**Data Management Hand Over Document**

**Project: The Sixth South African National HIV Prevalence, Incidence and Behaviour Survey (SABSSM VI)**

**Date: 28 June 2024**

**Prepared by: Mr Phaleng Maribe**

1. **Introduction**

This document serves as a comprehensive guide for the handover of data management responsibilities for the Sixth South African National HIV Prevalence, Incidence and Behaviour Survey (SABSSM VI). It includes an overview of the data, methodologies, key responsibilities, and essential contact information to ensure a smooth transition.

1. **Meta data**

* **Title**: SABSSM VI 2022
* **Description**: A comprehensive dataset capturing SABSSM VI.
* **Creator**: Phaleng Maribe/Data Management team]
* **Date of Collection**: [January 2022] - [April 2023]
* **Final laboratory results**: October 2023
* **Geographical Coverage**: South Africa
* **Target Population**: Households in South Africa
* **Unit of Analysis**: Individual, District level, Provincial level, National level
* **Summary of response rate** 
  + 29 447 VPs were approached, where 27 005 (91.7%) were valid.
  + Of 27 005 valid households, 80.0% completed a household interview.
  + Of the eligible 76 134 individuals, aged 0+ years and older, 94.1% were interviewed, and 62.7% provided blood for HIV and additional testing.
  + Of the 30 718 eligible women, aged 15 years and older, 94.9% were interviewed, and 69.0% provided blood for HIV and additional testing.
  + Of the 22 665 eligible men, aged 15 years and older, 90.6% were interviewed, and 60.8% provided blood for HIV and additional testing.

1. **Overview of the Data**

* **Data Type**: Household survey data on HIV prevalence, Incidence, and related behavioral determinants.
* **Data Files**:
* Raw data: STATA format : Link ([SABSSM VI raw datasets](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/EtWLGmyHaZxPgFiOaXvq6IMBr2mZ7k9VmILe6kjhqIZe1Q?e=WfUZtZ))

*Household dataset, Individual dataset, Adults (15+), Guardian (12-14), children (0-11)*

* Cleaned data: STATA format: Link ([SABSSM VI Final](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/Erg6r62_7t5JiPaoNN2YX5cB_3PkG9lvw-snLqM_zosXtA?e=C4fUcj))

*Household dataset, Individual dataset, Adults (15+), Guardian (12-14), children (0-11)*

* Documentation: Link ([Doc\_reports](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/En0T65rPiiZEmegBxf6ps_4BzrOjKxiZ70C1LreNp66xbA?e=gyE01C))

*Metadata, data dictionaries, codebooks*

* Tables and figures: Link ([SABSSM IV tables and figures](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/Es0eroXs62FNqm3plZtitZEBb5HQkIoh_zvDJp_RdgKUZw?e=V8fGbY))

*SABSSM VI combined tables and figures, SABSSM VI Executive summary figures, SABSSM VI core tables and figures, SABSSM VI provincial tables and figures*

1. **Data Collection Methodology**

* **Survey Instrument**: Questionnaire data were captured using the REDCap (version 10.6.3) mobile application on electronic tablets. <https://redcap.hsrc.ac.za/redcap/index.php>
  + *Household Questionnaire* ***(for 9 provinces)***
  + *0 to 11 years old individual questionnaire* ***(for 9 provinces)***
  + *12 to 14 years old individual questionnaire* ***(for 9 provinces)***
  + *15 years and older individual questionnaire* ***(for 9 provinces)***
* **Data Collection Period**: Survey data collection took 16 months to complete from **January 2022 to April 2023**.
* **Data Collection Team**: Data Management Specialist, Data Analysts, Data Monitors, IT technicians, Developers.

1. **Data Management Process**

* **Data Entry**: The data were captured in an electronic data capturing template and records were uploaded to a database
* **Data Cleaning**: See QC and deduplication SOPs ([Data Quality Control SOP.docx](https://hsrcacza-my.sharepoint.com/:w:/g/personal/apmaribe_hsrc_ac_za/EbvoeZpG2utOjnHhI6XHYW4BLGv5V2UaXZ63rBUo9IgvEA?e=cTfwmO)) **and** [SABSSM VI Deduplication SOP.docx](https://hsrcacza-my.sharepoint.com/:w:/g/personal/apmaribe_hsrc_ac_za/EQKAWRA0th5Lt5wkPoC10V4BEnL6TXawYgq25MP6nZV4wQ?e=Kl8hyW)
* **Data Analysis**:
* **Software Used**: Tools used for data analysis Excel (for graphs), Stata 18.0, R).
* **Analysis Procedures**: Data were analysed using STATA 18.0 and the figures were prepared in Microsoft Excel. Each set of results was independently run and validated by a pair of statisticians/analysts. Analysis for estimating HIV incidence and the number of new infections was undertaken in R and in STATA.

1. **Documentation**

* **Metadata**: Link: [SABSSM VI Meta data.docx](https://hsrcacza-my.sharepoint.com/:w:/g/personal/apmaribe_hsrc_ac_za/EXkCVcHZ-N1GguvlAU_fmYABJ-6eat9yXXauNKxvuCom4w?e=35cahF)
* **Data Dictionary**: Link: [Data dictionary and questionnaires](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/Eg3SzCiPnBRKuFmsbxCM_C4BalNWC4bbR30qzdT6AqQRKg?e=NHtWpr)
* **Codebooks**: [Data dictionary and questionnaires](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/Eg3SzCiPnBRKuFmsbxCM_C4BalNWC4bbR30qzdT6AqQRKg?e=NHtWpr)
* **Data Flow:** Link: [Data Flow.docx](https://hsrcacza-my.sharepoint.com/:w:/g/personal/apmaribe_hsrc_ac_za/EVG4YUOwUudEsJJxuxof28cBnnyqopwb9wUkpsXW7zRGug?e=SjKRLS)

1. **Roles and Responsibilities**

* **Project Manager**: Dr Lehlogonolo Makola
* **Data Manager**: Mr Phaleng Maribe
* **Data Analysts**: Ms Goitseone Maseko, Dr Inbarani Naidoo and Ms Ronel Sewpaul
* **Other data team members**: Lesiba Molopa, Thembelihle Ginyana, Kutloano Skhosana, Sbonelo Chamane, Prudence Chambale

1. **Next Steps**

* **Pending Tasks**: Testing of samples were the ART was not found and the viral load is suppressed, finalisation of main report and provincial reports upon receiving comments from reviewers.
* **Future Updates**: XXXXX
* **Transition Plan**: Goitseone Maseko will take over data management, assisted by Ronel Sewpaul, Lesiba Molopa, Sbonelo Chamane.
  + Sean Jooste will lead data analysis team assisted by Goitseone Maseko, Inbarani Naidoo and Ronel Sewpaul

1. **Contact Information**

* **Primary Contact**: Goitseone Maseko, email: [gmaseko@hsrc.ac.za](mailto:gmaseko@hsrc.ac.za)
* **Secondary Contact**: Sean Jooste, email: [sjooste@hsrc.ac.za](mailto:sjooste@hsrc.ac.za)
* **Technical Support**: [Contact information for IT or software support]

1. **Signatures**

* **Outgoing Data Manager**:
  + Name: Phaleng Maribe
  + Date: 28 June 2024
* **Incoming Data Manager**:
  + Name: Goitseone Maseko

This document should be reviewed and updated regularly to ensure all information remains current and accurate. Please refer to the attached files and documentation for further details on specific aspects of the data management process.

For any questions or further assistance, please contact Ms Goitseone Maseko email: [gmaseko@hsrc.ac.za](mailto:gmaseko@hsrc.ac.za), Dr Sean Jooste , email: [sjooste@hsrc.ac.za](mailto:sjooste@hsrc.ac.za) or Prof Edmore Marinda email: [emarinda@hsrc.ac.za](mailto:emarinda@hsrc.ac.za)

**Attachments:**

1. Overall link to all the files: [Final Report](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/Em_M-WQgRXdHmMuOXevnLkQBepHgr12x8A6ESrMIWOykPg?e=pTWr7j)
2. Raw Data Files [SABSSM VI raw datasets](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/EtWLGmyHaZxPgFiOaXvq6IMBr2mZ7k9VmILe6kjhqIZe1Q?e=H3B4Xk)
3. Cleaned Data Files : [Data Cleaning do files](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/EhUIsAs3IWVBj3sVFdc1kqABi1alDWGqByhA9K1uofaM-A?e=L2wBDa)
4. Data Dictionary : [Individuals Data Dictionary.xlsx](https://hsrcacza-my.sharepoint.com/:x:/g/personal/apmaribe_hsrc_ac_za/EYV0Tlr3GhBAsmNCYs21FYkBFeJn5imMj8DoW-QtrG2jUA?e=dndPyK)
5. Codebooks : [Codebooks](https://hsrcacza-my.sharepoint.com/:f:/g/personal/apmaribe_hsrc_ac_za/Eorl2IZSP85FkmzsiR_ikyMBw0DAydJ7UMdNcbVvAdjrvg?e=B9fXpn)

**End of Document**