International Journal for Quality in Health Care 2006; **Volume 18, Number 4:** pp. 299–305
Advance Access Publication: 4 May 2006
10.1093/intqhc/mzl009

The influence of observation and setting on community health workers' practices

S. Y. ROWE[1]*, M. A. OLEWE[2†], D. G. KLEINBAUM[1], J. E. MCGOWAN JR[1], D. A. MCFARLAND[3], R. ROCHAT[3] AND M. S. DEMING[4]

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA, [2]CARE Kenya, Siaya, Kenya, [3]Department of Global Health, Rollins School of Public Health, Emory University, and [4]Division of Parasitic Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA

Abstract

Objective. To determine whether results from an evaluation that involved observation of community health workers while they performed patient consultations in a hospital reflected normal everyday practices.

Design. Comparison of two samples of ill-child consultations: (i) consultations performed during an evaluation in which we observed community health workers in a hospital in-patient and outpatient department from February to March 2001 and (ii) consultations performed under no observation in villages and documented in clinical registers within the 90 days before the hospital evaluation.

Setting. Siaya District Hospital and villages in Kenya.

Study participants. Community health workers.

Main outcome measure. Treatment error indicator, defined as the percentage of consultations where at least one recommended treatment (where recommended treatments were those that were indicated based on community health worker assessments of the child's condition) was not prescribed.

Results. We analyzed data on 1132 consultations (372 from the hospital evaluation and 760 from the community) performed by 103 community health workers. For all types of consultations combined, the difference between treatment error indicators (hospital minus community) was –16.4 [95% confidence interval (CI): –25.6, –7.1].

Conclusions. We found that community health workers made treatment errors less frequently when they were observed in a hospital in-patient or outpatient department than when they were not observed in the community. Evaluations that involve the observation of community health workers in a hospital setting might overestimate the quality of care that they normally give in their villages.

Keywords: case management, child health, community health worker, Hawthorne effect, Kenya, treatment error

Community health worker programs are considered a major component of the strategy to reduce the burden of childhood illness in developing countries [1,2]. However, for these programs to be effective, community health workers must provide high-quality care. To reach and maintain high-quality care, programs need information on gaps in performance (where performance means adherence to clinical guidelines). One method to help identify these gaps is to observe community health workers while they perform patient consultations. Because of logistical difficulties in observing community health workers while they perform consultations in their villages (e.g. on a typical day, patient volume might be low; patients might be treated at any time of day), programs often evaluate community health workers in hospital settings.

Programs often need information on how well community health workers perform under usual conditions. To validly assess this, performance assessments should reflect normal everyday practice [3]. Evidence suggests that when health care providers, such as paramedics or facility-based health workers, are aware they are being observed or their performance is being assessed, they alter their practices (i.e. the phenomenon known as the 'Hawthorne effect') [4–6]. Determinants of community health worker performance may be different from those for other types of health workers [7]; analogously, community health workers may react differently when they

Address reprint requests to S. Y. Rowe, Division of Parasitic Diseases, Centers for Disease Control and Prevention, 4770 Buford Highway, Mailstop F-22, Atlanta, GA 30341, USA. E-mail: say9@cdc.gov
*Present address: Division of Parasitic Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA.
†Present address: Inter Country Programme for Malaria, World Health Organization, Kampala, Uganda.

are being observed or are asked to work in a non-community setting. To our knowledge, no data exist on how the presence of an observer or being evaluated in a non-community setting affects community health workers in developing countries.

A community health worker program, initiated in 1995 by the non-governmental organization, CARE, in Siaya District, Kenya, trained volunteers to assess, diagnose, and treat children <5 years old according to the CARE Management of the Sick Child (MSC) guidelines, a simplified version of the WHO/UNICEF Integrated Management of Childhood Illness guidelines [8]. To identify weaknesses in clinical skills, this program conducted evaluations where community health workers were observed as they managed ill children at Siaya District Hospital [9]. To interpret the results from these evaluations, one must know whether the assessment in the hospital setting reflects normal everyday practices.

We conducted this study to determine whether treatment practices during consultations in a hospital where community health workers were observed were different from practices during non-observed consultations in villages.

Methods

Study population

We studied two samples of ill-child consultations performed by the same community health workers: (i) consultations performed under observation during a 1-day evaluation at Siaya District Hospital during February–March 2001 [9] and (ii) consultations performed under no observation in villages during the 90 days before the hospital evaluation (November 2000–March 2001).

Data sources

Hospital evaluation data. We asked randomly selected community health workers to examine ill children in the outpatient and in-patient department of Siaya District Hospital and to document on standardized forms the treatments (including referral) they would recommend for these patients. The format of the recording forms was identical to that of the community health workers' clinical registers, which are books containing records of consultations performed in villages. Registers contain printed columns that prompt the community health worker to record the presence or absence of specific symptoms, classifications (i.e. diagnoses), and treatments; thus, the register is also a job aid. A trained observer documented each outpatient department consultation. In the in-patient department, each consultation was not directly observed; however, evaluation staff was present in the in-patient department to supervise data collection.

Village-based clinical register data. We collected clinical registers of community health workers who attended the hospital evaluation and abstracted data from registers onto a standardized form.

Ethics review

The human subjects research Institutional Review Boards at the US Centers for Disease Control and Prevention (CDC) and Emory University approved the study protocol.

Definitions and population for analysis

We analyzed three indicators of treatment errors: (i) the proportion of all patients who were not prescribed all recommended treatments, (ii) the proportion of all patients with ≥1 severe classification who were not prescribed all recommended treatments for all severe classifications, and (iii) the proportion of all patients with ≥1 moderate classification who were not prescribed all recommended treatments for all moderate classifications. Severe classifications for patients <2 months old included very severe disease, severe pneumonia, dehydration, bloody diarrhea, and persistent diarrhea. Severe classifications for patients 2–59 months old included very severe disease, severe pneumonia, persistent cough, severe dehydration, bloody diarrhea, persistent diarrhea, and complicated measles. Moderate classifications (for patients 2–59 months old only) included pneumonia, some dehydration, measles, and malaria.

Because a 'gold standard' clinician did not re-examine the consultations recorded in registers, the true signs and symptoms are unknown for patients treated in the village. However, registers contain data on community health workers' findings of signs and symptoms. By applying the definitions of classifications of the MSC guidelines to these findings, we identified for each patient one or more 'internally consistent' classifications that should have been identified. We defined recommended treatments as those that should have been prescribed based on community health worker assessments. In effect, the adherence score measured the community health worker's ability to process clinical information to choose recommended treatments.

From the register data, we analyzed consultations that (i) community health workers recorded in registers during the 90 days before their hospital evaluation date, (ii) involved an ill child whose age was 0–59 months, (iii) were initial consultations, (iv) occurred in the village, (v) were recorded by the community health worker, (vi) involved an ill child for whom the documentation of community health worker assessments was sufficient to allow the determination of recommended treatments, and (vii) involved an ill child who required ≥1 treatment according to these assessments.

From the hospital evaluation data, we analyzed consultations that involved (i) an ill child whose age was 0–59 months, (ii) an ill child for whom the documentation of community health worker assessments was sufficient to allow the determination of recommended treatments, and (iii) an ill child who required ≥1 treatment according to these assessments.

Data analysis

To test statistical hypotheses regarding the difference between hospital evaluation and register results, we constructed a 95%

confidence interval (CI) for the estimate of the difference in treatment error indicators. We concluded that treatment error indicators were different if the 95% CI did not include zero.

We analyzed data for three groups of patients: (i) all patients, (ii) patients with ≥1 severe internally consistent classification, and (iii) patients with ≥1 moderate but no severe internally consistent classification. To investigate the influence of different types of observation or hospital setting, we also compared treatment error indicators from register data with those for consultations performed in the hospital outpatient and in-patient departments separately.

We calculated point estimates of treatment error indicators and their standard errors using SUDAAN software [10], which accounts for the correlation between treatment errors for patients seen by the same community health worker, sampling weights, and a finite population correction (FPC) factor. The sampling weights adjusted for the following aspects of the study design: we studied only a sample of all active community health workers; community health worker participation rates varied by geographic areas of Siaya District; not all participants in the hospital evaluation actually examined four patients (we planned for each community health worker to examine four patients; however, the number of patients examined depended on the patient volume at the hospital on the evaluation day); not all community health workers performed the same number of consultations in their communities during the 90 days before the hospital evaluation; and the proportion of consultations with severely ill children was higher in the hospital evaluation than in the register data. We adjusted variance estimates by an FPC factor to account for the fact that a large proportion (45%) of the active community health worker population in Siaya District participated in our study, and we considered the registers to contain a 'census' of all consultations community health workers performed in villages. For surveys with a large sampling fraction, adjusting results by an FPC factor can improve the precision of estimates [11].

Results

We abstracted data on 760 consultations from registers and 372 consultations from the hospital evaluation, all performed by the same 103 community health workers. The median age of these workers was 39.0 years, and most were females, had completed at least seven levels of education (equivalent to an elementary school graduate in the United States), were of Luo ethnicity, and had been working as a community health worker since 1997.

Table 1 summarizes the completeness of documentation of assessment tasks in both data sources. We assumed that an unfilled column in the register or hospital evaluation recording form for an assessment task meant it was not performed; thus, data for that assessment were incomplete. For most assessment tasks, the completeness of the documentation was ≥85%, and documentation in registers was comparable to or slightly better than that in the hospital data. The assessment task of counting a respiratory rate had a moderate level of incompleteness in both data sources.

From the register and hospital data, we excluded 90 (11.8%) and 55 (14.8%) consultations, respectively, because documentation of assessments was insufficient to allow the determination of recommended treatments or the patient did not require treatment. Thus, we used 670 consultations from registers and 317 consultations from the hospital evaluation in the analyses.

Among all patients, the difference in treatment error indicators (hospital minus register) was significantly less than zero (difference estimate: –16.4; 95% CI: –25.6, –7.1) (Table 2, comparison 1). We saw similar results when we divided patients by illness severity [for patients with ≥1 severe internally consistent classification, difference estimate: –31.0; 95% CI: –52.9, –9.1 (Table 2, comparison 2); for patients with ≥1 moderate but no severe internally consistent classification, difference estimate: –18.9; 95% CI: –28.6, –9.1 (Table 2, comparison 3)]. When we compared treatment error indicators from register data with those for consultations performed in the hospital outpatient (Table 2, comparisons 4, 5, and 6) or in-patient department (Table 2, comparisons 7, 8, and 9), we found that differences generally were significantly less than zero.

Discussion

We found that treatment error indicators in the hospital evaluation data were lower than those in the register data. This implies that community health workers made treatment errors less frequently in the hospital, both under direct observation in the outpatient setting and when they were not observed but knew they were being studied in the in-patient setting.

The ability of community health workers in this study to process clinical information during consultations in their villages was comparable to or somewhat worse than that of community health workers in other studies. For example, we found that among all patients who required treatment according to community health workers' assessments recorded in registers, 51% of patients were not prescribed all recommended treatments. In a study of volunteer health workers in West Java, all responded that they gave oral rehydration solution to treat the most recent ill child with diarrhea they encountered, but 85% failed to refer patients whom they perceived to be severely ill [12]. In a survey in Nigeria, community health assistants who were not observed as they examined children did not give appropriate treatment to 29% of children whom they identified as having moderate acute respiratory infections [13].

Our findings are consistent with two explanations. Firstly, as suggested in other research [4,6,14], our findings suggest that health workers perform better than usual when they know they are being studied. Our finding of lower treatment error indicators in the in-patient setting suggests that community health workers, even without direct observation, alter their practices when they know they are being studied. Furthermore, we compared data from registers and the 2001 hospital evaluation in which each community health worker performed only a few consultations. Additional research is

Table I Completeness[1] of community health workers' documentation of assessment tasks performed for initial consultations with ill children aged 0–59 months, clinical register versus hospital-based evaluation data, Siaya District, Kenya, 2001

Assessment task	Clinical register		Hospital-based evaluation	
	Number of patients[2]	Percent with complete data	Number of patients[3]	Percent with complete data
For all patients				
Unable to drink or breastfeed	760	95.1	372	96.8
Vomiting everything	760	95.1	372	94.9
Convulsions	760	93.7	372	92.7
Lethargy	760	92.9	372	90.6
Cough or difficulty breathing	760	100	372	100
Diarrhea	760	100	372	100
Fever	760	100	372	99.7
For patients with cough or difficulty breathing				
Duration of cough	259	88.0	218	89.0
60 seconds respiratory rate	259	82.6	218	78.4
60 seconds respiratory rate (repeat)	259	79.5	218	75.7
Chest indrawing	259	93.8	218	88.1
For patients with diarrhea				
Duration of diarrhea	156	84.0	141	79.4
Diarrhea for ≥14 days	156	91.7	141	86.5
Blood in stool	156	93.6	141	85.8
Restless/irritable	156	90.4	141	83.7
Drinks eagerly	156	90.4	141	82.2
Slow skin pinch	156	89.1	141	83.7
For patients with fever				
Duration of fever	658	90.7	301	81.4
Generalized rash[4]	643	92.7	285	88.4
Red eyes or runny nose[4]	643	92.2	285	84.2
Pus in eyes[4]	643	92.2	285	84.6
Deep mouth sores[4]	643	92.7	285	84.9

[1]We assumed an unfilled column in the clinical register for an assessment task meant that the community health worker did not perform the assessment; thus, data for that assessment were incomplete.
[2]Consultations performed in the community by 103 community health workers during the 90 days before they participated in a hospital-based evaluation.
[3]Consultations performed during a hospital-based evaluation by 103 community health workers who performed at least one initial ill-child consultation in their communities during the 90 days before they participated in the hospital-based evaluation.
[4]Assessments only performed for patients 2–59 months old.

needed to assess whether the effect of observation (either direct or not direct, but one is aware of being studied) changes with repeated observations during a single evaluation (as suggested in a study of the influence of video observation on physicians' practices [15]) or with repeated evaluations over time. These results would have practical implications as programs may assess community health worker performance frequently over time as part of quality improvement efforts.

Secondly, our findings are consistent with the explanation that hospital setting might have altered community health workers' behaviors. For example, if community health workers perceived that the hospital drug supply was different (e.g. with respect to quality, quantity, or cost) from that in their own communities, this might have influenced prescribing practices. However, adequacy of medicine supplies might have had a limited effect because community health workers were trained to record all treatments they would recommend, regardless of the availability of medicine supplies. Additionally, conditions in the hospital (e.g. lighting in the consultation room and interaction with patients' caretakers) might have promoted better performance. Another factor that might explain our findings is the difference in the distribution of illness types among patients (i.e. case mix) between the two settings. We attempted to account for case mix by conducting weighted analyses, where the weight adjusted for the higher proportion of consultations with severely ill children in the hospital. Furthermore, our findings are less likely to be the result of differences in case mix because analyses that were restricted to specific subgroups of patients (i.e. patients who had ≥1 severe classification or patients who had ≥1 moderate

Table 2 Comparison of treatment practices of community health workers for initial consultations with ill children aged 0–59 months, hospital-based evaluation versus clinical register data, Siaya District, Kenya, 2001

Disease classification[3]	Treatment error indicator[1]				Arithmetic difference between treatment error indicators[2] (hospital minus register)	
	Hospital evaluation		Clinical register			
	N	Estimate (SE[4])	*N*	Estimate (SE)	Estimate	95% CI[5]
For community health workers who performed ≥1 consultation involving a patient with the specified disease classification during the hospital-based evaluation and in the community during the 90 days before the hospital-based evaluation						
(1) All classifications	317	34.3 (3.7)	670	50.7 (2.9)	–16.4	–25.6, –7.1
(2) All severe classifications	38	47.3 (8.7)	39	78.4 (7.0)	–31.0	–52.9, –9.1
(3) All moderate classifications	186	30.6 (3.7)	542	49.5 (3.3)	–18.9	–28.6, –9.1
For community health workers who performed ≥1 consultation involving a patient with the specified disease classification in the outpatient department during the hospital-based evaluation and in the community during the 90 days before the hospital-based evaluation						
(4) All classifications	152	33.0 (4.8)	649	50.4 (2.9)	–17.4	–28.3, –6.4
(5) All severe classifications	16	45.8 (11.2)	31	81.8 (6.0)	–36.0	–61.0, –11.0
(6) All moderate classifications	108	29.3 (4.8)	481	48.4 (3.6)	–19.1	–30.7, –7.4
For community health workers who performed ≥1 consultation involving a patient with the specified disease classification in the in-patient department during the hospital-based evaluation and in the community during the 90 days before the hospital-based evaluation						
(7) All classifications	165	35.8 (5.2)	634	50.6 (3.0)	–14.7	–26.5, –3.0
(8) All severe classifications	22	49.2 (13.7)	18	80.3 (10.3)	–31.1	–64.7, 2.4
(9) All moderate classifications	78	32.8 (5.4)	375	49.5 (4.1)	–16.7	–29.9, –3.4

CI, confidence interval.

[1]Treatment error indicator for all classifications is the percentage of patients for whom the community health worker did not prescribe all recommended treatments (including referral) for all internally consistent classifications. Treatment indicator for all severe classifications is the percentage of patients with ≥1 severe internally consistent classification for whom the community health worker did not prescribe all recommended treatments for the severe classifications. Treatment indicator for all moderate classifications is the percentage of patients with ≥1 moderate but no severe internally consistent classification for whom the community health worker did not prescribe all recommended treatments for the moderate classifications. 'Internally consistent' classifications were classifications that community health workers should identify based on the processing of their own assessment findings according to the CARE guidelines.

[2]Point estimate of treatment error indicator for hospital-based evaluation data minus point estimate of treatment error indicator for clinical register data.

[3]Severe classifications for patients <2 months old included very severe disease, severe pneumonia, dehydration, bloody diarrhea, and persistent diarrhea. Severe classifications for patients 2–59 months old included very severe disease, severe pneumonia, persistent cough, severe dehydration, bloody diarrhea, persistent diarrhea, and complicated measles. Moderate classifications (for patients 2–59 months old only) included pneumonia, some dehydration, measles, and malaria.

[4]Standard error.

[5]95% CI for the difference in treatment error indicators between the hospital-based evaluation and the clinical register data. The treatment error indicators are significantly different if (point estimate of treatment error indicator for hospital-based evalution data – point estimate of treatment error indicator for clinical register data) $\pm$ 1.96 $\left(SE^2_{\text{hospital}} + SE^2_{\text{clinical register}}\right)^{0.5}$ excludes 0 percentage points.

classification but no severe classification) revealed similar results.

Our study design had specific limitations. Firstly, the results might have been biased if failure to document one assessment task or to record an entire consultation was related to treatment errors or data source; however, similarity in the frequency of incomplete documentation in each data source would limit the effect of this bias. Secondly, the definition of treatment errors in this study did not consider drug dosages, whether caretakers (e.g. patients' mothers) understood how to administer treatment at home, or prescription of unnecessary treatments. Thirdly, we were unable to assess the influence of the hospital setting independent of observation because all observation occurred in a hospital. To test the hypothesis that observation leads to better performance, we need to collect data on observed consultations and consultations from the same setting in which community health workers are unaware that they are being studied.

As previously stated in the methods, we evaluated community health workers' abilities to process clinical information to choose treatments. Although this process indicator is important to evaluate, another important indicator is the 'true' quality of treatment (i.e. errors based on the treatment recommended according to a gold standard clinician's assessments).

If we understand the extent to which community health worker assessments during consultations in the village are correct, we theoretically could adjust treatment quality measurements from register data and estimate the true quality of care in villages [16]. However, given the current evaluation methods, measurement of these adjustment factors is problematic.

One method, commonly considered the 'best' method, to determine the accuracy of health worker assessments is the direct observation of consultations at a health facility along with the re-examination of the patient [17–20]. This method has weaknesses, which may have applied to the hospital evaluation in Siaya; it is highly resource-intensive, and the patient's condition might change between the time of the health worker's examination and the re-examination. Furthermore, as suggested by this study, the presence of an observer might influence health workers' practices. Other reasons exist why this method might not be appropriate for evaluating community health workers. Firstly, community health workers and their patients usually live in the same communities; thus, community health workers are often familiar with their patients. In contrast, health facility health workers often encounter patients with whom they are not familiar. Secondly, the hospital environment is quite different from the village setting in which community health workers normally perform consultations.

Evaluations have been done, which involve direct observation of community health workers in villages and re-examination of patients [21]. However, such evaluations have drawbacks. Not only do the weaknesses listed for the health facility-based studies apply but also village-based evaluations can be impractical if daily patient volume is low (e.g. in this study, on average, community health workers performed only 2.5 consultations in their villages per month based on the number of initial consultations recorded in registers during the 90 days before the hospital evaluation).

However, despite its weaknesses, the method of direct observation with re-examination of the patient is useful. The strength of this method is that patients' true signs and symptoms are known. Additionally, this method can be used to further investigate patterns in quality of care found in studies involving no observation. For instance, in this study, hospital evaluation results supported the finding from register data that community health workers' abilities to choose treatments for severely ill children were considerably worse than for moderately ill children. Furthermore, if an evaluation using direct observation found that the quality of care was poor, any bias introduced by the presence of an observer would not likely change the interpretation of this finding. For example, only about 40% of hospital outpatients were prescribed all treatments that were required according to gold standard clinician assessments [7]. Even if this were an overestimate, the conclusion would be that the quality of treatments given by the community health workers was poor. We acknowledge that if an evaluation using direct observation finds fair or good quality of care, interpretation of the findings is more complicated. The results might reflect the true quality of care, or the results might be overestimates (and thus, the true quality of care is unknown).

In conclusion, we found that community health worker treatment practices during the hospital evaluation did not reflect those in the village setting. When interpreting results from evaluations where community health workers are observed in a hospital in-patient or outpatient setting, programs should be aware that results might be overestimates of the quality of care usually given. We believe this study emphasizes the need to (i) more thoroughly examine how observation or evaluation setting influences community health worker practices and how this influence might change over time and (ii) develop adjustment methods to estimate the true quality of care usually given by community health workers using data collected from a non-observed setting.

Acknowledgements

The authors thank the project staff of the *Community Initiatives for Child Survival in Siaya* for their assistance in conducting the hospital evaluation and abstracting data from clinical registers. The authors also thank Dr Jane Kelly for her assistance with data collection and her editorial contributions to this manuscript. This research was funded by a grant from the RW Woodruff Foundation to CARE and the CDC Foundation as part of the CDC-CARE Health Initiative.

References

1. Walt G, ed. *Community Health Workers in National Programmes: Just Another Pair of Hands?* Philadephia: Open University Press, 1990.
2. Winch PJ, Gilroy KE, Wolfheim C *et al.* Intervention models for the management of children with signs of pneumonia or malaria by community health workers. *Health Policy Plan* 2005; **20:** 199–212.
3. Spies TH, Mokkink HGA, De Vrie Robbé PF, Grol RPTM. Which data source in clinical performance assessment? A pilot study comparing self-recording with patient records and observation. *Int J Qual Health Care* 2004; **16:** 65–72.
4. Campbell JP, Maxey VA, Watson WA. Hawthorne effect: implications for prehospital research. *Ann Emerg Med* 1995; **26:** 590–594.
5. Casalino LP. The unintended consequences of measuring quality on the quality of medical care. *N Engl J Med* 1999; **341:** 1147–1150.
6. Rowe AK, Lama M, Onikpo F, Deming MS. Health worker perceptions of how being observed influences their practices during consultations with ill children. *Trop Doct* 2002; **22:** 166–167.
7. Rowe SY, Kelly JM, Olewe MA *et al.* Effect of multiple interventions on community health workers' adherence to clinical guidelines in Siaya District, Kenya. *Trans R Soc Trop Med Hyg* (in press).
8. Gove S. Integrated management of childhood illness by outpatient health workers: technical basis and overview. *Bull World Health Organ* 1997; **75** (**suppl. 1**): 7–24.
9. Kelly JM, Osamba B, Garg RM *et al.* Community health worker performance in the management of multiple childhood illnesses: Siaya District, Kenya, 1997–2001. *Am J Public Health* 2001; **91:** 1617–1624.

10. Research Triangle Institute. *SUDAAN User's Manual, Release 8.0*. Research Triangle Park: Research Triangle Institute, 2001.

11. Lohr SL. *Sampling: Design and Analysis*. Pacific Grove: Brooks/Cole Publishing Company, 1999.

12. McDivitt JA, McDowell J, Zhou F. *Evaluation of the HEALTHCOM Project in West Java: Results from Surveys of Mothers and Volunteer Health Workers*. Philadelphia: Academy for Educational Development, The Annenberg School of Communication, University of Pennsylvania, 1991.

13. Fagbule D, Parakoyi DB, Speigel R. Acute respiratory infections in Nigerian children: prospective cohort study of incidence and case management. *J Trop Pediatr* 1994; **40:** 279–284.

14. Verstappen WHJM, van der Weijden T, ter Riet G, Grimshaw J, Winkens R, Grol RPTM. Block design allowed for control of the Hawthorne effect in a randomized controlled trial of test ordering. *J Clin Epidemiol* 2004; **57:** 1119–1123.

15. Pringle M, Stewart-Evans C. Does awareness of being video recorded affect doctors' consultation behaviour? *Br J Gen Pract* 1990; **40:** 455–458.

16. Kalter H. The validation of interviews for estimating morbidity. *Health Policy Plan* 1992; **7:** 30–39.

17. World Health Organization. *Health Facility Survey: Tool to Evaluate the Quality of Care Delivered to Sick Children Attending Outpatient Facilities*. Geneva: WHO, 2003.

18. Hermida J, Nicholas DD, Blumenfeld SN. Comparative validity of three methods for assessment of the quality of primary health care. *Int J Qual Health Care* 1999; **11:** 429–433.

19. Franco LM, Franco C, Kumwenda N, Nkhoma W. Methods for assessing quality of provider performance in developing countries. *Int J Qual Health Care* 2002; **14** (**suppl. 1**): 17–24.

20. Winch PJ, Bhattacharyya K, Debay M *et al. Improving the Performance of Facility- and Community-Based Health Workers*. Arlington: BASICS, 2003.

21. Hadi A. Management of acute respiratory infections by community health volunteers: experience of Bangladesh Rural Advancement Committee (BRAC). *Bull World Health Organ* 2003; **81:** 183–189.

Accepted for publication 1 April 2006