Strengthening HIV/AIDS Commodity Forecasting

Factors and Considerations for Increasing Accuracy

JANUARY 2011

This publication was produced for review by the U.S. Agency for International Development. It was prepared by the USAID | DELIVER PROJECT, Task Order 1.

Strengthening HIV/AIDS Commodity Forecasting

Factors and Considerations for Increasing Accuracy

The authors' views expressed in this publication do not necessarily reflect the views of the U.S. Agency for International Development or the United States Government.

USAID | DELIVER PROJECT, Task Order 1
The USAID | DELIVER PROJECT, Task Order 1, is funded by the U.S. Agency for International Development under contract no. GPO-I-01-06-00007-00, beginning September 29, 2006. Task Order 1 is implemented by John Snow, Inc., in collaboration with PATH; Crown Agents Consultancy, Inc.; Abt Associates; Fuel Logistics Group (Pty) Ltd.; UPS Supply Chain Solutions; The Manoff Group; and 3i Infotech. The project improves essential health commodity supply chains by strengthening logistics management information systems, streamlining distribution systems, identifying financial resources for procurement and supply chain operations, and enhancing forecasting and procurement planning. The project also encourages policymakers and donors to support logistics as a critical factor in the overall success of their health care mandates.

Recommended Citation
Akhlaghi, Laila, Marasi Mwencha and Naomi Printz. 2011. *Strengthening HIV/AIDS Commodity Forecasting: Factors and Considerations for Increasing Accuracy.* Arlington, Va.: USAID | DELIVER PROJECT, Task Order 1.

Abstract
Forecast consumption, and actual consumption data on HIV and AIDS commodities, from five separate countries were analyzed for error rates in forecasting. The forecasting process was reviewed, as well as the in-country context to determine possible causes for more or less accurate forecasting results. Findings and conclusions were determined based on the analysis.

Cover photo: Staff of the Supply Chain Management System, participate in a PipeLine workshop in Georgetown, Guyana, on July 12, 2007.

USAID | DELIVER PROJECT
John Snow, Inc.
1616 Fort Myer Drive, 11th Floor
Arlington, VA 22209 USA
Phone: 703-528-7474
Fax: 703-528-7480
Email: askdeliver@jsi.com
Internet: deliver.jsi.com

Contents

Tables

Acronyms

3TC	lamuvidune
ABC	abacavir
ACT	artemisinin-based combination therapy
AIDS	acquired immunodeficiency syndrome
AMC	average monthly consumption
ART	antiretroviral therapy
ARV	antiretroviral
AZT	zidovudine (also ZDV)
CD4	cluster of differentiation of 4 cells
CDC	Centers for Disease Control and Prevention
CHAI	Clinton Health Access Initiative
CHAZ	Churches Health Association of Zambia
CIDRZ	Centre for Infectious Disease Research in Zambia
CMS	central medical store
CU-ICAP	Columbia University-International Center for AIDS Care and Treatment Programs
D4T	stavudine
DACA	Drug Administration and Control Authority, Ethiopia
ddI	didanosine
DHO	district health offices
EFV	efavirenz
EHNRI	Ethiopia Health and Nutrition Research Institute
FDC	fixed-dose combination
FEFO	first-to-expire, first-out (warehouse management)
FHAPCO	Federal HIV/AIDS Prevention and Control Office (Ethiopia)
FHI	Family Health International
FTC	emtricitabine
GFATM	Global Fund to Fight AIDS, Tuberculosis and Malaria
HCSP	HIV Care and Support Project (MSH)

HIV	human immunodeficiency virus
HMIS	health management information system
IDV	indinavir
IOCC	International Orthodox Christian Charities
IP	Infection Prevention (Ethiopia)
I-TECH	International Training & Education Center for Health
JICA	Japan International Cooperation Agency
JHU/TSEHAI	Johns Hopkins University-Technical Support for the Ethiopian HIV/AIDS ART Initiative
LMIS	logistics management information system
LPV	lopinavir
M&E	monitoring and evaluation
MAPE	median absolute percentage error
MMIS	Making Medical Injections Safer
MOH	Ministry of Health
MSH	Management Sciences for Health
MSL	Medical Stores Limited
NFV	nelfinavir
NUFU	Norwegian Programme for Development, Research and Education
NVP	nevirapine
PEPFAR	President's Emergency Fund for AIDS Relief
PFSA	Pharmaceutical Fund and Supply Agency
PI	Pathfinder International
PMTCT	preventing mother-to-child transmission
PRA	Pharmacovigilance Action Plan (Zambia)
PSI	Population Services Incorporated
PSP	Private Sector Program (Abt Associates)
/r	low dose version of ritonavir
RTV	ritonavir
RTK	rapid test kit (for HIV)
SCMS	Supply Chain management System
SDP	service delivery point
SPS	Strengthening Pharmaceutical Systems (MSH)

SQV	saquinavir
STGs	standard treatment guidelines
TA	technical assistance
TB	tuberculosis
TDF	tenofovir
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
USG	United States Government
WFP	United Nation's World Food Program
WHO	World Health Organization
WMS	warehouse management system
ZDV	zidovudine (also AZT)
ZPCT	Zambia Prevention, Care and Treatment

Acknowledgments

This publication is dedicated to the many individuals from communities, nongovernmental organizations (NGOs), faith-based organizations, Ministries of Health, and other organizations that have consistently fought for access to essential medicines, diagnostics, and health services. The publication is also dedicated to friends and counterparts who have worked with the USAID | DELIVER PROJECT and its predecessor projects, the John Snow, Inc., DELIVER Project and the Family Planning Logistics Management I, II, and III projects, and to the thousands of committed professionals in Ministries of Health and NGOs who work daily to supply their customers and programs with essential public health commodities.

Sincere thanks go to the field office staff (JSI and SCMS) from the countries that provided the data for this exercise on forecast accuracy, forecasting reports, and background information and reviewed their respective sections. The lessons drawn from the experience of the USAID | DELIVER PROJECT and its predecessor projects in conducting quantification exercises in multiple countries across commodity categories would not have been possible without these valuable contributions.

Thanks also go to the staff at SCMS for sharing reports and data, as well as commenting on the final draft of the paper.

U.S. Agency for International Development (USAID) contracts funded the technical assistance, in-country projects, and research that produced the experience in this paper. We are deeply grateful to the team of professionals in the Commodity Security and Logistics Division in the Office of Population and Reproductive Health of the USAID Global Health Bureau's Center for Population, Health, and Nutrition for their encouragement and advice and for their commitment to improving human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) laboratory and public health programs through logistics.

Executive Summary

Forecasting (the process of estimating the quantity of each health commodity that will be consumed during a specified period in the future) has an enormous impact on the entire logistics system's ability to fulfill all six rights to ensure commodity availability. It is directly linked to upstream logistics functions such as financing and procurement and downstream activities such as service delivery. Information for forecasting is drawn from a plethora of sources, the most important of which is logistics management information systems (LMIS). Forecasting for all products, in particular public health commodities, remains a combination of an art and a science. Actual consumption is *almost always* different from forecasted consumption, as it is virtually impossible to determine the exact quantity of tablets of a particular medicine will be dispensed to clients in a future period.

The USAID | DELIVER PROJECT conducted a review of forecasts in order to analyze different countries' and programs' experiences in forecasting, with the aim of identifying factors that could assist in reducing forecasting errors. Forecasted consumption and actual consumption data on certain HIV and AIDS commodities (namely adult and pediatric ARVs, HIV rapid test kits [RTKs], and select laboratory commodities), in four countries (Countries X and Y, Ethiopia, and Zambia) were analyzed for forecast error rates. Mean absolute percent error, or MAPE, one of the most common measures of error rates and the easiest to interpret, was used as the calculation. MAPE is the absolute difference between forecasted and actual values, expressed as a percentage of the actual values and thus is a measure of the variance of the forecast from the actual consumption. Additionally, for each country the forecasting process was reviewed, as well as the in-country context, to determine possible causes for the forecast accuracy results.

Several key findings were identified. In reviewing the quantifications conducted in the four countries, it appears that high forecast accuracy (low error rates) can be attributed to:

- Well-developed and appropriately staffed logistics system and the resulting improvements in availability of quality data
- A mature and institutionalized program that is staffed by trained personnel
- A rigorous and regular review process that incorporates multiple data sources
- Large and diverse forecasting teams
- Standardization of laboratory supplies
- Regular maintenance of testing equipment

On the other hand, high forecasting errors may be attributed to:

- Delays in implementation of policy changes
- Prescribing practices not adjusting as anticipated
- The complexities of pediatric ART

- Discrepancies in available data for use in forecasting
- Abrupt move from a small, vertical ART program and logistics system to a large, integrated and decentralized ART without requisite coordination
- Uncertainty regarding leadership for the continued implementation and management of ARV LMIS during decentralization process
- Significant supply chain management challenges at the CMS
- Significant changes by MOH to STGs without proper coordination with the parties responsible for forecasting, procurement, and distribution
- Unpredictability of timing of fund disbursement
- In-country distribution challenges.

Although each country and situation is unique, we can draw some conclusions about factors that contribute to the error in forecasting. The cross-country findings are:

1. Planning for program scale-up and forecasting for procurement should be approached as two distinct processes.
2. If there is a policy change concerning any products, develop a realistic timeline for implementation and adjust forecast accordingly.
3. The better the data, the more reliable the forecast.
4. Routinely updating forecasts results in more accurate forecasts.
5. The more experience with forecasting, the more experience with the program, and the more mature the program, the more reliable the forecast.
6. The more prevalent a product or equipment, the less error there is in forecasting its demand.
7. Focus should be on most-consumed products, while keeping large errors low for those consumed least.
8. Regardless of the type of tool used, its accuracy and results are a reflection of the inputs to the quantification, the methodology used, and the management of the quantification process.

Background

Quantification is the process of estimating the quantities and costs of products required for a specific health program (or service), and determining when the products should be delivered to ensure an uninterrupted supply for the program. It is a critical supply chain activity that links information on services and commodities from the facility level with program policies and plans at the national level. Quantification results are used to inform higher-level decisionmaking on the financing and procurement of commodities. The process includes both a forecasting and a supply planning step, as demonstrated in figure 1.

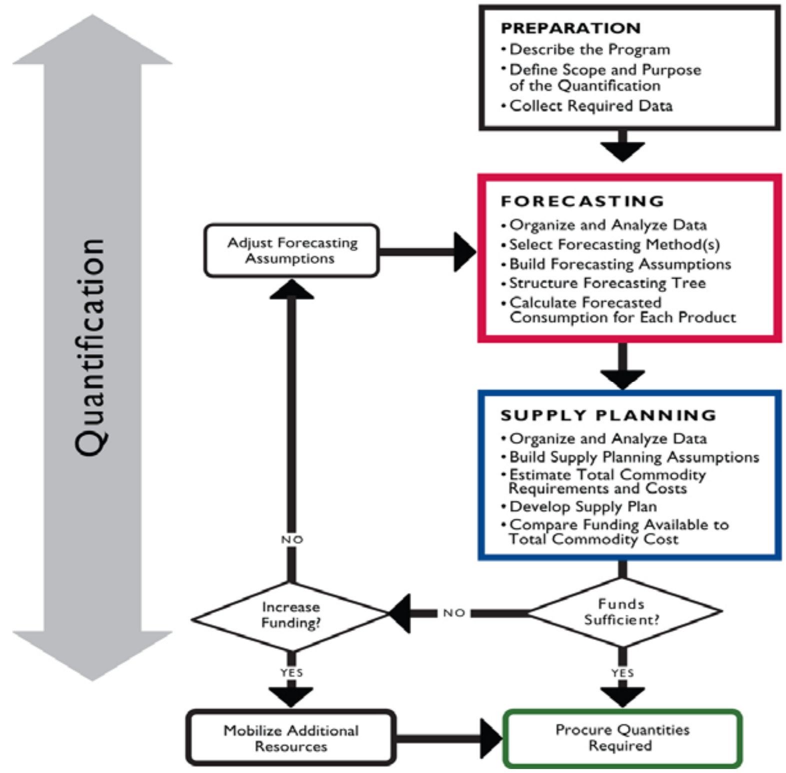


Figure 1. Quantification Process

Forecasting is the process of estimating the quantity of each health commodity that will be dispensed or used during a specified period of time in the future. These quantities then become the basis for calculating the total commodity requirements in the supply planning step. Supply planning starts where forecasting stops. The forecasted quantities to be consumed are adjusted to account for stock on hand, quantities on order, and established minimum and maximum stock levels to determine the quantities of each commodity that should be procured. For the purposes of this report, the focus is on the forecasting step—the quantities of products estimated to be actually consumed by or dispensed to clients.

The goal of any supply chain is to ensure commodity availability, or the six rights: the right goods, in the right condition, in the right quantities, at the right place, at the right time, for the right cost. Forecasting has an enormous impact on the entire logistics system's ability to fulfill all six rights and is directly linked to other logistics functions, such as LMIS. Forecasting attempts to better match up supply with demand. It takes a significant amount of time to conduct a forecast, prepare a supply plan, initiate and follow through on the procurement process, and eventually to receive and distribute the products. This is why it is essential to have advance knowledge of the quantities of products that are likely to be consumed during a period of time.

Forecasts can be made on a short-term basis or a long-term basis. Short-term forecasting is more tactical; it provides vital data on commodity requirements and costs for annual budget allocations on which specific supply plans and procurement contracts are developed, and actual orders are placed. Long-term forecasts are more strategic; they provide critical information for national governments and international funders for multi-year budgeting and resource mobilization. In addition, they can be less precise and larger in scope. The focus of this paper is on short-term forecasts.

Quantification, the process that encompasses both forecasting and supply planning, is not a one-time, annual event. One of the outputs of a quantification exercise should be an implementation plan for routine monitoring, review, and updating of the supply plan at least every six months for more stable, well-established programs, and more often if key data and assumptions change or the actual volume of services and commodity use differs significantly from the forecasted demand for services and consumption. Ongoing monitoring, updating the forecast assumptions, and adjusting the supply plan are critical to keep program managers, donors, and other stakeholders informed about the availability of drugs and diagnostics. This monitoring is vital for timely decisionmaking on product selection, financing, procurement, and delivery of commodities. It also allows changes to be made if assumptions on supply, financing, delivery times, and policy implementation do not occur as anticipated.

The environment in which forecasts are conducted is marked by significant increases in funding and, correspondingly, the numbers and types of commodities supported. At the same time, the number and range of funders, suppliers/manufacturers, and intermediary organizations (such as procurement agents) has also grown. The increase in funding and the number and type of supply chain stakeholders has required consistent and concerted coordination and has highlighted the need to produce accurate forecasts. Forecasts and supply plans are often national in scope, which includes a range of funding sources and procurement mechanisms as well as in-kind donations. The range of stakeholders that must be engaged during the quantification process necessitates significant teamwork and coordination among stakeholders that may not have worked closely together previously and often do not have the human resources required for such coordination.

Funding is not always predictable in terms of levels or the timing of disbursements. Funds for the procurement of commodities, particularly for some donors, can be subject to rapid and/or irregular fluctuations. Relying on a range of sources of funding may entail risk when some are unreliable, or when there are weak systems with no legacy of transparency and accountability. Commitments are not always reflected in disbursements, and funding can be curtailed if and when there are major governance concerns. While many new funding instruments are designed to create a more predictable flow of funding, the risks to a supply chain associated with reliance on donor funding imply a major challenge for forecasting demand.

Available funding does not necessarily translate to available products. Increased funding levels underscore the need for greater capacity in procurement and distribution, improved warehousing and storage procedures, and greater logistics capabilities. Funding mechanisms such as the Global Fund require recipient countries to deliver products quickly, efficiently, and on a large scale. Countries are expected to show measurable results in a short period of time to justify continued disbursements. For example, Global Fund grants are initially approved for five years, but after the first two years of the grant cycle, recipients must demonstrate good performance against targets to continue to receive funds. One way of demonstrating good performance is by measuring forecast accuracy. More accurately estimating the quantities of products to be consumed can increase overall confidence in the system, justifying investments to help strengthen the system overall, and minimizing risks of stockouts or overstocks.

Inaccurate forecasts can have far-reaching consequences in the logistics system specifically, as well as on the overall health situation in the country. If forecasts are not conducted, products may not be procured at all. If products are procured, they may not be the right type, or in sufficient quantities, or in the correct packaging.

Ultimately, clients may not be served. For ARVs, this can mean that a patient's treatment is interrupted, leading to drug resistance. For malaria treatment, this could mean that a patient does not receive the ACTs that are needed, leading to higher morbidity and mortality. For HIV test kits, this may mean that someone who wants to learn his or her status has been denied that service. Additionally, patients and clients have lost time and money because of unsuccessful visits to health facilities. The program can be discredited when timely services are not provided, resulting in program dropouts or increases in disease.

In most developing countries, financial resources are limited. These resources must be used as efficiently as possible, so that products are not oversupplied and consequently expire before they are used. At the same time, supply chain managers work to ensure that quality products are ultimately available. To this end, those conducting forecasts strive to minimize the error between forecast consumption and actual consumption. Accurate forecasting can help to plan for and justify investments.

Objective

Forecast accuracy is a measure of how closely the quantities forecasted to be consumed correspond with the quantities that were actually consumed. Forecasting remains a combination of an art and a science. Actual consumption is *almost always* different from forecast consumption, as it is virtually impossible to determine down to the exact tablet the quantity of a particular medicine that will be dispensed to clients. If the differences between forecasts and actuals are significant, the reasons should be investigated. Monitoring the differences between the forecast and actual consumption is a critical activity that enables program managers to take corrective action swiftly to avoid imbalance of supply. A review of forecasting accuracy may also show patterns in when some products or programs are more accurate than others, which may potentially lead to forecasting and/or review process changes.

This document presents an analysis of forecast accuracy both across commodity groups within a country (ARVs, HIV rapid test kits [RTKs], and select laboratory consumables) as well as across countries. Forecast and actual consumption data were provided from selected countries supported by the U.S. Agency for International Development (USAID) with technical assistance from the Supply Chain Management Systems Project (SCMS) and the USAID|DELIVER PROJECT. The factors, determinants, and levels of accuracy are described.

The purpose of this document is not to show how one country or program can do forecasting "better" than the others; rather, it analyzes different countries' and programs' experiences to identify some determining factors in reducing forecasting error. Specific attention is given to the process through which these forecasting exercises were conducted.

Forecasting Methods

Essentially, two types of forecasting methods exist: the consumption method and the morbidity method. The difference between the two methods is the basis from which the forecast starts. The consumption method uses past consumption data to estimate the future quantities of each product that will be dispensed or consumed during each year of the quantification. No matter which forecasting method is selected, the final outcome will be the quantity of each product expected to be dispensed or consumed during the forecast period.

Consumption Method

The consumption method uses data on consumption of products in the past as a basis for projecting future needs. Estimates of increases or other changes in consumption for each drug during the period of the forecast are based on past trends in consumption. This method requires the availability of data on the quantities of products actually dispensed to patients at service delivery points(SDPs) over a specified period of time. Because consumption must be as accurate as possible, data on quantities dispensed directly to patients are highly preferred over issues data, which reflect quantities distributed from a higher level in the system to a lower level.

Morbidity Method

When using the morbidity method, estimation of ARV drug needs is based on morbidity data on prevalence or incidence of HIV infection or AIDS cases, services data on the number of patients on ART, or patient targets that are then translated into the quantities of drugs that would be needed. This method involves estimating the number of patients expected to start, continue, and stop specific ART regimens during the period of the forecast. This method also requires the availability and use of standard treatment guidelines (STGs), which all health providers must adhere to.

In practice, forecasts may be conducted using two or more types of data and a combination of both forecasting methods. For example, the results of a consumption-based forecast and a morbidity-based forecast may be compared and adjusted to arrive at a best estimate of future commodity requirements. STGs or other policy guidelines should be clearly documented, disseminated, and assumed to be adhered to by all service providers who have been adequately trained. The accuracy of morbidity-based forecasts depends on the degree to which STGs are followed and on the availability of prescribed drugs or commodities.

Tools Used in Forecasting

Forecasting can be a detailed, time-intensive, and error-prone exercise. Calculations may need to be repeated based on revisions to assumptions, constraints in budget, or changes in available data. Several tools exist or have been developed to facilitate the process of determining quantities of products required and to assist in order planning and budgeting. Examples include individually created Microsoft Excel spreadsheets, Quantimed, ProQ, CHAI's pediatric forecasting tool, or other tools developed by field-based staff.

These tools minimize data entry error, improve accuracy by facilitating correct calculations, and offer report analysis. They do not, however, make human decisions, such as choosing which data to use, assessing the quality of the data used, determining which assumptions to make, or calculating final quantities. The assumption in this study is that specific tools do not provide better or worse results, and that forecasting results are very similar with the use of any of the tools to assist in calculating a forecast.

Measuring Forecast Accuracy

Methodology

Accuracy in forecasting is determined by comparing forecasted values (F) with the actual values (A). The forecast error is defined as:

Error = Value of Actual – Value of Forecast, or A – F

Several different measures are used in practice to calculate overall forecast error. For the purpose of this analysis, the Mean Absolute Percent Error, or MAPE, one of the most common measures and easiest to interpret, is used. MAPE is the absolute difference between the forecasted and actual values, expressed as a percentage of the actual values. The MAPE is calculated as:

$$\mathrm{M} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

For the purposes of this paper, we will use the formula below, which does not consider many points over time, just one point in time.

MAPE = |(A – F)|/A

Accuracy is the opposite of error. The closer the error gets to 0 percent, the more accurate the result. The further the error gets from 0 percent, the more inaccurate the result. An example of the calculation is given below.

If forecast is 10 and actual is 8, the MAPE is:
8-10 = -2 or |8-10| = 2

2/8 = 25% or 25% overforecasted

If forecast is 2 and actual is 8, the MAPE is:
8-2 = +6 or |8-2| = 6
6/8 = 75% or 75% underforecasted

For the purposes of this paper, when either actual or forecasted figures were equal to zero, we did not use them as part of the analysis. The analysis of these data would give either an inaccurate picture of the results or an undefined result. An example of each is given below.

Consider the case where there are quantities forecast, but the actual consumption is 0. In these cases, the MAPE cannot be determined. Below is an example where the forecast is 1 and actual is 0:

$$= 0-1 \text{ or } |0-1| = 1$$

$$1/0 = ?$$

Any number divided by a zero is meaningless or undefined and cannot be determined.

If forecast is zero, the MAPE cannot be determined accurately. In this case, there has been nothing forecast, but there is actual consumption. Consider a case where forecast is 0. In the first example, actual is 8, and in the second example, actual is 800:

1. |8-0| = 8
2. |800-0| = 800

In both of these cases, the MAPE is the same. That is:

1. 8/8 = 100%
2. 800/800 = 100%

When the forecast quantity is zero, the MAPE will always be 100%. This, however, does not mean that that the forecast is 100% inaccurate. Rather, the answer does communicate that one forecast is more inaccurate than the other.

Therefore, in cases where the forecast is zero or the actual is zero, an analysis of MAPE cannot be conducted; therefore these cases were excluded from the analysis.

Interpreting the Numbers

Forecast accuracy analysis has been conducted using data from four countries: Country X, Country Y, Ethiopia, and Zambia, from the period 2007–2010. Two of the countries included in the analysis did not want to be named, so they were included as Countries X and Y. For every country presented, at least two forecasting points (the time when the quantification exercise was conducted) were made available to allow for product analysis and accuracy comparison. These countries provided various forecasting methods (consumption or morbidity) and different types of data (for example, issues data or demographic data), so results are not directly comparable across countries. Detailed information on each analysis is provided under each country heading.

Dispensed-to-user data were used instead of issues data for actual consumption. Dispensed-to-user data are a much better reflection of what was used for a particular month (depending on reporting rates), rather than issues, especially since forecasts were determined in most situations on a monthly basis. Issues may include provisions for safety stock and lead time and are inherently more prone to error than dispensed-to-user data.

The analysis was performed product by product, and for most of the analysis, these product-by-product results were weighted equally when determining averages. However, the bulk of products used in countries are first line adult ARVs. Therefore, most of the results do not reflect these differences in quantity and are treated with the same value. In other words, a product with very high consumption quantities would be treated the same as a product with very low consumption quantities. In reality, it may be possible to get a more accurate forecast for products with more consumption and more difficult for the products used less often (that is, pediatric and second line formulations). At the same time, it may make a larger impact on a program to have greater error on products used most than on products seldom used.

To account for all ARV products being given equal weights, a second analysis was performed to relate the MAPE with the quantities of the product used. In this analysis, the MAPE of each ARV was multiplied by the percentage of the total ARVs (packages) consumed. This allows for a review of the impact or burden of the errors. With this analysis, the highest values are attributed to products with a combination of the highest error and quantities consumed, and the lowest value is attributed to the ARVs with a combination of the lowest error and quantities consumed. All others would lie somewhere in between based on their MAPE and volume. When reviewing these figures, the highest values should be the focus of concern and improvement. If no one product has a large proportion of this impact, and all products appear to have the same value, then no one product requires special attention; either they all require special attention or none does.

Since actual and forecasted figures of zero were eliminated from the analysis (see methodology section above), the average accuracies appear better than if we had included products that were not forecasted and consumed, or were consumed but not forecasted.

For ease of discussion, the cutoff of 25 percent error (over- and underforecasted) is used as a measure of being accurate, based on industry practices. In some of the results, products that have been over- or underestimated have been pointed out. For other analysis only the absolute error is discussed. Outliers in some graphs were eliminated in order to better visualize the results of the remaining products. In cases where this occurred, it has been noted.

Finally, the analysis compared product forecasted to product issued or dispensed, not regimen to regimen. A forecast developed as an input into procurement, by necessity, refers to specific formulations (ARVs) to be bought and not to regiments. Therefore, forecast accuracy will decline if fixed-dose combinations (FDCs) were forecasted and single formulations were dispensed or issued instead.

Country Examples

Country X

Forecasting Method:	ARVs: Morbidity method for both adult and pediatric ARVs RTKs: Consumption method
Responsible Party for Forecast:	ARVs: MOH Logistics Unit, MOH Department of Pharmacy Services, JSI-SCMS, and CHAI RTKs: same as above
Frequency of Forecast:	Annually
Frequency of Monitoring:	Quarterly
Monitoring Procedures:	Same as annual forecasts
Members of Monitoring Team:	Same as annual forecasts
Forecasts Used in Analysis:	ARVs: Quarter 4, 2008: reviewed October 2008 to December 2009 March 2009: reviewed March to December 2009 June 2009: reviewed June to December 2009 October 2009: reviewed October to December 2009 RTKs: March 2009: reviewed January to December 2009 July 2009: reviewed July to December 2009 November 2009: reviewed November to December 2009
Type of Data Used for “actuals”:	ARVs: LMIS reports RTKs: LMIS with some extrapolation

Analysis

In Country X, four ARV forecasts and three RTK forecasts were analyzed. The ARV forecasts were developed in October 2008 and March, June, and October of 2009. The RTK forecasts were developed in March, July, and November 2009. Quantimed was used to develop the ARV forecasts and ProQ the RTK forecasts. “Actual” data used for the ARV analysis are dispensed-to-user data from the LMIS reports. Reporting rates for the ARV sites range from 78 percent to 100 percent with an average of about 87 percent. The “actual” data for the RTKs come from LMIS reports, but since there was a system redesign for the distribution of RTKs in 2009, reports were not available for all months, and extrapolation was used for a few months.

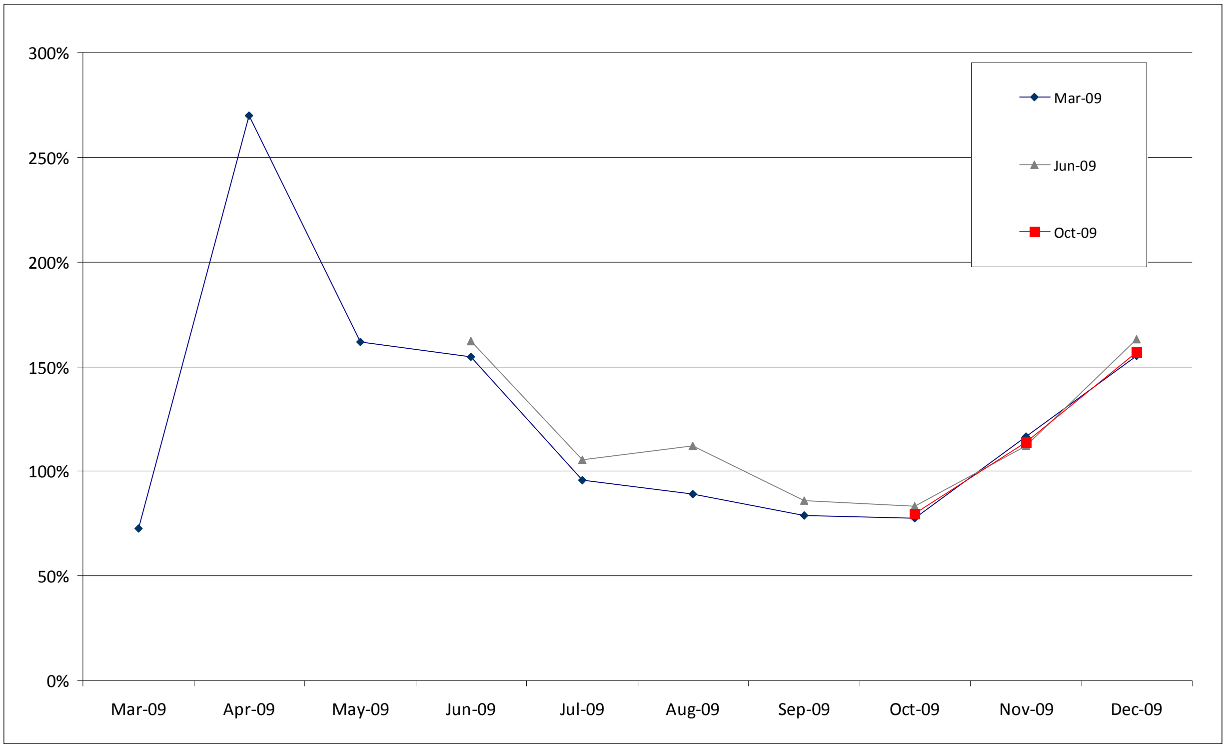
Upon review of several forecasting reports for Country X, it appears that the team responsible for forecasting uses all available data to come up with safe assumptions. The following reports are used: HIV Program Coordinator’s Annual Report, LMIS data, M&E Progress Reports, National Stock status, and supplier’s shipment records. For example, LMIS reports are compared with consumption records as well as M&E Progress Reports to determine patient enrollments and rate of scale-up. Not only does this provide a better sense of what is happening in country and greater ability to make better assumptions, this type of review also provides an opportunity to check the quality of the data from various sources, which in turn initiates support and supervision visits, revision of LMIS tools and registers, and training to increase the quality of data for future forecasts.

ARV Findings

Error increases due to prescribing practices not changing as anticipated. Due to extreme errors in all forecasts for two pediatric products (3TC/ZDV 30/60 and 3TC/ZDV/NVP 30/60/50), these outliers were removed from the analysis. The error was a result of two factors: first, the assumption that the two products would be available in-country earlier than they actually became and, second, that once they were made available, the prescribers would change their prescribing practices and prescribe this combination according the STGs. Unfortunately, prescribing practices did not change as predicted, and assumptions in January 2009 that first line pediatric ARVs should be split between D4T-based regimens at 73.5 percent and AZT-based regimens at 26.5 percent were revised downward in the January 2010 forecast to D4T-based regimens at 93.3 percent and AZT-based regimens at 6.7 percent to better reflect the reality of the situation.

The MAPEs for three of the four forecasts have been plotted in figure 2 below. The forecasts for Q4, 2008 were developed using quarterly forecasts for the adult ARVs and were not comparable with the rest of the 2009 forecasts on a monthly basis. This forecast was left out of the graph below. As seen in figure 2, forecasts for Country X do not appear to improve on average with each monitoring, and it appears that forecasts do not change significantly with each quarterly forecast, as the MAPEs in figure 2 tend to follow the same curves. Also, the first two forecasts appear to mostly improve with time—that is, the error rate decreases as time goes on. This is different from the cases of Zambia and Country Y, but similar to Ethiopia. Later in this section, when reviewing accuracy by product, it becomes apparent that the number of ARVs forecasted accurately does, indeed, increase with subsequent forecasts (table 3). One possible reason for the contradiction between these two results may be the extent to which the products that were forecasted inaccurately were inaccurate, thereby skewing the results of the averages seen below. The larger forecasting inaccuracies are seen mainly in the pediatric formulation.

Figure 2. MAPEs for Three Separate Forecasts in Country X



Data quality can impact forecasting accuracy. Another reason for the lack of overall improvement in forecasting may be attributed to data quality and how the number of patients assumed to be on treatment and rate of scale-up are determined for the ARV forecasts. Since there have been frequent discrepancies with the number of patients on treatment in the LMIS report and the M&E reports, the average of both is taken for the pediatric forecasting exercise. However, if these figures are very far from each other, and if one or both are greatly inaccurate, the average would be as well. This means that the patient numbers on treatment from the beginning of the forecast are inaccurate, making most or all of the subsequent forecasts inaccurate from that point as well.

The high error at the beginning of the March and June 2009 forecast is attributed to errors in several pediatric formulations, but mainly ZDV 100mg and ddI 25mg. These are further discussed below.

Greater error is seen with forecasting pediatric products than with adult products. Figures 3, 4, 5, and 6 (below) illustrate the MAPEs for each individual forecast as well as the differences between the adult and pediatric ARV forecast MAPEs. Like most other countries reviewed in this report, there is greater error when forecasting pediatric formulations than there is with adult formulations. And again, in Country X, as with other countries, pediatric formulations are being overforecasted. Several issues regarding pediatric forecasting are discussed in the quantification reports by Country X. Some of this error is attributed to children on adult formulations being counted as adults and to children who could have been switched to higher-dose pediatric formulations (when made available) but were not.

Collecting data on pediatric weights may increase forecasting accuracy. Country X does collect data on weights of pediatric patients on the LMIS reports and reviews these percentages as well as consumption patterns, with each monitoring/forecast. This has likely attributed to their increase in forecasting accuracy with subsequent pediatric ARV forecasts.

Figure 3. MAPEs for Country X Q4 2008 Forecast

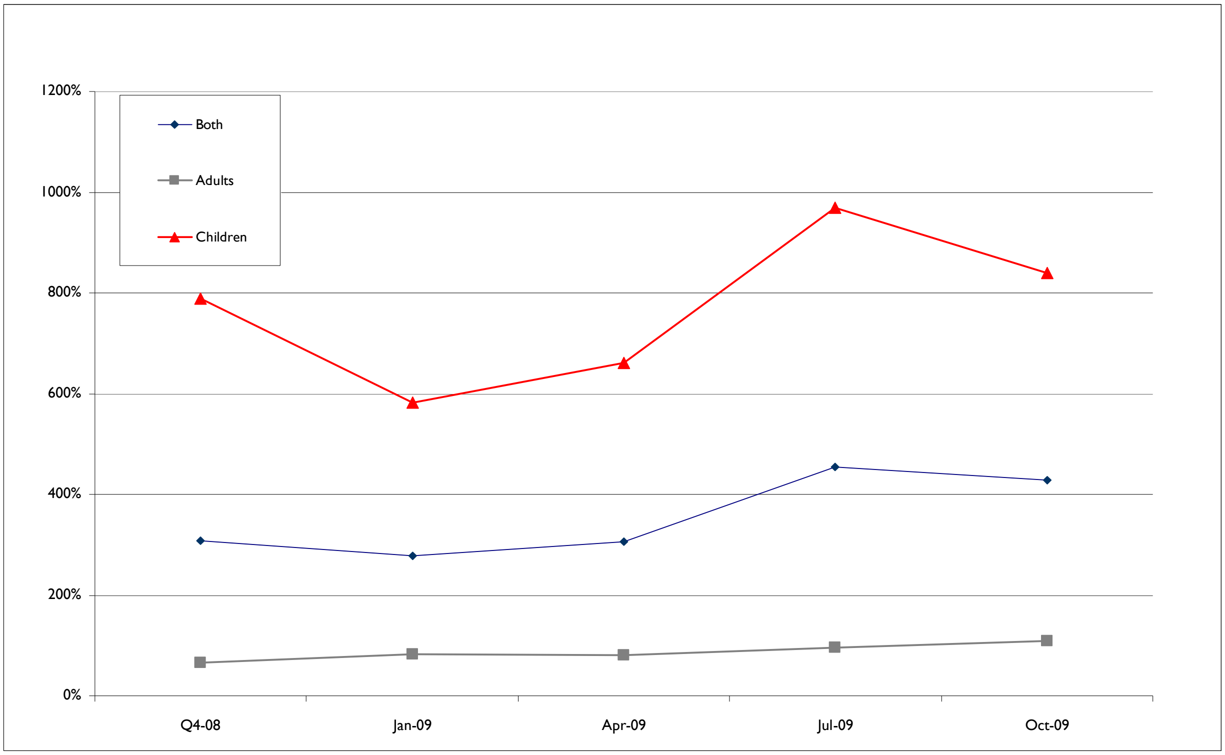
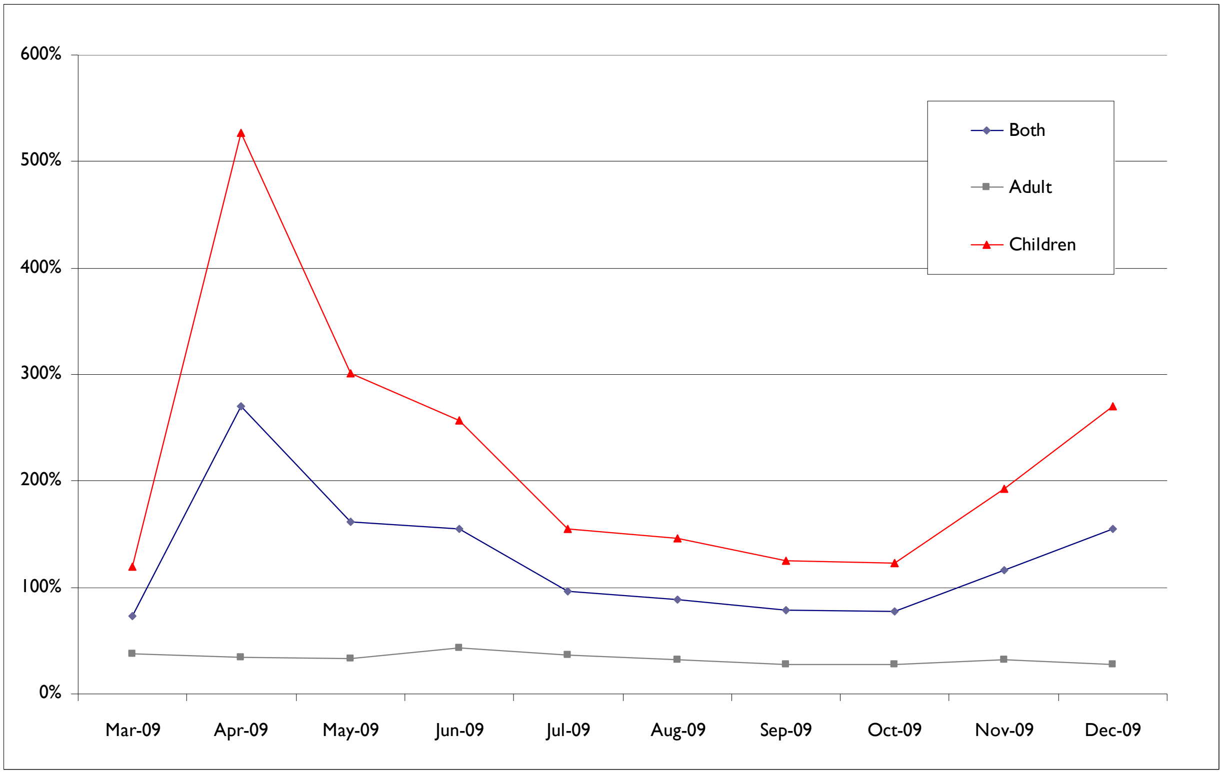
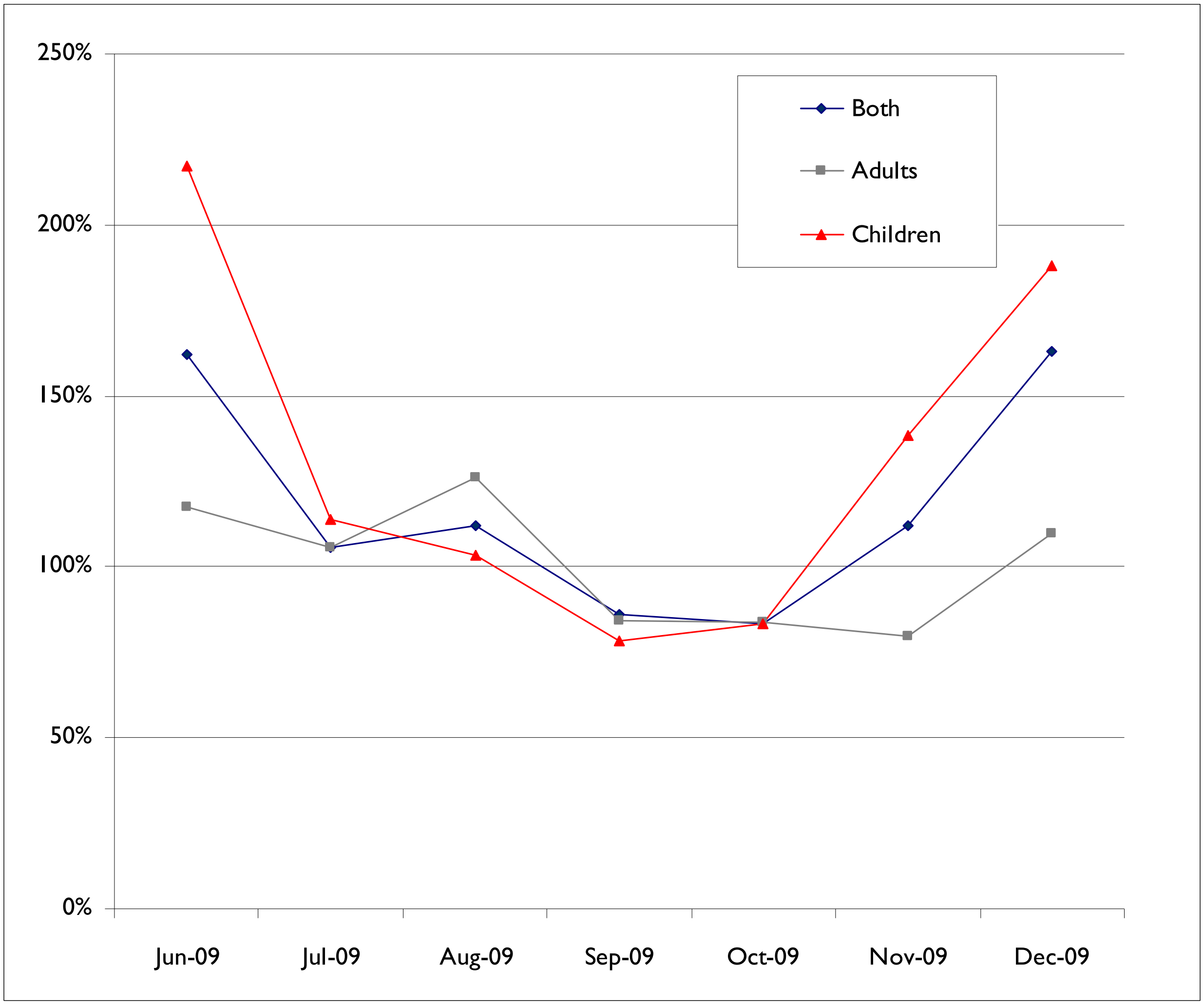


Figure 4. MAPEs for Country X March 2009 Forecast



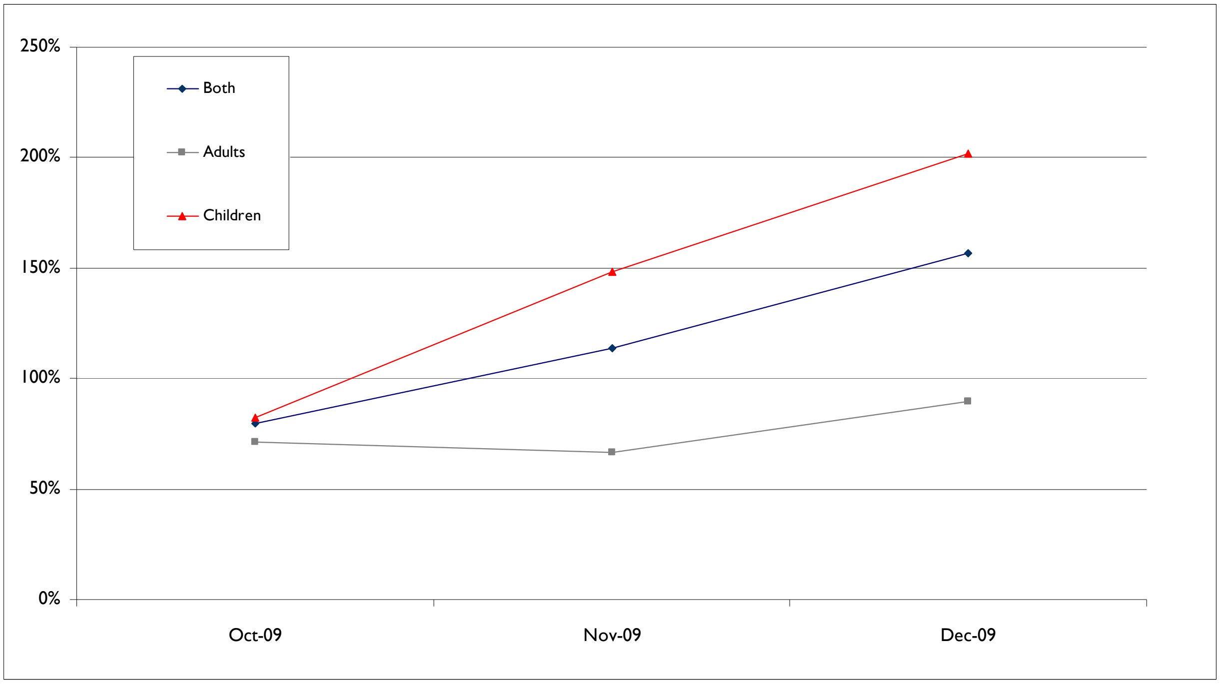
In figure 4, the high peak seen early in the forecast represents the high errors seen in the ZDV 100mg and ddI 25mg forecasts. These are both being overforecasted, with prescribing practices being below what was assumed. ZDV was forecasted for very small quantities (48 units each month) in the reviewed reports and is prescribed at about one-tenth of the percent assumed to be needed. This error, as well as the ddI 25mg error, may be a result of inaccurate assumptions on which formulations are used for which child weight range and/or an overestimation of prescribers prescribing a first line AZT regimen or second line regimens.

Figure 5. MAPEs for Country X June 2009 Forecast



The pediatric ARV, ddI 25mg, continues to attribute to the high error in June 2009 forecast.

Figure 6. MAPEs for Country X October 2009 Forecast



Forecasting error increases from the time of forecast. Figure 5 illustrates a trend in Country X that is similar to the ones seen in the other countries discussed in this paper. The MAPEs increase with time, and pediatric forecasted MAPEs are higher than those for MAPEs of adult forecasts. That is to say, forecast accuracy decreases with time from the point when the forecast was made.

In tables 1 and 2, the adult and pediatric ARVs have been grouped separately, and the error is presented without taking its absolute, revealing products accurately forecasted (within 25 percent of the actual), underforecasted (more than 25 percent), and overforecasted (less than -25 percent).

More adult ARVs are forecasted accurately than are pediatric ARVs. The difference between adult forecasting accuracy and pediatric accuracy becomes visually apparent again in these two tables, as it was in figures 3, 4, 5, and 6. More ARVs are forecasted accurately in the adult table (table 1) within the first three months following the forecast, than those forecasted for the pediatric formulations (table 2).

Table 1. MAPEs for adult ARVs for the first three months following forecasts in Country X

Product Name	Q4 2008	Mar 2009	Jun 2009	Oct 2009
3TC/ZDV 150/300mg	81%	-28%		
LPV/r 200/50mg	65%	28%		
NVP 200mg	57%	-35%	-667%	
ZDV 300mg	86%	5%	-685%	-435%
3TC/ZDV/NVP 150/300/200mg				3%
EFV 600mg	60%	-54%	-10%	6%
ABC 300mg	61%	55%	24%	-14%
3TC/D4T/NVP150/30/200mg	57%	-37%	-25%	-20%
3TC/D4T 150/30mg	38%	-5%	24%	98%
EFV 200mg	-131%	-11%	51%	25%
ddI 250mg, delayed-release	60%	-8%	-35%	-60%
ddI 400mg, delayed-release	-17%	-54%	-133%	-77%
TDF/3TC 300/300mg	29%	-74%	0%	-17%

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

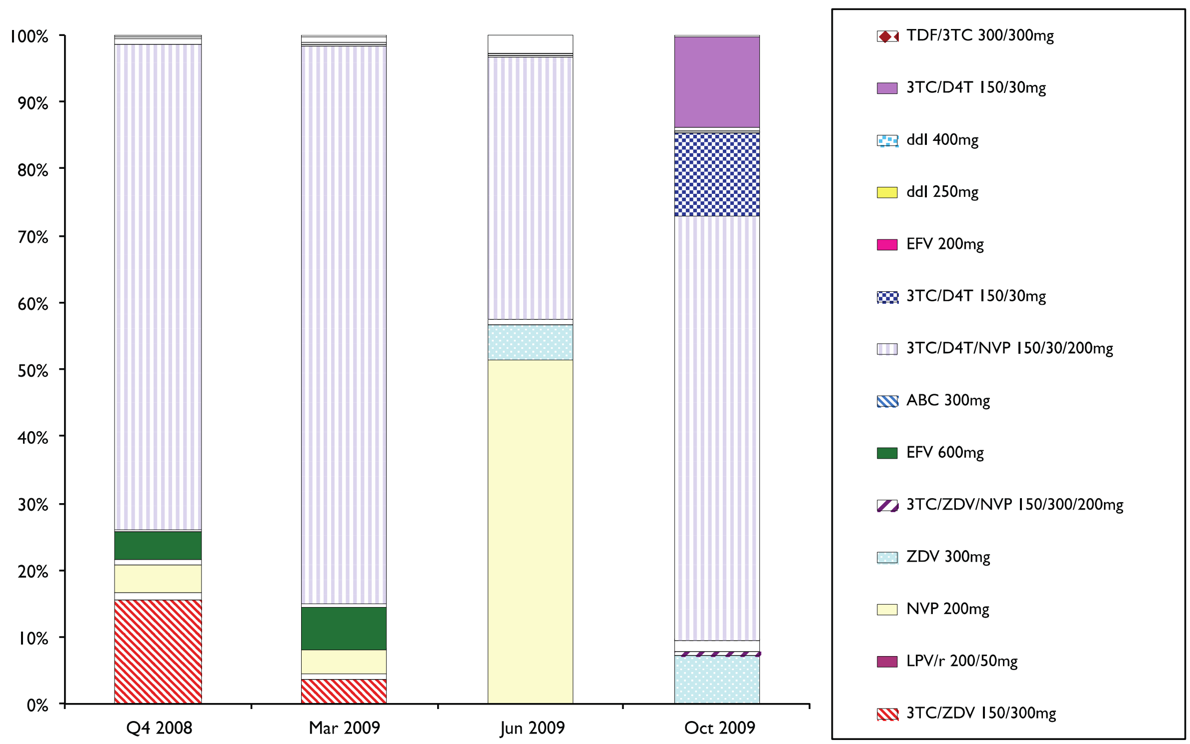
Table 2. MAPEs for pediatric ARVs for the first three months following forecasts in Country X

Product Name	Q4 2008	Mar 2009	Jun 2009	Oct 2009
NVP 10mg/ml			97%	98%
ZDV10mg/ml		93%	53%	39%
ZDV 100mg		-746%	-84%	-173%
ABC 20mg/ml	-80%	77%	44%	-15%
LPV/R 80/20mg/ml		98%	99%	99%
3TC/D4T/NVP 30/6/50mg		-218%	-36%	-31%
3TC/D4T/NVP 60/12/100mg	-145%	-3%	-46%	-10%
EFV 50mg		-217%	-50%	-155%
ddI 25mg	-3660%	-2227%		
ddI 50mg	-413%	-106%	-494%	-1021%
3TC/D4T 60/12mg	-148%	-66%	-15%	-31%
3TC/D4T 30/6mg	-288%	-128%	-34%	-11%
3TC/ZDV 30/60mg		-1943%	-368%	-6481%

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

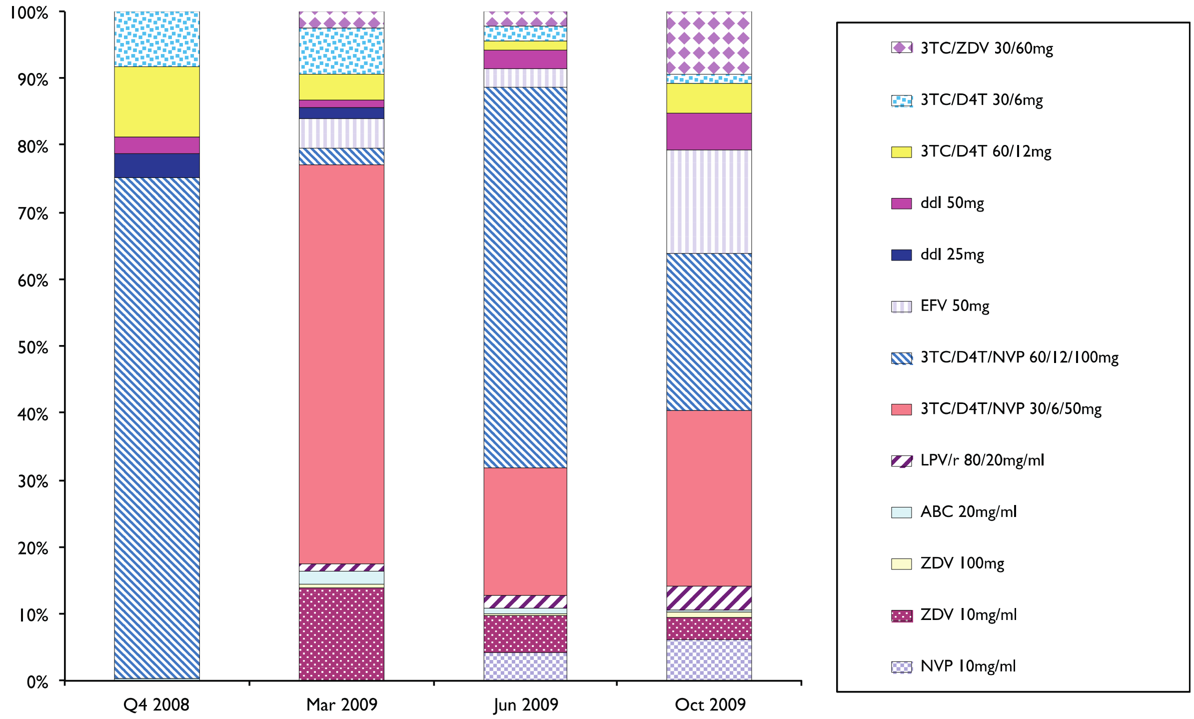
In figures 7 and 8, the impact of the errors is revealed. The adult and pediatric MAPEs from tables 1 and 2 have been multiplied by the percentage of the total ARVs consumed to reveal the impact of the forecasting errors. The highest values for each forecast are attributed to products with a combination of the highest error and quantities consumed, and the lowest values are attributed to the ARVs with a combination of the lowest error and quantities consumed. When reviewing these figures, the highest values should be the focus of concern and improvement.

Figure 7. Impact of Adult ARV MAPEs in Country X



The impact of MAPE for adult ARVs is highest for those consumed the most, even when MAPE is relatively benign. Although the MAPEs for the 3TC/D4T/NVP 150/30/200 ranged from -37 percent to 50 percent with two quarters within 25 percent in table 1, this product has the highest proportion in three of the four forecasts in figure 7. This is because this product was consumed at seven to10 times the rate of the second-highest consumed product. Although this trend may not continue in the future due to changes in D4T recommendations, Country X would want to concentrate on improving forecasts for this FDC. Attention should also be paid to NVP 200mg, which has the highest impact of error in the June 2009 forecast. The MAPEs for NVP was -667 percent, very comparable to the poor MAPE for ZDV 300 of -685 percent, but the quantities for NVP 200mg consumed for the first three months after the June 2009 forecast were a bit more than 10 times higher than those for ZDV 300mg. This is why the proportion or impact of MAPE is higher for NVP 200mg than it is for ZDV 300mg, and greater effort should be made to improve the NVP 200mg forecast.

Figure 8. Impact of Pediatric ARV MAPEs in Country X



The impact of pediatric MAPEs appears to be highest for the two pediatric FDCs, 3TC/D4T/NVP 60/12/100 and 30/6/50. The MAPEs for these products ranged from -3 percent to -218 percent, with a median MAPE of -36 percent, which is not too far from accurate. The reason for the high impact attributed to these products in figure 8 is that they comprise the bulk of the pediatric formulations consumed (from 50 percent to 89 percent of all packages consumed). Greater effort should be made to improve the forecast for these two pediatric products.

Table 3. Accurate, underforecasted, and overforecasted ARVs for the first three months following forecast in Country X

Forecast	within 25% error Accurate		> 25% error Underforecast		< -25% error Overforecast		Total ARVs
	Number	%	Number	%	Number	%	
Quarter 4 2008	1	6%	10	56%	7	39%	18
March 2009	5	20%	5	20%	15	60%	25
June 2009	6	26%	5	22%	12	52%	23
October 2009	9	39%	4	17%	10	43%	23

The number of ARVs accurately forecasted improves with each subsequent forecast. As discussed in figure 2 above, although there wasn't an overall average improvement in forecasting errors for each consecutive forecast, as seen in table 3, there are measurable improvements in the number of ARVs forecasted accurately with each subsequent forecast in Country X. Accurately forecasted ARVs increase from 6 percent of products to 39 percent of products. The percentage of products underforecasted decreases, and the percentage of products overforecasted averages around 50 percent. There appears to be a shift toward overforecasting rather than underforecasting, which makes sense, considering the nature of the product.

ARV Conclusion

The achievements in forecasting accuracy can be attributed to:

- Experienced staff, institutionalization, and a mature program
- Rigorous review process, consulting multiple data sources
- Well-developed LMIS and resulting quality data availability

The products that did have high MAPEs may have been subject to the following:

- Delays in implementation of policy changes
- Prescribing practices not changing as anticipated
- Difficulties in forecasting pediatric ARVs
- Discrepancies in available data for use in forecasting

RTK Findings

All three forecasts shared covered long periods of time and did not result in monthly forecasts as were done for the ARVs. Two of the forecasts (March 2009 and July 2009) were for a total of 12 months, and one forecast was for a total of seven months (November 2009). This made analysis by a monthly basis difficult, and some issues data had to be divided equally by the number of months covered. Also adding to the difficulty of analysis was that "actuals" data were incomplete for 2009, as the distribution of RTKs and the LMIS was in transition from one system to another.

Having three separate forecasts within one year does show that monitoring is occurring and that forecasts are being adjusted.

Table 4. RTK MAPEs for the first three months following forecast in Country X

Product	March 2009 Forecast	July 2009 Forecast	November 2009 Forecast
Determine chase buffer	0%	65%	37%
Determine 100	8%	53%	20%
SD Bioline 25	3%	55%	17%

From table 4 it is evident that, for the most part, forecasts are fairly accurate within the three months following the forecast. Although it is not possible to draw any other conclusions from these limited data, there does appear to be less error with Determine tests, which are used as the screening test than there is for the Determine chase buffer and SD Bioline, which have much less-predictable patterns of use. We might be able to assume from this that Determine chase buffer has errors primarily due to health worker behavior—that is, not using the amount of chase buffer required accurately. For SD Bioline, the higher error may be because it is used as the confirmatory test, which has a more unpredictable use pattern than does a screening test.

Country Y

Forecasting Method:	ARVS: Morbidity method Labs: Consumption method
Responsible Party for Forecast:	CMS with technical assistance from SCMS and participation of CHAI
Frequency of Forecast:	Quarterly
Monitoring Procedures:	Generally monthly to quarterly monitoring; however, due to funding delays, regimen changes, and looming gaps, forecasts were repeated on a monthly to quarterly basis during 2009 and much of 2010. Monitoring method and participants are as stated above.
Forecasts Used in Analysis:	ARVs January 2007: reviewed March 2007 to January 2009 December 2007: reviewed January 2008 to September 2009 December 2008: reviewed January to September 2009 Laboratory Reagents March 2009: reviewed April to November 2009
Type of Data Used for "actuals":	LMIS data

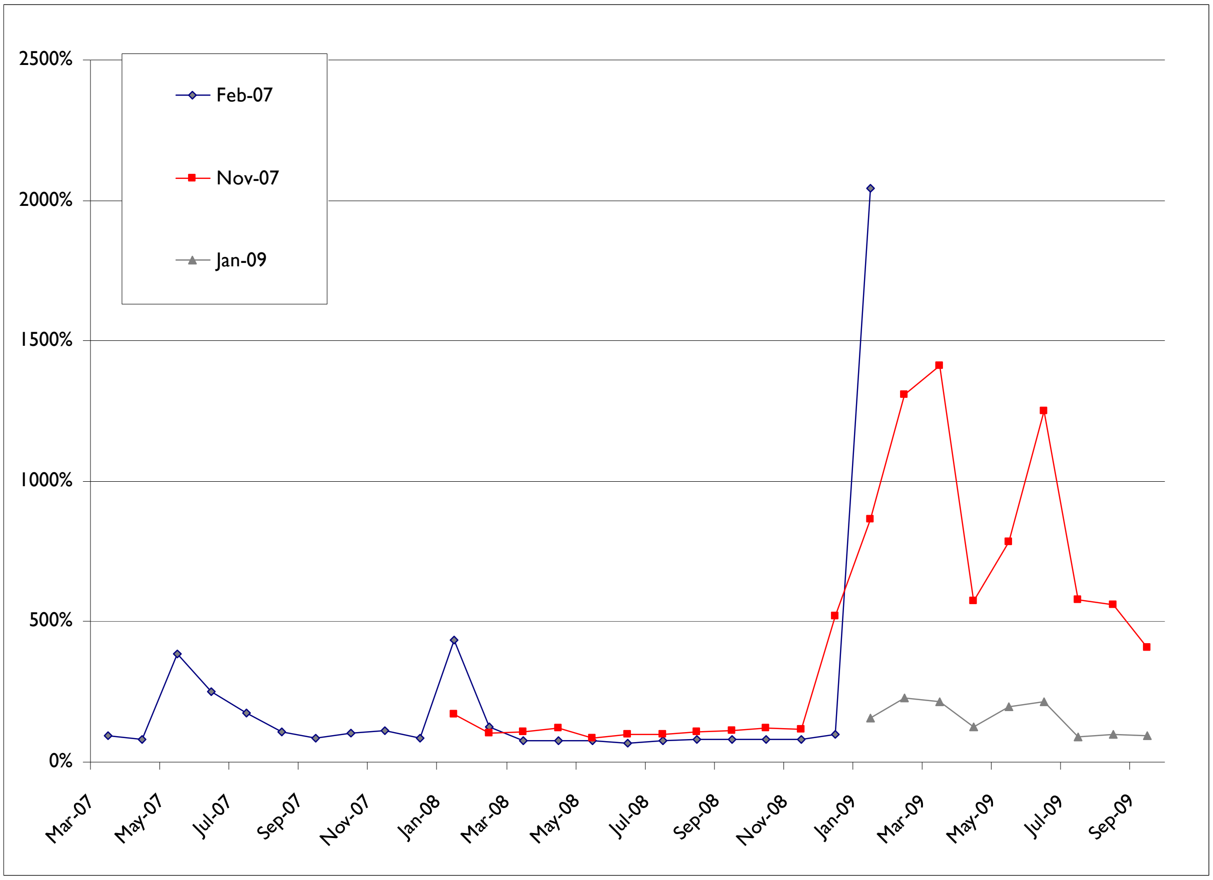
Analysis

In Country Y, three separate national ARV forecasts and one laboratory forecast were shared and analyzed. For ARVs, the forecasts were developed in January 2007, December 2007, and December 2008; for laboratory commodities, the March 2009 forecast was shared. The ARV forecasts were originally performed using a country-developed database tool, and the monthly forecasts were exported to Pipeline. "Actual" data represents dispensed-to-user monthly data from the LMIS reported by the SDPs and imported into PipeLine. Reporting rates for the LMIS reports were documented to be 100 percent of facilities for the beginning of the analysis period as there were few sites, but the reporting rates decreased as the number of sites increased.

ARV Findings

Errors increase around January 2009. The MAPEs for all three forecasts have been plotted in figure 9 below. Unlike most other countries discussed in this paper, forecasts do not improve with time. The errors sharply increase with the first two forecasts around January 2009, and the third forecast of January 2009 does not return to a low accuracy of the first two forecasts. Upon review of the country situation, there appear to be several reasons for this sharp increase in forecasting error during this time.

Figure 9. MAPEs for Three Separate Forecasts in Country Y



Rapid decentralization without planning increases forecasting error. The year 2007 was significant in the ramping up of ART services via the opening of specialized ART sites at existing public health facilities. A vertical LMIS for ARV medicines supported the program. Toward the end of 2008, the MOH decided that, to be able to continue to scale up, it needed to mainstream HIV care and treatment throughout public health services and to decentralize HIV care and treatment to the district level nationally and to the larger health centers in urban areas. The implementation of integration and decentralization happened very rapidly with the existing human resources in the system and without much orientation, training, or follow-up support.

During this time of change, various outcomes affected the logistics that supported the system. A scale-up from 50 sites to over 300 was undertaken with very little training in ARV LMIS or resupply; subsequent reporting decreased from 100 percent to about 85 percent, as did data accuracy. There was little laboratory and new provider capacity to follow treatment guidelines for patients with co-morbidities and/or those who had experienced side effects on particular ARVs, or those needing second line. Therefore, the bulk of patients remained on first line. The ratio of regimens prescribed did not reflect the forecast assumptions of informed clinicians and program managers. Before scale-up and decentralization, the number of LMIS reports and a push system were more manageable at the central level, and report feedback and follow-up was possible. Without increasing human resources at the central level to support this expansion, existing capacity was strained, and the quality of data for forecasting decreased.

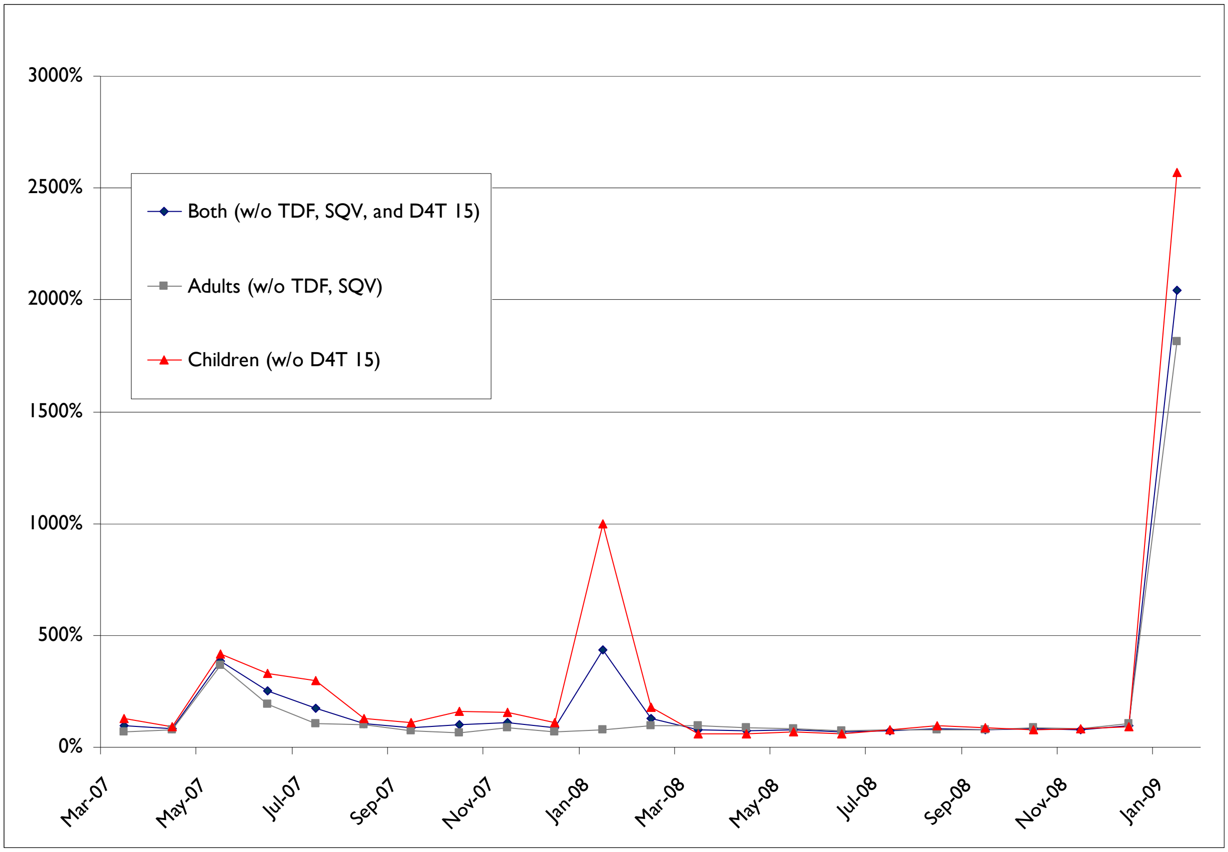
Improper storage and inventory practices may impact forecasting error. During this time, significant changes also occurred at the central warehouse level. Warehousing was subcontracted until the end of 2007. In 2008, new MOH staff with little experience in warehouse management became responsible for this operation. After a very difficult period for central warehouse management, in 2009 and 2010 staff received significant technical assistance in warehouse management systems, processes, and tools.

Unpredictable disbursement of funds may impact forecasting error. Disbursement of GFATM grants has been a significant problem in Country Y, reaching crisis proportions in 2009. Before then, CMS, with the assistance of SCMS, had been able to work with the USG and CHAI to close shipment gaps and keep the country stocked at desired levels. At this time, the country adopted a new first line and started planning the clinical and logistics phaseout of D4T-based first line, and phasein of AZT-based first line. A more effective PMTCT regimen was also adopted and implementation began immediately, with no coordination with the supply chain. Repeated delays in disbursements affected procurement and supply to the country. Over a period of many months, the supply chain and prescribers could only deliver existing stocks, which created a domino effect that resulted in shortages of one and then another ARV as people had to distribute and use what they had available. First, uptake of the new regimens was delayed, so stocks of the previous first line were used much quickly than forecasted. When a shipment of the new regimen arrived, clinicians had to shift everyone to the medicines that were available. Instead of a well-planned and coordinated phasein of the new regimen, change was abrupt and haphazard as people had to use the medicines that were available. The supply chain can only support the ART and PMTCT programs if the issues surrounding the disbursement of GFATM funds are resolved to allow for timely procurement, importation, and distribution. Prolonged funding crisis management, as experienced in Country Y, wreaks havoc with forecast accuracy.

Other contextual factors complicated supply chain management in Country Y, which, although not as significant as those above, still have an impact. Record high international fuel prices in 2008 resulted in annual transport budgets being spent by July. Furthermore, the national airline decreased cargo capacity. These two changes meant that Central Medical Stores needed to ship by road. Instead of resupply being monthly basis, at times it was as long as quarterly.

The simultaneous significant challenges described above resulted in the dramatic increases in MAPEs seen in figure 10. Figures 11, 12, and 13 below show the MAPEs for each of the three separate forecasts and the breakdown of these errors by adult and pediatric formulations. Like most other countries reviewed in this report, there is greater error when forecasting pediatric formulations than there is with adult formulations.

Figure 10. MAPEs for Country Y February 2007 Forecast



In figure 10, several large outliers needed to be removed from the chart to better visualize the data. These were TDF 300mg, SQV 200mg, and D4T 15mg, all of which were greatly overestimated in the February 2007 forecast. Since these products were consumed in low quantities during this period, the impact of the error is low. Further discussion of the impact of the MAPEs follows with figures 14 and 15.

Figure 11. MAPEs for Country Y November 2007 Forecast

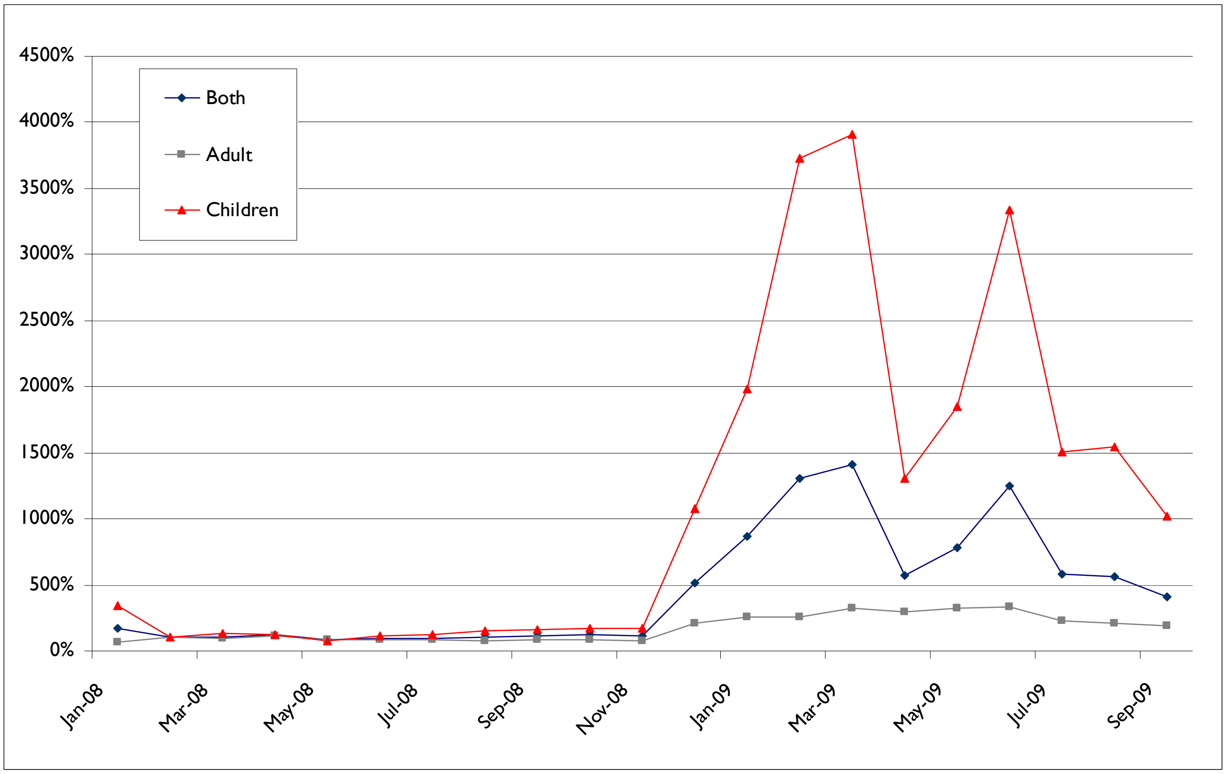
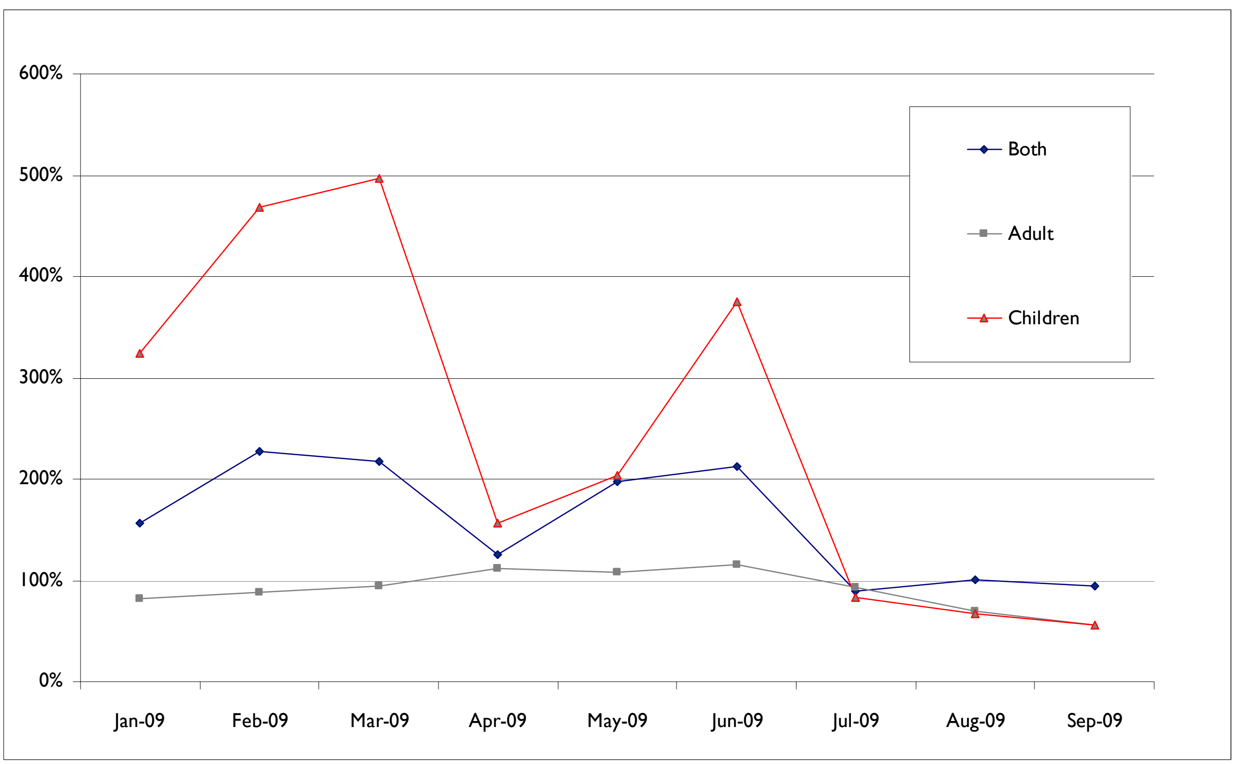


Figure 12. MAPEs for Country Y January 2009 Forecast



Greater forecasting error is experienced with pediatric ARVs than with is with adult ARVs. As well as seeing the rise in error in early 2009 in the three figures above, it becomes obvious that forecasts for pediatric ARVs have greater MAPEs than do the forecasts for adult ARVs. The same forecasting method was used for both the adult and pediatric ARVs. Most of these errors for the pediatric formulations are overestimates, as is visible when comparing tables 5 and 6 below.

In tables 5 and 6 the adult and pediatric ARVs have been grouped separately and the error is presented without taking its absolute, revealing products accurately forecasted (within 25 percent of the actual), underforecasted (more than 25 percent) and overforecasted (less than -25 percent).

Table 5. MAPEs for adult ARVs for the first three months after forecast in Country Y

Product Name	Mar 2007	Q1 2008	Q1 2009
ABC 300mg 60 Tabs	57%	-64%	-12%
ddI 250mg, delayed-release, 30 Caps	35%	-76%	-76%
ddI 400mg, delayed-release, 30 Caps	7%	-101%	-105%
EFV 600mg 30 Tabs	-75%	-76%	-41%
IDV 400mg 180 Caps	-48%	-51%	-32%
3TC/3TC/ABC 150/300/300mg 60 Tabs	-181%	-38%	
3TC 150mg 60 Tabs	96%	-101%	-18%
3TC/D4T 150/30mg 60 Tabs	-23%	-53%	
3TC/D4T/NVP 150/30/200mg 60 Tabs	-39%	-7%	-24%
3TC/D4T/NVP 150/40/200mg 60 Tabs	-27%		
3TC/ZDV/NVP 150/300/200mg 60 Tabs	-260%	-40%	
3TC/ZDV 150/300mg 60 Tabs		5%	-35%
Ritonavir 100mg 84 Caps	-1304%	-2%	
SQV 200mg 270 Caps	-4069%	-393%	-56%
D4T 30mg 60 Caps	63%	-113%	-150%
TDF 300mg 30 Tabs	-11578%	-116%	-241%
3TC 300mg 60 Tabs		-96%	
LPV/r 200/50mg 120 Tabs		-187%	

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

Pressure to scale up may increase forecasting error in favor of overforecasting. Almost all MAPEs for quarter 1 (Q1) 2008 forecast in table 5 above are in favor of overforecasting. The pressure to forecast according to targets was high at the beginning of the year due to a big push to scale up. The risk of overforecasting, and then overordering, was mitigated with the quarterly updates. Supply plan orders were only for the next six to nine months, thus enabling the program to have stock if, indeed, it reached its goals, or the shipment plans to slow down toward the end of the year, if monitoring showed the targets were not going to be met. In August 2008, the MOH ended up scaling back its targets to more realistic ones. But at the time of the initial forecast in January 2008, the official targets had to be used.

Table 6. MAPEs for pediatric ARVs for the first three months after forecast in Country Y

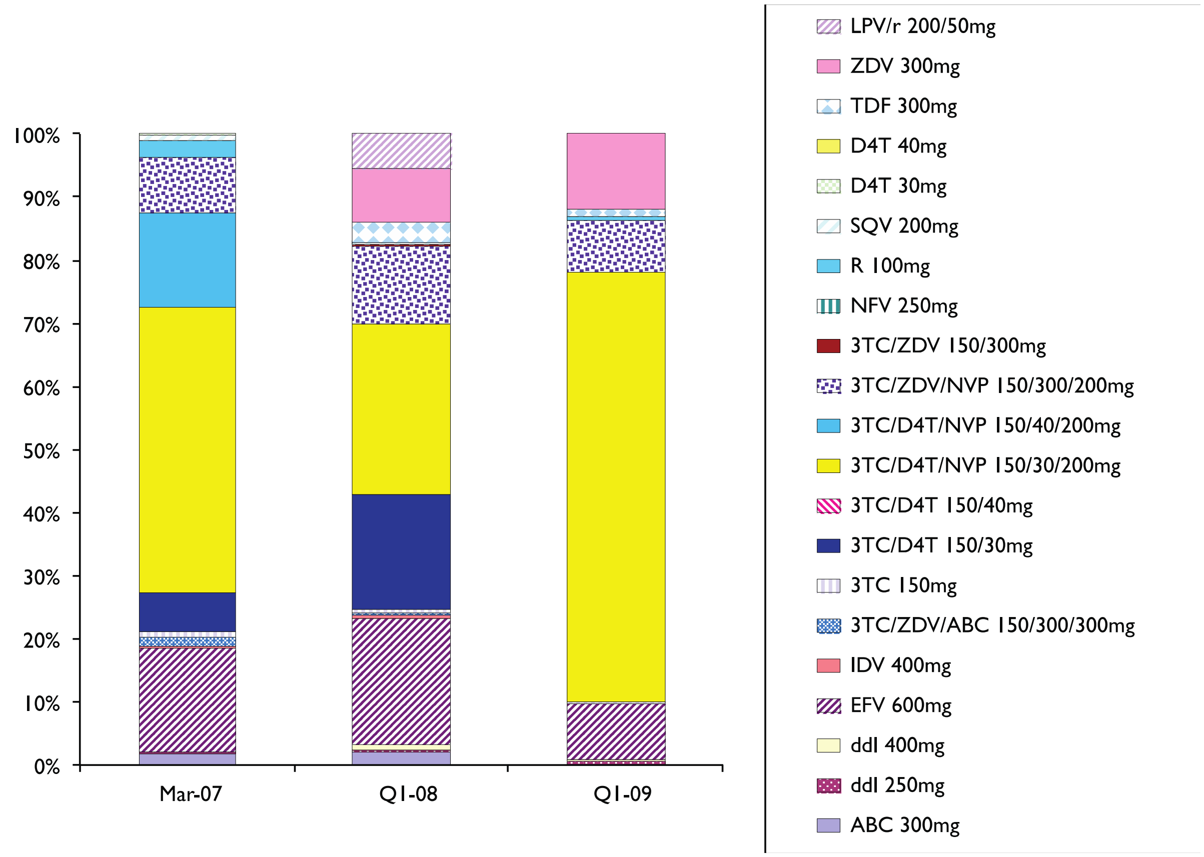
Product Name	Mar 2007	Q1 2008	Q1 2009
ABC 20mg/ml, bottle of 240ml	-32%	-97%	
ddI 50mg 60 Tabs	80%	-496%	
EFV 50mg 30 Caps	-2268%	39%	-1900%
EFV 200mg 90 Caps	44%	-125%	33%
3TC 10mg/ml 240ml	-69%	37%	-180%
3TC/D4T/NVP 20/5/35mg 60 tabs		-172%	
NVP 10mg/ml 240ml	56%	67%	
D4T 15mg 60 Caps	-7215%		22%
D4T 1mg/ml, 200ml	-387%	-2175%	
3TC 10mg/ml 240ml	-541%	-67%	
LPV/r 80/20mg/ml 60 ml	46%	12%	

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

Pressure to forecast a higher number of pediatric ART patients may increase forecasting error. One of the reasons for overestimating pediatric ARVs (as also seen in Zambia and discussed later) is the pressure to forecast based on an increased number of pediatric patients starting on ARVs without implementation of these policies. Although the forecasts for pediatric ARVs are overestimated, the team in Country Y reviews and monitors consumption patterns to manage procurement planning of these products to limit overstocking and expiration of pediatric products.

In figures 13 and 14, the impact of the errors is revealed. The adult and pediatric MAPEs from tables 5 and 6 have been multiplied by the percentage of the total ARVs consumed to reveal the impact of the forecasting errors. The highest values for each forecast are attributed to products with a combination of the highest error and quantities consumed, and the lowest values are attributed to the ARVs with a combination of the lowest error and quantities consumed. When reviewing these figures, the highest values should be the focus of concern and improvement. If no one product has a large proportion of this impact, and all products appear to have the same value, then no one product requires special attention; either they all require special attention or none does.

Figure 13. Impact of Adult ARV MAPEs in Country Y



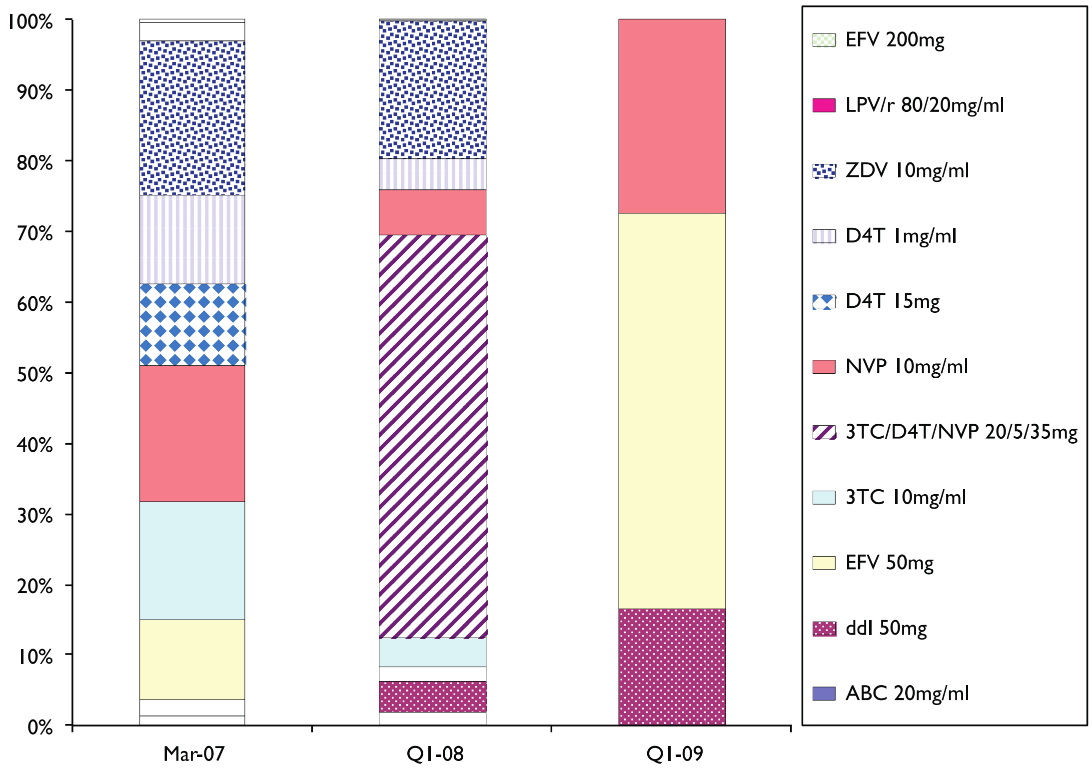
The impact of MAPE for adult ARV is highest with those consumed the most, but products with very large MAPEs and low consumption also have a great impact. Although the MAPEs for the 3TC/D4T/NVP 150/30/200 in the March 2007 forecast is only 39 percent, this product represents 46 percent of all products consumed, and this is the reason for its large share of the impact.

Again in the quarter 1 2008 forecast, the three products with the largest impact comprise the bulk of consumption, although their forecasts were fairly accurate. The forecast for 3TC/D4T/NVP 150/30/200 was more accurate this time with a MAPE of 7 percent but remained the highest consumed product for that quarter, with 74 percent of all ARVs consumed. The other three products, 3TC/D4T 150/30, EFV 600mg, and 3TC/ZDV/NVP had higher MAPEs of 53 percent, 76 percent, 40 percent, respectively, but also had the three next most-consumed products.

And yet again, the largest impact of forecasting accuracy is with 3TC/D4T/NVP 150/30/200, again the largest-consumed product in the Q1 2009 forecast, as it represented 74 percent of ARVs consumed.

Although SQV and TDF have the highest MAPEs in table 5, they are barely represented in figure 13 when their relative volume is taken into account. To improve forecast accuracy, Country Y should concentrate on improving forecasts for the highest-consumed products.

Figure 14. Impact of Pediatric ARV MAPEs in Country Y



The impact of MAPE for pediatric ARV is highest with those consumed the most, but products with very large MAPEs and low consumption also have a great impact. No one product appears to have a large impact in the pediatric forecasts for March 2007. They all appear to share in the burden of forecasting accuracy. The triple FDC of 3TC/D4T/NVP 20/5/35 appears to have the largest impact in the Q1 2008 forecast. It had a MAPE of 172 percent and comprised of 36 percent of the pediatric ARVs consumed. The second-most-consumed product, at 32 percent, had a lower MAPE of 6 percent and therefore a lower impact, but it remains the product with the second-highest impact. In the Q1 2009 forecast, the product with the highest MAPE had the highest impact. EFV 50mg had a MAPE of 1,900 percent and represented only 8 percent of ARVs consumed. Country Y should concentrate on improving forecasts for the highest-consumed products and on decreasing very large MAPEs on less-consumed products.

Table 7. Accurate, underforecasted, and overforecasted ARVs for the first three months following forecast in Country Y

Forecast	within 25% error Accurate		> 25% error Underforecast		< -25% error Overforecast		*Total Number of ARVs*
	Number	%	Number	%	Number	%	
March 2007	2	8%	9	35%	15	58%	*26*
January 2008	4	14%	4	14%	20	72%	*28*
January 2009	4	27%	2	13%	9	60%	*15*

Unfortunately, what becomes evident in tables 4, 5 and 6 is that forecasting in Country Y does not improve with time. Also, as seen in table 7, the count of items accurately forecasted increases from only two in March 2007 to four in January 2008 and January 2009. The percentage of accurate products may only increase because the data for available for analysis were lacking for products forecasted in January 2009. Possible reasons for this decrease in forecasting accuracy are summarized below.

ARV Conclusion

As explained above, the increase in forecasting error as represented by MAPE is attributed to the following major changes:

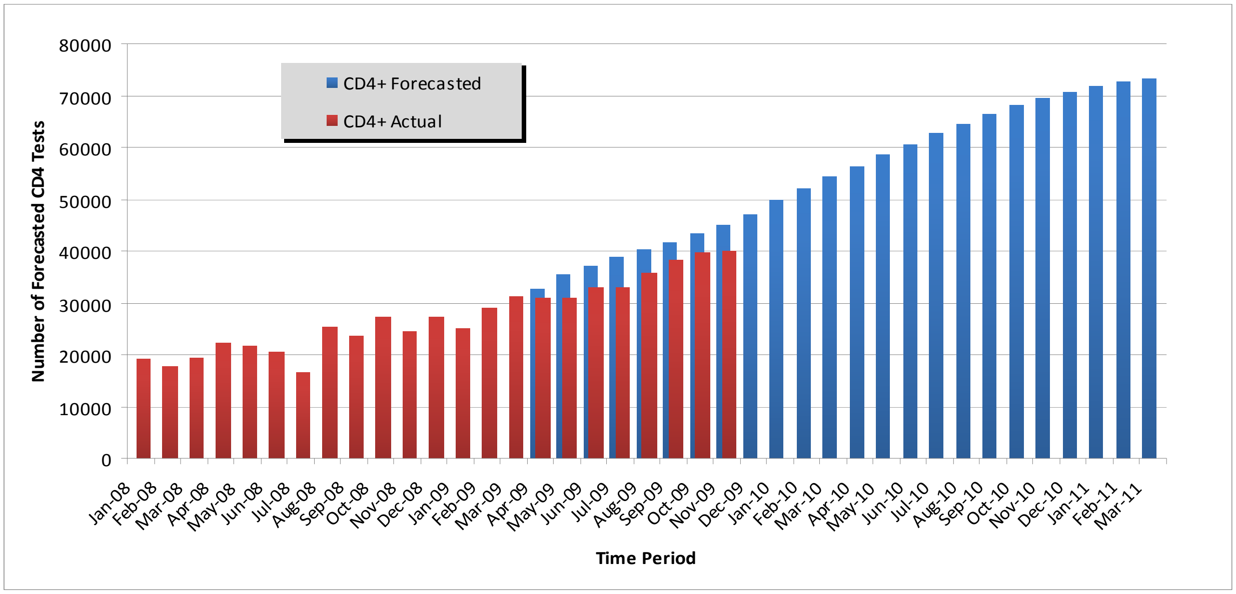
- Abrupt move from a small, vertical ART program and logistics system to a large, mainstreamed, integrated, and decentralized ART with little coordination with the supply chain, capacity building, and support of staff with logistics responsibilities.
- Significant supply chain management challenges at the CMS.
- Significant changes by MOH to STGs without coordinating implementation feasibility with parties responsible for forecasting, procurement, and distribution.
- Period of long uncertainty regarding leadership for the continued implementation and management of ARV LMIS during decentralization and mandated integration of services and supply chain without capacity readiness.
- Unpredictability of the disbursement of GFATM funds resulted in forecasted medicines not being available for implementation of new STGs as planned for phasein/phaseout.
- Increases in fuel costs and changes to transport options led to distribution challenges.

Laboratory Reagent Findings

The laboratory logistics system in Country Y is also a vertical system, but it differs distinctly from the ARV logistics system. Currently, the quantification process is led by a PEPFAR-funded project and includes forecast and supply planning updates, which occur on a regular basis. The project is also directly responsible for the procurement of HIV and AIDS-specific laboratory commodities. Laboratory supplies are procured from an in-country vendor who is responsible for shipping and handling these commodities as well as their storage and distribution. The vendor, who is well staffed, has adequate distribution capabilities that allow for the timely resupply of commodities to facilities. The PEPFAR program funds all of the HIV and AIDS-specific laboratory commodities for Country Y. The laboratory logistics system is also well staffed, and staff are trained on a regular basis.

Given the data available, a forecast error analysis was done looking at the forecasted number of CD4 tests against the actual number of CD4 tests carried out in a defined time period. In this analysis, the CD4 tests are representative of a total of 10 FACSCount and FACSCalibur reagents. The correlation between Country Y's laboratory forecasts and actual consumption has been plotted in figure 15 below.

Figure 15. CD4 Laboratory Reagent Forecast in Country Y



The MAPEs for the forecast have been listed in table 8 below. As one can see from the results, there is a very close correlation between the forecasted quantities of CD4 reagents and the actual quantities consumed.

Table 8. MAPEs for CD4 reagents in country Y

Timeframe	Actual Number of CD4 Tests	Forecasted Number of CD4 Tests	MAPE
April 2009	32,575	30,888	5.46%
May 2009	35,471	30,880	14.87%
June 2009	37,270	32,858	13.43%
July 2009	38,804	32,871	18.05%
August 2009	40,202	35,639	12.80%
September 2009	41,689	38,186	9.17%
October 2009	43,273	39,716	8.96%
November 2009	44,962	40,034	12.31%

Focusing on data quality, staff supervision, and training improves forecast accuracy. One of the reasons for the close correlation between the forecasted and actual results is because of the emphasis the country has placed on the collection of quality data, as well as on supervision and training on the logistics system. The country has a short, two-level pipeline, which facilitates the movement of commodities and data in a timely manner. As a result, annual forecasts and updates incorporate the latest consumption data, helping to minimize forecast error. The data that are used to inform the forecasts are reported monthly for 81 instruments, from the 42 individual labs, directly to the project office that manages the resupply. In addition, sufficient staff are allocated to LMIS report review, site follow-up, data entry, resupply quantity calculation, and vendor performance management. The LMIS data that are collected from the facilities are entered into a database for all client laboratories that are routinely resupplied with laboratory reagents and consumables. Data quality and reporting timeliness are well managed, and contact with sites is regular and positive.

Improved quality of data results in more confident use of these data for decisionmaking. Logistics data quality and visibility has helped Country Y in downstream decisionmaking for key logistics functions, such as forecasting, financial planning, and supply planning in close coordination with the local laboratory vendor who supplies each site on a monthly basis. Country Y also currently uses consumption data to carry out its HIV and AIDS laboratory reagent forecast. The low MAPE seen in table 8 is an indication that the quality of consumption data being used to carry out the forecasts is very reliable.

Comprehensive vendor responsibilities increase reliability of access to lab supplies and functional equipment. The vendor in Country Y's case is responsible for shipping, handling, storage, and distribution of laboratory supplies as well as providing regular service and maintenance to equipment installed at laboratories across the country. Such a comprehensive arrangement increases accountability and improves overall service.

Regular service and maintenance of equipment keeps instruments running better and results in more reliable forecasts of consumables. The relatively low downtime for the instruments within the PEPFAR network is an important reason for the low forecast error rates in Country Y. The third-party vendor that Country Y uses for the routine delivery to facilities is responsible for routine service and maintenance of the equipment. As a result of reliable equipment and supply chain, laboratories are able to operate regularly; therefore, they are able to report on their consumption on a monthly basis so that they can have a regular flow of commodities. Currently, sites send in their LMIS reports via fax or e-mail.

Consistent testing guidelines make it easier for providers to adhere correctly to protocols and for program managers to forecast demand. The testing guidelines and instrument platforms for Country Y have remained the same over the last few years, especially for CD4 testing. As such, the pattern of use for these products tends to stay constant, further improving the project office's ability to forecast demand accurately. Caregivers are also more likely to follow the guidelines more stringently when the testing protocols don't change frequently since they do not need additional training on new prescribing patterns for diagnostic tests.

Forecasts and budgets are used to help plan expansion of the lab network. Another reason for the low error rate in Country Y's laboratory forecasting is the close link among forecasting, funding, and scale-up of the lab network. There is very close collaboration among the CDC, the MOH, and the Project field office responsible for this CD4 supply chain. Expansion of the network is dependent on the available financial resources and MOH priorities. The forecast informs what is possible with existing financial resources, and policymakers decide how to increase the reach of the lab network more efficiently, be it by placing new instruments or referring samples to underused instruments, for example. The MOH and CDC's programmatic vision also feeds the assumptions in the forecast that inform planning and future financial resource availability.

Reliable funding enables vendor to provide services consistently. Funds for the procurement of laboratory reagents and instrument maintenance and repair are available and paid to the vendor as per contractual terms. Through accurate forecasts and consistent availability of funds, the vendor is able to ensure that the reagents are procured, stored, and distributed to the laboratories on a timely basis. There also have not been any known funding gaps for this group of products.

Ethiopia

Forecasting Method:	Morbidity method for both adult and pediatric
Responsible Party for Forecast:	PFSA, Federal Ministry of Health, FHAPCO, and EHNRI, DACA, Tikur Anbessa Hospital, Zewuditu Hospital, I-TECH, Intra Health, FHI, JHPIEGO, JHU/TSEHAI, USAID, WHO, WFP, CDC, MSH/SPS, USAID/HCSP, USAID/PSP, USAID | DELIVER PROJECT, SCMS, CU-ICAP, PSI, JSI/MMIS, PI, CHAI, Save the Children USA, IOCC, and others
Frequency of Forecast:	Annually
Monitoring Procedures:	Pipeline monitoring
Members of Monitoring Team:	PFSA, SCMS, and CHAI as part of a working group
Frequency of Monitoring:	Quarterly
Forecasts Used in Analysis:	January 2008: Reviewed June 2008 to April 2009 January 2009: Reviewed January 2009 to April 2009
Type of Data Used for "actuals":	Dispensed-to-user data from LMIS

Analysis

In Ethiopia, two separate national ARV forecasts were shared and analyzed. These forecasts, developed in January 2008 and January 2009, were originally performed in Quantimed for adult first line ARVs and CHAI forecasting tool for adult second line and pediatric ARVs with the resulting monthly forecasts exported into PipeLine. "Actual" data represent monthly dispensed-to-user data collected by MSH/SPS, shared through the LMIS and also imported into PipeLine. Reporting rates for the dispensed-to-user data are documented to be an average of 95 percent.

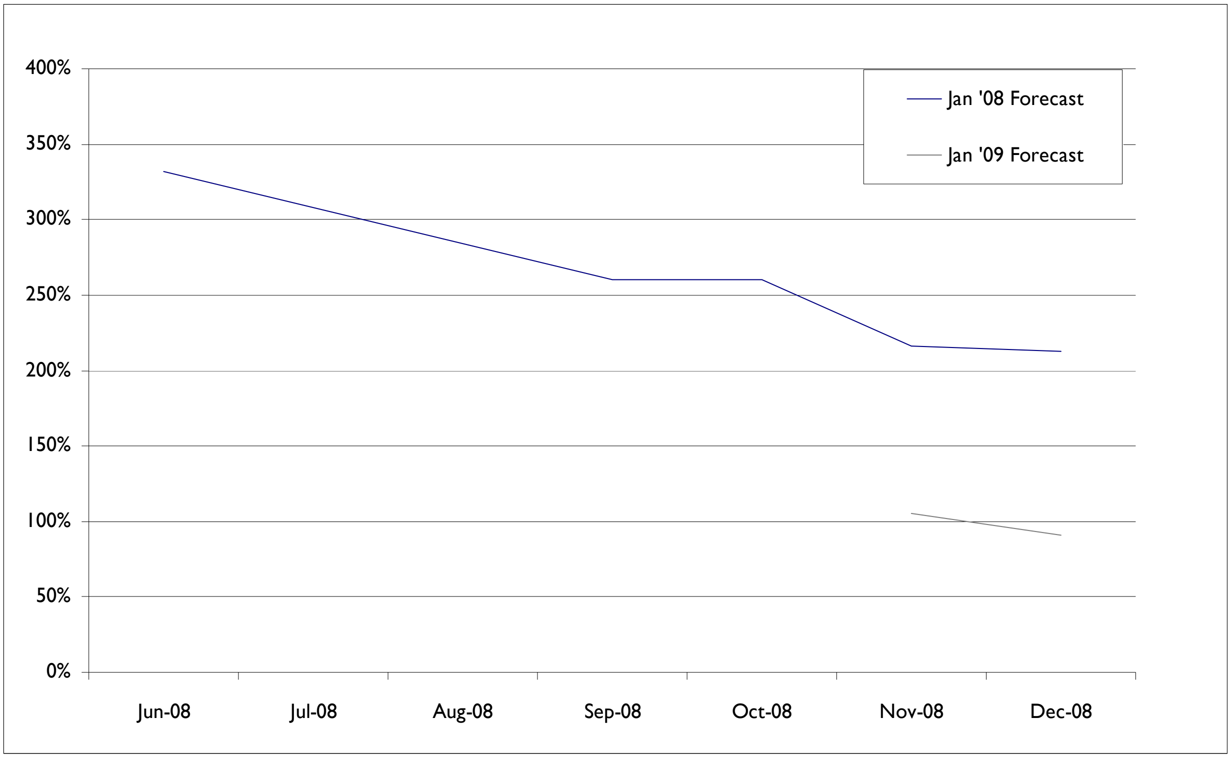
Some of the actual data used in the analysis were not reported bimonthly, and for most products were documented bimonthly or quarterly. Therefore, the analysis was performed accordingly, and monthly forecasts were added for some of the comparisons. In the January 2008 forecasts, forecasts for pediatric formulations was not available for 2009, so no analysis for pediatric products could be made for that period of time.

ARV Findings

Mature, institutionalized programs make for better forecasts. In Ethiopia, as in other countries reviewed, the second forecast is an improvement on the first. This is illustrated in figure 16 below. This result is also seen in Zambia (figure 12), and, to a lesser extent, in Country X (figure 2). With a review of only two forecasts, it is difficult to draw definitive conclusions; however, the decrease in MAPEs with the second forecast may be attributed to greater experience of the staff, an increase in forecasting skills, and further development of the LMIS. The Ethiopia program is also mature with institutionalized processes and structures for forecasting.

Adult ARV forecasts improve from the time of forecast in Ethiopia. The MAPEs for both forecasts have been plotted in figure 16 below. As you can see in this figure and in more detail in figures 17 and 18, unlike the results seen in the other countries reviewed, MAPEs in Ethiopia do not increase over time from the point of forecast. Accuracy actually improves for both forecasts during the period of time reviewed. The only exception to this is that pediatric ARV forecast errors do increase over time, as seen in figures 17 and 18. The CHAI tool was used to forecast all pediatric and adult second line ARVs, and this may have had an impact on these results.

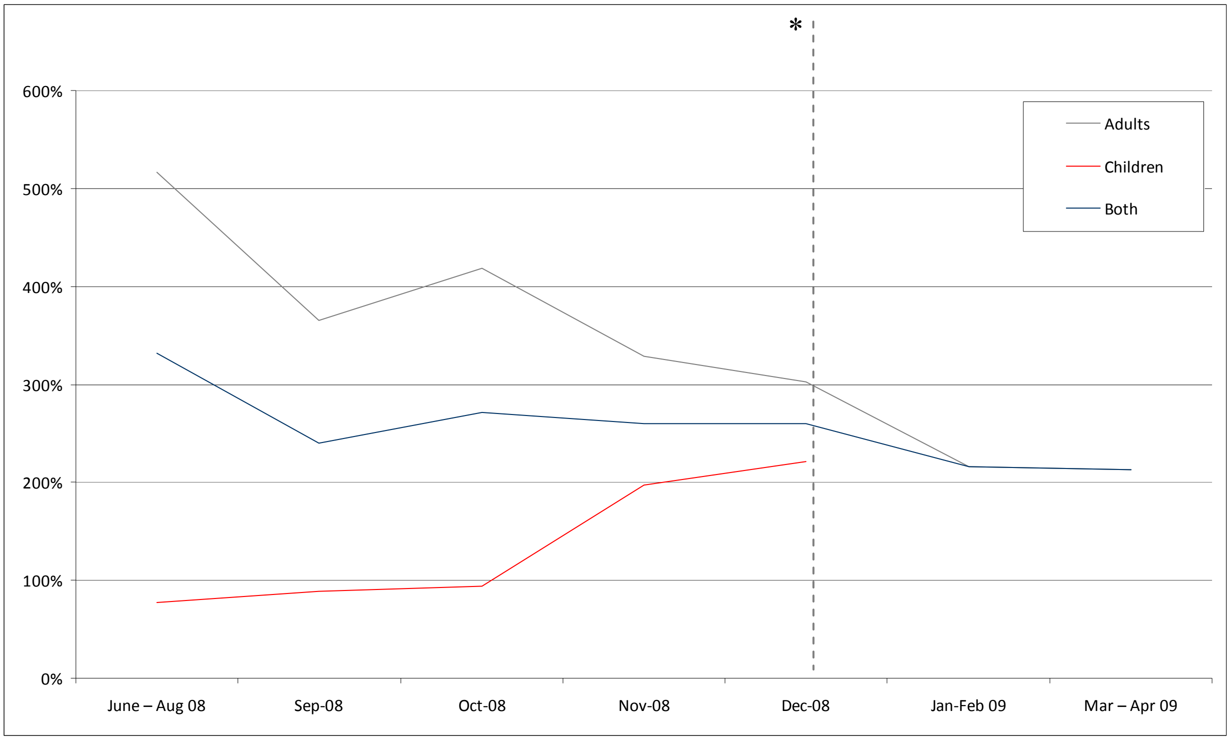
Figure 16. MAPEs for Two Separate Forecasts in Ethiopia



Forecasts for pediatric ARVs are more accurate than are forecasts for adult ARVs. Figures 17 and 18 represent the MAPEs for each of the two forecasts and the breakdown of these MAPEs by adult and pediatric formulations. Again, what we see in Ethiopia is in contradiction to what we saw in Country X, Country Y, and Zambia. Forecasts were more accurate for the pediatric products than they were for the adult products. The morbidity method was used for both forecasts, as they were for the other countries mentioned, so this is not a determining factor. What may be a determining factor is that CHAI is responsible for the pediatric forecasts in Ethiopia and revises them quarterly, given the greater experience with and understanding of the pediatric situation in country. The CHAI tool also has the added benefit of being able to adjust for changes in regimens better than Quantimed does.

Expert panels may contribute to greater forecasting accuracy. One reason for better accuracy of forecasting pediatric products in Ethiopia may be the large and diversified team that meets annually to assist in forecasting ARVs in Ethiopia. The larger team divides into expert panels to focus on different classes of products, such as adult and pediatric ARVs, PMTCT, RTKs, etc. This smaller team of experts, uniquely qualified to discuss pediatric ARVs, may have accounted for the better assumptions and analysis of available data and, therefore, may have contributed to more accurate pediatric ARV forecasting.

Figure 17. MAPEs for Ethiopia January 2008 Forecast



*Pediatric ARV forecasts were not available for review to the left of the dotted line.

Figure 18. MAPEs for Ethiopia January 2009 Forecast

In tables 9 and 10, the adult and pediatric ARVs have been grouped separately, and the error is presented without taking its absolute, thereby revealing products accurately forecasted (within 25 percent of the actual), underforecasted (more than 25 percent), and overforecasted (less than -25 percent). It is again evident from these two tables that forecasts for pediatric ARVs had more accurate results than did those performed for adult ARVs.

Table 9. MAPEs for adult ARVs for the first three and four months after forecast in Ethiopia

Product Name	Jan 2008 Forecast	Jan 2009 Forecast
ABC 300mg 60 Tabs		-350%
ddI 100mg 60 Tabs	98%	-25%
ddI EC 250mg blister (3x10) Caps	-207%	-39%
ddI EC 400mg blister (3x10) Caps	17%	31%
EFV 200mg 90 Caps	66%	41%
EFV 600mg 30 Tabs	-24%	-103%
3TC 150mg 60 Tabs	86%	-89%
3TC/D4T 150/30mg 60 Tabs	-406%	-38%
3TC/D4T/NVP 150/30/200mg 60 Tabs	-397%	-71%
3TC/ZDV/NVP 150/300/200mg 60 Tabs	-524%	-44%
3TC/ZDV 150/300mg 60 Tabs	56%	-24%
LPV/R 133.3/33.3mg 180 Caps	-18%	
TDF 300mg 30 Tabs	-139%	-378%
ZDV 300mg 60 Tabs	-223%	

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

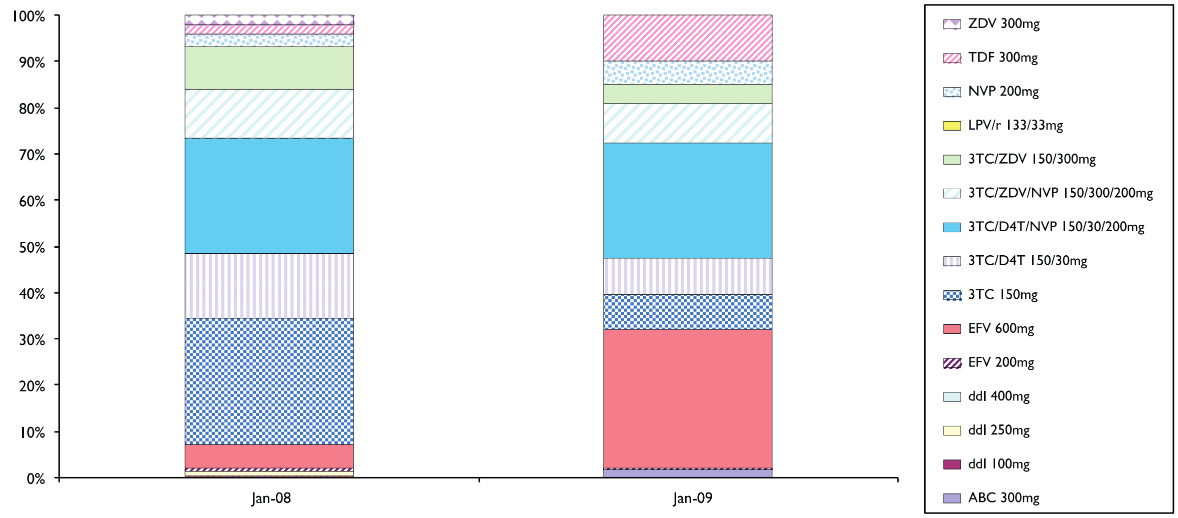
Table 10. MAPEs for adult ARVs for the first three and two months after forecast in Ethiopia

Product Name	Jan 2008 Forecast	Jan 2009 Forecast
ABC 20mg/ml 240 ml	-95%	-169%
ddI 25mg 60 Tabs	-84%	-59%
ddI 2g, powder for oral solution 10mg/ml	-786%	43%
EFV 30mg/ml 180ml	16%	-1%
EFV 50mg 30 Caps	54%	12%
3TC 10mg/ml 240 ml	13%	-1%
3TC/D4T/NVP 30/6/50mg 60 Tabs		-23%
3TC/D4T/NVP 60/12/100mg 60 Tab		-68%
LPV/R 80/20mg/ml 300 ml	64%	66%
NVP 10mg/ml 240 ml	89%	
D4T 15mg 60 Caps	-138%	
D4T 1mg/ml 200 ml	-187%	-17%
D4T 20mg 60 Caps	77%	59%
3TC 100mg 100 Caps	-55%	-22%
3TC 10mg/ml 240 ml	-32%	-2%

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

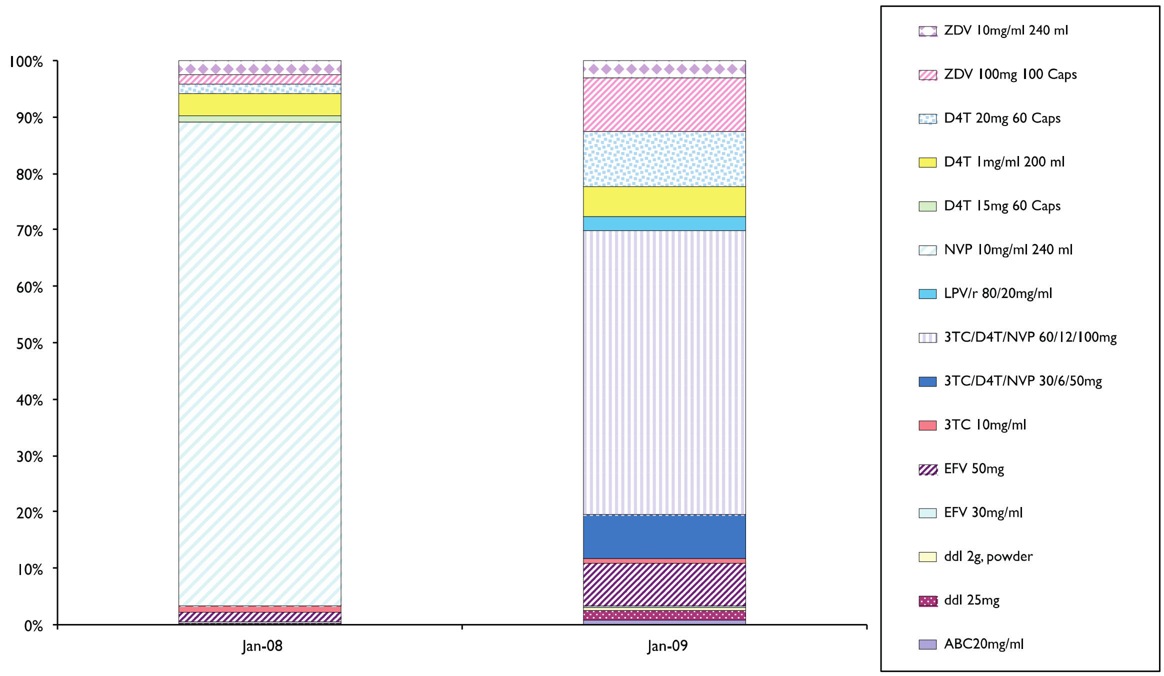
In figures 19 and 20, the impact of the errors is revealed. The adult and pediatric MAPEs from the tables 9 and 10 have been multiplied by the percentage of the total ARVs consumed to reveal the impact of the forecasting errors. The highest values for each forecast are attributed to products with a combination of the highest error and quantities consumed, and the lowest values are attributed to the ARVs with a combination of the lowest error and quantities consumed. When reviewing these figures, the highest values should be the focus of concern and improvement.

Figure 19. Impact of Adult ARV MAPEs in Ethiopia



High MAPE and high volume account for largest impact for the adult ARV forecasts. For both forecasts, the triple FDC 3TC/D4T/NVP 150/30/200 appears to have the highest impact. It is interesting that this product represents a high impact for two different reasons in each forecast. In the January 2008 forecast, the reason is that it had a high MAPE of 397 percent shortly after the forecast with a small percent of the volume consumed (7.31 percent of ARVs). In the January 2009 forecast, however, it had a lower MAPE of 71 percent, but its share of the ARVs consumed increased to 25.7 percent. The other products competing for highest impact are 3TC 150mg in the January 2008 forecast and EFV 600mg in the January 2009 forecast. Both had large MAPEs—86 percent and 103 percent, respectively, and represented large portions of the volume consumed—36.7 percent and 21.4 percent, respectively. Ethiopia should concentrate on improving forecasts for the highest-consumed products and on decreasing large MAPEs for all other products.

Figure 20. Impact of Pediatric ARV MAPEs in Ethiopia



High MAPE and high volume account for largest impact for the pediatric ARV forecasts. The two products that have a high impact in the January 2008 and January 2009 forecasts are NVP 10mg/ml and the pediatric FDC 3TC/D4T/NVP 60/12/100mg, respectively. NVP 10mg/ml had both a large MAPE (89 percent) and a large percentage of volume (76.7 percent) consumed shortly after the 2008 forecast, resulting in the very large impact viewed in figure 20. It continued to be a large percentage of all products consumed in the following year, yet data on its forecast were not available for analysis for January 2009. NVP 10mg/ml also represents 37.8 percent of the products consumed for January and February 2009, it's likely that its forecasting accuracy would have continued to have a high impact. Although the FDC 3TC/D4T/NVP had a rather small volume consumed shortly following the Jan 2009 forecast of 8 percent, it did have a large MAPE of 68 percent, in comparison with the other MAPEs for the pediatric products. Ethiopia should concentrate on improving forecasts for the highest-consumed products and on decreasing large MAPEs for all other products.

Table 11. Accurate, underforecasted, and overforecasted ARVs for the first three and two months following forecasts in Ethiopia

	within 25% error Accurate		> 25% error Underforecast		< -25% error Overforecast		Total ARVs
	Number	%	Number	%	Count	%	
January 2008	5	19%	8	30%	14	52%	27
January 2009	8	31%	5	19%	13	50%	26

The number of ARVs accurately forecasted improves with subsequent forecasts. As discussed earlier in figure 16 above and as seen in table 11, there is an improvement in the number of ARVs forecasted accurately with each subsequent forecast in Ethiopia. Accurately forecasted ARVs increase from 19 percent of products to 31 percent of products. The percentage of products underforecasted decreases, and the percentage of products overforecasted averages around 50 percent.

ARV Conclusion

The achievements in forecasting accuracy can be attributed to:

- Experienced staff, institutionalization, and a mature program
- Well-developed LMIS and resulting quality data availability
- Large and diverse forecasting teams
- Using the CHAI tool for pediatric and second line ARVs may have had an impact on better pediatric forecast accuracy; this result should be reviewed further.

Zambia

Forecasting Method:	**ARVs:** Morbidity method for both adult and pediatric **RTKs:** Morbidity method **Labs:** Service statistics method
Responsible Party for Forecast:	**ARVs:** MOH (with representatives from the central, provincial, district, and service delivery levels), USAID, CDC, UNICEF, USAID | DELIVER PROJECT, AIDSRelief, PRA, CHAI, ZPCT, and CIDRZ **RTKs:** MOH, USAID | DELIVER PROJECT, CHAI, JICA, Kara Counseling, Planned Parenthood Association of Zambia, ZPCT, National HIV/AIDS/STI/TB Council, Chreso Ministries, CIDRZ, Latkings, Virology lab-NUFU project, NAC, Society for Family Health, ZVCT (MOH) **Labs:** MOH, SCMS, CDC, AIDSRelief, ZPCT, PRA, CHAI, Lusaka DHO, and CIDRZ
Frequency of Forecast:	Annually
Frequency of Monitoring:	Quarterly
Monitoring Procedures:	**ARVs:** Same as in annual forecasts, with greater emphasis on review of consumption patterns and changes to ART guidelines or treatment practices **RTKs:** Same as in annual forecasts, with greater emphasis on review of consumption patterns **Labs:** Same as in annual forecasts, with greater emphasis on review of issues patterns and update of new equipment to be procured by all partners
Members of Monitoring Team:	Same as in annual forecasts
Forecasts Used in Analysis:	**ARVs:** January 2009: reviewed January 2009 to March 2010 July 2009: reviewed July 2009 to March 2010 November 2009: reviewed November 2009 to March 2010 **RTKs:** Q1 2009: reviewed January to December 2009 Q2 2009: reviewed April to December 2009 Q4 2009: reviewed October to December 2009 **Labs:** November 2007: reviewed April 2008 November 2008: reviewed April 2009 November 2009: reviewed April 2010
Type of Data Used for "actuals":	**ARVs:** LMIS data with a reporting rates averaging 97% **RTKs:** Issues data from CMS to SDPs **Labs:** Issues data from MSL to SDPs

Analysis

In Zambia, three separate national ARV forecasts, three RTK forecasts, and three lab forecasts were shared and analyzed. The ARV forecasts were developed in January 2009, July 2009, and November 2009. The RTK forecasts were developed for Quarters 1, 2, and 4 in 2009 and lab forecasts in November 2007, 2008, and 2009. The ARV forecasts were originally performed in Quantimed, and monthly forecasts were exported to PipeLine. "Actual" data represent dispensed-to-user monthly data reported by the SDPs and were exported from the Supply Chain Manager software to PipeLine. Reporting rates for the dispensed-to-user ARV data are documented to be an average of 97 percent of facilities during 2009.

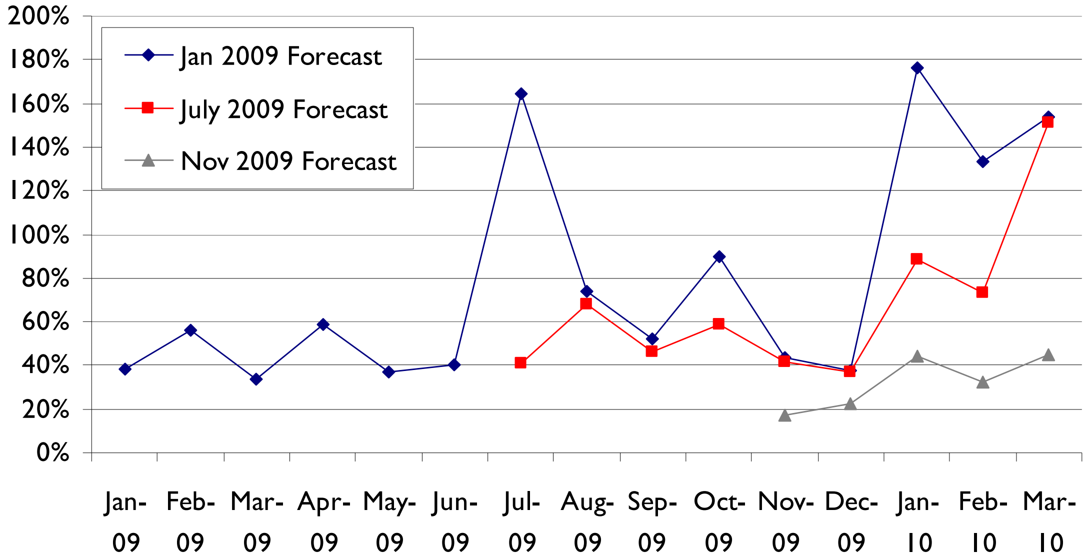
The RTK forecasts were done in ProQ and exported to Pipeline. Issues from the Medical Stores Limited (MSL) were the "actuals" in this analysis, as reporting rates from the LMIS were averaging around 73 percent.

The laboratory forecasts were carried out using service statistics data in Excel models. The forecast data were made available in six-month increments before being exported to PipeLine. "Actual" data represent issues data reported from MSL's warehouse management system software MACS. Prior to the quantification exercise, data collection was done with issues data collected from MSL. Zambia also has a laboratory LMIS, and reporting rates from that system were 79 percent, due, in part, to the fact that the laboratory logistics system was in the process of being implemented during the mentioned forecast periods. Although forecasts were being done nationally during this time, no national consumption data existed to inform the forecast or subsequent update. It is expected, however, that future forecasts will begin to use consumption data to inform any required changes to the forecast or supply plan. Given the sheer number of laboratory commodities that are managed in the LMIS (185), only a small subset of the tracer commodities presented in table 17 were monitored for forecast accuracy in this analysis. In order to have a diverse group of products to monitor, products were chosen across testing groups, namely CD4, chemistry, and hematology.

Findings: ARVs

Forecasting error increases from the time of forecast. The MAPEs for all three forecasts have been plotted in figure 21 below. Each individual forecast shows an overall increase in MAPE from the time of forecast. Over time, the data and assumptions used to perform the forecast grow farther and farther away from present reality. In other words, the longer ago the forecast was made, the greater the error. This is one reason why ongoing monitoring and updating of quantifications is recommended. As a result of the Zambia forecasting monitoring meeting in July 2009 and its review and development of new forecasts, the MAPEs for the July 2009 forecast decreased for all of the months analyzed in comparison to the January 2009 MAPEs. With the November 2009 forecast, the MAPEs dropped again for all months analyzed, this time more significantly. This forecast happened to be the more extensive annual forecast (where more data sources are consulted) and, as a result, showed greater accuracy than did the previous two reviews/forecasts.

Figure 21. MAPEs for Three Separate Forecasts in Zambia



Mature, institutionalized programs make for better forecasts. Consistent with the other countries reviewed for this study, the second and third forecasts improved on the first. We see similar examples in Ethiopia and, to a lesser degree, in Country X. Based on these three forecasts, the decrease in MAPEs with each forecast may be attributed to the additional experience of the staff and their improved forecasting skills, and perhaps further development/strengthening of the LMIS. The Zambia program is also a mature program, where processes and structures for forecasting are institutionalized.

Greater forecasting error is experienced with pediatric ARVs than with adult ARVs. Figures 22, 23, and 24 below show the MAPEs for each of the three forecasts and the breakdown of these MAPEs by adult and pediatric formulations. Seeing the overall increase in the MAPEs in relation to time, it becomes evident that forecasts for pediatric formulations have greater MAPEs than do the forecasts for adult formulations.

Figure 22. MAPEs for Zambia January 2009 Forecast

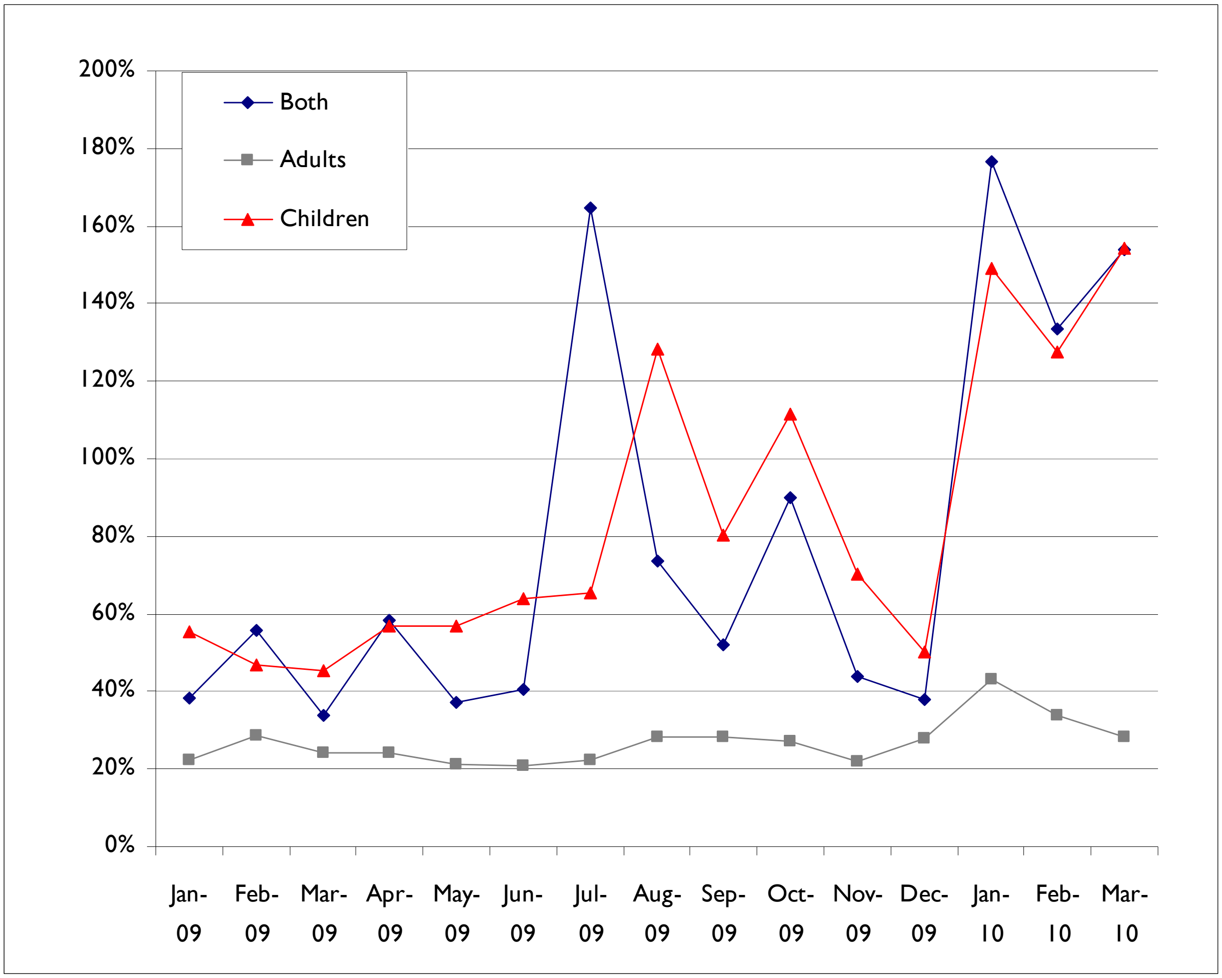


Figure 23. MAPEs for Zambia July 2009 Forecast

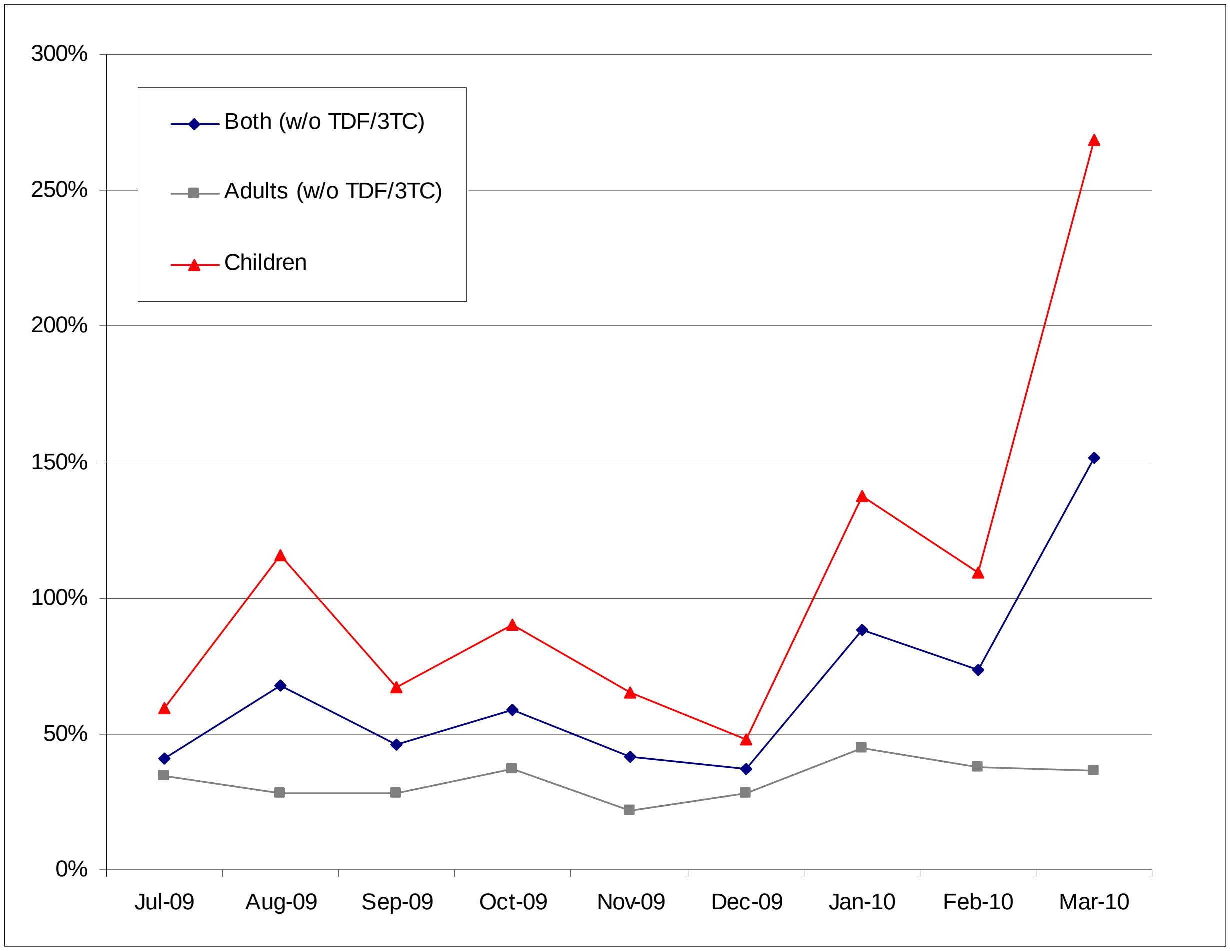
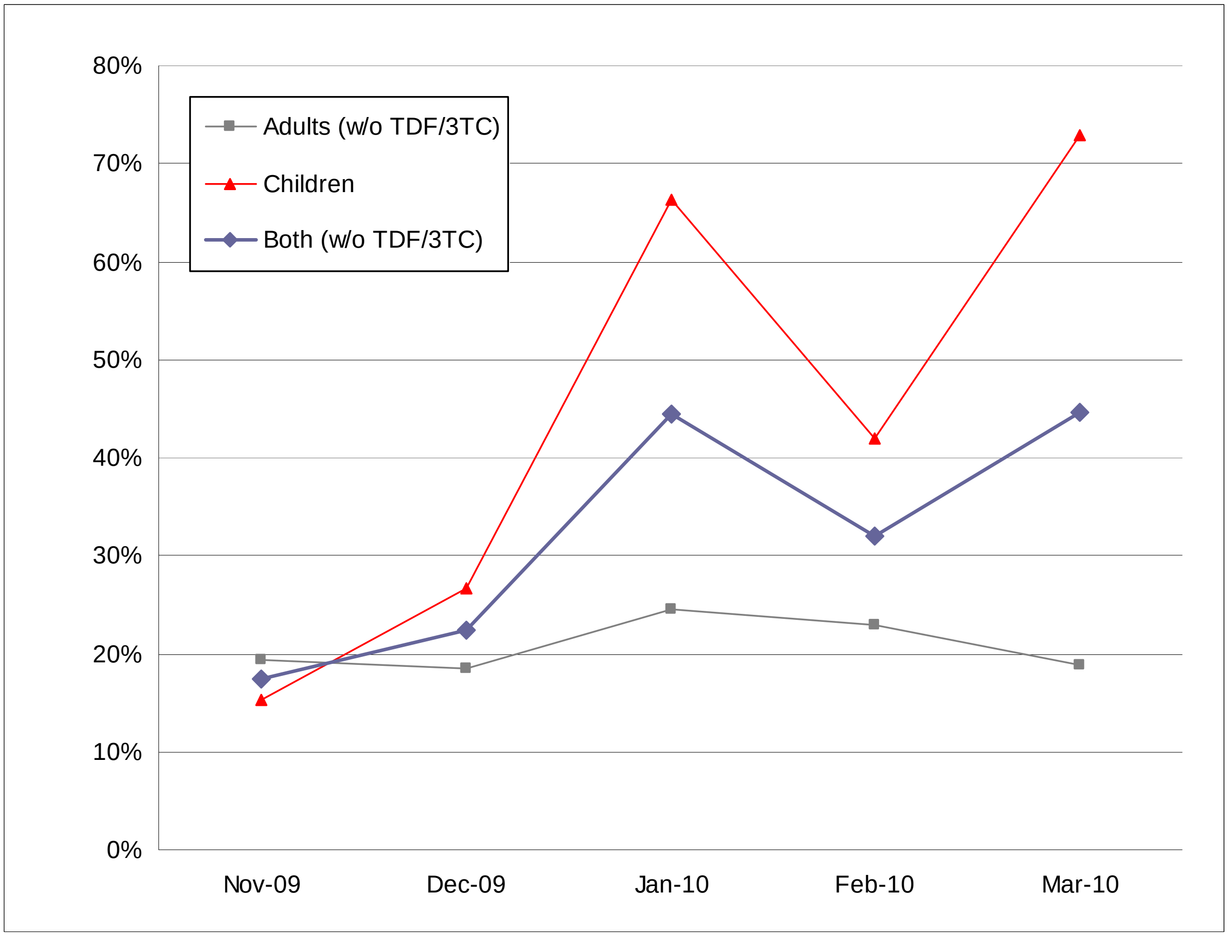


Figure 24. MAPEs for Zambia November 2009 Forecast



Note:
Not only is the error greater for pediatrics, but most of these errors are overestimates, as is visible when comparing tables 12 and 13 below.

Pressure to forecast higher number of pediatric ART patients may increase forecasting error. One of the reasons for overestimating pediatric ARVs (also seen in Country Y) may be the inability to agree on the actual numbers of current and expected pediatric patients. One issue in Zambia has been that the number of current patients does not compare with the actual consumption recorded by the sites. The number of patients is much higher than the consumption recorded, so either the number of patients is correct and they are not getting the medications, or the number of patients is over-recorded. Although the forecast for pediatric ARVs is overestimated, the team in Zambia reviews and monitors actual consumption patterns routinely, and, based on these figures, the forecast is adjusted in PipeLine in order to manage procurement planning of these products and limits overstocking and expiration of pediatric products.

Monitoring of consumption patterns may decrease forecasting error. The same can be said regarding all formulations, whether adult or pediatric; since monitoring is performed on a quarterly basis, adjustments can be made quickly to the procurement or pipeline plans without having to change forecasts to make sure that there is no overstocking or understocking. This was evident when reviewing the data made available for analysis. The forecast team reviewed the Quantimed forecasts, and if the morbidity data were very different from the actual consumption trends, the team agreed to follow consumption and override the Quantimed data and used the "projected trend" from PipeLine instead. This was seen mostly with pediatric products.

For figures 22 and 23 above, the MAPE for TDF/3TC was removed due to a very large error seen below in table 13. The details surrounding this product are discussed below.

In tables 12 and 13 the adult and pediatric ARVs have been grouped separately, and the error is presented without taking its absolute, thereby revealing products accurately forecasted (within 25 percent of the actual), underforecasted (more than 25 percent), and overforecasted (less than -25 percent).

Table 12. MAPEs for adult ARVs for the first three months after forecast in Zambia

	Jan 2009	Jul 2009	Nov 2009
Medicine	**Average of first three months**	**Average of first three months**	**Average of first three months**
ABC 300mg	-1%	32%	24%
ddI 100mg	-54%	-96%	27%
EFV 200mg	-5%	15%	-11%
EFV 600mg	-24%	-21%	-16%
D4T 30/3TC150/NVP200	13%	15%	8%
3TC 150mg	-6%	26%	34%
D4T 30/3TC 150mg	26%	26%	-1%
ZDV 300/3TC 150mg	-3%	8%	15%
LPV 200/r 50mg	0%	12%	12%
NVP 200mg	8%	3%	11%
TDF 300mg/FTC 200mg	-16%	22%	26%
TDF 300mg/3TC 300mg		42466%	-977%
ZDV 300mg	-57%	-36%	-20%

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

Error increases are due to prescribing practices not changing as anticipated. Table 12 shows that overall forecasts were fairly accurate, with the exception of TDF/3TC. This product was not forecasted in the January 2009 forecast, but was extremely underforecasted in July 2009 and overforecasted in the November 2009. The MOH believed that with the introduction of TDF/3TC FDC, the share of patients prescribed and taking TDF/FTC would decrease to half, with the rest being prescribed and starting on TDF/3TC. Because TDF was part of the main first line regimen, the quantity forecasted for it was rather high. Initially, the MOH did not take action to support the use of TDF/3TC, and it was not consumed at the rate assumed early in 2009. The consumption was actually zero until July, when its use slowly started to pick up. The assumptions of use changed, but they were not communicated and quantified for until the November update. By November, there was agreement that TDF/3TC consumption would not take off at the rate forecasted, and the MOH began taking a more active role in enforcing the new protocols to use this medicine.

Table 13. MAPEs for pediatric ARVs for the first three months after forecast in Zambia

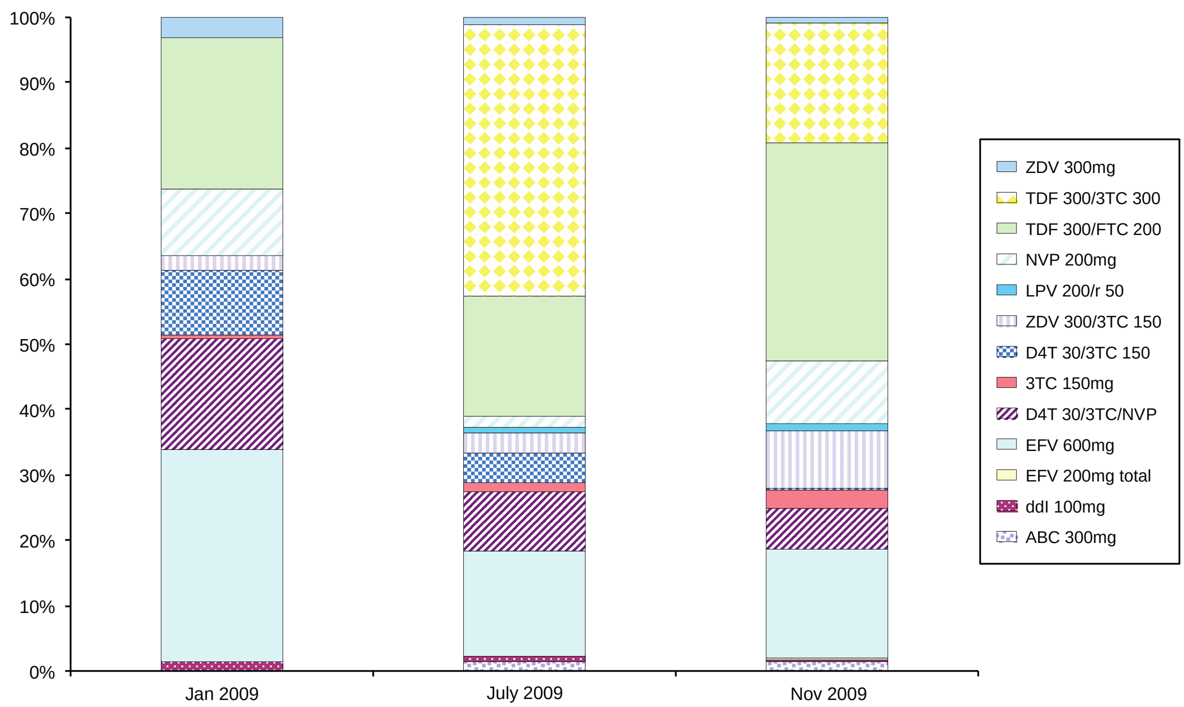
	Jan 2009	Jul 2009	Nov 2009
Medicine	**Average of first three months**	**Average of first three months**	**Average of first three months**
ABC 20mg/ml	-89%	-42%	-3%
ddI 25mg	-471%	-6%	4%
EFV 50mg	-140%	-159%	6%
3TC 10mg/ml	-3%	-44%	6%
D4T12/3TC60/NVP100	31%	33%	-12%
D4T6/3TC30/NVP50	22%	17%	
LPV 80/r 20 liquid	28%	-302%	-309%
NVP 10mg/ml	-5%	-45%	-1%
D4T 15mg	-158%	-190%	-14%
D4T 1mg/ml	-13%	-136%	10%
ZDV 100mg	-79%	-52%	-30%
ZDV 10mg/ml	-3%	-54%	11%

Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

Regarding pediatric ARVs, these also seem to have been forecasted fairly accurately, with a few exceptions. Policy changes regarding use of D4T led to assumptions of use that were not as predicted. D4T 15mg had a higher MAPE due to phase out of d4T 20mg as well as use of d4T 30mg. ddI 25mg was phased in to phase out ddI 50mg, and the uptake of ddI 25mg was lower than expected. Finally, the number of patients recorded taking EFV 50mg did not correspond with the reported consumption, which lead to higher MAPEs in forecasting.

In figures 25 and 26, the impact of the above errors is revealed. The adult and pediatric MAPEs from tables 12 and 13 have been multiplied by the percentage of the total ARVs consumed to reveal the impact of the forecasting errors. The highest values for each forecast are attributed to products with a combination of the highest error and quantities consumed and the lowest value are attributed to the ARVs with a combination of the lowest error and quantities consumed. When reviewing these figures, the highest values should be the focus of concern and improvement. If no one product has a large proportion of this impact, and all products appear to have the same value, then no one product requires special attention; either they all require special attention, or none does.

Figure 25. Impact of Adult ARV MAPEs in Zambia

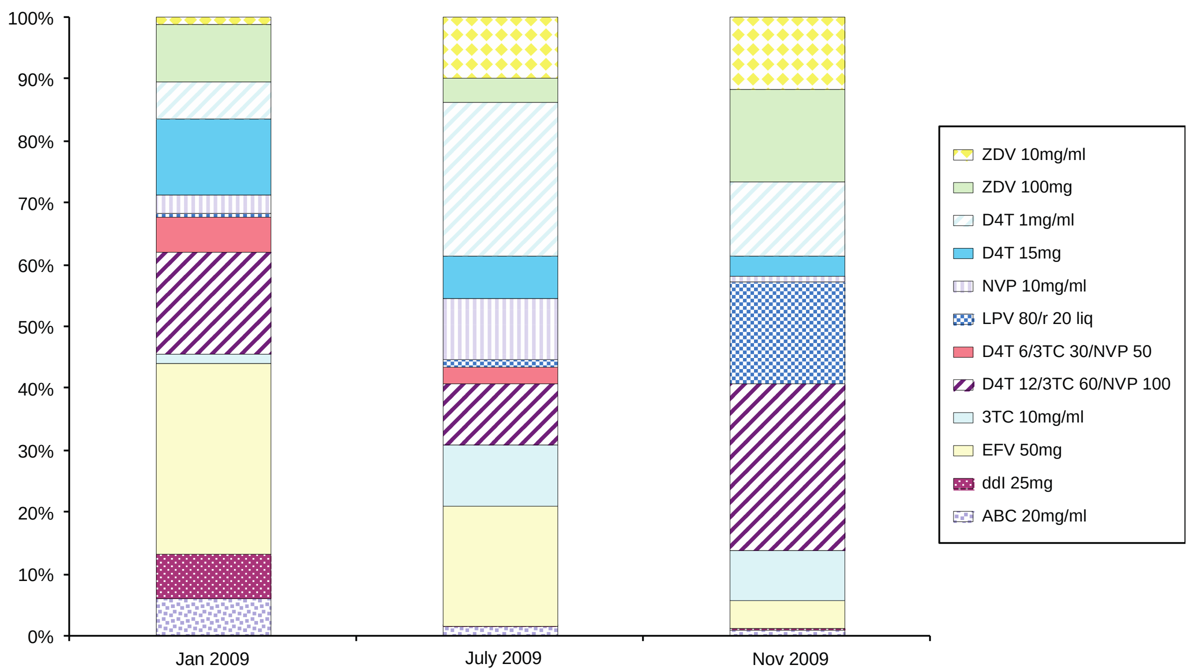


High MAPE and high volume account for largest impact for the adult ARV forecasts. TDF/FTC and TDF/3TC appear to be the two products with the largest forecasting impacts. The reasons for their forecasting errors were discussed earlier under table 12.

In the January 2009 forecast, the MAPE for TDF/FTC was 16 percent, and the volume consumed shortly after the forecast was 20.6 percent. The next product with a large burden of the error in this forecast was EFV 600mg, with a MAPE of 24 percent representing 19.6 percent of the total volume consumed. Other products had about the same level of consumption: D4T/3TC/NVP 150/30/200mg and NVP 200mg had 18.9 percent and 17.8 percent, respectively, but had slightly better MAPEs at 13 percent and 8 percent, respectively. It's interesting that when a country is forecasting accurately (based on this paper's definition of accuracy), a slight difference in MAPE and volume can increase one product's burden.

TDF/FTC continued to have a high impact in the July and November 2009 forecasts, with MAPEs of 22 percent and 26 percent, respectively, and consumption shortly after the forecast of 23 percent and 26.5 percent, respectively. TDF/3TC had an extremely high MAPE in the July 2009 forecast: 42,466 percent with a very small portion of consumption of 0.03 percent of ARVs. Its portion of the volume increased slightly to 0.37 percent with the MAPE improving, but this is still high in comparison to other products at 977 percent. Zambia should continue to improve on its forecasts for the highest-consumed products and decrease larger MAPEs for all other products.

Figure 26. Impact of Pediatric ARV MAPEs in Zambia



When a country has great forecasting accuracy, the impact of error is distributed fairly evenly. All products, with the exception of one, that represent the highest burden or impact of errors seen in figure 26 are neither the highest in MAPE nor in volume; they all lie somewhere in between. EFV 50mg has the highest burden of error in the January 2009 forecast, with a MAPE of 140 percent and a share of consumption of 6.68 percent shortly after the forecast. D4T 1mg/ml has the highest burden of error in the July 2009 forecast, with a MAPE of 136 percent and a share of consumption of 11.88 percent. The exception is D4T/3TC/NVP 12/60/100 in the November 2009 forecast, which was consumed at a higher rate than the rest shortly after the forecast, at 23.5 percent and a MAPE of 12 percent. Considering that most of the products are forecasted fairly accurately, the burden of error appears to be distributed fairly among all of the products. Zambia should continue to improve on its forecasts across all pediatric products.

Table 14. Accurate, underforecasted, and overforecasted ARVs for the first three months following forecast in Zambia

	within 25% error Accurate		**> 25% error Underforecast**		**< -25% error Overforecast**		***Total ARVs***
	Number	**%**	**Number**	**%**	**Number**	**%**	
January 2009	14	58%	3	13%	7	29%	*24*
July 2009	9	36%	4	16%	12	48%	*25*
November 2009	14	70%	3	15%	3	15%	*20*

The number of ARVs accurately forecasted improves with forecasts. As discussed earlier, and demonstrated in figure 21, overall the Zambia team had notable accuracy and improvement in accuracy in forecasting ARVs. The percentage of ARVs accurately forecasted during the three months following the forecast was 58 percent for January 2009 and 70 percent for November 2009.

ARV Conclusion

The achievements in forecasting accuracy can be attributed to:

- Experienced staff, institutionalization, and a mature program
- Review process
- Large and diverse forecasting teams
- Improvements in LMIS that resulted in the availability of quality data

For the products that did have high MAPEs, the reasons can be attributed to:

- Prescribing practices not changing as anticipated
- Difficulties in forecasting pediatric ARVs

RTKs Findings

For the HIV rapid test kits (RTKs), three forecasts were used: Quarters 1, 2, and 4 in 2009. The reporting rates for RTKs are lower than those of ARVs, at an average of about 73 percent in 2009. Therefore, issues data from MSL were used instead of dispensed-to-user data for analysis. Consequently, this increases the error rate because issues from the MSL will include provisions for future months as well as safety stock and will not reflect the quantities of actual tests used at the facility level. All three forecasts were for an annual quantity, with the exception of a quarterly forecast for SD Bioline tests performed in Quarter 1 and a monthly forecast performed for Determine tests in Quarter 2. This made analysis on a monthly basis difficult as forecasts had to be divided equally by the number of months covered. This decreases the value of the analysis since we are assuming that forecast would remain equal and stable during the forecasted period of time.

Having three separate forecasts within one year does show, however, that monitoring is occurring and forecasts are being adjusted.

Table 15. RTK MAPEs for the first three months following forecast in Zambia

Product	Quarter 1, 2009 Forecast	Quarter 2, 2009 Forecast	Quarter 4, 2009 Forecast
Bioline HIV 1/2 3.0 25 Tests	124%	143%	47%
Determine HIV-1/2 100 Tests	514%	30%	13%
Uni-Gold HIV 20 Tests	11%	29%	12%

In table 15, we can see that MAPEs are greater with Bioline and Determine tests and less so with Uni-Gold. Although it is not possible to make any other conclusions from these limited data, there does appear to be less error with Uni-Gold forecasts (used in confirmatory tests) than there is for the Bioline (used in tie-breakers) and Determine tests (used in screening).

For Determine, a large error in Quarter 1 is followed by a quick adjustment in Quarter 2 to bring the MAPE down from 514 percent to 30 percent and a very low MAPE in Quarter 4 of 13 percent. The errors seen in early 2009 for the Determine forecasts were due to underforecasting. The number of tests performed was increasing significantly, and after meetings with the different partners, it became apparent that some of the programs requiring HIV testing had not fully updated the quantification team about their needs.

Bioline forecasts also have high MAPEs in Quarters 1 and 2, with a decrease again in Quarter 4. More quality testing was assumed to be required than was actually used. An anticipated quality testing program never came into existence and the Bioline had to be donated to another country.

Laboratory Reagent Findings

Forecasting accuracy improves over time as staff gain experience and program matures. The MAPEs for the forecast of all of the tracer commodities analyzed have been plotted in table 16 below. From the analysis, chemistry and hematology reagents, some of which are used as tracer commodities in-country, had a relatively stable and low MAPE, aside from the Cobas reagent. On average, CD4 testing products had higher MAPEs as some of the reagents were being phased in during the forecast period assessed. For certain reagents, such as the FACSCalibur Calibrite beads, it is evident that the forecast distinctly improved over time. Overall, the forecast accuracy improved slightly from 2008 through 2009, though there are insufficient issues data for the latter half of 2009 and 2010 to determine how far that trend continues. One reason for this overall positive trend of a decrease in MAPEs is the additional experience of the staff in forecasting. Simultaneously, the roll-out of the laboratory logistics system during the same time period probably had a positive effect on the manner in which commodities were ordered from MSL and managed at the health facility level. The program has also begun to mature, so certain logistics functions, such as forecasting, are being institutionalized among key stakeholders.

Table 16. MAPEs laboratory reagents in Zambia

Product Name	Jan–Jun 2008	Jul–Dec 2008	Jan–Jun 2009	Jul–Dec 2009	Jan–Jun 2010
	Six-month average	Six-month average	Six-month average	Six-month average	Six-month average
FACSCount CD4 Absolute %	147%	10%	42%	110%	
ABX Minilyse	32%	43%	30%	55%	
ABX Minotone	11%	30%	17%	108%	
Olympus ALT	32%	7%	22%		
Olympus AST	63%	7%	65%		
FACSCount Contol Kit	178%	270%	93%		
FACSCalibur Calibrite 3 beads	13941%	2633%	82%	15%	
Sysmex Cell Clean	227%	17%	10%		
Cobas Creatinine	22%	374%	367%		
EDTA Vacutainer	155%	234%	255%	487%	67%

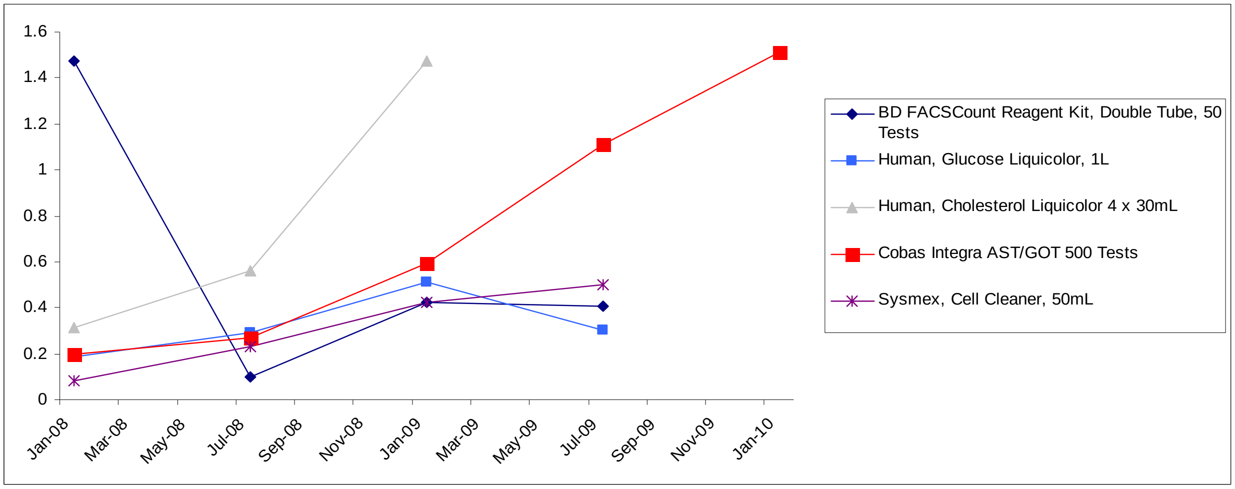
Note:
Red represents underforecasts (>25%), blue represents overforecasts (<-25%), gray represents insufficient data to complete the forecast, and no field color represents an accurate forecast, within 25%.

Lack of issues data can contribute to errors in forecasts. One of the challenges with the data noted during the forecast analysis was the inadequate amount of issues data used to inform the "actual consumption" at the SDP level. The country team noted this challenge in the various forecasting exercises it carried out. The reason for this is that certain commodities are not issued at the central level due to stockouts or to the warehouse management system's (WMS) failure to indicate that the item is in stock even when it is. As such, issues could be made and not noted by the WMS, subsequently falsely suppressing the "actuals." As seen in table 16, most of the errors noted represent overforecasts which could be a direct result of the under-reporting of the issues data used to carry out the analysis.

Introduction of new equipment can change assumptions on which forecasts are based. Another reason why the forecast error was on the higher side across all the testing areas was because of the ongoing procurement and installation of new testing equipment during the forecast period analyzed. Although the country team planned for this during the national forecast and the review exercise, it had no real control over the timing of this activity, which affected the issues data from MSL and the actual consumption at the facility level.

Figure 27 below shows the forecast accuracy of a few select laboratory reagents. The reagents chosen were the ones that had the fewest outliers and fairly reasonable forecast error results. The complete list of the forecast error is presented in table 16.

Figure 27. MAPEs for Select Laboratory Reagents Forecast in Zambia



Findings

Cross-Country Findings

Each country and situation is unique; however, some conclusions can be drawn about factors that contribute to errors in forecasting and the impact of those errors. Some of these are contextual, such as the environment in which forecasting is conducted, while others are due to the data and assumptions that are used to develop a forecast.

1. Consider the following as two distinct processes: planning for program scale-up and forecasting for procurement.

A forecast defines how the future is *likely* to look, whereas plans and targets are how we *want* the future to look. If a forecast and procurement plans are based solely on program plans, goals, and targets, the accuracy of the forecast will be compromised. Plans serve as inputs to forecasts and also are influenced by them. While program planning and forecasting are intertwined, they should be thought of as distinct processes when developing procurement plans based on a forecast. As a rule of thumb, forecasts should drive planning, not vice versa.

Consider the following policy shift in the context of forecasting: WHO recommends that ART should be initiated in HIV-positive patients with a CD4 count of <350 cells/mm3. This is a revision from the previous recommendation of initiating ART in patients with a CD4 count of <250 cells/mm3. This policy change could greatly influence the demand for ARV medicines in a particular country or program. With a higher CD4 count, more HIV-positive clients will qualify to initiate ART. In some high-prevalence countries, this policy change may result in exponential growth in demand for ARV medicines if it is implemented too swiftly. For forecasting purposes, it is important to determine how this policy would be implemented by collecting data on CD4 diagnostic capacity, human resource capacity, available financing for scale-up, and clarification of the relationship between diagnostic tests and initiation of ART. If we just assume that all patients who meet the CD4 count threshold for ART would immediately be placed on treatment as soon as the policy is announced, we would probably greatly overestimate the quantities of each ARV drug required.

One trend found in several of the countries reviewed for this paper was great pressure to increase scale-up of pediatric patients and forecasting based on overly ambitious goals. The results were higher MAPEs for these products than for the adult products. However, the teams responsible for the forecast and supply planning was aware of the ambitious goals and was able to decrease its risks of expiries and overstocking by monitoring consumption patterns and making appropriate changes to supply and procurement plans.

One way to address policymakers' desire to accommodate scaling up targets is to generate one forecast based on program plans and targets, and another based on historical services data or logistics data. Although both forecasts may be presented together, only the one based on data would be used for actual procurement planning.. For example, we could explain that, "If we achieve 80 percent of the target, demand for this ARV is likely to be 160,000; if we achieve 90 percent, it is likely to be 200,000. The likelihood that we will achieve 80 percent of the target is 70 percent, whereas achieving 90 percent of the target has only a 50 percent chance." This allows decisionmakers to understand and assess their risks in the context of other priorities. It will also allow procurers to decide how much risk they are willing to take when placing their orders.

Although this paper focuses on ARVs, this recommendation could also be applied to other kinds of commodity groups and for any program that intends to expand service drastically in a limited period of time. Targets and hopes must be tempered by the realities of the health systems on the ground.

2. If there is a policy change concerning any products, develop a realistic timeline for implementation and adjust forecast accordingly.

Even when program scale-up is realistic and attainable, there may be certain changes in policies that must be addressed and explained in the forecast. For example, many countries began their national expansion of ART using stavudine as a base in first line regimens. In 2009, the WHO removed stavudine as a recommendation from first line regimens; consequently, many countries adopted policies with either zidovudine or tenofovir as a base for a first line regimen.

Most countries had to consider this policy change during their forecasting exercise. All had ambitious plans to quickly switch existing patients; however, implementation of the new policy took longer than anticipated. This was due, in part, to delays in printing new updated standard treatment guidelines, rolling these guidelines out to all service providers, providing training to both clinicians and pharmacists on the use of these new products, registering new FDCs for the new regimes, releasing donor/government funds for procurement, and having these new products available for procurement and distribution. Furthermore, clinicians were reluctant to change the therapy if patients were stable. This resulted in overestimation and poor accuracy scores for the products that are the key elements in those new regimens. The other two medicines in the regimens were, however, forecasted accurately since they constituted the old regimens.

In addition to new regimens, new products (for existing regimens) must also be considered. Take as an example the introduction of chewable, dispersible pediatric ARV FDCs. Although the regimens did not change, the products that comprised the regimens changed. Many steps must occur from the time of policy change to the time of implementation, as mentioned above.

It is important to have the new products delivered when SDPs are ready and able to use them. Procuring new products must be done with careful consideration of preparing both the supply chain and service providers. There is a danger in procuring new products that arrive in country only to sit in the warehouse at the central level as their shelf life ticks away because the new policies have not yet been implemented, as happened in Country Y.

3. The better the data, the more reliable the forecast.

Every forecast has to deal with different kinds of data and data of varying quality. Understanding which data to collect and how to use them form the underpinning of good forecasts.

The type of data used can impact the forecast. In the analysis conducted here, countries primarily used services data on number of patients by regimen, morbidity data on incidence or prevalence of disease, consumption data on the number of products actually consumed, or a combination of all three. In general, it is recommended to consult and use as many data sources as possible during the forecast. In addition to morbidity, demographic, services, and consumption data, estimated program growth, provider prescribing practices, upcoming policy changes, and total available funding must be considered. A good example of this is in Country X, which uses several different sets of data, compares them, and then decides which to use. This review also assists with the evaluation of data quality.

When data are not available or are of very poor quality, significant assumptions must be made. When actual data are not available, it is not possible to conduct an analysis of forecast error or accuracy. Overall, the more a forecast is assumption-based, the more likely it is to have a greater rate of error.

It is also important to understand the sources of data and the particular biases of each source. Data collected for advocacy purposes to emphasize the importance of a disease and secure more funding for its treatment may be subject to biases that will need to be clearly addressed when applying these numbers to demand forecasts.

Although the type of data used is important, the quality of that data is just as critical. Some common data quality considerations include:

- The age of the data, especially in programs that are scaling up. The older the data, the less reliable they are to predict the future.
- If stockouts have occurred, especially depending on the duration of the stockout. The more stockouts occur, the less reliable are the data.
- The percentage of sites reporting. The lower the percentage reporting, the less reliable are the data and the greater is the forecast error generated, as noted in the Zambia laboratory reagent analysis. Higher reporting rates typically lead to more accurate forecasts, as noted in the Zambia ARV and Country Y laboratory reagents analyses. Furthermore, not all SDPs are equal in terms of number of client visits or quantity of product dispensed. However, if SDPs that have not reported are different from those that have reported (that is, they are known to dispense much larger or much smaller quantities to users or have different patterns of prescribing), this would also compromise the quality of the data.

The importance of available and quality service statistics data on the number and type of health services being provided and logistics data on the consumption and stock levels of commodities for informing the quantification cannot be overstated. The more uncertain the data (or the poorer their quality), the more difficult it is to forecast, which most often leads to greater errors in the forecast. A well-functioning health management information system (HMIS) and logistics management information system (LMIS) are central to improving the accuracy and usefulness of health commodity quantifications.

4. Routinely updating forecasts results in more accurate forecasts.

Quantification is an ongoing process of monitoring, reviewing, and updating the forecasting data and assumptions and recalculating the total commodity requirements and costs as needed. For the quantification exercise to be useful and effective, the forecasting assumptions and the supply plan should be reviewed and updated at least every six months, and more frequently for rapidly growing or changing programs. Ongoing monitoring and updating of the quantification is critical to keep program managers, donors, and other stakeholders informed on the availability of products, and is required for timely decisionmaking about product selection, financing, and delivery of commodities.

The longer in time the forecast projects, the higher the error rate will be. For example, if a forecast is conducted in January 2010, estimated consumption is closer to actual in February and March 2010 than in November and December 2010. If updates to the January forecasts are conducted routinely, then the difference between the estimated consumption and actual consumption will be smaller. In this example, if an update is done in October 2010, and the forecast consumption figures are adjusted appropriately, then November and December actuals will be more accurate than if no update had been done since January. We can see this visually in many of the country examples above, particularly Zambia in figures 21, 23, and 24. The MAPEs in all of these figures increase overall with time.

Many country programs have instituted a quarterly quantification review process for specific commodity categories. These updates should not just modify new shipment arrival dates, but they also should appropriately adjust estimated monthly consumption figures. These periodic updates of the forecast should be rigorous. If the forecast consumption is regularly and rigorously updated, then the overall error rate will be reduced.

Another consideration with routine forecast updates is that some of the assumptions made in the initial forecast may not remain valid in the long run. For example, one key assumption made is about the availability and functionality of the equipment. For forecasts that are for five or seven years, this assumption tends not be valid after the third year because so much has already changed. Most testing equipment needs replacement after three years or so. Therefore, if reagents were forecasted for a certain chemistry machine, such as the Cobas Integra, and it is replaced after three years by the ABX Pentra, the forecasted and procured reagents for the Cobas Integra can't be used with the ABX Pentra and may go to waste if an alternate user is not found. In addition to the uncertainty surrounding the availability of a particular piece of equipment, many programs scale up their testing services, and within the forecast period, the amount and type of equipment available could dramatically change and affect the forecast accuracy. As was seen in the Zambia scenario, uncertainty about the availability of new testing equipment seemed to have a negative effect on forecast accuracy.

5. The more experience with forecasting, the more experience with the program, and the more mature the program, the more reliable the forecast.

Generally speaking, with more time, experience, and maturity, the MAPE decreases. Capacity should be built within those conducting the forecasts, and processes and structures for forecasting should be developed and institutionalized. Over time, as more forecasts and updates are completed, skills in forecasting are increasingly honed, resulting in more reliable forecasts. Similarly, as those conducting the forecast garner more experience with the program, they will be better able to determine the validity of certain assumptions, better able to gauge the implementation time of policy changes, and more aptly adjust collected data to account for any issues with data quality.

The more mature the program, the more stable and predictable consumption will be. As programs mature (and, in fact, to help them mature), investments in LMIS should be made so that higher-quality national consumption and stock-on-hand data are obtained. A well-designed LMIS that is functioning as designed results in quality data being reported up the system. A stable program that is not growing exponentially, and where products are relatively constant, enables forecasts to be completed with fewer errors. Examples of this can be seen in the lower MAPEs with subsequent ARV forecasts in Zambia (figure 21) and in Ethiopia (figure 16), and the higher MAPEs seen in Country Y, where forecasting and LMIS responsibilities fluctuated due to policy changes (figure 9). More mature laboratory programs will also have less frequent changes, such as an increase in new facilities or equipment proliferation, which may in turn impact forecast accuracy.

6. The more prevalent a product or equipment, the less error there is in forecasting.

Generally, error rates are lower with products that are used more often and higher with products used less often. With products that have higher use, there is a more predictable pattern of use, increased experience, and more data collected, even if reporting rates are low. The opposite would be true with products used less often. For the ARVs, with the exception of Ethiopia, the lowest MAPEs were with adult first line ARVs and the higher MAPEs were with pediatric products, second line ARVs, and new FDCs. Higher errors for pediatric formulations may also be attributed to difficulties in predicting dosage forms and dosages prescribed for different age groups and weight bands. This information is difficult to collect with some LMIS.

In looking at the laboratory data from County Y and Zambia, one conclusion is that the more prevalent a piece of equipment, the less error there is in determining the reagent needs for a given period of time. In Country Y's case, the FACSCount machine is more prevalent than the FACSCalibur. Data on the consumption of the key testing reagents for both types of equipment are drawn from the same LMIS. In the forecast accuracy, however, forecasting for the FACSCount was more accurate than that for the FACSCalibur. The same is true in Zambia's case for the same equipment. Similar trends can also be observed across other testing areas such as chemistry and hematology testing in Zambia.

With more prevalent equipment, more consumption data points can be collected and subsequently used for forecasts. Additionally, if a piece of equipment stops functioning within a forecast period, supplies that were ordered can be sent from one facility to another to ensure that testing continues as scheduled. This can only occur when there is a large amount of similar equipment to be used for a particular test. As a result, consumption would stay close to what was forecasted, and wastage of previously forecasted reagents would be minimized.

7. Focus should be on highest-consumed products, while keeping large errors low for those that are consumed less.

The MAPE represents the error in forecasting accuracy, but this analysis does not take into account the quantity of the product consumed. Not all products with high MAPEs pose a problem. For example, in Country Y, SQV had a very large MAPE for all three forecasts (table 8), and its consumption was 11 packages for the month of April 2009. In December 2008, it had been forecasted to be 50 packages. This is a MAPE of 354 percent, but only represents an overforecast by 39 packages of the product. Compare this to a better MAPE of 20 percent for 3TC/D4T/NVP 150/30/200mg for the October 2009 forecast by Country X. The actual consumed quantity for this product for that month was 124,930 packages in comparison to the forecast of 146,885. This is a difference of 21,955 packages, just for one month. It is much easier for a pharmaceutical supply chain to manage the pipeline, storage, and distribution of 39 packages of SQV versus 21,955 packages of 3TC/D4T/NVP 150/30/200mg.

Although the analysis showed high MAPEs for pediatric and second line ARVs, these were consumed at much lower volumes and did not have the highest burden of error when comparing them to the products that were consumed the most, adult first line ARVs.

8. Almost any given tool can make for better—or worse—forecasts.

If a tool calculates accurately, it does not matter which one is used to calculate a forecast. The tool's accuracy and its results are a reflection of the inputs to the quantification (data quality and assumptions used), the methodology used, and how the quantification process is managed.

The various countries in this report used several different tools for their forecasts. For ARVs countries used MSH's Quantimed, CHAI's tool for forecasting pediatric ARVs, or a country-developed database tool. For RTKs, the tools were JSI's ProQ or Quantimed, and for laboratory commodities, there was an Excel spreadsheet in the tool. There was no pattern of one tool producing a better forecast than another. In fact, Country X performed its pediatric ARVs forecast using both CHAI's tool and Quantimed and determined that the results were similar enough not to warrant a change in the tool it had been using for its forecasts (that is, Quantimed). But, in Ethiopia, where the CHAI tool was used again, the forecasters did end up with better accuracy compared to the adult forecasts. These results may need further analysis with a greater number of datasets to make a conclusion.

Commodity-Specific Findings: Labs

1. The fewer the products, the lower the forecasting error.

One of the noted challenges associated with forecasting laboratory commodities is the sheer number of commodities that need to be forecasted at a given time. In countries where the laboratory logistics system is standardized, a national forecast will often exceed 150 unique commodities. In others, where standardization has not taken place, that list could run in excess of 300 line items.

A smaller number of laboratory commodities often leads to a lower forecasting error. In Country Y, for example, the LMIS manages data for fewer than 30 products. In Zambia however, the LMIS manages 185 commodities. One significant reason for this is the fact that fewer assumptions have to be made to cover the entire list of laboratory products to be forecasted. Furthermore, with fewer products, forecasters have a more accurate idea of how they are consumed and the functionality of the equipment on which they are used. A reduced list of commodities to forecast also allows for more accurate data to be collected, aggregated, and analyzed, which, in turn, leads to a more accurate forecast.

2. Improved equipment maintenance decreases the forecasting error.

One of the greatest challenges to maintaining accuracy of a laboratory forecast is ensuring that all of the equipment will be functional during the forecast period. In Country Y's case, all of the CD4 equipment is regularly maintained by the reagent supplier. The supplier is also responsible for fixing the equipment within 24 hours of any failures. Zambia has also made strides to improve the maintenance and service of testing equipment.

When equipment fails, and reagents for that specific equipment cannot be transferred from one facility to another, consumption decreases and the forecast error increases. Additionally, there is wastage of the forecasted reagents that cannot be used. If machines go for extended periods of time without being serviced or maintained, they are more likely to fail on a regular basis or for a longer period. Given the fact that tests cannot be run during the machine's downtime, the accuracy of the forecasts would be drastically affected. Forecasting should therefore take into consideration the state of equipment maintenance and ensure that overforecasting and potential wastage of reagents does not occur.

Conclusion

Several findings were determined based on the analysis conducted on forecasting accuracy of HIV and AIDS commodities, with context from their forecasting process as well as the in-country context. Forecasting accuracy increases as forecasting teams gain experience with and knowledge of the systems and health services and as forecasting functions mature and become institutionalized.

Strengthened LMIS also contributes to greater accuracy and increased ability of the forecasting team to perform effective monitoring. Monitoring forecasts and consumption, in turn, frequently improves forecasts by enforcing a periodic review of data quality, promoting prompt delivery of data, and supporting collection of the complete data needed for decisionmaking.

Greater error was seen in programs that did not have a realistic timeline for implementing policy changes or did not take into consideration the realities of the health systems. The involvement of diverse teams in forecasting may help to curtail inaccurate assumptions and improve forecasting accuracy.

Finally, the focus for increasing accuracy should be on highly consumed products; the greater the accuracy of forecasts for the highly consumed products, the less burden on the supply chain. But large errors should also be addressed for minimally consumed products. For laboratory products, standardization of the laboratory services and logistics systems as well as regular maintenance of testing equipment would decrease forecasting errors.

For more information, please visit deliver.jsi.com.

USAID | DELIVER PROJECT
John Snow, Inc.
1616 Fort Myer Drive, 11th Floor
Arlington, VA 22209 USA
Phone: 703-528-7474
Fax: 703-528-7480
Email: askdeliver@jsi.com
Internet: deliver.jsi.com