Field-testing multilingual clinical assessment using DDI standards

David K. Johnson, PhD.
University of Kansas
Epidemiology and the Development of Alzheimer’s Disease (EDAD)
A Collaboration of the University of Kansas with the University of Costa Rica
R21 TW009665
AD in Latin America

Low and middle income nations will experience an unprecedented growth of the elderly population and subsequent increase in age-related neurological disorders

- increased incidence (earlier detection/diagnosis becomes more available)
- introduction of effective life extending technologies that yield increased duration of survival with disease.

Worldwide prevalence and incidence of Alzheimer’s disease will increase as life expectancy increases across the globe
Why Costa Rica?

Nicoya Peninsula = Blue Zone (more older adults living 90+)
Equivalent Life Expectancy (CR=35th vs. US=33rd).

Hispanic Paradox: Rural older adult Costa Ricans live longer than we would expect given what we know (predict) about mortality

Popular culture of participation

Strong university connections for advanced training
University of Costa Rica (UCR) & University of Kansas (KU)

Century long tradition of international academic exchange
Share over 30 academic training programs
Trained over 1000 exchange students.
By June 2011 centralized IT services at KU will support most of the UCR library and administrative services.
Both identified health sciences as the next shared initiative and the KU-UCR endowment fund awarded $25,000 development grants
Identify faculty and resources at UCR to extend the KU AMP’s research success to include Costa Rican clinical samples collected by UCR affiliates.
KU Alzheimer Disease Center (P30 AG035982)

Leader in the US for the study of exercise and metabolism on brain health and lifestyles that prevent Alzheimer’s disease

• Trial for Exercise, Aging & Memory (TEAM; R01 AG034614)
• Alzheimer Disease Exercise Program Trial (ADEPT; R01 AG033673)
• KU Alzheimer Prevention Program (APP)
  • Alzheimer Prevention through Exercise Program (APEX; R01 AG043962)
  • Anti-Amyloid in Asymptomatic Amyloidosis (A4)
  • Aspirin in Reducing Events in the Elderly (ASPREE)
  • Trial Of Resistance Training for Increased AD Susceptibility (TORTIAS; KL2 TR000119)
• Clinical Translational Science Award (FRONTIERS CTSA; UL1 TR000001)
CHARM’s Aim to Build Infrastructure

Extend the clinical research infrastructure of KU AMP to the University of Costa Rica and build its clinical research expertise for neurological diseases.

• **Step 1:** Extend preexisting KU-UCR information technologies to underdeveloped Costa Rican clinical research centers; the IIP and CCP. KU will host data capture, management, and storage systems on the shared KU-UCR network system.

• **Step 2:** Train existing faculty at UCR for clinical assessment and research methods in the assessment and treatment of neurological diseases in the developing world to support future clinical research in Costa Rica and across Central America.

• **Step 3:** Create a baseline for a sustainable longitudinal study of Latin American aging and neurological diseases modeled on the KU Alzheimer and Memory Program.
CHARM Goal 1:

Create a multilingual applied clinical research library (to-date over 400 unique instruments in 850 different applications) that can be shared widely by investigators throughout the US and Latin America to facilitate high quality biobehavioral research on medical issues germane to Hispanic Americans.
BUT...
Low Resource World

Time & Money
Human fallibility
Turn over of trained students
Tracking significant events – Mid-study checks
Big messy library of instruments – A tornado of possible tests
Multiple existing studies with history and inflexible admin systems
Protocol management grew complex by adding Spanish
Perhaps...

1. I could use this opportunity to create tools that could index the tornado of instruments I use over and over again???

To succeed harmonizing data collection across 5 different research programs taking place many sites in 2 languages, we needed to create CHARM:

The Center for Hispanic American Research Methods (CHARM) is a cooperative of US and Latin American research laboratories interested in coordinating biobehavioral research.

CHARM relies on DDI and metadata standards to implement, collected, manage, and archive clinical research data
Clinical Research Data Lifecycle

Phase 1
- Methods
- Hypotheses and Aims
- Research Questions
- U/UCR Clinical Research Methods Training

Phase 2
- Research Project Defined
- User Friendly Front End to collect Paradata
- Battery/Assessment Strategy
- Instruments
- Summary Scores
- Items/ Stimuli

Phase 3
- Research Data Collection
- Focus on Existing Tools
- REDCap
  or
- LimeSurvey
  or
- queXF/OCR
  or
- CAT

Archiving
- Preservation metadata
- Confidentiality
- Additional processing

Processing
- Data specs
- Recodes
- Summary descriptive info

Distribution
- Terms of use
- Citation
- Packaging info

Analysis
- Replication code
- Publications

Repurposing
- Post-hoc harmonization
- Data transformations
KU Instrument Library (KUIL) - Philosophy

KUIL is a distributed research process
If the adoption of DDI is easy and committed to from beginning of many research studies then aggregating data reaches an economies of scale

KUIL is a large database of pre-formed items that are libraryed (i.e., rationally stored using standard conventions) in a large assessment database that can be accessed and implemented in a user friendly interface

Designed to aid grassroot researchers to economically implement a sophisticated study using best-practice assessment.

• Efficient
• Collaborative

Because the heavy burden of adopting DDI at the start up is carried by the KUIL then that cost is distributed across many broad shoulders.

• But Planning the interface becomes critical
• Instrumentation needs vetting

Participation in item/instrument/method validation studies need to shared by all users - part of the user agreement
KU Instrument Library (KUIL) Goals

Library of Ready-Mades
Purposing Instrument Library for Investigators
Translate Instruments (Project BabelFish)
  • Harmonize – Creating concordance tables
Create a actionable plan to Validate Instruments
Dream about Pipeline Processes
Nested Relational Database

Many Items to One Instrument

Many Instruments Belong to Many Batteries
  - Instruments often published in an applied battery
  - However, same instrument can be used over and over in many different investigator-initiated investigations
  - Rarely is the original published applied battery used in its entirety in research that targets sampling specific domains
Phase 1. Content Generation

Step 1. Data Entry in English
Step 2. Content Vetted by 2x RA
Step 3. Test Protocols Generated (Electronic, PDF, & OCR Paper)
Step 4. Content Vetted by Clinical Supervision Team
Step 5. Content Published to CHARM repository

Late Stage Revisions then incorporate user feedback revisions
Colectica/QueXF
Steps to create and use questionnaire

1. Data Dictionary
2. Import into Colectica
3. Export xml code from Colectica
4. Edit xml code, to compile quexml requirements
5. Import xml code into quexml to export a pdf
6. Fill up the questionnaires
7. Import filled questionnaires into quexsf
8. Export data from quexf in .csv format
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>D</th>
<th>J</th>
<th>T</th>
<th>Z</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data dictionary</td>
<td>Field Type (PRED_cj, ORGS)</td>
<td>Field Label (PRED_cj, ORGS)</td>
<td>Gender</td>
<td>USB_EURD</td>
<td>Missing</td>
<td>Implicit</td>
<td>Implicit</td>
<td>Field Type (PRED_cj)</td>
<td>Choices, Calculations, ORGS (PRED_cj)</td>
<td>Choices, implicit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who does a thorough job</td>
<td>survey</td>
<td>SURVEY1833</td>
<td>457.000</td>
<td>BRI</td>
<td>003</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is depressed, blue</td>
<td>survey</td>
<td>SURVEY1834</td>
<td>457.000</td>
<td>BRI</td>
<td>004</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is original and comes up with new ideas</td>
<td>survey</td>
<td>SURVEY1835</td>
<td>457.000</td>
<td>BRI</td>
<td>005</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is reserved</td>
<td>survey</td>
<td>SURVEY1836</td>
<td>457.000</td>
<td>BRI</td>
<td>006</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is helpful and can work with others</td>
<td>survey</td>
<td>SURVEY1837</td>
<td>457.000</td>
<td>BRI</td>
<td>007</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who can be depended on</td>
<td>survey</td>
<td>SURVEY1838</td>
<td>457.000</td>
<td>BRI</td>
<td>008</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is always on time</td>
<td>survey</td>
<td>SURVEY1839</td>
<td>457.000</td>
<td>BRI</td>
<td>009</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is stubborn about many different things</td>
<td>survey</td>
<td>SURVEY1840</td>
<td>457.000</td>
<td>BRI</td>
<td>010</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>Movemos algo que...</td>
<td>keerig</td>
<td>someone who is aloof and doesn’t care</td>
<td>survey</td>
<td>SURVEY1841</td>
<td>457.000</td>
<td>BRI</td>
<td>011</td>
<td>_</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Import into Colectica
Export xml code from Colectica
xml code
queXML to PDF and banding XML for queXF

If the file is valid queXML: a ZIP file containing a PDF and an XML file for banding using queXF will be returned

Choose a queXML file to upload:  

**Style:**

- Response text / sub question font size 10
- Response label font size (normal) 7.5
- Response label font size (small) 6.5

Upload File
**Example pdf**

**Section A:**

**Section B:**

<table>
<thead>
<tr>
<th>B1.</th>
<th>Nombre</th>
</tr>
</thead>
</table>

| B2. | Apellidos |

| B4. | Fecha de nacimiento |

**Section C:**

C1. Estimada señora, en estos momentos estamos realizando una investigación que tiene como objetivo conocer las características del envejecimiento saludable en Costa Rica.

C2. Los hemos seleccionado para que contesten de manera individual, le pedimos completar la información de la manera más sincera posible. Le recordamos que su participación es voluntaria y confidencial, es decir, que los datos recogidos se tratarán con la discreción necesaria para proteger su identidad, y solo se utilizarán para razones de investigación.

**Section D:**

<table>
<thead>
<tr>
<th>D1.</th>
<th>Me veo como alguien que... Es bien inmolador</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| D2. | Me veo como alguien que... Tiende a ser crítico |
|     | Mayormente | Un poco en mayormente | Ni de acuerdo ni en mayormente | Un poco de acuerdo | Muy de acuerdo |

| D3. | Me veo como alguien que... Es inminente en el trabajo |
|     | Mayormente | Un poco en mayormente | Ni de acuerdo ni en mayormente | Un poco de acuerdo | Muy de acuerdo |

| D4. | Me veo como alguien que... Es depresivo, melancólico |
|     | Mayormente | Un poco en mayormente | Ni de acuerdo ni en mayormente | Un poco de acuerdo | Muy de acuerdo |

| D5. | Me veo como alguien que... Es original, se me ocurren ideas nuevas |
|     | Mayormente | Un poco en mayormente | Ni de acuerdo ni en mayormente | Un poco de acuerdo | Muy de acuerdo |

| D6. | Me veo como alguien que... Es reservado |
|     | Mayormente | Un poco en mayormente | Ni de acuerdo ni en mayormente | Un poco de acuerdo | Muy de acuerdo |
Section A:

Section B:

B1. Id (para uso del entrevistador): 0004

B2. Nombre: [Llevar]

B3. Apellidos: [Llevar]

B4. Fecha de nacimiento: 09/01/1973

Section C:

C1. Estimada señorita, en estos momentos estamos realizando una investigación que intenta conocer con detenimiento las características del consumo de medicamentos en Costa Rica.

C2. Le hacemos unas preguntas que contará de manera individual, le pedimos completar la información de la manera más sincera posible. Le agradecemos que se participe ya que el voluntarismo es fundamental, es decir, con lo datos recogidos se trabajará con la discreción necesaria para proteger su identidad, y sólo se utilizarán para razones de la investigación.

Section D:

D1. Me veo como alguien que... Es bien hablado

D2. Me veo como alguien que... Tiende a ser crítica

D3. Me veo como alguien que... Es depresivo, melancólico

D4. Me veo como alguien que... Es original, se me ocurren ideas nuevas

D5. Me veo como alguien que... Es generoso y ayuda a los demás

D6. Me veo como alguien que... Puede a veces ser algo descuidado
Quexf interface

queXF Admin Functions

Form setup
- Test form compatibility with queXF
- Import a new form from a PDF file
- Import/Update banding from XML
- Delete a form (only if no forms yet imported)
- Touch-up a form
- Band a form
- Band a form using interactive banding
- Order variables on the form
- queXS and LimeSurvey integration

Users
- Add operators
- Assign forms to operators

ICR
- Train ICR
- Monitor ICR training process
- Import and Export ICR KB
- Assign ICR KB to questionnaire

Importing
- Import a directory of PDF files
- Successfully imported files
- Failed imported files
- Duplicate forms
- Reverify forms
- List page notes
- Pages missing from scan
- Handle undetected pages

Output
- Output unverified data
- Output data edit

Progress
- Display progress of form verification
- Display performance of verifiers (Completions per hour)

Clients
- Add clients
- Assign clients to forms

System setup
Round Tripping...

Instrument Content in Collectica.
Exported to Google Sheets with handles (names, tags, formats etc.)
Worked on collectively in developing world using common resources
Google Sheets has built-in function for machine translation
A process of quality controls
Authorship / Content signing moves the sheet through the process
End of Process is reimporting data slice back to CHARM repository
Phase 2. Translation Process

Step 1. Machine Translation (Google Sheets Function)
Step 2. Level 1 Editing by Graduate Students
Step 3. Level 2 Content Expert A Review
Step 4. Test Protocols Generated (Electronic, PDF, & OCR Paper)
Step 5. Level 3 Content Expert B Review
Step 6. Content Returns to CHARM repository

Late Stage Revisions then incorporate user feedback revisions
Clinical Research Data Lifecycle

Phase 1
- Methods
- Hypotheses and Aims
- Research Question

Phase 2
- Research Project Defined
- User Friendly Front End to collect Paradata
- Battery/Assessment Strategy
- Instruments
- Summary Scores
- Items/Stimuli

Phase 3
- Research Data Collection
- Focus on Existing Tools
  - REDCap
  - LimeSurvey
  - queXF/OCR
  - CAT

Archiving
- Preservation metadata
- Confidentiality
- Additional processing

Processing
- Data specs
- Recodes
- Summary descriptive info

Distribution
- Terms of use
- Citation
- Packaging info

Analysis
- Replication code
- Publications

Repurposing
- Post-hoc harmonization
- Data transformations

DDI More Common Focus