SUPPLEMENTAL INFORMATION 1:

Inclusion and Exclusion Criteria:

Inclusion criteria:

1.   Person diagnosed with tuberculosis (clinical or laboratory criteria) (TB case)

2.   Diagnosed or treated for TB in Ghanzi or Greater Gaborone cities

3.   Persons exposed to TB index case

4.   Consent to participate

Exclusion criteria:

1.   Patient (or patient's guardian if patient is <21 years old) does not provide consent.

2.   Patients aged 8 to <21 years old who do not provide assent.

3.   Patient (or patient's guardian if patient is <21 years old) declines to provide contact information for themselves.

4.   Patient (or patient's guardian if patient is <21 years old) declines to allow study staff to contact them by phone and in person if they miss a study visit.

5.   All prisoners.

SUPPLEMENTAL INFORMATION 2: DATA ELEMENTS

Data Captured for TB Cases:

a. Identifying information (name, Omang/passport number, telephone number, biometric TB strain genotyping with MIRU-VNTR)

b. Locator information (names of three friends/relatives , phone numbers of three friends or relatives)

c. Demographics (sex, age, educational level, occupation,income per month, workplace type)

d. Behavioral (alcohol, smoking, history of incarceration)

e. Location information - GPS Coordinates (workplace, home – primary household, additional homes, school, clinics- sites of inpatient hospitalization, outpatient general care, ART/HIV care, church, bar, prison)

f. Clinical History (TB History [previous diagnosis, previous diagnosis place, previous treatment, previous treatment place, previous treatment month/year])

g. Clinical findings (HIV serology, CD4 count [if applicable], TB treatment type [New, Retreatment, etc.], microbiology [baseline smear result, baseline culture result, follow-up smear results, follow-up culture results], chest X-ray findings [normal, cavity, miliary, abnormal-other])

h. Names of persons they spent time with during putative infectious period (phone [phone numbers stored in cell phone, recent dialed/received phone numbers in cell phone])

i. Household members (names, age, sex)

j. HIV testing is performed as an opt-out test in Botswana and national guidelines indicate the testing of all people with TB cases who are not known to be HIV-infected. Opting out from HIV testing will not preclude their participation in all other aspects of the study.

k. In order to maximize the efficiency of chart review and data collection process, we will perform periodic linkages of our database with the National HIV database. This process will allow us to identify people living with HIV that have not yet been linked to care and/or that have not been started on ART.

Data captured for close contacts of index patients with TB:

a. Identifying information (name, omang, fingerprint)

b. Locator information (phone number, friend/relative name, friend/relative phone)

c. Demographics (age, sex)

d. Relationship with index (parent, aunt/uncle, spouse/partner, cousin, child, friend, brother/sister)

e. Housing type (single home, compound of multiple independent units)

f. Clinical Findings (HIV serology, diagnosed with TB)

Data Captured from the laboratory:

a. Patient information: study identification number, clinic identification number, clinic name, sex, date of birth

b. Type of specimen: spot sputum, morning sputum, lymph node biopsy (specify site and method of biopsy), other biopsy (specify site and method of biopsy), fluid aspirate (specify type of fluid and source)

c. Method of sputum collection (if specimen is sputum): not induced, induced, gastric aspirate

d. Date and time of sputum collection: date (dd/mm/yyyy), time (e.g. 14h00)

e. Sputum transport: name of person transporting specimen, date (dd/mm/yyyy), time of departure (e.g. 14h00) , name of person receiving specimen , date (dd/mm/yyyy), time of receiving specimen (e.g. 14h00)

f. Tests performed at routine TB diagnostic lab: name of operator, date, time, result, date and time result returned to clinic staff

g. Tests performed at research lab: name of microscopist, date, time, result, date and time result returned to clinic staff

h. Culture-dependent tests performed at central lab: name of operator, date of culture start, time of culture start, result of culture, date and time of culture result at central lab, date and time result returned to clinic staff

i. Final outcome of sample: stored/disposed, date of storage/disposal, name of person responsible

SUPPLEMENTAL INFORMATION 3: GENOTYPING APPROACH

We will use two genotyping methods on all *M. tuberculosis* complex isolates: mycobacterial interspersed repetitive unit– variable number tandem repeat typing (MIRU–VNTR). These methods require small amounts of culture, yield digital results, and have a short turn-around time. Crude cell lysates may be prepared from isolates cultured on Lowenstein-Jensen slants, 7H10/11 agar plates, or non-radiometric 7H9-based broths such as BACTEC MGIT 96043. Twenty-four-locus-based MIRU-VNTR typing will be routinely applied using a four-capillary-based ABI 3130 genetic analyzer (Applied Biosystems, USA), calibrated with the MIRU-VNTR Calibration Kit (GenoScreen, Lille, France). PCR amplification of 24 MIRU-VNTR loci will be performed with the MIRU-VNTR Genotyping Kit (GenoScreen, Lille, France) as described in the manufacturer's manual. Briefly, PCRs will be performed using 96-well plates, each one including 10 samples, one positive (H37Rv) and one negative (water) control subjected to eight triplex PCRs. For each triplex PCR, two ml of isolated DNA will be added to 8 ml of ready-to-use PCR reaction mixture. PCR fragment sizing and assignment of the alleles of the 24 loci will then be done.   During the initial phase of the study, specimens will be shipped in batches for processing at the GenoScreen, Lille, France. Progressively, all specimens will be processed in-country. In addition, drug susceptibility testing for first- (and second-line) drugs will be performed.

SUPPLEMENTARY INFORMATION 4: GIS/GPS MAPPING

Specific procedures for the development of such maps will include:

a.   Field Work Preparation: Background Files. The Botswana Geological Survey has very well-developed planimetric files and maps that contain basic parcel information and are fully available to the public through the Office of Census and Statistics. Before beginning the data collection specific for the study, we will provide the study team with these maps to allow them to chart their production, ensuring that there will be no gaps in the data collection coverage.

b.   GPS Data Collection Procedures. Data collection will begin once the data dictionary has been created. Data will be organized in square mile segments that are arranged according to the sections in the Public Land Survey System (PLSS). Clear communication and supervision will be in place to prevent data gaps in the collection coverage areas. Daily field logs will be developed to assure QA/QC. This log will record information such as: (1) The date of field activity; (2) The number of field workers; (3) The names of the files that were logged; (4) The areas that were worked; (5) Base station information; (6) Notes and comments. After data has been collected in the field, it will be transferred to the office and becomes the responsibility of the office personnel. When the data is transferred to the office, it will be protected, checked, and preserved according to the procedures specified in the Data Management section.

c.   GPS Data Completeness. Downloaded data will be checked for completeness. Downloaded data will be checked using the GPS vendor's software. In addition, we will graphically display the positioned features against the scanned images of the source documents or against planimetric map files. This will allow our data managers to verify and document the areas in which data collection has been completed. The software will allow the data managers to browse the attribution of each feature and confirm that all attribution has been addressed.

d.   Data Accuracy. Once the data is checked for completeness, it will be checked for its accuracy. The control "checks" and "repeat points" will be examined. A spreadsheet will be used to compare the data. The comparisons will be analyzed to make sure the data meets 100% accuracy requirements. Any outliers will be flagged and checked. If discrepancies are identified, a field will be made to recollect all or a portion of that data to determine the cause of the inaccuracies.