

Workshop
Summary
Series

**Biological and Clinical
Data Collection in
Population Surveys in
Less Developed
Countries**

**National Academy of
Sciences**

January 24-25, 2000

Washington, D.C.



MEASURE
Evaluation

Biological and Clinical Data Collection in Population Surveys in Less Developed Countries

**Summary of a meeting held by
MEASURE Evaluation**

**National Academy of Sciences
Washington, D.C.
January 24-25, 2000**



**MEASURE Evaluation
March 2000**



Printed on recycled paper

Introduction

For three decades, developing countries have attempted to assess at least some of the characteristics of their populations' health through censuses and large scale household surveys that provided information on issues including fertility, health service use, nutrition and household expenditure. Many of these are undertaken regularly by the statistical offices of national governments. Others are carried out with international support, using relatively standardized questionnaires that provide information useful for comparing populations over time and between countries. The best known of these international programmes are the Demographic and Health Surveys (DHS) and their predecessors, the Contraceptive Prevalence Surveys and the World Fertility Surveys – programmes actively supported by the United States Agency for International Development (USAID).

As their names suggest, these surveys began with a very strong focus on fertility and contraceptive use. In recent years, they have broadened their scope to include other health and welfare issues including school attendance, sexual partnerships and HIV, and maternal and child nutrition. The addition of nutritional status to the surveys meant the inclusion of anthropometric measurement: the first time programmes had gone beyond simple survey questionnaires. Most recently, the DHS programme has added anemia testing to its surveys in some countries. Anemia tests, which require a finger prick to obtain capillary blood followed by an on-site rapid test, has added a new dimension to the survey programme – the collection of biological specimens and the measurement of biological markers of health status (biomarkers).

The technology continues to race forward. Many diagnostic tests, which until recently required complex equipment, excellent infrastructure and highly trained personnel, can now be carried out in an urban slum or a desert settlement by field staff with a minimum of training. Costs of existing tests are falling, and “field-friendly” tests for many more conditions are expected to become available in the next few years. The availability of these new technologies begs a question: how might they best be used to improve the health of people in countries where the conditions they identify are most prevalent?

Challenges

Some of the challenges of biological and clinical data collection in population-based surveys in less-developed countries are:

- What is the right public health question?
- What is the appropriate population to collect data from?
- Will policy decisions flow from the measurement of disease or markers for sub-clinical disease?
- Should surveys provide treatment or counseling to respondents?
- What are the ethical dimensions related to anonymous and unlinked data collection?
- How do we assure the consent process is as informed as possible?
- What criteria should be used to identify the most appropriate test?
- How precise should the prevalence estimate be?

Excerpted from a presentation and paper by Elizabeth Holt, 'The challenges of biological and clinical data collection in large scale population-based surveys in less-developed countries'.

There is a general belief that biological specimen collection – partly because it is more intrusive than just asking questions – crosses an invisible line into a territory that carries added responsibilities, as well as added risks. Certainly the collection of blood has added logistical and ethical complexity to DHS surveys. These complexities are likely to increase if the range of biomarkers is expanded to include markers of malaria, STIs, HIV or other conditions.

This document attempts to outline some of those complexities. It summarizes discussions on the subject held at the National Academy of Sciences in Washington, D.C., in January 2000. The meeting was organized by MEASURE Evaluation on behalf of a coordinating committee that included representatives of MEASURE DHS+, MEASURE CDC/Division of Reproductive Health, USAID, Johns Hopkins University and MEASURE Evaluation. It was sponsored by the USAID Office of Health and Nutrition.

The meeting was called in part to review what might be **possible** in the field of biological testing of specimens collected in general populations. Recognizing, however, that data collection may not be a valuable end in its own right, the more important task was to open a

discussion on what might be **useful**, and what might be **desirable**. This report focuses on issues raised during the meeting.

Need for population-based data in public health

In many countries a number of disease surveillance activities are ongoing to monitor the prevalence and incidence of disease. These include clinic-based surveillance, large national sample surveys with limited or no biological and clinical data collection (such as DHS), and special studies on specific interventions and diseases.

Inevitably, there are major gaps in knowledge about the distribution of health and health services within the country. Population-based surveys, including biological and clinical data collection, have at least one major advantage over all other methods – measurement of levels and trends in health inequalities (provided that a number of essential background characteristics can be linked to the biological and clinical test results). Data on inequality in health are needed for sound public health planning and implementation. Data are needed by major geographic region within the country, by urban and rural areas, by ethnic group or race, by level of education or occupation and so on.

Excerpted from conference presentations by Lindiwe Makubalo, 'Country perspectives on the need for biological and clinical data for health programs' (South Africa), Bernhard Schwartlander, 'The need for population-based data on HIV' and Henri Damisoni 'Current practices in surveillance in Malawi'.

What can biomarkers tell us?

Before charging headlong into the collection of specimens for biological testing, it is useful to clarify what the purposes of biomarker testing might be. In addition to the clinical purpose of diagnoses of ill health in individuals, the information generated by such testing can be used for three purposes:

- The assessment of needs and the planning of interventions to improve health
- The monitoring of changes in health and the evaluation of interventions
- Lobbying for changes in policies to address population health needs

An example of technological advances: Oral mucosal transudate

Oral mucosal transudate, a serum fluid that enters saliva from the gingival crevice and across oral mucosal surfaces, can be preferentially concentrated by a collecting system (such as OraSure) to yield detectable levels of immunoglobulins (e.g., IgG and IgM antibodies) against various bacterial and viral diseases. Assays based on oral specimen can be used for the diagnosis of diseases (antibody-based diagnosis) – HIV1/2, hepatitis A/B/C, *Helicobacter pylori*, measles, mumps, rubella, syphilis and CMV – chronic diseases (autoimmune disorders such as Sjogren's syndrome, rheumatoid arthritis, myasthenia gravis), some forms of colon, prostate and ovarian cancer, diabetes type 1 and 2, therapeutic drug monitoring (theophylline, phenytoine, digoxin, etc.), detection of other drugs (alcohol, cocaine, amphetamines, etc.).

Source: George RJ, Fitchen JH. Future applications of oral fluid specimen technology. Am J Med 1997; 102 (suppl 4A): 21-25.

Biomarkers as a planning tool

Resources available for improving health are limited in most countries. This means that choices have to be made about where to invest human energy and funds. The choices would be difficult enough for a government working with full information about the health needs of its population. It is harder still when little is known about who is suffering from what.

Much health planning currently relies on information about sickness and death reported through the regular health system. This includes information from hospitals and health centers, and from disease surveillance systems such as those that track HIV infection in pregnant women. This is supplemented by data generated by research studies, and by information reported by respondents (mainly women) in regular household surveys such as DHS.

One of the limitations of existing systems is that they tend to be very incomplete. Regular disease reporting systems are fragmented at best, chaotic at worst. Survey data rely mostly on self-reports. Studies that include clinical examination tend to suffer from wide variations in accuracy of diagnosis, and can in any case only detect symptomatic disease. This reduces the utility of these data for planning prevention.

Micronutrient deficiency, for example, is something that may be rectified relatively simply while it is still at a sub-clinical level. It is currently estimated that up to 250 million children suffer from vitamin A deficiency, and perhaps two billion people from iron deficiency. If countries had a better understanding of the level and distribution of these deficiencies, they would have the possibility of acting to prevent the higher morbidity and mortality associated with the conditions and to increase the intellectual and physical development capacity of those affected.

Testing of biological specimens can also identify a population's susceptibility to infectious disease. This information can act as an "early warning system," allowing health authorities to plan "catch up" immunization campaigns to avoid outbreaks of measles, rubella and similar diseases.

One of the main purposes of nationally representative surveys is to identify differences in health status and health services provision and utilization within the country by socio-economic, geographic and

demographic variables. Biomarkers are likely to make a major contribution to better documentation of inequitable distribution of health and health services within and between countries.

On the downside, there is a possibility that biomarkers will tend to overestimate the health threats in a population, precisely because they can often detect sub-clinical or asymptomatic infection. For many conditions, including some STIs, the morbidity and infectivity associated with asymptomatic infection is unknown.

Anemia testing in DHS surveys

Anemia testing is now commonly included in DHS surveys. Using Tenderlett devices, capillary blood is taken from the finger (or heel for young children) of the respondent (women and children in most surveys) by a health worker who is part of the survey team. Hemoglobin is measured with the Hemocue system which detects the level of hemoglobin within a minute. The Hemocue consists of a battery-operated portable photometer and a disposable cuvette. To assess the main cause of anemia in the survey population a small sub-sample is tested for the level of serum ferritine.

Thorough training is given to reduce individual variability as much as possible. Prior to the test, the woman is asked to give written consent on a form that explains the procedure and purpose of the test (to determine the rate of anemia among women and children) and the confidentiality of the results. If anemia is identified, the respondent or her child is referred to the nearest clinic. If anemia is severe, the respondent is asked for consent to allow the survey team to inform a local doctor about the test result. Non-response has not been an issue in any of the surveys. On the contrary, the respondents were keen on knowing whether or not they had anemia. DHS has had similar experience in taking the anthropometric measures.

The surveys in Central Asian countries showed high levels of anemia among children and women and also important differentials in the prevalence of anemia by region, level of education and ethnicity. The survey results led to the formulation and implementation of a regional integrated programme by UNICEF to reduce the prevalence of anemia.

DHS has published a manual for anemia testing in population-based surveys. It lays out a standardized approach for hemoglobin testing using the HemoCue system and pays particular attention to the biohazardous waste disposal and safety precautions when taking blood.

Excerpted from presentations by Martin Vaessen, 'Issues in collecting biological and clinical data in population-based surveys in developing countries' and Almaz Sharmanov, 'Experience with anemia testing in population-based surveys'. See also Sharmanov Almaz, 2000. Anemia testing manual for population-based surveys. Calverton, Maryland, USA: Macro International Inc.

Biomarkers as an evaluation tool

Precisely because of the resource constraints described above, it is important for governments (and the taxpayers and international partners that support their efforts) to be able to demonstrate that they are making progress towards improving health and well-being. It is not enough to have interventions designed to reduce malaria or improve nutritional status – those interventions have to be shown to make a difference and to be cost-effective.

Evaluation of interventions is usually among the weakest points of a health system. This is partly (though by no means exclusively) because it is difficult to know whether interventions are making a difference unless there is a clear picture of the prevalence of a condition both before and after the intervention. Biomarker data give potentially the most accurate assessment of prevalence levels for many conditions, and are therefore useful for evaluation purposes. Single round cross-sectional surveys may also be able to indicate whether national or international targets have been met by measuring, for example, the proportion of children with immunity to diseases covered by the immunization programme.

There are drawbacks, however. Evaluation depends on measuring changes in prevalence (or, more rarely, incidence) over time, so if cross-sectional surveys are to be used, the measure must be comparable over time. What's more, the sample size must be large enough to detect statistically significant changes between one survey and the next.

The real power of biological and clinical data collection for the purpose of evaluation lies in combining such data with interview data. The latter can be a good measure of exposure to the interventions, which may range from receiving vitamin A supplements to condom use.

If collected in general population surveys, biomarkers will be able to give a general idea of the impact on health status of the national response to a health problem. A much more sophisticated study design would be necessary to attribute particular changes to a particular intervention, net of external influences such as environmental, economic or security conditions.

Lobbying for a new health agenda

Data provided by tests on biological specimens are often more persuasive than interview data, which often rely on self-reports and are sometimes considered less “scientific.” This makes them particularly useful for lobbying for more attention to a neglected health problem. Lobbying (or “advocacy”) is similar to planning in as much as it argues for the prioritization of resources according to the gravity of a health problem. It differs, however, in that it presupposes resistance – usually political - to tackling the problem at hand. Resistance may exist because the health problem or its potential solutions are unpalatable to powerful groups such as religious leaders – HIV prevention including condom promotion is an example. Or it may exist because of a reluctance to recognize inequities in a country – differences in health status between people of different ethnic backgrounds or regions with different political loyalties, for instance.

Lobbying can happen at international as well as national levels. Indeed, data collected in international survey programmes such as DHS are often used for within and cross-country comparisons and contribute to building an international agenda in reproductive health and other issues. Ultimately this can help to change conditions in a country not immediately amenable to that agenda.

Vaccine-preventable diseases

There is a renewed interest in sero-epidemiological surveys in the context of immunization programmes. Serological techniques to assess immune status have become better and easier to apply. Specific public health and epidemiological issues require data at the population level. These issues include timing of mass vaccination to eliminate measles, introduction of rubella vaccine into immunization programmes, assessment of tetanus immunity in a population as a tool to control neonatal tetanus, assessment of diphtheria immunity, and disease burden studies for hepatitis A and human papilloma virus. Most antibody tests are based on serum samples, but oral fluids samples have also been used for determination of measles, mumps and rubella immunization status.

Rubella is still an important public health problem in many countries, and congenital rubella syndrome is more common than often thought. Sero-epidemiological studies (using serum antibody detection, although urine can also be used) have shown that in some countries as many as 30 percent of women of childbearing age are not immune. In such situations it may be cost-effective to introduce rubella vaccine as a companion to measles vaccine.

Major epidemics of diphtheria occurred in Eastern Europe and Russia during the early 1990's. Serological surveys showed major deficiencies in immunity in adult populations, partly due to waning immunity in adults and a faltering immunization programme.

Prevention of neonatal tetanus can be achieved through immunization of the mother during pregnancy with a cheap and highly effective toxoid. Four or five doses provide lifetime immunity. However, the actual immunity status of women is difficult to evaluate, as recording is often poor and incomplete or recall is flawed. In the Central African Republic capillary blood was collected by finger-stick blood sample on filter paper from one-fourth of the mothers in a UNICEF Multiple Indicator Cluster Survey (MICS). Tetanus antitoxin seroprevalence was found to be higher than estimated based on recalled doses of tetanus toxoid vaccine, although recall data were found to be fairly good (88% had antitoxin seroprevalence, 76% by vaccination recall). Presently, reliable tests are possible on filter paper aliquots, even at low titers of antibodies.

Excerpted from conference presentations by Marc La Force 'Immunisation: State-of-the-art' and Michael Deming 'Experience with evaluating a tetanus toxoid immunisation programme'.

A note on testing technology

As noted, this document cannot do justice to the range of information presented about existing testing technologies and those under development.¹ However since the current state of technology has important implications for logistics, a few broad observations might be made.

Clearly, testing technology is moving forward by leaps and bounds. For many markers of health status, including anemia, vitamin A deficiency and HIV, tests using dried blood spots on filter paper, obtained through a finger prick, have replaced serum which requires a venapuncture. Other tests, including HIV, can easily be performed on saliva or urine. Rapid tests for malaria that require no refrigeration and little technical training are becoming available to replace microscopy. These tests are particularly valuable for population surveys aiming to establish prevalence, since they register a positive result for up to two weeks after a person has cleared the parasite from their blood.

There are many trade-offs that must be considered in choosing an appropriate technology. Choices will be influenced by the purpose of the test, by decisions made about informing and treating participants, and by resources available.

For example, if a decision has been made to give on-the-spot results to those providing samples, then it is important that both sensitivity and specificity of the test are adequate under field conditions. These conditions may include use of equipment that has spent hours bumping along a dusty road in an open jeep, extremes of temperature, and performance of the test by someone with limited training. If, on the other hand, specimens will be shipped to a lab for unlinked anonymous testing, then specimen transportation and possibly cold storage requirements may take priority among selection criteria.

The time taken to process a test can also be an important consideration, especially if tests are being performed in the field. DHS recently studied the feasibility of using a number of biomarkers tests in the field. Lead testing was considered a failure because under

¹ An annotated bibliography on biological and clinical tests that can be used in surveys in developing countries was prepared by MEASURE Evaluation and may be obtained from the project.

field conditions the equipment performed poorly, making the tests more time-consuming than was acceptable to a survey team.

The choice of test may depend on the choice of specimen. Non-invasive specimens may be chosen in order to minimize refusal bias in a linked part of a survey or study, or tests may be chosen to get the maximum information out of a single specimen type. Storage of samples may allow tests to be used on previously collected specimens as they become available or affordable.

One of the drawbacks of this constantly improving technology is that trends over time may be difficult to interpret. Does higher population prevalence of a specific micronutrient deficiency signal a deterioration of nutritional status between surveys, or does it just mean that the test used on the second survey was more sensitive than the test used on the first? If biomarkers are to be measured to evaluate the impact of interventions, care must be taken to choose tests (as well as reference periods, cut-off points for definition of ill-health etc.) that are comparable over time.

Malaria

Malaria is a major public health problem in many countries. In general, population-based surveys have not been very successful in measuring the burden of malaria or in the evaluation of interventions to reduce malaria morbidity or mortality, using the clinical diagnosis. Recent technological progress however may make possible an easier and more accurate diagnosis under field conditions in surveys, using rapid diagnostic tests. The diagnosis of malaria can be made under field conditions using several approaches.

- Clinical diagnosis: reporting of recent fevers is often highly sensitive (those who have malaria will report fevers), but specificity is low (more than 50% diagnosed with malaria will not have malaria).
- Microscopic diagnosis: this is the current 'gold standard' based on a drop of capillary blood obtained by finger stick. The diagnosis is highly specific and sensitive, but requires trained staff, microscope, supplies and time (about 1 hour per test). The results are highly operator dependent.
- Rapid diagnostic tests: these are based on dipsticks or cards coated with monoclonal antibody. The tests are highly sensitive and specific for high-density infections, but less so for low-density infections. A threshold of 50-100 parasites per microliter is required for the tests to turn positive. Most tests are able to detect *Plasmodium falciparum* infection, but not other types of malaria. Costs are on average \$1.00 (range \$0.60-4.00). Examples of such diagnostic tests are Histidine Rich Protein (HRP-2) and parasite lactate dehydrogenase (pLDH - also for *Plasmodium vivax*). The tests can be performed by health workers with little or no training.
- Other diagnostic methods: including QBC (microscopy based, expensive) and nucleic acid amplification (PCR - most sensitive, but highly skilled staff required and expensive) are less suitable for large-scale surveys.

Excerpted from a conference presentation by Lawrence Barat 'Malaria: what can be done to better measure the burden of the disease and to evaluate interventions'.

Adding biomarkers to population-based surveys: is it possible?

The limited experience so far of adding biological specimen collection for anemia testing to DHS and other surveys in developing countries has raised some interesting issues surrounding feasibility. The possibility of adding different tests raises more issues, as experience with population-based research studies has shown.

Logistics

Even when they are limited to getting staff with a pen, a questionnaire and a clipboard into the homes of a representative sample of the population, surveys are complex logistical operations. The addition of anthropometric measurements adds to transport, staffing and training costs, and equipment failure is not uncommon. In fact, some countries have decided to exclude these measurements from population-based surveys for logistical reasons.

Adding biomarkers increases logistical complexity still further. In most cases, additional staff will be included in the survey team, adding again to recruitment, training and transport costs, regardless of the test type used. For more complex tests, cold chains may be needed, and specimen transportation becomes a major issue. The range of the survey team may in some cases be restricted by transport time from a reliable cold store. With the progress in technology, however, it is increasingly possible to avoid cold chain requirements.

Universal precautions

It is obviously critical that a survey does not become a vector of disease, but maintaining universal precautions has proven difficult. Gloves and other materials needed to ensure the safety of respondents and staff, and the biowaste disposal equipment needed to protect communities, pile up yet again on costs and organization. Existing guidelines on universal precautions are extremely thorough, and compliance may increase the financial and logistical burden on the survey (and the survey staff) unreasonably, especially in countries where prevalence of HIV and hepatitis B are low. It is suggested that minimum guidelines be developed for these countries.

Universal precautions in the collection of blood and disposal of biohazardous waste

A set of precautions are required to prevent transmission of HIV, hepatitis B virus and other blood-borne pathogens when providing first aid or health care. Such precautions also apply to biological data collection in population-based surveys (see, for example, DHS manual on anemia testing for population-based surveys) and should minimize the risk for the survey team, study participants and community against exposure to blood or blood-contaminated materials.

To protect the participants:

- Single sterile needles are used
- Worksites are clean, implying prevention of exposure to contaminated materials and use of appropriate disinfectants
- Access to worksite is restricted.

To protect the community:

- Adequate infectious waste management (sharps, contaminated materials)
- Treatment of instruments and waste (autoclaving and incinerating)
- Safe transport of specimens

To protect the survey team:

- Application of universal precautions in laboratory environment, home settings and field conditions by using safer medical devices (shielded needle or needle-less devices), plastic capillary tubes; personal protective equipment (gloves, gowns, aprons, etc.); use of sharps containers; periodic reviews (practices quality control)

Excerpted from conference presentation by Jonathan Richmond 'Responsibility to protect the participants, community, and survey team using universal precautions' . Further reading: Richmond JY, Mc Kinney RW (editors). Biosafety in microbiological and biomedical laboratories. US Department of Health and Human Services, Public Health Service. Centers for Disease Control and Prevention and National Institutes of Health. Fourth edition. May 1999; Sharmanov Almaz. 2000. Anemia testing manual for population-based surveys. Calverton, Maryland, USA: Macro International Inc.

Human and financial resources

Obviously transport, universal precautions, training and other factors associated with the addition of biomarkers to population-based surveys carry costs. Add to that the costs of performing the tests themselves. These vary greatly – as little as 50 cents for some tests, as much as 22 dollars for others. The overall cost will depend enormously on sample size, and will also be influenced by where the processing is being done. Establishing a local laboratory where none exists carries high capital and training costs but may contribute to the capacity of a country to carry out other types of research. Shipping samples to overseas labs is also costly, however, and may represent a biological hazard.

Human resource needs will also depend on the level of training required to carry out a test. In tests (or indeed in clinical diagnoses) where the assessor variability is typically high, surveys will only succeed if they have highly qualified staff or very active supervision. Human resource constraints are often underestimated; the problem is likely to grow more acute as health systems restructure and opportunities in the private research and health sectors grow. Where human resources are extremely scarce, it may be detrimental to a country's overall health system to pull too many people out of service delivery for the purposes of surveys or research.

Quality control

While some of the issues raised above may seem petty, they are critical to the success of any survey that includes biomarkers. Unless samples can be collected correctly, handled appropriately, transported at the right temperature and in the right time frame and tested carefully by people with the right equipment and training, the data generated are likely to be worthless. In the DHS pilot study mentioned above, 12 drops of capillary blood were collected for vitamin A testing. Several women had difficulty generating the needed volume of blood, and difficulties were encountered with the cold chain. When the tests were performed, just nine percent of women were found to suffer from vitamin A deficiency. The study team felt this was unreasonably low given the characteristics of the population, and the data were not published.

Where testing and laboratory conditions cannot be controlled, clinical diagnosis may actually give better results, since field conditions may

greatly alter sensitivity and specificity of a test. In the words of one presenter: “There are some pretty good tests, but some very bad lab conditions”. In a field laboratory in one area where malaria is highly endemic, the positive predictive value of “gold standard” microscopy sank to 43 percent, while the negative predictive value was just 33 percent. These results are not substantially different from those achieved by clinical diagnosis. For survey purposes, the latter, which eliminates most of the logistical hurdles cited above, may be preferable.

Sampling

A great deal of time and effort goes into sample selection in most population-based surveys. Sample size calculations are made on the basis of representativeness, as well as on the expected prevalence of key variables in the survey data. Surveys intended to track trends over time need larger sample sizes than a cross-sectional survey to establish prevalence levels.

The addition of biomarkers to a population-based survey may well have an impact on sample sizes, and this may in turn have cost and other implications. For conditions that are highly prevalent (more than 5-10%), this is unlikely to pose a problem. For conditions with low prevalence, household surveys may not have adequate sample sizes, because even if a test is highly specific (having few false negatives), the positive predictive value of the test at very low prevalences will be low.

A further issue is the demographic make-up of the survey population. Most population and health surveys have traditionally concentrated on women and children. However, biomarkers of interest may require entirely different sample populations. For example, levels of HIV prevalence among pregnant women are well known in many countries through antenatal clinic-based surveillance, but little or nothing is known about prevalence of the virus in men. If a country were to go to the trouble and expense of including HIV testing in a population-based survey, surely it would make sense to sample men in significant numbers. And yet this may well have wider implications for the survey, because, for example, men are much more frequently away from home, and therefore, harder to include in household surveys.

Micronutrient deficiencies

Micronutrient deficiencies are often hidden problems in the population, with only the most severe forms showing symptoms. Research during the past decade however has shown that mild to moderate deficiency may lead to increased risks of morbidity and mortality. Assessment of micronutrient status in a population through a survey can be done for the purpose of estimating prevalence. Other purposes include screening for interventions, evaluation of interventions, surveillance and research. Considerable progress has been made in the assessment of micronutrient deficiencies.

Vitamin A

Clinical data – Bitot's spots or keratomalacia and xerophthalmia. Trends in the prevalence of symptoms of severe vitamin A deficiency are useful to assess the possible magnitude of the problem and the evaluation of interventions and have been used in population-based surveys in developing countries. A new method is pupillary threshold dark adaptometry. A box is used to measure the tendency of the pupil to restrict under light, but the test is still somewhat awkward to use in field settings.

Biological assessment – retinol-binding protein is not a good indicator of vitamin A deficiency at the individual level, but is at the population level and can be determined from a dried blood spot on a filter paper, although the test is not rapid and easy. Assays for retinol binding protein are under development. PATH (Program for Appropriate Technology for Health) has developed an ultraviolet strip reader for field use, which is currently undergoing field testing.

Anemia

Clinical signs – pallor can be assessed in conjunctiva, palm and nail beds and can be a useful screening method, but sensitivity and specificity are relatively low.

Biological assessment – several methods are available. Hematocrit needs a centrifuge to separate cells from plasma and is not suitable for surveys. The WHO Colour Scale – is a low-cost filter paper method. The HaemoCue is most commonly used in field conditions, using a portable hemoglobinometer.

Iron and folate status

Serum transferrin receptors (sTfR) is a better indicator of iron status than ferritin levels which are affected by infection. Several kits are available and require a small amount of serum. Dried blood spots on filter paper can be used to assess folate status.

Zinc

No clinical methods have been used. Biological assessment of zinc status can be done using serum zinc concentration (better indicator at population level than at individual level), hair and breastmilk zinc concentration, and serum zinc metallothionein (is better indicator of zinc stores). No rapid method is available.

Iodine

A simple test to assess the iodine content of salt has been used in many household surveys. Goiter palpation is the most common clinical method to assess chronic iodine deficiency in surveys. A rapid urinary test based on a color scale has been developed to assess urinary iodine excretion. Blood spots can be used to assess levels of thyroid stimulating hormone (TSH).

Excerpted from conference presentations by Keith West 'State-of-the-art of bio-assessment: micronutrients', Paul Arthur 'Experience with micronutrient testing in population-based surveys', and Umesh Kapil 'Experiences with micronutrient testing in population-based surveys'.

Adding biomarkers to population-based surveys: is it ethical?

Like the national census, large household surveys that look at fertility, health and service use are generally viewed as planning activities rather than as research activities. The knowledge that they generate is intended to be translated directly into more effective programmes and better health in the country conducting the survey. In other words, they proceed on the assumption that the results will be used for the benefit of the whole population. This assumption justifies the intrusion on the time and privacy of individual respondents, even in the absence of any benefits specific to the individuals that respond.

The addition of biomarker testing to surveys, and the specimen collection that it implies, may change the ethical parameters. Issues of informed consent come to the fore, as do questions about what personal data can be linked to biological test results. Decisions must be reached about responsibilities for treatment or referral for those found to be suffering from infection or nutrient deficiency.

Securing consent

Consent for any survey or research study must obviously be secured at a national level through the appropriate ethical board or other authority. If a survey is externally funded, funding country authorities may also expect to review and clear the proposed work. While many countries have clear guidelines on ethical requirements for research studies, it is often not clear whether these guidelines cover household surveys, which are more often thought of as programme activities. The addition of biomarkers to population-based surveys carries them closer to the realm of research, and national and international clearance procedures for such activities need to be clarified.

At the individual level, consent for response to a survey questionnaire is usually sought from each participant selected from the sample frame. In general, literate participants are asked to read and sign a consent form, while the form is read out to illiterate participants, who then give verbal consent. In past DHS studies, this consent has been deemed to cover anthropometric measures as well as question responses. With the addition of anemia testing, two further consent forms were added. One seeks consent to draw blood and test it for anemia, while the second seeks consent to pass the names of individuals with severe anemia on to health authorities for follow-up.

In past research studies, questions have arisen about the extent to which “informed consent” is truly informed. People may consent to testing or other procedures in expectation that it will lead directly to personal benefit, or out of peer pressure or a desire to please authorities. Researchers have also pointed out that it is essential to make it absolutely clear under what conditions, if any, participants will be notified of their test results. People may assume that if they receive no notification of a negative result, they are in good health, even though notification was never part of the study protocol.

Non-response is a major concern in population-based surveys, and it is possible that the addition of specimen collection will increase non-response. Experience of the few DHS studies that have collected blood on filter paper for anemia testing have so far not recorded any major increase in non-response bias. Less invasive specimens such as saliva or urine may minimize non-response. Specimen collection may have to differ for different respondents. For example, adults may be prepared to give blood samples but may be reluctant to have blood drawn from their. Perhaps more widespread is the prospect of non-response associated with the fear of HIV testing. Clearly, this is likely to be highest in countries where HIV prevalence and associated stigma are highest.

The rapid march of technology creates the temptation to store specimens so that they can be tested for other pathogens or conditions once the technology becomes available or affordable. But, if specimens are to be stored in any way that is linked to individual data, respondents must consent both to the storage and to the possible use of their blood, urine or saliva for testing for other, unnamed conditions. This may well increase non-response bias.

Clearance, consent, confidentiality

Are national surveys research or programmatic activities? Most national surveys clearly have programmatic goals but should also be considered research.

If it is a research activity, clearance from an outside review board is required, preferably both within the donor country and with the host country. Review boards should have adequate representation of the intended beneficiaries of the survey or research. The clearance process will require documentation of how consent is obtained and how confidentiality is protected.

With regard to consent, there is always potential for the participant to misunderstand the purpose of the activity. There is a tendency for participants to think that research activities are clinical care and for personal benefit. This is stronger when a clinical procedure is part of the activity and when people have no familiarity with research. Survey participants might also assume that they have a clean bill of health if they do not hear back.

Consent can be oral or written. The consent must stress that the purpose is to understand the health of the people in the region and not to help the participant, and that no feedback to the individual is given.

Breach of confidentiality can result in harm, especially for stigmatizing conditions. Study procedures need to minimize this risk. Easy safeguards include clever use of codes to complicate linking. Safeguards that are harder to implement concern the survey staff and their relationship to participants and local authorities. This may also pertain to safeguarding communities as a whole, e.g., minorities.

Excerpted from conference presentation by Nancy Kass 'Responsibility to obtain clearance and informed consent, and maintain confidentiality' and Guillermo Figueroa 'Responsibility to counsel and treat the survey participants'.

Sharing results and the treatment dilemma

One of the great challenges presented by biomarker surveys is what to do about advising people of their test results, and treating where necessary. It is the norm in research studies to provide participants with, at a minimum, the existing standard treatment when needed. This has not, however, been the case in population-based surveys. In over a decade of anthropometric measurement, mothers were informed of their children's height for weight, but no survey team provided treatment or even referrals if the anthropometric data indicated malnutrition. In fertility surveys, staff have never provided information on contraception to women who report more children than they want and no contraceptive use.

With the collection of biological specimens, these issues of providing information and treatment arise again. DHS teams testing women for anemia give results on the spot and refer women to health clinics where necessary. Severe cases are also referred directly to clinic staff, but there is no further follow-up. Iron supplements are not provided by the survey team.

Any decision to provide treatment within a survey would raise massive difficulties. For a start, it would add to training, transport and commodity costs. Secondly, it raises questions about accuracy of diagnosis. No test is perfect under field conditions. For some conditions this is not wildly important, since an incorrect diagnosis or even unindicated treatment for anemia or vitamin A deficiency is unlikely to have devastating consequences. An incorrect result for other conditions, notably STIs and particularly HIV, are much more worrying. Confidentiality must be maintained, and confirmatory tests for HIV are critical. Once laboratory testing is involved, it becomes difficult to follow up individuals for counseling or care, and the potential for breaches of confidentiality is high.

Decisions would also have to be made about who qualifies for treatment. Studies comparing levels of anemia suggest that if a standard definition of anemia were used to qualify for treatment, up to 25 times as many people would have to be treated than if a more stringent definition of "moderate to severe" anemia were used. What should one do about people who report symptomatic infection (for example urethral discharge or genital warts) but for whom no biological confirmation is available? What about people whose sexual history puts them at clear risk for HIV or STI infection? Should they

be counseled or provided with condoms? And, importantly, what does one do about individuals in the community who suffer the same prevalence of ill-health as study participants but who were not selected for inclusion in the sample survey?

Considering these questions leads one to reconsider the importance of specimen collection. Is taking someone's body products really so different from taking their body measurements or their sexual and reproductive history? Should a drop of blood change the procedure that has served for many years? Should this change the assumption that participants can be expected to provide information that will contribute to the improvement of health services as a whole, without expecting any individual gain? Or have countries conducting surveys (and the institutions that back them) been in dereliction of duty for decades because they have not provided counseling, information and treatment to survey participants with unmet needs? There are no simple answers.

STIs and HIV

Survey data on self-reported symptoms of sexually transmitted infections (STIs) generally do not provide reliable estimates of STD prevalence or incidence in the community. In addition to a clinical examination, a range of diagnostic tests can be carried out to diagnose recent or past STIs. Many of these tests have become easier and cheaper and some are now suitable for use in household surveys. For instance, in Uganda self-administered vaginal swabs were well accepted.

HIV antibody testing can be done on saliva, urine and blood. For instance, a population-based survey in Zambia used saliva, collected by an OraSure device. Urine-based HIV testing has been used in the evaluation of an adolescent sexual health program in Tanzania. In the USA two ELISA tests with different levels of sensitivity have been used to estimate HIV incidence from a single sample (detuned ELISA). The method has been field tested in developing countries.

A summary of some of the tests by type of body fluid:

Blood

- Syphilis: non-specific tests such as VDRL, RPR and Trust for active syphilis (costs about \$1 per test)
- Syphilis: specific tests such as TPHA, FTA-Abs, and more recently, an immunochromogenic dipstick (\$4) measure prevalence of antibodies against syphilis (indicating recent or past exposure)
- Herpes simplex: serum antibody test is easy and measures ever-exposure (\$22).
- HIV: large number of ELISA tests with very high sensitivity and specificity available. Blood-spotted filter paper is sufficient. An increasing number of on-site simple and rapid tests are on the market. Most devices are of adequate sensitivity and specificity. The best rapid tests are supplied with standardized reading devices.

Urine

- Genital discharge syndromes: leucocyte esterase dipstick in urine has low specificity and sensitivity.
- Gonorrhoea and chlamydia: nucleic acid amplification tests (PCR, LCR) (expensive)
- Herpes simplex virus: antibody test, measures ever-exposure
- HIV: accurate ELISA tests for urine samples are available

Saliva

- Syphilis: saliva based tests are under development.

(Self-administered) Vaginal swabs

- Bacterial vaginosis – Gram stain (\$0.50) or rapid card test (\$3)
- Trichomonas vaginalis: wetmount (\$0.50), In-pouch TV culture (\$3.00) or PCR in urine (expensive)

Excerpted from conference presentations by Caroline Ryan 'Data needs and current practices (RTIs/STIs)' and Myron Cohen 'State-of-the-art of bio assessment: HIV'.

Testing for HIV

HIV is an incurable disease that is highly stigmatized in most societies. A positive test result can have dire consequences for individuals, including job loss, abandonment, violence and denial of basic services such as health care. This clearly puts it in a category apart from anemia or vitamin A deficiency, and the issues it raises merit separate consideration.

There are essentially two possibilities for HIV testing in a survey context: voluntary anonymous testing, and unlinked anonymous testing. Voluntary confidential testing, in which an individual gives informed consent to be tested for HIV and receives pre-test and post-test counseling, is not feasible in a large scale survey. Voluntary anonymous testing also requires informed consent, but a person is not necessarily informed of the results. In most cases, the individual is referred to a local facility for pre-test counseling, confidential testing and post-test counseling, should they choose to know their results. One of the advantages of these methods is that HIV status can be

linked to a range of demographic and behavioral characteristics, but only if all the usual safeguards on confidentiality can be maintained. A major problem arises if local HIV counseling and testing services are dysfunctional or of poor quality (or not available as in most rural areas in developing countries). A potential limitation of these methods is that they may result in very high non-response rates that may affect the quality of the regular survey and other data collection.

The second option is unlinked, anonymous HIV testing. This is the method most commonly used in HIV surveillance, particularly among pregnant women and STD patients attending clinics. In unlinked, anonymous HIV testing, blood or other specimens are taken for another purpose (e.g. syphilis or anemia testing). Consent is given for that primary purpose, but no consent sought for HIV testing. Leftover samples are stripped of identifying markers, although age, parity and sex data are usually retained. It is critical at this stage that there is no way of tracing a specimen back to the individual who donated it. The specimens are then tested for HIV.

Unlinked anonymous testing is easily the most feasible method, logistically. Drawbacks include the inability to inform individuals of their HIV status or to provide counseling on prevention or care. Indeed, for this reason a few countries consider unlinked anonymous testing for HIV unethical, and some have made it illegal.

Since only a minimum of data can be linked to the specimen, this method gives little information other than age/sex/geographic area prevalence of HIV in the population. In general, other socio-demographic or behavioral factors that might yield useful information about the epidemiology of the virus cannot be investigated. It has been suggested that a wider set of variables may be linked to the data without compromising the principle of absolute anonymity (in Germany a minimum of five respondents would have to have the same characteristics to be considered unlinked anonymous data).

A further caveat of unlinked anonymous testing for HIV is a potential breach of population trust. Many survey programmes put a great deal of time and energy into building up relations with a population in order to maximize participation. If consent is not sought for HIV testing, and respondents later see that the study in which they participated generated information for HIV, they may not look favorably on future surveys. This however has not occurred in antenatal clinics where unlinked anonymous testing has been carried

out for more than a decade. Careful attention to international guidelines, such as those developed by WHO/GPA in 1989, can help address some of the weaknesses and limitations of unlinked anonymous testing.

Unlinked anonymous screening

One of the primary epidemiological objectives of the public health surveillance of HIV infection is to obtain information on the prevalence and incidence of the infection in selected population groups in a manner that is as free as possible of participation and selection bias. Unlinked anonymous screening (UAS) or testing is generally considered to be an accurate and effective method for public health surveillance of HIV infection. However, UAS has inherent limitations, so that it cannot be considered a complete solution to all problems of public health surveillance of HIV infections. Whenever UAS is being contemplated as part of a comprehensive HIV/AIDS prevention and control programme, careful attention should be paid to the criteria and points for consideration:

Criteria

Prior to implementing public health surveillance through UAS, it is essential to conduct a thorough discussion of the ethics of UAS in the social and cultural context of the country where it is to be implemented. If it is against established national public health policy UAS should not be implemented. UAS can be regarded as being consistent with the existing global guidelines on human rights in biomedical research. If the proposal for **UAS originates in one country, but is conducted in another**, it should be reviewed by both an ethical review committee of the country of origin, as well as its equivalent in the host country.

Specimens for UAS should have been taken with appropriate consent for other purposes (e.g., anemia testing). To take blood primarily or solely for UAS would raise serious ethical concerns. The volume of blood taken should be the minimum necessary and should be, at most, only marginally greater than that required for the other tests for which the blood was originally obtained.

No information should be requested in addition to that normally collected for the primary purpose for which the blood specimen was obtained.

All data that could potentially identify the individual must be removed from the specimens set aside for UAS before they are tested by the laboratory.

Protocols for UAS should be carefully reviewed to ensure (a) that there is no possible way in which test results could be traced back to individuals, (b) that studies are designed to maximize the likelihood of obtaining data useful for surveillance purposes, given the estimated prevalence in the population under surveillance, and (c) that staff are trained to adhere to the UAS protocol and supervised to avoid breaches of anonymity.

Voluntary testing (confidential or anonymous) with counseling should be available wherever possible to populations in which UAS is being carried out, so that those individuals who wish to know their HIV-infection status can do so. This is particularly important if the population is estimated to have a moderate to high prevalence of HIV infection. However, such testing should be offered through a separate system.

The resources devoted to UAS should be commensurate with its value for surveillance, as one part of a comprehensive HIV/AIDS prevention and control programme. UAS should not detract from other important public health objectives, including the primary purpose for which the specimens were obtained.

Health care providers should be made aware that the specimens drawn by them from patients might be used for unlinked anonymous HIV screening.

In areas with low HIV prevalence, pooling of sera collected for UAS might be considered.

Points for consideration

- How will the public be informed of UAS in a way that will not deter people from using health care services where specimens may be obtained for UAS?
- How will health care providers and the public be informed and assured of the appropriateness and anonymity of UAS?
- How will services be targeted to population groups found to include HIV infected individuals?
- What information (e.g., age and sex) will be retained with the blood sample, given the need to guarantee anonymity and yet to obtain the most useful data for surveillance purposes? In general, the maximum useful information should be retained without jeopardizing anonymity. Aggregation of information retained with the sample (e.g., age information by age group only) may be of value.
- How will UAS findings be presented in order to reinforce other HIV/AIDS prevention and control activities?

Excerpted from conference presentation by Stefano Lazzari, 'When is it ethical to collect unlinked data?'. And WHO/ GPA. Unlinked anonymous screening for the public health surveillance of HIV infections. GPA/SFI/89.3. Geneva. June 1989.

Adding biomarkers to population-based surveys: is it useful?

The experience of the DHS with anemia testing has demonstrated that it is indeed possible to do at least some biomarker testing within large household surveys, but possible does not equal useful. Added data collection can be deemed useful only if the data add substantially to what is already known, to allow for better health policies to be formulated. Ideally, there should be both the possibility of better policies being formulated, and a real likelihood that the data will lead to those policies being implemented.

Adding value to existing data sources

The addition of biomarkers to household surveys must achieve two things if it is to add value to existing data. It must improve substantially on biological data already available from other sources, and it must improve substantially on data which might be obtained from a survey without the collection of biological specimens, for example through clinical diagnosis or self-reporting of symptoms.

The extent to which biomarker surveys add to existing data depends to a large extent on what data are already available, as well as what programmes are in place. The prevalence of mild and moderate levels of micronutrient deficiencies, well-hidden and undetected except for its most severe forms in many populations for decades, can only be measured accurately through biological and clinical data collection. Population-based surveys have multiple advantages over data from clinic populations for the assessment of the size and the distribution of the problem and for the subsequent planning and evaluation of interventions.

If a well-functioning and extensive sentinel surveillance system for HIV gives a good overview of the prevalence of the virus among pregnant women of different ages in different areas of the country, the usefulness of adding unlinked anonymous testing for HIV to a household survey is debatable. However, in most countries there are hardly any surveillance sites in the rural areas (where most of the population lives) and antenatal clinic-based systems do not provide data on men. Furthermore, multiple biases may affect antenatal clinic data (e.g., low attendance by pregnant women, selection bias for sexual activity, fertility-reducing effect of HIV). At this point, it is not clear how much can be gained by HIV prevalence data collection

in population-based surveys, but potentially, HIV prevention and AIDS care programs would benefit substantially from having more accurate data on levels, trends and differentials in HIV prevalence.

Biological and clinical data collection in population-based surveys is particularly important in the assessment of inequalities in health and health status. Surveys are the most important instrument to reveal inequalities by socio-economic status, geographic or demographic characteristics of individuals and households.

The “value added” equation is especially important when tests are expensive. Questionnaire data from a recent survey in Central African Republic recorded around 75 percent coverage of tetanus toxoid immunization in the target population. This was around ten percentage points lower than the level of immunity registered in a sero-survey of a sub-sample of a quarter of the same women. The survey data underestimated the country’s progress towards internationally-agreed coverage goals for tetanus toxoid immunization in women of reproductive age, although not radically. Policymakers in the country, which spends about 11 dollars per person per year on health, must evaluate whether the extra information was worth the extra cost of the tests (around 11 dollars per test), and the opportunity costs that the provision and training of staff represented to the rest of the health system.

It is important, too, to consider sources of data other than biological or clinical data, particularly where information is being used to evaluate programming. These include programme data and behavioral data. It is not reasonable to expect a supplementation intervention to produce any differences in, say, vitamin A deficiency if the intervention never actually reaches the target population. While this seems a statement of the obvious, it is remarkable how frequently process data are overlooked, how strong the temptation is to jump straight into looking for biological markers of the success of an intervention, without reviewing the process data that would indicate whether the intervention was even implemented as planned. The same is true of behavioral data: it is unlikely that HIV prevalence data will reveal the success of an HIV prevention programme if there have been no changes in sexual networking or condom use. And it is a lot less complicated in a survey to ask about changes in the behaviors that lead to HIV infection than it is to collect blood and test for antibodies to the virus.

The likelihood of policy changes

Population-based surveys are frequently justified by the fact that they produce information that will lead to better health programmes, and ultimately to a healthier and more fulfilled population. The information they produce should therefore be actionable, and there should be a reasonable expectation that they will be acted upon.

Involvement of decision-makers in the choice of information to be collected (including biomarkers) should increase the likelihood that the data will be used. This is especially true when governments are contributing all or a substantial part of the resources used in data collection. Specifying how data generated will be used – including the implications it will have for health programmes at the national and district level – is an important part of planning data collection and choosing biomarkers. Without local political commitment to data collection, it can be extremely difficult to secure the resources needed to ensure the success of a population-based survey. Political leaders in democracies are unlikely to support data collection of any sort unless they understand how the results can be used to benefit their constituencies.

Data are more likely to lead to policy changes if the data and survey results are released soon after the data collection exercise, when interest and momentum are still high and before other political priorities have taken over. This is an important consideration when weighing up the costs and benefits of adding biomarkers to a general population survey. Specimen collection and laboratory analysis can add many months to survey work. For example, data collected in 1997 in a survey of STIs, HIV and sexual behavior in several African cities had still not been published by early 2000, partly because difficulties with lab work in STI testing and subsequent re-testing led to long delays, holding up the publication of other information that had important policy implications. Similar delays are less likely where on-the-spot testing is possible. In India, a large-scale survey of family health conducted in seven states and including HemoCue tests for anemia finished data collection in November and published results by April.

Biomarker testing in large population-based surveys can, demonstrably, lead directly to policy changes. For example, a national survey in Pakistan determined that measles immunity was actually only 60 percent, far lower than the 90 percent expected by

the national immunization programme. This finding led directly to major changes in immunization policy. Another example comes from the Central Asian region. Here, anemia testing included in DHS studies in three countries has led to a large UNICEF-backed programme identifying and promoting local iron-rich foods, and iron fortification and supplementation are planned. The data may also have implications for other programme areas. The DHS surveys found on the one hand that short birth intervals and high parities were associated with high maternal anemia, arguing for a strengthening of birth-spacing programmes. On the other hand, the surveys found that women who used intrauterine devices (IUDs) were twice as likely to be anemic as other women, arguing for a change of method mix within the family planning programme.

Adding biomarkers to population-based surveys: is it desirable?

If it can be established that the addition of biomarkers to population-based surveys is feasible and useful, a final check is needed. Is it desirable? This is essentially an exercise in weighing the likely benefits of biomarker data to the health of a population (including the likelihood that the data will actually be used) against the costs of conducting the data collection. These costs include financial costs and opportunity costs to the health system of using personnel and resources in this way, considered above, but include also the possibility that existing data collection systems will be compromised.

The danger of increasing non-response to regular surveys has been discussed. The provision of treatment by a survey team may minimize non-response but raises a host of other difficulties. It is worth repeating that the effect of specimen-collection on survey response rates will almost certainly vary from place to place. The prevalence of and stigma surrounding HIV, in particular, may affect the acceptability of specimen collection for biomarker testing. It is not safe to assume that what works in Kazakhstan will work in Zambia.

The addition of biomarkers to population surveys may affect their quality in other ways. The extraordinary effort required to organize a successful biomarker survey may detract from time spent ensuring the quality of interviewer training and questionnaires. Interviewers trained to establish a personal rapport to put people at ease in order to maximize the reliability of sensitive information about sexual behavior may find their efforts usurped by specimen collectors in surgical gloves brandishing lancets and consent forms.

A further cause for concern in internationally driven survey programmes is the desirability of highlighting regional, ethnic or other inequalities in health status. This may be considered undesirable by some governments that may fear that highlighting inequities could exacerbate tensions within the country. Such fears are likely to lead to low political commitment, undermining the success of the survey.

Introducing biomarker testing in other ways

Needless to say, there are other ways of adding biomarker tests to a national health planning exercise. Testing every respondent in a national household survey such as DHS is one option, and an important one, but there are others.

One alternative option is to opt for clinical examination, taking biological specimens only from a sub-sample of participants in a regular survey and using the results to calibrate clinical examination data. This may help to reduce cost and logistic implications, and minimize any negative impact on the regular survey data. Limited experience suggests that sub-samples are viable for high prevalence conditions. In Kazakhstan, anemia testing was performed on all DHS respondents in 1995. Four years later, only a sub-sample of DHS respondents was tested for iron deficiency. The results of sub-sample testing were encouraging.

To date, biomarker testing has most commonly been used in research studies. These are often relatively well resourced, and are conducted on a smaller scale than national household surveys. This allows for careful quality control, and because research studies are generally rather tightly focussed, sampling frames can be constructed to maximize returns in terms of data generated.

Research studies may be particularly valuable for highly stigmatized conditions such as HIV and STIs. They are able to put more time and effort into ensuring informed consent and providing parallel counseling and voluntary testing than are regular household surveys. This should allow biological data to be linked with a larger and more useful set of demographic and behavioral variables. They are also better able to target samples according to the state of the HIV epidemic in a country. Such studies may well choose a household-based sample frame at a district or other level in high HIV prevalence countries, but opt to focus on populations with behaviors that carry high risk for HIV transmission in countries where prevalence in the general population is low.

Research studies are often international collaborations, and it may be difficult for developing countries to rely on such studies as part of their regular health planning and monitoring system. However countries can do much to reduce barriers to international collaborative research (often greatly to the benefit of the local research community)

and can solicit research projects in areas they feel present important information gaps in their health planning. Research studies can provide invaluable information for health planning even when they are not conducted on a national scale, and the information they generate may well obviate the need for large household biomarker studies.

At the other extreme are national health examination surveys. These are comprehensive studies of health and well-being, conducted on a nationally representative sample of households. Typically, households are interviewed about health status and service utilization, and household characteristics such as sanitation are recorded. Then, household members are invited to attend a mobile examination center where specimens are drawn and a battery of measurements and other tests are performed. Separate clinics are set up for the treatment of clinically diagnosed conditions for all community members, regardless of whether they were included in the sample. Such studies produce a mine of information, but are extraordinarily time-consuming and resource intensive. One such survey in Pakistan took three years to perform around 18,000 examinations, at a cost of at least 1.5 million dollars.

National health examination surveys

Four developing countries (Papua New Guinea, Columbia, Egypt and Pakistan) have completed national health examination surveys that were broad in scope. In such surveys, a national sample of clusters is taken and in each cluster a sample of households is taken. All eligible individuals are requested to come to a static or mobile clinic serving as a survey site. In addition to an interview, a wide range of clinical and biological diagnostic tests are performed in the survey site by medical personnel. Several tests are performed on site (anemia, cholesterol, creatinine, etc.). A typical health examination survey requires two to three years to complete. Its main outcome is a national health profile that may include nutritional status, prevalence and extent of disability, infectious disease prevalence, chronic disease prevalence, and assessment of selected national health programs such as immunizations. An example of a very extensive national health examination survey is the NHANES in the USA (National Health and Nutrition Examination Survey). The survey examines a nationally representative sample of about 5,000 persons each year and includes biological and clinical data collection for a large number of diseases and conditions.

Excerpted from a conference presentation by Greg Pappas 'Experience with health examination surveys in developing countries'. See also Fischer G, Pappas G, Limb M. 'Prospects, problems, and prerequisites for national health examination surveys in developing countries.' Social Science Medicine 1996;42(12):1639-50.

It is worth mentioning that while considerable private resources are being committed to improving diagnostic tests for a number of biomarkers, far less effort is being put into improving clinical diagnoses. This is for obvious reasons: companies cannot patent or sell a clinical diagnosis. It may be worth considering investing more public funds in improving clinical diagnostic algorithms that can be added to household surveys at much lower cost.

Conclusions

It is by now clear that it is possible to add biological and clinical data collection to household surveys, and it has been shown that in many instances the data generated are both useful and used. The decision on whether or not to add biological and clinical data collection to general population surveys, and what kind of data collection is most needed and useful, needs to be made on a country-by-country basis.

One of the great contributions of international programmes such as DHS is the production of standardized data that allows for comparison across countries and over time. Inclusion of biological and clinical data in DHS and other national surveys may lead to new perspectives on public health at the national and international levels and is likely to have a major impact on health programmes. Our knowledge of the burden of disease, inequality in the distribution of health and disease within the populations, and the health impact of interventions is limited in developing countries. This drive for standardization should not, however, lead to the indiscriminate addition of biomarker testing to all country surveys. Rather, specimen collection and testing might best be adopted as a module (similar to modules on AIDS and female circumcision) in countries where the process is least likely to compromise regular survey information, and where the data are most likely to be translated into policy changes.

Obviously, there are many issues that need further thought and discussion. However, some preliminary conclusions may be drawn from the meeting held at the National Academy of Sciences in January 2000.

- **Data assessment**

In deciding whether to include biomarker surveys in a given country, it is crucial to begin by reviewing the data that are already available, and the extent to which those data have been used. This also includes an assessment of usefulness and adequacy of self-reported data from surveys.

- **Cost assessment**

The potential (and likely) benefits of any data collection must be weighed against the costs. These include the danger of compromising survey data quality and the opportunity costs to the health system of tying up funds and qualified staff in specimen collection and analysis rather than in programme implementation.

- **Treatment and counseling**

Counseling and treatment have not traditionally been provided in general population surveys. In case of rapid on-site testing, results on treatable conditions are reported to the respondent who is advised to go to the local clinic in case of results indicating disease or deficiency. The rationale that has allowed for needs assessment without the provision of services in past surveys still holds: the survey should lead to improvements in policy and programme implementation, that should in turn lead to better health for the entire population, survey participants and non-participants alike.

- **HIV testing**

The stigma attached to HIV and the harm that can come to people known to be HIV positive put this infection in a different category from other health conditions for which biomarker testing might be considered. The ethical and logistic complications of adding HIV testing to regular household surveys in any way that will allow for HIV status to be linked to data appear insurmountable. Unlinked, anonymous testing is more feasible and the most practical option for HIV testing in surveys in countries where there is no extensive infrastructure for voluntary testing and counseling.

This summary and the preliminary conclusions provide a starting point, rather than an end point, for further discussion of the important issues raised. A working group will develop a set of guidelines for international organizations, such as USAID, and for countries. These guidelines will describe the issues and criteria to be considered in the planning of surveys with biological and clinical data collection.

Appendix A: List of participants

Jennifer Adams	USAID	Washington, DC
Marie-Christine Anastasi	Population Council	Washington, DC
Paul Arthur	Kintampo Health Research Centre	Kintampo, Ghana
Wendy Baldwin	National Institutes of Health	Bethesda, MD
Lawrence Barat	Centers for Disease Control and Prevention	Atlanta, GA
Sarah Bassett	MEASURE Evaluation, University of North Carolina	Chapel Hill, NC
Karen M. Becker	Centers for Disease Control and Prevention	Raleigh, NC
George Bicego	MEASURE DHS+, MACRO International Inc.	Calverton, MD
Robert Black	Johns Hopkins University	Baltimore, MD
Naomi Blumberg	USAID	Washington, DC
Ties Boerma	MEASURE Evaluation, University of North Carolina	Chapel Hill, NC
John Borrazzo	USAID	Washington, DC
Bates Buckner	MEASURE Evaluation, University of North Carolina	Chapel Hill, NC
Minki Chatterji	USAID	Washington, DC
Eunyong Chung	USAID	Washington, DC
Myron Cohen	University of North Carolina	Chapel Hill, NC
Henri Damisoni	National AIDS Control Programme	Lilongwe, Malawi
Michael Deming	Centers for Disease Control and Prevention	Atlanta, GA
Nancy Dole	University of North Carolina	Chapel Hill, NC
Sambe Duale	Support for Analysis and Research in Africa (SARA) Project	Washington, DC
Erin Eckert	MEASURE Evaluation, JSI Research & Training Inst., Inc.	Arlington, VA
Robert Emrey	USAID	Washington, DC
Juan-Guillermo Figueroa	El Colegio de Mexico	Pedregal de Santa Teresa, DF, Mexico
Fannie Fonseca-Becker	MACRO International: Measure Evaluation	Calverton, MD

Dorothy Foote	Human Nutrition Institute	Washington, DC
Karen Foreit	The Futures Group International	Washington, DC
Ciro Franco	MOST, The USAID Micronutrient Program	Arlington, VA
Anastasia Gage	USAID	Washington, DC
Duff Gillespie	USAID	Washington, DC
Lawrence M. Grummer-Strawn	Centers for Disease Control and Prevention	Atlanta, GA
Phil Harvey	MOST, The USAID Micronutrient Program	Arlington, VA
Yusuf Hemed	Adult Morbidity & Mortality Project (AMMP – 2)	Dar es Salaam, Tanzania
Mai M. Hijazi	USAID	Washington, DC
Elizabeth Holt	Johns Hopkins University	Baltimore, MD
Douglas Huber	Pathfinder International	Watertown, MA
Ishrat Z. Husain	HIV/AIDS Africa HRD	Washington, D.C.
Seham Hussein	Ministry of Health and Population	Cairo, Egypt
Joanne Jeffers	USAID	Washington, DC
Neeraj Kak	The Futures Group International	Washington, DC
Saidi Kapiga	Harvard University	Boston, MA
Umesh Kapil	All India Institute of Medical Sciences	New Delhi, India
Mihira Karra	USAID	Washington, DC
Nancy Kass	Johns Hopkins University	Baltimore, MD
Kevin Kinsella	National Research Council	Washington, DC
Marc LaForce	BASICS II	Arlington, VA
Steve Landry	USAID	Washington, DC
Deborah Lans	USAID	Washington, DC
Stefano Lazzari	World Health Organization	Geneva, Switzerland
Rose Maria Li	National Institutes of Health	Bethesda, MD
Yihong Li	University of Alabama	Birmingham, AL

Mark Lurie	The Africa Centre for Population Studies and Reproductive Health and South African Medical Research Council	Mtubatuba, South Africa
Ellen MacLachan	Centers for Disease Control & Prevention	Atlanta, GA
Lindiwe Makubalo	Health Systems Research and Epidemiology	Pretoria, South Africa
Kate McIntyre	Tulane University	New Orleans, LA
Caryn Miller	USAID/Johns Hopkins University	Washington, DC
Dayton T. Miller	Centers for Disease Control and Prevention	Atlanta, GA
Tersia Mitchell de Wilzem	University of Pretoria	Pretoria, South Africa
Melody Moore	MEASURE Evaluation, University of North Carolina	Chapel Hill, NC
Jose Mora	MOST, The USAID Micronutrient Program	Arlington, VA
Allisyn Moran	MNH Program	Baltimore, MD
Ritu Nalubola	MOST, The USAID Micronutrient Program	Arlington, VA
Christopher B. Nelson	World Health Organization	Geneva, Switzerland
Penelope Nestel	OMNI Research	Washington, DC
Susan Newcomer	NICHD/NIH	Bethesda, MD
Emma Ottolenghi	Population Council	Washington, DC
Gregory Pappas	Department of Health & Human Services	Washington, DC
Elizabeth Pisani	MEASURE Evaluation	Nairobi, Kenya
Timothy C. Quick	USAID	Washington, DC
Mizanur Rahman	Pathfinder International	Watertown, MA
Serena Rajabiun	Academy for Educational Development	Washington, DC
Usha Ramakrishnan	Emory University	Atlanta, GA
Thomas Rehle	Family Health International	Arlington, VA
Charlene Reiss	MEASURE Evaluation, University of North Carolina	Chapel Hill, NC
Jonathan Richmond	Centers for Disease Control & Prevention	Atlanta, GA
Cynthia Ronzio	National Center for Health Statistics	Hyattsville, MD
Amanda Rose	MACRO International: Measure Evaluation	Calverton, MD

Caroline A. Ryan	Centers for Disease Control & Prevention	Atlanta, GA
Bernhard Schwartzlander	UNAIDS	Geneva, Switzerland
Audrey S. Seger	Population Reference Bureau	Washington, DC
Almaz Sharmanov	MEASURE DHS+, MACRO International Inc.	Calverton, MD
James Shelton	USAID	Washington, DC
Buppha Sirirassamee	Mahidol University	Nakom Pathom, Thailand
Geoffrey R. Somi	National AIDS Control Programme	Dar es Salaam, Tanzania
Christine Sow	Family Health International	Arlington, VA
Jeff Spieler	USAID	Washington, DC
Kai Spratt	USAID	Washington, DC
Karen A. Stanecki	US Census Bureau	Washington, DC
David Stanton	USAID	Washington, DC
Ellen Starbird	USAID	Washington, DC
Robert Steinglass	BASICS	Arlington, VA
Krista Stewart	USAID	Washington, DC
Monica Tandon	All India Institute of Medical Sciences	Fairfax, VA
Natalie Tomitch	National Institutes of Health	Bethesda, MD
Martin Vaessen	Macro International	Calverton, MD
Johannes van Dam	Population Council	Washington, DC
Virginia Vitzthum	Binghamton University	Binghamton, NY
Kirsten Vogelsong	USAID	Washington, DC
Joshua Volle	Tulane University	New Orleans, LA
Dora Ward	CORE Group Polio Eradication Initiative	Atlanta, GA
Maxine Weinstein	Georgetown University	Washington, DC
Sharon Weir	MEASURE Evaluation, University of North Carolina	Chapel Hill, NC
Keith West, Jr.	Johns Hopkins University	Baltimore, MD
Regina Winkelmann	World Health Organization	Geneva, Switzerland

Emily Wong
Basia Zaba

Family Health International
London School of Hygiene & Tropical Medicine

RTP, NC
London, UK