



# “Ya Tsie” - Botswana Combination Prevention Project (BCPP)

The Botswana Combination Prevention Project (BCPP) was a research project conducted by the Botswana Ministry of Health (MOH), Harvard School of Public Health/Botswana Harvard AIDS Institute Partnership (BHP), and the U.S. Centers for Disease Control and Prevention (CDC). BCPP was a community randomized trial that examined the impact of prevention interventions on HIV incidence in 15 intervention and 15 control community. The interventions included extensive HIV testing, linkage to care, and universal treatment services. To reduce HIV incidence in the intervention communities, the UNAIDS 90-90-90 goals were used: 90% of HIV-positive persons know their status; 90% of persons who know status are to be on ART; 90% of persons on ART are to be virally suppressed.

The BCPP study is composed of 2 interlocking protocols: *Evaluation Protocol* and *Intervention Protocol*.

- The *Evaluation Protocol* of the BCPP evaluated the primary endpoint (HIV incidence), as well as some key related secondary endpoints. This protocol focused on the Baseline Household Survey; the HIV Incidence Cohort; and an End of Study Survey.
- The *Intervention Protocol* of the BCPP implemented the combination prevention (CP) intervention package in CPCs and measures the uptake of these interventions (expanded HIV testing and counselling, strengthened male circumcision, and expanded HIV Care and Treatment).

Additional information regarding the BCPP study can be found here: <https://clinicaltrials.gov/ct2/show/NCT01965470>

The public release data is available separately by protocol. Note that the public release data cannot be combined across protocols to reduce the risk of statistical disclosure concerns regarding risk of participant identification.

# Evaluation Protocol Data Files

The following data files contain all enumeration and survey data collected during the Evaluation Protocol. Plot, household, and member data is separated according to data type and population sample (20% vs. 80%). Survey data from the longitudinal cohort is separated according to study year collected (years 1, 2, and 3), with each file containing data for any survey administered that year. ESS data collected in study year 3 is contained in separate files. A .csv file and data dictionary are available for each type of data. The data dictionaries are a direct reflection of the data and contain a description of each variable, as well as the frequencies of each categorical response or number of non-missing continuous responses. All data file names contain the sample (\_20SAMP, \_80SAMP) and study year in which the data was collected (\_YR1, \_YR2, \_YR3). Refer to the BCPP Evaluation Protocol Data Manual for specific-use details.

## Enumeration Data

Enumeration data is comprised of plots, households, and members. Plot datasets contain 1 row per unique plot selected by the evaluation protocol field team, including those that were not residential/habitable and those not used during the study. Residency status and number of households is available for each unique plot. Household data contains 1 row per unique household selected regardless of enumeration status and includes information on the number of members and final study status as well as the household structure data. Members data consists of 1 row for each unique household member residing in the household regardless of age, presence at the time of the study visit, or eligibility status.

- Plots: 20% sample
  - PLOTSYR1\_20SAMP\_open.csv
  - PLOTSYR1\_20SAMP\_open\_dictionary.doc
- Plots: 80% sample
  - PLOTSYR3\_80SAMP\_open.csv
  - PLOTSYR3\_80SAMP\_open\_dictionary.doc
- Households: 20% sample
  - HOUSEHOLDSYR1\_20SAMP\_open.csv
  - HOUSEHOLDSYR1\_20SAMP\_open\_dictionary.doc
- Households: 80% sample
  - HOUSEHOLDSYR3\_80SAMP\_open.csv
  - HOUSEHOLDSYR3\_open\_dictionary.doc
- Members: 20% sample
  - MEMBERSYR1\_20SAMP\_open.csv
  - MEMBERSYR1\_20SAMP\_open\_dictionary.doc

- Members: 80% sample
  - MEMBERSYR3\_80SAMP\_open.csv
  - MEMBERSYR3\_80SAMP\_open\_dictionary.doc

## **Longitudinal Survey Data**

Each survey data file contains responses to most questions on each survey (BHS, AHS-1, AHS-2) along with endpoint data for those participants included in HIV incidence, coverage, and mortality analyses. Endpoint data for analysis participants is available at each study year in which the individual took a survey.

- Year 1: BHS data only
  - LONGYR1\_20SAMP\_open.csv
  - LONGYR1\_open\_dictionary.doc
- Year 2: AHS-1 and BHS data (from late-enrollees)
  - LONGYR2\_20SAMP\_open.csv
  - LONGYR2\_20SAMP\_open\_dictionaries.doc
- Year 3: AHS-1 and AHS-2 data
  - LONGYR3\_20SAMP\_open.csv
  - LONGYR3\_20SAMP\_open\_dictionary.doc

ESS data collected during study year 3 is separated from data collected as part of the longitudinal cohort. These datasets contain responses to most questions on the ESS as well as endpoint data for coverage analysis participants

- Year 3: ESS data
  - TRUEESS\_YR3\_20SAMP.csv
  - TRUEESS\_YR3\_20SAMP\_dictionary.doc

# Intervention Protocol Data Files

The following data files contain enumeration data from Round 1, home-based and mobile testing through HTC, and clinic data collected during the Intervention Protocol. A csv file and data dictionary are available for each type of data. All data is from the Combination Prevention Community arm (CPC) only. Refer to the BCPP Intervention Protocol Data Manual for specific-use details.

Dataset Name	File Name	Dataset Description
HTC Enumeration Plot	<ul style="list-style-type: none"> <li>▪ bcppr1_cpc_plotlist_open.csv</li> <li>▪ bcppr1_cpc_plotlist_open_dictionary.rtf</li> </ul>	<ul style="list-style-type: none"> <li>▪ Round 1 plots identified in the communities and assessed by HTC. Includes the Evaluation Protocol 20% sample and Intervention Protocol covering the remaining 80% in the community.</li> </ul>
HTC Enumeration HH	<ul style="list-style-type: none"> <li>▪ bcppr1_htc_household_open.csv</li> <li>▪ bcppr1_htc_household_open_dictionary.rtf</li> </ul>	<ul style="list-style-type: none"> <li>▪ Round 1 households identified on residential and habitable plots, whether or not the household was enumerated.</li> <li>▪ Refer to BCPP Intervention HTC HH Enumeration Questionnaire to reference the data collection form.</li> </ul>
HTC Enumeration Member	<ul style="list-style-type: none"> <li>▪ bcppr1_htc_member_open.csv</li> <li>▪ bcppr1_htc_member_open_dictionary.rtf</li> </ul>	<ul style="list-style-type: none"> <li>▪ Round 1 members enumerated in the visited households.</li> <li>▪ Refer to BCPP Intervention HTC HH Enumeration Questionnaire to reference the data collection form.</li> </ul>
BCPP HIV Testing and Counselling (HTC) Intake Forms	<ul style="list-style-type: none"> <li>▪ bcpp_htcintake_open.csv</li> <li>▪ bcpp_htcintake_open_dictionary.rtf</li> </ul>	<ul style="list-style-type: none"> <li>▪ Mobile and home-based testing through HTC in the CPCs</li> <li>▪ All HTC intake forms for those who tested or refused during home or mobile testing.</li> <li>▪ Event-level dataset (i.e. multiple tests per person).</li> <li>▪ Refer to BCPP Intervention HTC Individual Intake Questionnaire to reference the data collection form.</li> </ul>
BCPP Participant Summary in Intervention Arm	<ul style="list-style-type: none"> <li>▪ bcpp_summaryintervention_open.csv</li> <li>▪ bcpp_summaryintervention_open_dictionary.rtf</li> </ul>	<ul style="list-style-type: none"> <li>▪ Participants in the intervention arm of BCPP</li> <li>▪ Contains study outcomes relative to specific-study research questions.</li> <li>▪ This dataset is derived from the business rule requirements document that outlines the algorithms that define each variable.</li> <li>▪ Person-level dataset (i.e. 1 row per person).</li> <li>▪ Only participants with a valid identifier (i.e. omang or passport) are included to allow for deduplication across all sources to represent the individual, not event.</li> </ul>