduced transmission of information and small talk.[26] However, relationships between interpretation mode, frequency of errors, and outcomes have not been adequately described. Five studies have examined these issues through rigorous coding of audiotaped transcripts of interpreted medical encounters, however, four involved ≤ 16 encounters and 2 used scripted, not actual encounters.[23, 24, 26–28] None have examined videoconferencing interpretation. Our study extends this work by examining accuracy of interpretation across three modes (in-person professional interpreter, remote professional interpreter via videoconferencing, and ad hoc interpreter) and the clinical significance of interpretation errors, in a larger sample of adult primary care patients. The aims of this study were to compare the frequency of inaccurate

and accurate interpretation (primary outcomes) and the clinical significance of inaccurate (errors of) interpretation (secondary outcome) across modes.

Methods

Setting and sample

We recruited a consecutive sample of language-interpreted medical encounters of Spanish-speaking adult patients from a public hospital internal medicine clinic in Northern California from June to September 2005. We varied the days of the week that research associates were on-site and captured approximately 75% of eligible Spanish language LEP visits on the days a research associate was present. Inclusion criteria were: 1) medical visit with a Spanish-speaking monolingual patient; 2) clinician who was an attending physician, third year resident, nurse practitioner, or physician assistant; and 3) presence of a third-party to facilitate oral language interpretation between clinician and patient. The clinic provides care to primarily low-income persons; over 75% are race/ethnic minorities. Approximately 60% of outpatient visits involve LEP patients. Professional interpreters had completed a formal 40-hour training course and passed an oral and written proficiency examination in English and language of interpretation. The same professional interpreter staff provided videoconferencing and in-person interpretation.

Procedures

A trained bilingual-bicultural research associate obtained written informed consent. Medical encounters were audio recorded and transcribed. Immediately post-visit, the research associate administered a brief patient survey that asked about quality of interpretation and communication. Clinicians completed a brief self-administered survey regarding the same visit shortly after the encounter. Patients received $10 for participating. The public hospital and academic health center institutional review boards approved the study.

Survey Measures

Type of interpreter mode, the primary predictor of interest, was assessed by asking patients: “What type of interpreter did you use?” with response options of none, we did not need one; none, but we probably needed one; family member/friend; nurse or clerk who is not a professional interpreter; professional interpreter in-person; professional interpreter on a video screen; or professional interpreter on the telephone (none of the sampled visits used telephone interpretation). Mode of interpretation was classified into 1 of 3 categories: professional interpreter in-person, professional interpreter on a video screen, or ad hoc interpreter (if answered family member/friend or nurse/clerk).

Patients were asked about their age in years, sex, level of education (none, grades 1–5, grades 6–11, high school diploma, some college, or college graduate or higher), health insurance (none, Medicaid only, Medicare only, Medicare and Medicaid, Medicare and private, or private only), place of birth, marital status (married/living with partner or not married), employment status (employed, homemaker or unemployed), self-rated health (“In general, would you say your health is...” with response options of poor, fair, good, very good, or excellent), whether they had a chronic medical condition that required ongoing care

(yes or no), and the number of times including the present visit, that the patient had seen this clinician.

Clinicians were asked their age in years, sex, race/ethnicity (White, African American, American Indian, Asian, Pacific Islander, Latino, or other), specialty or level of training (family medicine or internal medicine physician, nurse practitioner, or physician assistant), and whether they had ever received training on working with interpreters (none, 1 lecture or workshop, several lectures, a course).

We developed parallel patient and clinician versions of a 4-item Quality of Interpretation Scale, based on the literature,[18, 29–31] input from internists and behavioral scientists, and pretesting with Spanish-speaking patients. Based on these sources, we defined high quality interpretation as listening and conveying accurate information and responsiveness to potential barriers to communication. Thus, items asked patients and clinicians to rate how well the interpreter listened to what they had to say, explained what the patient/clinician said to the clinician/patient, helped the patient/clinician understand what the clinician/patient said, and overall quality of interpretation for that encounter. Response options were: poor, fair, good, very good, or excellent.

Measures Based on Coding of Audiotape Transcripts

Coding procedures and measures derived from the coding are described next. Four bilingual, experienced qualitative data coders independently analyzed the transcripts; two were behavioral scientists and two were general internists (henceforth referred to as “clinician coders”). Coding employed a modified grounded theory approach [32, 33] in which a preliminary coding scheme was based on the literature, [22, 23, 28, 34, 35] but modifications and additions to these preliminary codes were extracted directly from the data. The four coders reached consensus on two final coding schemes applied in separate phases, one for coding instances of accurate and inaccurate interpretation and one for coding the clinical significance of inaccurate interpretation (errors). Coders were blinded to interpretation mode of encounters and all coding was adjudicated, listening to audiotapes as needed, until consensus was reached among the four coders.

Phase 1: Coding of accuracy of interpretation

The final coding scheme for accuracy of interpretation consisted of eight codes (primary outcome measures), two for accurate and six for inaccurate interpretation (Table 1). The two accurate interpretation codes were *accurate interpretation* (provides accurate interpretation of word/phrase uttered by clinician/patient) and *asks for clarification to ensure accurate interpretation* (asks additional question of clinician/patient that ensures accurate interpretation). The six inaccurate interpretation codes were: *additions* (adds word/phrase not uttered by clinician/patient), *substitutions* (substitutes word/phrase that differs from that uttered by clinician/patient that changes meaning of original statement), *answers for patient or clinician* (provides preemptive answer to question posed by clinician/patient rather than interpreting the question), *omission* (fails to interpret word/phrase uttered by clinician/patient), *editorializing* (adds their personal opinion to interpretation of word/phrase uttered by clinician/patient), and *false fluency* (interprets with an incorrect word/phrase). The final

accuracy of interpretation codes, their definitions, and examples of each code are presented in Table 1.

Five of these codes, *additions*, *substitutions*, *omission*, *editorializing*, and *false fluency*, were used in prior studies and constituted our preliminary coding scheme. [22, 23, 28, 34, 35] Based on the work of Laws B, et.al., [22] we added a code for *accurate interpretation*. Two additional codes emerged from our data: *asks for clarification to ensure accurate interpretation* and *answers for patient or clinician*.

In Phase 1 of coding, two behavioral scientists independently coded transcripts using the accuracy of interpretation coding scheme. The unit of analysis was an identifiable segment of continuous speech or text unit (TU). A new TU occurred when a speaker paused after a statement or a new speaker initiated; these ranged in length from a phrase to several paragraphs. Each TU was assigned a numeric value of 0 or 1, indicating the absence or presence of one of the eight accuracy of interpretation codes. In infrequent cases when more than one error of interpretation occurred in a TU, the most serious error code (deviated the most from the speaker's intended meaning) was assigned.

Phase 2: Coding of clinical significance of inaccurate interpretation, complexity of visit and quality of interpretation

In Phase 2 of the coding, the clinician coders independently applied the coding scheme for the clinical significance of inaccurate interpretation (secondary outcome), adapted from Gany, F., et al. which consisted of four mutually exclusive codes assessing the extent to which interpreter errors potentially affected clinical decision-making or outcomes (e.g., understanding of diagnostic, therapeutic or follow-up care): clinically insignificant, mildly clinically significant, moderately clinically significant, and highly clinically significant. [24] The coding scheme was identical to Gany's except that we dropped the category of "potentially life threatening" because errors of this severity were not observed in our data. Additionally, based on their clinical experience, after review of the visit transcripts, clinician coders independently rated the clinical complexity of visits as slightly complex, moderately complex, very complex, or extremely complex. Clinician coders met to review these ratings, adjudicating any differences until reaching consensus. Finally, for each visit, clinician coders rated the overall quality of interpretation, interpreter's ability to convey the intended meaning of clinician's statements, and interpreter's ability to convey the intended meaning of patient's statements. Response options were poor, fair, good, very good, or excellent.

Data analyses

For the clinician and patient survey data, the unit of analysis was the visit. Analyses were performed using SAS software, Version 9.2. We conducted psychometric analyses and confirmed the unidimensionality of the quality of interpretation scales using multi-trait scaling methods.[36] Scores for the Patient-rated Quality of Interpretation Scale (Cronbach's alpha = 0.85) and Clinician-rated Quality of Interpretation Scale (Cronbach's alpha=0.97) were computed as the mean of non-missing items. We computed least square means of the quality of interpretation ratings, with pairwise comparisons using Scheffe adjustment for

multiple comparisons, to examine differences in clinician, patient and clinician coder quality ratings by interpretation mode.

For the accuracy of interpretation codes, the unit of analysis was the TU. Coding was performed using QSR NVivo 8 software. Chi-square tests examined the frequency of the accuracy of interpretation codes by interpretation mode (Fisher's exact test was used with expected cell counts < 5). We used generalized estimating equations (GEE) to examine relationships between the odds of inaccurate interpretation and interpretation mode, adjusting for clinical complexity of visit, length of visit, patient age, patient sex, and nesting of patients within clinicians.

Descriptive statistics were used to examine the frequency and clinical significance of errors. We used GEE to examine relationships between the odds of a moderately/highly clinically significant interpreter error and interpretation mode, adjusting for clinical complexity of visit, length of visit, patient age, patient sex, and nesting of patients within clinicians.

Results

Patient and clinician characteristics

Of 32 encounters, 5 used professional in-person interpreters, 22 used professional interpreters via video-conferencing, and 5 used ad hoc interpreters (4 adult children and a nurse with no interpreter training). Encounters involved 32 unique patients and 14 clinicians, with 2.3 patients per clinician, on average. Patients' mean age was 53 years, most were women, almost all had less than a high school education, more than half had no health insurance, and most were born in Mexico (Table 2). About a fourth of visits were the first visit between the patient and clinician. Clinicians' mean age was 51 years, most were women, half were non-Latino Whites, almost all were internal medicine physicians, and half reported no training in working with interpreters.

Quality of language interpretation

The mean score on the Patient-rated Quality of Interpretation Scale was 4.0 (SD=0.80) (Table 3). Patient quality ratings were higher ($p < 0.05$) for professional video-conferencing interpretation (mean=4.3) than for professional in-person (mean=3.4) and ad hoc interpretation (mean=3.4). The mean score on the Clinician-rated Quality of Interpretation Scale was 3.4 (SD=0.80). There were no significant differences ($p = 0.70$) in clinicians' quality ratings by mode of interpretation.

Clinician coders rated the ad hoc interpreters' ability to convey the intended meaning of clinicians' statements as being of poorer quality than professional video-conferencing interpretation ($p < 0.05$).

Accuracy of interpretation codes

A total of 2,944 TUs were coded with one of eight accuracy of interpretation codes, 518 with professional in-person, 1,836 with professional videoconferencing, and 590 with ad hoc interpretation (Table 4). Accurate interpretation (includes *accurate interpretation* and *asks for clarification to ensure accurate interpretation*) constituted 70% and inaccurate

interpretation made up 30% of all coded TUs. Inaccurate interpretation occurred at twice the rate for ad hoc (54% of all coded TUs; p <0.001) versus professional in-person (25%) and videoconferencing (23%) interpretation. *Omissions* were the most common type of error overall; the rate of *omissions* was twice as high for ad hoc interpretation (33% of TUs, p <0.001) than for professional in-person (16%) or videoconferencing interpretation (16%). The rate of *answers for patient or clinician* was higher for ad hoc interpretation (16%, p <0.001) than for professional in-person (1%) or videoconferencing interpretation (1%). Other errors occurred infrequently (≤ 4% of all coded TUs).

In the adjusted model, the odds of inaccurate interpretation were significantly lower for professional in-person (OR = −1.25; 95% CI −1.56, −0.95) and videoconferencing (OR = −1.05; 95% CI −1.26, −0.84) than for ad hoc interpretation.

Clinical significance of interpretation errors and visit complexity

On average, there were 27 errors of interpretation/visit (Table 5). Among the TUs coded as errors, *omissions* made up 65% of errors, followed by *answers for patient or clinician* (14%) and *substitutions* (12%). Overall, 7.1% of errors were rated as moderately/highly clinically significant; mean rating of the clinical significance of errors was 1.67 (SD 0.61). The rate of moderately/highly clinically significant errors was higher for *substitutions* and *additions* than for other types of errors. An example of a moderately clinically significant error was when the patient said (in Spanish): “Yes, I feel shortness of breath. My chest does not hurt, but what happens is I feel as if I can’t breathe. I can’t run; I can’t walk quickly because I feel as if I can’t breathe.” The interpreter then said: “He has pain, chest pain, but he cannot run or he cannot walk fast, because then he has, he’s short of breath.” An example of a highly clinically significant error was when the physician said: “And this is Tylenol, extra strength. It says she can take 2 tablets every 6 hours for pain.” The interpreter translated this as (in Spanish): “And that is Tylenol that is stronger, you can take 2 tablets every 4 hours for pain,” which could be as much as 6 grams per day of acetaminophen.

Although moderately/highly clinically significant errors occurred more often in visits with ad hoc (8%) than with professional in-person (3%) or professional videoconferencing (7%) interpretation, differences were not statistically significant (p=0.14). In the adjusted GEE model, the odds of a moderately/highly clinically significant interpretation error were significantly lower for professional in-person (OR −0.06; 95% CI −1.05, 0.92) versus ad hoc interpretation.

Conclusions

This study found that inaccurate interpretation rates were comparable for professional in-person and videoconferencing modes and about half that of ad hoc interpretation. Omissions or providing preemptive answers for clinicians or patients were the most common types of interpretation errors. Although there were 27 errors on average per encounter, only 7% of these were moderately or highly clinically significant. Regardless, interpreter errors were common and disproportionately occurred in encounters with ad hoc interpreters, with, on average, 1–2 moderately or highly clinically significant errors per encounter. The likelihood of a moderately or highly clinically significant error was significantly lower for professional

in-person than ad hoc interpretation. These findings emphasize the importance of establishing a health systems quality metric for professional interpretation or language-concordant clinicians for LEP patient encounters.

Although the number of errors per encounter in our study was similar to that found in a pediatric study,[23] our specific error rates for professional and ad hoc interpreters, were somewhat lower than those observed for their counterparts in pediatric visits.[22, 23] Such differences could be due to variations in severity and acuity of conditions seen and the presence of another person (e.g., parent) in the encounter. Overall error rates and rates of clinically significant errors for professional interpretation in our study were similar to rates found in another primary care study of professional in-person interpretation.[28]

Our study is consistent with 3 studies that found that omissions were the most common type of error,[23, 28] which could be improved through better interpreter training. In our study, clinician coders rated the ability of ad hoc interpreters to convey clinicians' statements as being of lower quality than professional video-conferencing, highlighting the importance of training in medical terminology. In our study, only half of clinicians reported training on working with interpreters, which can improve clinician knowledge and attitudes about use of interpreters and intent to overcome language barriers.[37] Thus, standardized training of interpreters and clinicians is needed.

In our study, patients rated professional video-conferencing interpretation as being of higher quality than the other two modes, while clinician ratings did not differ by mode. It could be that for busy clinicians, merely having some type of interpretation is adequate, while LEP patients may be more sensitive to variations in the quality of interpretation than clinicians. Higher patient ratings of video-conferencing versus in-person professional interpretation could be due to shorter wait times for video interpreters. Consistent with the higher error rates observed for ad hoc versus professional interpretation, clinician coders indicated that professional video-conferencing interpreters were better able to convey the meaning of clinicians' statements than ad hoc interpreters, and tended to rate their overall quality as better. Clinician coders, who reviewed visit transcripts, may have provided more objective or nuanced ratings than the visit clinicians.

Demand for professional interpreter services exceeds availability and will continue to grow. In 2008, a national study found that while 97% of physicians reported having non-English-speaking patients, only 56% were in practices that had interpreter services, and nearly 1 in 5 reported an inability in the past year to secure an interpreter when they thought one was medically necessary.[38] Videoconferencing offers promise for addressing shortages of professional medical interpretation services, with demonstrated acceptability among patients and clinicians,[31, 39] and preference over telephone interpretation.[40, 41] Our study found that patients rated the quality of video-conferencing interpretation higher than in-person interpretation, while clinicians rated both modes equally.

A limitation of our study was the inability to randomize patients to interpretation mode, thus, results will require confirmation in another study. Implementation of videoconferencing interpretation increased system capacity of professional interpretation in

this setting, therefore, clinical staff felt that randomization to ad hoc interpretation was unethical because it could compromise patient confidentiality and quality of care. Future randomized studies might compare various professional interpretation modes without the ad hoc group. Another limitation was imbalance in the number of encounters by interpretation mode, which was a product of increased system capacity for professional videoconferencing interpretation. The inability to randomize and the imbalanced sample sizes may have introduced residual confounding beyond that controlled for in the analyses. Because we did not track the identity of the interpreter for each visit, we were unable to control for clustering effects within interpreter, which should occur in future studies. Other limitations include that the study involved only Spanish-speaking patients and one primary care clinic and results may not generalize to other languages or settings. Furthermore, the data were collected about 10 years ago. However, our comparison of the accuracy between professional videoconferencing and in-person interpretation in particular makes our findings very relevant to current clinical practice, which is increasingly turning to professional videoconferencing interpretation. A national trend among the major phone interpreter vendors now is to offer videoconferencing interpretation options.

Despite these limitations, our study demonstrates that interpretation errors are common and occur more frequently when untrained interpreters are used, compared to both in-person and videoconferencing professional interpretation. Professional video conferencing interpretation in particular, appears to be well accepted by patients and may be a cost-effective method for expanding access to professional interpretation services to meet quality standards. Using a mode of language interpretation with LEP patients that doubles error rates is unacceptable. Full deployment of professional interpretation capacity for LEP patients is a quality of care issue whose time has come.

Acknowledgements

We are indebted to the clinicians, staff and patients of Alameda County Medical Center Primary Care Clinic who made this study possible.

Funding: The California Endowment (grant number 20061003) and the Resource Centers for Minority Aging Research program of the National Institute on Aging grant number P30 AG15272.

References

1. Whatley M, Batalova J. Limited English Proficient Population the United States. Migration Policy Institute. 2013 Available at: http://www.migrationpolicy.org/article/limited-english-proficient-population-united-states.
2. Chen AH, Youdelman MK, Brooks J. The legal framework for language access in healthcare settings: Title VI and beyond. J Gen Intern Med. 2007; 22(Suppl 2):362–367.
3. Federal Register. TitleVI: Prohibition against National Origin Discrimination Affecting Limited English Proficient Persons. 2004; 69:1763–1768. Available at: http://www.archives.gov/eeo/laws/title-vi.html.
4. Chen A. Doctoring across the language divide. Health Aff. 2006; 25:808–813.
5. DeCamp LR, Kuo DZ, Flores G, et al. Changes in language services use by US pediatricians. Pediatrics. 2013; 132:e396–e406. [PubMed: 23837185]
6. Ramirez D, Engel KG, Tang TS. Language interpreter utilization in the emergency department setting: a clinical review. J Health Care Poor Underserved. 2008; 19:352–362. [PubMed: 18469408]

7. Schenker Y, Perez-Stable EJ, Nickleach D, et al. Patterns of interpreter use for hospitalized patients with limited English proficiency. J Gen Intern Med. 2011; 26:712–717. [PubMed: 21336672]
8. Flores G. The impact of medical interpreter services on the quality of health care: a systematic review. Med Care Res Rev. 2005; 62:255–299. [PubMed: 15894705]
9. Jacobs E, Chen AH, Karliner LS, et al. The need for more research on language barriers in health care: a proposed research agenda. Milbank Q. 2006; 84:111–133. [PubMed: 16529570]
10. Divi C, Koss RG, Schmaltz SP, et al. Language proficiency and adverse events in US hospitals: a pilot study. Int J Qual Health Care. 2007; 19:60–67. [PubMed: 17277013]
11. Karliner LS, Jacobs EA, Chen AH, et al. Do professional interpreters improve clinical care for patients with limited English proficiency? A systematic review of the literature. Health Serv Res. 2007; 42:727–754. [PubMed: 17362215]
12. 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; 16:468–474. [PubMed: 11520385]
13. Manson A. Language concordance as a determinant of patient compliance and emergency room use in patients with asthma. Med Care. 1988; 26:1119–1128. [PubMed: 3199910]
14. Crane JA. Patient comprehension of doctor-patient communication on discharge from the emergency department. J Emerg Med. 1997; 15:1–7. [PubMed: 9017479]
15. Perez-Stable EJ, Napoles-Springer A, Miramontes JM. The effects of ethnicity and language on medical outcomes of patients with hypertension or diabetes. Med Care. 1997; 35:1212–1219. [PubMed: 9413309]
16. Baker DW, Hayes R, Fortier JP. Interpreter use and satisfaction with interpersonal aspects of care for Spanish-speaking patients. Med Care. 1998; 36:1461–1470. [PubMed: 9794340]
17. David RA, Rhee M. The impact of language as a barrier to effective health care in an underserved urban Hispanic community. Mt Sinai J Med. 1998; 65:393–397. [PubMed: 9844369]
18. Rivadeneyra R, Elderkin-Thompson V, Silver RC, et al. Patient centeredness in medical encounters requiring an interpreter. Am J Med. 2000; 108:470–474. [PubMed: 10781779]
19. 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; 4:171–176. [PubMed: 16228770]
20. Hampers LC, McNulty JE. Professional interpreters and bilingual physicians in a pediatric emergency department: effect on resource utilization. Arch Pediatr Adolesc Med. 2002; 156:1108–1113. [PubMed: 12413338]
21. Butow PN, Goldstein D, Bell ML, et al. Interpretation in consultations with immigrant patients with cancer: how accurate is it? J Clin Oncol. 2011; 29:2801–2807. [PubMed: 21670462]
22. Laws MB, Heckscher R, Mayo SJ, et al. A new method for evaluating the quality of medical interpretation. Med Care. 2004; 42:71–80. [PubMed: 14713741]
23. Flores G, Laws MB, Mayo SJ, et al. Errors in medical interpretation and their potential clinical consequences in pediatric encounters. Pediatrics. 2003; 111:6–14. [PubMed: 12509547]
24. Gany F, Kapelusznik L, Prakash K, et al. The impact of medical interpretation method on time and errors. J Gen Intern Med. 2007; 22(Suppl 2):319–323. [PubMed: 17957418]
25. Pham K, Thornton JD, Engelberg RA, et al. Alterations during medical interpretation of ICU family conferences that interfere with or enhance communication. Chest. 2008; 134:109–116. [PubMed: 18347204]
26. Aranguri C, Davidson B, Ramirez R. Patterns of communication through interpreters: a detailed sociolinguistic analysis. J Gen Intern Med. 2006; 21:623–629. [PubMed: 16808747]
27. Gany FM, Gonzalez CJ, Basu G, et al. Reducing clinical errors in cancer education: interpreter training. J Cancer Educ. 2010; 25:560–564. [PubMed: 20390395]
28. Jackson JC, Nguyen D, Hu N, et al. Alterations in medical interpretation during routine primary care. J Gen Intern Med. 2011; 26:259–264. [PubMed: 20922496]
29. Stewart AL, Napoles-Springer A, Perez-Stable EJ. Interpersonal processes of care in diverse populations. Milbank Q. 1999; 77:305–339. 274. [PubMed: 10526547]

30. Karliner LS, Perez-Stable EJ, Gildengorin G. The language divide. The importance of training in the use of interpreters for outpatient practice. J Gen Intern Med. 2004; 19:175–183. [PubMed: 15009797]

31. Jones D, Gill P, Harrison R, et al. An exploratory study of language interpretation services provided by videoconferencing. J Telemed Telecare. 2003; 9:51–56. [PubMed: 12641894]

32. Glasser, BG.; Strauss, AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Hawthorne: Aldine Publishing Company; 1967.

33. Strauss, A.; Corbin, J. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park: Sage Publications, Inc; 1990.

34. Gany F, Leng J, Shapiro E, et al. Patient satisfaction with different interpreting methods: a randomized controlled trial. J Gen Intern Med. 2007; 22(Suppl 2):312–318. [PubMed: 17957417]

35. Hornberger JC, Gibson CD Jr, Wood W, et al. Eliminating language barriers for non-English-speaking patients. Med Care. 1996; 34:845–856. [PubMed: 8709665]

36. Hays, RD.; Hayashi, T. Beyond internal consistency reliability: rationale and user's guide for multitrait analysis program on the microcomputer. Santa Monica, CA: Rand; 1990.

37. Jacobs EA, Diamond LC, Stevak L. The importance of teaching clinicians when and how to work with interpreters. Patient Educ Couns. 2010; 78:149–153. [PubMed: 20036480]

38. Reschovsky, JD.; Boukus, ER. Issue Brief: Findings from HSC. Washington, DC: Center for Studying Health System Change; 2010. Modest and uneven: physician efforts to reduce racial and ethnic disparities; p. 1-6.Available at: http://www.hschange.com/CONTENT/1113/ [Accessed April 13, 2014]

39. Napoles AM, Santoyo-Olsson J, Karliner LS, et al. Clinician ratings of interpreter mediated visits in underserved primary care settings with ad hoc, in-person professional, and video conferencing modes. J Health Care Poor Underserved. 2010; 21:301–317. [PubMed: 20173271]

40. Locatis C, Williamson D, Gould-Kabler C, et al. Comparing in-person, video, and telephonic medical interpretation. J Gen Intern Med. 2010; 25:345–350. [PubMed: 20107916]

41. Locatis C, Williamson D, Sterrett J, et al. Video medical interpretation over 3G cellular networks: a feasibility study. Telemed J E Health. 2011; 17:809–813. [PubMed: 22011055]

Table 1

Measures of Accurate and Inaccurate Language Interpretation Observed in Audio-taped Visits of Spanish-speaking Adult General Medicine Patients.

Accuracy of Interpretation Code	**Definition** ***Interpreter….***	**Example**[1] **C=Clinician; I-Interpreter; P=Patient; O=Other person**
Accurate interpretation	Provides accurate interpretation of word or phrase uttered by clinician or patient	P: *Me dio un tratamiento ella por 3 meses y ya me ha mejorado bastante el dolor, pero sigo sintiendo poquito.* I: *She is still feeling a little bit of the pain but the doctor gave her a treatment for three months which has helped her quite a bit with the pain.*
Asks for clarification to ensure accurate interpretation	Asks clarifying question of clinician or patient that improves communication	I: *Entonces esta es Robitussin PM. Es dos cucharaditas, dos veces al, al día. Two teaspoons two times? Four times a day?*
Additions	Adds a word or phrase that was not uttered by clinician or patient	P: *Los oídos están bien, pero no escucho bien. Hay días que se me tapan más, otros días menos.* [The ears are fine, but I can't hear well. There are days when my ears are more plugged up, and other days when it is less.] I: *His ears are OK. It's just that he is deaf. He can't hear well.*
Substitutions	Substitutes word or phrase that is different from that uttered by clinician or patient, which changes meaning of original statement	P: *No. Me dieron una medicina.* [They gave me a medicine.] I:*Seems that they changed her medicine.* C: *Better, better than before the surgery?* I: *¿Qué si te sientes mejor ahorita?* [Do you feel better now?]
Answers for patient or clinician	Provides preemptive answer to question posed by clinician or patient rather than interpreting the question	C: *What is the problem?* I: *Now she is complaining about a pain in the chest.* C: *So, can you tell me about this chest pain, how often you are having it?* I: *She says that for a week she is having a pain in the chest.*
Omission	Fails to interpret word or phrase uttered by the clinician or patient	C: *But he has no infection. Everything else looks good. And he went to surgery for the problems on his face. What happened with that?* I: *Que lo mandaron al cirujano para el problema en la cara. ¿Qué pasó con eso?* [So they sent you to the surgeon for the problem with your face. What happened with that?]
Editorializing	Adds personal opinion to the interpretation of a word or phrase uttered by the clinician or patient	C: *Okay. And when was the last time she saw an eye doctor?* I: *¿Cuándo fue la última vez que vió el doctor de los ojos?* [When was the last time you saw the eye doctor?] O: *No, no la han mirado.* [No, they have not seen her.] I: *¿Hace cuánto?* [How long ago has it been?] P: *No, no, en la verdad nunca.* [No, truthfully, never.] I: *They don't think that…never.* C: *Okay.* I:*That doesn't make sense for diabetes though.*
False fluency	Interprets with an incorrect word or phrase	C: *OK. The nurse took her vital signs, and her blood pressure is one hundred over sixty. Normal.* I: *Ok. Le tomó, éste la… La enfermera le tomó sus vitales y éste, su presión arterial es cien sobre sesen, seiscientos diez y es normal.* [Ok. She took…the nurse took your vitals and your, your arterial pressure is 100 over six…six hundred ten and it is normal.]

[1] Underlined text illustrates lower quality interpretation

Table 2

Demographic Characteristics of Spanish-speaking Adult General Medicine Patients and their Clinicians, Alameda County, California, 2005.

Characteristic	N (%)
Patients (N=32)	
Age in years, mean (SD)	52.9 (15.8)
Sex	
Women	24 (75)
Men	8 (25)
Highest grade of schooling	
Never attended	4 (13)
Grade 1–5	11 (34)
Elementary/Middle school/Some high school (6–11 years)	13 (41)
High school diploma or higher	4 (12)
Health insurance	
Only Medicaid	11 (34)
Medicaid and Medicare	3 (9)
No insurance	18 (56)
Place of birth	
Mexico	22 (69)
Central/South America	10 (31)
Marital status	
Married or living with partner	15 (47)
Not married	17 (53)
Employment status	
Employed	8 (25)
Homemaker	13 (41)
Unemployed, retired or disabled	11 (34)
Health Status	
Poor/Fair	28 (88)
Good/Very good/Excellent	4 (12)
Has an ongoing medical condition	26 (81)
Number of visits to this clinician	
First time	9 (28)
2 or more times	23 (72)
Clinicians (N=14)	
Age in years, mean (SD)[1]	50.6 (11.5)
Sex	
Women	10 (71)
Men	4 (29)
Ethnicity[1]	
Non-Latino White	5 (50)

Characteristic	N (%)
Southeast Asian	2 (20)
Asian or Asian American	1 (10)
Latino, Latin American or Hispanic Multi-ethnic	1 (10)
Type of clinician[1]	
Internal medicine physician	8 (80)
Physician assistant	2 (20)
Training on working with professional interpreters[1]	
None	5 (50)
Any training (a lecture, several lectures, a course)	5 (50)

[1]Data missing for 4 clinicians

Table 3

Patient, Clinician and Study Clinician-coder Ratings of the Quality of Interpretation in the Medical Encounters of Spanish-speaking Adult General Medicine Patients, by Mode of Interpretation, Alameda County, California, 2005.

Mean (SD)[1]	Total (N=32)	Professional In-person Interpretation (N=5)	Professional video-conferencing Interpretation (N=22)	Ad hoc Interpretation (N=5)	*p*-value
Patient ratings					
Patient-rated Quality of Interpretation Scale Score (range=1–5)	4.0 (0.80)	3.4 (0.55)	4.3 (0.70) *I,A*	3.4 (0.89)	< 0.05
Clinician ratings					
Clinician-rated Quality of Interpretation Scale Score (range=1–5)	3.4 (0.80)	3.1 (0.58)	3.4 (0.77)	3.5 (1.29)	0.70
Study Clinician-coder ratings					
Overall quality of the interpretation services provided	2.9 (1.25)	2.8 (1.64)	3.2 (1.05)	1.8 (1.30)	0.08
Ability of interpreter to convey intended meaning of clinicians' statements	2.9 (1.17)	3.0 (1.41)	3.2 (0.91)	1.6 (1.34) *V*	< 0.05
Ability of interpreter to convey intended meaning of patients' statements	2.9 (1.28)	3.0 (1.58)	3.1 (1.19)	2.0 (1.22)	0.23

[1] Response options were: 1 = poor; 2 = fair; 3 = good; 4 = very good; and 5 = excellent.

Letter superscripts indicate that the indexed parameter differed significantly from the corresponding parameter for the interpreter mode designated as I=In-person professional, V=videoconferencing professional, A=ad-hoc.

Table 4

Frequency of Accuracy of Interpretation Codes by Type of Interpretation for 32 Primary Care Clinic Encounters, Alameda County, California, 2005.

	Professional In-person Interpretation N (%)	Professional video-conferencing Interpretation N (%)	Ad hoc Interpretation N (%)	*p*-value
Total number of text units coded	518 (100)	1,836 (100)	590 (100)	---
Accurate interpretation codes	389 (75)	1415 (77)	269 (46)	---
Accurate interpretation	336 (65)	1194 (65)	222 (38)	< 0.001
Asks for clarification to ensure accurate interpretation	53 (10)	221 (12)	47 (8)	<0.05
Inaccurate interpretation codes	129 (25)	421 (23)	321 (54)	< 0.001
Omission	81 (16)	288 (16)	196 (33)	< 0.001
Answers for patient or clinician	7 (1)	15 (1)	93 (16)	< 0.001
Substitutions	22 (4)	68 (4)	18 (3)	0.57
False fluency	11 (2)	25 (1)	3 (0)	0.06
Editorializing	3 (1)	15 (1)	10 (2)	0.10
Additions	5 (1)	10 (0)	1 (0)	0.22

Table 5

Frequency and Clinical Significance of Interpreter Errors by Type of Error during 32 Adult General Medicine Encounters, Alameda County, California, 2005.

Type of Interpreter Error	N	N/visit	Percent of errors that were moderately or highly clinically significant [1]	Mean rating (SD) of clinical significance of error
Total number of errors	871 (100)	27.2	7.1	1.67 (0.61)
Omission	565 (65)	17.7	6.0	1.66 (0.59)
Answers for patient or clinician	115 (14)	3.6	8.7	1.64 (0.66)
Substitutions	108 (12)	3.4	13.0	1.80 (0.68)
False fluency	39 (4)	1.20	5.1	1.54 (0.60)
Editorializing	28 (3)	0.90	0	1.50 (0.51)
Additions	16 (2)	0.50	12.5	1.69 (0.70)

[1] Response options for clinical significance of errors: 1=clinically insignificant; 2=mildly clinically significant; 3=moderately clinically significant; and 4=highly clinically significant.