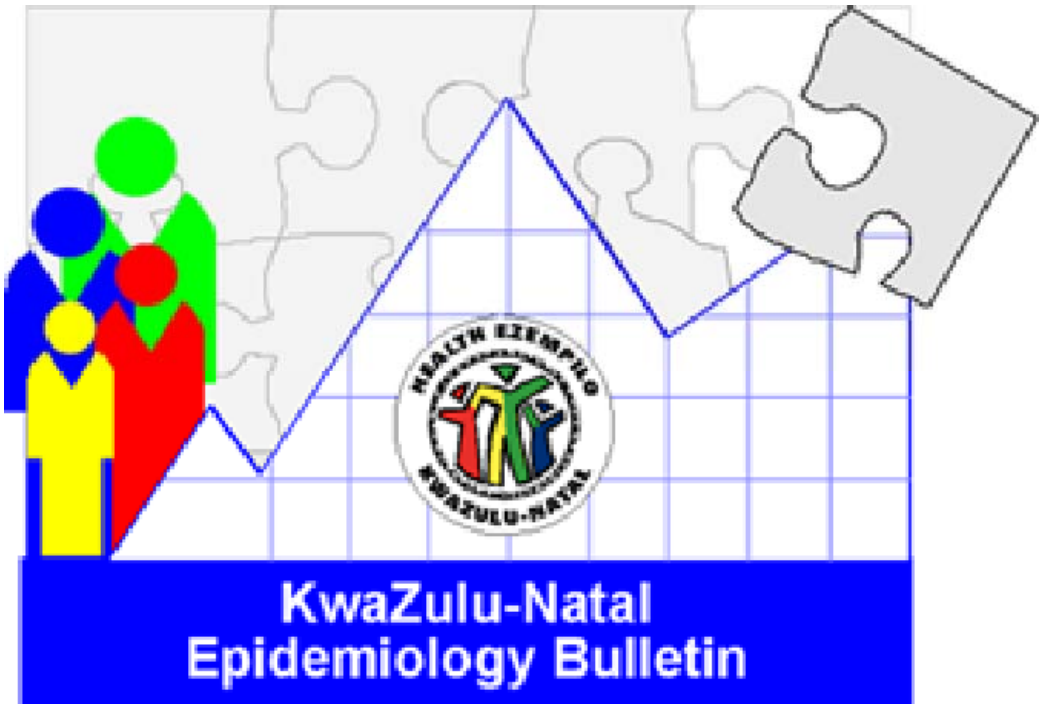


Issue 11

September 2005

Evaluation of the Hospital Information System in KwaZulu-Natal

Published Quarterly by:
The Epidemiology Unit
KwaZulu-Natal Department of Health
Private Bag X9051
Pietermaritzburg 3200

This issue has been sponsored by the Italian Co-operation

EDITORIAL

I feel truly honoured to write the editorial to the 11th issue of the Epidemiology bulletin. One of the greatest challenges facing the delivery of service in our hospital is the availability of reliable data that will inform us on how best to allocate our resources. In the last issue, we introduced the validation methodology, which served to provide a clear understanding of hospital indicators. We had set ourselves the task of ensuring that we are able to account for bed occupancy in our public hospitals. We have equally looked at the issue of admission, discharges, death, deliveries and incubators to inform us what we consider as pertinent to our definition of occupancy.

We have put various measures in place to ensure that the decision making process of hospital management is informed through a reliable system that will guarantee that we do not underreport nor do not over report on the state of our hospitals. In 2002 the Patient Throughput Statistical System (PTSSH) was introduced as a hospital data to assist us with reliable data.

We are hoping that this bulletin and the findings of the study will provide you with a better analysis of what to calculate and consider when counting useable beds and other data required to inform our hospital information system and at the same time, this analysis will help reduce misinterpretation. Over and above this, we are convinced that this analysis will once more reduce several data iterations between the district and the hospital, so that what is sent to Head Office is untainted data. This system, we believe, will create a more timely correction of data problems. Our hospitals are full to the brim and we are doing what is best; to strengthen Primary Health Care (PHC) and improve the efficiency of our hospitals.

We continue to pride ourselves with striving to provide optimal health status to all persons in KwaZulu Natal. I am very hopeful that in time, our work will begin to manifest itself in the reduction of the number of people that wind up taking up beds in our hospitals. Our objective is to ensure that we admit those patients that require specialized attention and it is our intention to discharge patients to be well enough to resume their normal life.

This bulletin provides a clear analysis of various correlations. Table 1 on page 15, provides Correlation coefficients between the validation and the data sources. You will also see the discrepancy in the calculation of adult beds versus juvenile beds. This analysis therefore will help with frequently asked questions and ensure that in future we disburse reliable information.

I am convinced that this bulletin will be informative as it is well thought after. The revelations that we may be overlooking valuable data in our reporting clearly indicate the need to engage each other through consultative research.

Nhlanhla Nkosi
Chief Operating Officer
Integrated Health Service Cluster

TABLE OF CONTENTS

ACRONYMS & DEFINITION OF TERMS

ALOS — The Average Length of Stay is the patient average number of days spent in hospital. From the formula below, it can be noted that an over-reporting of the numerator (inpatient days) will be associated with over-estimate of ALOS and vice versa. By the same token, an over-reporting of the denominator will be associated with underestimation of ALOS and vice versa.

$$\frac{(\text{½ Day Patients + Inpatient Days + Transfers-in})}{(\text{Discharges + Deaths +Transfers out})}$$

Case mix — This term is used to describe the complexity of the inpatients. Hospitals having a higher proportion of complex cases (i.e. Regional hospitals) need to be assessed separately from hospitals admitting less complex cases (i.e. district hospitals).

DHIS — District Health Information System

DIO — District Information Officer

DOH — Department of Health.

Efficiency — In this issue, efficiency is used to indicate the optimal utilization of resources in terms of high occupancy and high turnover.

FIO — Facility Information Officer

FY — Financial Year

KZN — KwaZulu-Natal

Occupancy — The proportion of occupied beds is measured in this issue by:

$$\frac{(\text{½ Day Patients + Inpatient Days + Transfers-in})}{(\text{Useable beds * days in a month or a year})}$$

Patient throughput — In this issue, the throughput is the number of patients transiting the hospital system. On average, the sum of the patients coming into the hospitals must balance the patients leaving the hospitals. The patients coming in are the sum of admissions+transfers-in and the patients leaving are the sum of discharges + deaths + transfers-out. The monthly variation of the entries and exits must also be consistent with the inpatient days.

PIC — Provincial Information Coordinator

PTSSH Patient Throughput Statistical system

.

Reliable A measurement is reliable if it provides the same value when it is repeated under identical conditions. For example, because the number of beds is unlikely to change from month to month, high monthly fluctuations in the hospital beds are almost always a sign of unreliability. If there are no real changes and the reported number of beds is unreliable, the occupancy and turnover will have false fluctuations because the number of beds will affect the denominator of these indicators.

Turnover In this issue turnover is the average number of patients per bed per month and it is obtained by:

$$\frac{\text{(Discharges + Deaths + Transfers out)}}{\text{Useable Beds}}$$

Validity Validity is the ability to report real values. For example, a hospital that has always had 300 beds across the years, may have reliably reported all the time the same invalid lower value of 250 beds.

AKNOWLEDGMENT

This issue of the Epidemiology Bulletin is a joint product of the KwaZulu-Natal Department of Health and the Italian Cooperation, which provided the technical support and paid for the printing costs. Dr Venanzio Vella, Epidemiologist of the Italian Co-operation planned the survey. Mr Alastair Van Heerden supervised the data collection, which was carried out by Miss Hlengiwe Mhlophe and Miss Lindiwe Farlane. The analysis and the write up were done by Dr Venanzio Vella and Mr Alastar Van Heerden. The collaboration of the nurses, matron and FIOs was critical to the success of the validation.

Abstract

This issue describes the fully-fledged application of the validation methodology, which was introduced in Issue 8. The validation was needed to find out the causes for the problems affecting the reliability of the hospital indicators and suggest solutions. After a description of how the hospital information system works and how the methodology was applied in all the public hospitals of KwaZulu-Natal, the Issue presents the results of the analysis.

Hospital Information System

The reliability of the information collected on the patients' throughput is critical for the correct estimation of the indicators of utilization. As discussed in previous issues, unreliable information is due to lack of standardization in data reporting. Training and supervision are never sufficient to ensure reliability because of the inevitable subjective interpretation of the many people involved. A validation was therefore conducted to assess what are the most frequent causes of unreliability and what can be done to improve the system.

Methodology

All the 67 public hospitals of KwaZulu-Natal were validated. The validation team counted the number of beds in each ward, the patient throughput related to a period of 48 hours, the number of operations and outpatients of the previous month. The same team visited each hospital after a few months from the first visit to compare valid with reported changes between two points in time.

Findings

Over-estimation of occupancy and turnover was more common than under-estimation. The numbers of admissions, discharges, deaths, deliveries and adult beds were more reliably reported than the numbers of transfers, juvenile beds, cribs and incubators. Most hospitals over-estimated their utilization because of an under-reporting of beds and or an over-reporting of inpatient days and discharges. Unreliable reporting was due to different interpretation about what should be included and excluded in the counting of each variable such as useable bed, and how consistently the criteria were applied within each hospital and across hospitals. Another source of variation was due to the file transformation when the data were sent from each hospital to the district level and from the district to the central level.

Solutions

The solutions include a better standardization of the data production within each hospital and the creation of a more direct link between the periphery and the centre. The over or under-estimation of occupancy and turnover across the hospitals impair the correct measurement of their utilization. This does not allow one to reliably rank the hospitals in terms of efficiency and measure trends in the indicators. This is critical to identify causes for low efficiency, find solutions and monitor changes. The efforts should be prioritised to reduce the variation of the reporting of the number of beds, inpatient days, discharges and deaths. To decrease errors and delays, besides

sending the files to the district level, the FIO should also send them directly to the central level. This will allow the central level to apply standardized criteria to identify outliers and to timely alert the DIOs and FIOs about data problems so that they will timely recheck the data. This will reduce the variation due to data manipulation between each hospital and the district level and will create a more timely correction of the data.

Recommendations

Improvement of reliability of the hospital statistics will depend on the following:

(a) Clearer guidelines on how the several data recording techniques should be applied;
(b) A survey carried out at the beginning of each FY should count the useable beds in each hospital, to provide constant denominators for occupancy and turnover for the whole FY;
(c) The planned move from the PTSSH into the DHIS should go ahead as soon as possible to decrease the variation caused by file manipulation;
(d) The FIO should directly send the monthly data also to the central level where one person should be assigned full time to identify outliers and timely alert the relevant DIO and FIO to recheck the outlier data;
(e) The central level should prioritize the effort to improve first the reliability of inpatient days, discharges and deaths before attempting to improve the reporting of the other variables; and
(f) The central level should make more use of occupancy and turnover in graphical forms (i.e. Pabon Lasso) to identify profiles of efficiency, formulate hypotheses on causes to find solutions and test them by monitoring their impact.

Introduction

The Patient Throughput Statistical System (PTSSH) processes the hospital data to inform the decision making process of hospital management. This is based on the reporting of individual variables, such as the number of useable beds and occupied beds that are then transformed into indicators of utilization such as in the case of occupancy. Because hospitals consume the lions' share of the budget, high utilization is an indicator of how efficiently resources are utilized. This is based on the assumption that everybody across the hospitals applies consistently the data collection criteria. If this does not happen, the estimation of the indicators will be inconsistent, impairing the comparison of efficiency between different periods of time within the same hospital and at a given point in time among hospitals. This is not without consequences for two reasons (a) the costs involved in setting up the information system will be wasted and (b) inaccurate indicators cannot be used to monitor efficiency, to find causes of low utilization, to apply solutions and to monitor their impact.

Because of the above reasons, the attention of the previous Issues of the Bulletin was focused on the issues of validity and reliability and the use of hospital indicators. As described in Issue 7 through 9 of the Epidemiology Bulletin, the reporting from the hospitals has been characterized by inexplicable variation and discrepancies. These have included wide monthly fluctuations in the number of beds and mismatches between the patients coming in and leaving the hospitals. This has resulted in inaccurate occupancy and turnover, impairing the interpretation of utilization.

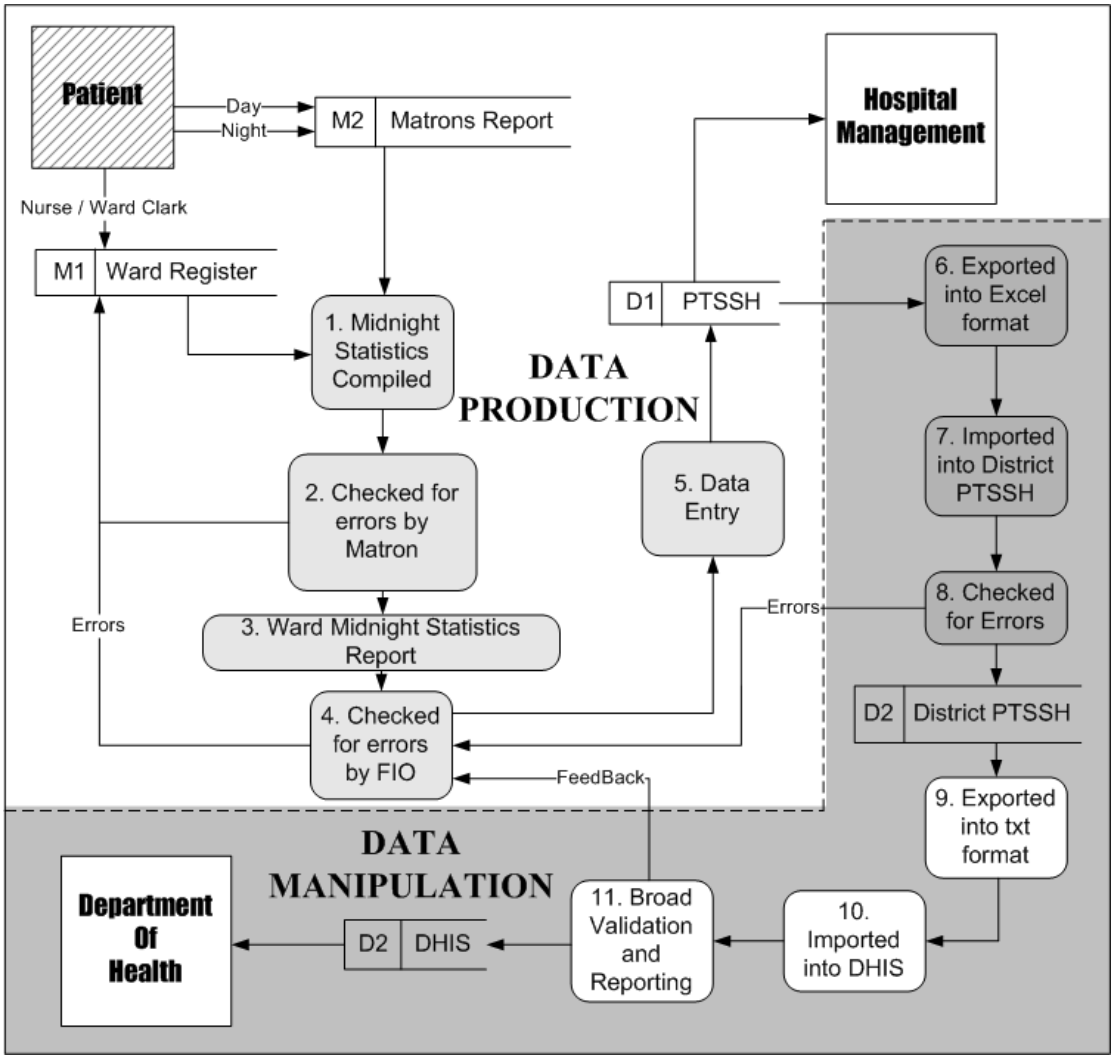
Errors can be generated within the hospitals and outside the hospitals. Figure 1 illustrates a generalized overview of how inpatients' data are captured at the hospital level and subsequently make their way through the system to the Head Office. The hospital level in Figure 1 begins with the patient and ends with the data entry into the PTSSH. It includes the data generation in each ward by the nurse, the editing of the information by the Facility Information Officer (FIO) and the data entry into the PTSSH. The data flow between the hospital and the Head Office includes export of files from PTSSH into Excel and again into PTSSH at the district level. This is followed by the transformation of the data into another format so that the file can be imported into the District Health Information System (DHIS). For clarity sake, the hospital information system has been divided into data production within the hospital and data flow and transformation outside the hospital.

Data production

To fully understand how errors can be generated at the hospital level it is necessary to understand how the system works. Data on admitted patients are entered on the ward register (M1 in the diagram). This is the first source of potential error because the method of recording is not standardized and has resulted in an assortment of techniques. For example one hospital may have only one register with both admissions and discharges recorded. Another hospital may have a system in which admissions are recorded on one register and discharges on a second register. In most wards the number of occupied beds is temporarily recorded on a white board, which gets overwritten every day. In addition to the ward register, a Matron's report (M2) is often maintained. At 7 pm and 7 am, the matron fills the data related to the previous

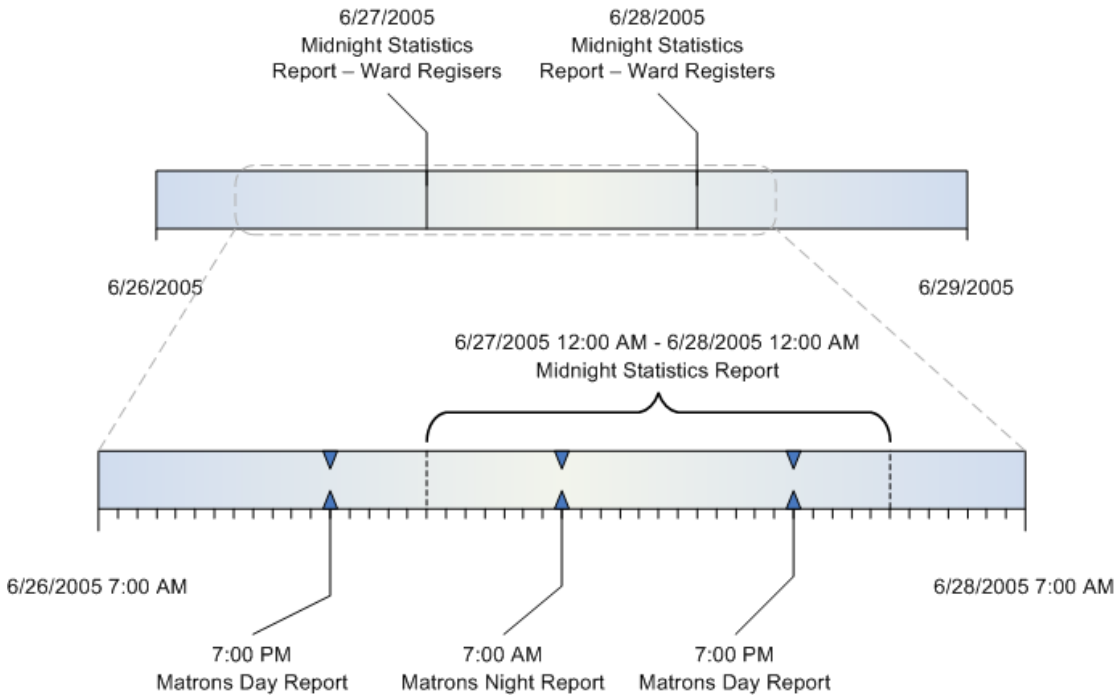
12 hours. Although there appears not to be a uniform implementation of the Matron's report, information about critically ill patients, admission, discharge, transfers and other variables may be recorded.

Figure 1 Data production and data flow



At midnight, in each ward, the nurse on duty fills the midnight statistics form related to the previous 24 hours. The nurse extracts the information from the ward and Matron`s reports and fills the midnight statistics form that will be sent to the FIO the following morning. Because the midnight statistics form is related to the previous 24 hours, the major limitation in using the Matron's report is related to its different time frame. While the first one is filled out at midnight and covers the previous 24 hours, the second one is completed at the end of the day shift at around 7 pm and night shift at about 7 am and covers the previous 12 hours. This mismatch between the time covered by the two complementary data sources creates obvious limitations to the consistent extraction of the same information as it can be seen in Figure 2.

Figure 2 Timeline contrasting Matrons report and Midnight statistics form



At midnight a nurse uses the ward register, along with the Matron's report, to fill out the midnight statistics form (as indicated by box 1 in Fig. 1). The nurse extracts the data from the ward register and Matron`s report and perform a manual head count of the patients and beds in the ward, and records the data on the midnight statistics form. Once the nurse in each ward has completed the midnight statistics form for the previous 24 hours, the night matron usually performs a check to verify that the data was entered correctly. If no errors are found the data is then forwarded on to the Facility Information Officer (FIO) the following morning.

The FIO is the corner stone of the entire information system. Each hospital has one FIO who enters the data submitted from the different sections of the hospital into the Patient Throughput Statistical System (PTSSH). His / her responsibility is to ensure that the hospital is collecting and providing accurate information by correcting the midnight statistics from all the wards (as indicated by box 4 in Fig. 1).

If errors are found during this process, most FIOs apply corrections without sending back the midnight statistics forms (FIO1) to the wards. Only a few FIOs send the forms back because the night staff that filled the forms leave at the beginning of the day and they are not available to correct the data. Once the data have been checked, either the FIO or a data capturer enters the data into PTSSH (D1 in Fig. 1). While the data from the wards are captured on a daily basis, the data on the number of operations and outpatients are captured monthly.

Data flow and transformation

By the 7th of each month, the FIO is required to send the data related to the previous month to the district office. In order to accomplish this, the data is exported from the PTSSH, which is a Microsoft access database, into a Microsoft Excel file, which is then imported into the district office PTSSH programme (box 6 in Fig. 1). Error checking is performed on the data and any anomalies reported back to the FIO for correction.

The data is then forwarded to the DOH Head Office. Since Head Office currently uses the District Health Information System (DHIS) to store and manipulate data whereas the hospitals and district office use PTSSH, a conversion of the data is once again needed. To accomplish this, data from PTSSH is exported into a comma delimited text file (box 9 in Fig. 1), which can then be imported into the DHIS (box 10 in Fig. 1). The data is broadly checked for errors, reports are written and feedback sent to the FIO (box 11 in Fig. 1) before finally being stored in the DHIS (box D2 in Fig. 1). This process will be much simplified when the hospitals move away from PTSSH in favour of the DHIS in 2006.

It should be noted that even though there are many checks and balances, the system is complicated with many opportunities for error. A one-off error is no large concern for worry such as a data capturer incorrectly entering the number of patients during a 24-hour period. What is of greater concern is if many errors enter the system; particularly if such errors occur earlier on in the data process.

Depending on whether the same error is made repeatedly or not, reliability and/or validity of the indicators may be compromised. Repeatedly making the same error would result in a reliably inaccurate (invalid) indicator while random errors would affect the reliability and possibly the validity of the indicator. For example if ten more beds than are actually in the ward are reported each day, the number of beds will remain constant even though it is incorrect. In other words the measurement is reliable (always returns the same value) but lacks validity. If, on the other hand, the number of beds reported each day varies depending on the nurse counting them, the measure will be unreliable.

Objective of the Validation

The primary aim of this validation was to collect a sample of valid and reliable data from each public hospital in KwaZulu-Natal and to compare them with the reported data. For this reason, the data collection was carried out with a standardized methodology in two different periods.

Methodology

The survey was designed specifically to address the issue of data validity and reliability. To achieve this aim, a standardized methodology was developed and consistently applied to each public hospital of KwaZulu-Natal in two rounds. A total of 67 hospitals were validated, of which 2 were central, 13 were regional, 40 were

district and 12 were specialized. To measure the reliability of the reporting between two different points in time, each hospital was visited twice. The first round lasted from the 20th October 2004 to the 24th March 2005, while the second round was conducted from 25th March to the 27th August 2005**.** Within each hospital, all the wards were visited and the standard methodology was applied to arrive to the most valid numbers related to the previous 48 hours, which is short enough to allow limited recall bias. The variables which were validated included admissions, discharges, intra and inter-hospital transfers, deliveries, live births, stillbirths, deaths, beds by type, occupied beds, day patients and pass out. Instead, the numbers of OPD headcounts and operations were validated for the month preceding each round. The source of information included interviews with the FIOs, the nurses and the matrons, the consultation of the ward and theatre registers and Matron`s report and the comparison with secondary sources such as the mortuary for the number of deaths. The validation methodology and the data collection forms are described fully in the Annex.

Such data included all major sources of variation to identify and track the major sources of error contributing to the unreliable and invalid data emerging from the hospital information system. Of particular interest was the stability of the number of useable beds within each hospital, as this figure is used to calculate many of the efficiency indicators its accuracy is critical. It was hypothesized that the total number of useable beds should remain roughly constant at least throughout a financial year (FY). To test this hypothesis, the number of useable beds as well as the other variables was manually counted on two separate occasions with between three and six months passing between countings. Collecting such a set of accurate data allowed to highlight discrepancies between reported and actual occupancy and turnover, and to identify practical solutions to improve the reliability of the data.

The analysis was based on the followings:

(a) The valid data collected on individual variables (i.e. number of beds) in each ward were compared with the data reported by the nurse on the midnight statistic form, with the correction applied by the FIO, and with the data entered into PTSSH;
(b) The indicators (i.e. occupancy) based on the valid data collected by the survey were compared with the indicators based on the data reported by the FIO;
(c) The consistency of the reporting was checked by comparing the valid changes of the indicators between the two rounds with the changes based on the data reported by the FIO;
(d) To check the variation introduced by the file transformation outside the hospitals, the valid data of the numbers of beds, OPD and operations were also compared with the monthly data produced by the DHIS.

Results

The results on the reliability of the reporting system are divided into the following sections:

- The first section deals with the validity and reliability of the individual variable;
- The second section shows the effect that the reporting of individual variable has on the validity and reliability of the indicators of utilization;
- The third section deals with the consistency of the indicators between the two rounds;
- The fourth section compares the DHIS reported values with the valid ones; and
- The last section aims to facilitate the understanding of some of the problems faced by the FIOs and the nurses.

Individual variables

The first measure to judge the reliability of reporting was the correlation between the numbers collected by the validation and the numbers reported by the various sources. The values measured by the survey were compared with the values reported by the nurses in the midnight statistics form, the numbers corrected by the FIO and the data entered into the PTSS. Figures 3-4 provide an example of how the above comparison was done for the numbers related to the adult and juvenile beds. In Figure 3, each hospital is plotted according to the number of valid adult beds on the Y-axis and the number reported by the FIO on the X-axis. For example in Figure 3, Greys is plotted according to 423 adult beds counted by the survey (Y-axis) and 422 by the FIO (X-axis). Because the valid value was almost the same as the reported value, Greys is almost on the 45-degree line. The few hospitals (i.e. Madadeni) that reported slightly more adult beds (over-reporting) than the valid ones are on the right of the line. Few others (i.e. Khan) that reported a lower number (under-reporting) compared with the valid one are on the left of the line. If there had been no difference between the valid and reported values, all the hospitals would have laid on the line and correlation would have been 1. As shown at the top of Figure 3, the correlation was 0.99 indicating that the reporting of the adult beds was almost identical to the valid values.

Figure 4 shows how the situation changed for the reporting of the juvenile beds. Compared with Figure 3, the dots are more scattered because the number of juvenile beds reported by the FIO were different from the numbers reported by the validation. As a consequence, the correlation declined to 0.77. This suggests that the reporting of the juvenile beds is less reliable than the reporting of the adult beds.

Figure 3 Adult beds according to the validation (Y axis) and the FIO (X axis)

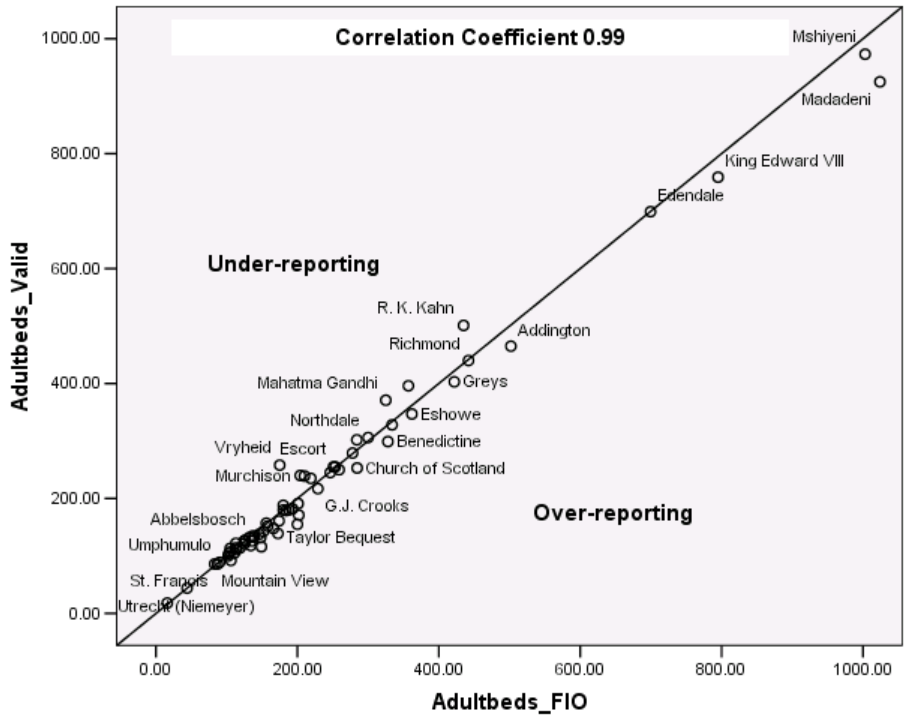
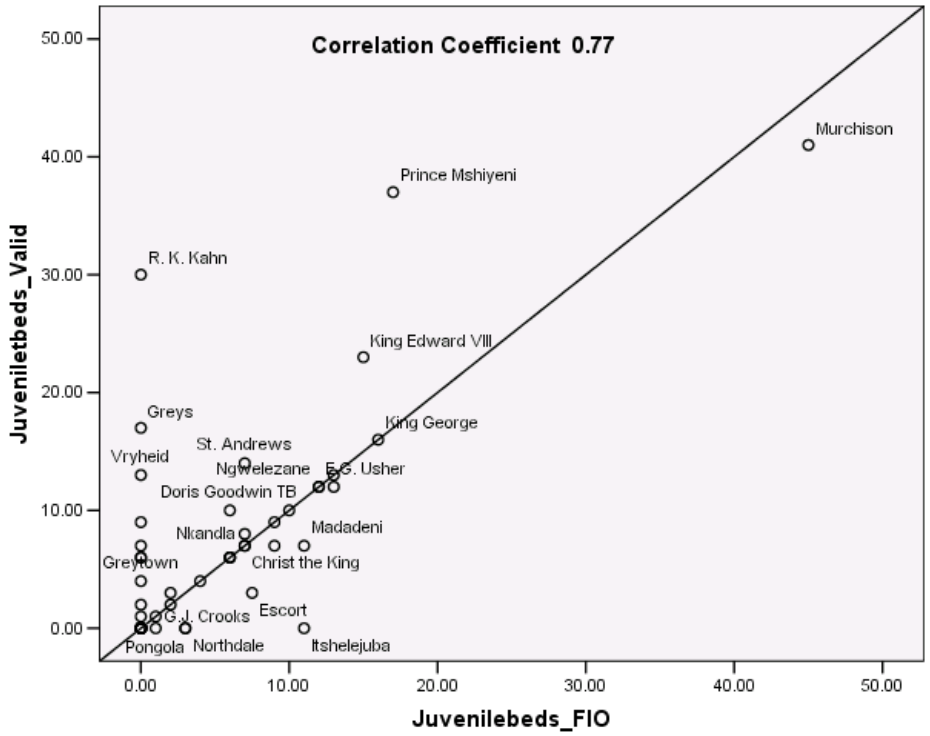


Figure 4 Juvenile beds according to the validation (Y axis) and the FIO (X axis)



Correlation Coefficients

The information provided in each of the above figures is summarized by the correlation coefficients of Table 1. For each variable in a given row, the correlation coefficients between the numbers validated by the survey and the numbers filled by the nurse, the FIO and by the data entry into the PTSS are provided in each column. For example, the correlations related to the adult and juvenile beds in Figures 3-4 can be found by intersecting the rows related to the adult and juvenile beds with the column under the FIO (0.99, 0.77). In this way it is possible to compare the degree of reliability associated with different data sources without resorting to too many graphs. Variables with correlations approaching 1 in every column such as in the case of adult beds indicate high validity for the nurse and FIO reports, because they approach the real values. However, because the PTSS correlation for the adult beds declined to 0.94 it is affected by declining reliability after the data are entered into the PTSS. Therefore, high correlation across all data sources indicates high validity and reliability, but a change in correlation indicates low reliability because the reporting is inconsistent across the nurse, the FIO and the PTSS. The correlation was higher for the number of admissions, discharges, deliveries, live births, deaths and adult beds, but even for these variables there was a slight decline in correlation after the data were entered into PTSS. Correlation was lower for transfers, stillbirths, juvenile beds, cribs, cots, incubators, day patients and pass outs.

The message from Table 1 is that the reliability of the reporting of the number of useable beds varied according to the type of bed. The adult beds are more reliably reported compared with juvenile beds, cots, cribs and incubators because it is more difficult to standardize their reporting. This is due to a higher variation in the interpretation of what types of beds should be included in the computation of the useable beds.

Table 2 shows the total number of beds by type for all the hospitals according to the different sources. The ward register has the lowest reliability for all types of beds, the nurses and the FIO slightly over-reported the number of adult beds, correctly reported the number of cots, under-reported the other types of beds. The fact that the reporting of the total number of juvenile beds, cribs and incubators are lower than the real values may have implication for the planning of the needs of paediatric wards. Another consideration is that there is an under-estimation of the border beds, which although according to guideline should be considered useable beds, are not taken into account by the nurses and the FIOs. Overall, there is an underestimation of more than 1000 beds according to what was reported by the nurses and slightly less than 600 according to the FIO compared with the 20448 total beds counted by the survey. Although this might not sound a big difference, it should be considered that this variation is based on a short period of 48 hours preceding the survey. If a longer period were considered, such as one month, the difference would add up to a substantially greater number.

Table 1 Correlation coefficients between the validation and the data sources

Variable	Validation versus Nurse	Validation versus FIO	Validation versus PTSS	Comments
Admissions	0.98	0.98	0.94	The correlation declines after the data entry, suggesting reliability problems.
Discharges	0.99	0.99	0.93	Same as above
Intra-hospital* Transfers-in	0.95	0.95	0.89	These are the transfers between wards within the same hospital. Validity is lower than with admissions and discharges and reliability declines after the data entry.
Intra-hospital* Transfers-out	0.94	0.94	0.87	Same as above
Transfers-out**	0.81	0.84	0.60	Low validity and low reliability
Transfers-in**	0.91	0.98	0.95	Low reliability
Deliveries	0.98	0.98	0.94	Reliability declines after data entry
Live births	0.98	0.98	0.93	Same as above
Stillbirths	0.64	0.84	0.97	Low reliability
Deaths	0.98	0.97	0.96	Declining reliability between the ward and the data entry
Adult beds	0.99	0.99	0.94	Same as above
Juvenile beds	0.75	0.75	0.59	Poor validity & reliability
Cribs	0.86	0.87	0.72	Poor validity & reliability
Cots	0.95	0.94	0.79	Poor validity & reliability
Incubators	0.65	0.65	0.86	Poor validity & reliability
Day patients	0.75	0.75	0.57	Poor validity & reliability
Pass out	0.95	0.96	0.90	Low reliability

***** Transfers between wards, ** transfers from and to other hospitals

Table 2 Variation in the counting of useable beds (First Round) *

Type of useable bed	Ward Register	Nurse report	FIO corrections	Validation
Adult*	10072	15346	15762	15522
Juvenile	209	272	268	371
Cot	962	2448	2490	2441
Crib	352	1007	1036	1293
Incubator	140	265	272	418
Border	86*	29*	32*	403
Total	11821	19367	19860	20448

* Most border beds were not considered by the FIOs as useable beds

Another consideration to be taken into account is that the numbers in Table 2 represent the total for all the hospitals, which hide a higher inter hospital variation. Some hospitals contribute a higher variation than others, leading to substantial comparability problems between hospitals and within the same hospital in different time periods, increasing the difficulty in evaluating the changes in efficiency. Last but not least, the results in Table 2 take into consideration the reliability of the reporting of the total beds within the hospital reporting system only. As already

mentioned, other sources of variation are added after the files are sent to the higher level.

Indicators of Utilization

This section describes how the reliability of individual variables influences the reliability of occupancy and turnover. As shown by the formulae, the occupied beds, discharges and deaths are the main variables influencing occupancy and turnover, while day patients and transfers are too small to have a relevant influence. Average length of stay (ALOS) is not taken into account in this analysis because ALOS becomes redundant when occupancy and turnover are already considered. In fact, a reliable occupancy and turnover is automatically associated with a reliable ALOS and vice versa.

Occupancy = (1/2 Day Patients + inpatient days + Transfers in)/useable beds

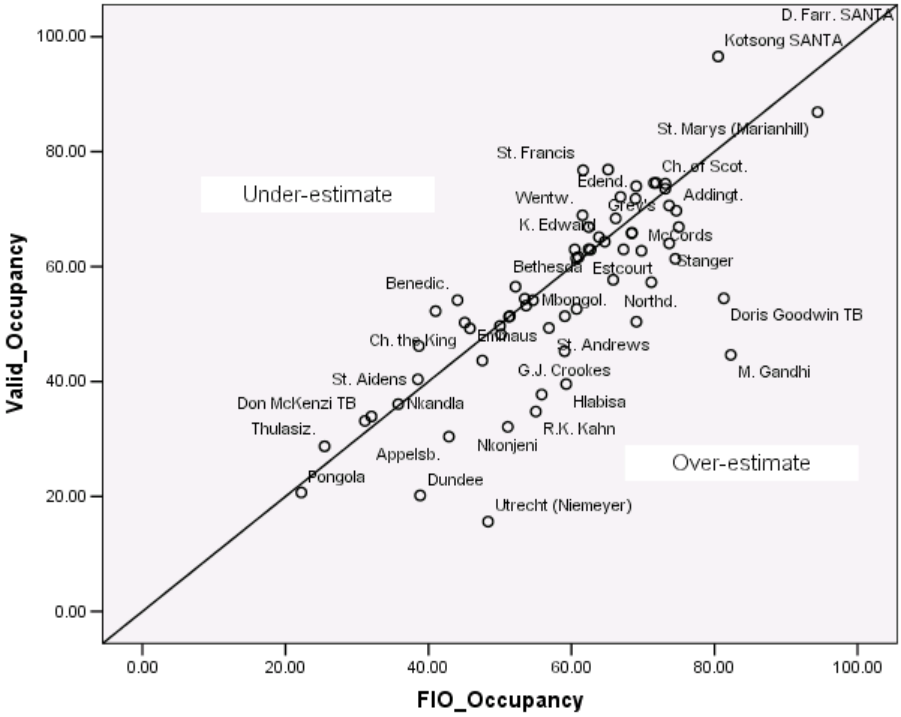
Turnover = (Discharges + Deaths + Transfers out)/Useable beds

Occupancy

Figures 5-7 show how the reporting of individual variables influence occupancy. The hospitals in Figure 5 are plotted according to the occupancy based on the data reported by the validation on the Y-axis and the occupancy based on the data reported by the FIO on the X-axis. For example, R.K. Khan is plotted according to its real occupancy of 35% (Y) and its occupancy based on the data reported by the FIO, which was 55% (X axis). Because Khan over-estimated its occupancy its position is at the right of the 45-degree line.

Figure 6 and 7 provide instead a visual representation of over and under-reporting of beds and inpatient days. For example, in Figure 6, Khan is plotted according to the 627 beds counted by the validation (Y) and the 523 reported by the FIO (X). Similarly, in Figure 7, Khan is plotted according to the 218 (Y) inpatient days counted by the validation and the 288 (X) reported by the FIO. Because the FIO under-reported the number of beds and over-reported the number of inpatient days, Khan is respectively at the left and at the right of the 45-degree line in Figure 6 and 7. Fewer hospitals like Benedictine are at the left of the line in Figure 5 (underestimated occupancy), at the right of the line in Figure 6 (over-reported beds) and at the left of the line in Figure 7 (under-reported inpatient days). Therefore, looking at the three Figures, it is possible to visualize the relationship between under and over estimation of occupancy and incorrect reporting of the numbers of beds and inpatient days. It can be noted that most hospitals were at the right of the line, indicating that most of the data from the FIO produced a higher occupancy compared with the data from the survey. This was due to under-reporting of beds and over-reporting of inpatient days.

Figure 5 Valid occupancy (Y axis) and reported occupancy (X axis)



Figures 6-7 Valid versus FIO number of beds and inpatient days

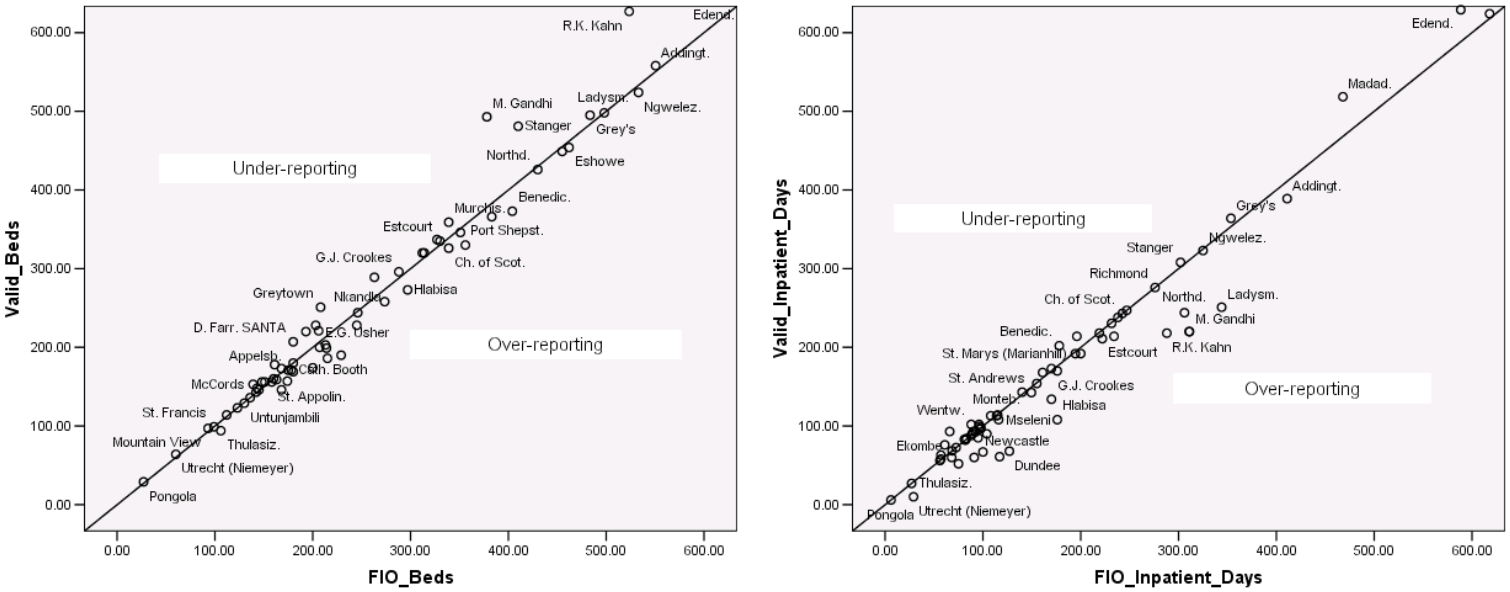
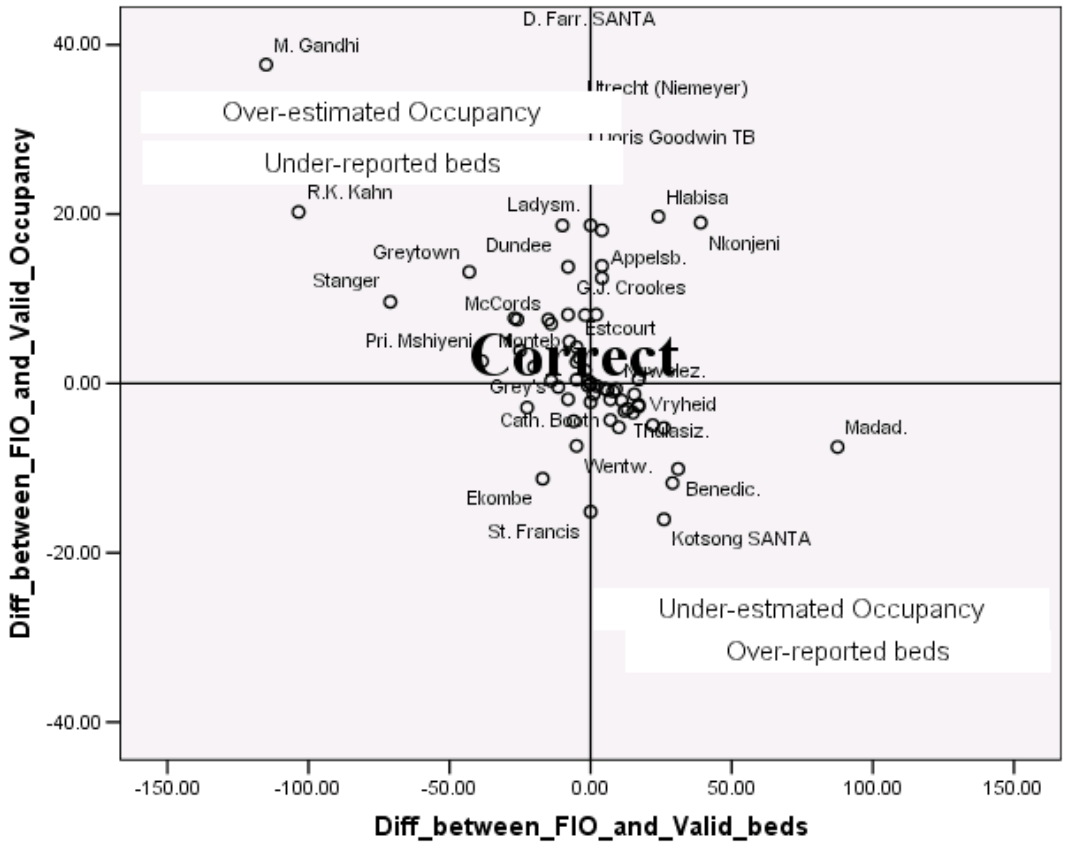


Figure 8 is a more efficient way of representing the relationship shown in Figures 5-6. In Figure 8 the hospitals are plotted according to the net difference between reported and valid occupancy on the Y-axis and the net difference between reported and real beds on the X-axis. For example, Khan is plotted in Figure 8 according to +20% occupancy on the Y-axis, which is the difference between reported and valid occupancy (55%-35%), and –103, which is the difference between the reported and valid number of beds (524-627). The net difference between the FIO and the valid

occupancy varied between +66% (over-estimation) and –16% (under-estimation). The net difference between reported and valid number of beds varied between +87 (over-reporting) and – 115 (under-reporting).

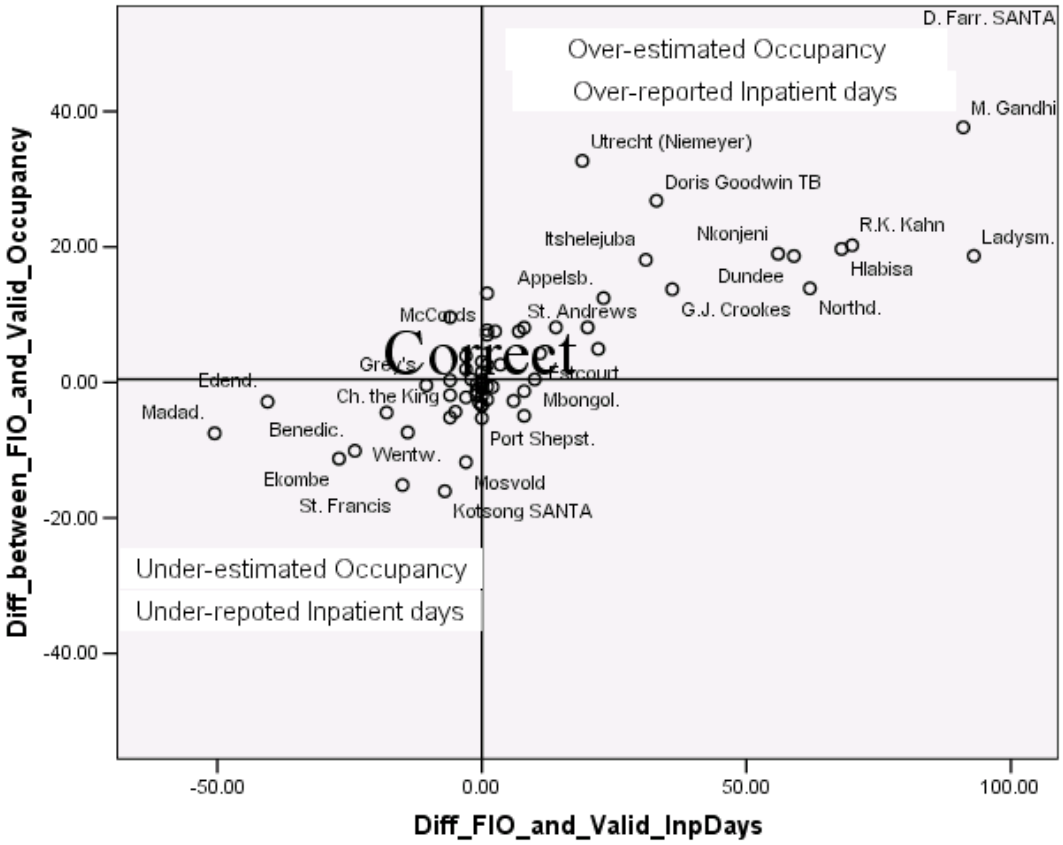
Figure 8 shows the negative linear relationship between over-reporting of the number of beds and underestimation of occupancy and vice versa. The hospitals in Figure 8 move on a linear declining pattern, with the hospitals on the upper right (i.e. Khan) being characterized by positive values on the Y axis (overestimated occupancy) and negative values on the X axis (under-reported beds). As the hospitals proceed towards the lower section of the graph, they approach 0 values for X and Y because there were no differences between the number of beds reported by the FIO and the validation, and between the reported and valid occupancy. Proceeding towards the extreme lower right, the Y-values become negative because the FIO occupancy was lower than the valid one (under-estimation) and the X-values become positive because the beds reported by the FIO were higher than the real one (over-reporting). Therefore, Figure 8 allows to visualize the same relation depicted in Figure 5 and 6, where over-estimation in occupancy is related to under-reporting of beds and vice versa.

Figure 8 Differences between FIO and valid occupancy (Y), and FIO and valid beds (X)



Similarly, Figure 9 represents the relationship between Figures 5 and 7. As it was done for Figure 8, the hospitals are plotted in Figure 9 according to the net difference between FIO inpatient days and valid ones (X-axis) and the difference between FIO occupancy and valid one (Y-axis). In this case, the linear relationship is positive because positive Y-values (overestimation of occupancy) are associated with positive X-values (over-reporting of the number of inpatient days).

Figure 9 Differences between FIO and valid occupancy (Y), and FIO and valid inpatient days (X)



Turnover

Figures 10-11 show the relationship between the reliability of reporting of individual variables and the correct estimation of turnover. As in previous figures, the Y-axis represents the net difference between the turnover based on the data from the FIO and from the survey. In Figure 10, the X-axis represents the net difference between the number of beds reported by the FIO and the validation, and the Y-axis represents the net difference between the reported and the valid turnover. In Figure 11, the net difference between the discharges plus deaths reported by the FIO and by the validation is on the X-axis, while the difference between reported and valid turnover is on the Y-axis. The relationships are similar to what already seen in Figures 8-9, with over-reporting of beds leading to under-estimation of turnover and over-reporting of discharges plus deaths leading to over-estimation of turnover.

Figure 10 Differences between FIO & valid values for turnover (Y) and beds (X)

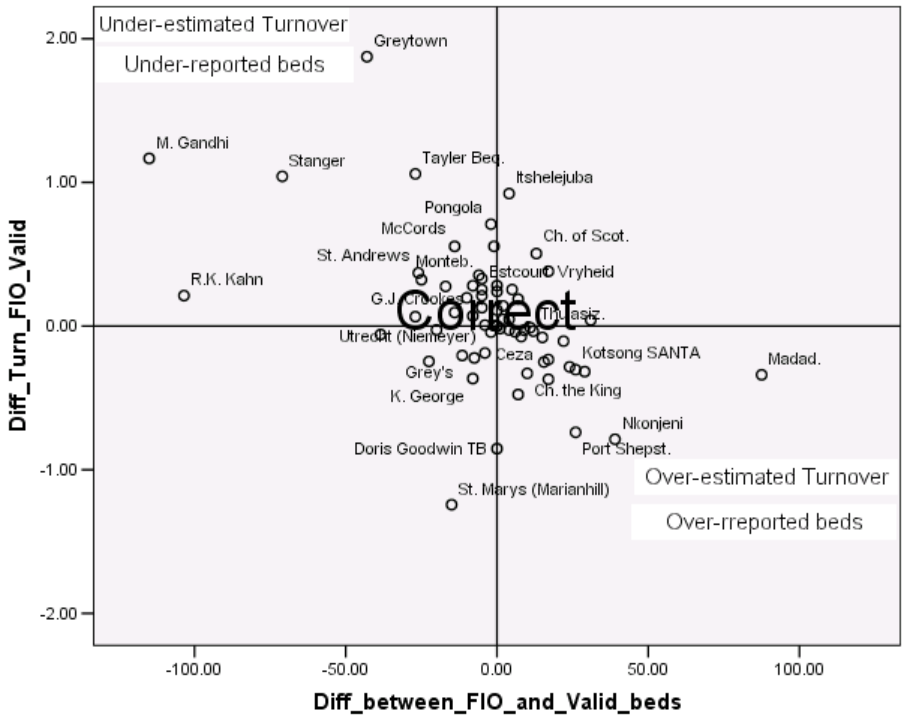
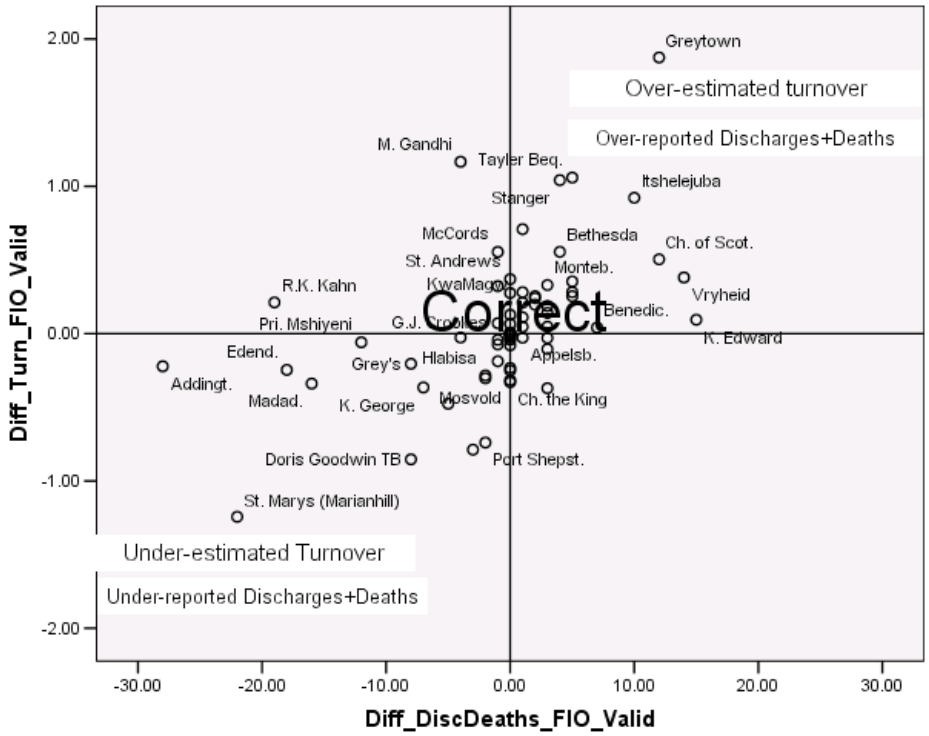


Figure 11 Difference between FIO & valid values for turnover (Y) and discharges + deaths (X)



To put the above figures in perspective, most hospitals produce wrong estimates for occupancy and turnover because of the incorrect reporting of the number of beds, inpatient days, discharges and deaths. The above graphs show a clear declining linear relationship between under-reporting of the number of beds and over-estimation of both indicators and vice versa. The under-reporting of the inpatient days, discharges and deaths leads to underestimation of the indicators of utilization and vice versa.

Comparison between the first and second round

The second round confirmed the same reliability findings of the first round. The correlation coefficients between the reported and valid numbers were similar to what is shown in Table 2. There was a higher correlation for admissions, discharges, deliveries and adult beds; and a lower correlation for transfers, day patients, juvenile beds, cribs, cots and incubators. Also in the second round the under-reporting of beds and the over-reporting of inpatient days and discharges caused a higher frequency of over/under-estimation of occupancy and turnover.

Efficiency

The changes in efficiency between the two rounds provided the basis for the reliability of the reporting system. Reliability was measured according to the degree of consistency between the changes in occupancy and turnover reported by the system and the changes measured by the validation team. Annex I shows that reliability varied across hospitals with some hospitals reporting more consistently than others. The changes shown in Annex I should not be used to judge efficiency of a certain hospital but to judge its reliability in reporting.

Mortality and stillbirth rates

Figure 12-13 and 14-15 show the reliability of the reporting on the mortality and the stillbirth rates. Except for specialized hospitals, the other categories reported almost the same rates as the validation with a slightly lower mortality for the FIO data. The stillbirth rates were higher for the validation in R1 while FIO and valid stillbirth rates were similar in R2. The rates could not be compared for individual hospital due to the very low number of deaths related to the short period considered, but it is likely that the variation will be much higher if individual hospitals were compared.

Figure 12-13 Mortality reported by hospital categories in R1 and R2

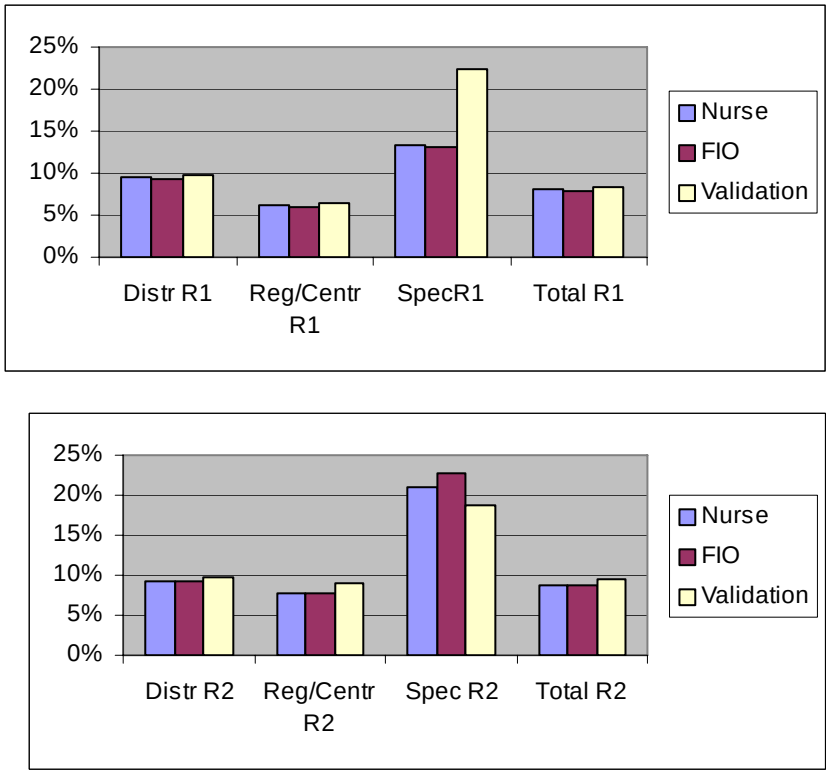
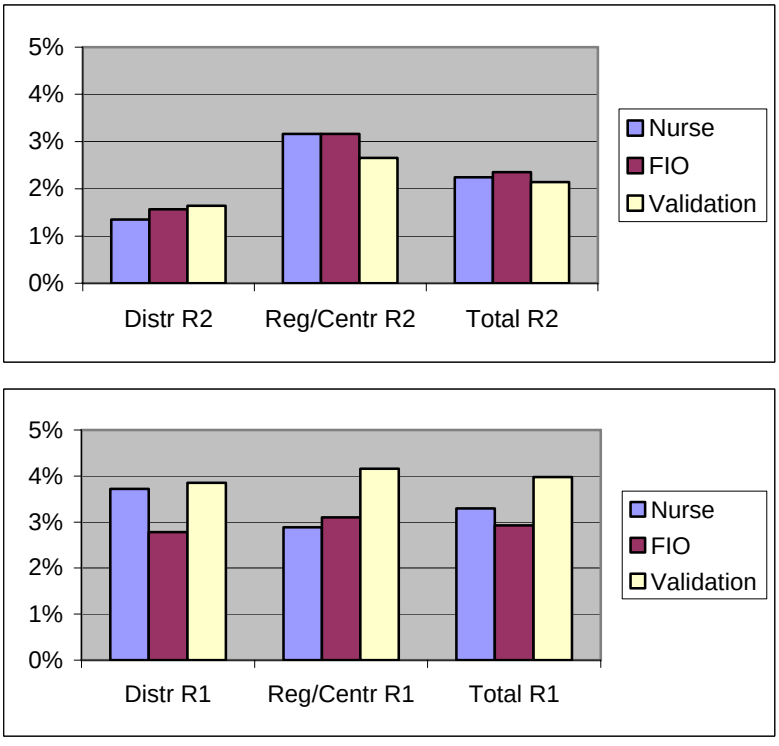


Figure 14-15 Still birth rate by hospital category in R1 and R2



Data received by the DOH

This section of the analysis deals with the data sent to the central level. The results presented up to this point are related to the variation within the hospital system only. As mentioned in the introduction, the complex data flow from the hospitals to the district and the central level introduces another source of variation. Because the data received by the central level can only be extracted in monthly aggregates, it was only possible to compare the number of beds, OPD and operations for the month, which was validated with the respective values reported by the central level.

As expected, the consistency between the valid and the DHIS data was lower than the consistency between valid and FIO data. Table 3 shows the correlation coefficients between the valid numbers versus the number reported by the FIO and between the valid numbers and the DHIS numbers. As usual, correlations approaching 1 suggest that the numbers related to the FIO or to the DHIS for each hospital were very near to the numbers produced by the validation, as already seen in Table 1. It can be noted that the correlation between the valid number of beds and those reported by the FIO was much higher (0.99) than the correlation between the valid number of beds and the DHIS number of beds (0.92). This decline in consistency was even higher for the OPD, which had a correlation of 0.97 for the FIO and 0.68 for the DHIS. The number of reported operations was characterized by a very low correlation with the valid values suggesting that both the FIO and the DHIS were highly unreliable in reporting the number of operations. These results suggest that extra variation is introduced by the data processing and manipulation outside the hospital and that the data on OPD and operations are so unreliable that they are unlikely to be useable to compare surgery rates or OPD workload within the same hospital or across hospitals. The unreliability of the OPD headcount might also affect the correct estimation of the equivalent patient days.

Table 3 Correlation Coefficients between the valid numbers and those produced by the FIO and DHIS

Variable	**Validation versus FIO**	**Validation versus DHIS**	**Comments**
Beds	0.99	0.92	Correlation declines for DHIS
OPD	0.97	0.67	Strong decline in correlation
Operations	0.57	0.54	Both sources were unreliable

Table 4 confirms that the number of beds was under-reported in both rounds and the under-reporting increased further when the data were integrated into the DHIS. While the FIO under-reported about 600 beds compared with the validation, the DHIS under-reported about 1000 beds. One positive note is that at least there was relative consistency in the numbers reported in R1 and R2, confirming the hypothesis that the number of beds did not change between different periods of time.

Table 4 Number of Beds between R1 and R2

	R 1	R 2
Validation	20488	20432
FIO	19861	19796
DHIS	19402	19491

Tables 3 and 4 confirm that extra variation is added between the FIO and the DHIS. The results show that when the final numbers are published by the DOH, they are different from the original one sent by the FIO because of the extra corrections at the district and central level and because of file transformations. This extra variation, which increases further the reliability problems, could be avoided if the FIO could also directly send the data to the central level where one person should be assigned to apply standardized techniques, like that introduced in Issue 9, to identify outliers. The fact that only one person would be involved would ensure that any statistic application to identify outliers would be done consistently. Once the outliers are timely identified, the DIO and FIO should be alerted about what data should be re-checked and corrected.

FIO

The FIOs were interviewed to capture some of the characteristics that may influence reporting. Sixty one FIOs were identified and interviewed to describe their job conditions, training and supervision, and problems that are represented in Figures 16 through 18.

Job conditions

Most FIOs started working in 2002 and are now well-established professional figures in the hospital system. About one fourth started working before 2000, another quarter started between the beginning of 2000 and the middle of 2002 and the rest after June 2002. Half of the FIOs were appointed and half were acting, but most of them were working full time as FIO. Slightly more than half had a copy of the job description and had an instruction manual on how to report the data.

Training and supervision

Most FIOs have been trained and supervised but they require more support. About three quarters received training, which usually lasted 1 day and included use of PTSSH, filling of forms and validation. Most of the supervision of the FIO within the hospital was done by the personnel dealing with accounting, only 20% mentioned the hospital manager as supervisor and 9% mentioned that they did not have a supervisor. Most of the supervisors coming from outside the hospitals were the District Information Officer (DIO). The feedback received during the last supervision was related to solving hardware and software problems and correcting the data, but 14% mentioned that they did not receive any feedback.

Some questions were used to crosscheck the clarity of understanding of key variables used to estimate utilization. About 60% of FIOs excluded border beds from the computation of the useable beds and 13% reported the authorized beds instead of the

useable ones. As far as the reporting of occupied beds, 7% included border mothers accompanying their children in paediatric wards and there was uncertainty if patients on floor beds were to be included or excluded from the count of occupied beds. Almost every FIO counted each transfer-in as "1 transfer-in + 1 occupied bed", which defeat the purpose of putting transfer-in and inpatient days at the numerator of occupancy because it leads to double counting. A minority (8%) of FIOs counted transfers-in as admissions. Transfers-out and day patients were categorized as discharges by 15%. Most FIOs categorized correctly the admissions and the discharges although slightly more than 10% of them included transfers-out, deliveries, births and day patients among the discharges. In terms of correctly counting the number of deaths, about 18% of the FIOs included stillbirths and most were uncertain if a death occurring in OPD had to be included. Less than half were crosschecking the mortuary registers and of those who did, one third found a mismatch between the mortuary and the number of deaths reported by the wards. Most counted OPD as headcount but 18% reported OPD visits.

Problems

Several problems were mentioned in the area of software and data reporting. About half of the FIOs had problems with the PTSS, including inconsistency in the outputs produced compared with the data entered by the FIOs and exportation of files. About 40% of the FIOs mentioned problems related to data reported from the ward, including late submission, incomplete on incorrect data, and lack of interest by the nurse. About three out of every four FIOs mentioned inadequate training of themselves but especially of the nurses and loss of trained staff. One fourth of the FIOs mentioned supervision as a problem including lack of interest and lack of background in statistics by the supervisor and inadequate support from him/her. About 40% mentioned that they were over-worked with backlog piling up because of insufficient time for the assigned tasks and lack of back-up support. Less than one third had a copy of the latest annual Statistic Bulletin, and although 80% had access to the intranet, only 27% have accessed the Statistics Bulletin through the Intranet and only 15% compared their figures with the data published in the Bulletin. The reasons mentioned for this low frequency of crosschecking against the Statistics Bulletin included no clear reason, lack of awareness of the existence of the Bulletin and lack of time.

Data check

One third of the FIOs checked the data less than one month before the survey, one third did it more than a month before and one third never did a data check. The major problems found during the last data check included inaccuracy of registers, miscounting of inpatients, beds and OPD headcounts. About 85% of the FIOs applied corrections on the data they received from the wards and this was based on balancing the patients present in the wards with the patients entering and leaving during the previous 24 hours. Some of the techniques employed by the FIOs to validate and correct the data included: (a) Checking that the number of patients entering and exiting the ward during the 24-hour period were matching that of the previous day; (b) Checking for the mistake of counting an intra-hospital ward transfer as a transfer coming from another hospital; (c) Altering the number of beds submitted by the ward

to a constant predefined number of beds; and (d) Checking for simple numerical mistakes made on the form.

Figure 16 Proportion of FIOs with certain characteristics

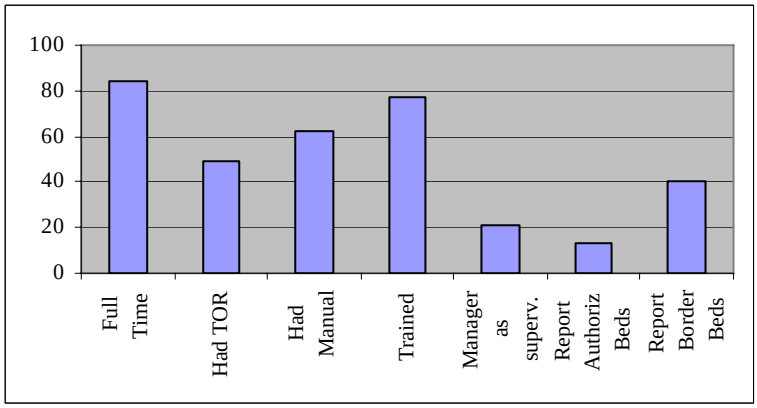


Figure 17 Percent of FIO mentioning types of problems

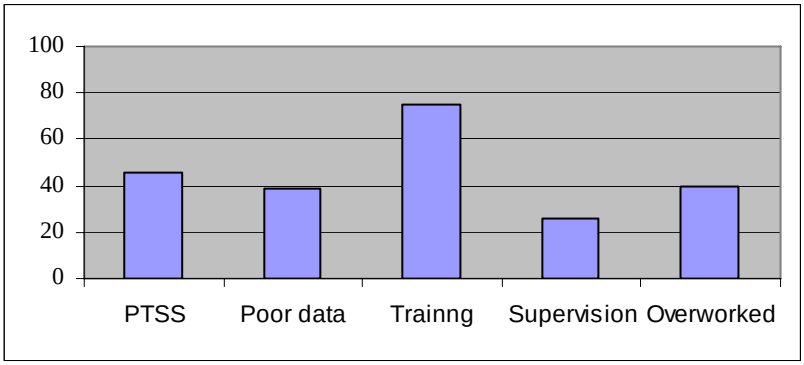
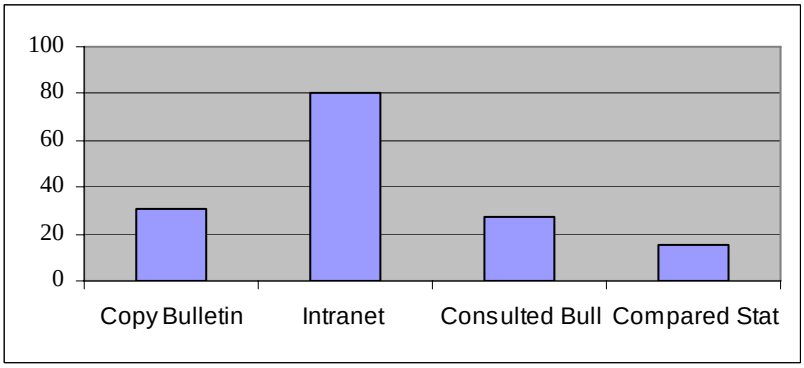


Figure 18 Proportion of FIOs with access to Statistics Bulletin



Nurse

Sixty one representative matrons were interviewed to check their knowledge and practice in data reporting. In district and specialized hospitals there is usually one matron while the regional and central hospitals have more than one matron. The Matron's register is one of the sources of information used by the nurse to fill the midnight statistics report. Therefore, it was important to check that the Matron has a clear understanding on what should be reported. Less than half of the matrons had an instruction manual on the criteria of the data reporting. About half and one third of the matrons confused transfers-in and out respectively with admissions and discharges and more than half categorized day patients as admissions and discharges. A minority included deaths and births among the discharges and border mothers among the inpatients.

Discussion

This validation has shed light on the problems affecting the hospital indicators. Although the problems affecting the hospital information system in KZN has been described in previous Issues, this is the first thorough analysis, which has tried to identify the causes and the potential solutions. The causes can be divided into data production and transformations, and the solutions need to be specific to these two different levels.

Data production

The first source of variation is located within the hospital. This is due to the fact that when many nurses across several wards are involved, there is inconsistency in reporting because of different interpretation of what should be included and excluded from the data. One positive note is that for most variables, the data collected by the nurses in the wards and gathered by the FIOs did not substantially differ from the data collected by the validation. This was reassuring, because having good quality data is considerably difficult when the counting is repeated daily across several wards in each hospital and across several hospitals in the province. Even the changes of the FIO indicators between the two rounds were pretty consistent with the validation changes for more than half of the hospitals, indicating that the consistency between reported and valid values persisted in most hospitals between different time periods.

The system can improve considerably if the counting of the number of beds is standardized. Reporting the number of useable beds might sound straightforward, but nurses and FIOs have their own subjective interpretation of whether to include or exclude border beds, beds without mattress, broken beds, floor beds, delivery beds and beds located in temporarily closed wards. By the same token, border mothers occupying a border bed may be included or excluded from the computation of the number of occupied beds. Considering that the beds are counted daily (monthly by some hospitals), small inconsistencies can add up to a substantial difference at the end of the month and even more if different years are compared.

The most problematic variables were the transfers. The common situation is that one patient who is a transfer-in is counted as '1 transfer-in + 1 occupied bed' creating duplication, because the numerator of the occupancy formula includes "1/2 Day patients + Inpatient Days + Transfer-in". Sometimes there will be triple counting because a certain proportion of nurses and FIOs count transfers-in as admissions. Transfers out and day patients were frequently counted as discharges.

Besides creating double counting, the data collection on transfers in not used to achieve the real objective of collecting this information. The objective of collecting transfers and day patients is not to estimate more precisely occupancy and turnover because their contribution is too small. The real use of these terms is to monitor the use of the referral system and of the day services. However, this does not seem to be the case at the moment because these data are not analyzed to monitor the referral system or the day services. Besides not achieving its objective, if the transfers are not properly categorized they can affect the reporting of other variables such as admissions and discharges.

There is a need to provide the FIO with standardized criteria to correct the data coming from the wards. Although the FIO is doing a commendable job in correcting the data reported by the wards, this is inevitably affected by subjectivity because no written criteria are available on how to identify and correct the data. Most FIOs applied their own judgment on how the numbers should match to balance the entries with the exits in each ward and few of them were routinely visiting the wards and checking secondary data sources such as the mortuary registers against the number of deaths reported from the wards.

Although most FIOs were trained, more support is required. Most of the training sessions lasted for one day and covered software and hardware components and data requirements. However, any training can only grasp the surface of the problem related to data reliability, validation, interpretation and use of the information. Most FIOs last received supervision visits more than three months before and it was frequently found insufficient to solve technical issues. It is also clear that most of the supervision is carried by the Systems and Finance manager and not by the hospital manager. Besides sending the files to the district level, the FIO should send them directly to the central level as well, where a full time person should identify outliers according to the methodology described in Issue 9. Once outliers are identified, the central person should alert the DIO and the FIO about the data problems so that a timely check on the raw data is carried out to correct the outliers.

The central level could also make more use of the types of graphs presented in this analysis to facilitate the comprehension of the relationship between over and under reporting of the individual variables and the over and under estimation of the relative indicators. The results of the data check could be presented in graphical form to relate the over or under-reporting of individual variables with over-under estimation of indicators. This issue has provided examples on how graphs could help visualize the relationship between incorrect reporting and estimation of indicators.

Data transformation

Further variation is introduced when the data are transmitted from each hospital to the upper level. The numbers reported by the FIOs explain only partially the reliability problems experienced by the DHIS. The comparison with the DHIS monthly numbers of reported OPD, operations and beds suggest that the complex data flow from the hospitals to the upper level adds another source of variation. This is generated through the file transformation carried out to make the PTSS files compatible with the DHIS and through the corrections carried out by the district and the central level. This source of unreliability was evident when the data reported and validated in the hospitals were compared with the DHIS monthly data, which differed substantially from the original data produced in the hospitals. Although this comparison was only possible for beds, OPD and operations, substantial differences are likely to be present also for the other DHIS variables.

Reducing the data manipulations between the hospital and the central level is likely to reduce the chance that the final data published by the central level might differ from the original data. The phasing out of the PTSS system is likely to decrease the problems of too many file transformations to render them compatible with the DHIS. However, as already mentioned there is a need to create a direct link between the FIO and the central level by allowing the FIO to send the files directly to the Epidemiology Unit where a full time person should timely identify the outliers to alert the FIOs and the DIO to check and correct specific data.

Implications for management

The indicators should be more frequently used for action. It is unclear if occupancy and turnover are correctly interpreted to rank hospitals in terms of efficiency, find the reasons for low efficiency and take the necessary actions to tackle the causes and monitor if these actions are sorting out any effect. Once it has been ensured that there are no reliability problems, the hospitals belonging to the same category could be plotted according to their occupancy and turnover. As described in Bulletins 7-9 this technique is commonly known as the Pabon Lasso graph. This could be used to cluster the hospitals according to their occupancy and turnover and to formulate hypotheses on the reasons for certain utilization profiles. Some hospitals may not be highly utilized compared with their average category because their patient mix may not reflect the category to which they belong and therefore they may require re-assignment to another category. One important issue is that the average utilization frequently hides a high intra hospital variation, with a few wards being over-utilized and few others being underutilized. Once it is confirmed that patient mix is not behind the problem, it may be hypothesized that certain wards may be inefficiently utilized and some changes may be required to shift the beds from under to over utilized wards and to reallocate resources within the hospital. If all the wards are under-utilized, some extra data may be collected to identify the reason, suggest solutions and verify their impact on utilization.

Conclusions and recommendations

The results of the analysis suggest a few pragmatic solutions to the complex problems affecting the hospital information system. Because improving the reliability of any hospital information system requires considerable efforts, it is necessary to establish priorities and practical recommendations. Although the system will continue to collect the present variables, the efforts should be prioritized first to improve the reporting of the number of beds, inpatient days, discharges and deaths. After the DOH will have succeeded in improving these priority variables, it will be possible to dedicate the attention and the energies to improve the other variables.

Number of useable beds

The number of useable beds should be counted once a year to provide a stable denominator for the whole FY. The problem is not so much whether to include or exclude border beds but to decide a criterion and consistently apply it. The major problem of unreliable reporting will not be solved through more training. Instead, ways must be found to avoid subjective interpretations that are repeated 365 days per year in several wards within a given hospital. Because each hospital is assigned a certain number of beds per year and they do not change daily, their function is to provide a fixed denominator for hospital utilization. If this concept were clear, the number of beds would be counted once a year at the beginning of each FY and not daily as it is currently done. The function of collecting the number of beds is to provide the potential utilization capacity of a hospital, which does not change if a ward is temporarily closed or if some beds are without mattress. Once the number of beds (potential capacity) is decided, it remains stable till when a clear decision is made to increase or decrease it, which is usually at the beginning of the FY. Therefore, counting the number of beds at every night shift does not have a solid conceptual justification and leads to mistakes due to the high variability with which different people count the different types of beds daily. This system defeats the purpose of providing a reliable denominator for the indicators of utilization. The only solution is therefore to count the number of beds at the beginning of each FY, preferably through a survey and to use this number for the whole year.

Occupied beds

In terms of occupancy there is a need to clarify inclusion and exclusion criteria for the occupied beds (inpatient days). At the moment there is a certain variation with which border mothers are included or excluded from the number of occupied beds. If it is decided that border beds are to be included among the useable beds, border mothers should be included as occupying beds and vice versa. Similarly, because a transfer is counted as '1 transfer-in + 1 inpatient day', transfers-in should be taken out from the numerator of occupancy to avoid double counting.

The transfers and day patients should be counted to judge the referral system and not to estimate overall occupancy. The major problem found in recording these types of variables is the lack of clarity about the objective of collecting this information with high potential for misclassification. Transfers and day patients are frequently misclassified with admissions, and discharges leading to misuse of these variables.

Because of the high variation with which transfers and day patients are counted, reporting these variables is at the moment not achieving any objective because of the lack of comparability across hospitals. Because the correct categorization of the above variables requires a much higher effort than with the other variable, collecting the numbers of day patients and transfers would be worth it only if this information were used to achieve their real objective, which is to monitor the day services and the referral system. Because the use of standardized categorization of transfers and day patients would require substantial efforts it should be first piloted to the regional and central hospitals.

Discharges and deaths

The FIO should send the monthly data to the central level where one person should be assigned to monitor the reliability of the data. This can be done through the methodology introduced in Issue 9, which includes (a) the crosschecking of the variation of admissions, discharges plus deaths and discharges, and (b) the identification of outliers. Once the data are checked, the central level should timely alert the DIO and the FIO of the hospitals with problematic data to guide them on the data check and correction. This includes the recounting in each ward of the daily admissions, occupied beds, discharges and deaths. The number of deaths reported by the wards should be compared with the number of deaths in the mortuary to check if they are consistent. Considering that the mortuary registers are the most valid source of information on mortality, this comparison will allow one to improve the reliability of the total number of deaths and stillbirths. There is a need to clarify the inclusion and exclusion criteria related to deaths occurring in emergency room and OPD, and to clarify that stillbirths should not be counted as deaths.

Use of the indicators

While monitoring the improvement of the reporting of the priority variables, the central level should use more graphical tools to increase the understanding of the indicators of efficiency. More use of graphs to plot the data could enhance the understanding of the relationship between individual variables and indicators and improve the chance of using the information for action. The graphs presented in this issue can be used by the central level (a) to understand the relationship between under and over reporting of individual variables and over and under-estimation of occupancy and turnover; and (b) to use the Pabon Lasso graph to cluster the hospitals according to efficiency profiles so that hypothesis on the causes behind different levels of efficiency can be formulated and tested.

Rest of the variables

Once the efforts have succeeded to improve the priority variables, the DOH could dedicate the attention in improving the rest of the hospital information system. The same technique used to identify outliers for the priority variables could be used to identify outliers for OPD, operations, deliveries, stillbirths, live births and the rest of the other variables. The DIO and FIO will then be alerted by the central level to recount the records on the relative registers and correct the data. Because collecting reliable data on the transfers and day patients, and because these data should be used to monitor the efficiency of the referral system, no effort should be wasted on trying

to improve the data collection in district hospitals. The efforts to improve the reliability in central and regional hospitals should be pilot tested first in a few regional and central hospitals.

Role of the central level

The central level has the role of standardizing the quality check and to direct the periphery on what data should be corrected. Because of the complexity of the data check and the need to apply it in a consistent manner, one highly qualified person should do it. The timely receipt of the monthly data from each hospital will allow it to implement a prompt monitoring of the reliability of the reporting so that corrective actions are carried out. Another critical role of the biostatistician will be to popularize the use of the hospital statistics through the mentioned Pabon Lasso and other graphic techniques.

References

Armitage P. Statistical Methods in Medical Research. Blackwell Scientific Publications.1980

Barnum H and Kutzin J. Public Hospitals in Developing Countries. The John Hopkins University Press.

Issue 7. Epidemiology Bulletin. Use of the indicators to assess hospital efficiency. June 2004. DOH.

Issue 8. Epidemiology Bulletin. Methodology to validate the hospital information. November 2004. DOH.

Issue 9. Epidemiology Bulletin. Reliability of hospital indicators in FY03/04 – FY04/05. March 2005. DOH.

ANNEX I Changes in Efficiency between Round 1 and 2

The changes in efficiency between round 1 and round 2 are presented in pairs of figures, such as Figures 1 and 2 for which the left figure represents the change in occupancy and turnover according to the FIO and the right figure shows the changes according to the validation. R1 and R2 besides the name of each hospital indicate the position of the hospital in round one and two in terms of occupancy (X-axis) and turnover (Y-axis). The line linking R1 with R2 indicates the direction of the change in efficiency between the two rounds.

About 60% of the hospitals reported reliable changes between R1 and R2. Only a few graphs are provided as example of consistent changes between round 1 and 2 according to both valid and reported data. The graphs should be seen in terms of the Pabon Lasso rationale, according to which a change from the lower left to the upper right part of the graph indicates increased occupancy and turnover and therefore increased efficiency. Vice versa, a change from the upper right to the lower left section of the graph indicates decreased efficiency. It has to be mentioned that the objective of these graphs is to show which hospitals had consistent changes in utilization and not to rank the hospitals according to their improved efficiency, which requires a longer period than the 48 hours covered in each of the two rounds.

Figures 1-3 show an example of the hospitals that consistently increased their utilization in terms of occupancy and turnover between the two rounds according to both the FIO and the validation. For example, between R1 and R2, St. Mary's (Marianhill) increased its occupancy from 61% to 94% and its turnover from around 3 to about 6 according to the FIO. Similarly, the changes according to the validation were from 68% to 87% for occupancy and from 3 to over 6 for turnover. Figures 3-4 show instead the hospitals, which consistently moved, from the upper right to the lower left and therefore deteriorated in occupancy and turnover according to both the FIO and the validation. For example, Pongola had a sharp decline in both occupancy and turnover according to the FIO and the validation. For about 40% of the hospitals, the changes in the indicators reported by the FIO were inconsistent with the changes of the validation. Figures 5-6 provide an example of less severe inconsistent changes. For example, while between R1 and R2 Ladysmith reported an increase in occupancy between 57% to 69%, the real change according to the validation was a decline in occupancy from 56% to 50%.

Figure 1-2 Consistent improvement between R1 and R2

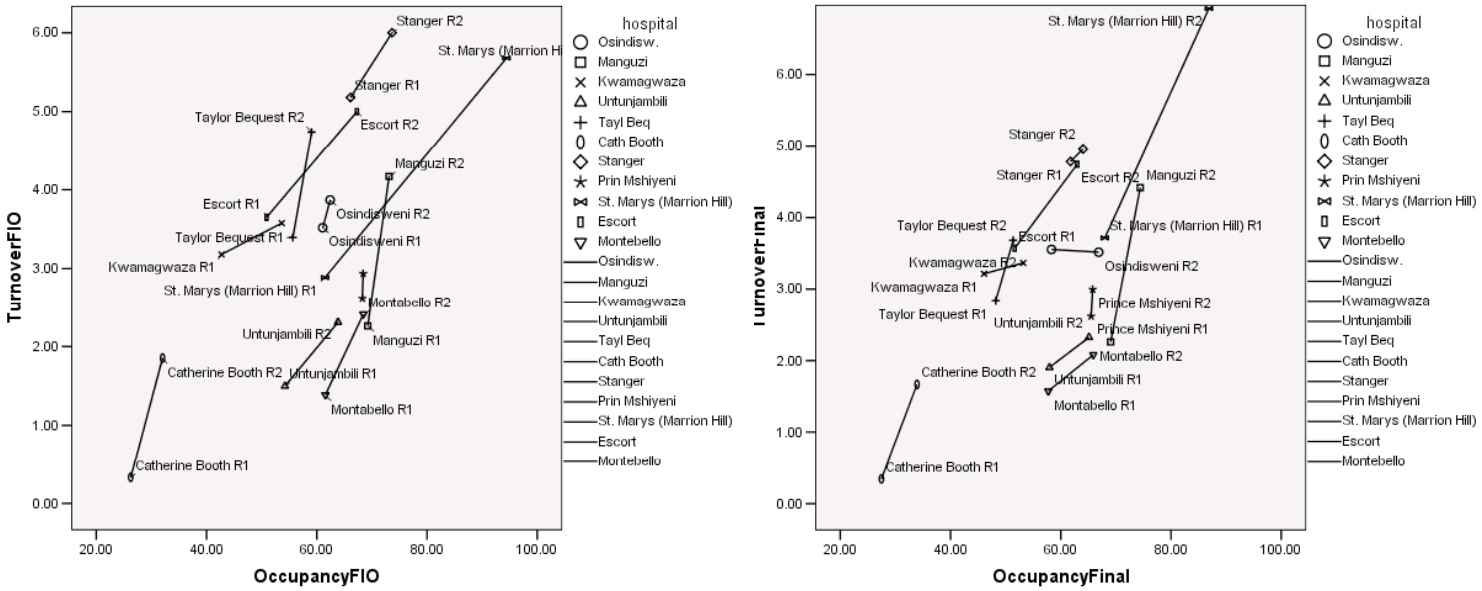
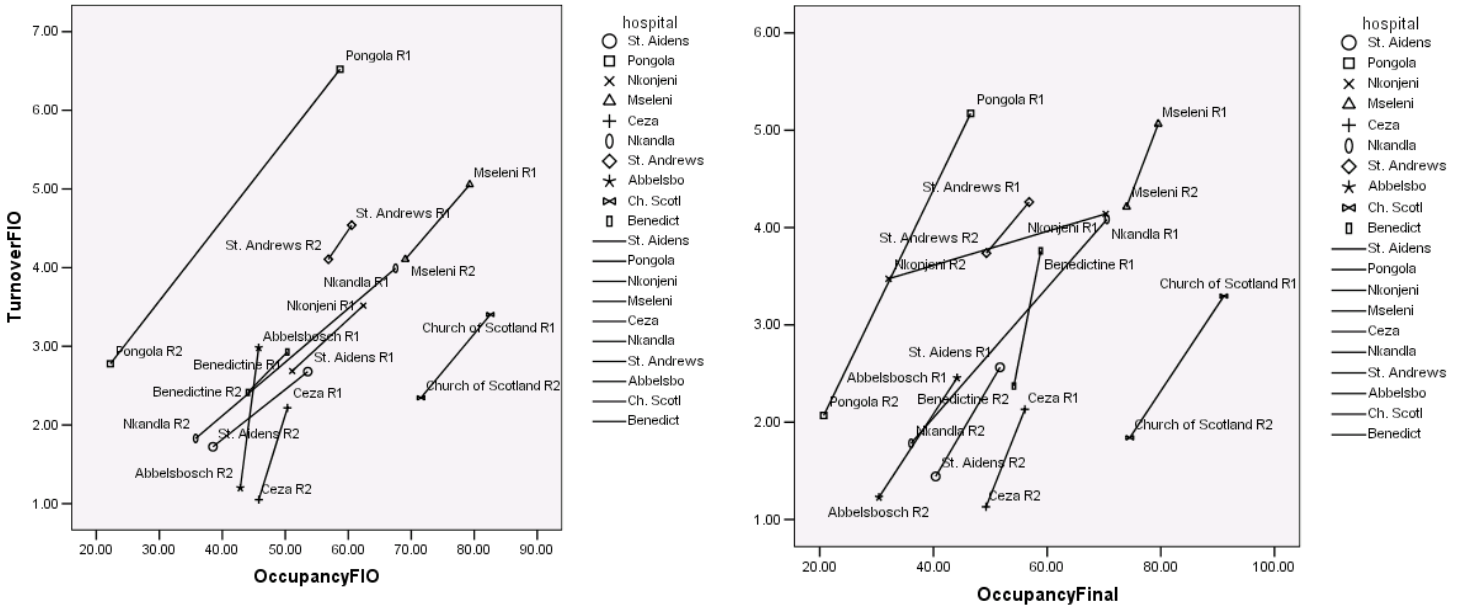
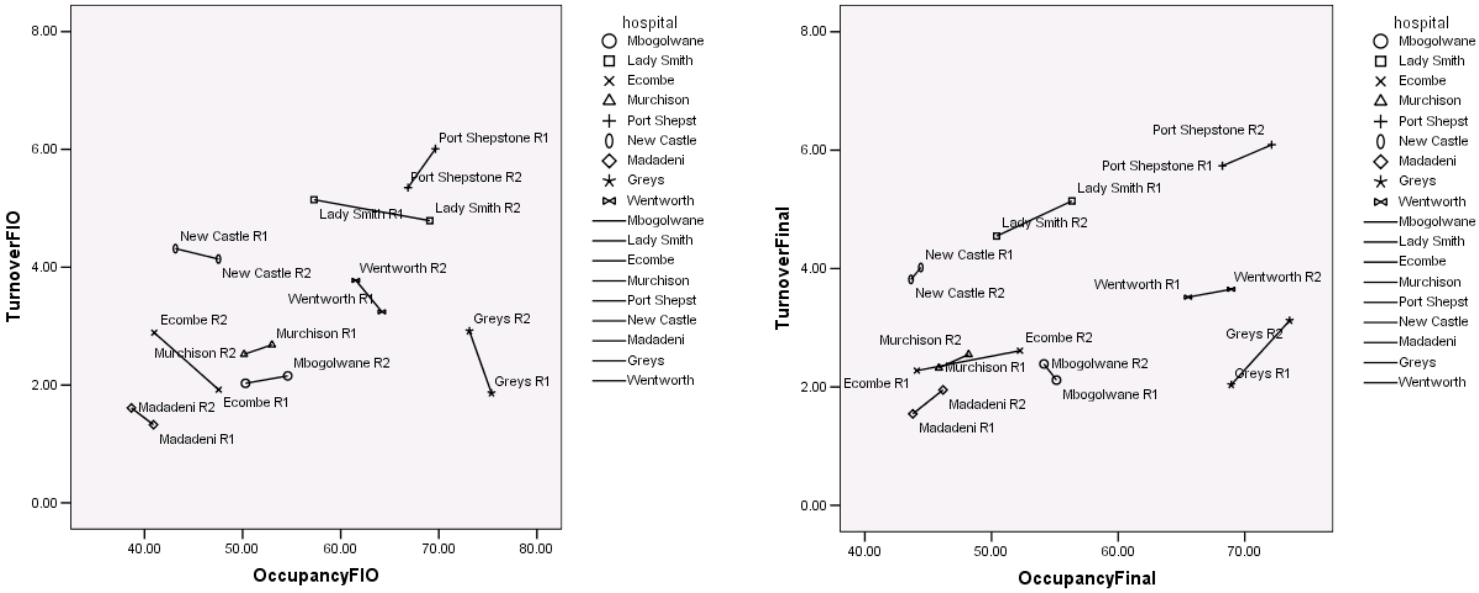


Figure 3-4 Consistent deterioration between R1 and R2

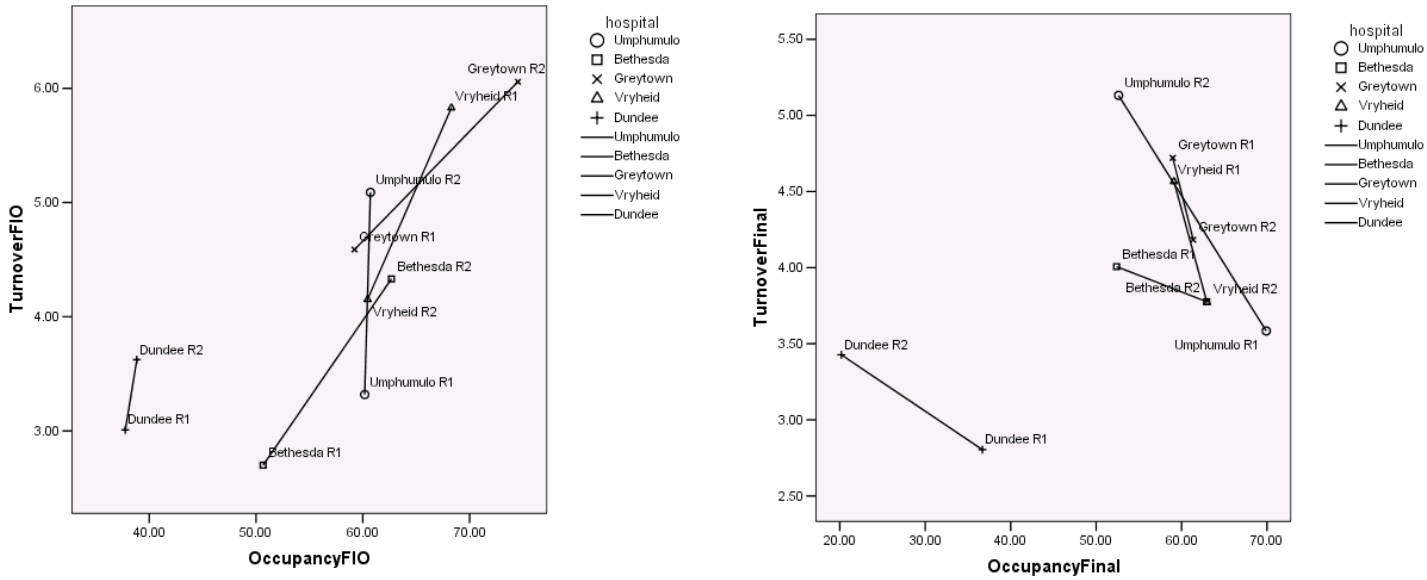


Figures 5-6 Minor inconsistent changes



Figures 7-8 provide an example of major inconsistencies where the changes in occupancy and turnover diverged substantially. It can be noted that Greytown and Bethesda improved substantially both in occupancy and turnover according to the FIO, but the validation showed a deterioration in turnover. Dundee and Umphumulo improved their turnover and maintained their occupancy according to the FIO, while the validation confirmed the increase in turnover but recorded a decline in occupancy. Vryheid declined substantially in both occupancy and turnover according to the FIO while the validation recorded a decline in turnover but not in occupancy.

Figure 7-8 Major inconsistent changes



It can be concluded that although more than half of the hospitals reported reliable estimates of change in efficiency between different points in time, there were still a sizable number of hospitals for which the reporting produced misleading changes. This has implications for the management, which relies on correctly estimating the changes in the indicators of utilization to judge how efficiently they are running their hospitals. Unreliable reporting provides a false impression that hospitals might have improved or deteriorated in utilization while the reality might be the opposite. It is therefore in the managers' interest to ensure that reporting is reliable.

ANNEX II – Methodology

Hospital management is based on the assumption that hospital indicators are reliable. If the data on which these indicators are based is inaccurate, management cannot make informed decisions and cannot measure the effect of such decisions. The methodology employed in this study was designed specifically to address the issue of data validity and reliability. To achieve this aim, a standardized methodology was developed and consistently applied to each hospital. Topics addressed in this section include sample selection and description, instrumentation, data collection procedures, data analysis, and limitations.

Sample Selection and Description

The validation was conducted in all public hospitals within KwaZulu Natal. To have a comprehensive picture of the high variation affecting the data processed by PTSSH, all the public hospitals with the exception of Albert Luthuli were included. Inkosi Albert Luthuli hospital was omitted because the system is completely computerized and does not follow the same pattern of data collection, data entry and flow of the rest of the hospitals. Therefore a study of its own with a different methodology would be needed to validate the information produced by Albert Luthuli.

To better capture the variation in data reliability and to test the hypothesis that the number of useable beds does not change considerably during a financial year, each of the hospitals was visited twice. The first round lasted from the 20th October 2004 to the 24th March 2005, while the second round was conducted from 25th March to the 27 August 2005.

Data collection forms

During the course of round one and two, six instruments were used to collect the necessary data. Four survey forms and two questionnaires were designed and developed specifically to come up with the most valid numbers against which to compare the data collected and processed by the PTSSH.

i. FIO Questionnaire

The FIO questionnaire was designed to give a profile of the FIO. The objective was to cover fundamental information about length of service, training, understanding of various data definitions and problems faced doing their work. The questionnaire was revised in round two to probe at issues that were highlighted as problematic in the first round. For example round one suggested confusion amongst FIO's regarding useable, authorized and border beds. This issue was explicitly addressed in the second round to FIO questionnaire.

Table 1 List of Hospitals Visited for the Validation Survey

Location	Tertiary	Regional	District	Specialized
Sisonke (D43)			EG & Usher Memorial Taylor Bequest St. Apollinaris Christ the King	Khotsong SANTA
Ugu (DC 21)		Port Shepstone	G.J. Crooks Murchison St Andrew's	D. Farrell SANTA
ethekwini (Durban)	King Edward VII	Prince Mshiyeni R.K. Khan Mahatma Gandhi Addington St. Aidens	Wentworth St. Mary's (Marian Hill) MCords Osindisweni	Charles James SANTA Don McKenzie SANTA FOSA TB King George V
uMgungundlovu (DC 22)	Grey's	Edendale	Northdale Appelsbosch	Doris Goodwin SANTA Richmond Chest
Ilembe (DC 29)		Stanger	Montebello Umphumulo Untunjambili	
Umzinyathi (DC 24)			Church of Scotland Greytown Dundee Charles Johnson Memorial	
Uthukela (DC 23)		Ladysmith	Estcourt Emmaus	
Uthungulu (DC 28)		Ngwelezane Lower Umfolozi	Catherine Booth Eshowe Mbongolwane Nkandla Ekombe KwaMagwaza	
Amajuba (DC 25)		Madadeni Newcastle	Niemeyer	
Zululand (DC 26)			Nkonjeni Ceza Vryheid Benedictine Itshelejuba Pongola	St. Francis Thulasizwe Siloah Lutheran Mountain View
Umkhanyakude (DC 27)			Manguzi Mosvold Mseleni Bethesda Hlabisa	

ii. Nurses / Matrons Questionnaire

The nurses' questionnaire was designed to assess the nurses' understanding of the data definitions. Nurses are the primary source of all data entering the PTTSH. They admit patients, record information about patients and fill out the midnight statistics reports. Their understanding of the basic data definitions is key to accurate reporting.

During round one, one nurse per ward was selected at random to have a representative sample of the nurses involved in the data collection. However this was fraught with problems because many nurses claimed to have never worked on night shift when the midnight statistics are captured (FIO1) and therefore they were unaware of how to fill out the midnight statistics form. Of those nurses who were successfully selected randomly, only a few were able to answer the entire questionnaire.

In an attempt to circumvent this problem, the matron was interviewed instead of a nurse in the second round. This approach met with much greater success and achieved the initial objective of finding a way to validate how correctly and standardized is the use of terms related to the variables being collected.

Both questionnaires were constructed using multiple choice, yes / no and some open ended questions. In order to help improve the reliability of the answers some questions were asked twice in a different format. For example in the case of boarders such as mothers accompanying their children it was asked how consistently they were correctly classified through two questions.

iii. Ward Capture Forms

This form was at the heart of the validation. It recorded five data sources that were then used to create one final set of validated data. How this valid number was decided upon will be covered in the next section. These five data sources were:

- Ward Register – Used to capture data gleaned from the ward registers.
- FIO1 – Used to capture 48 hours worth of data from the FIO's midnight statistics forms.
- FIO2 – Used to capture any obvious changes made to the data by the FIO before data entry into the PTSSH.
- Other Data sources- Used to capture any additional data that was available, e.g. matrons reports, death notification books, summary reports, etc.
- Manual Count –the number of beds present in the ward were manually counted and compared with the number recorded by the FIO.

Table 2 provides the data elements that were captured for each ward from the five data sources. Some data elements were not always applicable to every ward, e.g. deliveries in a surgical ward, and some data sources were not available for every data element, e.g. manual count of discharges.

Table 2 Data Elements captured from each ward (where applicable)

Data Element	Validation Source
Admission	Ward Register / Matrons Report
Discharges	Ward Register / Matrons Report
Ward Transfers-in	Ward Register / Matrons Report
Ward Transfers-out	Ward Register / Matrons Report
Hospital Transfers-in	Ward Register / Matrons Report
Hospital Transfers-out	Ward Register / Matrons Report
Deliveries	Ward Register / Delivery Register
Still births	Ward Register / Delivery Register
Live births	Ward Register / Delivery Register
Total births	Ward Register / Delivery Register
Number of Live births who died	Ward Register / Delivery Register / Mortuary
Deaths (excluding those among live births)	Death Register / Mortuary Register
Useable boarder beds counted as Boarder beds	Ward Register / Sister in Charge / Manual Count
Useable boarder beds counted as Useable Adult beds	Ward Register / Sister in Charge / Manual Count
Useable adult beds	Manual Count of Beds
Occupied adult beds	Ward Register / Sister in Charge
Useable juvenile beds	Manual Count of Beds
Occupied juvenile beds	Ward Register / Sister in Charge
Useable cribs	Manual Count of Beds
Occupied cribs	Ward Register / Sister in Charge
Useable cots	Manual Count of Beds
Occupied cots	Ward Register / Sister in Charge
Useable incubators	Manual Count of Beds
Occupied incubators	Ward Register / Sister in Charge
Boarders counted as Borders	Ward Register / Sister in Charge
Boarders counted as Inpatients	Ward Register / Sister in Charge
Day Patient	Ward Register / Matrons Report
Pass out/absconded	Ward Register / Matrons Report
Floor Beds / Mattresses	Manual Count of Beds
Beds with No Mattress	Manual Count of Beds

iv. Outpatients Department (OPD) Data Collection Form

OPD data were captured from the service statistics forms used in the hospitals. During the first round the team became aware that the method used to calculate the headcount is different from one hospital to the next. Some hospitals use a register in the admissions office to record OPD headcount data, some use the register found at the nurses sorting station while yet others use till slip readings from the revenue department. The total OPD headcount figure was captured for the 48 hours, which corresponded to the dates chosen for the inpatient validation.

Two problems arose during the first round using this methodology:

- The 48 hours period used to validate most of the data could not be used for OPD data because most district hospitals only collect OPD headcount data once a month. This meant that OPD data was unavailable for the 48 hours covered by the validation. This was taken into account in the second round where it was decided that the period of validation for OPD would be the most recent month. Therefore the monthly OPD data was validated for all hospitals.
- During the first round it became clear that each hospital was collecting the OPD data from a different source. Therefore to improve comparability between hospitals the method employed in the second round was to use the same data source for each hospital. It was decided that the most accurate data source, and the one least likely to be reporting visits rather than headcount, was the registers found in the admitting office. Most hospitals have an admitting office that acts as a central entry point into the OPD system for all patients. Here first case and repeat case registers were used to compute the total OPD headcount for the month.

v. Theatre Data Collection Form

As with the outpatient headcount data, the theatre data is reported on the service statistics forms used by the hospitals. Prior to the April 2005 update of PTSSH, it recorded information about caesarian sections, thoracic, orthopedic and other operations, etc. Since the update many of these categories have been changed or removed.

As for the OPD data, round one collected theatre data for a 48-hour period while in round two the monthly figure was instead validated. Both rounds validated data pertaining to caesarian sections and other operations.

vi. Mortuary Data

The mortuary register was used to validate data on deaths and stillbirths from the FIO. A record is kept of every body brought to the mortuary and the data source was compared to the data being provided by the wards.

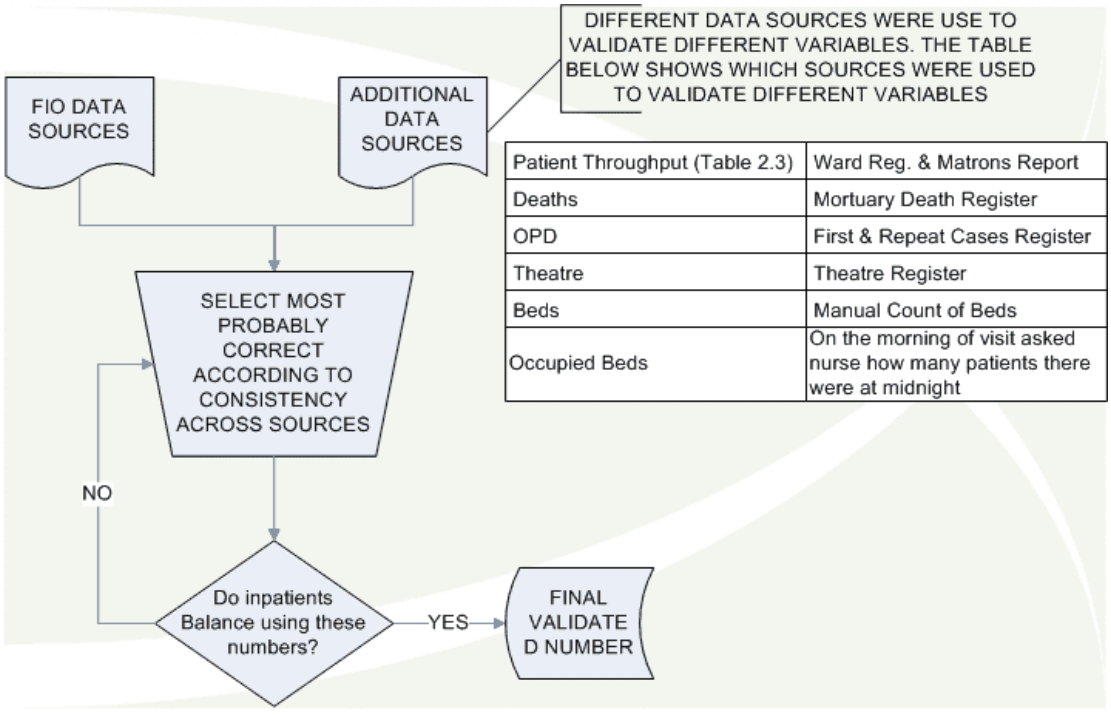
Data Collection and Standardization Procedures

The highest priority in this study was to ensure the validity of the final data against which the PTSSH data were validated. The method used to arrive at the final valid number needs to be systematic and consistent. This section will describe in detail, the procedure followed when validating the hospital data.

i. Data Capture Forms

The ward capture Form was used to collect the same variables from different sources in a standardized format to arrive at the final validated number. With five or more potential sources of information all often reporting a different value, it was imperative to apply the same validation rationale to each ward in every hospital. Figure 1 outlines the process used to decide on the final valid number. The first step was to collect all the relevant information from all the relevant sources onto the ward capture form. The sources included Ward Register, Ward summary book, FIO1 (midnight statistics before correction), FIO2 (midnight statistics after correction), Matrons Report and any other sources pertaining to the patient throughput data.

Figure 1 Decision making process used to validate the data.



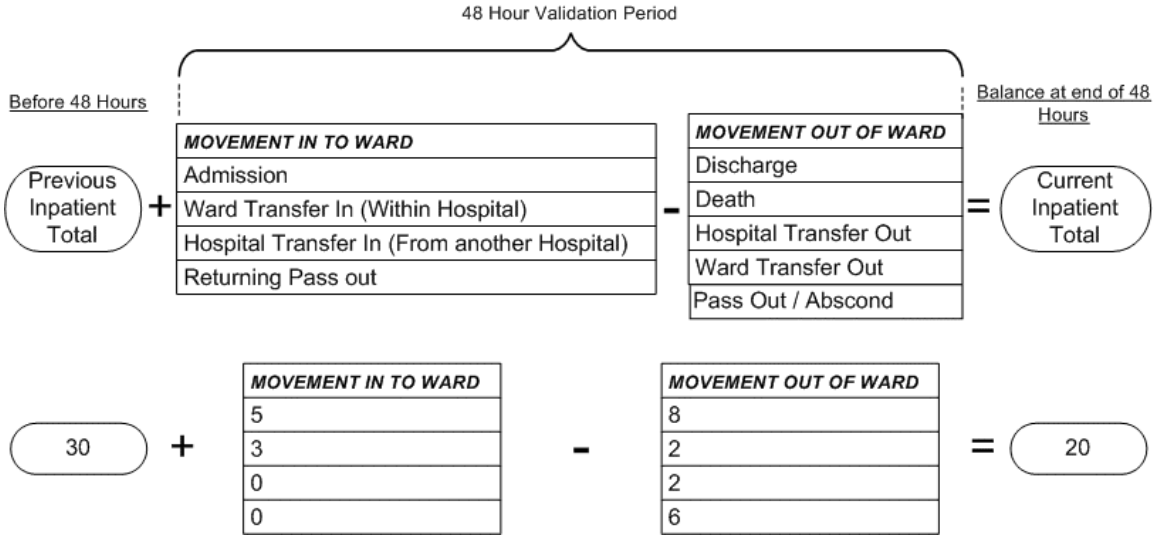
Patient Throughput (Table 2.3)	Ward Reg. & Matrons Report
Deaths	Mortuary Death Register
OPD	First & Repeat Cases Register
Theatre	Theatre Register
Beds	Manual Count of Beds
Occupied Beds	On the morning of visit asked nurse how many patients there were at midnight

The summarized data was collected. The previous step was the basis for the validation. This allowed one to select the valid numbers for each variable related to the patient throughput. The data sources mentioned above provided all the possible combinations in which the flow of patients entering the ward in 24 hours is balanced by the patients exiting the wards.

The validation was based on the consistency between the number of patients who entered and left each ward in the previous 48

hours. Figure 2 shows a hypothetical situation in which at the beginning of the 48-hour period there were 30 patients. Of the initial 30 patients, 18 left and 8 new patients arrived producing a balance of 20 inpatients at the end of the 48 hours. This was the valid number against which to compare the patients' throughput reported by the FIO.

Figure 2 Example of Balancing the Inpatient Headcount.



The number of deaths reported to the FIO was validated by using the mortuary register to record the date, time and ward from which bodies arrived during the 48 hour validation period. These numbers were then tied back to the deaths reported in the wards. If a discrepancy arose further investigation was done to uncover the problem.

The number of useable boarder, adult, juvenile, crib, cot and incubator beds were validated by doing a manual count of all the beds in each ward. This procedure was repeated in each of the two rounds making it possible to gauge the stability of useable bed numbers.

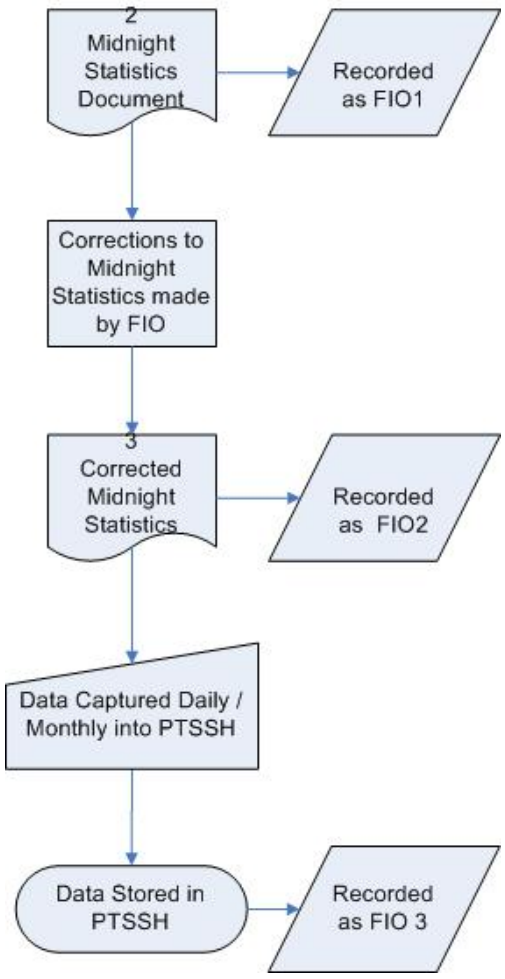
ii. Screen Capture Forms

During the second round, the 48 hours covered by the first round were traced to check the numbers entered by the FIO. The first round was able to cover the previous 48 hours, capturing all the different steps in which errors could be generated. This included the number reported by the ward to the FIO and the final correction applied by the FIO. Because the team left before the data were entered, the team checked the data, which were entered by the FIO when the team revisited the same hospitals. This was done by

querying the PTSSH to find all the variables, which were entered for the 48 hours covered by the first round.

Data are captured daily into PTSSH by the FIO and monthly printouts can be obtained. However, the data related to the specific day can be retrieved on the computer screen but they cannot be printed. Therefore, during the second round, the validation team retrieved the data entered by the FIO into the PTSSH for the 48 hours covered during the first round. The screen capture forms (FIO3) introduction was intended to supplement the previously collected FIO1 (midnight statistics hard copies) and FIO2 (changes made to the hard copies by the FIO) data. The relationship between FIO1, FIO2 and FIO3 can be seen in Figure A.3.

Figure A.3 Relationship between data and Recorded FIO variables



iii. Outpatients Department (OPD) Data Collection Form

The method of validation was slightly different in the two rounds. In the first round an attempt was made to validate the FIO's Total OPD headcount number. This meant that if they included casualty, the team included the casualties. If the FIO used the doctor's registers, the team used the doctor's registers.

In the second round, a different approach was taken. In an attempt to improve interhospital standardization, the OPD headcount was calculated by using the registers found in the central admitting office. This source of information was thought to be the most reliable and accurate measure of Outpatient headcount data.

iv. Theatre Data Collection Form

Theatre data was validated using the theatre register. It was one of the most well kept document data sources in the hospital making validation more clear-cut.

Limitations

Though every care was taken to produce a solid methodology, a few problems emerged during the survey.

- For most variables, the validation period was related to the previous 48 hours. This was recent enough to allow a good recollection by the nursing staff and access to any temporary written information stored in the registers of the wards. This short period was used to obtain the most valid data against which to compare the data reported to the FIO by the nurse on the midnight report and to the final data entered into the PTSSH after any correction done by the FIO.
- Validating the Inpatient data posed a challenge to the methodology. It was discovered during the first round that very few wards kept additional information about occupied beds. A count is done at midnight and recorded on the midnight statistics forms. The figure is rarely recoded in any permanent form. The general availability of only one data source made validation difficult. This limitation was addressed in the second round asking the nurse in charge how many patients there were that evening. As it was only eight or nine hours after midnight, the number reported by the nurse has a high probability of being valid. This data would not yet have reached the FIO and so they were contacted by fax a few weeks later to ask for the inpatient data for that specific day.

- In the first round of the survey daily data was collected for both the theatre and Outpatients department. This proved problematic in some of the smaller hospitals as this information was only collected on a monthly basis. In order to improve comparability, these data was collected on a monthly timeframe from all hospitals in the second round.

The intricacies of the Outpatient department data can cause a wide variation in the reporting of outpatient headcount. The data could be derived from a number of different sources, contain different services and clinics and record either the number of visits or headcount data. This makes comparisons between hospitals difficult. To overcome this short fall the procedure was altered slightly in the second round. Where available, data was collected from the hospital Outpatients admitting office. All cases recorded in the OPD registers related to the month preceding the survey were counted.