Risk and protective factors for two types of error in the treatment of children with fever at outpatient health facilities in Benin

Alexander K Rowe,¹ Faustin Onikpo,² Marcel Lama³ and Michael S Deming¹

Accepted 4 December 2002

Background In developing countries, health workers often do not follow clinical practice guidelines. However, few studies have examined why different types of errors occur.

Methods We analysed a sample of consultations of children with non-severe malaria (defined as fever without signs of severe illness) from a health facility survey conducted in Ouémé Département, Benin. Treatment was defined as correct (recommended antimalarial), a minor error (non-recommended antimalarial), or a major error (no antimalarial).

Results In all, 85 health workers and 289 children were studied. In a multivariate logistic regression analysis, the following factors were significantly associated with major errors: treatment by a physician (adjusted odds ratio [aOR] = 13.57, 95% CI: 1.45–126.75), child’s age <12 months (aOR = 3.41, 95% CI: 1.15–10.07), and child’s temperature (aOR = 0.58 per °C, 95% CI: 0.34–0.97). Factors significantly associated with minor errors were: child’s temperature (aOR = 1.43 per °C, 95% CI: 1.07–1.92), electricity at the health facility (aOR = 3.10, 95% CI: 1.05–9.17), ≥1 supervision visit in the past 6 months (aOR = 0.33, 95% CI: 0.14–0.77), fever treatment wall chart in the consultation room (aOR = 0.29, 95% CI: 0.12–0.73), and number of non-fever chief complaints (aOR = 0.67 per complaint, 95% CI: 0.48–0.93). In-service training in malaria treatment was not significantly associated with either error type.

Conclusions Many factors may influence health worker performance, and factors such as pre-service training may influence performance in unexpected ways. Identifying different errors and analysing them separately can help reveal potential causes that may be masked by combining errors into a single category.

Keywords Health services research, developing countries, Benin, malaria, child health services, epidemiological methods

In developing countries, clinical practice guidelines are increasingly used to improve health workers’ management of the leading causes of childhood mortality: pneumonia, diarrhoea, malaria, measles, and malnutrition.¹ Despite their potential for preventing deaths, studies have found that health workers often do not follow guidelines.²⁻⁸ To improve health worker performance (defined by adherence to guidelines) in developing countries, a better understanding is needed of why health workers do not follow guidelines.

Some studies have been published on this topic, however, most have important methodological limitations.⁹ Furthermore, among studies done with rigorous methods, performance was usually analysed as a dichotomous variable: correct or incorrect. In other words, all errors are combined. All errors, however, are not alike; and thus, such an analytical approach may be overly simplistic. For example, results from Benin on the treatment of illnesses requiring an antimicrobial revealed two common errors with very different clinical consequences: treatment with an...
antimicrobial that was effective, but not recommended by national guidelines (i.e. a ‘minor’ error since the error probably did not increase the child’s risk of dying); and treatment without any effective antimicrobial (i.e. a ‘major’ error since the error probably did increase the child’s risk of dying). 3

We reasoned that since errors are different, the causes of the errors might be different. To identify factors associated with major and minor errors, we analysed a sample of consultations of children with fever from a study conducted in southeastern Benin. The study had been originally designed as a baseline study to measure gaps between current practices and a new set of national guidelines that were about to be introduced. Because we were unsure which factors might lead health workers to make different errors, we examined a broad range of child, health worker, and health facility characteristics that seemed likely to affect health worker performance. Results from such a hypothesis-generating analytical strategy were expected to guide future research.

Methods

Study design and data collection

We analysed data from a cross-sectional survey of the quality of care received by ill children 1 week to 59 months of age at all licensed health facilities in Ouémé Département, Benin from July to October 1999. The methods are described elsewhere in detail. 3 Briefly, we used a cluster sampling design, with a cluster defined as all ill-child consultations occurring at a health facility on a single day. One cluster for each health facility was selected by systematic sampling. This sampling plan yielded an equal probability sample of consultations.

Three survey teams undertook training that continued until the agreement of practice results of surveyors and study investigators was greater than 90% (1 week). We did not give advance notice of our visit to health facility staff. Survey teams arrived at health facilities before the official opening time (8 a.m.) and stayed until the official closing time (6:30 p.m.). We collected data using five methods: (1) consultations were silently observed by a surveyor who used a checklist to record which clinical tasks were performed; (2) standardized, structured interviews were conducted with caretakers (usually the mother) when they were ready to leave the facility; (3) children were re-examined by a study clinician, out of the health worker’s view, to obtain a ‘gold standard’ clinical evaluation; (4) medical supplies and equipment at the facility were assessed by a surveyor at the end of the day using a checklist; and (5) standardized interviews were conducted with health workers at the end of the day to obtain information on health workers’ training, supervision, and opinions about their work. After children were re-examined, the study clinician gave appropriate medications free of charge to any child who had been inadequately treated. Surveyors measured temperatures with an electronic ear thermometer (Model HM3, Thermoscan Inc, San Diego, California). Health workers’ knowledge of fever treatment was assessed with three hypothetical case scenarios in which a child had fever and other symptoms.

The inclusion criteria for children in this analysis were that they were enrolled in the health facility survey, brought for an initial consultation with non-severe malaria (see definition below), and not referred for hospitalization. We studied non-severe malaria because it was the most frequently encountered disease in our study sample and because correct treatment prevents progression of this illness to severe disease and death.

Definitions

Definitions for malaria diagnosis and treatment were based on guidelines recommended by Benin’s national malaria control programme9 that were in place when the study was conducted (Box 1). Because we did not collect all the information needed for establishing a diagnosis according to these guidelines and because the guidelines sometimes lacked precision, we used a modified definition (Box 1).

Benin guidelines recommended treating non-severe malaria with oral chloroquine or, if a child had already received a full course of chloroquine, oral sulfadoxine-pyrimethamine (SP). We classified treatment quality as correct, a minor error, or a major error. ‘Correct’ meant the antimalarial essentially agreed with Benin’s guidelines: only oral chloroquine or, if the caretaker reported already treating with chloroquine, either oral chloroquine or oral SP (we accepted either medicine because we did not ask about chloroquine dosage and we assumed that most children would not have received a full course [Dr John

Box 1 Definitions of non-severe malaria

"Definition according to Benin national guidelines: fever (axillary temperature >37.5°C or a history of fever) and no sign of severe malaria (persistent vomiting, convulsions, impaired level of consciousness, severe anaemia,a renal insufficiency,a respiratory syndrome,a haemoglobinuria,a etc ... [sic])."

"Definition used in the analysis: fever (ear temperature >38°C or a complaint of fever from the caretaker or a history of fever from the caretaker in response to a health worker’s question) and no sign of severe malaria. Signs of severe malaria for children 2–59 months were: lethargy, unconsciousness, severe palmar pallor, or a history of ‘vomiting everything’ according to the survey clinician; and observed convulsions, a complaint of convulsions from the caretaker, or a report of convulsions from the caretaker in response to a health worker’s question according to the survey observer. Signs of severe malaria for children <2 months were: lethargy or unconsciousness according to the survey clinician; and observed convulsions, a complaint of convulsions from the caretaker, or a report of convulsions from the caretaker in response to a health worker’s question according to the survey observer."

a No explicit definition provided.

b Temperature cutoff value based on recommendations from the manufacturer of the ear thermometer used in the study.
Watson, Centers for Disease Control and Prevention, unpublished survey conducted in Ouémé Département, Benin, 1999). ‘Minor error’ meant the child was treated with an antimalarial (chloroquine, SP, or quinine), but the treatment did not match the guidelines; and ‘major error’ meant the child was not treated with an antimalarial.

These definitions have two potential limitations: dosage was not considered, and cotrimoxazole (trimethoprim-sulfamethoxazole), a commonly used medicine that can treat malaria, was not considered an antimalarial. Dosage was not considered because health workers’ documentation of dosage was often incomplete. Cotrimoxazole was not considered an antimalarial because Benin’s malaria control programme never recommended its use and never trained health workers to use it; thus, it was unlikely that health workers would have known that cotrimoxazole was an antimalarial (personal written communication, Dr Alban Quenum, Benin Ministry of Public Health, 12 August 2001).

Analysis

Data were double-entered into an electronic database and verified using EpiInfo. All analyses were performed using SAS. To identify predictors of major and minor errors, we first performed a univariate analysis with a series of binomial logistic regression models using the SAS GENMOD procedure with an exchangeable working correlation matrix. This procedure uses generalized estimating equations to account for the potential correlation of treatment quality among children treated by the same health worker. The factors examined were: time of the consultation (before 1:00 p.m. versus 1:00 p.m. or later); history of the child receiving previous antimalarial treatment for the same illness; health workers’ age, sex, and work experience (years working as a health worker and number of health facilities in which the person had worked); and all variables listed in Tables 1–4 (except where indicated). For supervision, we examined both quality (in three levels: supervision with observation during a consultation and feedback, supervision without observation and feedback, and no supervision) and frequency (number of visits involving observation of a child consultation in the past 6 months, total number of visits in the past 6 months, and >1 visit versus no supervision). For in-service malaria training, we examined whether the health worker received training, training duration, time since training occurred, and whether training included clinical practice. In addition, we evaluated interactions between in-service malaria training and pre-service training, in-service malaria training and supervision, and fever treatment knowledge and supervision.

To adjust for confounding, variables with a P-value < 0.10 from the univariate analysis were entered into three binomial logistic regression models using the GENMOD procedure. In the first model, the outcome was major error versus not a major error; in the second model, the outcome was minor error versus not a minor error; and in the third model, the outcome was correct treatment versus any error. Five variables meeting the entrance criteria for either the major-error model or the minor-error model was automatically entered into the other two models. The three models were made to contain the same variables to quantify how much each predictor ‘shifted’ the distribution of error types (see Results for an example). Hypothesis testing and CI estimation were done with an alpha level of 0.05.

Results

Enrolment and study participants

Altogether, 87 health facilities were visited, and 109 health workers and 584 ill children were enrolled (100% participation). Of (82.4%) nursing aides worked in small government facilities and never worked in non-governmental facilities, while most (83.3%) physicians worked in non-governmental facilities and never worked in small government facilities; and most nursing aides reported that they were underpaid and that staff were not motivated (94.1% and 88.2%, respectively), while physicians did not often report these problems (33.3% and 16.7%, respectively). The malaria diagnosis variable was excluded because it was believed to be in the causal pathway between other independent variables and the outcome; and the health worker temperature variable was excluded because another closely related factor, the temperature measured by the surveyor, had greater validity and no missing values.

To ensure that the three models contained the same variables, any variable meeting the entrance criteria for either the major-error model or the minor-error model was automatically entered into the other two models. The three models were made to contain the same variables to quantify how much each predictor ‘shifted’ the distribution of error types (see Results for an example). Hypothesis testing and CI estimation were done with an alpha level of 0.05.

Results

Enrolment and study participants

Altogether, 87 health facilities were visited, and 109 health workers and 584 ill children were enrolled (100% participation). Of...
Motivating factor (coded responses to the open-ended question: ‘What factors motivate you in your work?’)\(^a\)

<table>
<thead>
<tr>
<th>Motivating factor</th>
<th>No. of health workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training opportunities</td>
<td>38 (44.7)</td>
</tr>
<tr>
<td>Financial motivation (e.g. salary increase or bonus)</td>
<td>37 (43.5)</td>
</tr>
<tr>
<td>Altruism (e.g. helping others)</td>
<td>9 (10.6)</td>
</tr>
</tbody>
</table>

\(^a\) These opinions were not thought to be related to the quality of malaria treatment and therefore were not included in the univariate or multivariate predictors analysis; they are shown to reflect the complete list of potential problems included in the original health facility survey.

\(^b\) Up to three responses were recorded per health worker, and only the three most common responses are shown.

Health worker opinions and knowledge assessment

Health workers reported a variety of work-related problems. Lack of training and being underpaid were among the most commonly reported problems (Table 2). The most commonly reported motivating factors were training opportunities, financial motivation, and altruism.

In the knowledge assessment, the proportion of health workers whose response included the correct answer, chloroquine, varied according to the complexity of the case scenario. Health workers most often gave a correct answer (68/85, or 80.0%) for the scenario of a child who only had malaria. For scenarios describing malaria plus a second illness, the proportions were lower: 71.8% (61/85) for a child with malaria plus pneumonia, and 67.1% (57/85) for a child with malaria plus dysentery. The proportion of health workers who correctly answered at least two of the three scenarios was higher for trained health workers (21/23, or 91.3%) compared with untrained health workers (44/62, or 71.0%) (two-sided Fisher’s P-value = 0.082).

Predictors of fever treatment quality

Of the 48 variables studied as potential predictors of major and minor errors, 27 had univariate P-values ≥ 0.10 and were excluded from the multivariate analysis (selected results shown in Table 3). Two other variables, the availability of chloroquine and quinine, could not be studied because they were available at all or nearly all health facilities (Table 1). Of the 19 variables with a univariate P-value < 0.10, five were excluded from the multivariate analysis for other reasons (see Methods and Table 3 for details).

Multivariate analysis of the remaining 14 variables revealed a variety of associations with each error type. For major errors, significant risk factors (P-value < 0.05) were treatment by a physician and child’s age < 12 months; a significant protective factor was child’s temperature; and a protective factor of borderline significance (P-value 0.05–0.10) was a chief complaint of fever (Table 4). For minor errors, significant risk factors were child’s temperature and electricity at the health facility; a risk factor of borderline significance was treatment by a nurse; and significant protective factors were supervision, fever treatment wall chart in the consultation room, and number of non-fever chief complaints. None of the interaction terms tested was statistically significant.

To further examine the effect of in-service malaria training, a factor of particular programmatic interest, we forced the variable into the multivariate models. In this analysis, malaria training was not associated with making either a major error (adjusted odds ratio (aOR) = 1.99, 95% CI: 0.41–9.64) or a minor error (aOR = 0.58, 95% CI: 0.21–1.57). Results for other variables in the model were similar to those shown in Table 4.

The last column in Table 4, which compares correct treatment versus any error, illustrates the consequences of not separating treatment error types. Significant factors were treatment by a physician or nurse (risk factors for any error), supervision, and presence of fever treatment wall chart (protective factors for any error); factors of borderline significance were the health worker opinion that financial incentives were a source of motivation and the number of non-fever chief complaints (protective factors for any error). Thus, the analysis that combined errors ‘missed’ the significant associations of temperature, child’s age, and electricity at the facility.
In addition to describing the statistical associations between predictors and treatment errors, the three sets of modelling results in Tables 3 and 4 are designed to be interpreted together to show how much each predictor 'shifts' the distribution of error types. For example, in Table 4, results for the fever treatment wall chart suggest that, if there was truly a causal relationship between wall charts and treatment quality, the installation of wall charts in consultation rooms would increase the odds of correct treatment by 2.8-fold and decrease the odds of a minor error by 71%. In other words, ill children would 'shift' from having minor errors to being treated correctly, without changing the occurrence of major errors.

Given the moderate sample size and the possibility that the models contained too many variables, we repeated the analysis using smaller models that only contained variables meeting the entrance criteria for the outcome of the model (e.g. the major-error model only contained the five variables with \( P \)-value \(< 0.10 \)). Results from these analyses suggested that some associations disappeared and complex relationships were masked. Thus, when studying the determinants of health worker performance or evaluating interventions to improve performance, separating error types may be important. Moreover, such a conclusion is probably applicable for studies in both industrialized and developing countries.

Conclusions

A main finding of this study was that the predictors of major and minor errors were usually different. Not only were factors often associated with only one error, but in one case the associations with the two error types were in opposite directions: increasing temperature was a protective factor for a major error and a risk factor for a minor error. When errors were combined, some associations disappeared and complex relationships were masked. Thus, when studying the determinants of health worker performance or evaluating interventions to improve performance, separating error types may be important. Moreover, such a conclusion is probably applicable for studies in both industrialized and developing countries.

The decision to separate error types, however, can complicate studies in two important ways. First, it may be difficult to decide exactly what the error types should be. For example, if we had analysed medicine dosage (defined by the amount of drug and duration of treatment) as well as selection, the quality of care outcome may have had more than three categories: recommended drug correctly dosed, recommended drug overdosed, recommended drug underdosed, non-recommended drug correctly dosed, non-recommended drug underdosed, etc. In addition, terms such as overdosed and underdosed must be defined, which may be difficult because dosage recommendations for oral medications are sometimes given in more than one way. For example, in a training manual from Benin, the chloroquine dosage for the first day of treatment for a child weighing 3.4–7.4 kg is described both as three-quarters of a 100 mg tablet and 10 mg/kg.12
categories generally requires a larger sample. With the example described above, the sample size needed may be so large that it may be necessary to combine errors into a smaller number of categories defined by the clinical consequences of each error, such as the major and minor errors used in this analysis. Such an approach was used in a recent study of errors in the management of paediatric inpatients.13

A second main study finding is that ‘common sense’ factors, such as pre-service and in-service training, knowledge, and caseload, were either not associated with performance or had unexpected associations. The most striking example was that pre-service training, which was closely related to health facility type and the opinions that health workers were underpaid and not motivated, appeared to have a dose–response relationship.
with both error types: the more advanced the training, the greater the odds of making an error. We found a similar relationship when examining the predictors of diarrhoea treatment in the same Benin survey. One explanation for the minor errors is that health workers with less training may have been aware of fewer treatment options (e.g. chloroquine may have been the only antimalarial nursing aides knew how to give). Second, nurses and physicians may have been more likely to think that chloroquine was ineffective. Third, the association may have been more a result of health facility type than pre-service training: non-governmental facilities may have been places where drugs such as quinine, which generate greater profits, were preferentially prescribed.

Although the sample size of consultations performed by physicians was small, the possible relationship between treatment by a physician and making a major error may be explained by physicians' confidence in their ability to differentiate between fevers caused by malaria and fevers caused by other pathogens. In other words, physicians may have rejected the simple guideline 'fever equals malaria'.

For factors such as supervision and in-service training, which are commonly used to improve health worker performance, there are several possible explanations why they were not found to protect against both error types. First, there may have been a measurement problem. Perhaps characteristics such as whether a health worker was trained or supervised is less important than the technical or interpersonal quality of the training or supervision. We tried to assess quality with indicators such as the inclusion of clinical practice (for in-service training) and observation and feedback (for supervision); however, these measures did not distinguish health workers who performed well from those who made errors. Second, the lack of association may have been a type 2 statistical error, although the major error OR for both factors were greater than one, suggesting that training and supervision might have been risk factors for major errors. Third, the in-service training and supervision may have been effective, but only for a short period of time. Thus, by the time of the survey, the effects were no longer measurable. Finally, it is possible that training and supervision were ineffective. This is not to say that training and supervision are useless as strategies for improving health worker performance; rather, the way they were conducted in the study setting may have been ineffective or insufficient.

We also found a variety of child characteristics that were predictors of health worker performance. Some associations, such as that of temperature and major errors, seem easily explainable (i.e. higher temperatures alert health workers to treat with an antimalarial). Additionally, the association of temperature and minor errors may be explained in part by the fact that an older malaria treatment guideline in Benin had stated that a temperature \( \geq 39^\circ C \) was a sign of severe malaria, which should have been treated with quinine.

Other associations, however, such as the number of non-fever chief complaints (a measure of case complexity) being a protective factor against minor errors, are more puzzling. Perhaps the presence of multiple complaints in a child with fever led health workers to consider aetiologies other than malaria, so when children were treated to 'cover' malaria, the least expensive antimalarial (chloroquine), used without other antimalarials, was considered adequate. We stress, however, that these explanations are speculative; clearly, more study is needed to confirm and understand the true causes for these associations.

**Usefulness of the findings**

The usefulness of our findings can be viewed from four perspectives. First, from a conceptual perspective, the results contribute to the limited understanding of the factors that influence health worker behaviour. Although our study raises as many questions as it answers, perhaps it will stimulate further research on a more focused set of questions. Second, from a methodological perspective, identifying factors that are associated with health worker performance is useful because when future interventions are tested, it may be important to measure and adjust for these factors, as they are potential confounders. In addition, the use of multiple binomial logistic regression models to examine shifts among treatment categories might be helpful in selecting or prioritizing interventions. For example, an intervention that shifts major errors to correct treatment might be given higher priority than an intervention that only shifts minor errors to correct treatment, assuming the cost and feasibility of the two interventions are similar. Third, from a programmatic perspective, our results can be used to develop messages or teaching points; for example, 'Remember to treat all children with a history of fever—even if they have a normal temperature'. Finally, from a policy perspective, the results raise important issues. For example, should physicians or health workers at private facilities be expected to follow national guidelines? And, what changes in supervision and training are needed to improve health worker performance?

**Limitations**

This study has several limitations. First, we evaluated performance by directly observing health workers, which probably altered their behaviour. Second, the definition of malaria was not exactly the same as the guidelines health workers were expected to use. Third, some potentially important factors could not be studied, such as health facility type, drug availability, caretaker demands for certain treatments, profit motives, and confidence in the guidelines. Finally, our definition of treatment quality considered neither dosage nor whether caretakers understood how to administer medications at home—both of which are important facets of health care quality.

This study highlights both the complexity of health worker behaviour and the need for improvements in the way we study it. First, we need better methods for measuring performance. An ideal method has three features: the health worker is unaware that he or she is being observed; the inclusion of children with real illnesses; and a gold standard measure of the child's illness against which the health worker's performance can be compared. No method we are aware of has all these features: the direct observation used in this study involves children with real illnesses and a gold standard, but lacks unobtrusive measurement; the simulated client method is unobtrusive and involves a gold standard, but usually does not use children with real illnesses; and record reviews involve children with real illnesses and are unobtrusive, but they lack a gold standard. Second, we need more precise methods for measuring predictors, especially the quality of training and supervision, and latent variables such as motivation. Third, we need to expand
the use of qualitative methods to guide quantitative studies. Finally, we need larger studies to have adequate statistical power to identify multiple and complex interactions (e.g. three-way interactions) and to study the determinants of less-common outcomes (e.g. treatment of severe disease).

Ultimately, the goal of conducting research on health worker performance is to improve performance with interventions that are affordable and have long-lasting effects. Opportunities for conducting randomized, controlled trials that test such interventions, however, are uncommon. Therefore, a better understanding of health worker behaviour is needed so that interventions can be developed with the greatest chance of success.

**Acknowledgements**

We thank Francois Cokou for his assistance with data management, Samantha Rowe and Dawn Osterholt for their helpful suggestions on the analysis and the manuscript, and Africare for its support of our research activities. This survey was funded by the US Agency for International Development’s Africa Integrated Malaria Initiative (project number 936-3100).

---

**KEY MESSAGES**

- Two types of error occur in the treatment of children with suspected malaria: minor errors, which are unlikely to increase the risk of dying, and major errors, which may increase the risk of dying.
- At outpatient clinics in Benin, minor errors were common (prevalence = 31.1%) and major errors were relatively uncommon (prevalence = 5.5%).
- Major errors were associated with children treated by a physician, <12 months old, and with a lower temperature; minor errors were associated with children with a higher temperature, fewer non-fever chief complaints, treated at clinics with electricity and without fever treatment wall charts, and treated by unsupervised health workers.
- In-service training in malaria treatment was not significantly associated with either error type.
- Identifying different errors and analysing them separately can help reveal potential causes that may be masked by combining errors into a single category; understanding the causes of treatment errors might lead to effective and affordable interventions that prevent the errors.

---

**References**


