Quality of case management of sexually transmitted diseases: comparison of the methods for assessing the performance of providers

L.M. Franco,[1] C.C. Daly,[2] D. Chilongozi,[3] & G. Dallabetta[4]

This article examines the reliability and validity of direct observation of patient–provider encounters, interviews with providers, and use of patients simulating sexually transmitted diseases (STD) as methods for assessing the quality of STD case management in developing countries. Data were collected during an STD health facility survey in Malawi; the performance of 49 providers was observed, and the providers were also interviewed; 20 of them were visited by a simulated patient complaining of urethral discharge. Agreement (based on the kappa statistic) was generally poor between direct observation and provider-interview data, and also between direct observation and simulated-patient data. In contrast, percentage agreements between direct observation and simulated-patient data were often high. Multiple observations on providers indicated that a provider's behaviour is not consistent across several patients. Simulated-patient data are probably the best in reflecting normal performance, but their feasibility for routine quality assessment is limited because the provider's behaviour is not consistent and would require multiple data points. Direct observation data are the best option for assessing quality if the results are assumed to reflect better than normal levels of quality of care. Data from interviews with providers should be viewed with caution, because they may reflect provider knowledge and not necessarily performance.

Introduction

The quality of the health care process in developing countries has been assessed by many studies using the following methods: direct observation of patient–provider encounters, review of records, exit interviews with clients, interviews with providers, and inventories of facilities, drugs and supplies (*1–7*). Assessing the quality of case management of sexually transmitted diseases (STD) is now receiving more attention because the treatment of STDs has become an important component of acquired immunodeficiency syndrome (AIDS) control programmes. WHO has included proper case management and counselling of STD patients among the proposed AIDS prevention indicators (*8*), and also developed a facility-based quality assessment protocol for STD case management (*9*). This protocol, which incorporates observation, provider interviews, simulated patients, and inventories of facilities, drugs and supplies, is similar in concept to other facility-based assessment surveys developed by WHO for other diseases (*7*).

One difficulty in quality assessment is judging the process of care. There are many untested assumptions about the validity of observation data (*5*), and nothing has been published on the comparison of reliability and validity of data-collection methods for quality assessment in developing and developed countries. One unpublished study (*10*), conducted in Peru, compared direct observation of case management of childhood diarrhoea with simulation using a doll. This study found that health workers performed similarly, both in simulated circumstances and when carrying out the same activities under observation in normal clinical conditions. Published studies are needed on the reliability and validity of the various methods so that quality assurance programmes can determine how best to assess quality, and supervisors can select appropriate methods.

[1] Consultant, Family Health International. At present, Senior QA Advisor, University Research Corporation, Bethesda, MD, USA. Requests for reprints should be sent to this author at the following address: 2862 Waverly Way, Livermore, CA 94550, USA.

[2] STD Advisor, Family Health International. At present, STD Advisor, John Snow Inc., Boston, MA, USA.

[3] STD Officer, National AIDS Control Programme, Ministry of Health and Population, Lilongwe, Malawi.

[4] Director, Technical Support, AIDSCAP Division, Family Health International, Arlington, VA, USA.

Reprint No.**5808**

Background

Quality assessment methods differ in their ability to obtain reliable data about what takes place during a normal provider–patient encounter. Summarized below are the known or presumed advantages and disadvantages for the most frequently used methods.

• *Review of patients' records.* The information provided in such records is easy to extract and readily available at the convenience of the compiler, and can be assessed objectively using explicit criteria (*11*). This information can be collected on a large sample, which allows assessment of the health care process for rare conditions. However, the information present in patients' records is often incomplete, particularly as regards counselling, and it is not possible to recover missing information. In addition, in outpatient settings in developing countries, facility-based patients' records are often scanty or not used at all.

• *Direct observation of provider–patient encounters.* A more complete picture of what providers do during case management can be ascertained by direct observation, but the diagnostic thought process of the provider is difficult to observe without discussion (*8*). Furthermore, the presence of an observer could alter a provider's behaviour for better or worse.[a] Finally, direct observation would generally be able to provide information about the most common types of cases only; it would be difficult to generate an adequate sample size for rare conditions.

• *Interviews with providers.* These interviews yield information about the person's knowledge, but may not reflect his/her actual performance. Such information can be incomplete if the providers have difficulty in visualizing abstractly what they do with an actual patient. However, interviews can provide information about knowledge related to more serious and less common conditions, as well as the provider's considerations in the diagnostic process.

• *Exit interviews with patients.* Interviewing patients after they have received health care can supply information about what the provider did and what the patient learned during the encounter. Such exit interviews are not intrusive to the actual encounter, but the patients themselves may not be able to distinguish or understand certain aspects of a physical examination, and they may not remember all the tasks the provider carried out.

• *Simulated patients.* The simulated-patient method involves a trained observer posing as a patient, who records what happened during the encounter with the provider. As the providers do not know they are being observed, the information gathered should accurately reflect their normal performance. This method has been used successfully in family planning, but presents problems for STD case management because the simulated patient is not really infected and has no symptoms to show. Using simulated patients is also resource-intensive.

Data collected to assess the quality of a provider's performance have several problems relating to reliability, validity, and feasibility. *Reliability*, in this case, refers to the possibility of getting the same results for measures taken on different patients (by the same provider). *Validity* refers to the measure's ability to reflect the "truth" of what the provider normally does. *Feasibility* refers to the possibility of using such data collection on a routine basis for supervision and quality assurance.

This article assesses the value of three methods employed during a nationwide survey on STD case management carried out in health facilities in Malawi in 1994 (*13*): direct observation of provider–patient encounters, interviews with providers, and use of simulated patients. These data were not collected expressly for this type of analysis, but for the management purposes of the Malawi National AIDS Control Programme.

Methods

Data collection

Data on STD case management in Malawi were collected from a total of 39 health facilities using survey instruments adapted from WHO's *Protocol for the assessment of STD case management through health facility survey* (*9*), as well as standards of care provided in the *Malawi standard treatment guidelines* (*14*) and the *Malawi prescriber's companion* (*15*). This survey provided data from direct observation of 150 provider–STD patient encounters (with 54 providers), interviews with 103 providers, and 20 single simulated-patient encounters (one encounter with 20 providers). Provider interviews had been considered especially important with regard to knowledge about correct STD treatment, since Malawi had been experiencing chronic shortages of correct drugs for treatment of most STDs. Inclusion

[a] These effects have been discussed in the framework of research design (*12*). Two possible threats to validity are commonly described. (1) The "Hawthorne effect", which produces a threat to "external" validity, i.e. one's behaviour changes if one knows one is part of an experiment or study. (2) "Instrument reactivity" refers to the effect the instrument (or method) itself has on the subjects in the study (i.e. the fact of being observed).

of simulated patients had been recommended in the 1993 WHO protocol.

The WHO protocol limits STD assessment to two syndromes: urethral discharge and genital ulcers. Direct observations provided data on history-taking, physical examinations, requests for laboratory tests, prescribing or giving treatment, and counselling of patients on partner notification and use of condoms. Parallel items were included on the provider-interview and simulated-patient data-collection forms.

Outpatient departments at three types of health facilities were included in the sample: those under the Ministry of Health, missions of the Christian Health Association of Malawi, and clinics associated with private industry. The choice of facilities for the sample was based on a random selection of those health facilities having at least 10 inpatient beds, or employing more than 300 persons.

Providers selected for observation were those who generally see STD patients, and who would be working full-time in the outpatient department when the data collection team was present. If there was only one provider, he/she was observed. If there were two or more providers, two were randomly selected for observation. Direct observations were conducted for a period of 1–3 days per provider, depending on how long it took to see a minimum of two new patients with urethral discharge or genital ulcers. The average stay at each facility was two days. After the end of the observations, the providers were interviewed.

In addition, simulated urethral-discharge patients (four males recruited from the pool of trained observers) visited a total of 20 providers several days after the providers had been observed. The simulated patients were instructed to abscond if laboratory tests were requested, as repeated laboratory testing could have been painful to the simulated patients and the results would have been negative. Recruitment of symptomatic patients was not considered feasible or acceptable in this setting, and the currently revised WHO protocol no longer suggests the use of simulated patients.

Of the 54 providers observed, 49 were also interviewed. All 20 providers visited by simulated patients were both observed and interviewed. The 49 providers observed and interviewed represented a total of 137 STD patient encounters. However, 16% (8/49) of the providers were observed seeing only one patient. The remaining 41 providers saw 2–9 patients, with an average of 3.3 patients (mode of 2 and median of 3). These numbers are typical of STD patient loads at health facilities in Malawi. This sample is larger than a similar study using the WHO protocol in Jamaica, which included 98 new patients seen by 27 providers (*16*).

Analysis

A central issue for assessing quality is determining which method gives the most valid data about the providers' routine performance. This article compares results from direct observation, provider interview, and simulated-patient data by using two measures: the percentage agreement between the two data sources (the percentage of cases in which both methods agree that the task either happened or did not happen); and the kappa (κ) statistic, which measures the level of agreement that could be expected beyond chance (*17*). A value of $\kappa = 1.0$ means perfect agreement, $\kappa = 0$ means no agreement beyond chance; values of $\kappa = 0.75$–1.0 are considered excellent agreement beyond chance, values of $\kappa = 0.40$–0.74 are considered fair-to-good agreement, and values of $\kappa < 0.40$ are considered poor agreement. The P values presented represent the possibility that the level of agreement was due only to chance and not to a real agreement. Values of κ and P were calculated using EpiCalculator in Epi Info Version 6.01 (*18*).

Comparison of direct observation data with information from a provider interview required creating a single score for each provider based on multiple observation data points (multiple patients). Since providers performed inconsistently and the number of observations per provider varied, the scoring system used was to assign a positive score if the provider was *ever* observed performing the task, and a negative score if the provider was *never* observed performing it.

Before a presentation of the comparative analysis of the three quality assessment methods, difficulties arising in using each of the methods will be discussed. These issues relate to the reliability of the data, the types of answers to be considered, and the feasibility of the method itself.

Difficulties in quality assessment

Direct observation data. Reliable observation data should reflect the provider's usual behaviour and permit a justifiable assessment of performance. However, a series of observations of STD patient–provider encounters for the same provider will encompass a variety of patients with different presentations, and the provider may not approach each patient in the same manner.

Table 1 presents an analysis of the consistency of performance of 41 providers in history-taking, physical examination, treatment, and counselling during STD case management. It shows the percentage of providers: never performing the task on any patient observed; performing the task on all patients

Table 1: **Consistency of performance based on direct observation of 41 providers who saw ≥2 patients**

Observed task	% of providers carrying out the task: Never	Always	Sometimes	% of providers with consistent performance
History:				
Onset of symptoms	0	98	2	98
Contact with high-risk partners	44	12	44	56
Recent new sexual partners	34	20	46	54
Physical examination:				
Carried out	0	0	100	0
Exposed patient fully	31	44	25	75
Retracted foreskin (male patients)[a]	10	57	33	67
Separated labia (female patients)[b]	0	33	67	33
Treatment:				
Gave antibiotic	54	5	41	59
Gave correct drugs/dosages	61	2	37	63
Counselling:				
Advised to finish treatment	29	34	37	63
Mentioned risk of AIDS	44	22	34	66
Advised to use condoms	51	17	32	68
Gave instructions on how to use condom	98	0	2	98
Told to have partner treated	10	29	61	39
Gave or prescribed condoms to patient	0	0	100	0

[a] Thirty providers who saw ≥2 uncircumcised or unexamined men.
[b] Six providers who saw ≥2 female patients.

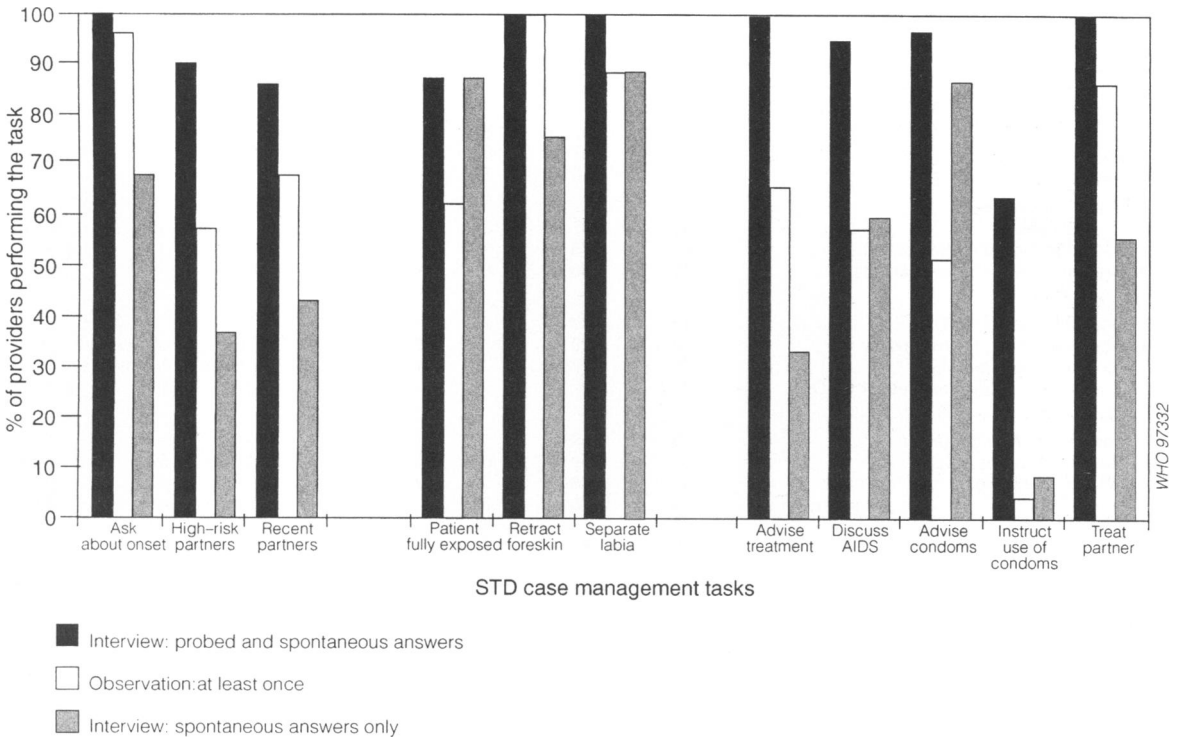
observed; performing the task on only some patients; and either *always* or *never* performing the task on all patients observed. This last indicator, which is a measure of consistency, would be high when the number of providers with consistent behaviour is high. Consistency of less than 100% implies that a single observation might not provide reliable data on the providers' performance.

The data in Table 1 show that many providers do not perform their tasks consistently for each STD patient they see. Providers were especially consistent in asking about the onset of symptoms, and in not giving instructions on the use of condoms. They were particularly inconsistent for many other tasks, such as conducting a physical examination, telling the patient to have the partner treated, and offering or prescribing condoms to the patient. This inconsistency was analysed for possible associations with the sex of the patient or with the chronological order in which patients were seen. Of the 20 providers who saw at least one female patient and at least one male patient, three-quarters did not manage one sex differently from the other on individual tasks. The results for first patients observed compared with those of the ensuing patients showed no trend, i.e., any effect or change in behaviour due to the presence of the observer did not get less with time.

Provider-interview data. The data from interviews with providers raise issues related to under- and over-reporting. The WHO protocol questionnaire starts by asking providers about their management of STD cases using open-ended questions: e.g. for history-taking, "*When a patient comes to you with a complaint of STD, what questions do you ask?*" After the provider had given his/her spontaneous answers, the interviewers are instructed to probe for the specific tasks not yet mentioned. Although probing may prompt providers to reply with what they think the interviewer is looking for (over-reporting), it is often necessary to assist providers in remembering all the things they do when physically confronted with a patient (to compensate for under-reporting).

Comparison of data from provider interviews and direct observations (Fig. 1) revealed both under- and over-reporting. These data are compared using the following measures: percentage of providers reporting that they do the task (probed *and* spontaneous responses); percentage of providers observed carrying out the tasks on at least one patient; and percentage of providers who *spontaneously* reported that they carry out the task. When only spontaneous interview responses were considered, the interview results were often much lower than those found during direct observation. When probed responses were considered together with spontaneous responses, the interview results were consistently higher than those observed. The observed rates for many tasks fall about half way between the spontaneous and probed responses: examples include all history-taking tasks,

Fig. 1. **Comparison of observation and provider-interview data concerning a variety of STD case-management tasks.**



as well as advice on finishing treatment and treating partners. For the tasks in general physical examination, there was better concordance between probed and spontaneous responses, although for full exposure of the patient during examination, both interview data measures were higher than those observed. For some counselling tasks (mentioning AIDS, promoting condoms, and giving instructions on condom use), the direct observation rates were also lower than the spontaneous interview responses, indicating that providers are well aware of what they should do, but still do not even when observed.

Simulated-patient data. These data present some problems on their validity (i.e. their ability to reflect what providers normally do). Our simulated patients did not have any STD symptoms to show, and could be treated in a different way compared with symptomatic STD patients. In fact, 35% (7/20) of simulated patients were sent for laboratory diagnostic tests, as against 19% (12/64) for actual urethral discharge patients observed. This difference is not statistically significant due to the small sample size, but it raises concern about the validity of simulated patient data for assessing the quality of STD case management. In addition, 5 of the 20 encounters ended with the simulated patients absconding (as instructed) when sent for laboratory tests. Consequently, there was no information on treatment and health education from these encounters.

Using simulated patients for quality assessment has an additional limitation: a single simulated patient could visit an individual provider only once, without arousing suspicion. Although it would be possible to organize a series of simulated patients to visit a single health provider, this would cost much more.

Results

Comparisons of direct observation, provider-interview, and simulated patient data

Comparison of direct observation and provider interviews. In Table 2, provider-interview data are

Table 2: **Comparison of direct observation and provider-interview data from 49 providers**

Tasks to be carried out	Observation (%)[a]	Provider interview (probed Y = N) (spontaneous answers only) %	Agree (%)	κ value[b]	Provider interview (probed Y = Y) (probed & spontaneous answers) %	Agree (%)	κ value
History:							
Onset of symptoms	96	67	63	−0.078	100	96	0.000
Contact with high-risk partners	57	37	63	0.292[c]	90	59	0.079
Recent new sexual partners	67	43	55	0.144	86	65	0.078
Physical examination:							
Exposed patient fully	62	87	58	−0.029	87	58	−0.029
Retracted foreskin (male patients)[d]	100	75	—	—[e]	100	—	—[e]
Separated labia (female patients)[d]	88	88	88	*0.433*[c]	100	—	—[e]
Treatment:							
Gave treatment for UD[f]	42	37	79	*0.561*[g]	These questions were not probed in the interview		
Gave correct drugs/dosages for UD[f]	32	29	82	*0.564*[g]			
Gave treatment for GUD[f]	20	27	76	0.321[c]			
Gave correct drugs/dosages for GUD[f]	20	22	76	0.267[c]			
Counselling:							
Advised to finish treatment	65	33	59	0.262[c]	100	—	—[e]
Mentioned risk of AIDS	57	59	61	0.204	94	59	0.067
Advised to use condoms	51	86	57	0.130	96	47	−0.082
Gave instructions on how to use condom	4	8	92	0.295[c]	63	41	0.048
Told to have partner treated	86	55	57	0.075	100	—	—[e]
Gave or prescribed condoms to patient	19	86	31	0.025	Not probed in the interview		

[a] These percentages represent the percentage of providers (observed and interviewed) who were observed carrying out the task at least once.
[b] Kappa values in italics signify fair, good or excellent agreement.
[c] *P* value for $\kappa \leq 0.05$.
[d] These measures are only for those patients who were examined.
[e] Kappa cannot be calculated when one group is 0% or 100%.
[f] For urethral discharge (UD), $n = 38$; for genital ulcer disease (GUD), $n = 45$.
[g] *P* value for $\kappa \leq 0.001$.

compared in two ways with direct-observation data: considering only spontaneous "yes" answers; and considering both spontaneous and probed "yes" answers. These results indicate that, for most STD tasks, the agreement between direct observation and interview data was poor. Values of κ were usually <0.40 and percentage agreements were generally not high. Only treatment of urethral discharge patients and separation of the labia in female patients showed fair to good agreement.

Comparison of simulated-patient data with direct observation. Table 3 compares data on the providers' performance on simulated patients with direct observations of the same providers, and uses two measures of direct observation: the percentage of providers carrying out the task on ≥1 observed patients ("some patients"); and the percentage of providers who carried out the task on all the observed patients ("every patient"). The number of providers here was smaller ($n = 20$) than with provider interviews, so the κ values are lower than for comparison of direct observation and interview data. Percentage agreement, however, is not dependent on sample size. The comparison of simulated patients to the measure "some patients" shows fairly low percentage agreements for all tasks. When compared to the measure "every patient," the κ values were fair-to-good for two tasks: asking about recent new sexual partners, and advice on the use of condoms.

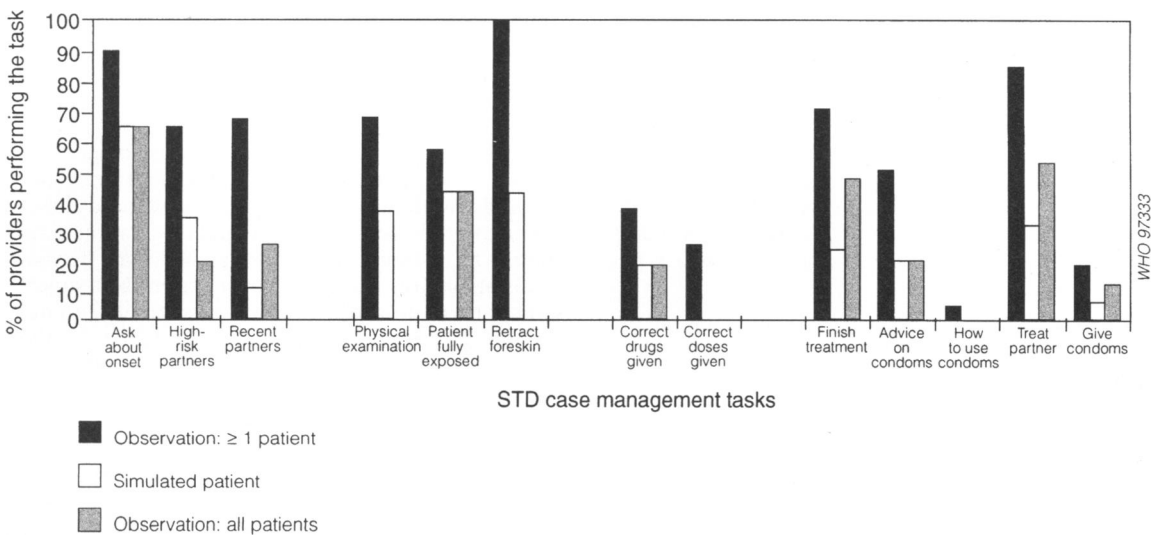
Fig. 2 shows graphically the range of values for the two measures of observation ("some patients" and "every patient"), compared with the simulated-patient data. These data reflect differences in conclusions that might be drawn from the various measures and data sources. When compared to the observation measure "every patient", simulated-patient data generate a picture of similar or lower performance for most tasks. The exceptions are as follows: asking about contact with high-risk partners, and conducting a physical examination. The fact that these tasks were more likely to be performed on a single simulated patient than under direct observation for "every patient" implies that health workers normally do these tasks only on selected patients. Few providers do them every time, even when observed.

Table 3: **Comparison of direct observation data from 20 providers seeing simulated patients**

Tasks to be carried out	Simulated patients (%)	Direct observation (carried out on some patients)			Direct observation (carried out on every patient)		
		%[a]	Agree (%)	κ value	%[b]	Agree (%)	κ value
History:							
Onset of symptoms	65	90	55	−0.182	65	65	0.205
Contact with high-risk partners	35	65	60	0.266	20	55	−0.098
Recent new sexual partners	11	68	56	0.158	26	84	*0.496*[c]
Physical examination:							
Carried out any physical examination	37	68	58	0.232	50	63	0.273
Exposed patient fully[d]	43	57	43	−0.120	43	57	0.125
Retracted foreskin (male patients)[d]	43	100	—	—	100	—	—
Treatment:							
Gave correct drugs (syndromic management) for UD[e]	19	38	69	0.259	19	75	0.180
Gave correct drugs/dosages for UD[e]	0	25	—	—	0	—	—
Counselling:							
Advised to finish treatment	24	71	53	0.227	47	65	0.271
Advised to use condoms	20	50	50	0.000	20	80	*0.375*[c]
Gave instructions on how to use condom	0	5	—	—	0	—	—
Told to have partner treated	32	84	47	0.159	53	47	−0.033
Gave or prescribed condoms to patient	6	19	75	−0.103	13	81	−0.091

[a] This represents the percentage of providers who were observed carrying out the task on at least one patient.
[b] This represents the percentage of providers who were observed carrying out the task on all patients they saw.
[c] κ values in italics signify fair agreement ($P \leq 0.05$).
[d] For patients who received any physical examination.
[e] For treatment of urethral discharge (UD), $n = 16$.

Fig. 2. **Comparison of observation and simulated-patient data concerning a variety of STD case-management tasks.**



The comparison between what providers did when observed and what they did when confronted with a simulated patient (not knowing that they were being observed) indicates that providers were most probably improving their normal level of performance when observed. The percentage of providers carrying out the task on "some patients" when observed is about twice that of providers carrying out the task on the simulated patient, while the percentage of providers carrying out the task on "every patient" when observed is either similar or about one third higher.

Discussion

Our study has investigated the validity and reliability of data used for assessing the quality of STD case management in outpatient settings. The results suggest that the providers' performances during history-taking, physical examinations, and counselling were better when they were observed than under normal circumstances. Differences were even larger between observed case management behaviour and what the providers say they do when interviewed, which casts doubt on the ability of provider interviews to furnish valid data about routine performance. One surprising result was the good level of agreement between observation and provider-interview data with respect to STD treatment, because it had been felt that drug shortages would reduce the value of the direct observation data in terms of the provider's knowledge.

These results on the validity, reliability, and feasibility of various data sources provide some important cautions when using the data to assess the quality of health care. The data presented in this article permit the following conclusions.

Validity

- The performance of providers in the case management of STD is more complete when they are observed than when they do not know they are being assessed, even after a period of 1–3 days of an observer being present. This implies that the providers do not "forget" the observer's presence over a period of time.

- Providers, when interviewed about their management of STDs, say they do more than they actually do, even when interviewed *after* the observer/interviewer has been present observing them in the consultation room for a period of 1–3 days.

- Performance rates for STD case-management tasks measured by performance on at least one patient from direct observation data generally lie between the rates for spontaneous and probed provider-interview responses.

- Performance rates for STD case-management tasks measured by performance on every patient from direct observation data appear to be similar to or slightly higher than the unobtrusive assessment method of using simulated patients.

Reliability

- Providers do not treat all STD patients in the same way, which suggests that a single observation or a single simulated patient's visit would not be sufficient to draw reliable conclusions about their performance. Inconsistency cannot be explained by their getting used to the observer's presence, since it does not appear to be related to the order in which the patients are seen. Other factors, perhaps individual patient's characteristics, must explain the inconsistency of a provider's performance. In the case of STDs, providers are most likely to be inconsistent on the tasks which are considered essential to good case management — physical examination, partner notification, and distribution of condoms.

Feasibility

- Although simulated STD patients would appear to provide the most accurate results about what providers normally do, this approach is very resource-intensive and it is difficult to ensure more than a single encounter per provider. In addition, it is difficult to adequately simulate all the physical signs of STDs.

In countries where patients' records are not used or they provide very little information, the options for obtaining STD quality-assessment data are limited. Simulated-patient data are theoretically the best measure of a provider's normal performance, but their use presents problems of cost and accuracy compared with real STD patients. Use of simulated patients in quality assessment of STD case management should be restricted to occasional evaluation of possible discrepancies or over-reporting from other methods. Without adequate patients' records, direct observation is the best option. Observation data should be assumed to reflect higher than normal levels of quality of care. However, if done by a supervisor, observation provides an opportunity for on-the-job training. Owing to inconsistency in the performance of tasks, providers need to be observed more than once in order to get a reliable picture of their performance. Provider interviews should be viewed with caution, since they do not necessarily measure actual practice. However, interviews have a role to play in supervision because they provide a picture of what health workers know, and how they

might manage types of cases that were not observed. Exit interviews, which were not used in this study, provide another possible non-intrusive method for obtaining information about a provider's performance, but comparative research on the reliability and validity of this method is needed.

This study provides new information about how well various methods are able to assess the quality of STD case management in developing countries. The conclusions offer some guidelines to supervisors and programme managers in choosing how to evaluate a health worker's performance and what weight to give to the results from the various quality assessment methods. In sum, direct observation is the best general method, bearing in mind that it exaggerates actual performance.

Acknowledgements

Support for this project was provided by Family Health International with funds from the U.S. Agency for International Development (USAID), under cooperative agreement No. 623-0238-A-00-4-31-00. The project provided support to the Malawi AIDS Control Programme under the Ministry of Health and Population in the areas of AIDS prevention and STDs. Views expressed in this article do not necessarily represent those of the funding agencies.

Résumé

Qualité de la prise en charge des cas de maladies sexuellement transmissibles: comparaison des méthodes d'évaluation de la qualité des prestations des dispensateurs de soins

L'une des tâches les plus difficiles de l'évaluation de la qualité consiste à juger le processus de prestation de soins de santé. Le présent article examine la fiabilité et la validité de trois méthodes d'évaluation de la qualité de la prise en charge des cas de maladies sexuellement transmissibles (MST): observation directe des rencontres patient-dispensateur de soins, entretien avec les dispensateurs de soins, et recueil de données par des malades simulés. Les données analysées ici ont été obtenues dans le cadre d'une enquête réalisée en 1994 dans un établissement de soins spécialisé dans les MST au Malawi, suivant le protocole d'enquête OMS (*Protocol for the assessment of STD case management through health facility survey, version 5*). Au total, 137 rencontres patient-dispensateur de soins ont été observées. Elles étaient assurées par 49 dispensateurs de soins, lesquels ont aussi été interrogés. Vingt d'entre eux ont en outre reçu la visite d'un malade simulé se plaignant d'écoulement urétral. Les résultats de ces observations ont été comparés (test kappa) afin de vérifier la concordance entre *a)* l'observation directe et les données de l'entretien avec les dispensateurs de soins, et *b)* l'observation directe et les données fournies par le malade simulé.

Le taux de concordance était en général faible entre l'observation et les données fournies par le dispensateur de soins, que l'analyse ait porté sur les réponses spontanées ou sur les réponses au choix. Les réponses spontanées donnaient souvent des valeurs plus faibles que l'observation directe, et les réponses au choix des valeurs plus élevées. Les pourcentages de concordance entre les données fournies par le malade simulé et les données de l'observation directe étaient souvent plus élevés. La qualité de la prise en charge n'était cependant pas régulière d'un patient à l'autre et les données du malade simulé ne concordaient avec les données d'observation que pour les prestations que le dispensateur de soins fournissait pour chaque patient.

L'analyse de ces données indique que les dispensateurs de soins assurent une prise en charge plus complète des cas de MST lorsqu'ils se savent observés que lorsqu'ils ignorent qu'ils sont évalués (recueil des données par un malade simulé). Le travail de prise en charge déclaré lors de l'entretien, même lorsque celui-ci suivait une observation directe, était également supérieur à celui effectivement réalisé. De plus, il est apparu que la prise en charge n'était pas régulière d'un patient à l'autre, d'où un manque de fiabilité des données résultant d'une observation unique ou d'une seule consultation d'un malade simulé pour en tirer des conclusions quant à la qualité du travail.

Dans les pays où les dossiers individuels de malades n'existent pas ou ne contiennent que des données très sommaires, le choix des méthodes permettant d'obtenir des données d'évaluation de la qualité est limité. Si les données recueillies par un malade simulé peuvent refléter correctement la pratique normale d'un dispensateur de soins, il semble que l'observation directe constitue la meilleure option pour l'évaluation de la qualité de la prise en charge des MST. Le recueil de données par un malade simulé est difficile à mettre en pratique et ne présente qu'un intérêt limité en raison de la variabilité des prestations d'un patient à l'autre et de la nécessité de multiplier les visites pour dresser un tableau exact de la situation. Les résultats obtenus par observation directe doivent cependant être considérés comme représentant une pratique de qualité supérieure à la pratique courante. Les données des entretiens avec les dispensateurs de soins doivent être interprétées avec prudence car

elles risquent de refléter davantage les connaissances théoriques de l'agent que la qualité de ses prestations.

References

1. **Forsberg BC, Barros FC, Victora CG.** Developing countries need more quality assurance: how health facility surveys can contribute. *Health policy and planning*, 1992, **7**: 193–196.
2. **Wishik SM, Bertrand JT.** Field supervision for quality control of local family planning workers in developing countries. *International quarterly of community health education*, 1990–91, **11**: 111–122.
3. **Begum S, Salahuddin AKM.** Study on quality assessment of neo-natal tetanus cases admitted in the Infectious Diseases Hospital, Dhaka. *Journal of Academy of Hospital Administration*, 1991, **3**: 25–28.
4. **Nicholas D, Heiby J, Hatzell T.** The quality assurance project: introducing quality improvement to primary health care services in developing countries. *Quality assurance in health care*, 1991, **3**: 47–165.
5. **Björck M, Johansson R, Kanji N.** Improving the quality of primary care services in Angola. *Health policy and planning*, 1992, **7**: 290–295.
6. **Kim Y et al.** Improving the quality of service delivery in Nigeria. *Studies in family planning*, 1992, **23:** 118–127.
7. **Bryce J et al.** Assessing the quality of facility-based child survival services. *Health policy and planning*, 1992, **7**: 155–163.
8. **Mertens T et al.** Prevention indicators for evaluating the progress of national AIDS programmes. *AIDS*, 1994, **8**: 1359–1369.
9. *Protocol for the assessment of STD case management through health facility survey (measurement of HIV/STD PPIs 6 and 7).* Unpublished document WHO/GPA/CNP/EVA/HFSURV/TM, 5th version, August 1993. A revised version has since been issued: unpublished document WHO/GPA/TCO/SEF/94.1, Section 5, 1994. (Both these documents are available upon request from: Joint United Nations Programme on HIV/AIDS, 1211 Geneva 27, Switzerland).
10. **PRISM Group.** *Peru country study: final report.* (Volumes 1–5). Columbia, MD, PRISM, 1989.
11. **Donabedian A.** The quality of care: how can it be assessed? *Journal of the American Medical Association*, 1988, **260**: 1743–1748.
12. **Spector P.** *Research designs.*. Beverly Hills, CA, Sage Publications, 1982.
13. **Chilongozi D et al.** Sexually transmitted diseases: a survey of case management in Malawi. *International journal of STD and AIDS*, 1996, **7**: 269–275.
14. *The Malawi standard treatment guidelines,* 2nd edit. Malawi, Ministry of Health, Essential Drugs Programme, 1993.
15. *The Malawi prescriber's companion*, 1st edit. Malawi, Ministry of Health, Essential Drugs Programme, 1993.
16. **Bryce J et al.** Quality of sexually transmitted disease services in Jamaica: evaluation of a clinic-based approach. *Bulletin of the World Health Organization*, 1994, **72**: 239–247.
17. **Fleiss J.** *Statistical methods for rates and proportion*, 2nd edit. New York, John Wiley & Sons, 1981.
18. **Dean AG et al.** *Epi Info, Version 6.01. A word processing, database, and statistics program for public health.* Atlanta, GA, Centers for Disease Control and Prevention, 1994.