Fundamentals of Implementation Science

Routine data sources and quality for implementation science: opportunities and complexity

Week 1 methods lecture

Welcome to the first week of the fundamentals of implementation science course. The purpose of this lecture is to provide an introduction to *routine* data sources commonly used in implementation science practice. Available data sources differ by country, and the data sources that you will choose to use depend upon your research question of interest, the purpose of the intervention you are implementing or program you are evaluating, and the inherent biases or missing data you are willing to assume. So, let's get started.

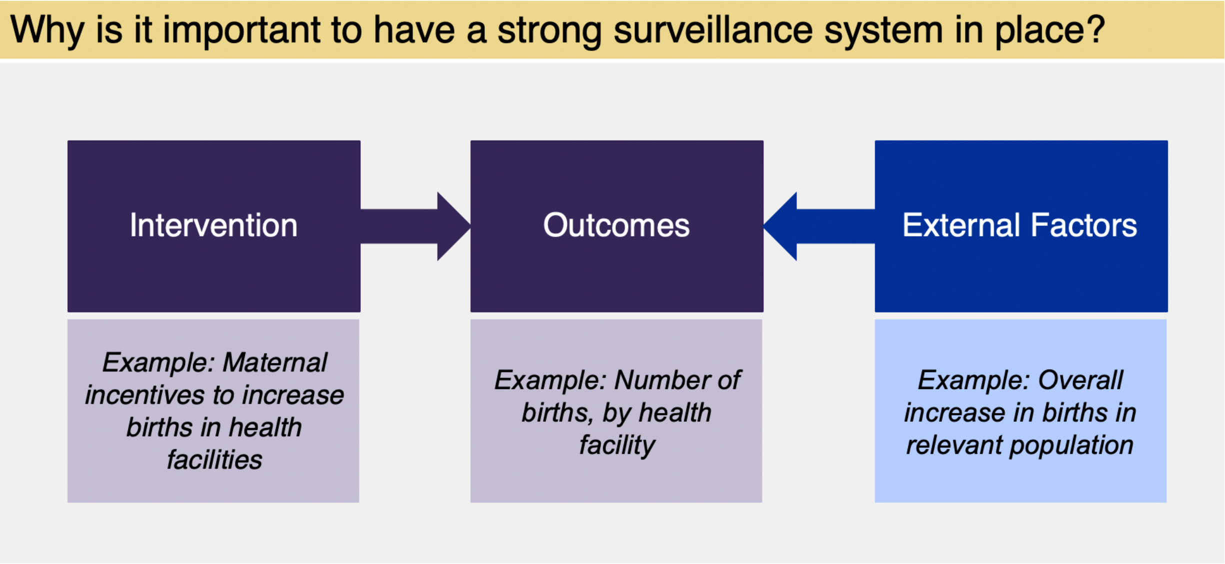
We will begin this lecture by introducing you to the concept of surveillance and discuss why it's important to focus on data quality. Next, you will learn about six surveillance data sources (typically available for implementation science work) with a particular focus on administrative data. Then, you will be introduced to two strategies, used to evaluate data quality. And we will finish this lecture with a few points about how to approach surveillance system evaluation. Along with some key take away messages.

Surveillance Basics

First, let's talk about surveillance basics. Most people are familiar with the concept of surveillance. Take a moment to think about how YOU define surveillance within your work.

- When you talk about conducting surveillance, what data specifically are you referring to?
- Why is it important that you collect these data regularly?
- What do you do with the data once collected, and why?

According to the Centers for Disease Control and Prevention, the definition of public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health. In other words, public health surveillance provides routine data needed for public health action.



Why is it important to have a strong surveillance system in place?

Intervention

Example: Maternal incentives to increase births in health facilities

Outcomes

Example: Number of births, by health facility

External Factors

Example: Overall increase in births in relevant population

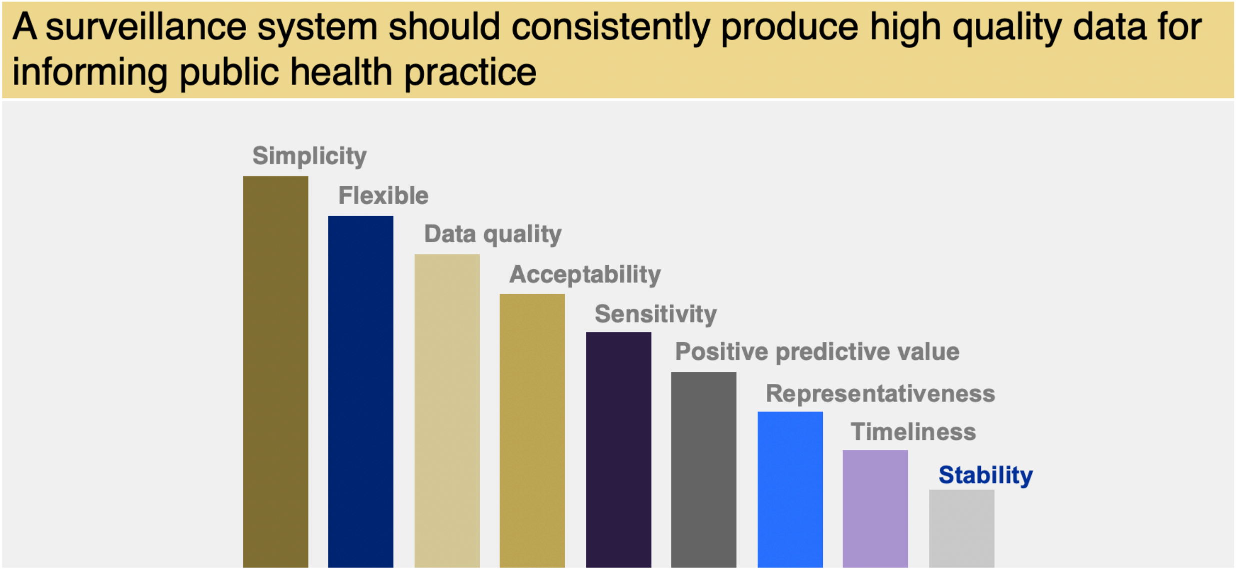
MMWR (2001, July 27) Updated Guidelines for Evaluating Public Health Surveillance Systems 50(RR13);1-35

Implementation scientists depend upon strong surveillance systems and data sources for a number of reasons. One reason might be in order to provide outcome data to assess if an intervention or program of interest had the desired outcome. For example, if a new program provides incentives to pregnant women to deliver in health facilities they would want to use facility service delivery data to understand if the number of births at health facilities changed as a result of the program. At the same time, data sources could be used to monitor other trends that might influence observed outcomes. For example, *population surveys* could be used to determine if there is an overall increase in births in the population around the time of the intervention, which might influences changes in the number of births observed in the facilities.

In addition to evaluating interventions or programs, other important reasons for why strong, routine information systems are necessary include: to assess the need for various interventions; to detect changes in public health practice; to monitor disease morbidity and mortality; to provide a basis for predicting the future course of an infectious or non-infectious disease; to improve resource management and monitor costs; to provide a basis for epidemiologic research; to support policy formulation and resource allocation; and to evaluate policy.
So surveillance data are important for understanding health trends and the outcomes of our interventions, but also as the global health community embarks on ambitious goals such as those set out in the Sustainable Development goals, the ability to document and track disease control efforts and expended resources is increasingly important. Additionally, funding and support for public health activities – such as prenatal syphilis screening or immunization programs – are usually contingent upon a demonstration of sufficient coverage using routine statistics. Thus for implementation scientists, surveillance data are critical to guiding program design, evaluation, and policy.

There are three broad categories of surveillance which should be differentially utilized according to the specific goal of the disease control program at hand.

- Data produced by **passive surveillance** are mainly what you will learn about in this lecture. These data are produced by standardized reporting systems including disease case reports, such as HIV test positives, and service utilization reporting, such as attendance at antenatal care. So note that passive surveillance is not a proxy for disease prevalence in the broader population. But rather it provides a really important opportunity to measure the changing burden of disease in a specific population sample, typically individuals who access or seek care.
- **Active surveillance** involves outreach by public health workers to identify prevalent cases of a disease. For example, community health workers may be asked to conduct active case detection for measles cases in the community in order to identify patients in the early stages of the disease or who may not be coming into a health facility for care.
- **Sentinel surveillance** is conducted in a random sample of the population. For example, in a generalized HIV epidemic you might conduct passive surveillance, but also set-up sentinel surveillance among pregnant women in antenatal care, thus conducting a clinic-based sero survey.



A surveillance system should consistently produce high quality data for informing public health practice

German et. al. 2001. Updated Guidelines for Evaluating Public Health Surveillance Systems. MMWR.

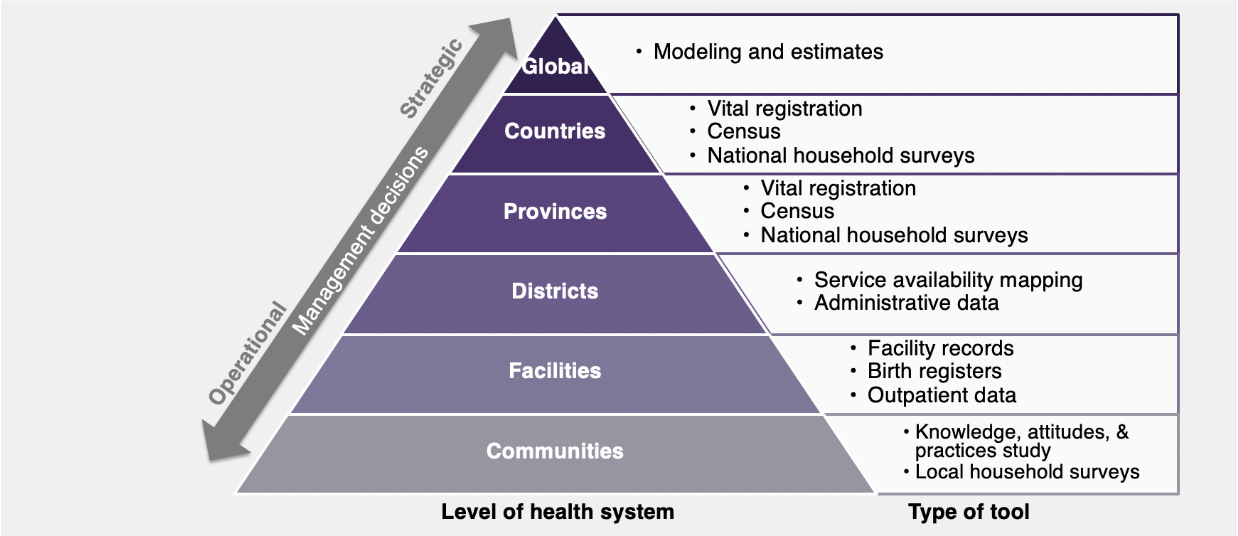
A surveillance system should consistently produce high quality data for informing public health practice. There are nine key attributes that a surveillance system should possess in order to effectively produce high quality data.

1. First, the system should exhibit **simplicity** in both its structure and its ease of operation. Surveillance systems should be as simple as possible while still meeting their objectives.
    a. For example, for reportable diseases, such as meningitis, are their automatic mechanisms in place that raise a flag to supervisors if a case of this disease is detected?

2. Second, the system should be **flexible**, which means the system can adapt to changing information needs or operating conditions with little additional time, personnel, or funds.
    a. For example, can the data system adapt to a changing case definition for a specific disease?
3. Third, the system should produce high **quality data**. Meaning the data are complete and valid; we will focus much of the rest of this lecture on this attribute in particular.
4. Next, the system should be **acceptable**. This means that people and organizations are willing to participate in the surveillance system by providing accurate, consistent, complete, and timely data.
5. Next, the system should exhibit adequate **sensitivity**, meaning that it accurately detects people who have the disease.
    a. For example, a health facility-based passive surveillance system may not have high sensitivity for identifying some diarrheal diseases because, one, people may not seek care for diarrheal episodes and, two, if the person does seek care there may not be available laboratory tests to diagnose the specific diarrheal pathogen.
6. Also, the system should have adequate **positive predictive value**, meaning that the cases identified actually have the disease.
    a. For example, are malaria cases detected true positives, or were they actually cases of other febrile diseases?
7. Next, the system should have **representativeness**, meaning that it accurately depicts the occurrence of a health events over time and their distribution in the population. As you can imagine, this attribute can be quite challenging to achieve.
8. Next, the system should be **timely**, which means that there are minimal delays between data collection and the availability of information for disease control actions. If the system is being used for research purposes, you can afford a bit of a time lag. However, if it is being used to respond to an acute public health outbreak, like Ebola, the timeliness needs to be as soon as possible.
9. Lastly, the system should be **stable** in that it collects and manages data in a dependable and functional manner. As we continue this presentation on data sources, keep these nine attributes in mind.

These attributes are important, because without high quality data, we will face challenges in evaluating and improving health care delivery, advocating for specific policies, making the case for funding, and enumerating the needs of the populations we are attempting to serve.

Data sources for public health practice
Next, let's take a look at specific sources of surveillance data that can be used in implementation science research and public health practice.



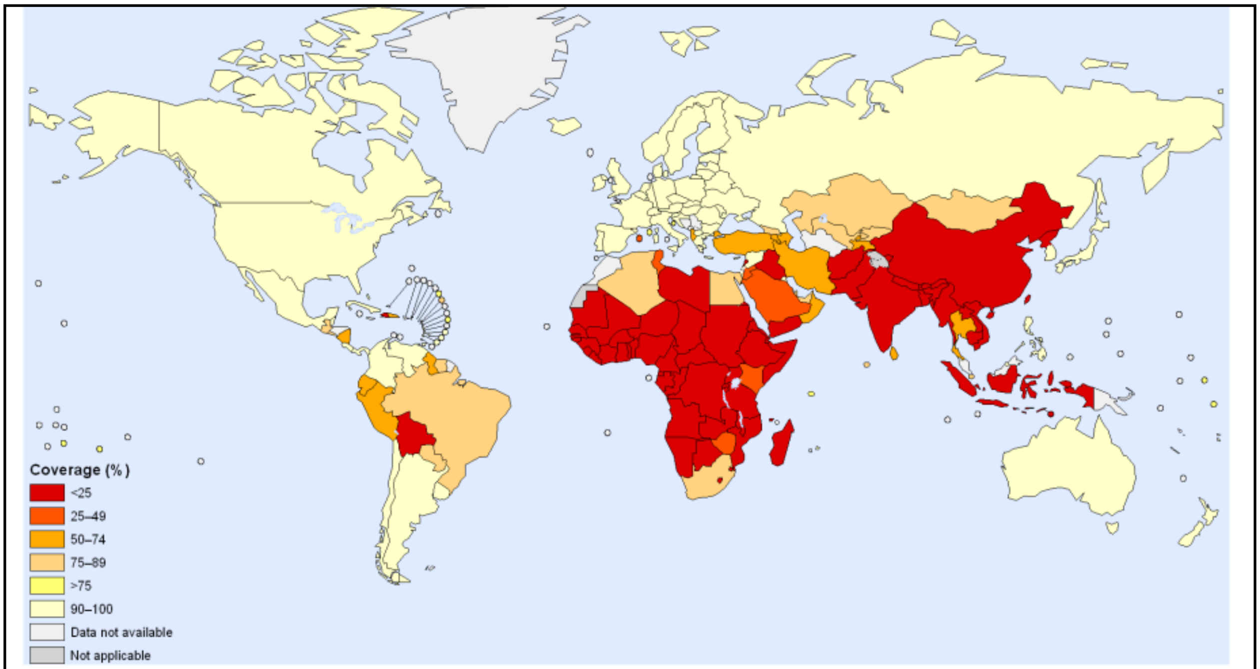
AbouZahr C et al. (2005) "Health information systems: the foundations of public health". Bull World Health Organ.

Depending upon whether strategic or operational decisions are being made and at what level of the healthcare system, different data tools become relevant. For example, OPERATIONAL decisions regarding interventions at the facility level, such as the introduction of a new protocol for prevention of mother to child transmission of HIV, might be best guided by administrative facility records, perhaps in combination with birth and death records. Because tools are not always complete or of sufficient quality, you might need to draw from more than one tool to answer your research question or corroborate your findings.

The data tools described fall into six broad categories of surveillance data systems which are very important to implementation science research and public health practice**.** These data sources may be population-based, such as mortality surveillance in select populations, like children under five. Data sources may also be institution based, such as observed trends in outpatient disease visits for malaria in health facilities. Population based data sources include censuses, civil registration, and population surveys. Institution-based data sources include individual records, service records, and resource records. Depending on the program you are engaging with, you may utilize one or more of these systems in your implementation science work. We'll now briefly review each of these data sources.

The first data source is the **census**. The population and housing census is the primary information source for determining the size of a population and its geographical distribution, plus the social, demographic and economic characteristics of its people. Censuses have been undertaken in most countries in recent decades, and in some places for more than a century. Censuses should be conducted at least every 10 years, and can provide vital statistical data on population and housing situations, births and deaths, and priority pieces of information, like water and sanitation access, at even the smallest administrative levels. So the representativeness of a census is a strength of this data source, so long as there aren't large migratory populations that might be systematically missed by a census. The disadvantage as a health information source, however, is the small number of health questions that can be included. Since it aims to capture everyone in a population, the amount of indicators that can be included is sometimes limited.

The second data source is **civil registration,** also known as vital events. This is the universal recording of occurrence and characteristics of events such as live births, deaths, fetal deaths, causes of deaths, marriages and divorces. This data source is really important because timely knowledge of the size and characteristics of a population is a prerequisite to planning of socioeconomic and health programs. For example, information on the number of live births over time, classified by various characteristics of the mother, constitutes the basis for analyzing reproduction dynamics in a population. Ideally, this is a properly functioning system with high coverage that does not systematically under or over-represent particular populations, or certain causes of death versus others. However, in many settings, particularly low- and lower-middle income countries, civil registration is weak or non-existent and cannot serve as the source of vital statistics. In this map you can see that in countries shaded red, vital registries were estimated to cover less than 10% of all deaths from 2000-2008.



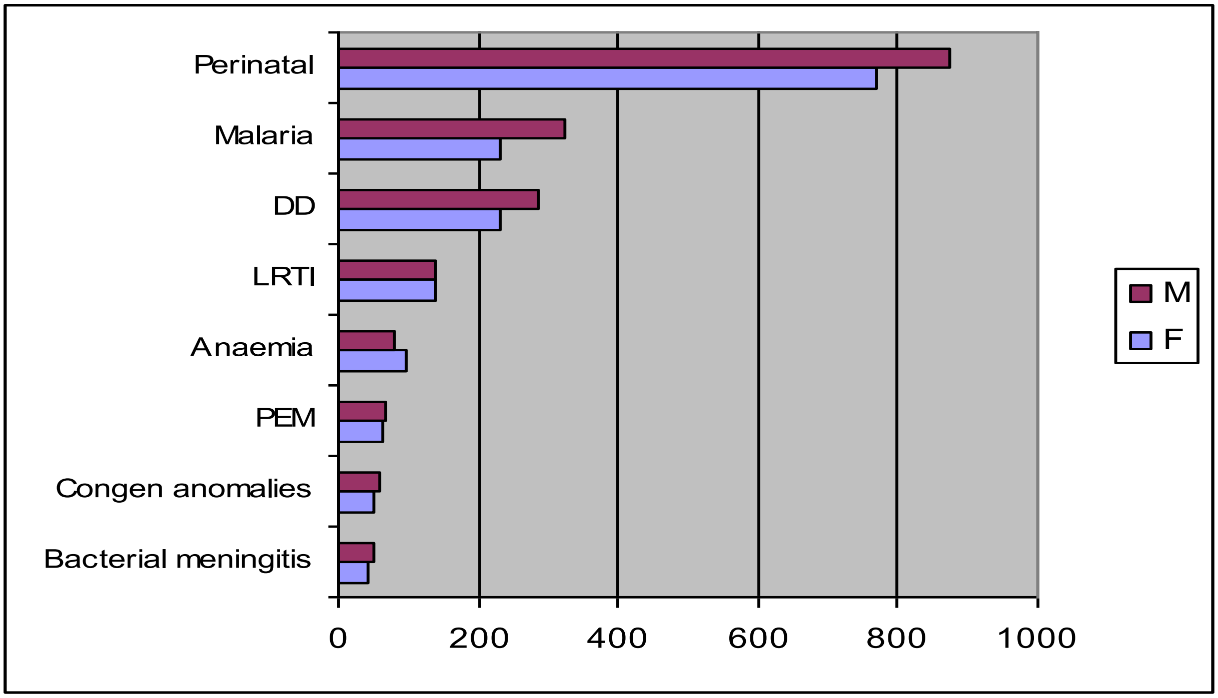
The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: World Health Organization
Map Production: Public Health Information and Geographic Information Systems (GIS)
World Health Organization

World Health Organization

If you look at this graph from Maputo Mozambique, it looks like perinatal issues are the leading cause of death for children from birth to 14 years of age. *Why do you think this might be the case?* A probable reason is that deaths associated with perinatal issues might be most frequently recorded because they are more likely to occur *in the actual health facility*. Thus, we have to be careful when using registries because of under-reporting of mortality events that occur in communities and over reporting of mortality events that occur in facilities.

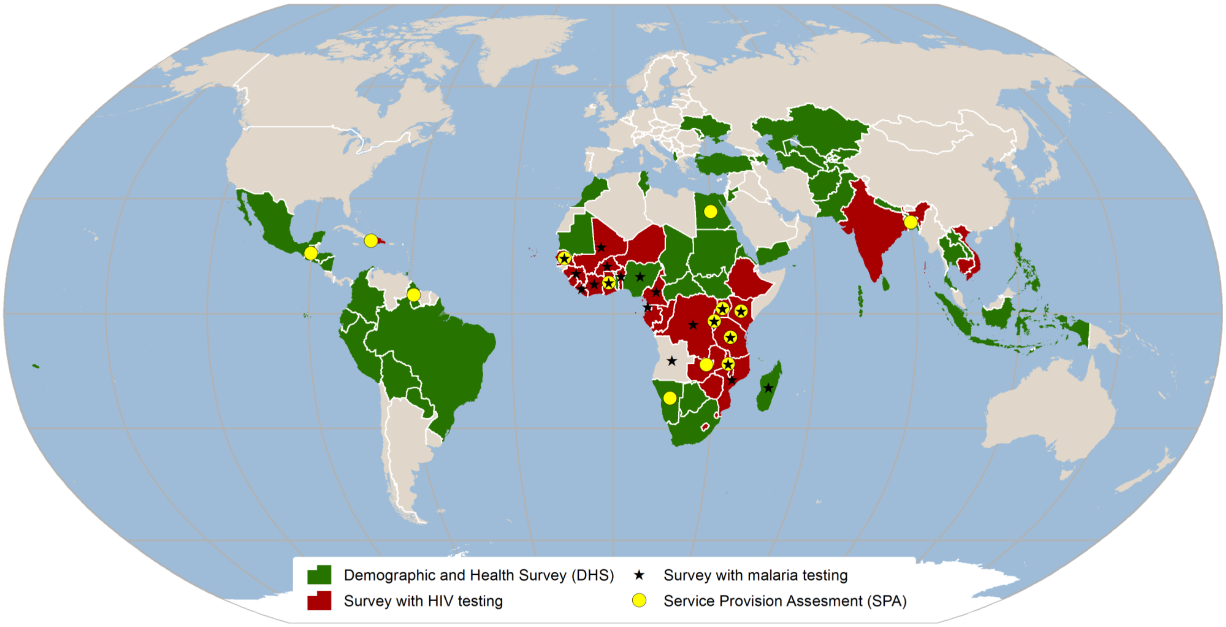
Leading Causes of Registered Death, 0-14 Years *Central Hospital of Maputo, 1994*



Gimbel, S. (2014) . "Surveillance Systems & IS Implementation Science in Health"

In many developing countries, **population surveys** are the single most important source of population health information. Of the 23 health-related millennium development goal indicators, 17 were generated through household surveys, such as the USAID-supported Demographic and Health Surveys (known as the DHS) and the UNICEF-supported Multiple Indicator Cluster Surveys (known as MICS). Household surveys collect information on health measures, knowledge, beliefs, and inequities of populations in certain geographic areas. So, surveys are the prime data sources on risk factors, such as unsafe sex, smoking, substance abuse and poor nutritional status in large populations, and they demonstrate a lot of flexibility as new questions can be easily added over time. Many household surveys have recently added biological and clinical data collection (via health examination surveys) providing much more accurate and reliable data on health outcomes than self-reports. But of course they are less efficient when it comes to measuring relatively rare health events. And, it is important to note that large-scale national household surveys are costly and complex undertakings that are rarely feasible or cost-effective to conduct more than once every 3–5 years. Household survey data are also subject to margins of uncertainty owing to sampling and other errors, illustrating the importance of triangulating data sources, a method we will return to in this presentation. An advantage of population survey data like DHS data is that it is a standardized base survey, which allows a data user to aggregate data at a regional or global level to assess major global health trends.

Standardized DHS surveys have been performed in over 90 countries



Arnold, Fred. Accessed online July 2015. "The state of the art: DHS and MICS"

As you can see in this map, to some degree many of the same indicators are being collected from over 90 countries. At the same time, a limitation is that in some countries it can be very difficult to get enough power to complete analyses below the provincial (or equivalent) level, so the unit of analyses is often at a higher aggregated level than one might want particularly if evaluating a specific health intervention. However, if there is careful preparation, this can be compensated for during survey implementation by oversampling in certain geographies or populations.

Now, let's move on to institution-based data sources including **individual records**, also known as case reports, **and disease registries.** These health records typically consist of case reports produced by health workers or special disease registries (like a Tuberculosis registry). These records can be sensitive where there is high utilization of health services, because disease conditions are typically being recorded at the point of diagnosis. These data often take the forms of paper medical records, electronic medical records, and patient-retained cards. But of course there are significant limitations on the amount of information that can be recorded on client-retained records. Consider, other potential drawbacks of measuring disease trends based on individual records or registers. Estimating HIV prevalence, for example, from a TB registry might over estimate disease prevalence in the population because this is a specific population at *particular* risk of co-infection.

Administrative records are based on **service-generated data** derived from health facilities or other routine settings capturing care offered, treatments administered, and so on. A major strength of service records is their local use for facility management. Additionally these data aren't typically affected by recall bias or social desirability bias because they are based on services delivered, rather than patient-reports and in many settings this can result in timely data. For this reason, some researchers assume that any biases present in the service record data are non-differential over time, so you can still estimate the effect of an intervention at the facility level despite imperfect systems. BUT, service records can be imprecise for extrapolating to the overall population due to the need to estimate denominators (and, remember censuses are conducted only every 10 years). But its worth noting that service records provide opportunities that population surveys do not:

- Administrative service records provide health facility level data which is really very important in implementation science research.
- The data typically also have more data time points available, which allows for more analytical options.
- And the data tend to exhibit attributes such as consistency, availability, and geographic granularity.

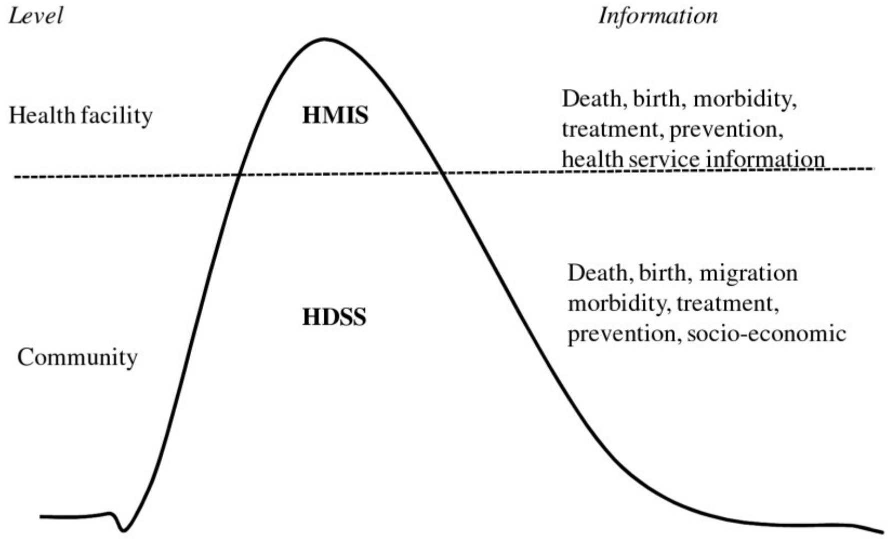
The last data source is **resource records**, which focus on the quality, availability and logistics of health service inputs. Some people may not consider these data surveillance data, but we think they are an important routine information source for implementation scientists and public health practitioners. For example in most countries, National Health Accounts (or NHAs) monitor trends in health spending for all sectors, public and private, and include data on activities, providers, diseases, population groups and regions. NHAs provide information on the amount of financial resources available for health and their flows across the health system. These systems are generally acceptable to governments, who aim establish databases of health facilities and the key services they provide. But more developed resource records are desirable, particularly those that would provide information regarding human resources, budgets and expenditures, core commodities, and key services at national and sub-national levels.

	Determinants of health	Health systems		Health status
		Inputs and outputs	Outcomes (coverage & use)	
Census	✓	✓		✓
Civil registration	✓			✓
Population surveys	✓	✓	✓	✓
Individual records	✓	✓	✓	✓
Service records		✓	✓	✓
Resource records		✓		

WHO (2008). Health Metrics Network: Framework and Standards for Country Health Information Systems. Second Edition

Depending upon the type of information that you need, different data sources will be relevant. For example, information on immunization coverage, a health systems outcome, could be generated from facility service records OR from population surveys. However, different sources may produce different results for a specific indicator of interest. For example, immunization coverage estimates derived from health facility service records would likely be different than those derived from self-reported immunization history recorded in population surveys.
This can be confusing to policymakers trying to use the data to make decisions. So, as implementation scientists, you have to be careful to educate users of the information regarding the relative strengths and weaknesses that might be present, as well as why information from data sources may vary.

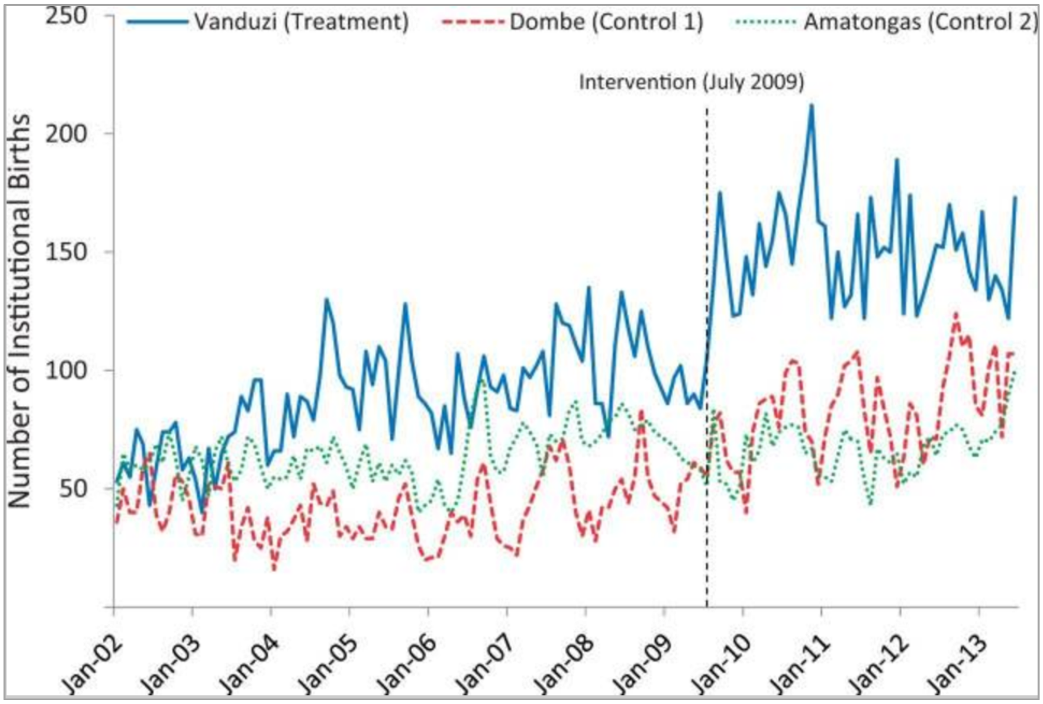
Complementary roles played by Health Management and Information Systems (HMIS) and health and Demographic Surveillance Systems (HDSS)



Ye et. al. (2012). Health and Demographic Surveillance Systems: A Step Towards Full Civil Registration and Vital Statistics System in Sub-Saharan Africa? BMC Public Health.

Choosing your data source also depends on the type of intervention or program you are evaluating. Service records at a health facility level are available on a continuous basis but MAY be subject to bias because they only reflect the populations *using* health facilities and are limited in their relevance to community level factors or events that occur outside of the facility, such as nutritional interventions, safe water interventions, bed net distributions, etc. However, since every time someone comes into a health facility there is a record indicating if they were tested for malaria, treated for malaria, tested for HIV, and so forth, these data can be extremely strong as they represent an overview of activities happening at health facilities. Therefore anytime you want to do an intervention at a health facility level these data can be very helpful. Population surveys are considered to be more representative but they are time-consuming and expensive., as previously discussed As a result they are generally conducted only occasionally, so the data are not continuous and they generate results often spanning several years rather than the immediate past. However if you are conducting a community-wide or community-based intervention, these data may be particularly helpful.

Example of data used in facility-level time series analysis



Wagenaar, Bradley H., et al. "Using routine health information systems for well-designed health evaluations in low- and middle-income countries." Health policy and planning 31.1 (2015): 129-135.

And it is very important to note that the data source you use can determine the study designs available to you (and vice versa, of course). For example, administrative data from the health facility allow you to use quasi experimental designs such as an interrupted time series analysis. Because you have more granular time points than other data sources might allow, you can evaluate facility-level interventions, using one comparable facility as a control for another. In the figure on the right, you can see a time-series of the number of institutional births from 2002 to 2013 for two control clinics and one clinic that underwent an intervention to increase institutional births in 2009, using routine data collected at the facility-level. Study designs such as interrupted time series, in additions to the ones that you will learn about in the impact evaluation module of this course, are really helpful in implementation science because they allow you to test or evaluate facility level interventions while also gaining insight into the strengths of the information system available.

No matter the data source used, existing surveillance systems often have a number of challenges. Challenges include:

- Lack of trained personnel, diagnostic labs, and funds;
- Overly complex systems;
- Delays in reporting urgent events;
- Incomplete reporting;

- A lack of dissemination and feedback of information; and
- A perception that data users did not have input into the collection.

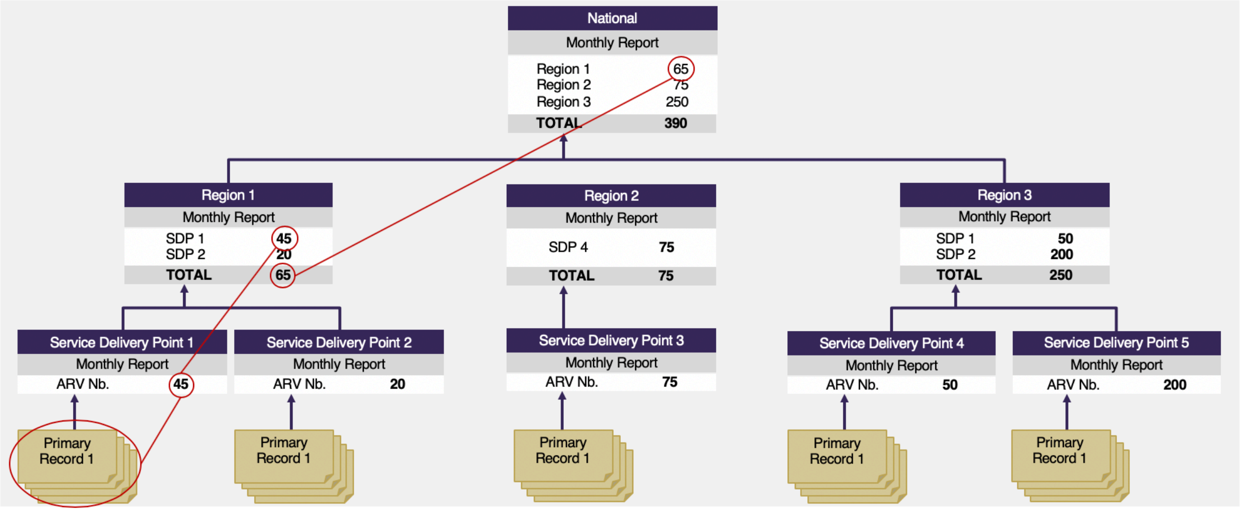
And yet these incomplete systems are often the only data sources available for the continuous, routine monitoring of health programs or as evidence for further public health programing and funding decisions.

Methods to assess the quality of surveillance data

So now we will look at strategies to assess or improve data quality. Data quality auditing (or, DQA) is one approach that can be used to assess the quality of data gathered during surveillance. DQA is an approach to enhance data integrity after collection, and before use.

1. The process uses a Bottom-Up Audit Trail where you first Verify the availability of primary records at service delivery and summary reports in which the data are aggregated.
2. You then assess the accuracy of recorded events in primary records by looking for outliers, impossible numbers of events, or other red flags.
3. Lastly, you re-aggregate data from primary records, and compare with the summary reports across multiple levels.

Let's look at an example.



GFATM. (2008) Routine Data Quality Assessment Tool (RDQA). Manual`. Adapted from: Gimbel, S. (2014). "Surveillance Systems & IS Implementation Science in Health"

Here you can see the multiple levels through which data moves in a health system, from the point of service delivery to aggregated national reports. Data quality problems may and – likely will – occur at any level. In the DQA process you start with primary records and verify if the first level of aggregation matches primary data sources, and then you move up the health system comparing aggregated numbers. Here it is reassuring that the aggregation cascade appears normal. A DQA procedure can be used on data from any number of healthcare levels depending upon the research question of interest: It can also be conducted on a random sample of facilities, patient records, geographies, etc. Primary data sources that could be used in DQA include:

- Number of people reached with services, as indicated by medical records.
- Number of commodities distributed as indicated by distribution log sheets or inventory statements.
- Or number of people trained as indicated by attendance sheets or per diem sign-up sheets.

Number of months when facility registers matched monthly facility reports in Sofala Mozambique (06/01/08-12/31/08), by indicator

Facility Type	**1st ANC N (%)**	**Institutional Birth N (%)**	**DPT3 N (%)**	**HIV Testing N (%)**	**Outpatient Consults N (%)**	**TOTAL N (%)**	**Global Fund Grade**
Urban	1/6 (17)	2/6 (33)	4/6 (67)	0/6 (0)	4/6 (67)	11/30 (37)	**B2**
Peri-Urban	5/6 (83)	6/6 (100)	6/6 (100)	NA	6/6 (100)	23/24 (96)	**A**
Rural	6/6 (100)	6/6 (100)	6/6 (100)	NA	6/6 (100)	24/24 (100)	**A**

Global Fund DQA Rubric	A	Less thank 10% error margin
	B1	Between 10-20% error margin
	B2	Above 20% error margin
	C	No system in place

Gimbel, et al. An assessment of routine primary are health information system data quality in Sofala province, Mozambique. Population Health Metrics. 2011;9:12.

As an additional example, researchers in Sofala Province, Mozambique conducted a DQA in which they compared the number of months when facility registers matched monthly facility reports. They looked at a number of indicators, including first antenatal care visit, number of institutional births, diptheria pertussis and tetanus vaccines delivered, HIV testing, and outpatient consults. And across 3 districts they found a median percent difference of 4%, with 86% of the months differing by less than 10%, which is very impressive. The Global Fund to Fight AIDS, Tuberculosis and Malaria actually has a rubric for data quality audits, which these researchers applied, as you can see here, and which could be used in any DQA initiative.

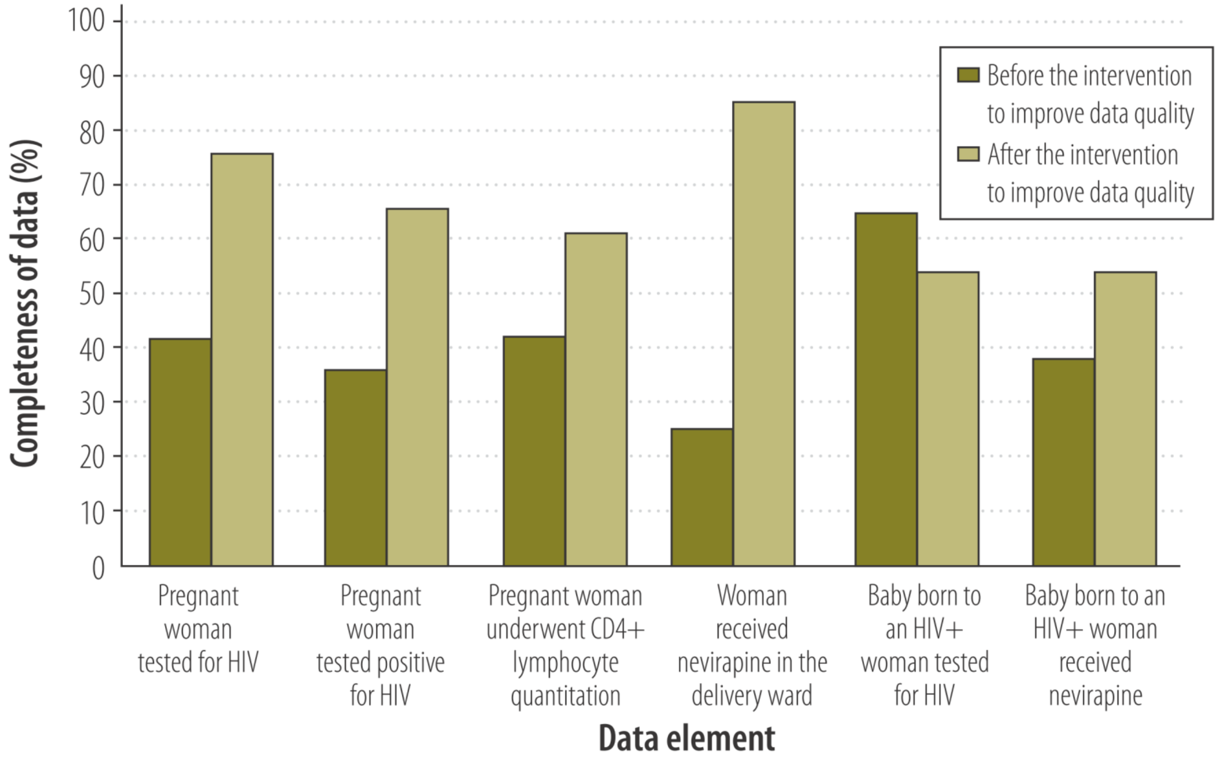
However very often the data don't align as well as that described in the previous example. In another example of a research study conducted in South Africa, researchers aimed to improve quality and completeness of data used in PMTCT programs. The intervention included a series of 3, one day trainings on data management, routine data review at the district level, and DQA at individual facilities.

DQA methods have also been used to measure intervention outcomes

78 health facilities (of 222) randomly selected to assess results of a data quality intervention for prevention of mother to child transmission (PMTCT) of HIV in South Africa

- Series of three, 1-day trainings on data management
- Routine data review at the district level
- DQA at individual facilities

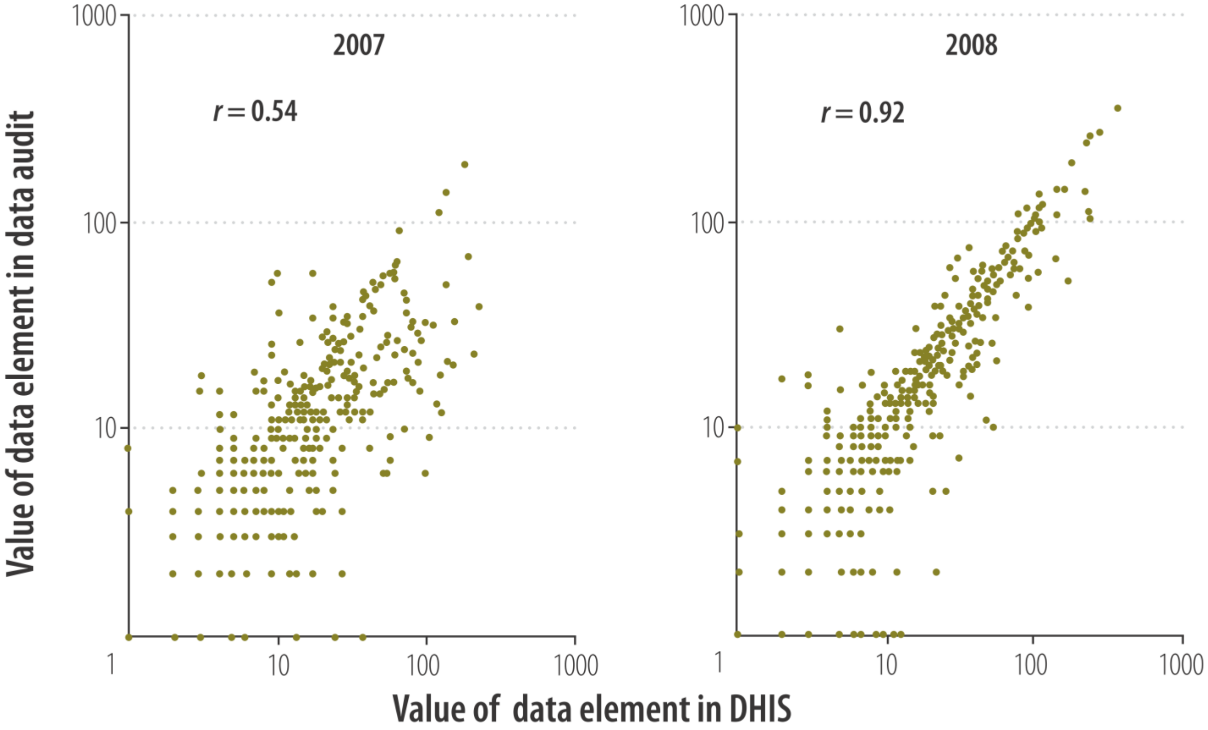
Completeness of data increased from 26% to 64%



Mphatswe, et al. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. Bull WHO. 2012;90:176-182

The researchers found that engaging health workers in a training on DQA vastly improved data quality post-intervention with the completeness of data increasing from 26% to 64%.

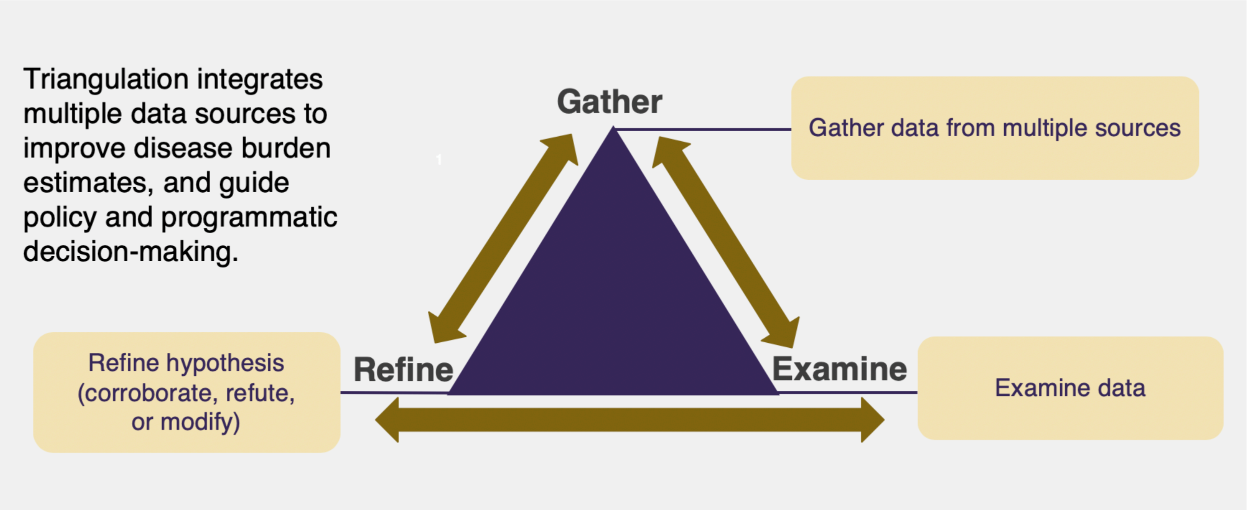
Accuracy between patient registries and aggregated reports: correlation coefficients increases from 0.54 to 0.92



DHIS: District Health Information System

Mphatswe, et al. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. Bull WHO. 2012;90:176-182

They also compared the data in the registries to aggregated reports and found that the correlation between the two increased from 0.54 pre-intervention to 0.92 post-intervention. So this simple, implementation science intervention vastly improved data completeness and accuracy in the short-term, and provides the opportunity for a number of *other* intervention outcomes to be assessed with greater confidence.



Triangulation integrates multiple data sources to improve disease burden estimates, and guide policy and programmatic decision-making.

WHO. HIV triangulation resource guide: Synthesis of results from multiple data sources for evaluation and decision-making. February 28, 2009

Another method for assessing data quality in surveillance systems is **triangulation**. Triangulation integrates multiple data sources to improve disease burden estimates and guide policy and programmatic decision-making. Triangulation is an iterative process in which data are gathered from multiple sources and examined, and hypotheses are continually generated, debunked, or corroborated. Say for example, you are interested in the number of malaria cases in children under 5 within a certain region. You might use aggregated service record data but then triangulate by comparing those estimates to those from the DHS or other population surveys. The key is to observe what discrepancies or similarities exist and then propose hypotheses for reasons why those discrepancies occur. However, you have to use your knowledge of how data are generated in different data sources to determine if data from each source are comparable.

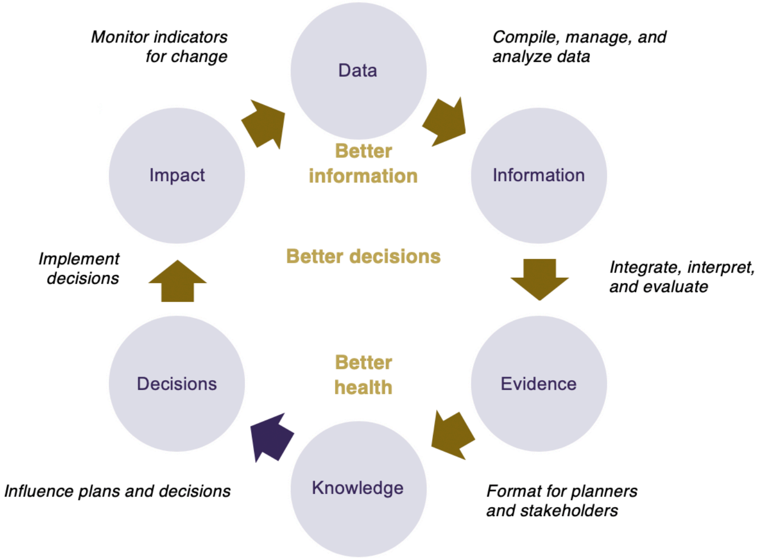
Evaluation of surveillance systems

Lastly, we'd like to point out that not only should you assess the quality of the data being produced within your surveillance system, you should periodically be assessing the performance of the entire surveillance system at large.

- In evaluating the surveillance system, one should consider: is the health event under surveillance of continuing public health importance?
- Is the surveillance system useful and cost-effective?
- Are the attributes of a high-quality surveillance system, that were discussed in this presentation, adequately addressed?
    - Sensitivity, specificity, acceptability, timeliness, simplicity, flexibility, and representative
- Are the informatics used appropriate and effective?
    - Hardware/software, interface, data format/coding, and quality checks
- And, are surveillance system data routinely disseminated to key stakeholders?

Main Messages

In conclusion, let's discuss a few takeaway message from this presentation. We routinely collect data so that we can use it to define research, to inform policy, to prevent emergencies, and to better meet the needs of the populations that the system is tasked to protect. In other words, the purpose of surveillance is to produce data that guides ACTION. Approaches that collect data regularly for multiple purposes are particularly helpful in implementation science research (for example, routine service delivery data). You should use the data system most appropriate for your research question at hand, and also select a quality assurance strategy to understand strengths and weaknesses of that data system. If there are concerns about data quality, there are available strategies to adjust or improve data quality, which is helpful for long-term health systems strengthening. And, lastly, data systems should be periodically evaluated and results should be disseminated to all relevant stakeholders including community members, practitioners, and policy makers.



WHO (2008). Health Metrics Network: Framework and Standards for Country Health Information Systems. Second Edition

In conclusion, data generated by the different available surveillance systems discussed today can be used to transform routine data into information and evidence for action. For implementation science and public health practice. It can't be emphasized enough that although reliable information is an invaluable tool for improving health, its availability alone does not guarantee that it will be used for improved decision-making. The data must be used in meaningful ways to understand disease trends and the functionality of the health system, measure the effects of targeted interventions, and to inform policy and priorities. Strategies to use and evaluate data for these purposes will be addressed in several modules, throughout this course. Thank you for listening to this presentation on data sources and quality for implementation science.

Acknowledgements: Arianna Rubin Means