

## Appendix 3.6 Data extraction process

### Method followed to build the database using the mental health information system stored in Microsoft Access

#### *Patient Identification for Addition to MHIS Database*

1. Designate Access file as up-to-date until Dec 12<sup>th</sup>, 2017 and add to Access Files folder. This will be the Access file that we use for all query editing and data extraction.
2. Edit Hojas Diarias commands.
  - a. Open query qry\_hojas\_diarias\_1 in Edit mode.
    - i. From Consultas table, add the following fields: PHQ9, Nota. From Pacientes table, add the following fields: CesID, Comunidad, Diabetes, Hipertensión, Asma, Embarazada, Depresión, Epilepsia, Desnutrición.
    - ii. Remove Group By variable for each field. May also be able to change Group By to Expression in the Notas field, need to double check. The purpose of this step is to make sure the Notas field keeps Long Text data type.
    - iii. Save query to enable editing of hojas\_diarias\_2.
  - b. Open query qry\_hojas\_diarias\_2 in Edit mode.
    - . Add all fields added to qry\_hojas\_diarias\_1 from above.
    - i. Make sure there are 25 fields are in the following order:
      - 1 CesID
      - 2 ConsID
      - 3 Fecha
      - 4 Apellido
      - 5 Nombre
      - 6 Comunidades
      - 7 Diagnostico
      - 8 PHQ-9
      - 9 Nota
      - 10 Diabetes
      - 11 Hipertensión
      - 12 Asma
      - 13 Embarazada
      - 14 Depresión
      - 15 Epilepsia
      - 16 Desnutrición

17 FN\_Dia  
 18 FN\_Mes  
 19 FN\_Ano  
 20 Edad  
 21 Sexo  
 22 SPSS  
 23 Oportunidades  
 24 Migrante  
 25 Indígena

- ii. Save query
  - iii. Run query with these dates: **sd: 01/12/16, ed: 31/12/17**
  - iv. Export resulting table to Excel, save in AccessHojasDiarias folder with name "qry\_hojas\_diarias\_CommunityX"
3. Open exported spreadsheet. Change name of tab to "all". Create new tab called "dx" with the same headers. Filter by diagnosis and copy and paste the resulting patient data to the "dx" tab.
- Diagnoses to filter by: depresión, episodio depresivo leve, ansiedad, trastorno de ansiedad generalizada, trastorno de estrés postraumático, trastorno obsesivo-compulsivo, trastorno psicossomático, trastorno de adaptación, trastorno bipolar
- TIP: you can filter by two things at once, and can also type trastorno to get most of the above diagnoses
- a. Create new column and use VLOOKUP to compare CES-ID in "dx" tab to CES-ID in corresponding community.
  - b. In MHIS database, use VLOOKUP to see whether CES-ID appears in "dx" tab of pulled hojas diarias data.
  - c. In MHIS, use VLOOKUP to see whether CES-ID appears in "all" tab of hojas diarias data
4. Compare growing list of service users with mental health census for each community.
- a. Open census, create new tab in census and copy CES-IDs
  - b. Use VLOOKUP to compare CES-ID to list of CES-IDs in MHIS Database. Copy and paste any CES-ID from censo that do not appear in MHIS Database into the MHIS Database.
  - c. Use VLOOKUP in the "En el censo?" column to return status of patient (ACTIVO, ALTA, or BAJA) or 0 if patient does not appear in census.
5. Compare growing list of service users with Consultas Recientes from Access
- a. Access Pantalla Inicial ➤ Consultas Recientes ➤ Filter by Depresión ➤ Export to Excel

- b. Add column in exported spreadsheet, VLOOKUP to see if present in MHIS Database, add service users that are included
  - c. VLOOKUP in MHIS Database in Consultas\_recientes column to see if present or not
6. Resolve inconsistencies
- a. Go through individual service users with lead investigator
  - b. Use extracted hojas diarias and filter by CES-ID
  - c. Common problems: no diagnosis in diagnosis box but in clinical notes
7. Exclude service users according to exclusion criteria
- a. Any service users that had no CMD diagnosis in the time frame of the study
  - b. Any service users under the age of 18 by Dec 1<sup>st</sup>, 2016
  - c. Any service users with symptoms of psychosis
  - d. Any service users who did not actually come to clinic even though there is an entry in Access during the time frame of study (this is particularly a problem in Salvador Urbina)
  - e. Include age, gender, community to all excluded service users

### *Patient Data Entry*

1. Demographics
  - a. Name, Date of Birth, Community, Sex, Migrant, Indigenous origin
    - i. Most data manually entered from Access, or entered using VLOOKUP on hojas diarias Access data
  - b. Time to clinic
    - . When in communities, meet with trusted member of community (usually nurse or CHW) to ask about estimated travel times from communities to the clinic
    - i. If not able to visit community, reach out to pasantes to ask trusted member of community to receive estimates of time to clinic from different communities in catchment area.
    - ii. Travel time should be converted into hours and reflect walking time. If the majority of people from certain communities come in car or motorcycle, input time but add note that says travel time reflects travel in car, truck, moto, etc.
2. Diagnosis
  - a. Diagnosis
    - i. Months diagnosis before dec2016: number of months between diagnosis and December 1<sup>st</sup>, 2016. Value should be an integer and should be zero if diagnosis occurred after December 1<sup>st</sup>, 2016.
  - . =IF(CELL<DATEVALUE("1/12/2016"),ROUNDUP(YEARFRAC(CELL,DATE(2016,12,1))\*12,0),0)
3. Diagnosis description
  - a. Sexual abuse/domestic violence/alcoholism
    - Standardized search through all clinical notes using Find tool
    - Domestic violence/sexual abuse search terms: "violencia", "violento/a", "abuso", "golpe", "violacion", "acoso", "sex", "sexual"
  - b. Diagnosis Fidelity

Most data entered manually accessing patient data from Access  
Remaining service users: all clinical notes extracted. Clinical note from consultation of diagnosis analyzed.

#### c. Comorbidities

Formula for comorbidity using hojas diarias. Create 2 new columns

-To say whether patient has one of the chronic comorbidities that CES monitors use this formula “=IF(OR(I2=TRUE,J2=TRUE,K2=TRUE,L2=TRUE,N2=TRUE,O2=TRUE),1,0)”

-To get list of chronic comorbidities use this formula “=TRIM(IF(I2=TRUE,\$I\$1,”) &” “& IF(J2=TRUE,\$J\$1,”) &” “& IF(K2=TRUE,\$K\$1,”) &” “& IF(L2=TRUE,\$L\$1,”) &” “& IF(N2=TRUE,\$N\$1,”) &” “& IF(O2=TRUE,\$O\$1,”))”

#### 4. Community Health Worker

Use following formula

=IFNA(VLOOKUP(I4,'C:\...\LookupFiles\[AcompañantesDbDic2017.xlsx]INDICADORES 2017'!\$B:\$O,14,FALSE),".")

“1” : patient has CHW assigned for depression or other mental health complaint

“0” : patient has CHW assigned for other chronic disease

“.” :patient does not have CHW assigned

#### 5. Monthly data

- Export prescriptions in order to fill in monthly treatment data

##### 1. qry\_recetas2

- a. Add “Fecha” from tbl\_recetas and add Criteria: Between [sd] And [ed]
- b. Input sd: 1/12/16 and ed: 31/12/17 when running query
- c. Run and export to Excel saved as qry\_recetas2\_communityname

#### 6. Alta/Baja

#### *Ensure Database Accuracy and Consistency*

1. Double check Fidelity1 and Fidelity2 diagnosis scores
2. Missing data
3. Clean data