# EVALUATING FORECASTS

Below are answers to questions on how to evaluate forecasts.

## Evaluating Consumption-Based Forecasts

**Question: Were Dispensed to User data used to make the forecast? If not, what level issues data were used?**

Dispensed to User data provide by far the best estimate. The higher the level from which issues data are used, the less reliable the projection; service delivery site visits should have been made to determine whether issues data are an adequate substitute for Dispensed to User data. Except in the smallest distribution systems, projections based solely on central level issues data are unacceptable.

**Question: What percentage of reports from SDP's or warehouses was missing over the period covered by historical data? What adjustments were made for incomplete reporting?**

The higher the percentage of facilities not reporting, the less reliable the projection. The greater the variability in data which *were* available, the more uncertainty the adjustments for missing data introduce.

**Question: How many time periods of data were used for the projection? What percentage of reporting periods was missing from the period covered by historical data? What adjustments were made for missing time periods?**

The more historical data used for the extrapolation, the more reliable the forecast. For an annual projection, a minimum of 2 to 3 years' of quarterly data (i.e., 8 to 12 data points) should be used. The higher the number of missing reporting periods, the less reliable the projection. The greater the variability in data among time periods which *were* available, the more uncertainty the adjustments for missing time periods introduce.

**Question: Were losses reported separately from consumption or issues data? If not, what adjustments were made to account for system losses?**

All logistics systems have product loss in storage and transit, and any LMIS which reports zero losses is immediately suspect. Adjustments for losses should be made based at least on surveys of a sample of storage facilities at all levels, but in practice this is very difficult to accomplish.

**Question: Were there stockouts during the time period covered by the data? If so, what adjustments were made to estimate true demand?**

The greater the percentage of time during which one or more products were stocked out, the less reliable the forecast. The greater the variability in data among time periods in which stocks *were* available, the more uncertainty the adjustments introduce.

**Question: Were there special circumstances affecting past demand which no longer effect the program? If so, what adjustments were made to consumption estimates?**

Political, economic, or other external circumstances may have affected historical consumption trends either positively or negatively, and may similarly affect future consumption. If the projection has been adjusted to account for changes in such external circumstances, a rational justification for both direction and size of the adjustments should be given.

**Question: What was the basis for projection of future consumption? What adjustments, if any, were made to the extrapolation of historical data?**

*Simple extrapolation* of historical data may not be appropriate, depending on the answer to the above questions. If based on *program plans*, significant differences from the historical pattern must be satisfactorily explained. If based on *government or donor targets or policies*, projections which differ from historical patterns are very suspect. If based on *funding or other resource constraints*, justification for anticipated consumption levels should be given, along with an explanation of plans to cover demand which will not be met by currently available resources.

## Evaluating Services Data-Based Forecasts

**Question: Were Visit data (either in total or broken down by visit or client type) used to make the forecast? If not, what services data were used? Are service data definitions written down? Do service delivery staff understand them?**

If any data other than Visits/Revisits were used, the conversion from services data to consumption estimates will be very difficult, and the conversion process must be documented completely. In any case, data definitions must be absolutely clear to staff who record and report services data. If definitions are not well documented and understood, recording inconsistencies should be suspected which will make the forecast less reliable.

**Question: Are prescribing protocols documented and understood by service delivery staff? What evidence is there that such protocols are routinely followed?**

The assumption regarding quantities of each product dispensed at each visit critically affects forecast accuracy. If the protocol specifies that 10 condoms should be dispensed to each client but staff in fact give 12, the forecast will be off by 20%. If protocols are not written and disseminated to staff, the basis for the conversion factor should be carefully investigated. If there has been a history of shortages or over-supply in the program, the likelihood that prescribing protocols are not strictly followed is higher.

**Question: What percentage of reports from SDP's was missing over the period covered by historical data? What adjustments were made for incomplete reporting?**

The higher the percentage of facilities not reporting, the less reliable the projection. The greater the variability in data which *were* available, the more uncertainty the adjustments for missing data introduce.

**Question: How many time periods of data were used for the projection? What percentage of reporting periods was missing from the period covered by historical data? What adjustments were made for missing time periods?**

The more historical data used for the extrapolation, the more reliable the forecast. For an annual projection, a minimum of 2 to 3 years' of quarterly data (i.e., 8 to 12 data points) should be used. The higher the number of missing reporting periods, the less reliable the projection. The greater the variability in data among time periods which *were* available, the more uncertainty the adjustments for missing time periods introduce.

**Question: Were there special circumstances affecting past service levels which no longer affect the program? If so, what adjustments were made to service activity estimates?**

Political, economic, or other external circumstances may have affected historical service levels either positively or negatively, and may similarly affect future service levels. If the projection has been adjusted to account for changes in such external circumstances, a rational justification for both direction and size of the adjustments should be given.

**Question: What was the basis for projection of future service levels? What adjustments, if any, were made to the extrapolation of historical data?**

*Simple extrapolation* of historical data may not be appropriate, depending on the answer to the above questions. If based on *program plans*, significant differences from the historical pattern must be satisfactorily explained. If based on *government or donor targets or policies*, projections which differ from historical patterns are very suspect. If based on *funding or other resource constraints*, justification for anticipated service levels should be given.

## Evaluating Services Data-Based Forecasts for ARV Drugs

**Question: What source(s) were used to estimate the number of treatment episodes of the health program?**

Collecting this data requires either a well-functioning HMIS with good reporting rates or routine sentinel surveys. Conducting routine sentinel surveys requires a great deal of resources, both human and financial, to yield useful data. Therefore, services-based forecasts based on such low quality or non-existent data do not prove to be useful.

However, due to more recent stringent reporting requirements, countries usually do have data on the number of AIDS cases and number of people on ART.  And as such, a services-based forecast for ARVs can be fairly accurate.

**Question: How was the scale-up rate, program targets, or rate of forecasted increase determined?**

The process for calculating the rate of increase for a services-based forecast for ARVs can be quite complicated depending on the number of variables to factor in (program expansion, number of patients initiating treatment, remaining on treatment, switching regimens, patients discontinuing treatment, etc.)  Usually a percentage increase is applied or target numbers of patients are used to then figure out what the percentage increase would be.

Depending on how thorough the process was, what data (and the quality of the data) that was used for determining the increase, the forecast's validity rests greatly on how realistic the rate of increase is.

**Question: What source(s) were used to identify the quantity of drug specified for a standard course of treatment?**

Countries develop Standard Treatment Guidelines (STG) for a variety of diseases and health problems, but the bigger question is not so much if STGs exist, rather it is if the STGs are followed by prescribers. If the STGs are not adhered to by prescribers then the application of the STGs to the number of cases or incidents may either over or underestimate what is needed. Another challenge, especially related to HIV and AIDS, is that STGs for ART and algorithms for HIV testing may change often as new drugs and tests become available or as prices for drugs and tests lower.

**Question: Does the services-based forecast for treatment of a specific disease or health condition represent the same use or all uses of the product as in other forecasts (such as consumption)?**

Services-based forecasts are computed for a treatment of an estimated number of patients with a specific disease or an estimated frequency/recurrence of a health condition that will require treatment during the period of the quantification. Consumption forecasts would include/account for all uses of the product.

When comparing forecasts for products with multiple uses such as ciprofloxacin, ensure that you remember that the consumption forecast represents all uses of ciprofloxacin whereas the services-based forecast could represent only the use of ciprofloxacin for a specific STI.  Care must be taken to ensure that comparisons between forecasts represent the same use of the commodity.

## Evaluating Demographic (Population Data-Based) Forecasts

**Question: What source(s) were used for Total Population, Prevalence rates, Source Mix, etc? Are there known problems with or biases in these surveys?**

DHS's and Reproductive Health and Family Planning Surveys are the most carefully designed and executed demographic surveys, though local surveys may also contain data appropriate for local program forecasts. Local surveys may be especially useful for programs which do not provide nationwide coverage. Where local surveys are used, survey instruments, sampling plans, and data processing protocols should be examined for appropriateness, with the help of demographers or survey researchers as needed. Survey figures for other contraceptive methods are likely to be more accurate.

**Question: How old were these demographic data? If adjustments were made to older data to obtain base year estimates for the forecasts, what methodologies were used?**

The older the data, the less reliable the forecast. Ideally, the date of the primary survey source will have been used as the base year for the survey. If adjustments to survey data were required to obtain base year figures, Adults 15-49 estimates should preferably be taken from the U.S. Census Bureau's International Database or from the U.N.'s World Population Prospects. Other methodologies and assumptions should be fully explained.

**Question: What was the basis for the estimate of Program Proportion of Each Method?**

Population data surveys rarely cover exactly the same target population as a service program. Few surveys other than the DHS's provide source breakdowns, and the DHS source breakdowns may not relate directly to particular service programs. This assumption is critical, and its basis should be carefully examined; the forecast will be in error by the same percentage that this figure is in error.

**Question: What regimen profiles were used to calculate commodity requirements?**

Many programs may have no choice other than worldwide average estimates. Those from WHO and UNAIDS are reasonably reliable. SYP assumptions should be examined closely for forecasts of AIDS prevention programs which serve specific target populations; there are at present few hard data on which to base SYP estimates.

**Question: If regimen-specific projections were made, what was the basis for disaggregating the commodity data into individual regimens?**

Regimen-specific projections are always needed for forecasts made for procurement purposes. If consumption data on regimen -specific distribution are available, they should be used as the basis for disaggregating by regimen (and they should also be used to prepare a

consumption data based forecast). If such data are not available, regimen-specific estimates will be less reliable, though method-specific aggregates may still be acceptable.

**Question: How were population data changes over the forecast period estimated?**

Adults 15-49 estimates should preferably be taken from the U.S. Census Bureau's International Database or the U.N.'s World Population Prospects. If these sources were not available, the methodology used to estimate this number should be fully explained.

**Question: How were prevalence and incidence of disease or frequency of health conditions estimated?**

E.g. HIV and AIDS prevalence should preferably be taken from the UNAIDS reports. Projected changes in the AIDS prevalence rate of more than one percentage point per year should be reviewed carefully, and compared to the appropriate UN and local country estimates. These ranges may not apply to AIDS prevention programs serving particular target populations, but strong justification for higher estimates should be given. Projections based on *government or donor targets or policies* which differ from historical patterns are very suspect.