Reviews/Analyses

From efficacy to effectiveness: insecticide-treated bednets in Africa

C. Lengeler¹ & R.W. Snow²

Insecticide-treated bednets and curtains (ITBC) have proven in recent large-scale trials to have a high efficacy in reducing morbidity and mortality from malaria in African children. However, it is unlikely that the efficacy measured in trials can be entirely sustained under programme conditions. This has important implications for the cost-effectiveness of the intervention. Furthermore, there is a need to assess the long-term impact of ITBC.

This article traces the history of ITBC and the different phases of their assessment, especially the determination of efficacy in randomized controlled trials (phase III assessment). It then outlines the reasons for continued assessment of their effectiveness under programme conditions (phase IV assessment). The methodologies for measuring effectiveness are discussed, and a critical review of the issues reveals that it is impractical to measure effectiveness directly.

A simple effectiveness model, allowing for differentiation between individual and community effectiveness, provides a useful conceptual framework. First, individual effectiveness is measured through a case-control study. This estimate is then combined with a coverage indicator to estimate community effectiveness. This approach could provide programme managers with a powerful tool to monitor the impact of health interventions at the community level.

Introduction

Malaria is the major challenge to further improvements in child survival in sub-Saharan Africa. Plasmodium falciparum accounts for >25% of childhood mortality outside the neonatal period (1), and malaria has been ranked first by the World Bank in terms of disability-adjusted life years (DALYs) lost in sub-Saharan Africa (2). Drug and insecticide resistance as well as insufficiently developed and financed health services have hampered efforts over the past 20 years to improve the situation. As a result, the malaria burden has remained largely unchanged.

The search for a safe, effective, and cheap vaccine continues. However, it is unlikely that a vaccine will play a major role in the control of malaria among African children during the next decade, despite recent results in the United Republic of Tanzania with the SPf66 vaccine (3). Prompt diagnosis and treatment therefore remain the basis for controlling malaria morbidity and mortality (4). As a result, the search for cheap and effective antimalarials must continue to be a priority, given the diminishing efficacy of chloroquine in many parts of Africa. In parallel with these efforts, a renewed emphasis on disease prevention is required. Over the past decade, significant advances have been made in the prevention of malaria using insecticide-treated bednets and curtains (ITBC). The technology was developed through small-scale experiments and trials which led to efficacy testing in larger-scale trials (with morbidity and mortality as the major measured outcomes). Plans for the large-scale implementation of ITBC are now being considered by ministries of health, with the support of donor agencies, as integral components of national malaria control programmes. In view of this, it is crucial to assess how much of the efficacy (measured under carefully controlled trial conditions) can be retained as effectiveness (under programme conditions). High efficacy does not necessarily imply high effectiveness, and there is a need to examine closely the steps that lead to effectiveness in the community (5).

This article reviews the rationale and the principal assessment stages which lead from efficacy to...
effectiveness assessment. The concept of effectiveness is then discussed in a way that renders it measurable. Finally, the methodological approaches that are available for assessing the impact of ITBC programmes are reviewed. The article focuses on Africa and *P. falciparum* malaria, excluding important work done in China and elsewhere on the impact of ITBC on *P. vivax* malaria.

**Historical review of bednet studies in Africa**

The application of a residual insecticide to fabrics as a means of personal protection against vector-borne diseases began during the Second World War with the impregnation of bednets and combat fatigue by the Soviet, German, and U.S. armies (6). In the late 1970s pyrethroids were used for this purpose; their high insecticidal activity combined with low mammalian toxicity made them ideal for treating fabrics. A scientific panel convened in 1983 by WHO reviewed the first laboratory evidence and recommended the initiation of field trials to assess the potential of this technology for disease control (6).

Although ITBC are not required to go through a formal licensing process, the steps in the development of this intervention have analogies to those followed for vaccines and drugs (pre-clinical and phase I–IV trials). The first two stages for ITBC were carried out largely in the 1980s. These studies showed that this use of pyrethroids was safe and that ITBC had a significant impact on entomological parameters such as vector feeding success, vectorial capacity, and man-biting rates. These studies were instrumental in defining the mechanism of action (repellency and killing), determining the optimal dosage for different insecticide and netting combinations, and revealing the negative influence of washing upon insecticide activity (6–8).

Between 1986 and 1994 a number of small-scale trials (which could be classified as phase II trials) were conducted. These established the feasibility, acceptability, safety, and apparent efficacy of ITBC. Few of the early studies in Africa used a formal clinical trial design, i.e., having a sufficient number of intervention and control units to be randomized (at either the individual or community level) and a prospective longitudinal assessment of outcome. However, trials in the Gambia (9), Kenya (10), Guinea Bissau (11), the United Republic of Tanzania (12), and Sierra Leone* met basic requirements for randomized clinical trials. At least 10 other trials were conducted in Africa, and although they did not meet these requirements, they contributed valuable knowledge on the implementation of ITBC and their impact on entomological and morbidity parameters. Direct comparison between the studies is difficult because of the wide range of malaria endemicity, types of control group, and measurement method. However, most showed a protective efficacy against malaria in the range of 20–60%.

Larger phase III trials (having a reduction in childhood mortality as the main outcome of interest) began with a study in the Gambia (13). As with phase II trials, investigators supervised the delivery of ITBC to ensure optimal coverage of the target population and correct dosage of insecticide. Following 2 years of continuous demographic surveillance, a protective efficacy of 63% against childhood mortality (due to all causes) was demonstrated. This was sufficient for the Gambian government to introduce insecticide treatment of existing bednets into its national primary health care (PHC) system.

However, results from a single country, with its tradition of bednet use (14) and markedly seasonal malaria transmission of low-to-moderate intensity (15), were insufficient for pan-African recommendation of ITBC for malaria control. The WHO Special Programme for Research and Training in Tropical Diseases therefore initiated further phase III trials in areas where bednets were uncommon and the malarial ecologies different: Ghana (Navrongo District), Burkina Faso (Oubratenga Province), and Kenya (Kilifi District). These studies are randomized controlled trials in which efforts to achieve maximum compliance have been made. Results from these large trials (between 60000 and 120000 participants each) became available in 1995 (Kenya and Ghana) or will become available in 1996 (Burkina Faso). Combined with the Gambian data (13), these trials provide a range of estimates of the protective efficacy of ITBC against childhood mortality across a variety of settings in Africa.

**The need for efficacy estimates**

The randomized clinical trial is the most powerful and bias-free means of providing accurate and critical evaluations of the efficacy of interventions. Furthermore, accurate data on efficacy allow the calculation of the number of clinical cases or deaths averted by an intervention in a given population. This can be combined with the cost of implementation to obtain cost-effectiveness estimates.

The need for mortality trials as well as morbidity trials stems from our inability to infer


---

C. Lengeler & R.W. Snow
malaria-related mortality rates on the basis of malaria morbidity rates. It can be assumed that in most stable endemic areas every child will suffer at least one febrile episode due to malaria each year. Yet only a small proportion of these children develop severe, life-threatening disease or die (16). The mechanisms that predispose these children to severe malaria are poorly understood. Death is the last stage in a chain of events affected by many factors (including access to treatment). The relation between infection, disease, and death is therefore likely to be variable. Consequently, mortality cannot be predicted from the results of morbidity trials. However, it is unlikely that an intervention which does not reduce host infection rates or fever episodes attributed to malaria will have an impact upon mortality.

Unfortunately, large-scale efficacy trials are expensive and usually need to be conducted over at least 2–3 years. The cost of the three WHO-funded trials exceeded US$ 3 million over a 4-year period. Consequently, phase III trials need to be planned carefully, so as to cover as much as possible the spectrum of transmission, disease, and sociodemographic characteristics across different settings in Africa. Finally, the data need to be reviewed systematically and critically in order to provide a scientific basis for the promotion of ITBC as a public health intervention.

From efficacy to effectiveness

Phase II and III efficacy trials should form the basis for decisions to implement an intervention through national or regional pilot programmes. Once such programmes have begun, Ewbank (17) has defined two complementary stages of evaluation: operational (process) monitoring, and impact (effectiveness) monitoring.

Operational monitoring is the routine assessment of process indicators (e.g., the number of nets delivered and treated with insecticide at the correct dose, the number of personnel trained, the number of educational posters produced, the coverage and compliance of the target group). Operational monitoring is usually integrated into project management.

By contrast, the monitoring of impact indicators, such as a reduction in mortality, has rarely been carried out in health programmes. Effectiveness (phase IV) assessment relates to measuring the health impact under programme conditions and differs from efficacy (phase III) assessment in several respects. The implementation of health programmes in the developing world is fraught with problems (delivering essential components to remote areas, limited access to them by the community, inadequate training of service providers, ineffective application of interventions, poorly designed or implemented promotional campaigns, “leakage” of supplies and equipment, insufficient resources, and so on). Furthermore, in phase IV assessments an intervention might not be delivered to the same target group or with the same frequency as in phase III trials because the delivery capacity of health services is limited. It is therefore unlikely that efficacy estimates derived from tightly controlled phase III trials will ever be realized in programme settings. These shortcomings affect cost-effectiveness, which is important for informed resource allocation (2). It is therefore essential to document effectiveness in a phase IV assessment.

Example of a phase IV trial of ITBC

A recent study in the Gambia (18) was the first in Africa to carry out phase IV assessment of ITBC. The investigators assessed the impact on childhood mortality of the government’s delivery of free insecticide for use on bednets already owned by the population. The delivery was integrated into the PHC system, and nets were treated by village health workers with minimal control by the investigators. A prospective mortality evaluation with a contemporaneous control group was carried out. Two operational limitations were identified: despite overall high use of bednets in the Gambia, several areas showed low net use among children; and a minimum target dose of 200 mg/m² permethrin was achieved on only 46% of sampled nets (U. d’Alessandro, personal communication, 1995). Notwithstanding these deficiencies, childhood mortality showed a statistically significant reduction of 25%, ranging from zero in areas where net use was low to >40% in areas where a high percentage of children slept under bednets. The contrast between a mortality reduction of 25% in the phase IV assessment and a 63% mortality reduction in the phase III trial (13) demonstrates that efficacy and effectiveness are not synonymous. A comparative analysis of the cost-effectiveness of the two trials also shows a substantial difference in the cost per death averted (in 1992 US$ equivalent): US$ 223 in the phase III trial (19) and US$ 600 in the phase IV programme. In light of these changes in both impact and cost, it is clear that phase IV assessment provided data that are more likely to reflect the results of integration of ITBC into the PHC system.

Defining effectiveness

Before reviewing the methodology of measuring programme impact it is useful to discuss briefly the concept of effectiveness. It will be shown that community impact can be measured either directly or estimated through its components. Definitions of the terms used below are given by Last (20).

Interventions have an impact at two levels: community and individual. There is therefore a need to define both community and individual effectiveness.

Community effectiveness is the impact of a health programme on the overall death rate in the community. It can be expressed in one of two ways.

- The ratio of death risk in the intervention group to that in the control group expresses relative risk (RR). With this ratio the community protective effect (CPE = 1 − RR × 100) can be obtained, which expresses the percent reduction in the overall death rate as a result of the intervention (63% and 25% in the two Gambian studies discussed above, 13, 18).

- The number of deaths averted as a result of the intervention (attributable risk) is measured by the difference in risk between the two groups. In the Gambian studies discussed above, attributable risks were 15.2 and 2.2 deaths per 1000 children, respectively.

By contrast, individual effectiveness measures the reduction in disease or death risk that a compliant individual can expect compared to a non-compliant. It can be defined only by a risk ratio (relative risk) and hence as a protective effect (see section Measuring individual effectiveness with case–control studies).

Community and individual effectiveness are not equivalent, unless the intervention coverage is 100%. In an ITBC programme, coverage refers to the number of target children with access to ITBC. For an ITBC intervention:

\[
\text{Community effectiveness} = \frac{\text{Individual effectiveness} \times \text{Coverage}}{} \quad (\text{eq. 1})
\]

This model provides a useful framework, since directly measuring community effectiveness may not be easy (see sections Measuring effectiveness using historical or non-compliant controls and Measuring community effectiveness directly). Individual effectiveness is likely to be adequately measured with a case–control approach (see section Measuring individual effectiveness with case–control studies). Coverage can be measured either from data that are routinely obtained or from limited cross-sectional surveys (in-process monitoring).

The other parameters in the more complex model proposed by Tugwell et al. (21) can be ignored. Although user compliance is an important parameter, it can be excluded as long as a similar definition of “protection” is used in both coverage surveys and in defining “exposure” in a case–control study.

The time frame of effectiveness measurement

There is a need to assess the impact of ITBC programmes over a period longer than 1–2 years for two reasons. First, the Gambian phase IV trial assessed impact on mortality while the programme was relatively new and the morale and enthusiasm of providers and recipients were likely to be highest. Short-term evaluations are therefore unlikely to reflect sustained impact.

Second, there is the impact on the functional antidental immunity of reducing human–vector contacts over long periods. Comparative studies of high- and low-transmission areas in East Africa show similar rates of severe, life-threatening disease below the age of 5 years, although the peak age for severe disease is younger in areas of high transmission (22). This fact has been used to explain the apparent similarities in malaria-specific mortality across a wide range of transmission intensities (23). The data suggest that under natural conditions in Africa, a >90% reduction in infection rates could alter only the age at which severe disease and death occur, without a net cohort gain by the fifth birthday. Current trials of 1–2 years are unlikely to detect this because they are conducted with a large proportion of children who have some immunity due to high transmission conditions; the most important group, young, immunologically “naive” children, will be initially protected by the intervention, but the trial will have been completed when they enter the period of deferred risk. Although evidence pointing towards this phenomenon is scant, it is critical for a recommendation on ITBC programmes in Africa. Only through long-term programmes (with phase IV assessment) can such questions eventually be answered, since randomized controlled trials are unsustainable over long periods.

Measuring effectiveness using historical or non-compliant controls

The principal objective of programme delivery is to offer the intervention to the whole target population each year. This pre-empts the use of contemporaneous control populations for impact assessment. The usual alternative in such situations, a stepped-wedge
design, is unlikely to work for ITBC programmes for the following reasons: the uptake rate of the intervention would have to be rapid, so that the steps would not last longer than 1–2 years, which is unlikely under programme conditions; and the marked differences in infection rates, morbidity, and mortality that exist within countries and even regions (24) would render the identification of suitable control groups difficult. The cost of large-scale demographic monitoring would also be high. Further discussion of problems associated with randomized controlled trials for programme evaluation can be found in the article by Smith (25).

There are two alternatives for evaluating health impact in the absence of a defined control population: use of historical control data, or use of non-compliers as controls.

**Historical controls.** In most endemic areas there is a marked seasonal and year-to-year variation in the number of malaria cases (24). As a result, even pre-intervention surveillance data over many years would not allow a convincing conclusion to be made as to whether the subsequent 5- or 10-year rates were different from those before the intervention. In most settings, there is also a marked secular trend in malaria mortality and morbidity as well as in all-cause childhood mortality (26). This problem is compounded, since most malaria control programmes include additional activities such as improved case-management. The use of historical controls to interpret changes in malaria morbidity and mortality is thus full of difficulties, and such data should be interpreted critically.

**Non-compliers as a control group.** Comparing compliant individuals (or communities) to those that are non-compliant provides a potential alternative but raises the problem of selection bias. Compliers are unlikely to be similar to non-compliers in subtle but important ways. Their socioeconomic background, their expenditure on health, and their beliefs about disease prevention might be different. This may account for a difference in malaria rates and survival that is independent of the assessed intervention.

**Measuring community effectiveness directly.** Mortality changes are the most unambiguous, objective, and tangible indicators of programme success. Overall mortality should be preferred to malaria-specific mortality as an outcome measure, since it is difficult to assess the latter in areas where acute respiratory tract infections and gastroenteritis compete with malaria as causes of childhood death. Two investigations of verbal autopsy in Africa (27, 28) confirm that it lacks sensitivity and specificity.

There are three main approaches to measuring overall mortality at the community level: routine civil registration, active prospective monitoring, and indirect demographic techniques. The discussion in the previous two sections applies equally to all three approaches.

**Routine civil registration.** It is accepted that routine civil registration of deaths is too incomplete in most developing countries to be of much use in programme monitoring (29, 30). It may be possible to upgrade such systems in defined areas for the purpose of evaluating the impact of a programme. However, there would be the problem of defining who is protected. This approach needs further investigation.

**Prospective surveillance.** The prospective monitoring of childhood mortality at the district or regional level using specialized demographic teams is usually the best approach, and is usually chosen in phase III trials. Unfortunately, it is expensive and difficult to justify in a programme. The supervision and management of modifications to this system, such as the use of village informants who report vital events prospectively, are also expensive and require extensive supervision and quality control.

**Indirect demographic techniques.** Indirect demographic techniques are a potentially cheaper approach to mortality surveys, especially within national programmes.

The required data can be collected in population-based censuses that are simple, rapid (1–3 months), and relatively inexpensive. These are proven methodologies for the assessment of mortality trends and have been used widely in national censuses. The first published technique was developed in the 1960s by Brass et al. (31). It involves asking women of child-bearing age about all their births and their children’s subsequent survival. A life-table is then used to reconstruct the mortality trend over the decade preceding the survey. Another technique, giving a more precise estimate of recent mortality relies on mothers’ recall of their last two pregnancies (32). This and related techniques could be used to evaluate health interventions (33). If the samples of women are large enough and if each event can be dated through local calendars, a mortality trend for the previous 2–10 years can be inferred, allowing the estimation of community effectiveness.

However, these techniques have limitations
Measuring individual effectiveness with case–control studies

A case–control approach for the evaluation of interventions against tropical diseases has been proposed by several workers in a variety of fields (25, 35–38). The standard reference remains that by Schlesselman (39), which discusses the basic issues of design and analysis.

In a case–control study, a group of children who have died or suffered from a disease (cases) are compared with healthy children (controls). The aim is to discover different exposure factors (such as ITBC use) between the groups that explain the difference in disease. Since exposure is determined individually for each child, this method measures individual effectiveness. The measure obtained (odds ratio, OR) is an approximation of relative risk. Individual protective effect (IPE) can be calculated as: IPE = 1 − OR × 100.

For rare outcomes such as death, this is a more efficient design, as there is no need to follow a large cohort of children over a long period of time in order to detect a sufficient number of events. A case–control study requires only a few hundred cases and controls (who can be identified by different approaches such as rapid interview surveys or routinely available statistics). As a result, investigators can ensure high quality of data collection for each participant.

Another significant advantage of this method for effectiveness monitoring is that several important exposures (such as access to treatment or health education) can be examined concurrently. In fact, this is necessary to control for confounding factors. Case–control studies are the method of choice for measuring at the individual level the impact of an intervention under programme conditions. Such studies work in situations in which there is no defined control population, provided that there is no excessive selection bias between compliers and non-compliers (see section Measuring effectiveness using historical or non-compliant controls).

Case–control studies are attractive, but their pitfalls should be borne in mind. The most important is that they are susceptible to bias, since the choice of cases and controls is determined by the investigators. The different types of bias have been described comprehensively by Sackett (40). The study design has to be carefully planned using expert assistance, and meticulous care has to be paid to its execution.

Estimating community effectiveness from individual effectiveness and coverage

Once the individual effectiveness of an intervention has been measured by a case–control study and coverage is determined, it is easy to use eq. (1) to estimate the CPE (see section Defining effectiveness).

In order to obtain the attributable risk (or the absolute number of deaths averted by the intervention), we need first to obtain an independent estimate of the mortality rate in the unprotected target population. This can be derived from different sources, such as a cross-sectional retrospective survey at the start of the programme or from available national data. Although these data need to be accurate, they do not need to be very precise. The mortality rate in the protected group can then be inferred by multiplying the mortality rate in the unprotected group by the risk ratio (which is defined as: RR = (100–CPE)/100). The difference between the estimates of mortality in the protected and unprotected groups gives the attributable risk.

Summary and conclusions

The main focus of this article is to outline the reasons for a continuous assessment of the impact of ITBC beyond randomized controlled trials (phase III assessment) and to discuss the methodological problems associated with phase IV assessment. The most powerful rationale for monitoring impact under programme conditions is that the efficacy measured under trial conditions is unlikely to be as sustained as effectiveness under programme conditions. This was demonstrated recently in the Gambia, where a three-fold reduction in the cost-effectiveness ratio was observed under programme conditions. Another powerful argument in favour of ongoing impact monitoring is the need to obtain information on the long-term impact of ITBC interventions.

Measuring community effectiveness directly represents an ideal reference standard. However, a
review of the issues involved (e.g., the time required for such an assessment and the nature of available control groups), as well as a discussion of existing methodologies, revealed that it is impractical.

A simple effectiveness model, allowing for differentiation between individual and community effectiveness, is an alternative. It involves estimating individual effectiveness through a case–control study and combining this estimate with a simple coverage indicator. This should be both adequate and feasible in most programme situations and provides programme managers with a powerful impact assessment tool, for ITBC as well as for other health interventions at the community level.

Acknowledgements
Don de Savigny, Anne Mills, Peter Smith, and Marcel Tanner are thanked for their comments on the manuscript. Christian Lengeler is supported by grants from the Overseas Development Administration (United Kingdom) and the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases. Bob Snow is funded as part of the Wellcome Trust's Senior Research Fellow Programme in Basic Biomedical Science. This article is published with the consent of the Director of the Kenya Medical Research Institute.

Résumé
De l'utilité à l'efficacité: moustiquaires de lit traitées par insecticide en Afrique
Les bons résultats obtenus récemment lors d'essais de moustiquaires de lit et rideaux imprégnés d'insecticide dans la réduction de la morbidité et de la mortalité dues au paludisme chez les enfants africains incitent fortement à adopter cette intervention à grande échelle. Cet article expose les raisons qui doivent conduire à poursuivre l'évaluation de l'impact de cette intervention au-delà de la phase des essais contrôlés randomisés (phase III) et examine les problèmes méthodologiques associés à une telle évaluation.

La justification majeure de la surveillance d'impact dans les conditions du programme (phase IV) réside dans le fait que l'utilité mesurée dans le cadre d'essais scientifiques est rarement retrouvée en termes d'efficacité dans les conditions réelles du programme. Ce phénomène a des répercussions profondes sur le rapport coût-éfficacité de l'intervention et doit être connu des décideurs en santé publique. Il est de plus nécessaire d'évaluer l'impact à long terme de l'intervention et cela n'est probablement possible que dans les conditions réelles du programme.

L'idéal serait de mesurer directement l'efficacité de l'intervention dans la communauté. Cependant, un examen critique des problèmes posés (temps nécessaire pour une telle évaluation et existence de groupes témoins) ainsi qu'une discussion des méthodologies existantes ont montré que cela était irréalisable.

Un modèle simple d'efficacité, permettant de distinguer l'efficacité au niveau individuel et communautaire, offre un cadre théorique utile pour l'évaluation de l'impact d'une intervention dans les conditions réelles du programme. Dans un premier temps, l'efficacité est mesurée par une étude cas-témoins. Cette estimation est ensuite associée à un indicateur de couverture approprié pour donner une estimation de l'efficacité au niveau communautaire. Cette approche donne aux directeurs de programme un outil puissant pour la surveillance de l'impact des interventions de santé publique au niveau communautaire.

References
C. Lengeler & R.W. Snow


