A Comparative assessment of ChatGPT *vs.* Google Translate for the translation of patient instructions

Pavithra Rao[1]^, Lauren M. McGee[2], Casey A. Seideman[2]

[1]Oregon Health & Science University School of Medicine, Portland, OR, USA; [2]Department of Urology, Oregon Health & Science University, Portland, OR, USA

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Pavithra Rao, BS. Oregon Health & Science University School of Medicine, 3266 SW Research Dr Mackenzie Hall 4th Floor, Portland, OR 97239, USA. Email: raopa@ohsu.edu.

Background: An estimated 30 million people in the United States demonstrate a non-English language preference (NELP). Professional interpreter services and NELP patient advocates can mitigate some barriers to quality care but are not readily available at all institutions. Many studies have explored the use of Google Translate (GT) for medical communication with mixed results, most concluding that GT alone should not be used in the translation of medical documents for patients. Studies of medical translation with newer artificial intelligence (AI) language models, particularly ChatGPT, do not exist to date. This study's purpose is to compare ChatGPT *vs.* GT in its ability to accurately translate instructional and educational medical documents across multiple languages.

Methods: ChatGPT-3.5 and GT were prompted to translate two English language documents: post-operative discharge instructions for care after circumcision and patient information for undescended testicles (UDT). The instructions were copied and pasted into ChatGPT and GT, and translated into Spanish, Vietnamese, and Russian. Reviewers were blinded as to which documents corresponded to ChatGPT and GT. Members of the Language Services department independently reviewed the translations and assessed them based on meaning, expression, and technical errors. Each translation error was recorded with the reviewer's comment and reviewer's preferred translation. The errors were categorized as meaning, flow, language, form, style guide, or terminology errors, and were labeled as minor, major, or critical. Primary outcome assessed was accuracy of the translation, largely based on retention of meaning.

Results: There were 132 sentences between the two documents. In Spanish, ChatGPT incorrectly translated 3.8% of all sentences, while GT incorrectly translated 18.1% of sentences. In Russian, ChatGPT and GT incorrectly translated 35.6% and 41.6% of all sentences, respectively. In Vietnamese, ChatGPT and GT incorrectly translated 24.2% and 10.6% of sentences, respectively.

Conclusions: ChatGPT excelled considerably compared to GT for Spanish translation. However, it was outperformed by GT for the Vietnamese translations, and both ChatGPT and GT produced low quality translations in Russian. ChatGPT has an unacceptably high rate of translation error in Vietnamese and Russian and should not be used alone to translate medical documents from English to these languages. While it shows promise in the translation of Spanish materials, its utility for additional languages requires further study and it remains unreliable for use without human oversight.

---

^ ORCID: 0000-0002-1550-635X.

Introduction

An estimated 9% of the United States population, or upwards of 30 million people, identify as having limited English proficiency. In the healthcare setting, this has major implications on the quality of care received by patients, with many studies demonstrating that when a patient's preferred language is discordant with that of their healthcare provider's, they receive lower quality care and are less satisfied with their healthcare interactions (1-7). Thus, the emergence of patient-centered terminology, "non-English language preference" (NELP), has been proposed to help identify individuals presenting to a healthcare setting who may benefit from additional resources. This language challenges the healthcare system to shift the perception of a patient's "lack of ability" to communicate in English and places the burden on the healthcare system itself, which is often poorly equipped to provide the tools needed to communicate effectively in a patient's preferred language (8).

Utilization of language-concordant providers, NELP

patient family advocates, and professional translation services are all tools hospitals have used with success to mitigate some of these barriers to quality care and improve patient satisfaction (9). However, these services are not always readily available and have associated costs; thus, institutions may turn to machine translation services to bridge these gaps (10,11). One important distinction is that translation involves written materials while interpreters are used during verbal communication. Google Translate (GT) has, to date, been a popular machine translation service used by researchers to evaluate accuracy in translation of preexisting documents. As of 2016, GT translates largely via Google Neural Machine Translation (GNMT), a deep learning artificial neural network that aims to translate whole sentences at a time instead of piece by piece, which was the basis of its initial statistical machine translation capabilities. Despite improvements in development, GT has not demonstrated independence in the translation of medical documents from English into target languages, performing especially poorly when translating into non-European languages (12-15).

This issue was well-demonstrated in a study by Taira *et al.* in 2021, in which volunteers assessed twenty commonly used discharge statements from a single Emergency Department that had been translated into seven commonly spoken languages (Spanish, Chinese, Vietnamese, Tagalog, Korean, Armenian, and Farsi). Their results showed that while overall meaning was retained for 82.5% of the translations, accuracy varied widely depending on the translated language, with Spanish and Tagalog maintaining 94% and 90% accuracy rates, while Farsi and Armenian only produced accuracy in 67.5% and 55% of all statements, respectively (13). Similar results have been noted in other studies, with Khoong *et al.* in 2019 showing that Emergency Department discharge instructions translated from English to Spanish and Chinese resulted in inaccuracies in 8% of the Spanish translations and 19% of the Chinese translations, with potential harm coming from 2% of the Spanish instructions and 8% of the Chinese instructions (14). Potential harm to patients may also be exacerbated by increasing complexity of text. A study by Cornelison *et al.* in

---

Highlight box

Key findings

• ChatGPT should not replace a professional translator and should be used with caution, given the high number of errors reported in this study.

What is known and what is new?

• Google Translate has, to date, been the machine translation service of choice used by researchers to evaluate accuracy in translation of preexisting documents. Since reaching consumers in 2022, ChatGPT has been marketed as a superior language processing tool, pulling from a larger corpus of data more equipped to excel at nuance and context across many languages.

• No studies exist comparing these translation engines in a medical context.

What is the implication, and what should change now?

• Given that ChatGPT was not primarily designed for translation purposes, further studies are needed to evaluate its utility in this field. Additionally, healthcare providers must continue to critically examine any artificial intelligence-driven initiatives to guarantee their quality and validity.

Table 1 Description of each translation error category as standardized by the Language Services department at our institution

Category	Description
Flow	Unnatural word order or translations using source-language syntax. Non-idiomatic use of target language
Form	Improper use of translation placeholders, special characters, variables, etc. Misalignments
Language	Grammatical, syntactic, or punctuation errors, including misspellings or typos
Meaning	Incorrect meaning has been transferred; an unacceptable omission or addition in the translated text. Misunderstanding of source language concept. Ambiguity in target text. Incorrectly suppressed content. Improper exact or fuzzy match
Style guide	Guidelines for translating the customer's brand (as indicated in the style guide) were not followed. Includes but is not limited to: language that is not suited to the customer's brand, the target market, or the tone of translation; incorrect handling of product names; incorrect mode/formality. Incorrect formatting of dates, addresses, measurements, etc.
Terminology	Violation of glossary approved by customer, use of banned term, translation of non-translatable term. Redundancy. Incorrect industry-specific terminology

2021 showed that when asking GT to translate counseling points for the top 100 drugs used in the United States from English into Arabic, Chinese, and Spanish, only 54.1% of counseling points in Arabic, 76.5% in Chinese, and 38.2% in Spanish were accurately translated, with 29.1% of inaccurate translations being identified as highly clinically significant or potentially life-threatening (16).

Since reaching consumers in 2022, ChatGPT has been marketed as a superior language processing tool, pulling from a larger corpus of data more equipped to excel at nuance and context across many languages. Patients are typically sent home from the inpatient and outpatient setting with written instructions not in their preferred language. We hoped to address this gap in care by comparing the accuracy of ChatGPT *vs*. GT in the translation of two forms of educational materials. We hypothesized that ChatGPT would produce fewer errors and outperform GT across multiple languages, and that it would be safe for use in the translation of discharge materials for patients. No studies exist comparing these translation engines in a medical context.

Methods

Data creation

In May 2023, ChatGPT-3.5 was prompted to translate post-operative discharge instructions for circumcision and patient information for undescended testicles (UDT) into Spanish, Vietnamese, and Russian. These languages were chosen as they are common to this institution's local population. These two handouts were chosen because they represent the most commonly utilized instructions given to patients in a pediatric urology practice. Pre-existing instructions were written in English and utilized by a single institution (see Appendices 1,2). The authors chose to use the following standardized prompt:

"Translate the following into [indicated language]"

This was followed by copying and pasting the instructions into the field. The pre-existing instructions in English were copied and pasted directly into GT as no prompt is needed for this software.

Error categorization and data review

In determining the parameters by which to judge translation errors, we turned to the field of Translation Quality Assessment (TQA). Popularized by Juliane House in the 1990s, TQA is a dynamic model by which the quality of translations can be assessed (17). At its core, it relies on three general parameters: meaning, consistency, and errors. By using pre-defined categories in concordance with TQA models, professional Spanish, Vietnamese, and Russian translators from the Language Services department at our institution independently reviewed the translations and assessed them based on meaning, expression, and technical errors (*Table 1*). The reviewers within the Language Department all have specialized training and certification in both written and oral translation, ensuring appropriate expertise to provide written translation services.

Based on an internally developed coding scheme from the Department of Language Services, the translated
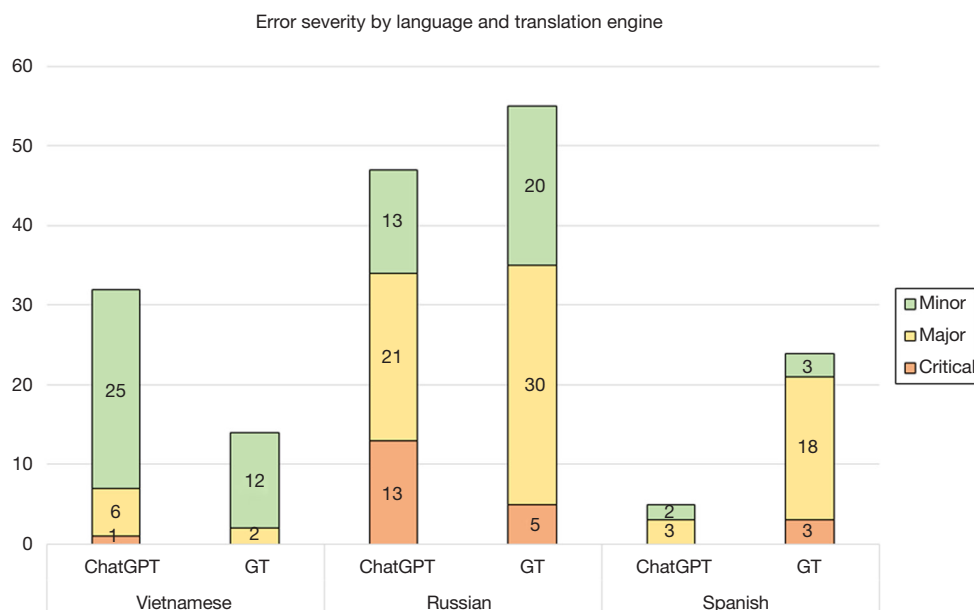
Figure 1 Total minor, major, and critical errors across all translated documents for ChatGPT *vs.* GT by language. GT, Google Translate.

documents were labeled as "Engine 1", which always corresponded to ChatGPT, and "Engine 2", which always corresponded to GT, in order to blind reviewers. Each translation error was recorded with the reviewer's comment and reviewer's preferred translation. Using a standardized template, the errors were categorized as meaning, flow, language, form, style guide, or terminology errors, and were labeled as minor, major, or critical. Minor errors often involved organization, flow, or grammar, and were those that did not detract from the overall message. Major and critical errors were those that led to significant alterations in the meaning of the text, with critical errors being those that could confer harm. The primary outcome assessed was accuracy of the translation, largely based on retention of meaning.

Results

There were 132 total sentences between the two documents. In Spanish, ChatGPT incorrectly translated 3.8% of all sentences, while GT incorrectly translated 18.1% of sentences. In Russian, ChatGPT and GT incorrectly translated 35.6% and 41.6% of all sentences, respectively. In Vietnamese, ChatGPT and GT incorrectly translated 24.2% and 10.6% of sentences, respectively. Proportions of critical, major, and minor errors can be seen in *Figure 1*.

In total, GT committed more errors compared to

ChatGPT (93 *vs.* 84). Most errors for both ChatGPT and GT came from Russian translation (47 *vs.* 55 total errors, respectively). ChatGPT outperformed GT in Spanish (5 *vs.* 24 total errors) but was surpassed by GT for Vietnamese translation (32 *vs.* 14 total errors). The majority of errors made by ChatGPT were related to meaning (46%) and language (26%), while most GT errors were due to flow (48%) and meaning (28%). The distribution of errors can be seen in *Figure 2*.

Examples of translation errors can be seen in *Tables 2,3*. Based on translator feedback, both the ChatGPT and GT Spanish translations were true to the source in terms of meaning, despite some unnatural sounding phrases, few mistranslations, and occasional disorganization from mixed verbal tenses. Overall, the two translations were comparable, though ChatGPT had significantly fewer total errors compared to GT.

Notably, the ChatGPT translation for both Vietnamese and Russian had the same critical error of taking the last paragraph of the circumcision file and completely replacing and making up information compared to the source. This did not happen with GT. While the ChatGPT and GT Russian translations both had several grammar and fluency issues, they were largely equivocal in number of errors and overall quality of translation. However, for Vietnamese, it was noted that GT provided acceptable adherence to the source text and a better writing style compared
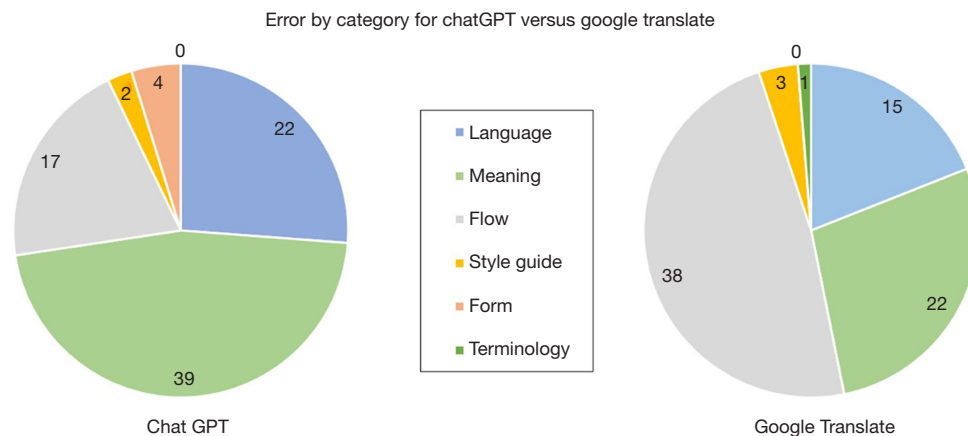
Error by category for chatGPT versus google translate

Figure 2 Translation errors by type performed by ChatGPT *vs.* Google Translate across all documents.

to ChatGPT, which was noted to have subpar fluency, unnatural translations, and many meaningless translations.

Discussion

This is the first study evaluating the accuracy of ChatGPT's translation capabilities in a medical context, and the only study available at this time directly comparing the functionality of ChatGPT to GT in this capacity. Although GT is currently the most popular translation machine, ChatGPT has the potential to deliver similar results. Our study shows that ChatGPT may be reliable in translating medical information to Spanish but was not as effective in translating to other languages. These results are similar to other studies which also note greater deficits in machine translation with regards to non-European languages (13). Given that ChatGPT was designed using more advanced training parameters, we expected it to perform with higher fluency and accuracy compared to GT across all studied languages, and for it to produce useable material for distribution to patients. However, these results provided a more varied and nuanced outcome than expected. Though the translated documents from ChatGPT into Spanish largely retained their original meaning with only small errors in grammar and tone, the number and severity of errors that ChatGPT produced in Russian and Vietnamese rendered these documents unsafe for use in patient-facing settings. While tempting to turn to large language models in an effort to improve the quality of care for NELP patients, especially with increasingly powerful artificial intelligence (AI) tools at our fingertips, this study suggests that such practices are likely unsafe, and should not yet be

utilized in practice.

Limitations

There are several limitations to this study. We opted to translate only the two most utilized instruction handouts within our clinic's subspecialized practice, and our findings may not be readily generalizable to broader medical contexts. Only three languages were translated, and ChatGPT's translation capabilities may vary even more significantly with other languages. We only utilized one translator per language, subjecting the study to potentially significant inter-rater variability. Additionally, the translators may vary in their levels of training and thresholds for assessing meaning and error. Consequently, the determination of errors and their severity is subjective by nature. For example, the higher percentage of critical errors seen in the Russian translations could be a result simply of individual interpretation of the text, rather than the true rate of critical errors. However, the threshold for defining an error was less subjective, and the overall number of errors identified in each text should not have been as susceptible to this bias. Since we used single coding, reviewers may have also noticed patterns and become more critical of either Engine 1 or Engine 2.

The accuracy of translations will likely differ depending on the type of medical documentation as well, as a technical operative report would be expected to have a lower quality translation compared to an "easy to read" patient hand-out. Additionally, pediatric urology is a highly subspecialized field, and further studies are needed in other specialty settings to determine generalizability of the results. Another

Table 2 Examples of errors by type

Variables	Source text	Original translation	Reviewer's comment	Reviewer's preferred translation
Flow	Any stitches will dissolve on their own. The doctor does not need to take them out	Любые швы сами рассосутся, врачу не нужно их удалять	The translation has unnatural flow due to being very literal	The translation has unnatural flow due to being very literal
Form	This is because having general anesthesia can affect their coordination	Điều này b ởi vì gây mê tổng quát có thể ảnh hưởng đến sự phối hợp của họ	The sentence is corrupted	Điều này bởi vì gây mê tổng quát có thể ảnh hưởng đến sự phối hợp của trẻ
Language	Give your child apple, pear or prune juice to drink	Давайте вашему ребенку яблочный, грушевый или сливовый сок на питье	Items in the bullet list should have identical grammar form. Also, the translation needs rephrasing for a more natural flow	Давать ребенку яблочный, грушевый или сливовый сок
Meaning	If your child is hungry give them a small, light meal	Нếu trẻ của bạn đói, cho họ một bữa nhẹ nhàng, ít nhiều	Wrong translation for "small". Not appropriate translation for "them", "light"	Nếu trẻ của bạn đói, hãy cho trẻ ăn một bữa nhẹ, nhỏ
Style guide	Call us at 503-494-4808 (8 a.m.–5 p.m., Monday–Friday) if you have questions, concerns or your child has any of the following	Позвоните нам по номеру 503-494-4808 (с 8:00 до 17:00 с понедельника по пятницу), если у вас есть вопросы, опасения или у вашего ребенка есть следующие симптомы	Time should not be converted into 24 h format	Позвоните нам по номеру 503-494-4808 (с 8 a.m. до 5 p.m. с понедельника по пятницу), если у вас есть вопросы и опасения или у вашего ребенка есть следующие симптомы
Terminology	Check your child's penis each time you change his diaper (every 3 to 4 hours) or when he goes to the bathroom	Проверяйте пенис вашего ребенка каждый раз, когда вы меняете ему подгузник (каждые 3–4 часа) или когда он идет в туалет	Inconsistent translation of the key term	Проверяйте половой член вашего ребенка каждый раз, когда вы меняете ему подгузник (каждые 3–4 часа) или когда он идет в туалет

consideration is that different dialects might impact translation outputs, especially if words used in a specific region are not commonly utilized in the "traditional" language model used for these translation services. It is important that the documents are U.S. Health Insurance Portability and Accountability Act (HIPAA) compliant, and precautions would be needed to ensure that patient-sensitive information was not accidentally used. Finally, we utilized ChatGPT-3.5 instead of the newer version of ChatGPT-4. ChatGPT-4 has a higher number of learning parameters, can read and understand images, reportedly experiences fewer instances of fact fabrication (which contributed to critical errors in this study), and is touted as more multilingual with improved accuracy. As such, if this study were repeated with the newer GPT model, we may see ChatGPT perform more favorably against our designated parameters. We used general prompts as well, and further studies are needed to evaluate how performance varies with increasing specificity.

Given that ChatGPT was not primarily designed for translation purposes, further studies are needed to evaluate its utility in this field. This study encompasses a small comparison of the language capabilities of ChatGPT and GT, and any future work should ideally include both a greater breadth of medical phrases for improved generalizability across specialties, and a wider variety of languages. Additionally, healthcare providers must continue to critically examine any AI-driven initiatives to guarantee their quality and validity. These safety measures are essential for adhering to professional and ethical standards and maintaining patient safety.

Conclusions

ChatGPT should not replace a professional translator and

Table 3 Examples of inaccurate translations, error type, and severity of error

Variables	Language	Original sentence	Specified error	Error type	Severity of mistake
ChatGPT	Spanish	In the womb, the testicles form inside the abdomen and then move down (descend) through a space between the groin muscles into the scrotum	Omission of "move down (descend)"	Meaning	Major
	Russian	• Swelling (a small amount is normal) • Bleeding (small dots of blood is normal) • Fever of 101 degrees or higher • Redness, pain that is getting worse or pus coming from the surgery area • Pain even after pain medicine is given • Not eating or drinking well, or unusually fussy or tired	The translation does not at all correspond to the source	Meaning	Critical
	Vietnamese	Give pain medications less often or not at all if you think your child is comfortable without medication	Incorrect translation of "give pain medications less often"	Language	Minor
Google Translate	Spanish	Babies who are born early are most at risk for this condition	"Born early" is not idiomatic but was written literally	Flow	Minor
	Russian	You need to call our office or go to the front desk during business hours to get an account	"Account" is translated as "invoice" which completely changes the meaning of the sentence	Meaning	Critical
	Vietnamese	Your child may have a bandage or clear plastic wrap on the surgery area. This can fall off on its own, shortly after surgery or several days later. Take it off if	Incorrectly translated "take off" as "turn off"	Meaning	Major

should be used with caution, given the high number of errors reported in this study. Although many of the errors were minor, involving organization, flow, or grammar, there were a significant number of critical and major errors which altered the meaning of the text. Such errors can have deleterious effects on patient education and outcomes. Patients are likely to lose confidence in their healthcare team if they are provided with documents that are written at subpar fluency or have numerous spelling errors. Therefore, trained translators must examine such AI-generated translations prior to providing them to patients in order to confirm accuracy and reliability. This is especially important for less widely used languages, since ChatGPT will have less data and likely generate poorer quality translations for these.

or integrity of any part of the work are appropriately investigated and resolved. IRB and informed consent are not applicable as this study doesn't involve any human experiment.

References

1. Flores G. The impact of medical interpreter services on the quality of health care: a systematic review. Med Care Res Rev 2005;62:255-99.
2. Yeheskel A, Rawal S. Exploring the 'Patient Experience' of Individuals with Limited English Proficiency: A Scoping Review. J Immigr Minor Health 2019;21:853-78.
3. Schenker Y, Karter AJ, Schillinger D, et al. The impact of limited English proficiency and physician language concordance on reports of clinical interactions among patients with diabetes: the DISTANCE study. Patient Educ Couns 2010;81:222-8.
4. Brooks K, Stifani B, Batlle HR, et al. Patient Perspectives on the Need for and Barriers to Professional Medical Interpretation. R I Med J (2013) 2016;99:30-3.
5. Ramirez N, Shi K, Yabroff KR, et al. Access to Care Among Adults with Limited English Proficiency. J Gen Intern Med 2023;38:592-9.
6. Latif Z, Makuvire T, Feder SL, et al. Experiences of Medical Interpreters During Palliative Care Encounters With Limited English Proficiency Patients: A Qualitative Study. J Palliat Med 2023;26:784-9.
7. Woods AP, Alonso A, Duraiswamy S, et al. Limited English Proficiency and Clinical Outcomes After Hospital-Based Care in English-Speaking Countries: a Systematic Review. J Gen Intern Med 2022;37:2050-61.
8. Ortega P, Shin TM, Martínez GA. Rethinking the Term "Limited English Proficiency" to Improve Language-Appropriate Healthcare for All. J Immigr Minor Health 2022;24:799-805.
9. Gil S, Hooke MC, Niess D. The Limited English Proficiency Patient Family Advocate Role: Fostering Respectful and Effective Care Across Language and Culture in a Pediatric Oncology Setting. J Pediatr Oncol Nurs 2016;33:190-8.
10. Davis SH, Rosenberg J, Nguyen J, et al. Translating Discharge Instructions for Limited English-Proficient Families: Strategies and Barriers. Hosp Pediatr 2019;9:779-87.
11. Kornbluth L, Kaplan CP, Diamond L, et al. Communication methods between outpatients with limited-English proficiency and ancillary staff: LASI study results. Patient Educ Couns 2022;105:246-9.
12. Patil S, Davies P. Use of Google Translate in medical communication: evaluation of accuracy. BMJ 2014;349:g7392.
13. Taira BR, Kreger V, Orue A, et al. A Pragmatic Assessment of Google Translate for Emergency Department Instructions. J Gen Intern Med 2021;36:3361-5.
14. Khoong EC, Steinbrook E, Brown C, et al. Assessing the Use of Google Translate for Spanish and Chinese Translations of Emergency Department Discharge Instructions. JAMA Intern Med 2019;179:580-2.
15. Chen X, Acosta S, Barry AE. Evaluating the Accuracy of Google Translate for Diabetes Education Material. JMIR Diabetes 2016;1:e3.
16. Cornelison BR, Al-Mohaish S, Sun Y, et al. Accuracy of Google Translate in translating the directions and counseling points for top-selling drugs from English to Arabic, Chinese, and Spanish. Am J Health Syst Pharm 2021;78:2053-8.
17. House J. Translation quality assessment: A model revisited. Gunter Narr Verlag; 1997.

Appendix 1 Source document—undescended testicle

Undescended testicle

An undescended testicle (*cryptorchidism*) is the absence of one or both testicles from the scrotum. In the womb, the testicles form inside the abdomen and then move down (*descend*) through a space between the groin muscles into the scrotum. Sometimes the testicles do not descend or only descend into the inguinal canal but not the scrotum. In most cases, undescended testicles will descend within the first 4 months after birth.

It is important to get treatment for an undescended testicle. Getting treatment will lower the chance of infertility and testicular cancer. Sperm production can begin as early as 12 months of age, so it is recommended that treatment occur between 6 and 12 months of age.

What are the causes?

The exact cause of this condition is not known. Causes may include:
- ❖ Hormone abnormalities.
- ❖ A blockage that prevents the testicles from dropping into the scrotum.
- ❖ Abnormalities in structures or muscles in the testicles.

What increases the risk?

Babies who are born early are most at risk for this condition. Other risk factors may include:
- ❖ Low birth weight.
- ❖ Cerebral palsy.
- ❖ Certain genetic syndromes, such as Down syndrome, Noonan syndrome, or prune-belly syndrome.
- ❖ Being born bottom-first. This is called the breech position.
- ❖ A family history of undescended testicles.
- ❖ Neural tube disorders, such as myelomeningocele.
- ❖ Having a mother who is older, had gestational diabetes, or was exposed to any of the following during pregnancy:
  - ◆ Estrogen hormones.
  - ◆ Cola drinks.
  - ◆ Pesticides.
  - ◆ Medicines for pain.

What are the signs or symptoms?

The main symptom of this condition is not feeling or seeing a testicle where you would expect it to be.

How is this diagnosed?

Undescended testicles are diagnosed with a physical exam. During this exam, a health care provider will check to see whether the testicle is still in the abdomen.

If the testicle is not felt during the physical exam, a procedure called laparoscopy may be done to find out where the testicle is located or whether there is a testicle at all. During this procedure, a small incision is made and a thin, lighted tube is used to look into the abdomen.

How is this treated?

- ❖ In children, this condition is treated with surgery to move the testicle down into the scrotum.
- ❖ In adults, the undescended testicle may be removed, as it is not likely to make sperm.

Follow these instructions at home

❖ Monitor the scrotum for any changes.
❖ Keep all follow-up visits. This is important.

Get help right away if

❖ The scrotum swells, becomes discolored, or appears bruised.
❖ There is severe pain in the scrotal area.

Summary

❖ It is important to get an undescended testicle treated to lower the chance of infertility and testicular cancer.
❖ Babies born prematurely are most at risk for this condition.
❖ If the testicle is not felt during a physical exam, a procedure may be done to find out where the testicle is located or whether there is a testicle at all.
❖ In children, this condition is treated with surgery to move the testicle down into the scrotum. In adults, the undescended testicle may be removed, as it is not likely to make sperm.

This information is not intended to replace advice given to you by your health care provider. Make sure you discuss any questions you have with your health care provider.

Appendix 2 Source document—circumcision

Pediatric urology

Home care after circumcision or penis surgery
What to expect

After surgery, your child will have some soreness, bruising, and swelling in the affected areas.

❖ Any stitches will dissolve on their own. The doctor does not need to take them out.
❖ Your child may have a bandage or clear plastic wrap on the surgery area. This can fall off on its own, shortly after surgery or several days later. Take it off if:
  ◆ Poop gets underneath it.
  ◆ It starts to bunch up anywhere on the penis.
❖ If the doctor used skin glue instead of a bandage, it usually falls off after one or two weeks. It is okay for the skin glue to get wet.

The head (tip) of the penis may have a yellowish or crusty film on it. Put antibiotic ointment (bacitracin) on the surgical cut 2 to 3 times each day. Or do it every time you change your child's diaper, if he wears diapers.

❖ Use the antibiotic ointment (bacitracin) for 2 weeks. Use it on the stitches, too, if your child has them.
❖ Check your child's penis each time you change his diaper (every 3 to 4 hours) or when he goes to the bathroom.

Eating and drinking

❖ Your child may have an upset stomach or throw up after anesthesia. Start off with things like juices, popsicles, water or ice. If your child is hungry give them a small, light meal. For example, they can try eating some toast without butter or some plain rice and applesauce. Make sure they are drinking plenty of liquids.

Activities

❖ Your child should be very careful walking and doing basic activities for 24 hours after the end of their surgery. This is because having general anesthesia can affect their coordination. Older children should not drive, take tests in school or be left alone for 24 hours after the end of surgery.

❖ No sports, bike riding or straddle toys, walkers, swings, jumpers or climbing structures for 2 weeks. Try to keep your child's activity level low during this time.
❖ Your child should use their car seat, booster seat or seat belt normally. You should not adjust or change anything after surgery.

Bathing and water
❖ Short tub baths or showers are fine 48 hours (2 days) after surgery. You can use regular soap and water.
❖ No swimming until 2 weeks after surgery.

Medication
For mild to moderate pain:
Give your child ibuprofen (Advil, Motrin) and acetaminophen (Tylenol). Give one, but not both, every 3 hours. For example, if you give ibuprofen first, give acetaminophen 3 hours later. Then 3 hours later give ibuprofen again. Keep switching back and forth during the day and night as needed.
For severe pain:
The doctor might have prescribed pain medication for your child. Only use it for severe (very bad) pain. Your child may not need this at all. If you think your child has severe pain, but you did not get a prescription for pain medication, call our office right away.
It is OK to:
❖ Give pain medications less often or not at all if you think your child is comfortable without medication. Make sure you are writing down the time and name of medications you give. Use pain control methods like comforting, breast feeding, bottle feeding, ice packs, distraction, and limiting your child's activity.
❖ If you have leftover medications after your child gets better, please talk to your pharmacist about getting rid of them. Do not put them in the sink or flush them down the toilet.

Constipation
❖ Your child may be constipated after surgery, or they may not. If they have small, hard or painful poop, this is from the anesthesia and pain medications they received.
❖ If your child is constipated before surgery, you may want to consider doing the following things.
  ◆ Give them more liquids, unless your doctor tells you not to.
  ◆ Give them more foods that are high in fiber, such as whole wheat bread.
  ◆ Give your child apple, pear or prune juice to drink.
  ◆ Give your child medications such as MiraLAX or suppositories, if approved by your doctor.
  If your child has never taken medications for constipation, talk with your doctor first.

Using MyChart for medical information
MyChart is a way to get your child's medical information online. Please set up a MyChart account for your child as soon as possible, if you do not have one.
  ❖ You need to call our office or go to the front desk during business hours to get an account.
  ❖ It is best to have MyChart set up before your child's surgery.
  ❖ You can use MyChart to send messages to your child's doctor. You can also send us photos after surgery and ask general questions.
  ❖ If you are sending a MyChart message with a photo or an urgent concern, you must call our office as well. This lets us know to check the message right away.

When to call your doctor
Call us at 503-494-4808 (8 a.m.–5 p.m., Monday–Friday) if you have questions, concerns or your child has any of the following:

- ❖ Swelling (a small amount is normal)
- ❖ Bleeding (small dots of blood is normal)
- ❖ Fever of 101 degrees or higher
- ❖ Redness, pain that is getting worse or pus coming from the surgery area
- ❖ Pain even after pain medicine is given
- ❖ Not eating or drinking well, or unusually fussy or tired

You may also call us toll free at 1-888-222-6478 and dial extension 4-4808. After hours and on weekends and holidays, call 503-494-8311 and ask for the pediatric urology doctor on call.