Can language translation applications be safely used in a health care setting?

A Critical Evaluation

Panayiotou, A.¹, Gardner, A.¹, Williams, S.², Zucchi, E.², Mascitti-Meuter, M.³, Goh, A.M.Y.¹,²,³, You, E.⁴, Chong, T.⁴,⁵, Loguidice, D.³, Lin, X.¹, Haralambous, B.¹, Batchelor, F.¹

1. National Ageing Research Institute, Parkville VIC
2. Northern Health, Epping VIC
3. St Vincent’s Health, Fitzroy VIC
4. Academic Unit for Psychiatry of Old Age, University of Melbourne
5. Melbourne Health, Parkville VIC

Abstract

Background: Currently, over 300 languages are spoken in Australian homes. People from non-English speaking (NES) countries without proficient English may not receive equitable care if their health care workers do not speak the same language. Use of professional interpreters is considered the gold standard, but, for a variety of reasons, is often limited to specific aspects of care (e.g. diagnosis and consent). People from NES backgrounds may be disadvantaged when health care workers try to “get by” with minimal means of communication. With the emergence of mobile technologies, health care workers are increasingly using digital translation tools to fill this gap. However, many of these technologies were not developed for or evaluated in health care settings, and may pose a risk if used for certain situations and conversations.

Objective: The aim of this study was to evaluate language translation applications that have the potential to be used in health care settings. The applications will be analysed for their capabilities and potential risks.

Methods: Translation applications were identified by searching the Apple iTunes Store, and published and grey literature. Applications that met inclusion criteria were reviewed in two stages. Stage 1: was conducted by two independent researchers who evaluated technology features (e.g. input and output methods, languages available). Stage 2: was conducted by two independent experts in translation and cross-cultural communication, who evaluated the potential risk when used in a health care setting. High-risk health care conversations involved clinical assessment, provision of diagnoses, conversations about treatment and care planning, discharge planning, and medico-legal information such as seeking consent for medical treatments.

Results: Fifteen applications were evaluated. Eight of the 15 applications contained voice to voice and voice to text translation options. Six applications were restricted to using pre-set health phrases only, while one application used a combination of free input and pre-set phrases. Five applications were excluded prior to Stage 2. Six of the 10 remaining applications reviewed in Stage 2 were specifically designed for health care translation purposes. Of these, two applications were rated as low risk for use in the health care setting - CALD Assist and Talk To Me. Both of these applications contained simple and appropriate pre-set health phrases and did not enable free input.

Conclusions: No translation application was without risk and none should replace professional interpreters. However, some applications may be suitable for low risk everyday conversations, such as those that enable pre-set health phrases to be translated on subject matters that pose little to no risk in the continuum of care. Further research into the use of translation technology for these types of conversations is needed and clinicians should use translation technology cautiously and consider the risks before use.

Keywords: translation, language, technology, applications
Introduction

The widespread prevalence of telemedicine and telehealth has led to increasing acceptance of technology in health care. While there is limited evidence of the effective use of translation technology in medical and health care settings, clinicians anecdotally report the use of internet and mobile applications for language translation purposes. This raises potential concerns, as most applications have not been specifically developed or validated for use in a medical or health care context. However, there is potential for technology to be used to improve communication between patients and staff in the health care setting when used as an adjunct to professional interpreters in low risk situations [1].

Australia is one of the most ethnically and culturally diverse countries in the world [2]. According to the Australian Bureau of Statistics [2] almost half of all Australians were either born overseas, or had at least one parent who was born overseas. In 2015, Australia had the 9th largest population of people born overseas worldwide and a higher proportion of overseas-born people (26%), compared to other countries founded on migration such as New Zealand (23%), Canada (22%) and the United States of America (14%) [2]. Net overseas migration continues to increase in Australia and has shown periods of influx linked to major world events. For example, following the Second World War, Australia saw a high proportion of European migrants. In more recent times, migration has predominantly been from China, India and the Middle East, with Asian countries now making up 8 out of the 10 top migration countries to Australia [3]. In 2016, there were over 300 different languages spoken in Australian homes and more than one-fifth of Australians spoke a language other than English at home [2]. After English, the ten most common languages spoken at home in Australia are Mandarin, Arabic, Cantonese, Vietnamese, Italian, Greek, Hindi, Spanish, Punjabi, and Tagalog [2].

The ability to convey essential care needs (e.g. addressing pain, help with hygiene), communicate simple safety messages and provide orientation cues are essential in health care settings. People from non-English speaking (NES) countries without proficient English may not receive equitable care if their health care workers do not speak their primary language [4, 5]. The use of professional interpreters is considered the gold standard [6]. However, due to issues related to cost, access, availability and time constraints [7, 8], use of professional interpreters in health care is often limited to specific aspects of care, such as comprehensive assessments, procedural consent, diagnosis and the development of treatment plans. Everyday communication between health care workers and clients without proficient English generally occurs without professional interpreters and has been described in the literature as “getting by”, where health care workers rely on gestures, facial expressions, and knowledge of minimal key words in the target language [9, 10]. The “getting by” approach has the potential for miscommunication and misdiagnosis, that may lead to inappropriate or inadequate care provision and patients’ needs being unmet. At worst, the “getting by” approach can result in inappropriate or non-beneficial treatments and care as highlighted by Runci et al 2012 [11], finding a higher frequency of prescription of antipsychotic drugs for Italian speaking residents in mainstream residential care facilities than their counterparts in language specific facilities.

While using an interpreter remains the gold standard for complex medical and legal discussions in all settings, in some situations it is not appropriate or feasible to use an interpreter, yet communication remains an issue. Through the widespread uptake of mobile devices, technology enabling translation has been identified as a potential way to improve communication between patients and staff in health care settings when used as an adjunct to professional interpreters [1]. Very few studies have evaluated the use of translation applications in medical and health care settings and even fewer have compared multiple translation applications or examined the contexts in which their use may be suitable.

While early studies of web-based language tools, such as Google Translate, highlight high levels of user satisfaction [12], the risks relating to accuracy when used in the clinical setting have become more apparent [12-15]. One study evaluated text translation of ten common questions relating to medical history
and assessment from English into ten languages, and found a wide discrepancy in the accuracy depending on the target language [13]. Vietnamese and Polish translation had the lowest accuracy (10% correctly translated), while Spanish had the highest accuracy (80% correctly translated). Another study evaluated the accuracy of ten medical phrases in 26 languages [14]. Of the total translations, approximately 58% were accurately translated from English. However, the accuracy between the different languages also showed variability with African languages scoring the lowest accuracy (42%), followed by Asian languages (46%), then Eastern European languages (62%), and Western European languages (74%). The authors reported the presence of some phrases where the translation resulted in considerable changes to the intended meaning when using complex medical terminology in high risk situations. For example, “Your child is fitting” was incorrectly translated to “Your child is dead” in Swahili and “Your husband has the opportunity to donate his organs” was incorrectly translated to “Your husband can donate his tools” in Polish. As a result, the authors cautioned against the use of Google Translate when obtaining consent for surgery or other medical procedures, and participation in research.

Beh and Canty [15] also reviewed the accuracy of Google Translate in a simulated pre-anaesthetic consultation between an English-speaking anaesthetist and a Mandarin-speaking patient. Twenty-four English phrases and 13 Mandarin phrases were tested, and an independent anaesthetist fluent in both English and Mandarin assessed the translation accuracy. The accuracy of translation from English to Mandarin was 72%, and from Mandarin to English was 67%, improving with short or simple phrases that did not contain technical information, or when speaking clearly and slowly. The authors concluded that Google Translate was not accurate enough to replace professional interpreters, but might be useful if an interpreter was not available.

Albrecht et al. [16] conducted a 6-week trial of a German translation application called “xprompt-multilingual assistance”, designed for use in health care settings. The application allowed for pre-set health phrases to be translated and was used for basic conversations. Nursing staff were then surveyed about their experiences. While most staff (92%) reported that the translation tool was helpful for communicating with patients who spoke another language, that it was easy to use, and that there were no obvious problems with the usability of the device, others reported that the technology was not practical, too time consuming and did not integrate well into existing workflows. They also reported difficulties using the technology with older patients who were unfamiliar with technology or unable to use the application due to visual impairment or illiteracy. The staff also reported the desired target language was not always available. Explaining the menu items in the application caused problems in some instances.

Given the availability and widespread acceptance of language translation applications by the general public and anecdotal evidence of their use in health care settings, more research is required to evaluate their use, particularly in low risk health care situations when professional interpreters are not normally used, and with particular cohorts, such as older people from NES backgrounds. This group makes up 20% of the older Australian population (50 years and over). Of those, 36% report not speaking English well or at all, with English proficiency tending to decline with increasing age and with the onset of dementia [17-19]. Research that includes older people is therefore needed to evaluate the use of translation technology where an increased incidence of auditory, visual, cognitive, and oral health difficulties may further impact on communication, and when unfamiliarity with technology may also have impacts. To date, research evaluating language translation technology in the health care setting has done so in high-risk situations, such as those that involve seeking consent [14], conducting medical assessments [13], or engaging in technical or complex medical conversations [15].

The current evaluation aims to provide an overview of available language translation applications and considerations for use in real-world health care settings. Our study uses experts in the field of health care translation and cross-cultural communication to evaluate the content of translation applications and provide expert opinion regarding the associated clinical risks. It was hypothesised that applications
specifically designed for the health care setting would have a lower level of risk that those designed
business, travel and other purposes.

Method

The study design involves 2 components:
(1) A search for available language translation applications
(2) An evaluation of retrieved applications consisting of 2 stages – feature analysis and risk analysis.

Component 1: Search for available language translation applications
Searching for translation applications was conducted by first searching the Apple iTunes Store (Apple Inc.,
California, USA) on 22nd August 2017 using the search terms in Figure 1. Following this, a grey literature
(Google Search and Google Scholar) and published literature (PubMed) search were conducted for
published articles relating to any smartphone or tablet applications used in health care settings for
translation purposes. Secondly, iMedicalApps, a website that reviews all medical applications, was
searched using the terms in Figure 2. Finally, any applications that the authors were familiar with from
professional experience that were not discovered in the previous searches, were included.

Figure 1. Search Terms for Apple iTunes Store

| technology, language, multilingual translate, translator, translation application, apps, app healthcare, health care, hospital |

Figure 2. Search Terms for Literature, iMedicalApps and Google Search

(Multilingual or Language*)
and translat*
and
app or apps or application* or technolog*
and
healthcare or health care or hospital* or medical or health or clinical

Inclusion criteria
Translation applications were included if they were developed for or used
for language translation purposes, were freely available, were
available on iPad (Apple Inc., California, USA) and enabled translation to or from English. Each
application had to include translation to at least one of the top 10 languages spoken in Australian homes
as at 2016 (Mandarin, Arabic, Cantonese, Vietnamese, Italian, Greek, Hindi, Spanish, Punjabi, and
Tagalog, excluding English) [2]. The applications had to operate on iOS 10.3 that was current for August
2017 on the iPad (iPad Air 2). Applications that were only available for use on Android devices were
excluded following consultation with health care professionals and IT departments from author-affiliated
health care organisations, after which it was determined that iPads were the devices predominately used
in health care settings.

Component 2: Evaluation of retrieved applications

Stage 1: Feature Analysis
Once identified, an objective analysis of all applications that met inclusion criteria was conducted by two
independent clinical researchers (AP and RTJ). The researchers evaluated the applications according to
the following key categories: offline use, input and output method, and languages available. Issues that
may arise for use in the health care setting were also recorded including ease of use based on whether
the application required a high level of user knowledge, or required many steps to navigate through the
application (Table 1). A consensus approach was adopted by the two clinical researchers on all aspects of
each application. Applications were excluded after the feature analysis stage if they required any in-
application purchases or subscriptions as this was considered a barrier to use in the health care setting.

Stage 2: Risk Analysis
The applications included were then evaluated on the basis of risk by two independent professionals (EZ
and MM) with expertise in translation and cross-cultural communication in health care settings. This stage
of the evaluation focused on the risk of using each application in a health care setting, particularly whether
the application had functionality that allowed high-risk conversations. High-risk conversations were defined
as those involving clinical assessment, provision of diagnoses, conversations about treatment and care
planning, discharge planning, and medico-legal information such as seeking consent for medical
treatments. These types of conversations are high risk because they require a high degree of accuracy,
need to confirm understanding from patients, and must allow patients to ask questions. Other risks were
also identified in relation to the type, content and structure of phrases available for selection in the
applications, taking into consideration the complexity and sensitivity of information and the ability to allow
open-ended or two-way conversation with appropriate responses available for selection by the patient.
Recommendations for use in a health care setting were also included in this stage (see Table 2). The risk
evaluation required the individuals to reach consensus about each aspect of each application.
Search for translation technology applications
Apple iTunes Store, searched followed by Published and Grey Literature & iMedicalApps searched

Stage 1: Feature Analysis review of applications (n=15)
Conducted by two independent clinical researcher
Application inclusion criteria:
Language translation purposes
Free
iPad compatibility (iOS 10.3 operations platform)
Able to translate to at least 1 of the top 10 language spoken in Australia

All applications were analysed by two independent researchers.

Excluded applications (n=5)
Required subscription purchase (n=3)
Captured Image translation only (n=1)
Unable to translate into any of the top 10 language spoken in Australia (n=1)

Stage 2: Risk Analysis eligibility (n=10)
Conducted by two independent professionals with expertise in translation and cross-cultural communication in healthcare settings.

Applications were rated on their level of risk when used in a healthcare setting. High-risk conversations were defined as those involving clinical assessment, provision of diagnoses, conversations about treatment and care planning, discharge planning, and medico-legal information.

Low Risk (n = 2)
CALD Assist
TalkToMe

High Risk (n = 8)
Google Translate
Microsoft Translator
MediBabble Translator
Universal Doctor Speaker
Canopy Speak
Say Hi Translate
TripLingo
Naver Papago Translate
Results

Fifteen applications met inclusion criteria and were evaluated.

Stage 1: Feature Analysis (Table 1)

Most applications enabled free voice or text input and this feature usually required an internet connection even once the language package had been downloaded. Applications were capable of voice to voice (V2V) translation (n=8, 53.3%), voice to text (V2T) translation (n=8, 53.3%), text to voice (T2V) translation (n=5, 33.3%) or text to text (T2T) translation (n=5, 33.3%). Seven applications (46.7%) enabled translation of pre-set phrases. Six of the seven applications did not allow for free input and five (33.3%) of these could be used offline, but may require language packages to be downloaded. While Triplingo was capable of multiple input and output functions, it was developed for travel purposes and contained very few pre-set phrases that would be suitable in the health care setting.

Six (40%) of the fifteen applications were related specifically to health care translation. These were CALD Assist, Canopy Speak, Dr. Passport - Personal, MediBabble Translator, Talk To Me, and Universal Doctor Speaker. All of these applications were limited to the use of pre-set phrases and did not allow free voice, text or image input. Of these applications, only two (13.3%) were capable of two-way conversation between a patient and health care worker - CALD Assist and Dr. Passport. As well as containing closed questions that required a simple “Yes” or “No” response, CALD Assist also enabled some open ended questions with limited selections to be made by the patient, and some follow up questions. An example of a follow-up question was “Have you lost weight in the last six months?” then “How much weight have you lost?” with several options available on the screen for the patient to select. Dr. Passport also allowed for two-way conversation between a patient and health care worker. However, this was only possible by enabling patients to select pre-set phrases to translate to their health care worker. This application is divided into pre-set health phrases for the patient and a separate section for health care workers. This application appeared to be intended for patient-led conversations, not conversations led by health care practitioners.

Table 1 – Feature analysis stage.

<table>
<thead>
<tr>
<th>Application name</th>
<th>Developer</th>
<th>Year</th>
<th>Offline use?</th>
<th>Function available (V=voice, T=text, a=application, b=application, c=application, d=application)</th>
<th># of languages</th>
<th>Comments/ issues identified by the researchers</th>
</tr>
</thead>
</table>
| 1. CALD Assist   | CSIRO, Australia | 2016 | ✓ | V a V b T c T d | 10 | ● All languages downloaded with application  
● Easy to use  
● iPad only  
● Developed for health care setting |
| 2. Canopy Speak | Canopy Innovations, Inc., United States of America | 2014 | ✓ | V b T c T d | 15 | ● Can be used offline  
● No free input, pre-set phrases only  
● Difficult to
<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Dr. Passport – Personal</td>
<td>MAIS Co., Ltd., Japan</td>
<td>2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>103</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. iTtranslate</td>
<td>iTtranslate, Austria</td>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Developed for health care settings
- Difficult to use
- Some languages require in-app purchases
- Developed for health care and medical setting
- Led by patient, not doctor

- Must be online for all functions
- Free 7 days, then $5.99/month USD
- Developed for business
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6. iTranslate VOICE</strong></td>
<td>iTranslate, Austria</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>8. Microsoft Translator</strong></td>
<td>Microsoft Corporation, United States of America</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Application Name</td>
<td>Developer/Owner</td>
<td>Year</td>
<td>Features</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------</td>
<td>------------------------------------------------------</td>
<td>------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Naver Papago Translate</td>
<td>NAVER Corp., Korea</td>
<td>2016</td>
<td>- Easy to use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Developed for business and travel purposes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Must be online for all functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Image translator option</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Developed for travel, business and study purposes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Say Hi Translate</td>
<td>SayHi, United States of America</td>
<td>2011</td>
<td>- Must be online for V2V and V2T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- No free text input</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Easy to use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Developed for general conversational purposes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Speak &amp; Translate</td>
<td>Apalon Apps, Belarus</td>
<td>2014</td>
<td>- Must be online for V2V and V2T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- 3 days free, then $9.99/month USD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Easy to use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Developed for travel and business purposes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Talk To Me</td>
<td>Datacom, New Zealand /Australia</td>
<td>2017</td>
<td>- pre-set phrases only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Easy to use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Developed for health care setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>TripLingo</td>
<td>TripLingo, LLC, United States of America</td>
<td>2011</td>
<td>- Must be online for V2V, V2T and image</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Stage 2: Technology Risk Analysis

Of the 15 applications evaluated in Stage 1, 10 continued to Stage 2. iTranslate, iTranslate VOICE and Speak and Translate were excluded as they required monthly subscriptions once the free trial period had ended. Waygo was excluded as it only translated captured images (i.e. text within images). Dr Passport was also excluded as it was only available for free when using it to translate English to Japanese. All other language translations required a fee.

Of the 10 applications evaluated for risk, none of these were entirely risk free (refer to Table 2). All 10 applications enabled conversations about assessment, and all applications, except for one (MediBabble Translator) enabled conversations about treatment / care planning and discharge. Three of the 10 applications did not enable conversations about diagnosis and medico-legal information. The applications that enabled conversations in the least number of high-risk situations were MediBabble Translator, CALD Assist and Talk To Me. This contributed to an overall risk rating of either high or low risk. Two applications - CALD Assist and Talk To Me - were rated as low risk on this basis.

### Table 2 – Technology risk analysis stage

<table>
<thead>
<tr>
<th>Application</th>
<th>Can the application be used for high-risk situations:</th>
<th>Other risks /</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>14. Universal Doctor Speaker</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal Projects and Tools S.L., Spain</td>
<td>✓</td>
<td>![ ]</td>
<td>![ ]</td>
</tr>
<tr>
<td><strong>15. Waygo</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translate Abroad, United States of America</td>
<td>✓</td>
<td>![ ]</td>
<td>![ ]</td>
</tr>
</tbody>
</table>

---

* a – Voice to Voice; free voice input with voice output; b – Voice to Text; free voice input with text output; c – Text to Voice; free text input with voice output; d – Text to Text; free text input with text output.
<table>
<thead>
<tr>
<th>name</th>
<th>Assessment</th>
<th>Diagnosis</th>
<th>Treatment / care plan</th>
<th>Discharge</th>
<th>Medico-legal conversations (inc. consent)</th>
<th>comments</th>
<th>risk rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CALD Assist</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Some phrases and questions are lengthy and/or complex. Only translates pre-set phrases. The phrases relating to assessment, treatment and discharge are considered low risk (e.g. “Do you have pain?”; “I need to do a scan of your bladder”; “You are leaving hospital today”).</td>
<td>Low</td>
</tr>
<tr>
<td>2. Canopy Speak</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Many questions are lengthy, highly detailed, complex, open ended, and cover a broad scope, which poses a high risk. There are also highly sensitive questions (e.g. “Do you have thoughts of killing others?”) and questions containing multiple components (e.g. “When you take that medicine, does it make you feel sleepy, give you a headache or make you feel nauseated?”; “Do you use tobacco now? In</td>
<td>High</td>
</tr>
</tbody>
</table>
the past? For how long? Type and amount daily?"
therefore posing a high risk.

<table>
<thead>
<tr>
<th>3. Google Translate</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>Free input allows for any information input to be translated. Therefore it is considered high risk.</th>
</tr>
</thead>
</table>
| 4. MediBabble Translator | Y | N | N | N | Y | Many questions are lengthy, highly detailed, complex, sensitive and/or cover a broad scope, which poses a high risk (e.g. “Are you allergic to any medication?”, “Do you think about harming yourself?”, “I’d like to know what the pain feels like”; “Do you experience recurrent or persistent thoughts, impulses or images that are inappropriate or upsetting?”; “Are you experiencing prolonged or excessive menstrual bleeding at irregular intervals or more frequently than your normal menstrual
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5. Microsoft Translator</strong></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Free input allows for any information input to be translated. Therefore it is considered high risk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>High</strong></td>
</tr>
<tr>
<td><strong>6. Naver Papago Translate</strong></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Free input allows for any information input to be translated. Therefore it is considered high risk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>High</strong></td>
</tr>
<tr>
<td><strong>7. Say Hi Translate</strong></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Free input allows for any information input to be translated. Therefore it is considered high risk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>High</strong></td>
</tr>
<tr>
<td><strong>8. Talk to Me</strong></td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Only translates pre-set phrases. The phrases relating to assessment, treatment and discharge are considered low risk (e.g. “Are you sad?”; “I will take your blood pressure”).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Low</strong></td>
</tr>
<tr>
<td><strong>9. TripLingo</strong></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Free input allows for any information input to be translated. Therefore it is considered high risk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>High</strong></td>
</tr>
<tr>
<td><strong>10. Universal Doctor Speaker</strong></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Only allows for limited pre-set phrases and questions to be translated. Also includes open ended</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>High</strong></td>
</tr>
</tbody>
</table>
questions, which poses a risk. Medical information about an individual can be saved and this poses a risk to confidentiality (e.g. “Are you allergic to any medication?; “You have the following illness or condition - depression / anxiety / chronic depression / obsessive-compulsive disorder, etc”).

Y = yes, N = no, U = unsure.

Discussion

The current study evaluated language translation applications on the basis of their features and provided expert opinion on the risk of using applications in a real-world health care setting. This is the first study to involve professionals with expertise in health care translation and cross-cultural communication in the evaluation of clinical risk. It was hypothesised that applications specifically designed for the health care setting would have a lower risk rating that those designed for other purposes, however our results show that only two could be considered low risk.

Fifteen language translation applications were identified from searches in the Apple iTunes Store, published and grey literature that met the criteria deemed to be suitable for use in health care settings. These applications were evaluated in two stages to analyse capability and risk level. The feature analysis identified that applications enabled translation of over 100 languages, that they enabled different input and output modes, which determined whether or not they could be used offline, and that the applications had been developed for a range of different purposes, most commonly for health care, travel and business purposes. Applications that had been specifically designed for travel or business purposes tended to have limited applicability to the health care setting. Of these applications, five were excluded from Stage 1 on the basis of requiring payment or not meeting inclusion criteria and ten applications were then evaluated for their suitability in a health care setting in relation to risk. None of the applications evaluated in Stage 2 were completely free from risk. However, the degree of risk varied across applications. For example, applications that allowed free input of information to be translated were inherently high risk as there were no limits to the way the applications could be used, whereas applications that only enabled translation of pre-set phrases had the potential to limit the contexts in which the applications were used and theoretically could be low risk. Despite this, all of the applications enabled conversations about topics considered to be high risk in the health care setting, such as clinical assessment and conversations about treatment and care planning. Translating these conversations requires high degree of accuracy and the ability to confirm understanding from patients and allow patients to ask questions, which is not met by current translation applications. Discussing high risk topics with translation technology could result in miscommunication which might lead to serious negative health outcomes for patients.
Applications with the lowest risk ratings were CALD Assist and Talk To Me. Both enabled conversations in the least number of high-risk topics and limited translation to pre-set phrases. While MediBabble Translator limited topics of discussion to two high-risk situations and restricted input to pre-set phrases, the phrases were deemed to be high risk as they included highly detailed, sensitive and/or personal open ended questions, such as “Do you think about harming yourself or others?”.

A recent study [20] identified mobile medical translation applications where the authors identified key features and scored each application. Applications that were low cost, able to be used offline, did not contain advertisements and in-app purchases, were compatible with multiple platforms (iPhone, iPad, Android and Nexus), were easy to navigate and were well presented scored highly. Canopy Medical Translator, Universal Doctor Speaker, and Vocre Translate were rated favourably. However, these applications were either excluded at Stage 1 of our evaluation on the basis of requiring payment, or were rated as high risk at Stage 2.

Considerations for Clinical Use
The current study evaluated translation technology features and provided expert opinion regarding their risk. These factors are important considerations for use in health care settings and for the development of new translation technologies. In addition to these considerations, other factors such as current policy regarding the use of translation technology, data security and confidentiality should be considered carefully in the design and use of applications for health care. Those that require access to the internet or that keep a record of conversations may require additional precautions. For example, not disclosing the patient’s identity with the application, avoiding collection of personal or sensitive information with the application, avoiding the use of personal devices when using online applications, and using a secure internet connection so that the individual device or location cannot be identified. Other considerations in design and development of applications may include avoiding use of a device with a small screen (e.g. smartphone) which may pose difficulties for patients with visual impairments or reduced motor dexterity.

It is not possible to provide access to professional interpreters for a patient’s entire health care episode and there is an urgent need for effective, accessible and safe tools, such as translation technology, to facilitate communication to improve health outcomes. However, there is a risk that translation technology may become the preferred means of communicating with patients with limited English proficiency due to the perceived simplicity, accessibility and timeliness. The over dependence and over reliance on technology may impact negatively on establishing rapport and providing high quality care. CALD Assist and Talk To Me were the only applications included in the current evaluation that provided users with a disclaimer about their limitations and stressed the importance of using professional interpreters, where possible. CALD Assist and Talk To Me were specifically designed for health care settings and both restricted conversations to pre-set phrases that were considered low risk. They did not include topics and situations that were considered high risk. While Canopy Speak and MediBabble Translate were also specifically designed for health care settings, the independent evaluators found these applications difficult to navigate and contained high risk content that would require a professional interpreter.

Limitations
It was beyond the scope of the current study to examine translation accuracy and cultural suitability. These are important aspects that would impact upon the effectiveness of use in health care settings and involve the suitability of translated words for the context, the syntax of the translated phrases (e.g. order of words and grammar) and the ability to recognise different accents and dialects (when using free voice input). Previous studies have identified poorer accuracy for the translation of non-Western languages in Google Translate [14-16]. However, further research involving experts, health professionals and consumers is required to evaluate the translation accuracy and cultural suitability of other applications and in other contexts.
Given how rapidly technology develops and changes, it was not possible to capture every available application. Therefore, the current study provides a snapshot of the available translation applications and considerations for use in the health care setting. As more applications become available, further research will be required. Additionally, if the use of tablets other than iPads becomes more common in Australian healthcare settings, further research that evaluates android-based apps is warranted.

Conclusions

Overall, the findings of this evaluation have identified that language translation technology is never without risk when used in health care settings and may be completely or partly prohibited by existing health care policies. The degree of risk varies on the basis of the content and features available within each application. Those that allow translation of free voice, text or image information contain the highest degree of risk when used in a health care setting. Of the applications evaluated, only two were considered to be low risk when used in a health care setting, based on their use of pre-set health related phrases. These were CALD Assist and Talk To Me. Pre-set health phrases posed the least risk when the content was appropriate because information was brief, simple, and suitable for the context. While many applications featured pre-set phrases, the content was frequently considered unsuitable because phrases were overly complex, lengthy, contained sensitive information, and did not allow for an appropriate answer. When considering the use of translation technology in health care settings, clinicians are encouraged to consider the risks involved in the interaction based on the translation technology itself, as well as the particular situation, the patient and any organisational policies. Translation technology is not an appropriate substitute for a professional interpreter and further research is required to evaluate its use for every day, low risk conversations. However, it is not logistically and financially possible to have access to a professional interpreter for every interaction in a patient's health care episode. Therefore, there is a pressing need to develop and research tools that facilitate this communication in a safe and effective manner. Translation technology can play an important role, but clearly our research shows the importance of considering risks and that translation technology cannot replace professional interpreters.

Acknowledgements

We thank Mr Ryan Townley-Jones for his assistance in the feature analysis stage of the research. The authors would also like to thank the Melbourne Ageing Research Collaboration.

Conflicts of Interest

Monita Mascitti-Meuter was involved in the initial development and testing of the Talk To Me application. All other authors declare that they have no competing interests.

Abbreviations

NES – Non-English speaking
CALD – Culturally and linguistically diverse
V2V- voice to voice
V2T- voice to text
T2V- text to voice
T2T – text to text

References