

# Training course on Household Surveys

A Presentation to NSO and other  
Participants

Suriname

October 2017

# Overview

- Introduction
- General survey organisation
- Purpose, Scope and objective of household surveys
- Survey Organisation
- Questionnaire Design (PAPI vs CAPI)
- Data transmission, data security/annonymisation and storage
- Quality Assurance in household surveys

# Household survey organisation

- Is a large scale and a complex process involving many stakeholders
- Influenced by decisions made both within NSO and outside the NSO
- Requires substantial amount of resources to successfully execute it

# Household survey .... cont'd

- Cover a wide range of topics (modules) with varied users and uses.
- Many users (both within NSO and those outside the NSO who may not have an idea how long it takes to complete the survey and share results

# Household survey .... cont'd

- It draws resources (human, hardware etc) from many departments within an organization/NSO
- Requires a dedicated team with different expertise (sample design, IT skills, field organization skills, analysis, data management and dissemination skills and archiving
- It is a process which starts from planning to data archiving

## **Household survey ... cont'd**

- The demand for data and information from household surveys has been on the increase since the era of structural adjustment programme
- Prior to SAP era, macro economic aggregates were largely considered adequate to inform planning processes in many developing countries

## Household survey ... cont'd

- With the just concluded MDGs and now the SDGs
  - Means more pressure on NSOs to provide more timely, more disaggregated, more faster and more detailed information
  - Continuous monitoring to increase to be able to update the statistics
  - Increase collaboration and coordination among statistics producers to deliver the product

## Household survey ... cont'd

- Even at country level,
  - There are many data requests/demands requiring household based data
  - Other producers who play in the space of statistics production (do other agencies conduct household surveys in Suriname?)



# Different Survey approaches

- Qualitative and Quantitative:
  - collected either through direct observation or through semi structured interviews
  - Increasingly being used to capture what cannot be quantified using quantitative surveys
- Different researchers opt for different approaches depending on what they are investigating
- However, the focus here is on quantitative approach using a structured set of interview questions

# Different survey approaches cont'd

- The quality of data and information produced varies greatly due to differences in methodology, coverage and scope
- Many (Actors) are not usually transparent about the methods and design issues of many of the surveys they undertake
- Duplication is rampant and coordination is limited.

# The different kinds of field based research cont'd

- Quantitative research conducted by interviewing selected respondents. The selection of the respondents may be either purposive or objective

Purposive surveys are sometimes used in market research

Objective surveys are based on scientific methods

# Survey Organisation

- Critically look at each of the following stages :
- **Planning:**
  - formulation of the Statement of Objectives;
  - selection of a survey frame;
  - -determination of the sample design;
  - questionnaire design;
- **Data collection, processing and analysis;**
  - data capture and coding;
  - editing and imputation;
  - estimation;
  - data analysis;
- **Data dissemination, archiving and documentation**

# Survey Organisation ..cont'd

- Decide whether a centralized or a decentralized approach to data collection is to be adopted
- Centralised approach:
  - Managed and coordinated from headquarters
  - All activities of data collection are monitored and validated centrally
  - Receives instructions and updates from one source
- Decentralized approach
  - regional/provincial centres managing some aspects of the survey
  - sometimes, all activities decentralized (recruitment, data entry and processing, transport etc)

# Why conduct household Survey ?

- Purpose/objective:
  - To inform the planning frameworks of the country
  - To inform the international frameworks to which Govt has committed to
  - To build the capacity of NSOs and key partners to undertake large scale household surveys capable of meeting most of the user demands
  - Overtime build a core team of survey statisticians, analysts, programmers, and message developers capable of addressing user requirements in real time

# Planning/Design of Surveys

- Identify the users/intended beneficiaries of the information to be gathered
- Determine the availability of resources
- Mapping of stakeholders
- Agree on the objective(s) of the field research
- Establish what policy question are you trying to answer?
- Capacity issues
- Is there secondary information in existence to facilitate sample design

# Survey Design

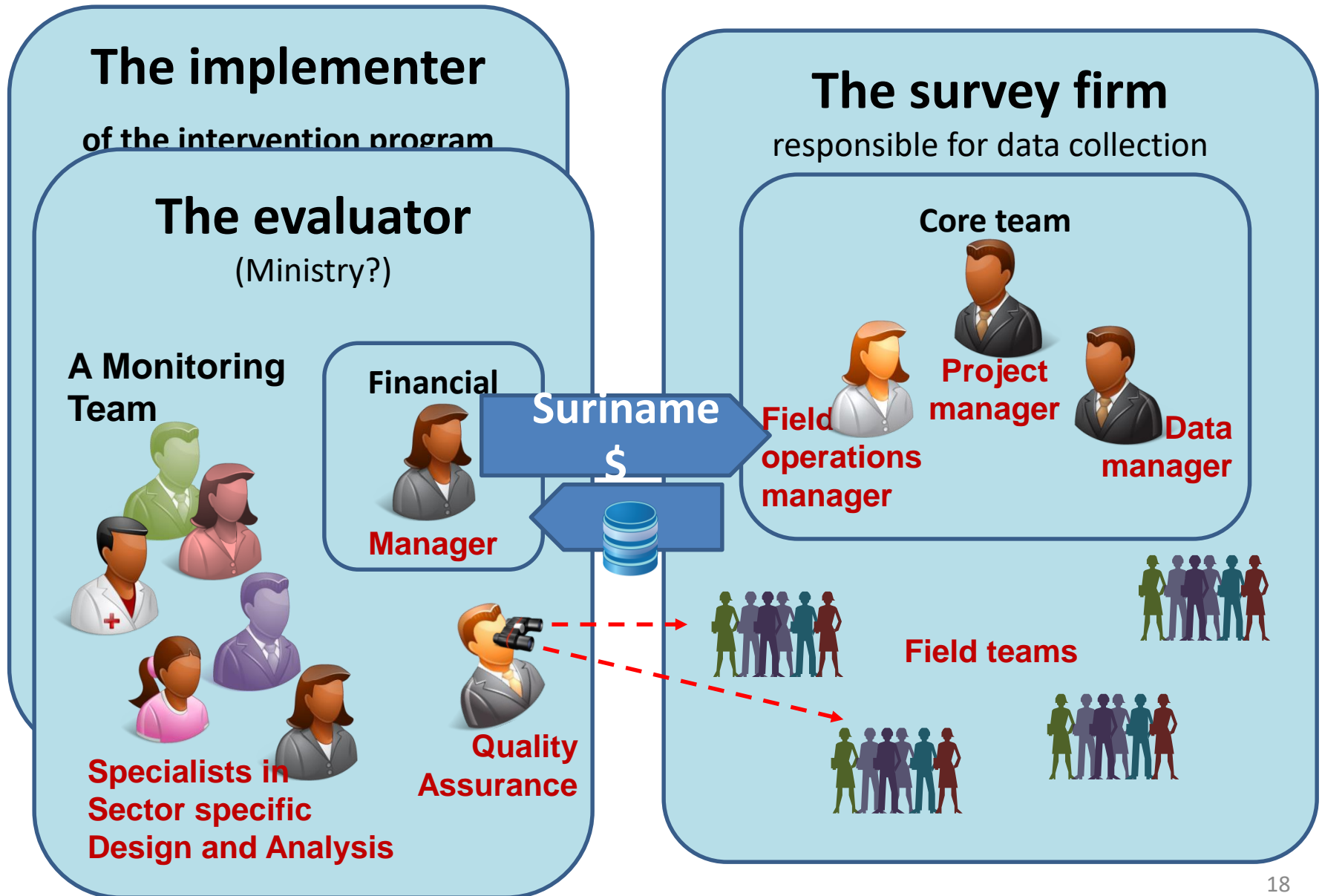
- Universe from which to base the field study
  - How comprehensive is it? Does it cover only part of the population? etc
- Unit of observation
  - Is it a household , an individual, household business, etc
- Coverage
  - Is it covering the entire country or targeted to one district/region



# Survey Design cont'd

- Sample design
  - Simple Random, Multi stage (Stratified or cluster)
- Sample Size
  - Number of observations and levels of disaggregation

# The stage and the actors



# Survey Plan

- Which team will visit each sampling point and when
- The survey plan should consider:
  - Time spent by a team at a sampling point (work plan)
  - Number of sample points
  - Total time available for fieldwork
  - Spatial and temporal distribution of the sample
  - Number of field teams
  - External constraints
    - Weather
    - Security
    - The calendar of the evaluated project

# The fieldwork plan at a sampling point

What is the team supposed to do in each sampling point, and how long does it take

- Things to consider:
  - Number of interviews per sampling point
  - Geographical scattering of households within each sampling point
  - Conformation of teams
  - Household listing operation, if applicable
  - Quality controls
  - Biometric and cognitive test, measurements, etc., if applicable
  - Interview duration, number of visits and correction visits
  - Other questionnaires, for example, at the community level
  - Rest time
  - Approximate travel time to the next sampling point

x

\$



Optimistic scenario

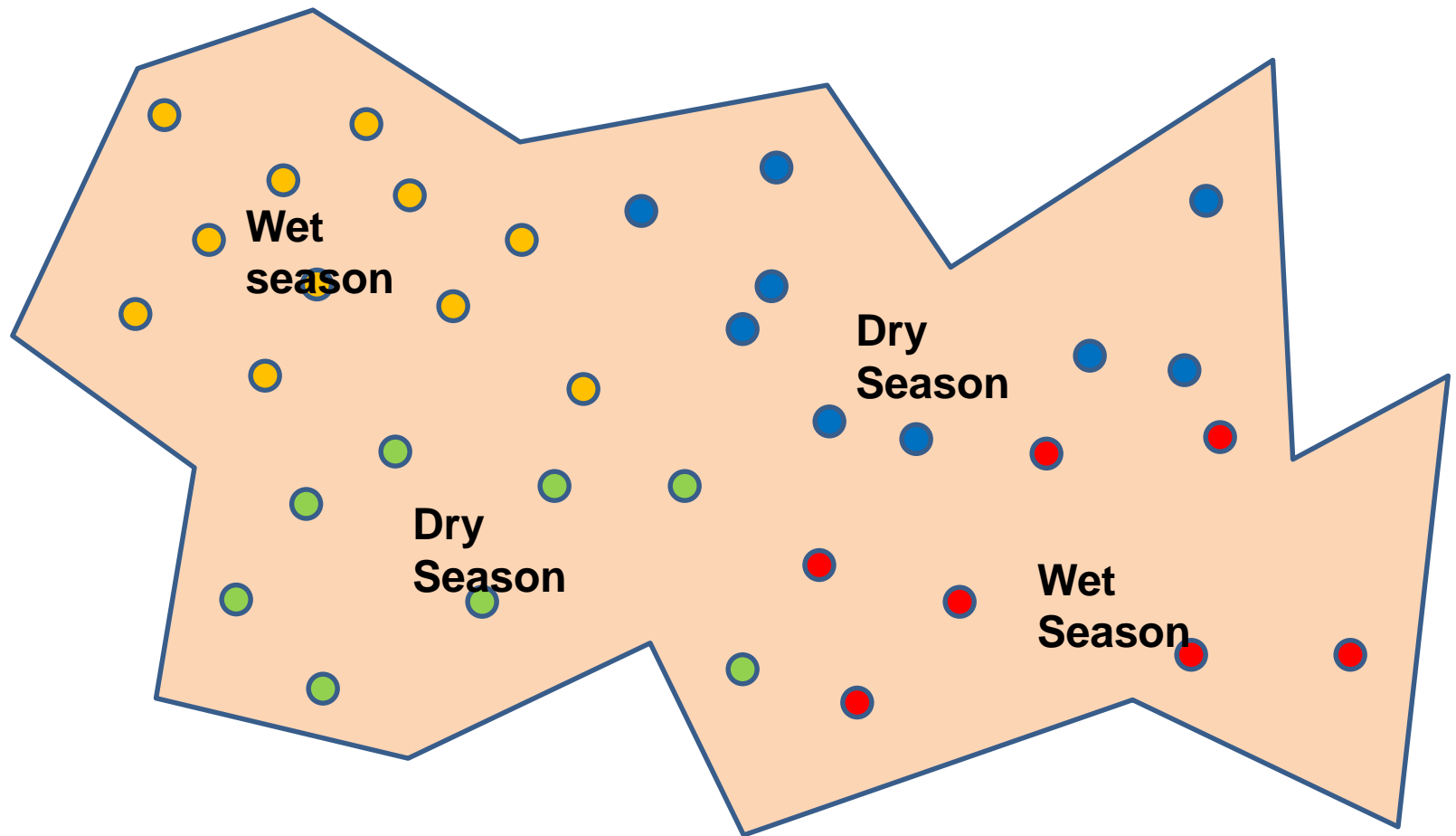
Pessimistic scenario



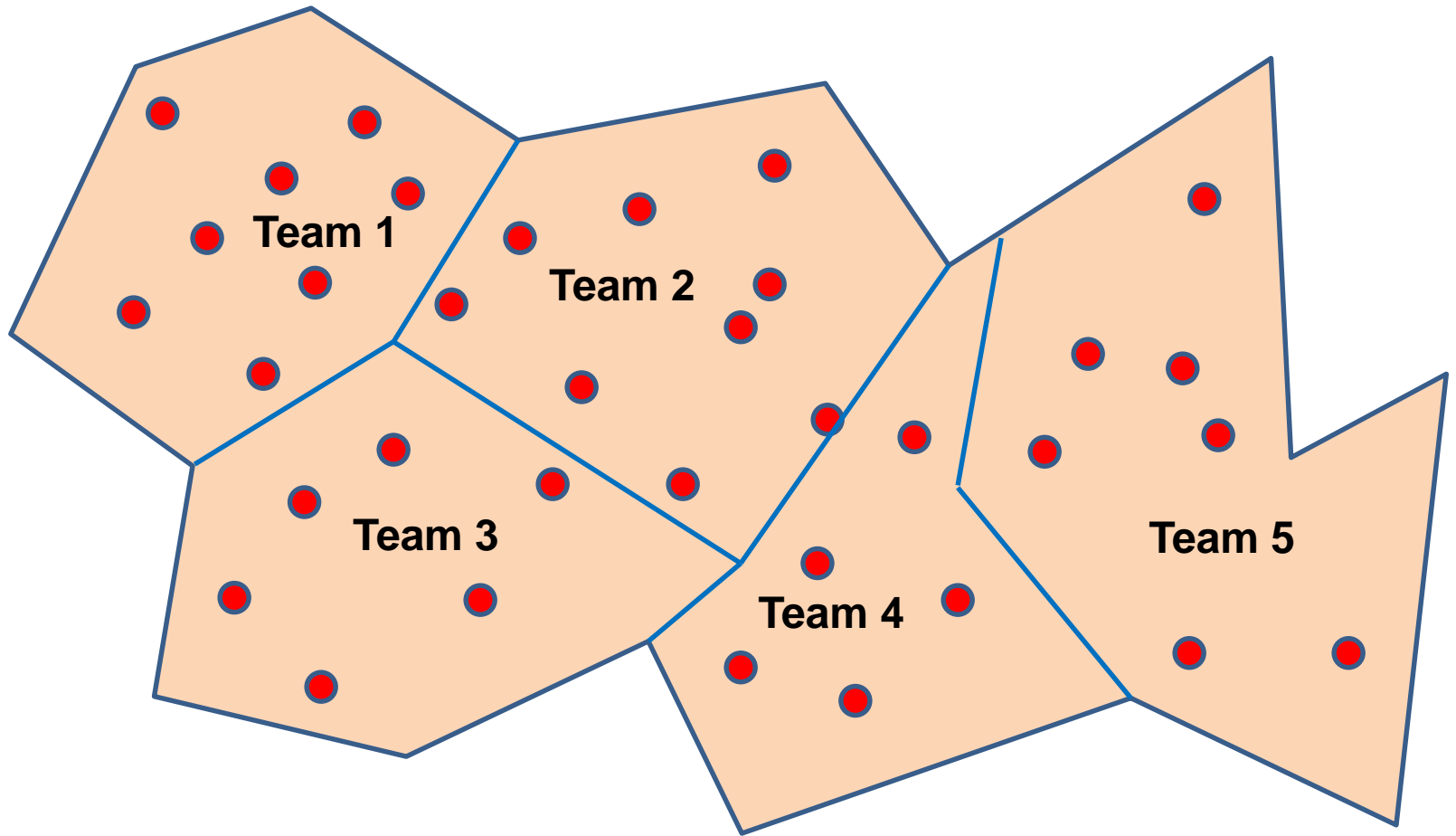
✓

\$

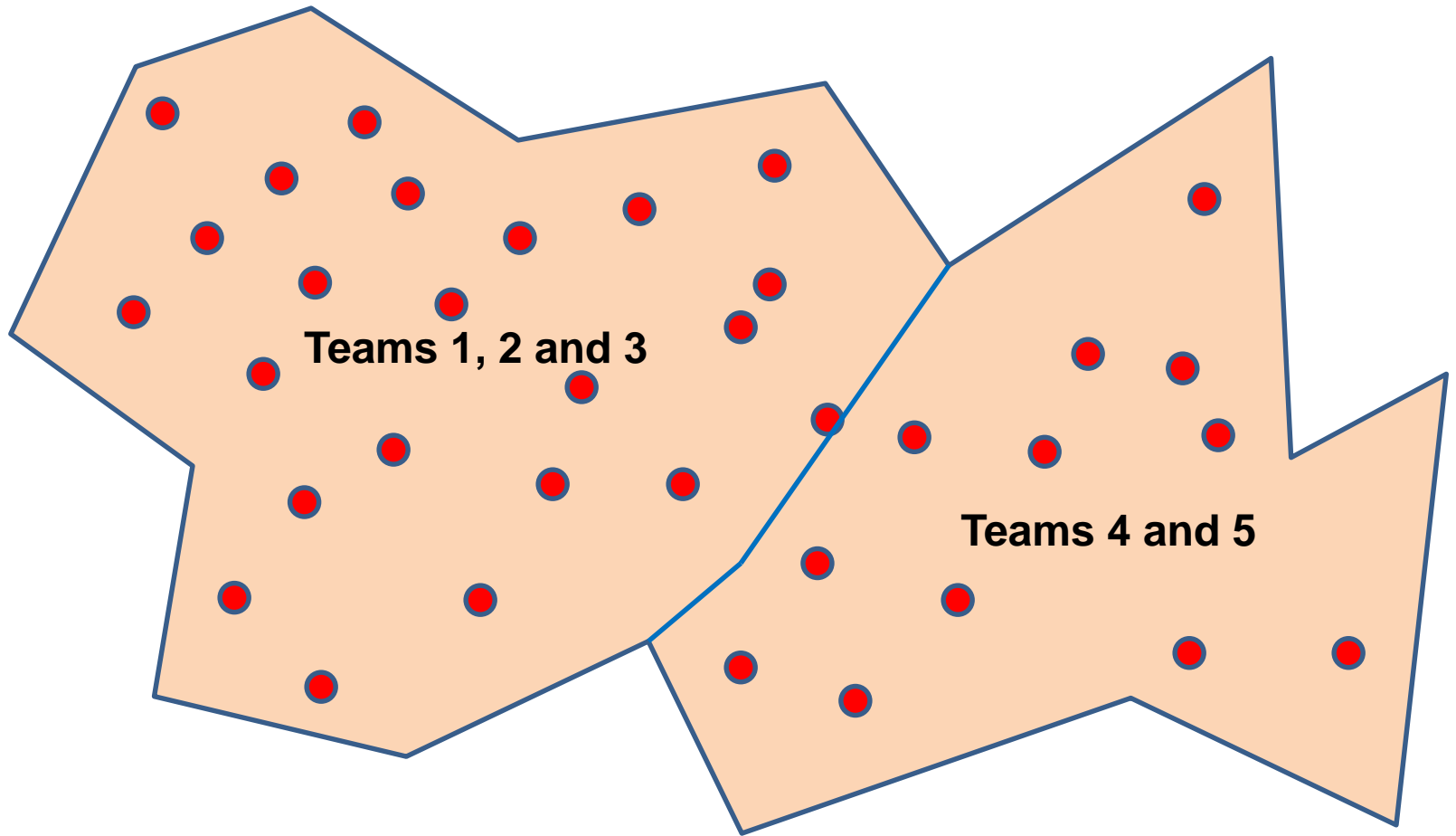
# A bad survey plan



# A better survey plan



# Even better (if you can)



Inter-penetrating samples (Mahalanobis, 1953)

# The key to good survey design and implementation

## Sampling

- Only random sampling permits the definition of standard errors, and estimations with confidence intervals
- Samples are often not SRS but complex (with two or more stages, and stratified). Complex designs have more error (Design Effect)
- Paneling can reduce the error of changes over time, but is vulnerable to ageing, attrition and manipulation
- In impact evaluation, we need to estimate the difference between treatment and control groups. Power is the probability of not misjudging an effective program

## Questionnaire design

- A process with many actors and many phases, often iterative
- It must be reality-checked

## Field work / Data management

- Organized on the basis of self-sufficient teams
- Training is essential and takes time (weeks, not days)
- Quality Assurance supported on three pillars:
  - Intelligent data entry in the field
  - Central monitoring of field workers
  - Human supervision through check-up visits

## High management

- The survey firm ToRs must specify the quality of the deliverables
- The Core Staff team must be devoted to the survey full time
- Realistic plan and budget

## Documentation



# Quality Assurance

# Quality Assurance in household surveys

- Some critical areas to keep watching closely
  - Design (both the sample and questionnaire)
  - Data collection and editing
  - Data management and storage

# Quality Assurance in household surveys cont'd

- Important to consider which institution is conducting the survey
- Issues to do with recruitment of field staff, sampling translation of questionnaires etc
- Overall survey implementation (monitoring-recall visits, re-interviews, field visits (announced and un announced etc)
- Data analysis (editing checks, treatment of missing data, data archiving and management

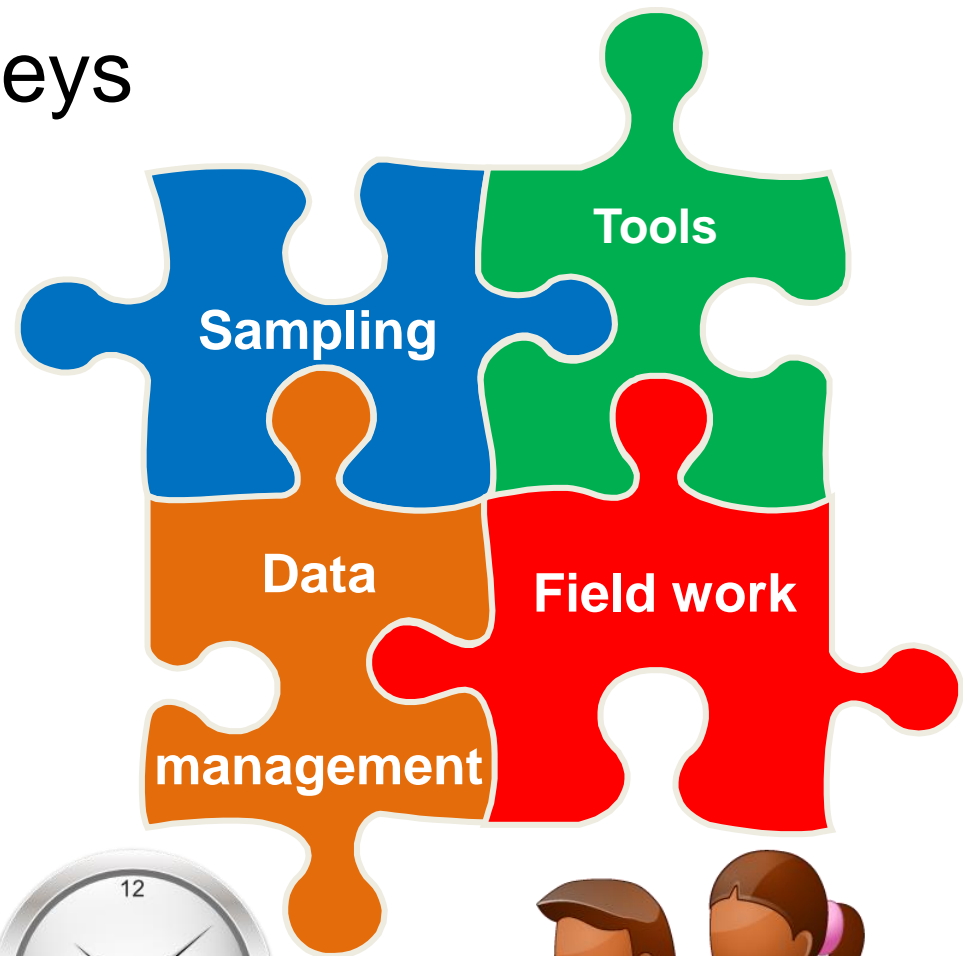
# Quality Assurance in household surveys cont'd

- Reports: are they descriptive, detailed or policy briefs (depending on mandates of NSOs or other agencies conducting the survey)
- How deep in analysis do you want to go?
- You may consider, a 'a quality audit of a few indicators at a frequency to be determined by the organisation
- DQAF Of IMF and SQAF of south Africa are

# Total quality in surveys

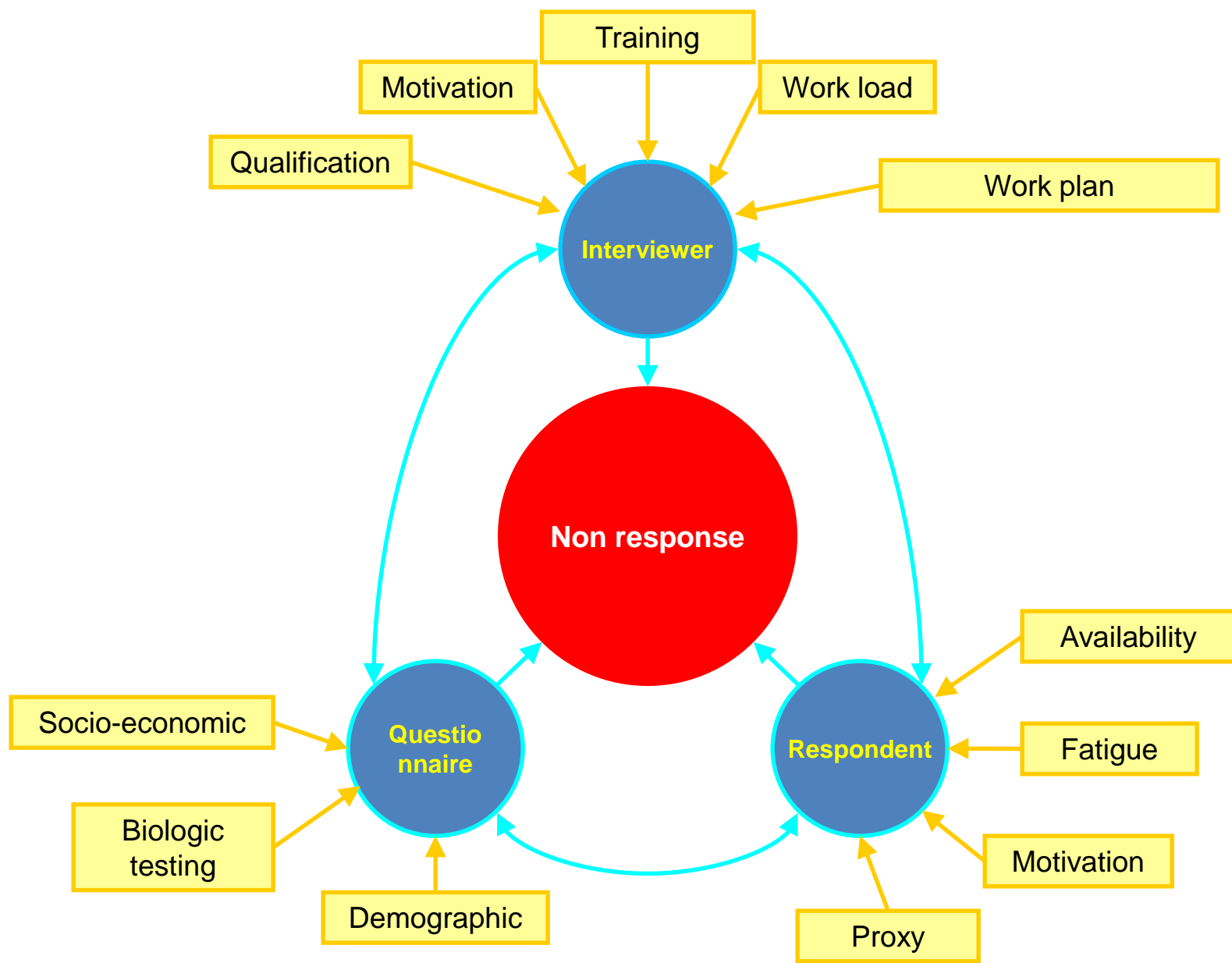
The integration of  
sampling design,  
questionnaires and tools,  
field work  
and data management

with the goal  
of delivering analysts  
a reliable database  
on time



**Data lose their value if they don't  
represent the reality of the day**

**The end product**



# Paper Questionnaire Vs Computer Assisted Personal Interview

Which way to go?

# Why Leave PAPI

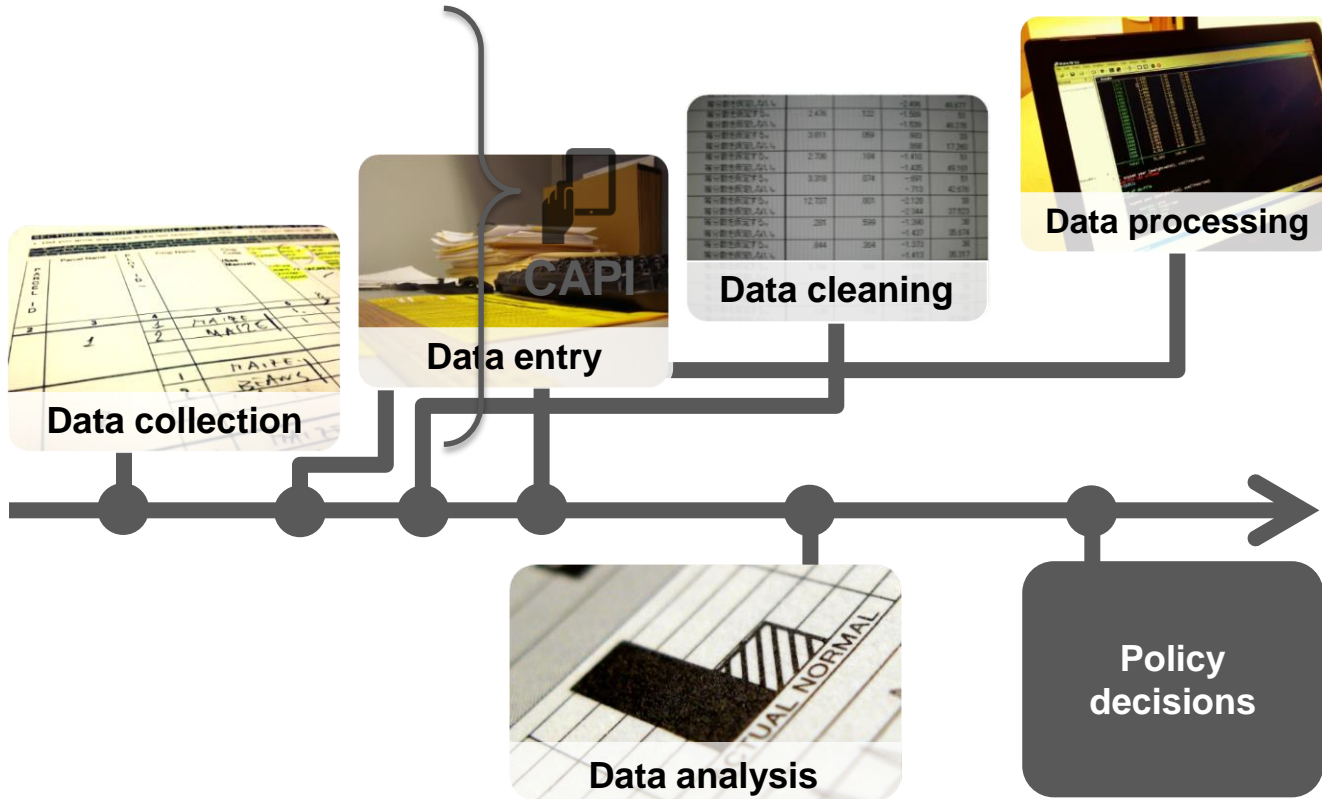


# Why leave PAPI ?

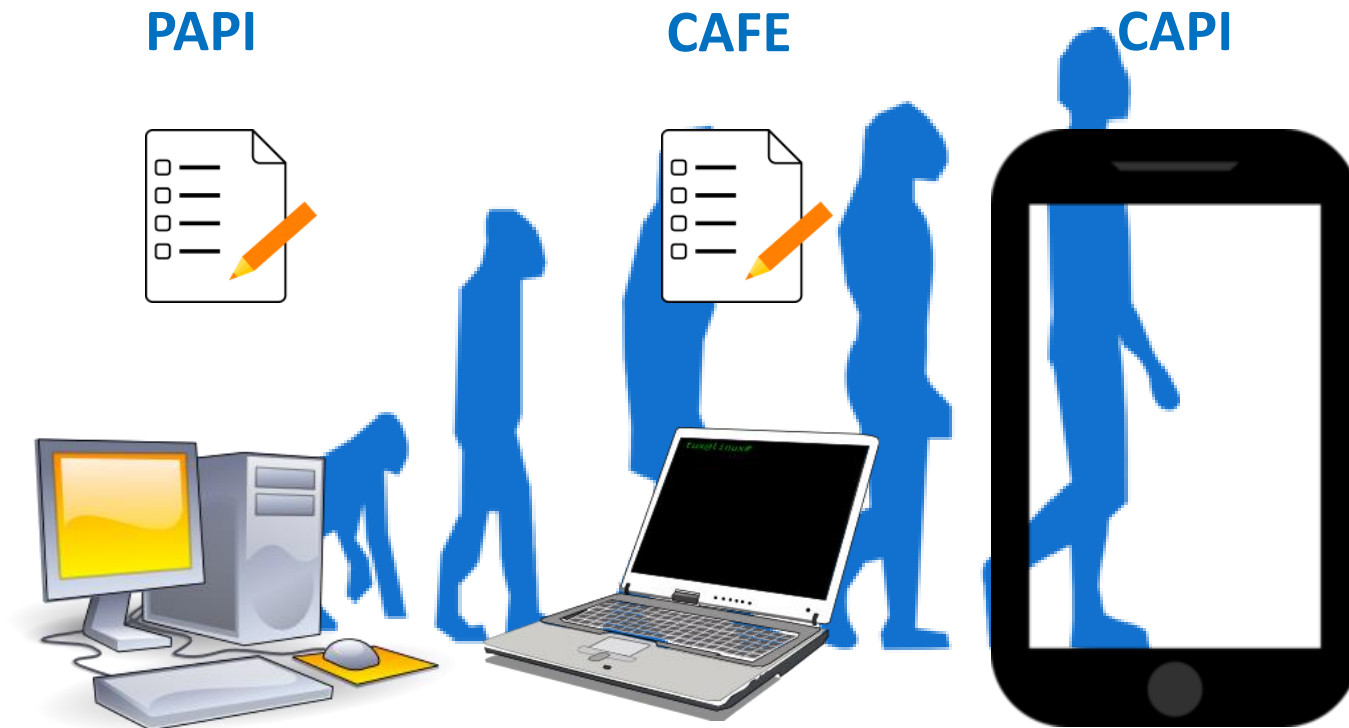
- The demand for data and information has increased tremendously and cannot cope with paper questionnaires
- Data dissemination takes a long time with PAPI
- Technology has evolved and some flexibility is needed to keep with the speed

Why use computer-assisted  
personal interviews (CAPI) ?

# Faster data



# Better quality data



# Cheaper data (maybe)

## No longer pay

- Printing
  - Paper (*as much*)
  - Packing questionnaires
- Data entry
  - Training a team
  - Time to type (twice)
- Transporting questionnaires
  - Fuel and vehicles
  - Drivers

## Now pay

- Hardware
- Software
- Survey programmer time
- Data transfer costs
- Extra (CAPI) training time
- Extra (pilot) testing time

# Cost, in other words

## No longer pay

- Disposable goods
  - Printed surveys
  - Limited printing
- One-time services
  - Printing
  - Data entry
  - Transportation for questionnaires

## Now pay

- Fixed costs, with benefits
  - Survey programmer time
  - Extra training and testing time
  - Internet bundles
- Investments
  - Hardware
  - Software
  - (Survey programmer time)

# Easier data

- Enablement
- Validations
- All automatically

8:19

CAPI experience

Do you have any experience with CAPI software packages?

Yes

No

With which of the following CAPI software packages do you have experience?

Of those packages, which do you like?

Experience with Survey Solutions

8:20

CAPI experience

Of those packages, which do you like?

CSPro

Open Data Kit

**Survey Solutions**

Surveybe

SurveyCTO

**None of the above**

**ANSWER IS INVALID**

*You have provided incompatible answers. You cannot like one and none of the listed software packages at the same time.*

# Different data pooled into one data set

- Relates quantities, individual records and household variables and integrates into one file
- Increasing adoption of CAPI in many countries



Which CAPI software to use?

# Which CAPI to use?

- Review existing solutions
- Ensure that you have access rights to all the processes and data
- Consider what is best for the NSO given the constraints we usually face (resources etc)
- `Some CAPI solutions on the market include:
  - Surveybe, Survey Solutions (WB) CsPro, OPEN DATA KIT among others

# Challenges with CAPI use

- Power challenges
- Technical failures due to use
- Machines with low capacity present a challenge
- Most solutions available on market are good for cross-sectional surveys as opposed to panel/longitudinal studies
- Internet connectivity affects for data transmission to office
- Machines are not an end in themselves, programmers and user inputs are needed
- Restriction to modification

# Some possible solutions

- Adequate preparation in design, testing, discussion and adoption
  - Programmer, statistician and field staff need to work together
- Programmers need to give room for modifications
- Capacity should be built in statistical offices to use the application
- Quality of the machines is key as panel surveys last longer periods

# -----Some possible Solutions

- Backup power plan is needed (like generators, power banks etc)
- Enough internet bundles allocation is needed
- CAPI misuse control

# The Sample frame as a foundation of household Surveys

- Irrespective of the type of research, the foundation of sound scientific studies is rooted in its design.
- Well designed sample surveys or assessments thus provide the best solution to most of the research questions

# Sampling frame cont'd

- Frame development for a stand-alone facility survey differs from frame development for a survey linked to a set of population-survey sampled areas
  - A list of facilities by type, together with figures on students and staffing, patient load, number of hospital beds and staff, would serve this purpose
  - A perfect list – an unattainable ideal – must be complete, accurate and current. Frame lists that have been tried in facility surveys have suffered from various flaws.
  - A list of public/ private facilities are available from authorities such as the Ministry of Health, education but these lists are often incomplete or outdated.

# Sampling Frame cont'd

- Sampling frame should be representative and have complete coverage, with appropriate measures of size (households or population, businesses, programs/projects, total value, etc.)
- Establish all outcome indicators identified for measuring programme performance-a possible sampling frame
- Generally these should be obtained from appropriate institutions
- Sampling frame by sector with appropriate measures of size



# Sampling frames cont'd

- Components should include:
- Outcome indicators to be used to measure programme performance
- Value of the program (highest value Vs highest impact)
- Institutions to make it happen (who are the key actors if they are many); remember, you do not need to monitor all
- Duration of implementation and the outputs and immediate outcomes to measure over the period of the program
- Source of data to measure performance
- Capacity to collect the data

# Construction of Sampling frame

- (a) start with an available list and update it
- (b) conduct an independent listing operation from scratch,
- (c) compile a list (dependently or independently) but restrict it to a sector such as large facilities and use it in combination with a geographic area sample.

## *List frame sample*

- a) Compile a list of the large (or otherwise significant) facilities for the study universe.
- b) Stratify the list in an appropriate way.
- c) Make it comprehensive and mutually exclusive
- d) Ensure you cover all units of interest without omission or duplication

## *Area frame sample*

- a) Compile a list of geographically defined areas that cover the entire study universe (example, census enumeration areas).
- b) Stratify the list of areas in an appropriate way depending on the anticipated use.
- c) include service delivery points (SDPs) and any other information that is likely to inform sample design
- d) Regularly update the frame to ensure its complete and exhaustive

# Data transmission data security and storage

## *Data transmission*

- a) Data transmission under the paper arrangement was not a problem because it was all undertaken at the head office
- b) With CAPI, data is electronically wired to servers (cloud or Office)
- c) The stability of power, internet, and access to sufficient bundles by all field teams is critical
- d) Need data transmission protocols (part of the design) ‘

## *Data encryption*

- a) Data transmission should follow properly laid out procedures
- b) Access rights and those with rights to edit must be clear
- c) Who anonymises the data (should be designed and programmed to automatically become activated once data transmission is activated)

# *Data Storage*

- a) Data storage is very important to safe guard confidentiality of respondents
- b) Access rights and who authorises(levels of authority to access data)
- c) Do you go for your own dedicated server or hire a cloud server?
- d) What are the experiences in the region regarding use of third party cloud servers
- e) Consider what is acceptable to the organisation. Internal discussions with IT team, Head of departments key for buy-in



**END**

**THANK YOU**