

## Lessons from the field

### Assessing the quality of data aggregated by antiretroviral treatment clinics in Malawi

Simon D Makombe,<sup>a</sup> Mindy Hochgesang,<sup>b</sup> Andreas Jahn,<sup>c</sup> Hannock Tweya,<sup>d</sup> Bethany Hedt,<sup>b</sup> Stuart Chuka,<sup>e</sup> Joseph Kwong-Leung Yu,<sup>f</sup> John Aberle-Grasse,<sup>b</sup> Olesi Pasulani,<sup>g</sup> Christopher Bailey,<sup>h</sup> Kelita Kamoto,<sup>a</sup> Erik J Schouten<sup>i</sup> & Anthony D Harries<sup>j</sup>

**Problem** As national antiretroviral treatment (ART) programmes scale-up, it is essential that information is complete, timely and accurate for site monitoring and national planning. The accuracy and completeness of reports independently compiled by ART facilities, however, is often not known.

**Approach** This study assessed the quality of quarterly aggregate summary data for April to June 2006 compiled and reported by ART facilities ("site report") as compared to the "gold standard" facility summary data compiled independently by the Ministry of Health supervision team ("supervision report"). Completeness and accuracy of key case registration and outcome variables were compared. Data were considered inaccurate if variables from the site reports were missing or differed by more than 5% from the supervision reports. Additionally, we compared the national summaries obtained from the two data sources.

**Local setting** Monitoring and evaluation of Malawi's national ART programme is based on WHO's recommended tools for ART monitoring. It includes one master card for each ART patient and one patient register at each ART facility. Each quarter, sites complete cumulative cohort analyses and teams from the Ministry of Health conduct supervisory visits to all public sector ART sites to ensure the quality of reported data.

**Relevant changes** Most sites had complete case registration and outcome data; however many sites did not report accurate data for several critical data fields, including reason for starting, outcome and regimen. The national summary using the site reports resulted in a 12% undercount in the national total number of persons on first-line treatment. Several facility-level characteristics were associated with data quality.

**Lessons learned** While many sites are able to generate complete data summaries, the accuracy of facility reports is not yet adequate for national monitoring. The Ministry of Health and its partners should continue to identify and support interventions such as supportive supervision to build sites' capacity to maintain and compile quality data to ensure that accurate information is available for site monitoring and national planning.

Bulletin of the World Health Organization 2008;86:310–314.

Une traduction en français de ce résumé figure à la fin de l'article. Al final del artículo se facilita una traducción al español. الترجمة العربية لهذه الخلاصة في نهاية النص الكامل لهذه المقالة.

#### Problem

By 31 December 2006, Malawi had enrolled 82 000 patients in its free national antiretroviral treatment (ART) programme. Each quarter, data from all ART clinics are aggregated for national

reporting on ART scale-up. This information is essential to monitoring site performance, guiding national planning and supporting sustained funding.

Despite increasing reliance on sites to aggregate data, the completeness and accuracy of sites' reports was unknown.

We therefore conducted an operational study during regular supervisory visits to assess the quality of data in the site reports. Specific objectives were to: i) determine the completeness and accuracy of key case registration and outcome data compiled by ART clinics,

<sup>a</sup> Clinical HIV Unit, Ministry of Health, PO Box 30377, Lilongwe, Malawi.

<sup>b</sup> US Centers for Disease Control and Prevention, Global AIDS Programme, Malawi.

<sup>c</sup> International Training and Education Centre on HIV, Seattle, WA, United States of America.

<sup>d</sup> Lighthouse Trust, Lilongwe, Malawi.

<sup>e</sup> Malawi Business Coalition against AIDS, Blantyre, Malawi.

<sup>f</sup> Taiwan Medical Mission, Mzuzu Central Hospital, Mzuzu, Malawi.

<sup>g</sup> Medecins sans Frontieres-Belgium, Thyolo District Hospital, Malawi.

<sup>h</sup> World Health Organization, Geneva, Switzerland.

<sup>i</sup> Management Sciences for Health, Lilongwe, Malawi.

<sup>j</sup> Family Health International, Malawi Country Office, Lilongwe, Malawi.

Correspondence to Mindy Hochgesang (e-mail: mindyho@gmail.com).

doi:10.2471/BLT.07.044685

(Submitted: 13 June 2007 – Revised version received: 8 November 2007 – Accepted: 12 December 2007 – Published online: 4 March 2008)

Simon D Makombe et al.

ii) compare national data summarized from site reports versus supervision reports, and iii) analyse characteristics associated with sites' capacity to compile quality data.

## Local setting

A standardized monitoring and evaluation (M&E) system is key to Malawi's national ART programme.<sup>1-3</sup> The system is based on WHO's recommended tools for ART monitoring, and includes one master card for each ART patient and one patient register at each ART facility. At each ART clinic visit, patients are monitored according to standard primary and secondary outcomes. Each quarter, sites are asked to complete cumulative cohort analyses, summarizing the characteristics and outcomes of all patients ever enrolled at the site. Each quarter, the Ministry of Health and its partners conduct supervisory visits to all public sector ART sites, during which the supervision teams review all patient master cards with clinic staff, update the ART patient registers, and independently generate cumulative cohort analyses for each site. These visits ensure the quality of reported data.<sup>4</sup>

## Approach

The Ministry of Health HIV/AIDS Unit and partners conducted supervisory visits, during July and August 2006, to assess case registration and patient outcome information as of 30 June 2006. At the beginning of the visit, site reports completed by facility staff were collected. All patient master cards were reviewed, registers updated as appropriate and the cumulative cohort analysis was aggregated by the clinic and supervision team together to generate the supervision report. Data from site and supervision reports were analysed with SPSS (Statistical Package for Social Sciences 2006, SPSS Inc. Chicago, Illinois, United States of America).

By 30 June 2006, 56 000 persons had started treatment in 94 public sector sites. Five sites (5%) had electronic data systems for ART reporting and were excluded from the analyses. Analysis of site report data was therefore conducted on the 89 sites that were expected to have completed a paper-based analysis.

Completeness and accuracy of data from site reports were compared to

that of supervision reports for six case-registration variables (total patients ever started on ART, and of those, the number who were men; children; teachers among all occupation categories; patients in WHO clinical stage 4 at ART initiation; and patients with history of tuberculosis) and two outcome variables (total patients alive on ART and the number on first-line regimen). Data were considered complete if sites had recorded any value for a specific field; data were considered accurate if none of the examined variables in the site report were missing or showed a difference of more than 5% compared to the supervision report. In addition to site-level analyses, national totals summarized from site reports were compared with totals from supervision reports.

We examined whether site characteristics such as facility-type, burden (maximum number of new patients the clinic can start each month), length of time providing treatment, and number of data clerks were associated with complete and accurate data in sites' reports using binary logistic regression.

## Findings and relevant changes

### Data completeness

Sixty-two (70%) sites provided complete data for all six case-registration fields in the site report. Aggregates for the number of patients starting ART because of tuberculosis history and patient occupation were least likely to be complete with 26 (29%) and 21 (24%) sites having incomplete information for these fields, respectively. Outcome data were complete at 71 (80%) sites.

### Data accuracy

The accuracy of many site reports was classified as unacceptable (data missing or > 5% difference compared to supervision report). Case registration data were accurate in only 36 (40%) sites. Fields less accurately reported included numbers of patients in clinical stage 4 at ART initiation, patients with history of tuberculosis, and male patients, with 33 (37%), 29 (33%), and 12 (25%) sites reporting inaccurate data respectively.

Sixty-four (72%) sites reported accurate outcome data on the cumulative number of patients alive and on ART and number of patients on first-line regimen at the end of the quarter.

### Accuracy of national summary figures using site report data

Outcome data were added for the 89 sites and summary totals from site and supervision reports compared. The national summary of persons alive and on ART, based on site report data, resulted in a 5.4% undercount (22 390 patients in site reports versus 23 665 from supervision reports.) The national figures for patients on first-line regimen were 20 500 patients in the site report versus 23 380 patients in the supervision report, an underestimate of 12.3%.

### Association between facility characteristics and data quality

Table 1 presents the relationship between facility characteristics and the completeness and accuracy of outcome data. Several factors were associated with quality data, including a higher burden (starting more patients each month), having dedicated clerks for recordkeeping, having a visit by a zonal ART supervisor before the national supervision visit, being located in the central or southern region, having provided ART for a longer period of time, and being a type of facility other than a rural hospital or health centre. It should be noted that all rural hospitals and health centres are low burden, thus have similar odds ratios. Further analysis revealed that the length of time a facility had provided ART explained the association between burden and data quality due to the correlation between site burden and length of time that sites had provided treatment.

## Lessons learned

In many national health information systems, sites routinely collect and report summarized data. This was the first systematic audit of routinely reported data generated by ART sites using Malawi's national ART M&E system. Based on thresholds of accuracy considered critical for programme planning, this audit shows that, at the national level, the ART programme in Malawi still requires skilled assistance to compile reliable data. Data quality varied considerably between sites and was associated with facility characteristics. While some factors, such as the length of time sites have been providing treatment, are not easily changed, other "actionable" factors, including supervision and human resources for record keeping (clerks)

Table 1. Sites with acceptable data quality

Clinic characteristic	Total number of sites ( <i>n</i> = 89)	Data completeness		Data accuracy	
		Number (and proportion) of sites with complete outcome data	Odds ratio (90% confidence interval)	Number (and proportion) of sites with accurate outcome data	Odds ratio (90% confidence interval)
<b>Burden</b>					
Medium or medium-high burden (max. 50 or 75 new patients/month)	30	28 (0.93)	ref	27 (0.90)	ref
Low burden (max. 25 new patients/month)	59	43 (0.73)	0.19 (0.05–0.70) <sup>a</sup>	37 (0.63)	0.19 (0.06–0.56) <sup>a</sup>
<b>Type</b>					
Central or district hospital	22	20 (0.91)	ref	19 (0.86)	ref
Mission hospital	26	21 (0.81)	0.42 (0.10–1.83)	19 (0.73)	0.43 (0.12–1.50)
Rural hospital or health centre	20	14 (0.70)	0.23 (0.05–1.00)	11 (0.55)	0.19 (0.06–0.68) <sup>a</sup>
Other (NGO, military)	21	16 (0.76)	0.32 (0.07–1.41)	15 (0.71)	0.40 (0.11–1.44)
<b>Region</b>					
Southern	35	33 (0.94)	ref	29 (0.83)	ref
Central	37	28 (0.76)	0.19 (0.05–0.73) <sup>a</sup>	28 (0.76)	0.64 (0.24–1.70)
Northern	17	10 (0.59)	0.09 (0.02–0.37) <sup>b</sup>	7 (0.41)	0.15 (0.05–0.43)
<b>Number of clerks for record-keeping<sup>c</sup></b>					
None	6	4 (0.67)	ref	3 (0.50)	ref
1 clerk	66	52 (0.79)	3.90 <sup>d</sup> (1.30–11.69) <sup>a</sup>	47 (0.71)	3.55 <sup>d</sup> (1.32–9.55) <sup>a</sup>
2 clerks	15	15 (1.00)		14 (0.93)	
<b>Length of time site providing treatment</b>					
0–3 months	33	20 (0.61)	ref	15 (0.46)	ref
4–12 months	27	24 (0.89)	5.20 (1.62–16.68) <sup>a</sup>	24 (0.90)	9.60 (2.61–21.56) <sup>b</sup>
13 or more months	29	27 (0.93)	8.77 (2.30–33.51) <sup>b</sup>	25 (0.86)	7.50 (3.01–30.62) <sup>a</sup>
<b>Previous supervision visit in the quarter</b>					
No supervision	61	47 (0.77)	ref	41 (0.67)	ref
Supervision	28	24 (0.86)	1.79 (0.64–4.95)	23 (0.82)	2.24 (0.89–5.68)

NGO, nongovernmental organization; ref, reference.

<sup>a</sup> *P* < 0.05.

<sup>b</sup> *P* < 0.01.

<sup>c</sup> No information on clerks at two sites.

<sup>d</sup> Odds ratio shows the increase in odds for every additional clerk.

were associated with increased data quality. Sites with longer experience in providing ART were more likely to generate quality data, which may be attributed to the intensive quarterly supervision that sites have received. With rural hospitals and health centres found to have lower quality data than other hospital types, a better understanding of what specific characteristics (e.g. burden, staffing levels, management, incentives) impact data quality at the facility is needed.

Errors in site report data did impact national totals: national totals of persons currently on treatment and persons on first-line regimen were undercounted by more than 5% and 12%, respectively, when national figures were derived

from site reports. Left uncorrected, the difference between reported and true figures would likely increase over time. Since drug orders are based on these national summary data, errors in the number of patients on first-line regimen would underestimate the quantity of first-line ARV drugs needed annually by nearly 35 000 tins. Malawi has not yet stocked out of ARV drugs, and an adequate drug supply is critical to the continued scale-up of the programme.

Analyses were based on data from supervisory visits as the gold standard; given the number of records to be reviewed during supervision visits (30 sites have started more than 500 patients), some errors may have been made during supervision. Still, the effect is likely

to be small. This study assessed data completeness and accuracy of aggregated facility-level data; the accuracy of individual patient information, including inaccurate diagnoses or recording of data, could not be determined retrospectively.

Some important lessons have been learned (Box 1). Results of this analysis show limited accuracy and completeness in clinic-level reporting, and so support the WHO recommendation that ART M&E tools be kept to a minimum and focus on critical indicators. Ongoing supervision is critical to ensure quality data during scale-up, and decentralization of such capacity-building activities may help to ensure sustainability of these efforts. Allocation

Simon D Makombe et al.

of dedicated trained personnel is critical to maintaining quality patient records and M&E data collection. An electronic data system that includes mechanisms to improve data quality (e.g. numerical validation, bounds checking, alerts) may facilitate site-level data management and reporting; this is currently being piloted in four district hospitals in Malawi. Finally, Malawi has had good success in simple incentives (e.g. certificates) for sites with good data quality. Such factors appear to motivate sites to maintain or improve data quality. As ART continues to scale-up, Malawi's ART programme must ensure systems are in place to generate critical national information. ■

### Acknowledgements

Supervisory and monitoring visits were financed through WHO. Andreas Jahn

### Box 1. Lessons learned

#### Recommended factors to ensure quality data

- A simple, prioritized list of key reporting needs that does not overburden the monitoring and evaluation system;
- Proactive ongoing on-site assessment and capacity building (through supervision and on-the-job training);
- Adequate numbers of skilled human resources (e.g. dedicated data clerks) for collecting and maintaining data;
- Possible use of technology (appropriately designed computer systems) with built-in mechanisms to check completeness and validity of data to facilitate data collection and aggregation;
- Rewards for high quality data or other incentives for clinic staff.

is a technical adviser in monitoring and evaluation placed at the Lighthouse Trust. Erik J Schouten is a technical adviser to the Clinical HIV Unit at the Ministry of Health, Malawi. Anthony D Harries is the technical adviser to the ART programme within the Clinical

HIV Unit at the Ministry of Health, Malawi. He is also a professor at the London School of Hygiene and Tropical Medicine.

**Competing interests:** None declared.

## Résumé

### Evaluation de la qualité des données agrégées par les centres de traitement antirétroviral au Malawi

**Problématique** Les programmes nationaux de traitement antirétroviral (ART) étant en phase d'élargissement, il est essentiel de disposer en temps utile d'informations complètes et exactes pour la surveillance des sites et la planification nationale. On ne sait souvent pas si les rapports compilés indépendamment par les établissements dispensant des traitements ART sont complets et exacts.

**Démarche** Les auteurs de cette étude ont évalué la qualité des données résumées et agrégées par trimestre pour la période avril-juin 2006 provenant des rapports des établissements dispensant des traitements ART (« rapports de site ») par comparaison avec celle des données résumées de référence, compilées indépendamment pour ces établissements par l'équipe de supervision du Ministère de la santé (« rapport de supervision »). Ils ont également comparé la complétude et l'exactitude des principaux paramètres d'enregistrement des cas et d'issue du traitement. Les données ont été considérées comme inexacts s'il manquait des paramètres dans les rapports des sites ou si les valeurs des paramètres différaient de plus de 5 % des valeurs indiquées par les rapports de supervision. En outre, les auteurs ont comparé les résumés nationaux établis à partir de deux sources de données.

**Contexte local** La surveillance et l'évaluation du programme national de traitement ART du Malawi s'appuient sur les outils recommandés par l'OMS pour la surveillance ART. Elles prévoient une carte informatisée pour chaque patient sous traitement ART et un

registre des patients pour chaque établissement dispensant ce type de traitement. Chaque trimestre, les sites réalisent des analyses de cohortes cumulatives et des équipes du Ministère de la santé effectuent des visites de supervision dans tous les sites ART du secteur public pour s'assurer de la qualité des données rapportées.

**Modifications pertinentes** La plupart des sites disposent de données complètes pour l'enregistrement des cas et les issues du traitement ; néanmoins, nombre d'entre eux ne fournissent pas des données exactes sur plusieurs points critiques, et notamment les raisons du placement sous traitement, l'issue du traitement et le schéma thérapeutique. Le résumé national à partir des rapports des sites a fait apparaître une sous-comptabilisation de 12 % du nombre total de personnes dans le pays sous traitement de première intention. Plusieurs caractéristiques au niveau des établissements sont associées à la qualité des données.

**Enseignements tirés** Si de nombreux sites sont en mesure de générer des résumés de données complets, leurs rapports ne sont pas encore suffisamment exacts pour la surveillance nationale. Le Ministère de la santé et ses partenaires doivent continuer à définir et à appuyer des interventions comme la supervision de soutien afin de renforcer la capacité des sites à conserver et à compiler des données de qualité et de garantir ainsi la disponibilité d'informations exactes pour la surveillance des sites et la planification nationale.

## Resumen

### Evaluación de la calidad de los datos reunidos por los centros de tratamiento antirretroviral en Malawi

**Problema** A medida que se expanden los programas nacionales de tratamiento antirretroviral (TAR), es fundamental contar con información completa, puntual y exacta para la vigilancia de los centros y la planificación nacional. Sin embargo, a menudo se desconoce la exactitud e integridad de los datos reunidos independientemente por los centros de TAR.

**Enfoque** Se evaluó la calidad de los datos agregados trimestrales que habían compilado y notificado los centros de TAR para el periodo de abril a junio de 2006 (« informe del centro »), comparándolos con los datos resumidos « de referencia » reunidos independientemente por el equipo de supervisión del Ministerio de Salud (« informe de supervisión »). Se compararon la integridad

y exactitud de variables clave del registro de casos y de los resultados, y se decidió considerar inexactos los datos de los informes de los centros cuando las variables analizadas no constasen o difiriesen en más de un 5% de las consignadas en los informes de supervisión. Procedimos asimismo a comparar las cifras recapitulativas nacionales obtenidas a partir de esas dos fuentes de datos.

**Contexto local** La vigilancia y evaluación del programa nacional de TAR de Malawi está basada en los instrumentos recomendados por la OMS para la vigilancia del TAR. Ello incluye una ficha maestra para cada paciente sometido a TAR y un registro por paciente en cada centro de TAR. Cada trimestre, los centros realizan análisis de cohortes acumulativos y equipos del Ministerio de Salud realizan visitas de supervisión a todos los centros de TAR del sector público para verificar la calidad de los datos notificados.

**Cambios destacables** La mayoría de los centros disponían

de datos completos en lo relativo al registro de casos y los resultados; sin embargo, muchos centros no aportaron datos precisos en varios campos cruciales, como la razón para comenzar el tratamiento, los resultados y el régimen. El resumen nacional basado en los informes de los centros subestimó en un 12% el total nacional de personas bajo tratamiento de primera línea. La calidad de los datos estaba relacionada con diversas características de los centros.

**Enseñanzas extraídas** Si bien muchos de los centros pueden generar datos recapitulativos completos, los informes de los centros aún no tienen la exactitud exigible para la vigilancia nacional. El Ministerio de Salud y sus asociados deberían seguir identificando y apoyando intervenciones tales como la supervisión de apoyo a fin de desarrollar la capacidad de los centros para mantener y reunir datos de calidad de modo que pueda disponerse de información precisa para la vigilancia de los centros y la planificación nacional.

## ملخص

### تقييم جودة المعطيات المجمعة في عيادات معالجة المصابين بالفيروسات القهقرية في ملاوي

من الفيروسات القهقرية. وكل ثلاثة شهور تستكمل المواقع التحليل التراكمي للأثر، وتجري أفرقة من وزارة الصحة زيارات إشرافية على مواقع المعالجة من الفيروسات القهقرية التابعة للقطاع العام للتأكد من جودة المعطيات المبلغ عنها.

**التغيرات:** لقد استكملت معظم المواقع تسجيل البطاقات والمعطيات الخاصة بالنتائج، إلا أن الكثير منها لم يبلغ عن المعطيات الدقيقة لعدد من الحقول البالغة الأهمية في المعطيات والتي تشمل سبب البدء، والنتيجة، ونظام المعالجة. وقد أدى استخدام الخلاصة الوطنية للتقارير الميدانية إلى نقص حسابي مقداره 12% في العدد الوطني الإجمالي للأشخاص الذين يعالجون بأدوية الخط الأول، وهناك علاقة بين العديد من خصائص مستوى المرفق وجودة المعطيات.

**الدروس المستفادة:** رغم أن باستطاعة العديد من المواقع الحصول على خلاصات كاملة حول المعطيات فإن دقة التقارير الصادرة عن المرافق غير كافية للرصد الوطني. وينبغي على وزارة الصحة وعلى شركائها مواصلة العمل للتعرف على المداخلات من قبيل الإشراف الداعم بغية بناء القدرات في المواقع للاستمرار في جمع المعطيات الرفيعة الجودة وتقديم الدعم لتلك المداخلات لضمان توافر المعلومات الدقيقة للرصد الميداني والتخطيط الوطني.

**المشكلة:** في إطار النهوض بالبرامج الوطنية لمعالجة المصابين بالفيروسات القهقرية لا بد من أن تكون المعطيات مستكملة ودقيقة وتعبّر عن الوقت الملائم لموقع الرصد والتخطيط على الصعيد الوطني. وكثيراً ما يتعذر معرفة دقة ومدى كمال التقارير التي تجمعها مرافق معالجة الفيروسات القهقرية بشكل منفرد.

**الأسلوب:** يقيم الباحثون في هذه الدراسة جودة المعطيات الملخصة والمجمعة خلال ثلاثة شهور هي نيسان/أبريل وأيار/مايو وحزيران/يونيو، والتي تبلغ عنها مرافق معالجة الفيروسات القهقرية (تقارير ميدانية) ومقارنتها بالمعايير الذهبية للمعطيات الملخصة التي جمعها المرافق بشكل مستقل من قبل فريق الإشراف من وزارة الصحة (تقرير الإشراف). وأخضعت للمقارنة مدى دقة واكتمال السجلات الرئيسية للحالات والمتغيرات الخاصة بالنتائج، وقد اعتبرت المعطيات غير صحيحة إذا لم تكن المتغيرات متوافرة أو كانت مختلفة بمقدار يزيد على 5% من تقارير الإشراف، ثم قارن الباحثون الملخصات الوطنية التي حصلوا عليها من مصدرين من مصادر المعطيات.

**الموقع المحلي:** يرتكز رصد وتقييم البرنامج الوطني لمعالجة المصابين بالفيروسات القهقرية في ملاوي على أداة توصي باستخدامها منظمة الصحة العالمية لرصد تلك المعالجة، تتضمن بطاقة رئيسية لكل مريض يتلقى المعالجة للفيروسات القهقرية وسجلاً آخر للمريض في كل مرفق للمعالجة

## References

1. Libamba E, Makombe S, Mhango E, de Ascurra Teck O, Limbambala E, Schouten EJ, et al. Supervision, monitoring and evaluation of nationwide scale-up of antiretroviral therapy in Malawi. *WHO Bull.* 2006;84:257-336.
2. Harries AD, Gomani P, Teck R, de Teck OA, Bakali E, Zachariah R, et al. Monitoring the response to antiretroviral therapy in resource-poor settings: the Malawi model. *Trans R Soc Trop Med Hyg* 2004;98:695-701. PMID:15485699 doi:10.1016/j.trstmh.2004.05.002
3. Harries AD, Libamba E, Schouten EJ, Mwansambo A, Salaniponi FM, Mpazanje R. Expanding antiretroviral therapy in Malawi: drawing on the country's experience with tuberculosis. *BMJ* 2004;329:1163-6. PMID:15539674 doi:10.1136/bmj.329.7475.1163
4. Lowrance D, Filler S, Makombe S, Harries A, Aberle-Grasse J, Hochgesang M, et al. Assessment of a national monitoring and evaluation system for rapid expansion of antiretroviral treatment in Malawi. *Trop Med Int Health* 2007;12:377-81. PMID:17313509