A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

1) Protocol Title
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.
Funded by AHRQ. R01 HS024949 2016-2021
Protocol Version Date: 4-3-2019

2) Objectives

This is the second stage (years 2 through 4) of a 5-year project funded by AHRQ. This stage analyzes video recorded data from new recruitment of patients — in addition to videos to be analyzed from the UCD IRB approved clinical trial 522696 “A controlled trial of Patient Centered Telepsychiatry Interventions.” This stage constitutes a clinical trial of the automated online language interpreting tool developed by us in 2016/17.

In this randomized cross-over study with patient recruitment occurring during years 2 through 4, up to 100 Spanish speaking patients will be recruited from several outpatient clinics in Sacramento, and will receive two psychiatric interviews: Method A (conventional in-person interview with a psychiatrist with a live human interpreter) and Method B (asynchronous telepsychiatry — that is, video-recorded interviews that are subsequently processed with automated speech recognition and machine translation technologies). 80 of these patients will have a previous psychiatric diagnosis, and 20 of the 100 patients will have no previously-diagnosed psychiatric disorders, but a chronic medical condition.

In addition to this interview, the patient will complete a battery of questionnaires and a Spanish version of the Structured Clinical Interview for DSM (SCID)i,ii. Both clinical interviews will be digitally video recorded for comparative analysis of language and translation accuracy, as well as for repeat diagnostic and inter-rater reliability assessments by Spanish and English speaking psychiatrists in year 5. Patients and assessing psychiatrists will all be enrolled as subjects.

3) Background

There is a pressing national need to provide higher-quality, more effectively accessible language interpretation services to improve the health outcomes of the 4.7% of the US population who have limited English proficiency (LEP)iii and who currently, as a result, have increased rates of hospital admissions, misdiagnosis, improper treatment and poorer health comprehension and outcomes.iv This project addresses a critical component of this problem: the need to improve access to high quality, mental health services for diverse populations by improving the flow of clinical work across care settings (primary care and specialty care) through the use of online asynchronous methods of communicating. We have already created and demonstrated an efficient, provider compatible, administratively simple health IT solution: Asynchronous Telepsychiatry (ATP). We are now building an automated language translation process into our ATP consultations to allow clinical evaluations to occur across languages without the use of human interpreters.

An 8-minute summary of our past work leading to this study is on YouTube at https://www.youtube.com/watch?v=g8_5hCF2q6I.
In this study, after completing the first phase of tool development and testing, we plan to compare two methods of cross-language psychiatric assessment:

- **Method A** (current gold standard of in-person real-time interpreting practice). A Spanish-speaking patient is diagnostically assessed in-person by an English-speaking psychiatrist through a Spanish-speaking interpreter.
- **Method B** (comparative practice – ATP). A Spanish-speaking patient is interviewed in Spanish by a trained mental health interviewer. The interview is recorded in real-time, translated into English with sub-titles added to the video file, and sent to an English-speaking psychiatrist to asynchronously — that is, at a later time — review the video and make a diagnosis.

All patients will undergo evaluation by both methods. Half of the patients will be randomized to be assessed by Method A first, followed by Method B and half to be assessed by Method B first, followed by Method A. The specific aims of the study are:

- **Aim 1:** To iteratively evaluate and refine the automated asynchronous interpretation tool already developed in prior studies.
- **Aim 2:** To compare patient satisfaction of Method A vs. Method B.
- **Aim 3:** To compare the diagnostic accuracy and psychiatrist inter-rater reliability of Method A vs. Method B and demonstrate psychiatrist inter-rater reliability for Method B.
- **Aim 4:** To compare the interview and language interpretation quality and accuracy of Method A vs. Method B.
- **Aim 5:** To evaluate the diagnostic accuracy of facial recognition software compared with the psychiatrists.

This highly innovative proposal builds logically on our prior feasibility studies and our current randomized clinical trial. The potential impact of this study is far-reaching if we find automated asynchronous language interpretation to be an improved, more easily accessible method of clinical interpretation that is acceptable to, or preferred by, LEP patients than in-person interviews. The automated translation approach we are using has the potential to dramatically expand clinical capacity nationally and worldwide by reducing the need for on-site interpretation services. We will be video recording all patients to collect their language and will use the facial interview data to examine the diagnostic accuracy of facial recognition software. While we are testing and evaluating our approach with Spanish-speaking patients in the field of psychiatry, this technical approach can be generalized towards primary care and other medical disciplines and could be applied to many other languages, as well as beyond the healthcare field.

4) **Inclusion and Exclusion Criteria**

We will only recruit Hispanic individuals with significant LEP — a critically underserved population — comprising 5% of the local Hispanic population at the Sacramento County Adult Psychiatric Support Services (APSS) Clinic and other Northern California mental health and primary care clinics. To help reach a broader Spanish speaking population, we will also provide flyers to additional organizations (e.g. faith-based, social services, and etc.) that predominantly work with Spanish speaking individuals. Finally, we will be placing advertisements in both web based media (online journal) and paper based media (e.g. local
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

We have learned that we are likely to be able to enroll only 2 to 3 such patients per month on average, so we have set up this project with the expectation that it will take us up to 3 years to enroll all study patients. We have study information in English and Spanish, have a bilingual physician (Dr. Odor) as our project co-investigator, a bilingual psychiatrist (Dr. Sciolla), and have already recruited a bilingual project coordinator for our current study to work in the clinics and manage day-to-day study issues. The clinics are well served with bilingual staff, interpreters and documentation.

Participants will be adults with significant LEP such that they prefer being interviewed in Spanish, aged 18 or older and referred by County clinical providers or those who self-refer via the flyers. We will ask referring providers and self-refers to be two types of LEP Spanish-speaking patients: 1) patients who have a non-urgent psychiatric issue — a mood disorder, psychotic disorder, anxiety disorder, or substance or alcohol use disorder(s) — and/or 2) patients who have a chronic medical condition. Patients will be asked whether they are primarily concerned about a behavioral/psychiatric issue or a chronic medical condition issue. Many patients will have comorbid conditions and multiple diagnoses. These diagnoses will be noted as would happen under routine clinical circumstances, and patients with multiple diagnoses will not be excluded.

We will exclude any patients less than 18 years of age, patients with imminent suicidal ideation and/or plans (see Protection of Human Subjects Section), patients who have immediate violent intentions or plans, patients who have significant cognitive deficits and any patient whose primary care provider or psychiatrist do not recommend participating. Patients who meet the exclusion requirements due to safety concerns or for other exclusion criteria will be identified and referred through appropriate channels for emergency or primary care.

5) Study Timelines

In this randomized cross-over study with patient recruitment occurring during years 2 through 4, we plan to enroll 100 Spanish-speaking participants (80 with psychiatric symptoms, 20 with chronic medical disease) to partake in this study starting in October of 2017. After recruitment, participants will undergo a one-day, 4-5-hour process of completing various study surveys, including a Spanish version of the Structured Clinical Interview for DSM (SCID) and the two digitally-recorded clinical interviews. Year 5 will be spent exclusively in data analysis.

6) Study Endpoints

At baseline, all patients will be evaluated in person using the SCID by our trained research physician. This is the “gold standard” diagnostic tool for psychiatric disorders and is widely used in psychiatric research.

Clinical outcome measurements (all presented in Spanish) will focus on disorders likely to be seen in primary care, notably anxiety, depression and substance abuse, and include:

- The SF-12, which is a widely validated and used self-report health survey consisting of 12 questions that produces a functional health, well-being, physical and mental health summary;
- The PHQ-9 is a multipurpose instrument that is widely used for screening, diagnosing, monitoring and measuring the severity of depression (Scores greater than 9 have sensitivity and specificity of 88% for major depression).
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

- The GAD-7 is a widely used screening tool and severity measure for generalized anxiety disorder with high sensitivity and specificity. \(^{xiv}\)
- The Alcohol Use Disorders Identification Test (AUDIT) was developed\(^{xv}\) for the World Health Organization to identify persons whose alcohol consumption has become hazardous or harmful to their health and has been widely used in many studies. The AUDIT takes under 2 minutes to administer and is commonly used in primary care. Based on our previous research we have found a great deal of substance abuse comorbid with other disorders.
- The Clinical Global Impression – Severity scale (CGI-S) \(^{xxvi}\) is a 7-point scale that requires the clinician to rate the severity of the patient's illness at the time of assessment, relative to the clinician's past experience with patients who have the same diagnosis. The analogue scale that is routinely used by UCD psychiatrists as an outcome measure ranges from 1 (Normal) to 7 (Extremely ill, and likely to be an inpatient).
- Global Assessment of Functioning (GAF) used by clinicians to rate patients’ social, occupational and psychological functioning.

The satisfaction and interview comparison measures will be:

- Patient Telepsychiatry Satisfaction Questionnaire. The provider questionnaire was used in our preliminary studies. The patient questionnaire is a modified version of the Parent Telemedicine Satisfaction Survey.\(^{xvi}\)
- Patients will report their perceived quality of the interpretation using a modified version of a UC San Francisco-developed tool measuring satisfaction of patient-centered tools through of quality of communication and visit satisfaction. These questions include the 4-item Quality of Interpretation Scale and the 4-item Patient Engagement Scale.
- Patients will be asked to quantify on a number of visual analogue rating scales of 0-10 the two interview methods in terms of their likeability and effectiveness, using questions developed using the STAR technique\(^{xviii}\) (Situation, Task, Action, Result) which assess the effectiveness of the interview process in the recruiting sector.

The satisfaction, quality of interpretation and interview comparison questions have been integrated into a single document which all patients will complete in 3 components (Attached English and Spanish versions), after each of the interviews separately, and then after the second interview to enable comparison.\(^{xviii, xix, xx, xxi}\)

The questionnaires, self-report measures and interviews will be conducted in person by our trained clinic provider in Spanish. The clinical functioning CGI and Global Assessment Of Functioning GAF will be completed by the interviewing psychiatrist Dr. Chan. We will also assess how clinicians and the care delivery systems work to help LEP patients make the best decisions about their health care. Specifically, we will assess

- provider diagnostic accuracy, inter-rater reliability and other psychiatrist related clinical outcome measures using Kappa as in our previous research\(^{11}\)
- total encounter time, comprising of time for the recorded patient-clinician encounter and time for the consulting psychiatrist to view, evaluate, and write assessments for the patient
- language accuracy and cultural competence: All 200 interviews in both methods will be video-recorded (as below) and the audio transcribed. We will measure in 3 ways:
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

1. Two researchers\textsuperscript{xxii} will measure the medical errors, which we will define as having the potential for clinical consequence,\textsuperscript{xxiii, xxiv} either modifying the history or clinical outcomes. Past research has rated such errors in five groups: “clinically insignificant, mildly clinically significant, moderately clinically significant, highly clinically significant, and potentially life threatening.” Clinically insignificant errors can include a family history of heart attacks at 40 years old being interpreted as 45 years old. Highly clinically significant can include errors of medication dosing.

2. Two research assistants will quantify and compare the error rate of methods A and B. The accuracy will measure particular translation errors following the research of Flores et al in medical interpretation\textsuperscript{xxiii, xxiv}
   i. omission, in which the interpreter (human or machine) did not interpret a word/phrase uttered;
   ii. addition, in which the interpreter added a word/phrase not uttered by the patient;
   iii. substitution, in which the interpreter substituted a word/phrase for a different word/phrase; and
   iv. false fluency, in which the interpreter used an incorrect word/phrase that does not exist in that particular language.

7) Procedures Involved
All research study assessments will take place at the UC Davis outpatient psychiatry clinic at 2300 Stockton Blvd, Sacramento. This is a fully equipped academic outpatient clinic with a reception and waiting area and approximately 20 clinical offices that is designed for both patient care and clinical research studies.

Patient Recruitment:
All patients will be recruited from Northern California primary care clinics or self-referred though flyers placed at various organizations with a Spanish speaking population. We will arrange information sessions for Northern California medical providers to inform them of the study, so that they can discuss it with patients, and will place flyers in waiting areas of relevant clinics so patients can self-refer. We will also ask the local providers, to review clinic lists of primarily Spanish speaking patients who are likely to meet study criteria so that they can then more easily introduce the study to the patients and ask them if they would be prepared to receive a phone call from the research team to discuss the study. We request a HIPAA waiver for this recruitment and screening.

The 100 patients recruited will consist of 80 who have psychiatric disorders and who their primary care providers consider require continuing psychiatric treatment. These will typically be patients with anxiety, depression or substance related problems. Exclusion criteria are any patients who are acutely dangerous to themselves or others, or who have cognitive deficits. The other 20 patients will have chronic medical illnesses, such as diabetes, chronic heart failure or arthritis, and who specifically do not have known psychiatric disorders. This group of patients is being included at the request of the funder to allow us to assess how generalizable this approach is to non-psychiatric medical disciplines.

Patient Screening: Patients from Northern California primary care clinics who are identified as potential participants or those who self-refer via the clinic flyers, will be contacted and asked to verbally consent to undergo an initial phone screening to ensure they meet broad
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

study criteria and understand the 4-5 hour one day commitment of time involved (see attached script).

Phone screening measures include:

(1) The PHQ-9 which is a multipurpose instrument that is widely used for screening, diagnosing, monitoring and measuring the severity of depression (Scores greater than 9 have sensitivity and specificity of 88% for major depression)xxv Patients will be screened out if they respond as being acutely suicidal on question 9.

(2) We are using questions 1,2,3,6 and 12 from the WHODAS 12-item survey which is the 12-item interviewer administered version of the World Health Organization Disability Assessment Schedule II (12-item WHO-DAS II) which asks about difficulties due to health conditions in the past 30 days by assessing six different adult life tasks: 1) Understanding and communication; 2) Self-care; 3) Mobility (getting around); 4) Interpersonal relationships (getting along with others); 5) Work and household roles (life activities); and 6) Community and civic roles (participation). From a screening perspective the questions we are using will ensure that the subjects are physically and psychologically able to undertake the half day research study and the cut off will be a score of 4 on one of them, or a score of 3 on two or more.

Patients who qualify for an enrollment visit will be scheduled for an appointment and at the beginning of this appointment will complete an informed consent form and HIPAA agreement (see attached) prior to assessment. Prior to their scheduled appointment, patients will be sent a reminder phone text message via study assigned cell phone with the day, time, and location of their appointment. Patients will be asked to respond via phone text message to confirm their appointment. The patient’s phone number will be deleted from the phone once the patient has confirmed their appointment.

Assessment and Randomization: Following consent all enrolled participants will be randomly assigned to “Method A first” versus “Method B first” in a 1:1 allocation, using random permuted blocks with a block size determined by the study statistician and stratified by gender (Female vs. Male).All subjects (n = 100) will partake in the Group A and Group B interviews which will take place first to minimize fatigue and possible repetition with other questions later. Both of these interviews will be video-recorded. Immediately following the two clinical interviews, they will be given the clinical questionnaires in Spanish that will assess satisfaction and interview assessment (see measures section below), and will finish with the SCID in Spanish. The SCID-I, as a diagnostic instrument, is designed to provide a primary DSM diagnosis as well as secondary Axis I diagnoses.

Method A (current gold standard of in-person real-time interpreting practice). A Spanish-speaking patient is diagnostically assessed in-person by an English-speaking psychiatrist through a Spanish-speaking interpreter. A report with treatment recommendations following American Psychiatric Association guidelines will be sent, with the consent of the patient, to the PCP or treating psychiatrist who will be able to have ad-lib telephone or e-mail consultations with the psychiatrist (Dr. Chan - Provider A) at UCDHS (consistent with the UCDHS services to community providers). Dr. Chan will have access to all previous clinical information about the patients. This entire interview will be video-recorded for comparison purposes and data analysis.
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

Later and will be the clinical interview that is used as a psychiatric consultation for the PCP’s to use to assist in follow-up treatment of their patients.

Method B (comparative practice – ATP). A Spanish-speaking patient is interviewed in Spanish by a trained mental health interviewer. This interview will be conducted by a bilingual Licensed Professional Counselor (LPC), Nurse Practitioner (NP) or a Licensed Clinical Social Worker (LCSW), Marriage and Family Therapist (MFT) or Licensed Trainee under supervision (provider B), whom we will employ. The interview is video-recorded in real-time, automatically translated into English with sub-titles added to the video file, and sent to Dr. Sciolla (provider C) and to 3 English-speaking psychiatrists (Provider D) to asynchronously review the video and make a diagnosis. This interview will be video-recorded. The interviewers will then fill out a standardized clinical template that will be reviewed at the end of the study for comparison and data analysis. This interview will not be used for clinical purposes during the study. At the conclusion of the interviews patients will be given a release of information to self or a release of information to their pcp to sign, if they wish to receive the research psychiatric recommendations or have them sent to their pcp. All patients will be diagnostically assessed by Dr. Chan (provider A) who is non-Spanish speaking (who will interview for method A). All video-recordings of both methods for all patients will be reviewed by Dr. Sciolla (provider C), who is Spanish speaking, and he will ensure that there are no major language related errors in Dr. Chan’s reports to the PCP’s that are based on method A. If Dr. Sciolla (Provider C) believes language interpretation has caused an error by Dr Chan in method A, the two psychiatrists will discuss the issue and arrive at a consensus recommendation for the PCPs, and we will note this as part of our analysis.

Data Analysis:
Our anticipated sample size (100 patients, 50 per group) in this two-period cross-over study will allow us to achieve our descriptive and analytic aims with appropriate statistical power. We specified that the study should provide us at least 80% power under standard testing conditions (two-sided, with a type-1 error rate of 5%) and performed sample size computations using SAS power and sample size procedure PROC POWER. For Aim 2, assuming a range of within-subject correlation of 0.3-0.6 for the primary outcome (summary measure of patient satisfaction (based on overlapping questions for method A and B questionnaires) measured twice, after each assessment) we would have 80-96% power to detect a 0.33 SD difference in satisfaction levels between Method A and Method B, a small effect size.
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

**Figure 1. Outcome Comparison of 2 Interview Methods Study Design**

8) **Data and/or Specimen Management and Confidentiality**
Federal clinical and research guidelines will be followed regarding screening, advertising, informed consent, confidentiality, and privacy issues. All procedures will be approved by the UC Davis Institutional Review Board (IRB) prior to implementation. We are not collecting any human specimens and are not performing any clinical intervention in this study. Data analysis is discussed above under “Procedures involved”.

9) **Data and/or Specimen Banking**
Our Internet-based system is fully HIPAA compliant and will be utilized to collect and manage multimedia and numeric data in this study. Data will be held in the same HIPAA compliant manner as for our UCD IRB approved trial 522696 “A controlled trial of Patient Centered Telepsychiatry Interventions” submitted 10-17-2013. The data will also reside on HIPAA-compliant Microsoft Azure instance, managed by UC Davis Health IT personnel, and all data
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

and transactions are transmitted using secure sockets layer encryption. Study data will be collected on REDCap database, which allows for direct analysis in common statistical packages such as Python/Jupyter/Anaconda, SAS and R. All videos are streamed securely and cannot be downloaded to other machines. In this study we will collect video and audio data and specifically wish to be able to use it in future studies, and potentially for other studies involving automated facial, movement and audio recognition systems.

10) **Provisions to Monitor the Data to Ensure the Safety of Subjects**

   N/A

11) **Withdrawal of Subjects**

   We will clearly indicate during consent procedures that the investigators have the right to withdraw a participant at any time without their consent. However, we do not expect this to be a significant issue and it is unlikely to occur. Instances that may require withdrawing participants includes failure to continue to meet inclusion criteria, safety, or inappropriate conduct on the part of the participant. Participants who are withdrawn from the study will also have their data withdrawn. Participants who withdraw from the study of their own accord will be asked if information can continue to be collected on them from their routine care records and only used with documented consent.

12) **Risks to Subjects**

   This is a “minimal risk” study. The primary risk involves the potential for subjects to become distressed, or to have psychiatric symptoms increase, when undertaking the interviews. All interviews are being performed by experienced interviewers trained and supervised by a licensed and experienced mental health provider. Interviewers or PCP’s may contact Dr. Yellowlees or Dr. Odor 24 hours a day, 7 days a week by pager, phone or email. Potential subjects may refuse to take part in the project or stop once involved or may require face-to-face treatment with a psychiatrist. Participants who drop-out or are determined to require face-to-face treatment by any of the providers or supervising psychiatrists will be referred to local mental health care resources and back to their referring PCP. Other potential risks include ensuring the participant’s ability to consent, the collection of sensitive mental health information, stigmatization in being involved in a psychiatric research project, and the time required to complete procedures. Overall, the risks are minimal with the use of the proposed study procedures.

   Use of technology: There are no significant risks in the technical procedures as subjects will only be involved in interviews which will be recorded, all transfer of data electronically will be in accord with HIPAA requirements. All study information will be stored on UC Davis IT HIPAA-compliant servers. Individual subject data will be securely encrypted or de-identified per HIPAA requirements, and will never be identifiable in reports or scientific papers generated from the study data.

13) **Potential Benefits to Subjects**

   Patients may benefit by improvement in continuity of care, improvement in psychiatric symptoms, and improved medication management. Providers may benefit from enhanced patient care with additional diagnostic recommendations provided by consultation notes.
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

Subjects will be paid a total of $75 each in the form of a credit card as compensation for their time and will be either given a Sacramento Regional Transit bus pass or reimbursed for any travel expenses incurred to attend the UCD Wong Clinic. Subjects will also be provided with snacks throughout the study session. This is a similar level of compensation that we have used in prior studies and has been judged to be non-coercive.

14) Multi-Site Research  
N/A

15) Community-Based Participatory Research  
N/A

16) Sharing of Results with Subjects  
As described in our prior protocol, all clinical data and medication recommendations from psychiatric staff from the interviews will be shared with patients’ PCP’s or psychiatrists, who may then choose to share it with their patients. Language and interpretation data obtained in this trial will not be shared with patients or PCP’s.

17) Prior Approvals  
We have already obtained signed support from John Onate, MD, Medical Director, Sacramento County Primary Care Clinic (see letter of support) and will commence recruitment of patients from that clinic once IRB approval is given.

18) Provisions to Protect the Privacy Interests of Subjects  
The physical, legal and social risks to subjects in this study are minimal. Researchers will describe the study process thoroughly to participants to ease any privacy concerns. All data and information will be stored on UCD HIPAA-compliant protected servers (REDCap, Microsoft Azure) and any paper data will be stored in locked file cabinets in a locked room at UCD facilities per HIPAA requirements. Only researchers directly involved in patient recruitment, consent or project oversight/audit will have access to PHI. All other data transmitted will be de-identified by using subject numbers (etc. subject001).

19) Compensation for Research-Related Injury  
This is a minimal risk study of psychological outcomes. No study related injury is expected to occur. Though highly unlikely, if injury did occur, the investigators would follow the standard legal procedures of the university and the IRB to settle the claim with the patient.

20) Economic Burden to Subjects  
There is no expected burden to subjects for participating in this research.

21) Drugs or Devices  
N/A
A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.

References


x Maghazil A, Yellowlees P. Novel Approaches to Clinical Care in Mental Health: From Asynchronous Telepsychiatry to Virtual Reality. Mental Health Informatics: Springer; 2014:57-78.


xvii www.recruitloop.com

A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency – Phase Two.


xxi [www.recruitloop.com](http://www.recruitloop.com)


