Errors of Medical Interpretation and Their Potential Clinical Consequences: A Comparison of Professional Versus Ad Hoc Versus No Interpreters

Glenn Flores, MD, Milagros Abreu, MD, MPH, Cara Pizzo Barone, MD, Richard Bachur, MD, Hua Lin, PhD

From the Division of General Pediatrics, Department of Pediatrics, UT Southwestern Medical Center, Dallas, TX (Flores, Lin); the Children's Medical Center, Dallas, TX (Flores); the Department of Epidemiology, Boston University School of Public Health, Boston, MA, and the Latino Health Insurance Program, Inc, Framingham, MA (Abreu); the Palo Alto Medical Foundation, Palo Alto, CA (Barone); and the Division of Emergency Medicine, Children's Hospital Boston, Boston, MA (Bachur).

Study objective: To compare interpreter errors and their potential consequences in encounters with professional versus ad hoc versus no interpreters.

Methods: This was a cross-sectional error analysis of audiotaped emergency department (ED) visits during 30 months in the 2 largest pediatric EDs in Massachusetts. Participants were Spanish-speaking limited-English-proficient patients, caregivers, and their interpreters. Outcome measures included interpreter error numbers, types, and potential consequences.

Results: The 57 encounters included 20 with professional interpreters, 27 with ad hoc interpreters, and 10 with no interpreters; 1,884 interpreter errors were noted, and 18% had potential clinical consequences. The proportion of errors of potential consequence was significantly lower for professional (12%) versus ad hoc (22%) versus no interpreters (20%). Among professional interpreters, previous hours of interpreter training, but not years of experience, were significantly associated with error numbers, types, and potential consequences. The median errors by professional interpreters with greater than or equal to 100 hours of training was significantly lower, at 12, versus 33 for those with fewer than 100 hours of training. Those with greater than or equal to 100 hours of training committed significantly lower proportions of errors of potential consequence overall (2% versus 12%) and in every error category.

Conclusion: Professional interpreters result in a significantly lower likelihood of errors of potential consequence than ad hoc and no interpreters. Among professional interpreters, hours of previous training, but not years of experience, are associated with error numbers, types, and consequences. These findings suggest that requiring at least 100 hours of training for interpreters might have a major impact on reducing interpreter errors and their consequences in health care while improving quality and patient safety. [Ann Emerg Med. 2012;60:545-553.]

Please see page 546 for the Editor's Capsule Summary of this article.

A **feedback** survey is available with each research article published on the Web at www.annemergmed.com.
A **podcast** for this article is available at www.annemergmed.com.

0196-0644/$-see front matter

doi:10.1016/j.annemergmed.2012.01.025

INTRODUCTION

Background

More than 6,900 known living languages are spoken in the world.[1] A sizable, growing body of literature documents that language barriers can have a major adverse effect on health and health care, including suboptimal health status; a lower likelihood of having a regular health care provider; lower rates of mammograms, pap smears, and other preventive services; greater likelihood of a diagnosis of more severe psychopathology and leaving the hospital against medical advice among psychiatric patients; increased risks of drug complications; and higher resource use for diagnostic testing.[2]

Importance

More than 59 million Americans speak a language other than English at home, and 25.2 million have limited English proficiency (ie, a self-rated ability to speak English less than "very well").[3] Despite the large number of Americans with limited English proficiency and federal policy[4] requiring providing adequate language assistance to patients with limited

Editor's Capsule Summary

What is already known on this topic

Limited English proficiency could lead to clinically important misunderstandings between patients and emergency department (ED) staff.

What question this study addressed

This study analyzed 57 audiotaped ED encounters to compare the accuracy of translation by 3 means: professional interpreters, ad hoc interpreters (such as family or staff members), and no interpreter (when patient and physician used no assistance).

What this study adds to our knowledge

Interpretation errors of potential clinical importance were common (18%); professional interpretation had roughly half as many errors as the other 2 modes.

How this is relevant to clinical practice

These results suggest the need for more robust and practically usable language support in EDs where patients with limited English skills can be reasonably expected.

English proficiency, many such patients do not receive professional medical interpretation, but rather must resort to using ad hoc interpreters, such as family members, friends, or strangers from the waiting room, or having no interpreter. One study of 530 Latino adult ED patients revealed no interpreter use for 46% of patients with limited English proficiency, and 39% of interpreters used had no training.[5]

Goals of This Investigation

To our knowledge, no study has compared errors in medical interpretation and their potential clinical consequences among professional hospital interpreters, ad hoc interpreters, and encounters with no interpreters. The aim, therefore, was to compare interpreter errors and their potential clinical consequences in pediatric encounters with hospital versus ad hoc versus no interpreters. A secondary aim was to determine whether, among professional interpreters, hours of previous interpretation training or years of interpreter experience are associated with error numbers, types, and potential clinical consequences, which might provide useful data for interpreter training programs, hospitals, and policymakers.

MATERIALS AND METHODS

Encounters were audiotaped in the 2 largest pediatric emergency departments (EDs) in Massachusetts. To be eligible for the study, the child's principal caretaker had to identify Spanish as the primary language spoken at home and had to have limited English proficiency, according to the US Census definition (self-rated ability to speak English of less than "very well").[6] During 30 months (May 2000 to November 2002), pediatric ED visits were purposively sampled between 8 AM and 11 PM on weekdays and weekends. A bilingual research assistant was present during the encounter only for audiotaping and did not interpret or produce transcripts. Bilingual verbatim transcripts were prepared from audiotaped encounters by professional transcriptionists fluent in both English and Spanish and unaware of the study aims. To ensure accuracy and reliability, each transcript was reviewed for errors 3 times by 3 research assistants by simultaneous comparison of audiotapes with transcripts.

Patients, families, clinicians, and interpreters were told only that this was a study of patient-physician communication, and were unaware that interpretation errors would be analyzed.

Interpreters were classified as (1) professional interpreters, who receive financial compensation and are employed by the hospitals' interpreter services department; (2) ad hoc interpreters, including family members, friends, nonclinical hospital employees, strangers from waiting rooms, and hospital clinical staff (including nurses and social workers) who had no formal medical interpreter training or screening; and (3) no interpreter, defined as encounters with no professional or ad hoc interpreter present (including when patients served as interpreters).

Professional interpreters were asked to identify the number of hours of previous medical interpreter training that they completed and years of experience as a medical interpreter.

Methods for identification and categorization of interpreter errors were identical to those detailed in previous research by our team on interpreter errors in outpatient clinics.[7] In brief, 5 categories were used to classify interpreter errors, based on 4 categories from previous work,[8,9] supplemented by an additional category (false fluency). These categories are:

- Omission: The interpreter did not interpret a word/phrase uttered by the clinician, parent, or child.
- Addition: The interpreter added a word/phrase not uttered by the clinician, parent, or child.
- Substitution: The interpreter substituted a word/phrase for a different word/phrase uttered by the clinician, parent, or child.
- Editorialization: The interpreter provided his or her own views as the interpretation of a word/phrase uttered by the clinician, parent, or child.
- False fluency: The interpreter used a word/phrase that does not exist in that particular language or an incorrect word/phrase that substantially altered the meaning. For example, a false-fluency error in interpreting the Spanish word for eye (*ojo*) would be "eyo" (a nonexistent word) or *oreja* (external ear).

Consistent with prior work,[7] an interpreter error additionally was considered to have potential clinical consequences if it altered or potentially altered the history of present illness,

medical history, diagnostic or therapeutic interventions, parental understanding of the child's medical condition, and/or plans for future medical visits (including follow-up visits and specialty referrals). Thus, a substantial number of interpreter errors would not be expected to have potential clinical consequences, including errors caused by natural conversational redundancy that might occur in conversations between 2 native speakers of the same language, such as errors in phrasing or word choice that generally contribute little to misunderstanding and have no potential clinical consequences.

This error classification framework was used when encounters with all 3 categories of interpreters—professional, ad hoc, and no interpreter—were analyzed. For encounters with no interpreters, in which it was hypothesized that there might be fewer interpretation acts per utterance, this error framework was considered suitable because the omission category would capture such interpretation "skips."

Consistent with previous research by our team,[7] the focus of error analyses was meaning, rather than word-for-word interpretation. Medical jargon, idiomatic expressions, and contextual clarifications may oblige interpreters to not interpret utterances word for word. Hence, any deviations from word-for-word interpretation because of jargon, idioms, or contextual clarifications were not classified as errors, nor was it considered an error when the interpreter utterance was intended to preserve or enhance meaning. Because medical interpreters also can serve as cultural brokers or advocates, any utterances that could be interpreted as cultural explanations or as patient or family advocacy were not classified as interpreter errors.

Data analysis required several years of effort by the research team because of the extremely labor-intensive nature of the analyses and the large volume of data (>2,300 transcript pages and >61,000 dialogue lines) that had to be analyzed word by word by at least 2 independent observers. The error classification framework has been shown to yield a mean agreement of 99% (SD 1.7; range 97% to 100%) and a mean κ of 0.99 (SD 0.03; range 0.94 to 1.0).[7] Details on validation of the classification tools for interpreter errors and their clinical consequence are provided elsewhere.[7]

Primary Data Analysis

Data were coded and analyzed with SAS (version 9.1; SAS Institute, Inc., Cary, NC). Univariate analyses were performed to identify missing values and outliers. The *t* test, nonparametric Wilcoxon test, and ANOVA were used to test for differences among interpreter categories in the median and mean numbers of dialogue lines. Two-tailed *P* values are reported, with $P<.05$ considered statistically significant.

To test for associations between professional interpreters' years of experience and the number of interpreter errors, Pearson correlation coefficients were calculated for each error category and the overall number of errors. To examine associations between the number of hours of previous professional interpreter training with numbers and types of errors and their clinical consequence, a sensitivity analysis was performed to identify the optimal threshold of hours of training by examining the effect on the median number of interpreter errors of varying the cut point from as few as 40 hours to as many as 150 hours of training. The optimal threshold was identified using the greatest magnitude of difference and the associated *P* value in comparing median interpreter errors at or above the cut point versus below that cut point.

An error-count approach is vulnerable to length-biased sampling in that longer utterances have more opportunities for errors than shorter ones. To adjust for possible differences among interpreter categories in the numbers of lines and pages, multiple linear regression was performed for the total number of interpreter errors and for error categories, adjusting for the length of the conversation. Because the findings of these analyses did not differ from those of bivariate analyses, only bivariate findings are reported.

Approval to conduct this study was obtained from the Boston Medical Center and Boston Children's Hospital institutional review boards. Written consent and assent (when the child was of the appropriate age) were obtained from all participants.

RESULTS

A total of 57 interpreted encounters were audiotaped, yielding 2,367 pages of transcripts and 61,478 dialogue lines. Interpreters present during these encounters included professional interpreters, 20 (35%); ad hoc interpreters, 27 (47%); and no interpreter, 10 (18%). There were no significant differences in the median number of lines of dialogue per encounter for professional versus ad hoc versus no interpreters (501 versus 630 versus 420, respectively; $P=.20$), consistent with previous work documenting no significant differences in uttered words per encounter by interpreter category.[7]

The mean age of the patients was 5.1 years (SEM 0.43 years); 53% were male patients, 14% were uninsured, and 46% of the patients' parents were not high-school graduates. The most common diagnostic categories included wounds, fractures, and other trauma (28%); viral illnesses (17%); gastroenteritis (10%); fever (5%); pharyngitis (5%); and otitis media (5%). Other less common diagnoses (<5%) included upper respiratory infections, asthma, observation, dehydration, pneumonia, cellulitis, swallowing a coin, and scoliosis.

A total of 1,884 interpreter errors were noted in the 57 encounters, with a mean of 33 errors (SD 4.7 errors) per encounter and a range of 2 to 246 errors per encounter. The mean errors per encounter by interpreter category did not significantly differ, and were as follows: professional hospital interpreters, 32.7 (SD 4.9); ad hoc interpreters, 33.7 (SD 4.7); and no interpreter, 32.3 (SD 23.9). The most common category of errors was omission (47% of all errors), followed by false fluency (26%), addition (10%), editorialization (9%), and substitution (9%).

Compared with professional interpreters, ad hoc interpreters and having no interpreter resulted in significantly higher proportions of omissions and false-fluency errors (Table 1). In

Table 1. Group differences in the proportions of interpreter errors by interpreter type.

	Type of Interpreter, %		
Error Category	**None**	**Ad Hoc**	**Professional**
Omission*	54.2	46.3	41.9
False-fluency error†	35.9	31.6	13.6
Addition†	1.5	6.5	17.9
Substitution†	2.5	9.0	13.3
Editorialization†	5.9	6.6	13.3

*P=.001 by χ^2 test for comparison of types of interpreters.
†P<.01 by χ^2 test for comparisons of types of interpreters.

contrast, professional hospital interpreters had higher proportions of errors in the less frequent remaining error categories.

A total of 344 errors (18%) had potential clinical consequences. The mean number of errors of clinical consequence per encounter was 6 (SD 1.1), ranging from 0 to 47 per encounter. The proportion of errors of potential clinical consequence was significantly lower for professional hospital interpreters versus ad hoc interpreters and no interpreter, at 12% versus 22% versus 20%, respectively (P<.01). Both ad hoc interpreters (odds ratio 2.2; 95% confidence interval [CI] 1.6 to 2.9) and having no interpreter (odds ratio 1.9; 95% CI 1.3 to 2.8) were associated with approximately double the odds of committing errors of potential clinical consequence, compared with professional hospital interpreters.

Two examples (Figure) are provided of errors of potential clinical consequences. In addition, multiple clear cases of potential clinical consequences were noted. One family with an ad hoc interpreter believed that they should administer 2 tablespoons instead of the prescribed 2 teaspoons per dose of amoxicillin prescribed for 10 days for their child's otitis media, and had no awareness of a second medication (Cortisporin otic solution for otitis externa) prescribed for their child, or the dosing instructions. In another encounter in which a family friend was the ad hoc interpreter, the interpreter told the physician that the child was not receiving any medications and did not have any drug allergies, although the interpreter never asked the mother these questions.

In an encounter without an interpreter in which the physician's Spanish and parents' English were substantially limited, the physician described an oral antibiotic for treatment of otitis media as for the "mouth" and provided the dose, dose frequency, and duration of the antibiotic only in English. The physician also prescribed eardrops, but explained their dosage and dose frequency only in English. The parents were told only in English to return to the ED for high fever, not drinking fluids, or experiencing no improvement in 2 days, but it was unclear whether the parents understood these instructions.

For professional hospital interpreters, the years of experience interpreting were not significantly associated with the overall number of errors, the number of errors by category of interpreter errors, or whether the errors were of potential clinical consequence (Table 2).

In contrast, the number of hours of previous training was significantly associated with the number and types of errors and their clinical consequence. The cut point in the number of hours of previous training that yielded the greatest magnitude of difference and the most significant P value was greater than or equal to 100 hours of training versus fewer than 100 hours of training (Table 3). The median number of errors committed by professional interpreters with greater than or equal to 100 hours of training was significantly lower, at 20 errors fewer, than the median errors committed by those with fewer than 100 hours of training (Table 4*A*). Professional interpreters with greater than or equal to 100 hours of training also committed significantly fewer false-fluency, substitution, and editorialization errors than those with fewer than 100 hours of training (Table 4*A*). Those with greater than or equal to 100 hours of training had a median of zero errors of potential clinical consequence, significantly lower than the median of approximately 4 errors for those with fewer than 100 hours of training (Table 4*A*). In addition, compared with professional hospital interpreters with fewer than 100 hours of training, those with greater than or equal to 100 hours of training committed a significantly lower proportion of errors of potential clinical consequence, both overall (2% versus 12%) and in every one of the 5 interpreter error categories (Table 4*B*). Indeed, for individuals with greater than or equal to 100 hours of training, 0% of the false-fluency, substitution, or editorialization errors had potential clinical consequences compared with 88%, 73%, and 87%, respectively, for those with fewer than 100 hours of training (Table 4*B*).

LIMITATIONS

This research was conducted in 2 urban EDs in Massachusetts, so the findings may not necessarily generalize to other nonurban settings, other states, and other countries. This study was restricted to children; similar work on interpreter errors in adult encounters is needed. Only families with limited English proficiency caretakers who spoke Spanish as their primary language were enrolled in the study; additional studies are needed to determine whether these findings pertain to limited English proficiency patients and families who speak a primary language other than Spanish. Because encounters were audiotaped and communication effectiveness was assessed through transcript analysis, important nonverbal and other contextual cues may have been missed, although research personnel were present throughout the audiotaping of all encounters, which they later analyzed, and all research personnel were instructed to comment on any contextual issues that might not have been appreciated solely through transcript analysis. Transcript reviewers could not be blinded to the interpreter category because of unavoidable cues in transcripts. The hours of training for professional interpreters were assessed by self-report, so it is possible that these self-reports overestimate or underestimate the actual hours of training completed. It has been several years since data on the last encounter were

A

Mother: *Yo veo que habla poco. Poco a poco y para comer también tiene dificultad y para tragar.*	I notice he talks little. Little by little and he has difficulty eating and swallowing.
Interpreter: She says he is talking like, like lazy, whatever he says all the time, and plus he is not eating at all.	
Doctor: And…why? Does he have a sore throat?	
Interpreter: *El está comiendo o…*	Is he eating or…
Mother: *Comiendo sí, pero…*	Eating, yes, but…
Interpreter: He eats just a little. Not a lot.	
Mother: *Con dificultad. El come pero…*	With difficulty but…
Interpreter: A little difficult but…	
Doctor: Okay.	

B

Mother: *La semana pasada a el le dio mucho mareo y no tenía fiebre ni nada, y la familia por parte de papá todos padecen de diabetes.*	Last week he had a lot of dizziness and he did not have fever or anything, and his dad's family all suffer from diabetes.
Doctor: Uh-hum	
Mother: *A mi me da miedo porque el lo que estaba mareado, mareado, mareado y no tenía fiebre ni nada.*	I'm scared because he's dizzy, dizzy, dizzy and he didn't have fever or anything.
Doctor: Ok. So she's saying you look kind of yellow, is that what she's saying?	
Patient: *¿Es que si me ví amarillo?*	Is it that I looked yellow?
Mother: *Estaba como mareado, como pálido.*	You were like dizzy, like pale.
Patient: Like I was like paralyzed, something like that.	

Dialogue from this encounter previously appeared in a published commentary (Language barriers to health care in the United States. *New England J Med.* 2006;355:229-231).

Figure. *A*, Multiple errors committed by an ad hoc interpreter (a family friend) during an encounter with a 6-year-old patient presenting with fever and a sore throat. *B*, Multiple omissions and false-fluency errors committed with no interpreter present in an encounter with a 12-year-old patient presenting with a chief complaint of dizziness.

Table 2. Correlation coefficient analysis of whether professional interpreters' years of experience interpreting are associated with numbers and categories of interpreter errors.

Error Category	Pearson Correlation Coefficient: Years of Experience Interpreting (95% CI)	*P* Value
All interpreter errors	0.28 (−0.20 to 0.63)	.24
Omissions	0.18 (−0.29 to 0.57)	.46
False-fluency errors	0.25 (−0.22 to 0.62)	.29
Additions	0.32 (−0.15 to 0.66)	.17
Substitutions	0.22 (−0.25 to 0.60)	.36
Editorializations	0.18 (−0.30 to 0.60)	.45
Potential clinical consequence	0.03 (−0.42 to 0.46)	.91

collected, so it is possible that the findings are out of date. Several years of effort by the research team, however, were required to complete the data analysis because of the substantially labor-intensive nature of the analyses and the large volume of data (2,367 pages of transcripts and 61,478 dialogue lines) that had to be analyzed word by word by at least 2 independent observers. In addition, the authors' clinical and policy observations and experience, along with recent studies,[10,11] suggest that little to nothing has changed from the time of completion of the study's data collection until now in terms of the ongoing lack of interpreters available for limited English proficiency patients and the variability of training that professional interpreters receive. Finally, telephone interpreters were not assessed as part of this study; we hope to do so in a future study.

Table 3. Analysis of varying cut points for hours of previous interpreter training for professional interpreters (N=20) and the effect on the median number of total interpreter errors.

Cut Point for Hours of Professional Interpreter Training	Median Number of Interpreter Errors for Cut Point		Magnitude of Difference (95% CI)*	*P* Value
	≥Cut Point	<Cut Point		
40	25	18	–7[†]	.63
60	23	30	7 (−28 to 26)	.55
72	24	25	1 (−18 to 23)	.71
80	12	30	18 (1 to 33)	.03
100[‡]	**12**	**32.5**	**20.5 (5 to 36)**	**.01**
150	12	30	18 (1 to 36)	.02

*Calculated as the difference between the median number of interpreter errors for less than cut point versus greater than or equal to cut point.
[†]No CI calculable for this value because there was only 1 observation in the group of previous interpreter training less than 40 hours.
[‡]A univariate analysis also revealed that 100 hours of training would be the optimal differentiating cut point for the mean number of hours of training in the data scatterplot.

Table 4. *A*, Median number of interpreter errors per encounter, by hours of training for professional interpreters.

Error Category	Median Number of Errors per Encounter, by Hours of Training for Professional Interpreter (N=20)		Magnitude of Difference (95% CI)	*P* Value
	≥100	<100		
All interpreter errors	12	32.5	20.5 (5–36)	.01
Omissions	8	10.5	2.5 (0–15)	.07
False-fluency errors	0.5	3	2.5 (0–7)	.047
Additions	2	5	3 (0–8)	.06
Substitutions	1	4	3 (1–6)	.01
Editorializations	0	3	3 (0–9)	.02
Potential clinical consequence	0	3.5	3 (0–9)	.03

Table 4. *B*, The proportion of errors of potential clinical consequence by hours of training for professional interpreters.

Error Category	Proportion of Errors of Potential Clinical Consequence by Hours of Training for Professional Interpreters (N=20), %		Magnitude of Difference, % (95% CI)	*P*
	≥100	<100		
All errors	2	12	10 (2 to 18)	.03
Omissions	3	29	26 (18 to 34)	.003
False-fluency errors	0	88	88 (78 to 98)	<.001
Additions	11	69	58 (50 to 64)	<.001
Substitutions	0	73	73 (52 to 95)	<.001
Editorializations	0	87	87 (65 to 99)	<.001

DISCUSSION

The study findings document that the proportion of errors of potential clinical consequence is significantly lower for professional hospital interpreters, compared with ad hoc interpreters and having no interpreter, both of which were found to have approximately double the odds of professional interpreters of committing errors of potential clinical consequence. These findings are consistent with a substantial body of research showing that optimal communication, patient satisfaction, and outcomes and the fewest interpreter errors occur when limited English proficiency patients have access to trained professional interpreters.[2] The findings also are consistent with those of recent studies documenting a higher likelihood of errors of potential clinical consequence for ad hoc versus professional interpreters in a variety of languages in adult oncology settings[12,13] and in a family medicine clinic.[14] This is the first study to our knowledge, however, to provide conclusive quantitative evidence that hospital interpreters commit significantly lower rates of errors of potential clinical consequence versus ad hoc and no interpreters. This study complements our earlier 2003 study,[7] which did not examine encounters in the ED or those with no interpreter, and did not analyze the associations of hours of interpreter training or years of experience with interpreter error numbers, categories, or potential clinical consequences.

Study results reveal that an average of 33 interpreter errors are committed in ED encounters, with as many as 246 errors committed in 1 encounter. About 1 in 5 errors had potential clinical consequences, with an average of 6 errors of potential clinical consequences per encounter and up to 47 in a single encounter. These findings indicate that interpreter errors are common in the ED, and a significant proportion of these errors have potential clinical consequences. These data are concerning because interpreter errors have been documented to compromise patient safety and to cause or be associated with preventable harm and serious injuries, including overdoses, misdiagnosis, and quadriplegia.[7,15-19] Recent studies indicate that the proportion of interpreter errors of potential clinical consequence can range from 5% in an adult primary care clinic[20] to as high as 77% in adult ICUs.[21] These findings, together with our study results, underscore that language issues routinely need to be included in discussions of quality of care and patient safety.[22] Two recent major reports from the Institutes of Medicine also highlighted that language and communication issues should be a part of monitoring and improving health and health care quality.[23,24]

Errors of potential clinical consequence were significantly more common with ad hoc interpreters and no interpreters, compared with professional hospital interpreters. These findings highlight the hazards of using ad hoc interpreters or no interpreter during encounters with limited English proficiency patients. Ad hoc interpreters can include family members, friends, untrained medical staff, custodians, or strangers recruited from the waiting room or the street. Previous research has documented that ad hoc interpreters are associated with patients being less likely to be told about medication adverse effects,[25] misinterpretation of up to half of all questions asked by physicians,[26] exclusion or distortion of key clinical information,[27] a greater likelihood in outpatient pediatric clinics of interpreter errors having potential clinical consequences versus those by professional interpreters,[7] and embarrassment by and tendencies to ignore questions about menstruation, bowel movements, and other bodily functions when the ad hoc interpreters are children.[26] Concerns about the use of children as ad hoc interpreters led to legislation being introduced in California to ban children as interpreters in health care settings.[28]

Study strengths include the largest sample to date to examine interpreter errors, and recruitment at the 2 largest pediatric EDs in Massachusetts. In addition, this is the first study (to our knowledge) to include encounters in which no interpreter was present, to compare professional hospital versus ad hoc versus no interpreters, and to examine associations of professional interpreters' hours of training and years of experience with interpreter errors.

The study findings demonstrate that interpreter errors of potential clinical consequence are significantly more likely to occur when there is an ad hoc or no interpreter, compared with a professional interpreter. Indeed, no significant difference was found between ad hoc interpreters and having no interpreter in the proportion of errors of potential clinical consequences. These findings are consistent with those of a recent comprehensive systematic review that concluded that optimal communication, patient satisfaction, and outcomes and the fewest interpreter errors occur when limited English proficiency patients have access to trained professional interpreters or bilingual providers.[2] The findings also indicate that ad hoc interpreters and having no interpreter are suboptimal for encounters with limited English proficiency patients and represent strong evidence for avoidance altogether of using ad hoc interpreters or no interpreter when caring for limited English proficiency patients. A presidential executive order issued in 2000 stated that "each Federal agency shall examine the services it provides and develop and implement a system by which [limited English proficiency] persons can meaningfully access those services," and also that, "programs that serve a few or even one [limited English proficiency] person are still subject to the Title VI obligation to take reasonable steps to provide meaningful opportunities for access."[29] Subsequent guidance issued in 2003 by the Department of Health of Human Services,[30] however, was criticized by many organizations[31] for weakening standards, reducing expectations that all recipients of federal funds must undertake at least some action to provide language services, allowing the use of ad hoc interpreters (including minors), and permitting costs or limited resources as mechanisms for essentially opting out of requirements to provide meaningful access to language services for limited English proficiency individuals. A recent review of the literature on use of interpreters in the EDs concluded that professional interpreters ". . . are largely under-utilized in ED settings."[32] The findings of that literature review, together with our study results, underscore the need for third-party reimbursement for interpreter services, which has been shown to result in double the odds of use of professional interpreters,[33] but currently occurs in only 13 states and the District of Columbia.[34]

For professional hospital interpreters, the years of experience interpreting were associated with neither the number or category of interpreter errors, nor their potential clinical consequences. The lack of association of years of interpreting experience with outcomes is consistent with the hypothesis that interpreters with little or no training might continue to commit the same numbers and types of errors over time. For example, one could speculate that an interpreter who did not receive sufficient initial and ongoing training on medical terminology and how to avoid common interpretation errors might easily continue to make the same numbers and types of errors over many years, without sufficient opportunities for correction and learning.

In contrast, the number of hours of previous training was significantly associated with the number and types of errors and their clinical consequences, with interpreters with greater than or equal to 100 hours of training experiencing the most favorable outcomes. One possible explanation for this finding is

that extensive interpreter training is more likely to include developing knowledge and skills about medical terminology, adhering to guidelines for effective interpretation, and learning how to avoid common interpreter errors. A recent National Council on Interpreting in Health Care report[35] highlighted that the duration of medical interpreter training programs in the United States varies from as few as 2 hours to 200 or more hours. In addition, this article did not specifically recommend how long introductory training programs should be. Our study findings therefore provide the first evidence-based guidance on a useful minimum number of hours (100) for medical interpretation programs to consider.

Additional research is warranted on the optimal content of medical interpreter training, particular in relation to how best to reduce interpreter errors and their potential clinical consequences. The study findings, however, suggest that requiring at least 100 hours of training for medical interpreters might have a major impact on reducing interpreter errors and their potential clinical consequences in health care, while potentially improving the quality of care and patient safety.

The authors thank Donglin Sun for his assistance with the initial statistical analyses. The authors are indebted to Guadalupe Pacheco for his unwavering support of and invaluable assistance with this research. The authors are grateful to Leonardo Medina, Roberto Beltran, Ivan Villon, and Patricia Francisco for their invaluable efforts in collecting and analyzing data. The authors are indebted to the many patients, families, and interpreters who participated in this study.

Supervising editor: Robert L. Wears, MD, PhD

Author contributions: GF had full access to all study data, takes responsibility for the data integrity and accuracy of the data analysis, was responsible for study concept and design, obtained funding, and provided administrative, technical, and material support. GF, MA, CPB, and RB were responsible for acquisition of data. GF, MA, and HL were responsible for analysis and interpretation of data. All authors drafted the article and were responsible for critical revision of the article for important intellectual content and final approval of the version to be published. HL was responsible for statistical analysis. GF and MA supervised the study. GF takes responsibility for the paper as a whole.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). The authors have stated that no such relationships exist. Funded in part by a grant from the Office of Minority Health, US Department of Health and Human Services.

Publication dates: Received for publication November 3, 2011. Revisions received December 21, 2011, and January 11, 2012. Accepted for publication January 25, 2012. Available online March 15, 2012.

Presented at the Pediatric Academic Societies annual meeting, May 2003, Seattle, WA; and the AcademyHealth annual meeting, June 2003, Nashville, TN.

Address for correspondence: Glenn Flores, MD, E-mail glenn.flores@utsouthwestern.edu.

REFERENCES

1. Lewis MP (ed). *Ethnologue: Languages of the World.* 16th ed. Dallas, TX: SIL International; 2009. Available at: http://www.ethnologue.com/. Accessed August 18, 2011.
2. 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.
3. US Census Bureau. Language spoken at home: 2010 American community survey 1-year estimates. Available at: http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_10_1YR_S1601&prodType=table. Accessed December 21, 2011.
4. Hayashi D. Guidance memorandum. January 29, 1998. *Title VI Prohibition Against National Origin Discrimination—Persons With Limited-English Proficiency*. Washington, DC: US Department of Health & Human Services; 1998.
5. Baker DW, Parker RM, Williams MV, et al. Use and effectiveness of interpreters in an emergency department. *JAMA*. 1996;275:783-788.
6. US Census Bureau. Language use and English-speaking ability: 2000. Census 2000 brief. October 2003. Available at: http://www.census.gov/prod/2003pubs/c2kbr-29.pdf. Accessed August 18, 2011.
7. 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.
8. Barik H. *Simultaneous Interpretation: Temporal and Quantitative Data* [unpublished PhD dissertation]. Chapel Hill, NC: University of North Carolina; 1972.
9. Hornberger JC, Gibson CD Jr, Wood W, et al. Eliminating language barriers for non-English-speaking patients. *Med Care*. 1996;34:845-856.
10. Flores G, Torres S, Holmes LJ, et al. Access to hospital interpreter services for limited English proficient patients in New Jersey: a statewide evaluation. *J Health Care Poor Underserved*. 2008;19:391-415.
11. Hudelson P, Vilpert S. Overcoming language barriers with foreign-language speaking patients: a survey to investigate intra-hospital variation in attitudes and practices. *BMC Health Serv Res*. 2009;9:187.
12. Gany FM, Gonzalez CJ, Basu G, et al. Reducing clinical errors in cancer education: interpreter training. *J Cancer Educ*. 2010;25:560-564.
13. 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.
14. Rosenberg E, Richard C, Lussier MT, et al. The content of talk about health conditions and medications during appointments involving interpreters. *Fam Pract*. 2011;28:317-322.
15. Flores G, Abreu M, Schwartz I, et al. The importance of language and culture in pediatric care: case studies from the Latino community. *J Pediatr*. 2000;137:842-848.
16. Flores G. Language barrier. AHRQ WebM&M morbidity and mortality rounds on the Web 2006; April. Available at: http://www.webmm.ahrq.gov/case.aspx?caseID=123. Accessed February 7, 2012.
17. Harsham P. A misinterpreted word worth $71 million. *Med Econ*. 1984;June:289–292.

18. Cohen AL, Rivara F, Marcuse EK, et al. Are language barriers associated with serious medical events in hospitalized pediatric patients? *Pediatrics*. 2005;116:575-579.
19. 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.
20. Jackson JC, Nguyen D, Hu N, et al. Alterations in medical interpretation during routine primary care. *J Gen Intern Med*. 2011;26:259-264.
21. 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.
22. Flores G, Ngui E. The impact of racial and ethnic disparities on patient safety. *Pediatr Clin North Am*. 2006;53:1197-1215.
23. Institute of Medicine. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, DC: National Academies Press; 2009.
24. Institute of Medicine. *Child and Adolescent Health and Health Care Quality: Measuring What Matters*. Washington, DC: National Academies Press; 2011.
25. David RA, Rhee M. The impact of language as a barrier to effective health care in an underserved urban Hispanic community. *Mount Sinai J Med*. 1998;65:393-397.
26. Ebden P, Carey OJ, Bhatt A, et al. The bilingual consultation. *Lancet*. 1988;1:347.
27. Launer J. Taking medical histories through interpreters: practice in a Nigerian outpatient department. *British Med J*. 1978;2:934-935.
28. Yee L. California State Senate. Assembly passes bill to end the use of children as interpreters in medical situations. Available at: http://dist08.casen.govoffice.com/index.asp?Type=B_PR&SEC=%7BDA377F1D-8DF2-4AE6-B3A0-7909404FC1D1%7D&DE=%7B986FEEC3-6987-4597-A5C8-56368EEB2D7B%7D. Accessed August 18, 2011.
29. Clinton WJ. Executive order: improving access to services for persons with limited English proficiency. August 11, 2000. Available at: http://www.justice.gov/crt/about/cor/Pubs/eolep.pdf. Accessed December 21, 2011.
30. US Department of Health and Human Services. Guidance to federal financial assistance recipients regarding Title VI prohibition against national origin discrimination affecting limited English proficient persons. Available at: http://www.hhs.gov/ocr/civilrights/resources/laws/revisedlep.html. Accessed December 21, 2011.
31. National Health Law Program. Final comments on OCR's LEP guidance. Available at: http://www.healthlaw.org/index.php?option=com_content&view=article&id=331:final-comments-on-ocrs-lep-guidance&catid=45. Accessed December 21, 2011.
32. 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.
33. Kuo DZ, O'Connor KG, Flores G, et al. Pediatricians' use of language services for families with limited English proficiency. *Pediatrics*. 2007;119:e920-927.
34. Youdelman M. Medicaid and SCHIP reimbursement models for language services 2009 update. National Health Law Program. Available at: http://www.scribd.com/doc/38667358/Medicaid-and-SCHIP-Reimbursement-Models-for-Language-Services-2009-Update. Accessed August 18, 2011.
35. National Council on Interpreting in Health Care. National standards for healthcare interpreter training programs. 2011. Available at: http://data.memberclicks.com/site/ncihc/National_Standards_5-09-11.pdf. Accessed August 9, 2011.