



Quality of a routine data collection system for health: case of Kinondoni district in the Dar es Salaam region, Tanzania

D.O. Simba

Department of Community Health
Muhimbili University College of Health Sciences (MUCHS)
Dar es Salaam, Tanzania.
dsimba@muchs.ac.tz

M.A. Mwangu

Department of Development Studies
Muhimbili University College of Health Sciences (MUCHS)
Dar es Salaam, Tanzania
mmwangu@muchs.ac.tz

Contents

1. [Introduction](#)
2. [Methodology](#)
 - 2.1 [Study site](#)
 - 2.2 [Existing health management information system](#)
 - 2.3 [Study type and procedures for data collection](#)
 - 2.4 [Study sample](#)
 - 2.5 [Data analysis](#)
 - 2.6 [Definition of terms and measurement methods](#)
 - 2.7 [Limitations of the study](#)
3. [Findings](#)
 - 3.1 [Data completeness](#)
 - 3.2 [Accuracy](#)
 - 3.3 [Timeliness of reporting](#)
4. [Discussion](#)
5. [Conclusion and recommendations](#)
6. [References](#)

Key words: Records management, information systems management, health sector

1 Introduction

Data that are accurate, complete and delivered on time to users (as information) are an important aspect in health planning, management and decision making. Evidence-based plans and decisions must, of necessity, be based on accurate, complete and timely data. Good planning and management depends on the availability of reliable, accurate and timely information (Garner, Harpham and Annett 1992; Robey and Lee 1990). Most often, decisions on whether or not to use routinely collected data for planning and

decision making are made subjectively. It was reported in the Pacific Islands that, although routine reports showed a gradual shift in the mortality pattern towards non-communicable diseases, the information was not used to effect policy changes, due to doubt over the quality of data (Finau 1994).

Despite concerns about the poor quality of data collected routinely through health facilities, the extent of reliability has not been well explored. Therefore, decisions to use or not to use these data are made subjectively. This study, therefore, was aimed at assessing the quality of the routine data collected at facility and district level under the Health Management Information System (HMIS). Specifically, the study sought to determine data completeness, and accuracy and timeliness in reporting to assess the impact on the provision of health services and to ultimately give recommendations for improvement.

[top](#)

2 Methodology

2.1 Study site

Kinondoni is one of the municipalities that form the city of Dar es Salaam. With a population of about 1088867, it is the most densely populated district in the country (National Bureau of Statistics 2003). The municipality is served by a district hospital, four health centres and 19 dispensaries owned by the government. There are also 10 hospitals and 148 dispensaries that are owned by the private sector. In addition, there are four parastatal dispensaries in the municipality. The municipal medical officer of health (MMOH) is in charge of health activities, assisted by other members of the council health management team (CHMT). A member of the team, known as the research coordinator, supports the MMOH in overseeing HMIS activities in the municipality. Although the municipal health office has a computer, there is no special software for processing routine data; instead data are processed through a manual system. It was further noted that members of the CHMT, including the research coordinator, lack spreadsheet and database software skills.

2.2 Existing health management information system

Data collected from registers, that is, the outpatient register (Book 5), are compiled on a monthly basis into a data book (Book 2) available in each health facility, whether public or private. Each health facility compiles 131 data items from the registers and forms on a monthly basis, while 505 data items are compiled annually. In addition, each facility calculates nine quarterly indicators and four annual indicators. A selection of raw data and the indicators are forwarded to the districts (municipalities) quarterly, using the report book (Book 10). This process is repeated at regional level and the compiled data are then sent to national level.

2.3 Study type and procedures for data collection

A cross-sectional study was conducted in the Kinondoni municipality, Dar es Salaam region, Tanzania, from January to May 2003. Data collection was done by documenting the available records. A review of the available HMIS-related documents (registers and data books) was done to determine the levels of accuracy and completeness of data processing and timeliness of reporting to the district. A selection of data items to be observed was made so that various data categories, such as outpatient, laboratory and maternal and child health, were included. In the observation schedule, data were tracked retrospectively according to the frequency of reporting as follows: past four weeks for daily recorded data; past four months for monthly reported data; past four quarters for data reported quarterly; and past four years, for data reported annually. A pre-test was done in three health facilities to check the relevance of the data collection tools and the overall research approach. The results of the pre-test were used to refine and prepare the final data collection tools.

2.4 Study sample

Kinondoni district was conveniently selected because it is within Dar es Salaam where the researchers are based. All public (23) and parastatal health facilities (5) were selected because of their small numbers. In addition, a random sample of 41 private health facilities, representing 25% of all private health facilities, was selected. A representative sample of 25% is considered to be sufficient in health

systems studies (Kielman 1995). Altogether, 69 facilities were covered in this study. Table 1 shows the number of health facilities by ownership.

Table 1 Health facility type by ownership

Health facility type	Health facility ownership			
	Government	Private	Parastatals	Total
Hospitals	1	2	0	3
Health centres	2	1	1	4
Dispensaries	20	38	4	62
Total	23	41	5	69

2.5 Data analysis

Data were analysed using EPI Info software. Frequency and cross tabulation were done and descriptive analyses were used in interpreting the data.

2.6 Definition of terms and measurement methods

Data completeness was defined as the extent to which the observed records were filled in. This was measured by determining the proportion of cells that were filled in among the selected tables.

The extent to which the table cells were completed when entered in the health facility data books for the year 2001 was used to assess the completeness of data. Of the 61 tables in the data book, [Tables 25a and 29](#) that deal with outpatient attendances and the top ten diseases respectively were deliberately selected for this purpose. The main reason for selecting these tables was the fact that outpatient services were offered by virtually all health facilities, and this enabled analysis to be done by type and ownership. In both tables, 25a and 29, the total number of cells filled in were counted against the total number of cells expected to be filled in. The tables had a total of 177 and 250 cells respectively. Data completion was calculated as follows:

$$\text{Completeness of data} = (A * 100)/B$$

Where A = Number of cells not filled in

B = Total number of cells in the table 25a and 29

Data accuracy was defined as the extent of variation between actual and recorded/reported figures. These were measured by calculating the difference between summation done by health workers as found in the data book and a re-calculated value from the registers done by research assistants. This is referred to in this study as *calculation variance*. The difference between the totals recorded in the data book and the number reported by health workers to the district was also calculated. In this study, this was referred to as *transcription variance*.

Calculation and transcription variance were calculated by comparing the number of attendances as found in the register books and summations recorded in the report book for the lowest and the highest attendance months. Calculation of variance was done using the following formula:

$$\text{Calculation variance} = ((B - \sum A) / B) * 100$$

$$\text{Transcription variance} = ((C - B)/C) * 100$$

Where A = Record made in the register from each patient

B = Total number (for the respective data item) as found in the data book (Book 2)

C = Total number (for the respective data item) as found in the report book (Book 10)

Timeliness of reporting was calculated using the reporting rate of the various forms to the Municipality

for the past four quarters or years, depending on the frequency of reporting of the respective form. Timeliness of reporting was calculated as follows:

$$\text{Reporting rate} = X * 100 / Y,$$

Where X = Number of reports received in the respective quarter/year

Y = Number of health facilities required to send the respective report

2.7 Limitations of the study

The few private health facilities that had all the required data denied researchers access to the valuable information, which limited the scope of observation and analyses for some of the indicators. For example, owing to the limited response of private facilities that had completed the data book and report book (4 out of 45), data accuracy had to be calculated for government facilities only.

Being an urban municipality, Kinondoni might not be a good representative of the majority of the districts in Tanzania that are mostly rural.

[top](#)

3 Findings

3.1 Data completeness

This study found that the average data completion rate was 64,2%. The completion rate was found to be higher for government facilities (71,6%) than for private facilities (54,8%). However, these findings were based on 34 (51,5%) health facilities, as the rest did not fill in the data books at all. When facilities that did not have data were included, the overall data completeness rate was reduced by half, from 64,2% to 32,1%. Further analyses showed that the lack of data was more pronounced among private facilities (66,7%) compared to government facilities (9,5%) (Table 2).

Table 2 Completeness of selected OPD tables by ownership

Health facility by type	Completion rate of Table 25a and 29				
	Table 25a		Table 29		Total
	Cells filled (a)	Total cells (b)	Cells filled (c)	Total cells (d)	% Completion rate = (a+c) *100/(b+d)
Government (N=19)	1829	2223	3162	4750	71,6
Private (N= 15)	1221	1755	1796	3750	54,8
Total/Average (N=34)	3050	3978	4958	8500	64,2

3.2 Accuracy

Calculation of data accuracy was done for government facilities only, because private facilities had inadequate data to allow analysis. Of the 21 government facilities, 10 had enough data to process an accuracy indicator. Table 3 shows that the *calculation variance* for the lowest month was 90,3% (ranging from -316,1% to 93,3%.), while for the highest month the variance was 92,0% (ranging from 33,6% to 94,6%). The *transcription variance* from the data book to the report book was 86,4 % (ranging from -52,8% to 59,9%).

Table 3 Level of accuracy of HMIS data among public health facilities

	Average	Range

Calculation variance (low months)	90,3	-316,1 to 93,3
Calculation variance (high month)	92,0	33,6 to 94,6
Transcription variance	86,4	-52,8 to 59,9

In general, there was about a two-fold variation between the actual and the recorded data for each step made in data processing. Since the steps are conducted in series, the resulting variation was about four-fold. For example, if 20 patients were recorded in the register, a two-fold increase will result in 40 patients being recorded in the data book. With a two-fold increase in transcribing data to the reporting book, 80 patients might be reported. The wide variation suggests the existence of significant inaccuracies in compiled and reported data, thus making the reported data unreliable.

3.3 Timeliness of reporting

Reporting rate for the various forms ranged between 0% (for the notifiable diseases quarterly report – F003) and 43,3% (for the annual report – F005) as shown in Table 4. The annual report form had the highest reporting rate (43,3%) compared to the rest of the forms. In fact, it was the only form that showed an increase in the reporting rate over the years, from 0% in 1998 to 40% in the year 2001.

Table 4 Reporting rate from the facility to the district

Form number	Number received last quarter/year	Percent (N = 270)	Number received two quarters/years before	Percent (N = 270)	Number received three quarters/years before 1999	Percent (N = 270)	Number received four quarters/years before 1998	Percent (N = 270)
F001 district summary report (annual)	14	5,2	73	27,0	27	10,0	0	0
F003 notifiable disease report (quarterly)	0	0	20	7,4	53	19,6	0	0
F004 quarterly report	15	5,6	17	6,3	6	2,2	14	5,2
F005 annual report	117	43,3	81	30,0	30	11,1	0	0

[top](#)

4 Discussion

Inadequacies in the HMIS data are of growing concern in Tanzania and many other developing countries (WHO/AFRO 1996). Data collected in these countries have been reported to be incomplete, inaccurate and not delivered on time. As a result, it is usually not used for the reason that it is of poor quality and unreliable (HERA 2000; Rajesh 1993). This study has shown the magnitude of incompleteness, inaccuracies and untimeliness in reporting data collected from the health facilities. Although this study has shown that the available data are inadequate for the municipality to use for decision making, the relatively high completeness rate among public health facilities (about 71,6%) justifies the use of HMIS data for decision making, if restricted to the public sector. However, this denies municipal authorities the opportunity of assessing the contribution of the private sector in the

delivery of health care services. In the era of health sector reforms, inadequate data from private health facilities undermine the public–private partnership agenda; bearing in mind that in the Kinondoni municipality, 85% (158 of 186) of the health facilities belong to the private sector. It has been reported that in the wake of reforms it will be increasingly difficult for any government to determine the allocation of resources without taking into account the interests of the private sector (Howden-Chapman and Ashton 2000).

This study has shown that for each step taken in data processing a two-fold magnitude of error occurs. Over and underreporting of data by up to 50% has been reported in various studies (Chongsuvivatwong and Nitimanop 1991). It should be noted, however, that this study addressed data accuracy from the registration of patients to reporting to the next level. It would be interesting to measure the completeness of reporting between the actual number of patients attended to and the number recorded in the registers to obtain a holistic picture on data accuracy.

One of the fundamental problems of the HMIS is that the system does not have a built-in mechanism for data validation. It therefore depends on the ability and competence of the data compilers and supervisors to detect errors and to rely on their personal initiatives to detect and rectify them. The reality, however, is that supervision at district and facility levels is inadequate and therefore it is unlikely that serious errors will be noticed as data flow up to the national level (Ministry of Health 2002).

Less than half of the annual reports were available one year after the expected date. Therefore, even if other quality parameters were satisfactory, the delay in reporting would have undermined prompt decision making by the municipal health authorities.

[top](#)

5 Conclusion and recommendations

The inadequate availability of data from the comparatively larger private sector, as found in this study, would undermine the quality of comprehensive plans and the overall management of health services in the municipality. The wide margin of error limits prediction of the future, rendering the planning exercise futile. With the margin of error being two- to four-fold, health managers cannot effectively monitor the performance of health workers or the efficient use of resources. This will undermine an evidence-based rewarding system by favouring those who can manipulate data in their favour. The resulting uncertainty would make health managers suspicious of data collected through the routine system, thus reinforcing reliance on the use of intuition that will further undermine evidence-based decision making and planning. In the long run, this situation will continue to suppress the already undermined information culture.

In each step of the compilation, transcription and reporting process, there is a two- to four-fold margin of error. Reducing the number of steps in compiling and transcribing data from one register or form to another can significantly improve the quality of the routine data collection system. Reporting to the higher level can be done by either sending all or sampled pages for analyses at the receiving end. In addition to reducing the margin of error, this will:

- reduce the burden on the local health staff in compiling data that are not used locally. Reducing the workload will ensure an improvement in the quality of the data;
- foster commitment and a sense of ownership to the HMIS, since health staff will only analyse data that are relevant at their respective levels;
- facilitate trimming and streamlining the system to a few indicators selected by the local staff. The fact that the higher level will not demand compiled data will allow more flexibility for local staff to decide on relevant indicators for monitoring performance; and
- expedite reporting of data from the lower level to the higher level, thus ensuring timeliness of data for decision making.

With the ongoing computerization of the HMIS at municipal/district level, analysis of raw data from facilities should be much easier.

The Kinondoni case study in Tanzania may provide useful lessons and insights for the rest of the countries in Africa and developing countries in general, as they all face similar conditions affecting the quality of routine health data collection. The key lessons and insights are the following:

- In general, this case study emphasized the importance of having simple and operationally feasible routine health information systems for purposes of generating basic data that are easy to collect, reliable and hence more useful for decision making purposes.
- The more complicated and multi-staged the system is, the greater the chance of introducing errors, either during the transcription or the reporting stages. The greater the margin of error, the less reliable the data and hence the reduced value of the data in management and decision making.
- Data validation and strict supervision of data collectors are crucial in minimizing errors in a routine data collection system.

[top](#)

6 References

Chongsuvivatwong, V. and Nitimanop, P. 1991. Voices from the grassroots. *International Journal of Epidemiology* 20:1133-1136.

Finau, S.A. 1994. National health information system in the Pacific Islands: in search of a future. *Health Policy and Planning* 9 (2):161-170.

Garner, P., Harpham, T. and Annett, H. 1992. Information support for urban primary health care. *World health forum* 13:244-248.

HERA. 2000. Review of the health management information system in Tanzania. Ministry of Health, Tanzania. (Unpublished report).

Howden-Chapman, P. and Ashton, T. 2000. Public purchasing and private priorities for healthcare in New Zealand. *Health Policy* 54(1): 27-43.

Kielman, A. 1995. *Assessing health needs, services and systems: protocol for rapid data collection*. London: AMREF and Macmillan.

Ministry of Health. 2002. Tanzania Joint Ministry of Health/Donor review. (Unpublished report).

National Bureau of Statistics. 2003. *2002 Population and housing census general report*. Dar es Salaam: President's Office (Planning and Privatization).

Rajesh, K. 1993. Streamlined records benefit maternal and child health care. *World Health Forum* 14:305-307.

Robey, J.M. and Lee, S.H. 1990. Information system development in support of national health programme monitoring and evaluation: the case of Philippines. *World Health Statistics Quarterly* 43:37-43.

WHO/AFRO. 1996. *A study of the rapid assessment of the health management information system in Mauritius*. Summary of health system research reports, 1994-1997.

Disclaimer

Articles published in SAJIM are the opinions of the authors and do not necessarily reflect the opinion of the Editor, Board, Publisher, Webmaster or the Rand Afrikaans University. The user hereby waives any claim he/she/they may have or acquire against the publisher, its suppliers, licensees and sub licensees and indemnifies all said persons from any claims, lawsuits, proceedings, costs, special, incidental, consequential or indirect damages, including damages for loss of

profits, loss of business or downtime arising out of or relating to the user's use of the Website.

[top](#)

ISSN 1560-683X

Published by [InterWord Communications](#) for Department of Information and Knowledge Management,
University of Johannesburg