Using Routine Data for OR/IS & Data Quality Audits

Sarah Gimbel, MPH, RN
HAI/DGH

Introduction: Measurement

- Measurement of exposures, outcomes, other characteristics are a key part of most epidemiologic studies
- Measures take many forms:
  - Response on self-administered questionnaire
  - Answer to interview question
  - Lab result
  - Symptom recorded in medical record
  - Physical finding
  - Diagnosis code in a database

Measurement Error

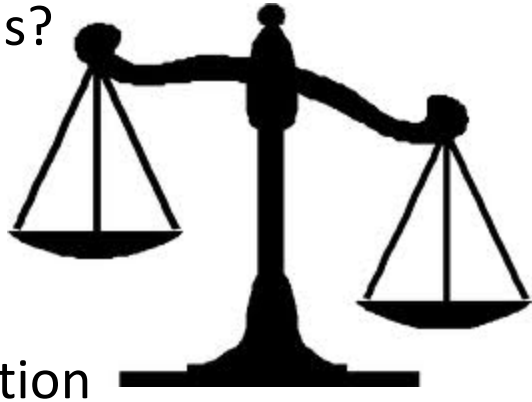
- Nearly all measures are imperfect
- Quantifying a measure's performance helps in
  - Choosing among alternative measures for same purpose
  - Interpreting study results

New data creation

- Collecting new data to determine information "more accurately" or "more completely"
  - Questionnaires, direct observation, blood samples
- Added procedures (not routine)
- Indicators: variable / infinite
- Pros:
  - Can often obtain more accurate data, data otherwise not available, or data not from appropriate time-frame
- Cons:
  - Expensive, time consuming, suffers from biases and limitations in generalizability of study

Routine vs. “New” data

- Balance availability vs. accuracy vs. cost
  - Does routine data have relevant indicators?
  - Is routine data collected at appropriate time-points for study?
  - Is routine data accurate enough to make decisions?
  - Is need for “new” data justified by cost/time?



- Remember:
  - Strengthening routine data systems always an option
    - Can be followed over time and used for variety of purposes
  - Information is rarely “perfect”
    - Avoid paralysis by analysis

Lots of types of data

- Available data (routine)
  - Health system (administrative)
    - Vital registration
    - Disease registries
    - Sentinel surveillance systems (prevalence, resistance)
    - Health facility data (HIS, TIMS and vertical programs)
    - Patient files
    - NGO databases, other partner databases
  - Community surveys
    - Sample vital/civil registration (birth, death, marriage)
    - Demographic Surveillance Sites (DSS)
    - Multinational health surveys (WHO WHS, DHS, MICS)
    - Census
- Not routine
  - New study

Can health systems data be used?

- Some concerns about validity, so verify!

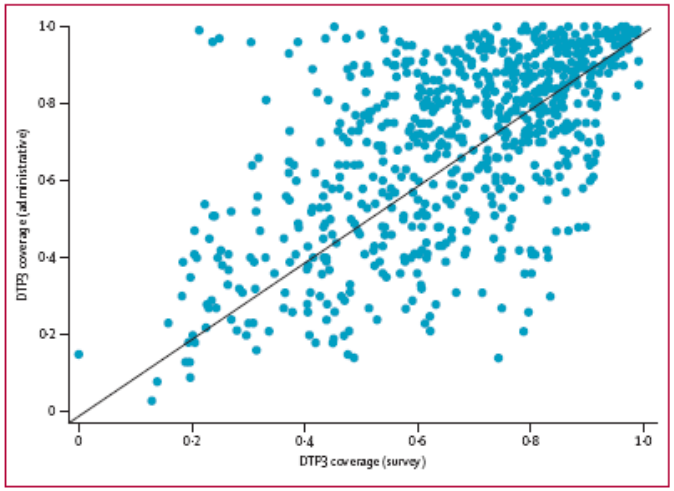
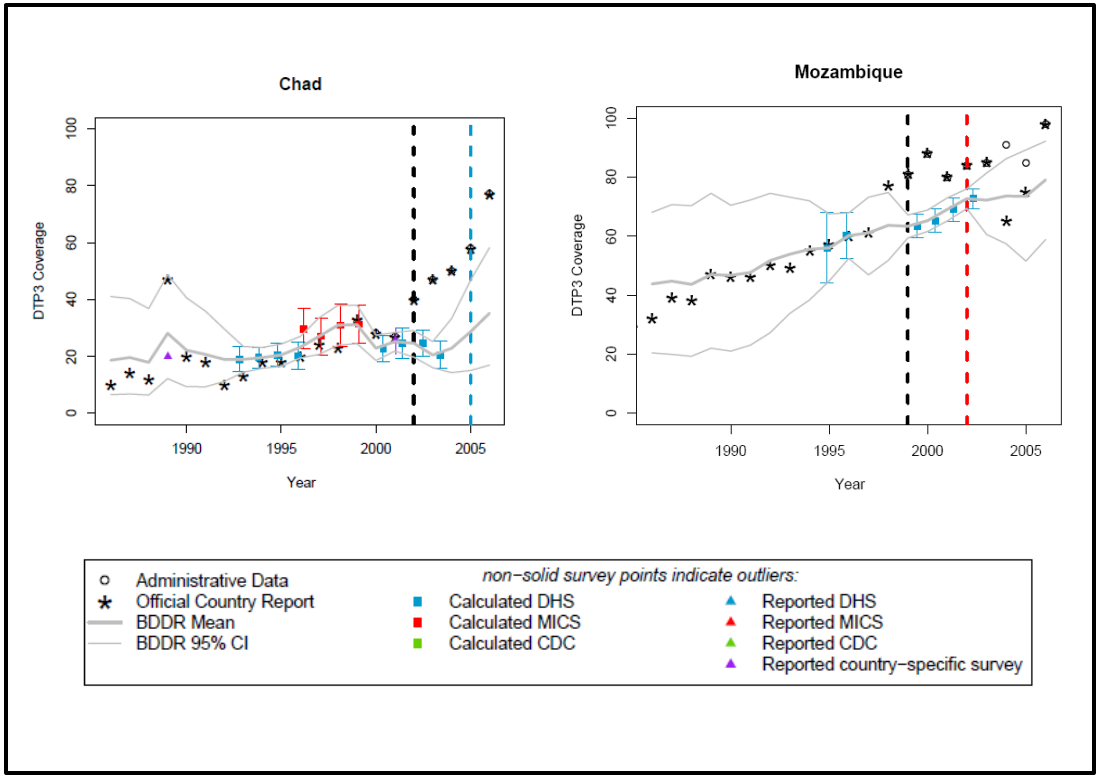


Figure 2: Comparison of three doses of diphtheria, tetanus, and pertussis (DTP3) immunisation coverage estimated from administrative data and estimated from surveys



From: Lim SS, et al, Lancet 2008; 372:2031-46

Snapshot review: Reliability & Validity

Poor reliability
Poor validity

Good reliability
Poor validity

What sorts of primary records will you use? Depends on what you want to measure….

- For people reached: Medical records, Registers, Tally sheets, etc

- For commodities distributed: Distribution log sheets, Inventory statements, etc.

- For people trained: Attendance sheets. Per diem sign-up sheets, etc.

Ways to check health systems data

- Look for consistency over time
- Disaggregation
  - Outliers
  - Missing data (can cause wide fluctuations)
- Compare to other data sources
  - Surveys: often “gold standard” but also have methodological limitations
  - Compare facility reports to other health systems data (patient charts, prenatal cards, pharmacy records)
  - Directly observe clinical services, and compare to point-of-care registers

Example: ANC and pMTCT coverage 2 districts in Northern Cote d'Ivoire

	District A	District B
% ANC coverage	95%	96%
% pMTCT	88%	103%

ANC and HIV testing data, by health facility

District	Pop (2010)	HF	Number of women attending 1st ANC												Total	Est Preg	% Coverage	HIV test	% HIV test
			Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec					
A	28298	1	79	76	78	61	77	68	72	101	52	85	68	98	915	1273	72%	801	88%
	5676	2	38	39	42	31	49	37	30	30	31	29	26	22	404	255	158%	431	107%
	9202	3	49	32	38	29	40	39	29	23	32	30	24	39	404	414	98%	487	121%
	3399	4			26	19	20	19	22	17	20	22	25	21	211	153	138%	204	97%
	6031	5	36	27	36	23	43	26	36	14	23	0	0	0	264	271	97%	129	49%
	4027	6				24	12	11	15	15	9	29	11	8	134	181	74%	153	114%
	3967	7			18	29	30	10	30	23	31	29	29	17	246	179	137%	280	114%
	60600	**TOTAL**	**202**	**174**	**238**	**216**	**271**	**210**	**234**	**223**	**198**	**224**	**183**	**205**	**2578**	**2726**	**95%**	**2485**	**96%**
B	31425	1	60	55	115	81	94	87	48	90	67	40	54	45	836	1414	59%	950	114%
	11966	2	80	62	77	61	52	43	54	57	39	42	41	36	644	538	120%	690	107%
	2304	3	53	41	52	59	43	33	49	49	39	54	48	25	545	104	524%	519	95%
	10544	4	37	22	35	25	34	21	24	20	13	24	22	23	300	474	63%	285	95%
	13658	5	55	69	72	59	71	38	38	68	37	49	48	46	650	615	106%	674	104%
	13891	6	36	41	44	36	38	30	29	16	20	21	4	37	352	625	56%	311	88%
	4135	7	40	12	9	11	17	10	16	11	8	11	0	0	145	186	78%	139	96%
	87923	**TOTAL**	**361**	**302**	**404**	**332**	**349**	**262**	**258**	**311**	**223**	**241**	**217**	**212**	**3472**	**3956**	**88%**	**3568**	**103%**
Both	**148523**		**563**	**476**	**642**	**548**	**620**	**472**	**492**	**534**	**421**	**465**	**400**	**417**	**6050**	**6682**	**91%**	**6053**	**100%**

3 Types of Data Verifications

Type 1: Cross Verifications

- Cross verifications involve comparing programmatic results, for a given period, with other sources of information such as financial or inventory data.
- Identified discrepancies may suggest data-quality issues (for example, stock levels of ARVs during the period may contradict the number of patients reported on ARVs)

Ex. Assessment of MOH Mozambique ART program data sources

- Routine analysis found discrepancies in # patients on ART in clinic versus pharmacy records (128,330 clinic versus 96,858 pharmacy reports)

Province	Mean
Manica	99%
Sofala	87%
Tete	84%
Cabo Delgado	78%
Maputo	75%
Maputo Cidade	73%
Nampula	73%
Zambezia	70%
Gaza	68%
Niassa	67%
Inhambane	66%
Total	77%

Problem

- What would be the next steps?
  - What do you want to know?
  - What approach would you take?

Investigate Causes of Data Discrepancy

Dr. Ema Chuva, Manager of AIDS Treatment Section, Ministry of Health, Mozambique, et al.

- Objectives: To investigate the causes of discrepancy between the pharmacy and clinic reports, and offer recommendations to bridge the gap between these two sources
- Design: Cross sectional, 2 health facilities each in 5 provinces (total 10 sites), one with low discrepancy of data and the other with high discrepancy. Compared quality of gold standard (all ART patient charts) to CMAM (Pharmacy Unit) and DNAM (Clinical Medicine Unit) reports
- Population: 11,091 patients currently on ART at 10 HF across 5 provinces in Mozambique

Preliminary Results

- Smaller facilities had better results
- Regions with consistent, singular partners did better
- Major differences among sites in how LTFU patients were identified and categorized
  - Sites with HIV+ activists did better, as this strategy addressed the major building block challenge of HRH in the health system

3 Types of Data Verifications-con't

Type 2: Spot checks

- Spot checks of actual delivery of services and commodities involve sampling individuals or households from records or registers and conducting short interviews

3 Types of Data Verifications-con't

Type 3: Bottom-up Audit Trail

- Verify the availability of primary records (at the service delivery points) and of summary reports (at national and other relevant administrative levels where data are aggregated-provinces, districts)

- Check the accuracy or probability of recorded events in primary records (e.g outliers, impossible # of events recorded in one day)

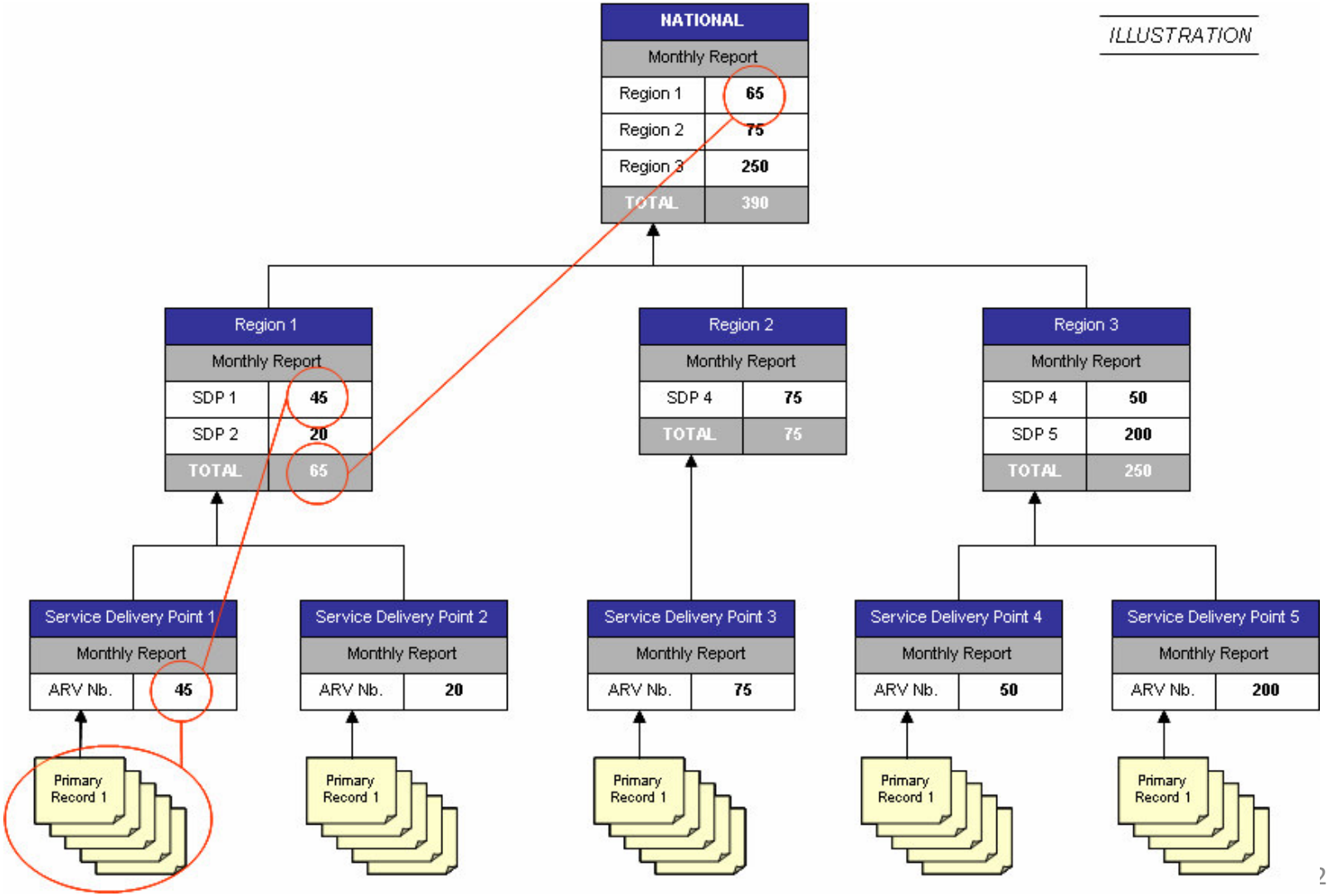
- Re-aggregate data from primary records and compare with results in summary reports up to the National level.

Data Audits/Verification: General Norms

- Takes between 2-5 days (up to 7 in high risk situations)
- Audit should concentrate on the most important indicators related to your area of research interest (or program)- 1 to 3 indicators is typically enough
- Concentrate on the most important regions or districts from the point of view of your proposed research/program

Bottom-up Audit Trail:

Recommended at least annually for LFA of GFATM



What is good enough?

- GFATM rubric says

Rating	Metric
A	Less than 10% error margin
B1	Between 10%-20% error margin
B2	Above 20% error margin
C	No systems in place

Note: the data verification rating only pertains to the sample of data verified; it is not meant to be representative of all the results report in the national HIS

Bottom-up audit in Mozambique*

- Pilot sample of PHC from 9 health facilities across 3 districts
- Determine the availability of monthly facility reports at the health facility and district health departments (12 months)
  - Presence/absence of monthly reports for (1) institutional births and (2) DPT3
- Determine the reliability (concordance) of monthly statistics obtained from facility clinical registries, monthly facility reports and the MOH electronic database (6 months)
  - Proportion of months where data were identical
  - Calculate % difference for months where data not identical
  - 5 key indicators: 1ANC, Institutional birth, DPT3, HIV testing, Outpatient Consults

** Gimbel S et al, An assessment of routine primary care health information system data in Sofala Province, Mozambique. Population Health Metrics. 2011, 9:12.*

Methods (con't): Comparisons with DHS/MICS

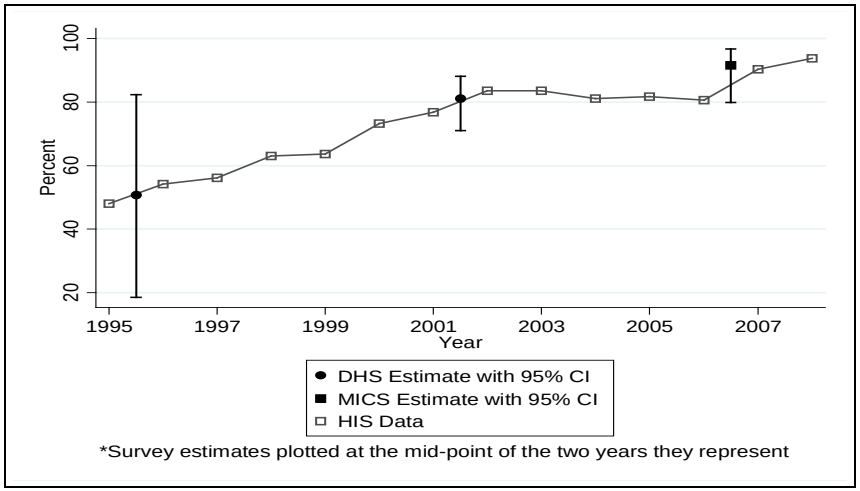
- Examine the validity of HIS data by comparison with population-level surveys
  - 3 key indicators: 1ANC, institutional birth, DPT3
  - Compared statistics from the provincial health department's annual reports (derived from MOH electronic database) with those obtained from the 1997 and 2003 DHS and 2008 MICS.

Summary of results

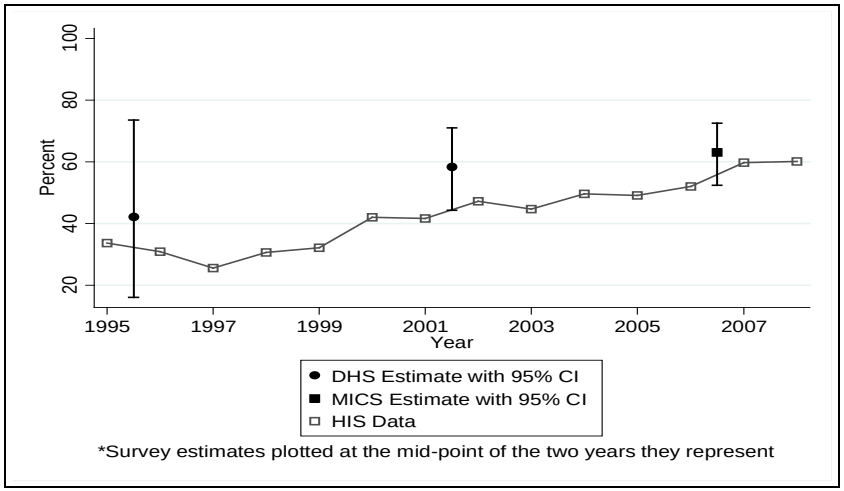
- Availability of monthly reports: 98.1% at health facility, 98.6% at district health department
- Concordance between monthly facility reports and MOH electronic databases = 98.0%
- Weaker concordance between monthly facility registers and facility reports = 80%
  - Clustered around 2 main facilities
  - 86% of differences were <10%

Comparison with DHS surveys

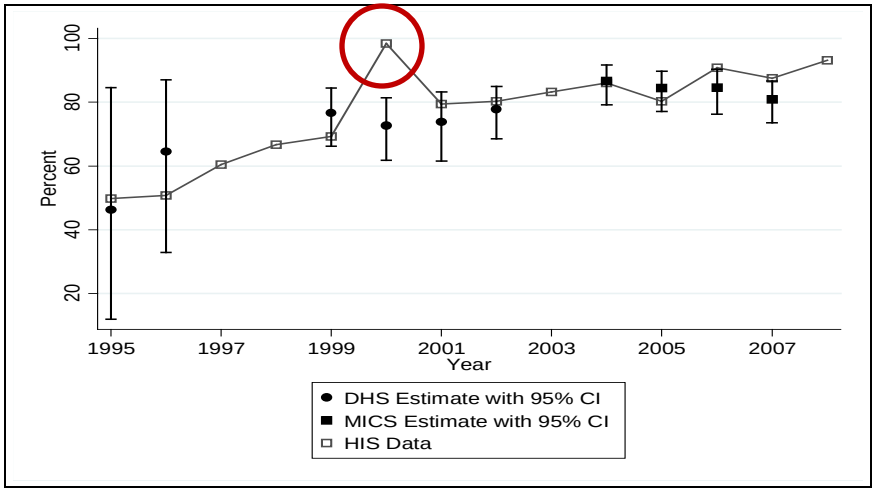
ANC care



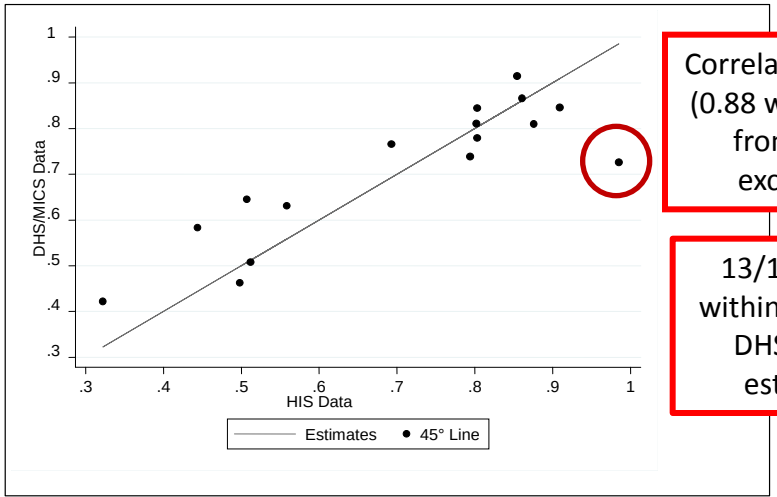
Institutional delivery



DPT3



Statistical comparison



Correlation = 0.73 (0.88 when DPT3 from 2000 excluded)

13/16 (81%) within 95%CI of DHS/MICS estimate

How did we use these findings?

- Obvious data weaknesses at district capital hospitals and in ANC and maternity clinics
  - Information used to target program strengthening work at those sites/clinics
  - QI collaboratives—one type of activity to foster health manager use of routine data
  - For OR this data was fairly strong and viable to use for future research
- Overall strength of data validity when compared with DHS/MICS is encouraging
  - Advocate at national level for increased inputs for routine data systems—it will be worth it!

Working together with MOH colleagues to collect and review health information builds capacity to use data to improve health services