

Data Quality Checks in a Paperless HDSS: Lessons from the Cross River HDSS, Nigeria

Iwara Arikpo, Ideba Mboto, Anthony Okoro, Martin Meremikwu

Cross River HDSS
Directorate of Research, University of Calabar
Cross River State, Nigeria



INDEPTH Network



Outline

- * Background
- * Introduction
- * Data Quality Checks in Paper-based HDSS
- * Problems with paper-based systems
- * Data Quality Checks in Paperless HDSS
- * Lessons Learned
- * Limitations
- * Conclusion



Background

- * CRHDSS established Dec. 2012 after pilot
- * Akpabuyo LGA (rural) & Calabar Municipal (urban)
– Cross River State.
- * Population under surveillance: **32,301**
- * Occupation:
 - * Rural cohort: subsistence farming, fishing
 - * Urban cohort: public service, business, etc



Introduction

- * HDSS – robust longitudinal systems for collecting & reporting health and demographic events.
- * Vital events include births, deaths, migrations & socioeconomic indicators
- * Information from update cycles assist stakeholders (govt. & NGOs) on policy formulation & implementation
- * Achievement of HDSS goals lies on the quality of data.
- * INDEPTH recommends spot-checks, 5% re-interviews, etc.



Quality checks in Paper-based HDSS

...1/2

- * Stage I – Error Checking:
 - * FWs log in their paper forms with Filing Clerk at FO each end of cycle
 - * Research Assistant (or equivalent) checks forms for completeness and errors
 - * Confers with relevant FW if errors are found
 - * FW goes back to the field if error cannot be resolved

Quality checks in paper HDSS ...2/2

- * Stage II – Quality Checks Interview
 - * Supervisor randomly selects 5% of all household events for re-interview
 - * Conducts re-interview using new data forms
 - * Compares new data with what was collected by FW for discrepancies
 - * A third person visits same households to verify data if there are discrepancies between FW & supervisor
- <see logical steps next>*



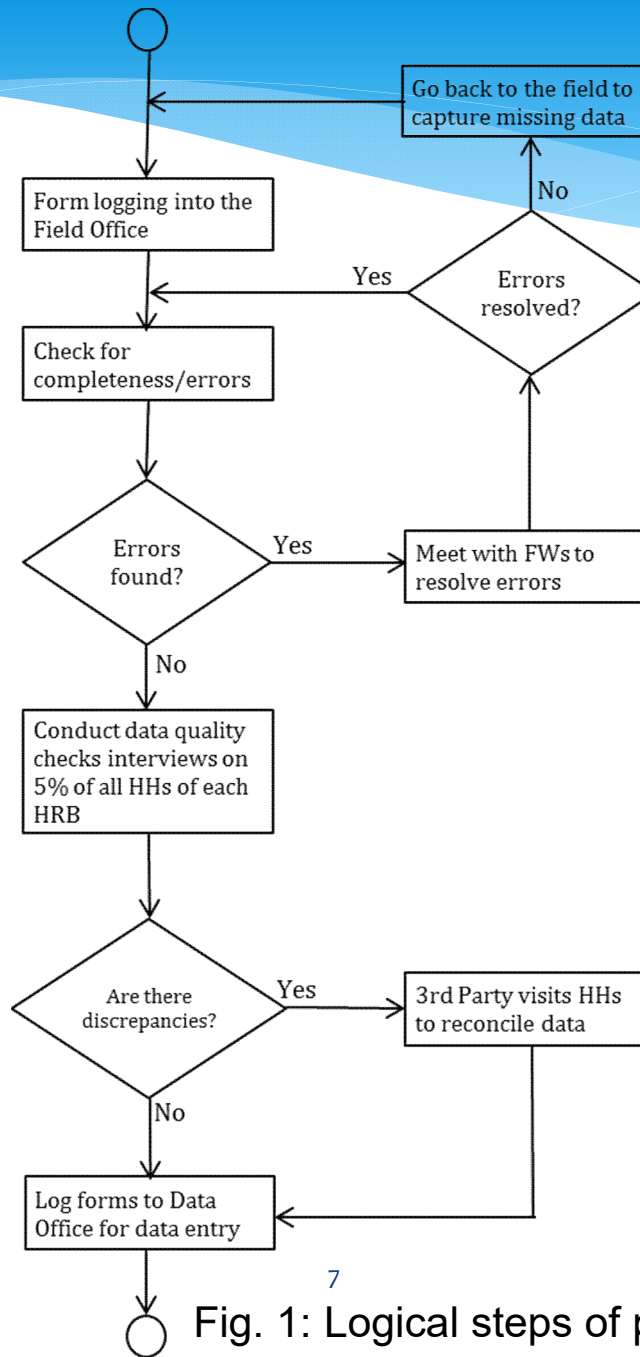


Fig. 1: Logical steps of paper-based quality checks



Problems with paper-based systems

- * High costs on time and resources
- * Other quality checks by DQA supervisors are delayed until cycle of field checks & data entry are completed
- * Data analysis and reporting is stalled until DQA is completed
- * Time-bound interventions may be missed or delayed due to late availability of reliable information
- * Blame-game syndrome (field workers vs. quality supervisors)
- * Shortcomings of the red herring technique



Data Quality Checks in Paperless HDSS

- * Events data is collected and transmitted directly from field (no logging in at FO)
- * No separate data entry required
- * First stage of quality checks is at data capture through in-built validation rules
- * Second stage of checks is at database update using middleware applications
- * Re-interviewing of 5% of each type of events from the cycle. Records randomly selected by software

<see logical steps of process>



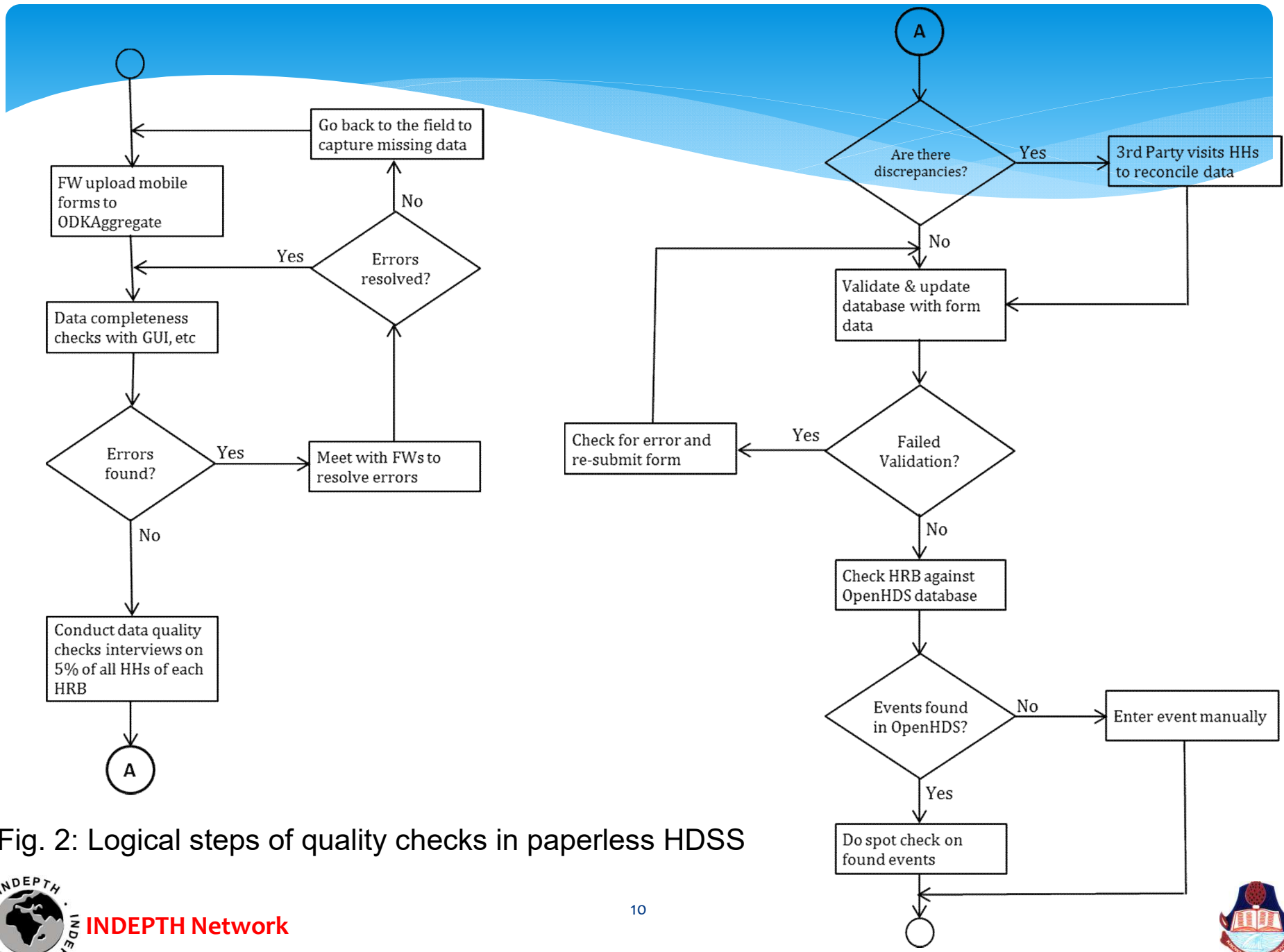


Fig. 2: Logical steps of quality checks in paperless HDSS

Lessons Learned from the use of electronic methods ... 1/2

- * Re-interviews are conducted after records have dropped onto the system
- * Record selection bias is avoided by using a randomizer (e.g. Research Randomizer – see *next slides*)
- * Re-interviews are done with paper forms and results compared with electronic data



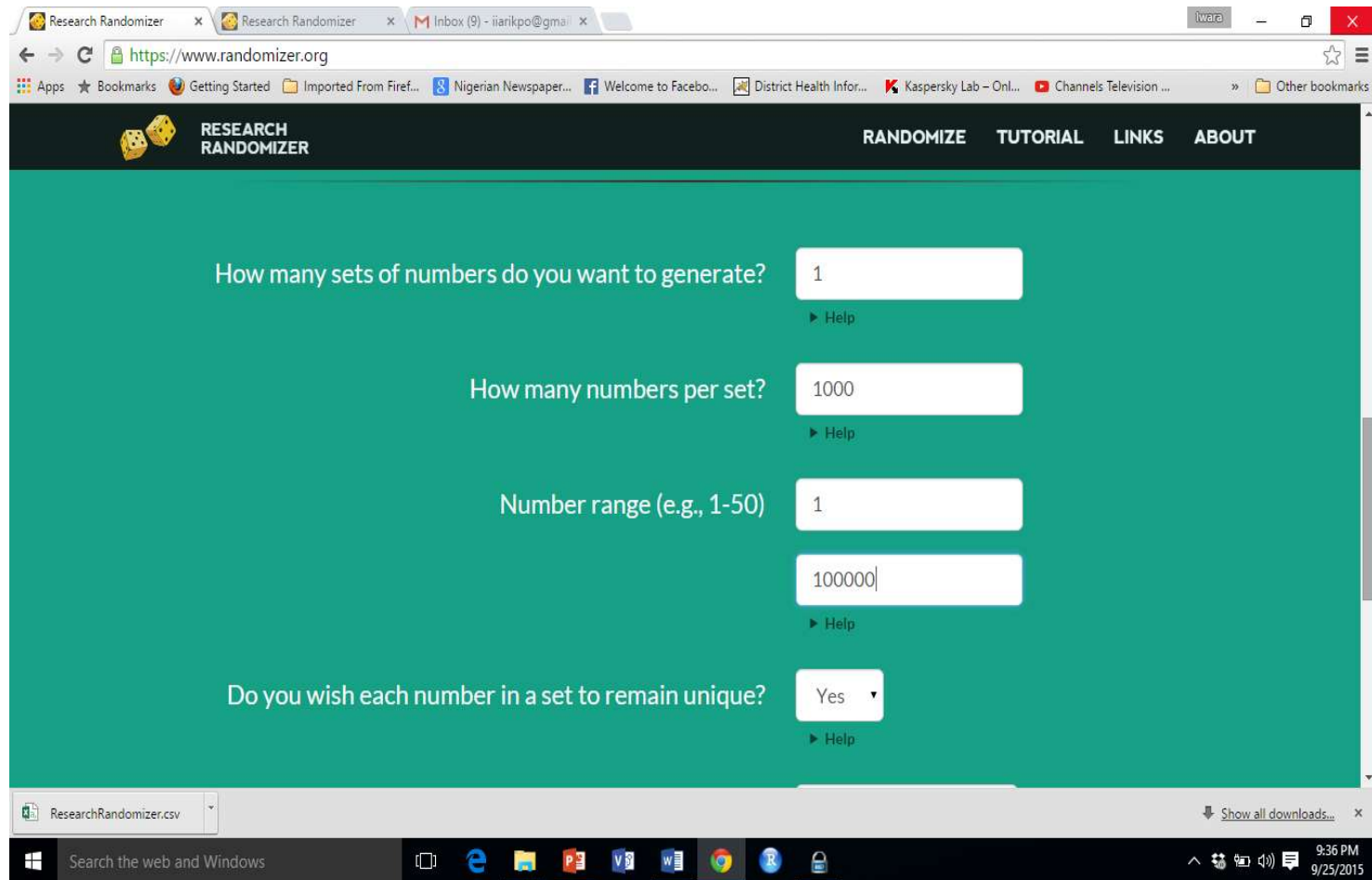
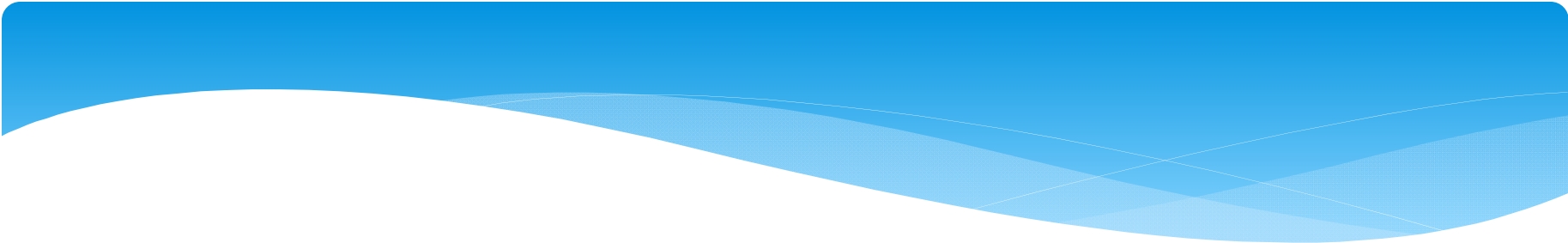


Fig. 3: Input interface for Research Randomizer





RESULTS PRINT DOWNLOAD CLOSE

1 Set of 1,000 Unique Numbers
Range: From 1 to 100,000— Sorted from Least to Greatest

Set #1

108, 313, 350, 377, 497, 761, 789, 800, 939, 1056, 1112, 1315, 1352, 1480, 1619, 1635, 1651, 1785, 1884, 1996, 2200, 2232, 2277, 2527, 2795, 2932, 3080, 3087, 3089, 3306, 3619, 4038, 4046, 4092, 4097, 4216, 4570, 4696, 4732, 4745, 4768, 4907, 4918, 5058, 5157, 5392, 5422, 5470, 5761, 5824, 5875, 6153, 6354, 6391, 6421, 6884, 6919, 7083, 7113, 7182, 7245, 7441, 7510, 7578, 7637, 7664, 7696, 7774, 7993, 8095, 8109, 8174, 8347, 8369, 8408, 8420, 8618, 8810, 8881, 8973, 9002, 9006, 9054, 9077, 9142, 9182, 9298, 9384, 9459, 9467, 9476, 9484, 9513, 9571, 9625, 9707, 9908, 10185,

Excel interface showing the results of a randomizer. The spreadsheet displays the following data in column A:

Row	Value
1	Research Randomizer Results:
2	1 Set of 1000 Numbers Per Set
3	Range: From least to greatest
4	Set 1
5	108
6	313
7	350
8	377
9	497
10	761
11	789
12	800
13	939
14	1056
15	1112
16	1315
17	1352
18	1480
19	1619
20	1635
21	1651
22	1785
23	1884
24	1996

Fig. 4: Result interface for Research Randomizer

Lessons from the use of electronic methods ...2/2

- * Data quality checks in electronic-based HDSS overcomes the time consuming error-prone approach in traditional paper-based HDSS
- * The random record generator requires active internet connection to use
- * The subset (5%) of records for random generation is manually computed and fed in as input



Limitations

- * The requirement for internet connection for the Research Randomizer to run can delay the random record generation
- * The subset (5%) of records to be randomly generated is computed manually
- * The re-interviews are still done with paper forms after extraction instead of mobile devices



Solution strategies

- * Our IT team recently wrote a script in R that computes the 5%, generates random records and exports the random numbers in Excel/CSV for lookups
- * Currently working on extending this script to include extracting the random records directly from the database
- * Future work will focus on conducting re-interviews with mobile devices



Conclusion

- * Quality of data collected from HDSS sites can influence various aspects of the healthcare and socioeconomic system, hence the need for premium attention to data quality assurance (DQA)
- * The use of electronic data collection approaches, although preferable, comes with new challenges for DQA
- * We have learned that, HDSS sites with good IT teams can leverage the full strength of paperless HDSS in terms of improved DQA process



Acknowledgements

- * **University of Calabar** – office space, utilities, salaries of core scientific team members, field workers, etc.
- * **Cross River State Government** – facilitates our work with MOH
- * **Akpabuyo LGA & Calabar Municipal** – cooperation and support
- * **INDEPTH Network** – Technical support and ISC 2015
- * **Cross River HDSS Field & Data Staff** – diligence and dedication
- * **University of Calabar Teaching Hospital** – technical personnel



Thank You



INDEPTH Network

