Prevention of Mother-to-Child Transmission of HIV data completeness and accuracy assessment in health facilities of the Nkangala District

Authors:
Motlatso G. Mlambo[1,4]
Karl Peltzer[1,2,3]
Annariina Koivu[5]

Affiliations:
[1]HIV/STIs and TB Research Programme (HAST), Human Sciences Research Council (HSRC), South Africa

[2]Department of Psychology, University of Limpopo, South Africa

[3]ASEAN Institute for Health Development, Mahidol University, Thailand

[4]Department of Family Medicine, University of the Witwatersrand, South Africa

[5]INDIGO International Doctorate in Global Health, Trinity College, Ireland

Correspondence to:
Karl Peltzer

Email:
kpeltzer@yahoo.com

Postal address:
Private Bag X41, Pretoria 0001, South Africa

Background: Even though significant progress has been made in the roll-out and quality of the prevention of mother-to-child transmission of HIV (PMTCT) services in South Africa, the quality of patient data recording remains a challenge.

Objectives: To assess PMTCT data completeness and accuracy at primary healthcare level to district level in order to assist with the improvement of the PMTCT data recording.

Methods: This is a retrospective record review study which involved collecting PMTCT data on indicators which was for the period of August 2009 to January 2010. We conducted baseline facility assessments which included 72 PMTCT sites in one health district, Nkangala. We assessed the data completeness and accuracy of the data values recorded on the seven PMTCT data elements.

Results: Data were only complete for less than a quarter of the time for most of the antenatal indicators (0.5% – 44%) and for the maternity indicators, data were only complete 11% of the time. Data inaccuracy was a result of recording of data values in the District Health Information System (DHIS) which were not within 10% of the data values recorded in the case registers. The results show that data were missing from the case registers, monthly summary sheets and DHIS between 30% and 99% of the time and that data elements had values recorded in the DHIS which were > 10%.

Conclusion: There is a need for ongoing training on data recording procedures at all levels. To maintain data quality, healthcare data must be appropriate, organised, timely, available, accurate and complete.

Agterground: Selfs al is beduidende vordering gemaak met die bekendstelling en die kwaliteit van die voorkoming van moeder-na-kind-oordrag van MIV (PMTCT) dienste in Suid-Afrika, bly die kwaliteit van die pasiëntdataopname 'n uitdaging.

Doelwitte: Om PMTCT datavolledigheid en akkuraatheid van primêre gesondheidsorg vlak tot distriksvlak te bepaal ten einde te help met die verbetering van die PMTCT data opname.

Metodes: Dit is 'n terugwerkende hersieningstudie wat die versameling van PMTCT data op aanwysers vir die tydperk van Augustus 2009 tot Januarie 2010, behels het. Ons het basislyn fasiliteitsramings en/of skattings uitgevoer wat 72 PMTCT webwerwe in een gesondheid distrik, Nkangala, ingesluit het. Ons het data volledigheid en akkuraatheid van datawaardes op die sewe PMTCT data-elemente, beoordeel.

Resultate: Data was slegs voltooi vir minder as 'n kwart van die tyd vir die meeste van die voorgeboortesorg aanwysers (0.5% – 44%) en vir die kraamaanwysers, data was slegs volledig vir 11% van die tyd. Onakkuraate data was die resultaat van die opname van datawaardes in die Distrik Gesondheid Inligting System (DHIS) wat nie binne 10% van die datawaardes in die gevalleregisters aangeteken was nie. Die resultate het getoon dat data in die gevalleregisters, maandelikse opsommings en DHIS tussen 30% en 99% van die tyd ontbreek het en dat data-elemente waardes in die DHIS opteken het wat > 10% was.

Gevolgtrekking: Daar is 'n behoefte vir voortgesette opleiding ten opsigte van datavasleggingsprosedures op alle vlakke. Om datakwaliteit te handhaaf, moet gesondheidsorgdata gepas, georganiseerd, tydig, beskikbaar, akkuraat en volledig wees.

Introduction

This study aimed to assess prevention of mother-to-child transmission of HIV (PMTCT) data completeness and data accuracy at the primary healthcare level to district level in order to assist

with the improvement of the PMTCT data recording in the Nkangala district of Mpumalanga Province, South Africa. This will be done by assessing the completeness and accuracy of data values recorded in the seven selected national PMTCT data elements. We assessed antenatal, postnatal, maternity and infant PMTCT data elements.

Background and rationale

The data coming from African countries do not reflect accurately the issues affecting this region (Ndira, Rosenberger & Wetter 2008:489; World Health Organization, Regional Office for Africa [WHO/AFRO] 2002). According to WHO/AFRO, the data provided by routine health systems in the African Region are generally not always available for most countries; and even when they are, they are not always comprehensive, complete or up to date (Ndira *et al*. 2008:489; WHO/AFRO 2002). The information flows around the prevention of mother-to-child transmission (PMTCT) of HIV in South Africa make no exception. In addition to the work done around the data quality issues in the PMTCT area (Mate *et al*. 2009; Mphatswe *et al*. 2012:176), studies from South Africa (Jacucci 2005) have raised questions about data quality regarding the South African District Health Information System (DHIS).

The data quality issues should be considered in the context of the South African health system. South Africa has a history of disease- or condition-specific approach, though the health system can be considered comprehensive (Lush *et al*. 1999:771), with integrated tuberculosis (TB) and HIV programmes, including PMTCT of HIV (Uwimana *et al*. 2012:658). However, it has been argued (Uwimana *et al*. 2012:658) that these programmes are not yet integrated fully into the health system functions as outlined in the WHO framework on health systems (WHO 2000) and other conceptual frameworks for integration analysis (Atun *et al*. 2010:104; Cairncross, Peries & Cutts 1997:20). Hence, the varying quality in collection, reporting and use of data should be discussed with reference to a limited integration of largely vertically-delivered programmes to other health system functions.

Research done in other resource-limited settings has indicated that the quality of data reporting is inconsistent (Freund & Kalumba 1985:137; Kumar 1993:305; Mavimbe, Braa & Bjune 2005:108). Weaknesses that contribute to compromised data quality, as well as suboptimal functioning of the health information system, are shown to stem from data incompleteness and untimely reporting (Chaulagai *et al*. 2005:375; Kimaro & Twaakyondo 2005:189), poor documentation of services provided within the facilities and overly-onerous reporting requirements (Brown 2007; Makombe *et al*. 2008:310; Mate *et al*. 2009). Furthermore, as evidenced by earlier research, the larger the number of data elements to be reported upon, the poorer the quality of the data (Stoops, Williamson & Braa 2003).

Numerous studies have shown that there is a need for improved information systems for monitoring healthcare

services that will assist with better service provision (AbouZahr & Boerma 2005:578; Chan *et al*. 2010; Nash *et al*. 2009:S58; Rommelmann *et al*. 2005:569; Shaw 2005:632). A patient register-based system has been utilised to a large extent by South African primary healthcare facilities in order to record patient data and it continues to experience challenges which point to many inaccuracies. Such has been the case in resource-limited settings whereby indicators for monitoring and evaluation are often hand-tallied from paper-based sources and are reported to people at different levels of the programmatic chain (Joint United Nations Programme on HIV/AIDS [UNAIDS] 2008; WHO 2005). Similar to other countries (AbouZahr & Boerma 2005:578), where routine health information systems have been introduced as a way of improving healthcare system performance, South Africa has also introduced a health information system (HIS) which has brought about some improvements in patient data quality sheets (Garrib *et al*. 2008:549) and data recorded on the DHIS.

Aim of the study

This study aimed to assess the PMTCT data completeness and data accuracy of the primary healthcare level to district level in order to assist with the improvement of the PMTCT data recording in the Nkangala district of Mpumalanga Province, South Africa. This was done by assessing completeness and accuracy of data values recorded in the seven selected national PMTCT data elements. We assessed antenatal, postnatal, maternity and infant PMTCT data elements.

Problem statement

Although improvements have been made in the coverage and quality of PMTCT services in South Africa, the quality of patient data recording remains a challenge, as evidenced by recent research (Mate *et al*. 2009; Mphatswe *et al*. 2012:176)**.** The DHIS was introduced as a routine system to track health service delivery in the public health sector of South Africa (Heywood & Rohde 2001) and to support the collection of data that have been used to calculate the indicators by site, subdistrict, district and province. However, several challenges remain which contribute to inconsistent data quality and limited utilisation of available data and information and data flow bottlenecks (Department of Health 2011). On a more detailed level, the challenges also include data recording discrepancies at facility level on case registers and data being recorded on monthly summary.

Therefore, collecting accurate information on the PMTCT data elements would not only strengthen the existing PMTCT services, but would also broaden areas of further intervention as part of comprehensive patient care. To make evidence-based decisions, high-quality data are required at all levels of the healthcare system. At the facility level, the information is used for care delivery and the details about prior events should be recognised and incorporated into current care decisions (Reid & Wagner 2008:987). Furthermore, data collected at facility level can be used to make decisions and improve the quality of healthcare services (Garrib *et al*. 2008:551–552). Indeed, data analysis is required for advocating, designing,

planning and evaluating public health actions (Sánchez, Lloyd-Smith & Getz 2010; WHO 2004) and to inform health policy and resource allocation (WHO 2003). The global interests in the monitoring of development, as illustrated by the Millennium Development Goals (MDGs), generate pressure for quality and timely data in order to demonstrate country progress (Boerma & Stansfield 2007:779).

Research method and design

Study setting

The study was conducted in the Nkangala district (one of the three districts of the Mpumalanga Province) in order to provide technical support and to strengthen PMTCT service implementation at all the subdistricts that fall under its jurisdiction. A similar assessment was conducted in another district (Gert Sibande district) and it also pointed to serious problems with regard to data recording. The study complemented the existing national PMTCT programme and heeds the international call to make women and children a priority in the fight against the epidemic. The PMTCT cascade in the Nkangala district includes antenatal uptake and HIV testing during pregnancy, provision of ART to pregnant women with HIV, safe childbirth practices, safe infant feeding, infant HIV testing and postnatal-care ARV adherence.

Design

This is a retrospective record review study of PMTCT data elements in Nkangala health facilities. The study was conducted between February 2010 and April 2010 and PMTCT data collected on indicators was for the period of August 2009 and January 2010. As part of the PMTCT record review, data for a period of six months (August 2009 – January 2010) were collected from the case registers in the healthcare facilities, using a tool that was developed with data elements for the PMTCT programme. A total of seven PMTCT data elements on antenatal and postnatal care, maternity and infant indicators were assessed.

Sampling criteria

For the baseline assessment, all 80 PMTCT sites in one health district (Nkangala) were identified. Of these 80 sites, four could not be assessed because the PMTCT site manager was unavailable, even after several visits were made to those healthcare facilities. An additional four sites did not agree to provide us with access to assess data recorded in case registers and monthly summary sheets. No facility was excluded from the facility assessments in Nkangala District. Thus, the final sample comprised 72 facilities (53 clinics, 12 community health centers and seven regional and/or district hospitals).

Data collection methods

With regard to the existing data management system, patient data in the Nkangala district healthcare facilities are primarily collected using a paper-based record system.

There are a number of registers that are used to collect patient PMTCT-related information, including the PMTCT register, tally sheet register, drug register, PCR register, HIV testing register and CD4 count register. The data flow from the facility to district level involves the collation of patient data from various registers and transferring it to the case registers. Data from case registers are then summarised on the monthly summary sheet, which is then sent to DHIS. Patient data are handwritten and the information is captured on the computer at district level.

In addition, separate books were used to record useful patient information for the clinic which was not necessarily submitted to the district for recording on the DHIS system, since this information was not regarded as providing indicators for PMTCT. Patient records were also not stored at a central site, which means that in preparation for data to be submitted to the district, the site manager had to take all the different registers from different points in order to record data for that particular month. In most of the healthcare facilities, the primary healthcare (PHC) site manager is the one responsible for collating all the information from different registers and sending it to the DHIS, with the exception of few facilities where there are clerks dedicated to do the data recording. For the current study, we asked a professional nurse in charge of the PMTCT programme at each site to provide the researchers with all the registers that had PMTCT data and monthly summary sheets. As part of the staff awareness regarding the effectiveness of the PMTCT programme, most of the healthcare facilities had monthly statistics for the PMTCT indicators displayed (using graphs) on the healthcare facility notice boards. A total of five Human Sciences Research Council (HSRC) researchers reviewed the case registers in each of the 72 facilities.

Assessment of data completeness

Data completeness was defined as the extent to which observed records were filled in with data values (Simba & Mwangu 2005). The data completeness process involved checking the proportion of data values recorded on the seven PMTCT data elements to see if it was reflecting on both the PMTCT registers (case registers and DHIS). A data element was considered complete if a value was recorded on both the case registers the DHIS for each of the six months of the study period. Although there are three levels of data recording (case registers, monthly summary sheet, DHIS) the acceptable limit for data completeness for this study was based on having recorded data values in both case registers and DHIS.

To assess the completeness of the data, we reviewed the seven PMTCT data elements, namely:

Antenatal and postnatal indicators:

1. antenatal care (ANC) clients pre-test counselled for HIV at first visit
2. ANC clients tested positive for HIV at first test (new)
3. HIV-positive ANC clients tested for CD4 count
4. ANC clients initiated on Azidothymidine (AZT) at 28 weeks or later.

Maternity indicators:

5. women receiving Nevirapine (NVP) in labour.

Infant indicators:

6. babies given AZT
7. polymerase chain reaction (PCR) test done on babies born to HIV-positive mothers at six weeks or later,that were reported routinely to the DHIS over a six-month period (August 2009 – January 2010).

In order to ascertain the full completeness level, the selected six months were based on the fact that data were already submitted to the DHIS, so they would not be missing from the case registers. Therefore, this means that if data on the case registers were already submitted to the DHIS using monthly summary sheets, then they are complete. We excluded the following indicators from the analysis because of unavailability of data values recorded on them at the time of the study: NVP given to HIV-positive women at the ANC; babies given NVP, HIV-positive mother; PCR test results given. A data element was considered complete if a value was recorded in the case registers and the DHIS for each of the six months of the study period.

Assessment of data accuracy

Data accuracy was defined as the extent of variation between actual and recorded figures in the PMTCT registers (Simba & Mwangu 2005). Assessment of data accuracy involved comparing data values recorded in the DHIS with data values obtained from the individual facility case registers (Mphatswe *et al.* 2012:176). Using the data elements, patient counts were done by the researcher and a professional nurse for the months that were being assessed (August 2009 – January 2010). We considered a data element to be accurate if the value of the data recorded in the DHIS system was within 10% of the value recorded in the case registers. The 10% accuracy cut-off point was based on previous studies regarding data accuracy levels (Garrib *et al.* 2008:549; Mate *et al.* 2009), although for the current study the 10% was calculated by comparing values in the case registers and the DHIS instead of focusing on DHIS data values only. In cases where data were missing from either the case registers or DHIS, we excluded these from the accuracy analysis of the data (Mate *et al.* 2009).

Data analysis

Data obtained from the primary healthcare facilities and the DHIS were entered using the Software Package for the Social Sciences (SPSS) program, version 20 (released 2011, IBM Corp, Armonk, NY). Descriptive statistics were used to calculate the frequencies for the data recorded on each of the seven PMTCT data elements assessed. Data were analysed according to data completeness and data accuracy. Data analysis was stratified by type of facility for the ANC and maternity indicators. For instance, when analysing data on ANC indicators, we excluded the hospitals and, when analysing maternity indicators, we excluded the clinics from the analysis. The degree of accuracy between data reported in

the DHIS and data obtained from individual facility registers was based on a 10% threshold.

Ethical considerations

Ethics approval was obtained from the HSRC ethics committee (Protocol REC 5/11/06/08) and the Mpumalanga Department of Health approved the study. Prior to facility visits, appointments were made with the district coordinator, who in turn liaised with the HIV/AIDS, STIs and TB (HAST) coordinators for all six subdistricts about the HSRC visit to the facilities. The HAST coordinators informed the PHC site managers of the various clinics about the PMTCT baseline assessment. Before conducting the study, the site managers or PMTCT nurses were given background information about the study for purposes of enhancing their understanding about the study and they were also requested to provide informed consent. The district information manager was also asked to provide DHIS data for the period of August 2009 – January 2010. The PMTCT record review took a minimum of two to three hours per site, depending on the ease of access to the records. The nature of the study was emphasised, making sure that the site managers understood that the study was not about their personal evaluation, but was rather an assessment of the PMTCT programme at large.

Results
Data completeness

We assessed the seven PMTCT indicators and found that data completeness varied from one indicator to another. The results show that the data were only complete less than 25% of the time for most of the ANC indicators (with percentages of completeness ranging between 0.5% – 44%) and for the hospitals, data were completed only 11% of the time. The data incompleteness was due to missing values in either the case registers or the DHIS. Of the seven data elements, the best-reported data elements, showing maximum completeness, were: 'ANC clients tested positive for HIV at first test (new)' (*n* = 171; 44%) and 'ANC clients pre-test counselled for HIV at first visit' (*n* = 158; 40.5%). The data elements with very minimal completeness were: 'babies given AZT' (*n* = 2; 0.5%) and 'PCR done on babies born to HIV-positive mothers at
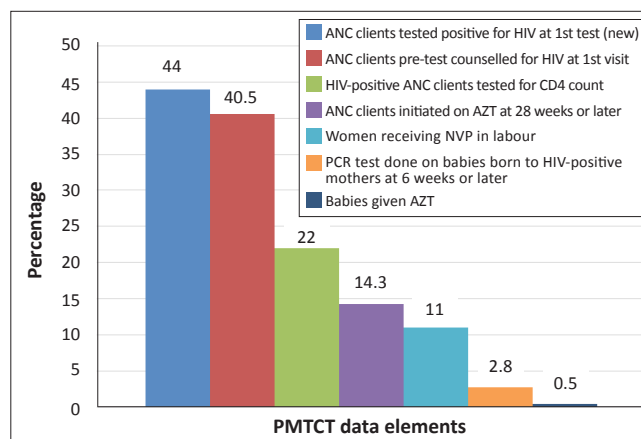
ANC, Antenatal Care; AZT, Azidothymidine; NVP, Nevirapine; PCR, Polymerase chain reaction; PMTCT, prevention of mother-to-child transmission of HIV.

FIGURE 1: Percentage of prevention of mother-to-child transmission of HIV data completeness.

six weeks or later' ($n = 11$; 2.8%). Therefore, data were only completed between 0.5% and 44%, as is reflected in Figure1.

Data completeness by month

The results show varying data completeness pattern of data recorded by month. The data element showing the highest data completeness was for the month of November 2009, on 'ANC clients tested positive for HIV at first test (new)' ($n = 36$; 55.4%), followed by 'ANC clients pre-test counselled for HIV at first visit' ($n = 29$; 44.6%), with the peak of data completeness being for the months of August andSeptember 2009. The data element on 'babies given AZT' was only recorded for two months, namely, August and September 2009. For this data element, the values recorded in the DHIS for the period of October 2009 – January 2010 were > 10% (Table 1).

Missing data

The results show that data were missing from case registers, monthly summary sheets and the DHIS between 30% and 99% of the time (Figure 2). The data elements with the highest missing data were: 'babies given AZT' ($n = 389$; 99%)', 'PCR done on babies born to HIV-positive mothers at six weeks or later' ($n = 379$; 97.1%) and 'ANC clients initiated on AZT at 28 weeks or later' ($n = 329$; 84.3%).

Data accuracy and inaccuracy

The results show that there were data elements which had accurate data values falling within 10% accuracy. The data accuracy in this case ranged from 0.5% to 44%. There were also data elements that showed values with > 10% deviation and these ranged from 0.5% to 17% (Figure 3). The data elements with more accurate data values were: 'ANC clients tested positive for HIV at first test (new)' ($n = 171$; 44%), followed by 'ANC clients pre-test counselled for HIV at first visit' ($n = 158$; 40.5%). The three data elements showing > 10% value recorded in the DHIS as compared to original source were: 'ANC clients pre-test counselled for HIV at first visit ($n = 66$; 17%), 'ANC clients tested positive for HIV at first test (new)' ($n = 62$; 16%) and 'HIV-positive ANC clients tested for CD4 count' ($n = 44$; 11.2%).

Discussion

The results revealed that patient data recording in the Nkangala district remains a challenge because of the lack of accurate data recording by the healthcare professionals. We found that the inaccurate data recording stems from the lack of a double-checking system for the data recorded on the monthly summary sheets before they are submitted to the district to be recorded in the DHIS. There are a number of registers that are used to collect patient PMTCT-related information including the PMTCT register, tally sheet register, drug register, PCR register, HIV testing register and CD4 count register. All of these registers cover the information that is recorded on the case registers and that is summarised on the monthly summary sheet which is then sent to DHIS. The results indicate that because of many PMTCT registers, there is non-verification of data that are being transferred from one register to another. When assessing data completeness and accuracy for the seven PMTCT data elements, we found that the best reported data element, showing maximum completeness and accuracy, was 'ANC clients tested positive for HIV at first test (new)'. Furthermore, the results show that, for the hospitals in the study, there was minimal completion of data. Our findings regarding the incompleteness and inaccuracy of the PMTCT data are in accordance to other studies conducted in South Africa (Mate *et al.* 2009; Mphatswe *et al.* 2012:176). This is worrying considering that the unreliability of data compromises the efforts to improve the healthcare systems that provide PMTCT interventions (Mate *et al.* 2009). The
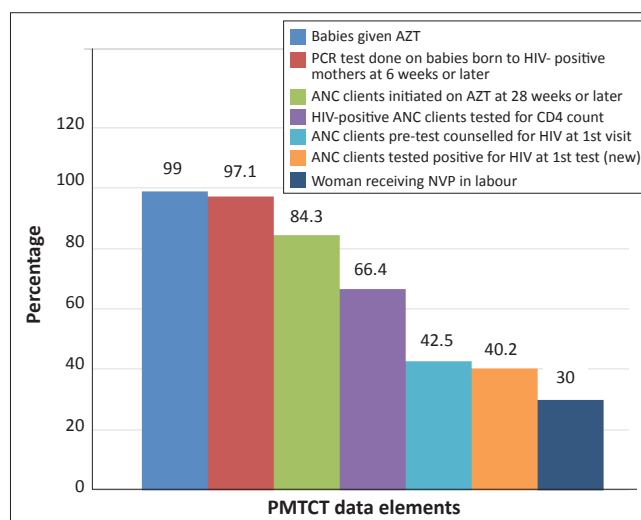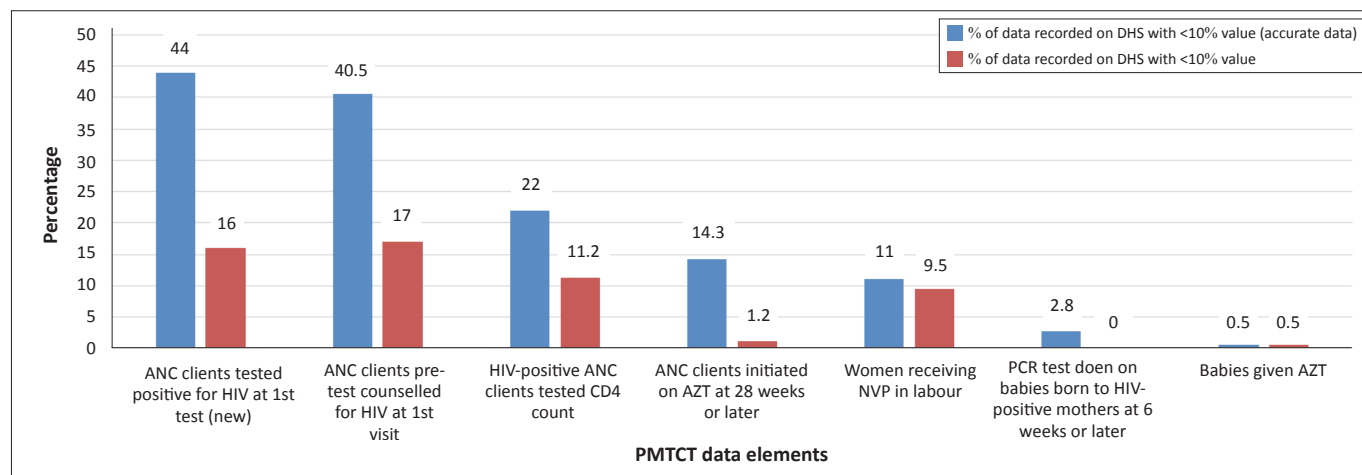

PMTCT, prevention of mother-to-child transmission of HIV; ANC, Antenatal Care; PCR, Polymerase chain reaction; AZT, Azidothymidine; CD4 Cluster of differentiation (CD4).

FIGURE 2: Percentage of missing data values on case registers, monthly summary sheet and District Health Information System.

TABLE 1: Data completeness by month August 2009 – January 2010.

Months	ANC clients pre-test counselled for HIV at 1st visit		ANC clients tested positive for HIV at 1st test (new)		HIV-positive ANC clients tested for CD4 count		ANC clients initiated on AZT at 28 weeks or later		Women receiving NVP in labour		Babies given AZT		PCR test done on babies born to HIV-positive mothers at 6 weeks or later	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
August 2009	29	44.6	26	40	14	21.5	9	14	1	14.2	1	1.5	4	6.1
September2009	29	44.6	28	43	15	23	9	14	2	28.6	1	1.5	2	3
October 2009	27	41.5	23	35.3	15	23	8	12.3	1	14.2	-	-	2	3
November 2009	28	43	36	55.4	15	23	11	17	2	28.6	-	-	1	1.5
December 2009	23	35.3	30	46.1	13	20	10	15.3	1	14.2	-	-	1	1.5
January 2010	22	34	28	43	15	23	9	14	2	28.6	-	-	1	1.5

ANC, Antenatal Care; PCR, Polymerase chain reaction; AZT, Azidothymidine; CD4, Cluster of differentiation; NVP, Nevirapine.

ANC, Antenatal Care; PCR, Polymerase chain reaction; AZT, Azidothymidine; CD4, Cluster of differentiation; NVP, Nevirapine; PMTCT, prevention of mother-to-child transmission of HIV; DHIS, District Health Information System.

FIGURE 3: Percentage of data accuracy and inaccuracy by data element.

current study concurs with findings from other studies (Garrib *et al*. 2008:549) which found that data validation to ensure data completeness and accuracy was limited.

We also found that the two data elements with minimal completeness were on infant indicators, namely, 'babies given AZT' and 'PCR test done on babies born to HIV-positive mothers at six weeks or later'. The lower data recording of values for the two data elements could be as a result of no service provision during the assessment period (August 2009 – January 2010). Most of the antenatal facilities were not providing AZT or even doing infant PCR testing. If they were providing that service, there was very little recording for the two data elements during our assessment period. However, it is worth noting that most of the healthcare facilities of the Mpumalanga Province are now giving AZT to the infants and they do infant PCR testing at six weeks. The study conducted by Sherman and Lilian (2011) showed improved PCR testing rates from 18.4% in 2008 to 39.3% in 2010. When assessing data completeness by month, the two data elements with the highest completeness were 'ANC clients tested positive for HIV at first test (new)' and 'ANC clients pre-test counselled for HIV at first visit', which do not show any significant difference when all six months' data are combined.

In this study we found that data recording inaccuracy was common at all levels (PHC to district). As with the reason for data incompleteness, data inaccuracy was also as a result of the over-recording of data values in the DHIS (by over 10% when compared with the data values recorded in the case registers and monthly summary sheets). We found that the over-recording of the values on data elements were mainly in the DHIS. Other studies also suggest that even though data are recorded in the DHIS, the quality is not good (Mphatswe *et al*. 2012:176). In order for patient data to be useful, they must be valid and accurate (WHO 2003). The results for this study show that patient data are clearly not accurate since DHIS data are collected from many sources and thus may not always be reliable. Another worrying factor is that we found that data were missing from case records, monthly spreadsheets and the DHIS, with the

highest missing data element being on 'babies given AZT'. In a study conducted in KwaZulu Natal, South Africa, of the total data values that had been collected, they found a mean of 2.5% missing data in the DHIS and healthcare facilities had not provided any explanation for the omission (Garrib *et al*. 2008:549). This suggests that patient data recording is a weakness which requires immediate intervention in order for the PMTCT service provision to improve significantly. We concur with the findings of other studies which point out that regular audit of health data could improve data recording significantly. A study conducted in South Africa found significant improvements in all key PMTCT indicators after PMTCT intervention, which included conducting a roadshow to orient staff on PMTCT protocol and monthly support visits by a senior experienced professional nurse in order to assess routine programme indicators (Doherty *et al*. 2009:406).

There is, therefore, a need to perform interventions aimed at improved data recording quality. Other studies have demonstrated that such interventions work in under-resourced countries (Mphatswe *et al*. 2012:176). After an intervention that included training on data collection, monthly reviews of data and data audits at individual facilities, a study from South Africa found that data completeness in the studied antenatal care clinics improved from 25% (before the intervention) to 63% (after the intervention) and in the study hospitals, from 28% to 86%. The overall data accuracy improved from 35% to 65%. Reviewing data quality and completeness could be done on a monthly basis during monthly district programme review meetings (Mphatswe *et al*. 2012:176). Other studies in under-resourced countries also reported on interventions that could lead to the improvement of data quality and completeness. For instance, in Kenya, the onsite assessments and feedback on data completeness and validity were used to improve the quality of routine health information. As a result, the number of HIV tests recorded, as well as the number of HIV-positive patients, doubled (Otwombe *et al*. 2007:315). In some settings, data improvements were linked to clinic visits by supervisors

(Huaman *et al*. 2009:16). Several studies (Berwick 2004:1124; Center for Applied Research on Population and Development 2003) advocated for quality improvement interventions using a data-driven approach to ensure sustainable improvements in health programmes. Other studies indicate that in order to ensure acceptable completeness and accuracy of patient data, ongoing review of monitoring and evaluation (M&E) indicators is required at multiple levels of the healthcare system (site, regional and national levels) (Nash *et al*. 2009:S58). The WHO recommends the development of ongoing quality assessment plans which involve identifying important aspects of data collection (accuracy, validity, reliability, completeness and timeliness) and obtaining a threshold to determine acceptable error rate (WHO 2003).

It is not sufficient to focus only on interventions that improve data quality. In many countries there remains an inadequate appreciation and use of available information in planning and management of health services (Chaulagai *et al*. 2005:375). In other words, if the use of the collected PMTCT data elements is limited, it undermines any effort to improve the different aspects of data quality. Moreover, further research may be beneficial with regard to determining whether the local health professionals perceive the current data recording system adequate or to investigate if any perceived deficiencies in the data recording system may contribute to the atmosphere which, in turn, contributes to suboptimal data entry into the DHIS. Kawonga, Blaauw and Fonn (2012) argue that the key drivers of the silo approach are specific reporting requirements linked to ear-marked funding, politically-driven imperatives and mistrust of DHIS capacity. They conclude that the parallel systems that bypass the DHIS represent a missed opportunity to strengthen system-wide M&E capacity. Finally, they recommend the integration of HIV M&E into the health system. M&E functions such as these would mobilise earmarked HIV funding toward improving DHIS capacity so as to produce quality and timely HIV indicators (Kawonga *et al*. 2012). Hence, it should be explored to what extent the possibly parallel data requirements of the PMTCT programme or by the non-governmental organisations working in PMTCT may contribute to the data quality issues in the DHIS. Although we found a lot of data inaccuracy and incompleteness in our study, we did not provide any interventions because of funding limitations. The study recommendations were made available to the district.

Limitations of the study

Study limitations include that the study was conducted in only one district of the Mpumalanga Province, thus findings cannot be generalised to every district in South Africa. It cannot, therefore, be assumed that similar issues with data incompleteness and inaccuracy are present in other districts.

Recommendations

In this study, we found that patient data recording varies between different PMTCT data elements whereby antenatal indicators seem to be more recorded than maternity and infant indicators. There is a need for ongoing training on data recording at all levels, which support on DHIS is critical for it to function properly.

Conclusion

To maintain data quality, healthcare data must be appropriate, organised, timely, available, accurate and complete (Abdelhak *et al*. 1996). Data entry checks are also critical for accurate data reporting (WHO 2003). For this reason, having dedicated personnel in each facility would improve data recording quality significantly.

Acknowledgements

Competing interests

The authors declare that they have no financial or personal relationship(s) which may have inappropriately influenced them in writing this article.

Authors' contributions

M.G.M. (Human Sciences Research Council) was the project manager and participated in the conceptualisation and design of the study. She also analysed and interpreted data and developed a draft article. K.P. (Human Sciences Research Council) was the Principal Investigator for the study. He participated in the design of the study, data analysis and interpretation and writing of the article. A.K. (Trinity College) participated in writing the draft for the article by conducting literature review. All authors corrected and approved the article.

References

Abdelhak, M., Grostick, S., Hankin, M.A. & Jacobs, E., 1996, *Health information: Management of a strategic resource*, WB Sanders Company, Philadelphia.

AbouZahr, C. & Boerma, T., 2005, 'Health information systems: The foundations of public health', *Bulletin of the World Health Organization* 83(8), 578–583.

Atun, R., De Jongh, T., Secci, F., Ohiri, K. & Adeyi, O., 2010, 'Integration of targeted health interventions into health systems: A conceptual framework for analysis', *Health Policy and Planning* 25(2), 104–111. http://dx.doi.org/10.1093/heapol/czp055

Berwick, D.M., 2004, 'Lessons from developing nations on improving health care', *BMJ* 328(7448), 1124–1129. http://dx.doi.org/10.1136/bmj.328.7448.1124

Boerma, J.T. & Stansfield, S.K., 2007, 'Health statistics now: Are we making the right investments?', *The Lancet* 369(9563), 779–786. http://dx.doi.org/10.1016/S0140-6736(07)60364-X

Brown, W., 2007, *Data quality assurance tool for program level indicators*, Measure Evaluation, Chapel Hill, NC.

Cairncross, S., Peries, H. & Cutts, F., 1997, 'Vertical health programmes', *The Lancet* 349 (Suppl III), 20–22. http://dx.doi.org/10.1016/S0140-6736(97)90079-9

Center for Applied Research on Population and Development, 2003, *Using data to improve service delivery: A self evaluation approach*, CERPOD, Bamako, Mali.

Chan, M., Kazatchkine, M., Lob-Levyt, J., Obaid, T., Schweizer, J., Sidibe, M., *et al*. 2010, 'Meeting the demand for results and accountability: A call for action on health data from eight global health agencies', *PLoS Medicine* 7(1), e1000223. http://dx.doi.org/10.1371/journal.pmed.1000223

Chaulagai, C.N., Moyo, C.M., Koot, J., Moyo, H.B., Sambakunsi, T.C., Khunga, F.M. *et al*., 2005, 'Design and implementation of a health management information system in Malawi: Issues, innovation and results', *Health Policy and Planning* 20(6), 375–384. http://dx.doi.org/10.1093/heapol/czi044

Department of Health, 2011, *District Health Management Information System (DHMIS) Policy 2011*, viewed 14 June 2014, from http://www.hst.org.za/sites/default/files/District Health Management Information System Policy.pdf

Doherty, T., Chopra, M., Nsibande, D. & Mngoma, D., 2009, 'Improving the coverage of the PMTCT programme through a participatory quality improvement intervention in South Africa', *BMC Public Health* 9, 406. http://dx.doi.org/10.1186/1471-2458-9-406

Freund, P.J. & Kalumba, K., 1985, 'Monitoring and evaluation of primary health care in rural Zambia. A comparative study', *Scandinavian Journal of Social Medicine* 13(4), 137–146.

Garrib, A., Stoops, N., McKenzie, A., Dlamini, L., Govender, T., Rohde, J. *et al*., 2008, 'An evaluation of the District Health Information System in rural South Africa', *South African Medical Journal* 98(7), 549–552.

Heywood, A. & Rohde, J., 2001, *Using information for action. A manual for health workers at facility level*, The EQUITY Project, Pretoria.

Huaman, M.A., Araujo-Castillo, R.V., Soto, G., Neyra, J.M., Quispe, J.A., Fernandez, M.F. *et al*., 2009, 'Impact of two interventions on timeliness and data quality of an electronic disease surveillance system in a resource limited setting (Peru): A prospective evaluation', *BMC Medical Informatics and Decision Making* 9, 16. http://dx.doi.org/10.1186/1472-6947-9-16

Jacucci, E., 2005, 'Duality of control in large-scale IS: The case of the health information system in South Africa', in G. Schreyögg, J. Sydow, S. Quack & H. Rudolph (eds.), *Proceedings of the 21st EGOS Colloquium*, Berlin, Germany.

Joint United Nations Programme on HIV/AIDS (UNAIDS), 2008, *Organizing framework for a functional national HIV monitoring and evaluation system*, UNAIDS, Geneva.

Kawonga, M., Blaauw, D. & Fonn, S., 2012, 'Aligning vertical interventions to health systems: A case study of the HIV monitoring and evaluation system in South Africa', *Health Research Policy and Systems* 10, 2. http://dx.doi.org/10.1186/1478-4505-10-2

Kimaro, H.C. & Twaakyondo, H.M., 2005, 'Analysing the hindrance to the use of information and technology for improving efficiency of health care delivery system in Tanzania', *Tanzania Health Research Bulletin* 7(3) 189–197.

Kumar, R., 1993, 'Streamlined records benefit maternal and child health care', *World Health Forum* 14(3), 305–307.

Lush, L., Cleland, J., Walt, G. & Mayhew, S., 1999, 'Integrating reproductive health: Myth and ideology', *Bulletin of the World Health Organization* 77(9), 771–777.

Makombe, S.D., Hochgesang, M., Jahn, A., Tweya, H., Hedt, B., Chuka, S., *et al*. 2008, 'Assessing the quality of data aggregated by antiretroviral treatment clinics in Malawi', *Bulletin of the World Health Organization* 86(4), 310–314. http://dx.doi.org/10.2471/BLT.07.044685

Mate, K.S., Bennett, B., Mphatswe, W., Barker, P. & Rollins, N., 2009, 'Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa', *PLoS ONE* 4(5), e5483. http://dx.doi.org/10.1371/journal.pone.0005483

Mavimbe, J.C., Braa, J. & Bjune, G., 2005, 'Assessing immunization data quality from routine reports in Mozambique', *BMC Public Health* 5, 108. http://dx.doi.org/10.1186/1471-2458-5-108

Mphatswe, W., Mate, K.S., Bennett, B., Ngidi, H., Reddy, J., Barker, P.M. *et al*., 2012, 'Improving public health information: A data quality intervention in KwaZulu Natal, South Africa', *Bulletin of the World Health Organization* 90, 176–182. http://dx.doi.org/10.2471/BLT.11.092759

Nash, D., Elul, B., Rabkin, M., Tun, M., Saito, S., Becker, M. *et al*., 2009, 'Strategies for more effective monitoring and evaluation systems in HIV programmatic scale-up in resource-limited settings: Implications for health systems strengthening', *Journal of Acquired Immune Deficiency Syndromes* 52(Suppl 1), S58–S62. http://dx.doi.org/10.1097/QAI.0b013e3181bbcc45

Ndira, S.P., Rosenberger, K.D. & Wetter, T., 2008, 'Assessment of data quality of and staff satisfaction with an electronic health record system in a developing country (Uganda): A qualitative and quantitative comparative study', *Methods of Information in Medicine* 47(6), 489–498.

Otwombe, K.N., Wanyungu, J., Nduku, K. & Taegtmeyer, M., 2007, 'Improving national data collection systems from voluntary counselling and testing centres in Kenya', *Bulletin of the World Health Organization* 85(4) 315–318. http://dx.doi.org/10.2471/BLT.06.033712

Reid, R.J. & Wagner, E.H., 2008, 'Strengthening primary care with better transfer of information', *Canadian Medical Association Journal* 179(10), 987–988. http://dx.doi.org/10.1503/cmaj.081483

Rommelmann, V., Setel, P.W., Hemed, Y., Angeles, G., Mponezya, H.,Whiting, D. *et al*., 2005, 'Cost and results of information systems for health and poverty indicators in the United Republic of Tanzania', *Bulletin of the World Health Organization* 83(8), 569–577.

Sánchez, M.S., Lloyd-Smith, J.O. & Getz, W.M., 2010, 'Monitoring linked epidemics: The case of tuberculosis and HIV', *PLoS ONE* 5(1), e8796. http://dx.doi.org/10.1371/journal.pone.0008796

Shaw, V., 2005, 'Health information system reform in South Africa: Developing an essential data set', *Bulletin of the World Health Organization* 83(8), 632–636.

Sherman, G. & Lilian, R., 2011, *Early infant diagnosis of HIV infection in South Africa: 2008 to 2010*, National Health Laboratory Service and Department of Molecular Medicine and Haematology, University of the Witwatersrand.

Simba, D.O. & Mwangu, M.A., 2005, 'Quality of a routine data collection system for health: Case of Kinondoni district in the Dar es Salaam region, Tanzania', *South African Journal of Information Management* 7(2), 8 pages.

Stoops, N., Williamson, L. & Braa, J., 2003, 'Using health information for local action: Facilitating organisational change in South Africa', in S. Krishna & S. Madon (eds.), *The digital challenge: Information technology in the development context,* pp. 156–169, Athenaeum Press Ltd., Gateshead.

Uwimana, J., Jackson, D., Hausler, H. & Zarowsky, C., 2012, 'Health system barriers to implementation of collaborative TB and HIV activities including prevention of mother to child transmission in South Africa', *Tropical Medicince & International Health* 17(5), 658–665. http://dx.doi.org/10.1111/j.1365-3156.2012.02956.x

World Health Organization, 2000, *The World Health Report 2000 – health systems: Improving performance*, WHO, Geneva.

World Health Organization, 2003, *Improving data quality: A guide for developing countries*, WHO, Geneva.

World Health Organization, 2004, *Guidelines for HIV surveillance among tuberculosis patients*, 2nd edn., WHO, Geneva.

World Health Organization, 2005, *National AIDS programmes: A guide to indicators for monitoring and evaluating national antiretroviral programmes*, WHO, Geneva.

World Health Organization, Regional Office for Africa (WHO/AFRO), 2002, *Report on strengthening national health information systems: workshop on the use of Geographical Information System HealthMap*, Harare, Zimbabwe, July 01–04, viewed 10 June 2014, from http://www.equinetafrica.org/bibl/docs/WHOehs02.pdf