UCSF

Title

Assessing the Use of Google Translate for Spanish and Chinese Translations of Emergency Department Discharge Instructions

Permalink

https://escholarship.org/uc/item/2md2s39x

Journal

JAMA Internal Medicine, 179(4)

ISSN

2168-6106

Authors

Khoong, Elaine C
Steinbrook, Eric
Brown, Cortlyn
et al.

Publication Date

2019-04-01

DOI

10.1001/jamainternmed.2018.7653

Peer reviewed

Letters

Assessing the Use of Google Translate for Spanish and Chinese Translations of Emergency Department Discharge Instructions

Patients with limited English proficiency experience communication barriers to health care in English-speaking countries. Written communication improves comprehension,[1] but pretranslated standard instructions cannot address patient-specific issues (eg, medication titration). Machine translation tools, including Google Translate (GT), have potential to improve communication with these patients, but prior studies showed limited accuracy; 1 study found that GT Spanish translations of patient education materials were 60% accurate, with 4% resulting in serious error.[2]

In 2017, GT changed its translation algorithm, claiming significant improvement.[3] In this study, we assess the use of GT to translate emergency department (ED) discharge instructions into Spanish and Chinese.

Methods | We abstracted 100 free-texted ED discharge instructions and oversampled for medication changes and common complaints.[4] We analyzed each sentence by content category; Flesch-Kincaid readability score; use of medical jargon,[5] such as atypical use of normal words (eg, positive test result) or medical terminology; and presence of nonstandard English (spelling or grammar errors, abbreviations, colloquial English, proper nouns). Content categories included explanation of diagnosis and/or results, follow-up instructions, medication instructions, return precautions, and greeting.

Using GT we translated instructions into Spanish and Chinese, and then bilingual translators translated the text back into English.

The primary outcome was sentence translation accuracy, assessed for overall content accuracy, not word-for-word accuracy, and coded as a binary outcome. Two clinicians coded accuracy independently; a third adjudicated disagreements. A second translator reviewed back-translations deemed inaccurate to ensure these were not back-translator error.

Potential for harm from inaccurate translations was assessed by 2 clinicians (with a third adjudicating) using an established rating system: clinically nonsignificant, clinically significant, and life-threatening potential harm.[6] For analyses, we used a binary variable (clinically significant/life-threatening vs clinically nonsignificant/no harm).

We used logistic regression analyses stratified by language to assess associations between sentence characteristics and accuracy and/or harm. Variables with significance of $P < .20$ in bivariate analyses were used in multivariable analyses.

Results | The 100 sets of patient instructions contained 647 sentences. Overall, 594 (92%) and 522 (81%) sentences were accurately translated into Spanish and Chinese, respectively, by GT (**Table 1**). A minority of inaccurate translations had potential for clinically significant harm: in Spanish, 15 (28%) of 53 inaccuracies and 15 (2%) of 647 sentences; in Chinese, 50 (40%)

Table 1. Characteristics of Inaccurately Translated Sentences and Clinically Significant Potential Harm From Inaccurate Translations

Characteristic	Accuracy		Harm	
	No. Inaccurate/No. Sentences (%)		No. Potential Harm/No. Sentences (%)	
	Spanish	Chinese	Spanish	Chinese
All sentences	53/647 (8)	125/647 (19)	15/647 (2)	50/647 (8)
Readability of full instructions				
Flesch-Kincaid >8	27/248 (11)[a]	53/248 (21)	11/248 (4)[a,b,c]	24/248 (10)[a]
Sentence content				
Explain diagnosis/results	23/307 (7)	55/307 (18)	5/307 (2)	16/307 (5)[a,b]
Follow-up instructions	14/160 (9)	35/160 (22)	8/160 (5)[a,b,c]	19/160 (12)[a,b]
Medication instructions	7/82 (9)	21/82 (26)[a]	5/82 (6)[a,b]	13/82 (16)[a,b]
Return precautions	14/107 (13)[a,b]	25/107 (23)	2/107 (2)	10/107 (9)
Greeting	1/23 (4)	0/23 (0)[a]	0/23 (0)	0/23 (0)[a]
Medical jargon				
Atypical use of normal words	3/31 (10)	8/31 (26)	1/31 (3)	1/31 (3)
Medical terminology	22/251 (9)	57/251 (23)[a]	7/251 (3)	29/251 (12)[a,b,c]
Nonmedical nonstandard English				
Spelling/grammar anomalies	9/44 (20)[a,b,c]	18/44 (41)[a,b,c]	4/44 (9)[a,b]	11/44 (25)[a,b,c]
Abbreviations	16/150 (11)	33/150 (22)	3/150 (2)	13/150 (9)
Colloquial English	3/10 (30)[a,b]	5/10 (50)[a,b]	0/10 (0)	4/10 (40)[a,b,c]
Proper nouns	4/44 (9)	7/44 (16)	2/44 (5)	2/44 (5)

[a] $P < .20$, and so included in multivariable regression analyses.

[b] $P < .05$ for bivariate analysis, and so included in multivariable regression analyses.

[c] $P < .05$ in multivariable regression analyses (all supporting data reported in the Results section).

 E1

Table 2. Examples of Inaccurate Translations and Associated Level of Potential Clinical Harm

Characteristic	Original Sentence	Back-Translated Spanish	Level of Potential Harm	Back-Translated Chinese	Level of Potential Harm
Sentence readability, >8th grade	Please follow up with your primary care doctor tomorrow to get a refill on all your medications and seek immediate medical care if you develop confusion, severe abdominal pain or nausea vomiting	Please perform a follow-up with your primary care doctor tomorrow to obtain a resupply of all your medications and look for immediate medical attention if you develop confusion, intense abdominal pain or vomits nausea	NA—accurate translation	If you feel confused, have serious stomachache or nausea and vomiting, please go to your family doctor to get your medications refilled tomorrow and seek for medical treatment immediately	Clinically significant
Atypical use of normal words	Hold the kidney medicine until you have a chance to speak with your kidney doctor	Keep the medication for the kidney until you have the chance to talk with your kidney doctor	Life-threatening	Keep taking kidney medicine until you talk to your kidney doctor	Life-threatening
Medical terminology	You do have an abdominal aortic aneurysm (an outpouching of the major blood vessel in your body) which will need to be followed by your primary care doctor regularly	Aortic aneurysm (an evacuation of the main blood vessel in your body) which your primary care doctor must follow up regularly	Clinically significant	Your stomach abdominal aortic aneurysm (exudation of the internal major blood vessel) needs to be examined regularly by PCP	Clinically significant
Spelling/grammar anomalies	You have a low back stain[a]	You have a low patch on your back[b]	Clinically nonsignificant	You have a low back spot[b]	Clinically nonsignificant
Spelling/grammar anomalies	You will be called with an appointment for tomorrow ophthalmology	You will be called with an appointment for morning ophthalmology	Clinically nonsignificant	You will be arranged to see an eye doctor tomorrow	Clinically nonsignificant
Spelling/grammar anomalies	Please take these medications as follows: Ibuprofen 400 mg every 6 h as needed for inflammation and pain Valium 5 mg every 8 h for muscle spasms Lidocaine patches on for 12 h, off for 12 h	Take these medications in the following way: 400 mg Ibuprofen every 6 h as it may be necessary for inflammation and pain 5 mg Valium every 8 h for muscular spasms Lidocaine is plugged during 12 h, it is turned off during 12 h	Clinically significant	Please take the following medicines: Ibuprofen every 6 h 400 mg, every 6 h according to infection and pain needs every 8 hours take 5 mg muscle spasm Ligudane supplement 12 h close 12 h	Clinically significant
Abbreviations	You were seen in the ED today for your weakness and difficulty speaking	You were seen today at the emergency service due to your weakness and difficulty to talk	NA—accurate translation	Today you come to the education department room because of your weaknesses and difficulties	Clinically nonsignificant
Colloquial English	Please return to the emergency department for worsening abdominal pain, inability to eat or drink due to vomiting, bloody diarrhea, if you pass out or any other concerning symptom	Please return to the emergency service for worsening of abdominal pain, inability to eat or drink due to vomits, diarrhea with blood, if it happens outside or any other concentrated symptom	Clinically nonsignificant	If you pass, vomiting, bloody diarrhea, unable to eat or drink, or any other symptoms; come back to the emergency department for treatment	Clinically significant

Abbreviations: ED, emergency department; NA, not applicable; PCP, primary care physician.

[a] Clinician mistyped "low back strain."

[b] Although these are accurate translations of the English words, overall content accuracy differs from the English language instructions, which would be clear in context.

of 125 inaccuracies and 50 (8%) of 647 sentences. Some errors were correct translations of errant English instructions, but overall, content was inaccurate owing to grammar or typographical errors (Table 2) that would readily have been overlooked or understood by a reader of the English text.

Only spelling and grammar anomalies were associated with inaccurate translations in multivariable analyses: Spanish (odds ratio [OR], 2.6; 95% CI, 1.1-5.8); Chinese (OR, 2.6; 95% CI, 1.3-5.0).

In multivariable analyses, potential harm was associated in Spanish with a Flesch-Kincaid reading level higher than eighth grade (OR, 4.0; 95% CI, 1.2-13.5) and follow-up instructions (OR, 3.5; 95% CI, 1.2-10.2); and in Chinese with medical terminology (OR, 2.4; 95% CI, 1.2-4.9), spelling or grammar anomalies (OR, 3.1; 95% CI, 1.4-7.2), and colloquial English (OR, 5.9; 95% CI, 1.4-24.7).

Discussion | Discharge instructions were translated by the new GT algorithm with higher accuracy and fewer seriously harmful inaccuracies than previously,[2] yet 2% of Spanish and 8% of Chinese sentence translations had potential for significant harm. While GT can supplement (not replace) written English instructions, machine-translated instructions should include a warning about potentially inaccurate translations.

Clinicians using GT can reduce potential harm by having patients read translations while receiving verbal instructions; being vigilant about spelling and grammar; and avoiding complicated grammar, medical jargon (eg, fingerstick), and colloquial English.

Study limitations include assessment of only 2 languages (though our inclusion of Chinese is a strength, since non-European languages are often less accurately translated by machines); no assessment of translation readability; and no comparison to human translators.

Google Translate can be used to translate clinician-entered, patient-specific ED instructions for Spanish- and Chinese-speaking patients. Potential for harm can be minimized by using clear communication practices. We recommend including English instructions and automated warnings regarding the use of machine translation.

Elaine C. Khoong, MD, MS
Eric Steinbrook, BA
Cortlyn Brown, MD
Alicia Fernandez, MD

Author Affiliations: Division of General Internal Medicine, Department of Medicine at Zuckerberg San Francisco General Hospital, University of California, San Francisco (Khoong, Fernandez); University of Michigan School of Medicine, Ann Arbor (Steinbrook); Department of Emergency Medicine, University of California, San Francisco (Brown); Center for Vulnerable Populations at University of California, San Francisco (Fernandez).

Corresponding Author: Elaine C. Khoong, MD, MS, Division of General Internal Medicine, Department of Medicine at Zuckerberg San Francisco General Hospital, University of California, San Francisco, 1001 Potrero Ave, 1M, San Francisco, CA 94122 (elaine.khoong@ucsf.edu).

Author Contributions: Dr Khoong had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Khoong, Brown, Fernandez.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: All authors.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Khoong, Brown.
Administrative, technical, or material support: Steinbrook, Fernandez.
Study supervision: Fernandez.

Conflict of Interest Disclosures: None reported.

1. Johnson A, Sandford J, Tyndall J. Written and verbal information versus verbal information only for patients being discharged from acute hospital settings to home. *Cochrane Database Syst Rev*. 2003;4(4):CD003716. doi:10.1002/14651858.CD003716

2. Khanna RR, Karliner LS, Eck M, Vittinghoff E, Koenig CJ, Fang MC. Performance of an online translation tool when applied to patient educational material. *J Hosp Med*. 2011;6(9):519-525. doi:10.1002/jhm.898

3. Wu Y, Schuster M, Chen Z, et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. https://arxiv.org/abs/1609.08144. Accessed January 17, 2019.

4. Weiss AJ, Wier LM, Stocks C, Blanchard J. Overview of Emergency Department Visits in the United States, 2011. Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project Statistical Brief #174. June 2014. https://www.hcup-us.ahrq.gov/reports/statbriefs/sb174-Emergency-Department-Visits-Overview.pdf. Accessed January 17, 2019.

5. Castro CM, Wilson C, Wang F, Schillinger D. Babel babble: physicians' use of unclarified medical jargon with patients. *Am J Health Behav*. 2007;31(suppl 1):S85-S95. doi:10.5993/AJHB.31.s1.11

6. Nápoles AM, Santoyo-Olsson J, Karliner LS, Gregorich SE, Pérez-Stable EJ. Inaccurate language interpretation and its clinical significance in the medical encounters of Spanish-speaking Latinos. *Med Care*. 2015;53(11):940-947. doi:10.1097/MLR.0000000000000422