Does shortening the training on Integrated Management of Childhood Illness guidelines reduce effectiveness? Results of a systematic review

Final Report

Alexander K. Rowe
Malaria Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention (CDC), Atlanta, United States

Samantha Y. Rowe
Data Management Activity, Division of Parasitic Diseases and Malaria, Center for Global Health, CDC, Atlanta, United States

Kathleen A. Holloway
World Health Organization (WHO)—formerly Headquarters, Geneva, Switzerland, now Regional Office Southeast Asia, Delhi, India

Verica Ivanovska
Department of Medicines Policy and Standards, WHO, Geneva, Switzerland

Lulu Muhe
Department of Child and Adolescent Health and Development, WHO, Geneva, Switzerland

Thierry Lambrechts
Department of Child and Adolescent Health and Development, WHO, Geneva, Switzerland

WHO Library Cataloguing-in-Publication Data

A systematic review of the effectiveness of shortening Integrated Management of Childhood Illness guidelines training: final report / Alexander K. Rowe…[et al].

1.Integrated management of childhood illness. 2.Child health services - manpower. 3.Health personnel - education. 4.Training - methods. 5.Program evaluation. I.World Health Organization. II.Rowe, A.K. III.Rowe, Samantha Y. III.Holloway, Kathleen. IV.Ivanovska, Verica. V.Muhe, L. VI.Lambrechts, T.

ISBN 978 92 4 159721 0 (NLM classification: WA 320)

RELATED PRESENTATIONS AND PUBLICATIONS

Preliminary methods and results were presented at a WHO meeting.

A review of the effectiveness of shortening IMCI training. Presented by Alexander Rowe at the: Technical consultation on IMCI training approaches and review of pre-service training. Geneva, Switzerland, 19–23 November 2007.

Report of technical consultation on IMCI training approaches and pre service IMCI. Geneva, Switzerland, 19–23 November 2007. Available at Internet address: http://www.who.int/child_adolescent_health/documents/imci/en/index.html

Final methods and results in this report have been submitted for publication.

Rowe AK, Rowe SY, Holloway KA, Ivanovska V, Muhe L, Lambrechts T. Does shortening the training on Integrated Management of Childhood Illness guidelines reduce its effectiveness? A systematic review. *Health Policy and Planning* (in press)

EXECUTIVE SUMMARY

Background

The Integrated Management of Childhood Illness (IMCI) strategy has been shown to improve care for ill children in outpatient settings in developing countries. A central component of the strategy is an 11-day in-service training course for health workers on IMCI clinical guidelines. The 11-day course duration is recommended by the World Health Organization, which developed IMCI. In some countries, the course has been shortened to reduce training costs and the time health workers are away from their clinics during training. However, it not known whether shortening IMCI training reduces its effectiveness.

Methods

We conducted a systematic review to compare the effectiveness of the IMCI strategy that used standard in-service training (duration $\geq$11 days) versus shortened training (5–10 days). Studies were identified from a search of MEDLINE, two existing systematic reviews, and by contacting investigators and content experts. We included published or unpublished studies that: 1) compared standard versus short training (direct comparison studies) or compared IMCI-trained health workers versus health workers without IMCI training (indirect comparison studies), 2) reported quantitative measures of health worker practices related to managing ill children less than five years old in either public or private facilities, and 3) were conducted in low- or middle-income countries. As we found very few studies that directly compared standard with short training approaches, we also performed an indirect comparison by contrasting the effects of "standard training versus no IMCI" in one group of studies and "short training versus no IMCI" in a different group of studies. We also examined the effect of other interventions to support IMCI (e.g., extra supervision), the effect of IMCI over time for the two training approaches, the overall effect of IMCI training, and the absolute level of healthcare quality delivered to ill children after IMCI training. Outcomes abstracted from studies included direct measures of health worker behavior (e.g., tasks related to treatment or counseling) and patient knowledge of how to administer therapy at home. Two summary measures were analyzed: the median of effect sizes (MES) for all outcomes from a given study, and the percent of patients needing an oral antimicrobial or oral rehydration solution (ORS) who received these treatments according to IMCI guidelines ("patient treated according to IMCI guidelines", or PTIG). Studies were classified as having either a "first-tier" design (randomized controlled trials, pre-post studies with a control, or interrupted time series) or a "second-tier" design (pre-post studies without a control, post-only studies with a non-randomized control, or case-control studies). A main analysis included only studies with a first-tier design, and a sensitivity analysis included studies with either a first- or second-tier design. We focused on studies with at least one effect size based on $\geq$20 consultations per study group and time point.

Results

The search strategy identified 232 reports, 59 of which were included. The 59 reports presented results from 31 distinct studies. Of these, our primary analysis focused on the 29 studies with at least one effect size based on $\geq$20 consultations per study group and time point. A secondary analysis included all 31 studies.

Five (17%) of these 29 studies had a first-tier study design, including two direct comparison studies (standard versus short IMCI training) and three indirect comparison studies (three studies comparing standard IMCI training to no IMCI). Twenty-four (83%) of the 29 studies had a second-tier design, including one direct comparison and 23 indirect comparisons (16 studies of standard training and seven studies of short training).

Direct comparisons revealed little difference between standard and short training, with median effects from different analyses ranging from a small advantage of standard training (by 5 percentage points [%-points]) to a small advantage of short training (by 3 %-points) (median result was a 2 %-point advantage of standard training). In all indirect comparisons, effect sizes for standard training versus no IMCI were greater than short training versus no IMCI, with differences ranging from 9 to 23 %-points (median result was 16 %-points). Our best estimate, which was a range bounded by the median results of direct comparisons and indirect comparisons, was a 2 to 16 %-point advantage for standard training. Analyses of training duration as a continuous variable generally showed that IMCI's effect increased with longer duration (by 2 to 6 %-points per additional day of training), although at best these associations were of borderline statistical significance.

Results for IMCI's effect over time revealed no evidence of deteriorating performance after training. The influence of other interventions on IMCI's effect varied substantially, ranging from –9 to 42 %-points. None of these results was statistically significant.

As a broad summary of the effect of IMCI training (standard and short combined), an analysis of the 26 indirect comparison studies showed that IMCI training generally had a moderate effect (median MES increase of 19 %-points, and median PTIG increase of 27 %-points). Performance levels after training from direct and indirect comparison studies usually revealed considerable room for improvement. The median of study-specific medians of post-training outcome measures (for all outcomes) was 75%, and the median of post-training PTIG measures was 66%. The latter result means that, in general, even after IMCI training (and sometimes other supports), 34% of ill children needing an antimicrobial or ORS were not receiving these treatments according to IMCI guidelines.

Conclusions

There were too few direct comparisons of standard and short training with first-tier study designs to conclude firmly whether, and to what degree, shortening IMCI training reduces its effectiveness. A review across all comparisons suggested that the standard in-service IMCI training course is somewhat more effective than short training; although the magnitude of the difference is unclear, ranging from –3 to +23 %-points. Our best estimate was a difference of 2 to 16 %-points. Given that a sizable performance gap often exists after IMCI training, countries should consider implementing other interventions to support health workers after IMCI training, regardless of training duration. Such complementary interventions, however, should be selected with care, as some appear to be highly effective while others seem to confer little or no benefit.

INTRODUCTION

To reduce child mortality and improve child development in developing countries, the World Health Organization (WHO), the United Nations Children's Fund (UNICEF), and other technical partners developed the Integrated Management of Childhood Illness (IMCI) strategy [Gove et al., 1997]. IMCI has three components: improving case-management practices of health workers (especially in outpatient health facilities), strengthening health systems, and promoting community and family health practices. Seventy-six countries have reportedly scaled-up IMCI training beyond a few pilot districts [WHO, 2010]. Studies have demonstrated the strategy can improve healthcare quality at health facilities [Armstrong Schellenberg et al., 2004a; Amaral and Victora, 2008; Arifeen et al., 2009; Rowe et al., 2009a], although its effect on mortality is uncertain [Armstrong Schellenberg et al., 2004b; Arifeen et al., 2009; Rowe et al., in press].

To improve healthcare quality at outpatient health facilities, IMCI includes a set of evidence-based guidelines [Gove et al., 1997; WHO, 2005] for managing the leading causes of child deaths (pneumonia, diarrhea, and malaria) [Black et al., 2010]. WHO recommends implementing the guidelines through an 11-day in-service training course for health workers, a follow-up visit to health workers' facilities 1 month later to reinforce new practices, and job-aids (e.g., a chart booklet and wall chart of clinical algorithms, and a 1-page form for recording patient assessments, disease classifications, and treatments). WHO also recommends the following quality criteria for IMCI training: a ratio of participants to facilitators no more than 4 to 1; completion of all training modules; distribution of the IMCI chart booklet to each trainee to keep as a reference; a minimum of 30% clinical practice and 20 sick children managed by each trainee; and no more than 24 participants [Lambrechts et al., 1999].

Despite favorable results from IMCI evaluations and evidence from two countries that training costs in districts implementing IMCI are similar to districts without IMCI [Adam et al., 2005; Adam et al., 2008], concerns have been raised that the 11-day in-service training is too expensive and that it takes health workers away from their clinics for too long [Goga et al., 2009; WHO, 2003a]. In many countries, the response has been to shorten the course. A recent survey of 24 countries found that all offered shortened courses, typically lasting 5–8 days [Goga et al., 2009]. It is not known, however, whether shortening IMCI training reduces its effectiveness. As part of WHO's efforts to re-examine IMCI training strategies to identify ways to scale-up IMCI coverage rapidly, we conducted a systematic review to compare the effectiveness of the IMCI strategy that used the 11-day in-service training course (or courses that lasted slightly longer) versus shortened training. A secondary objective was to use the opportunity of the review to examine: 1) the effect of other interventions (in addition to IMCI in-service training) to strengthen health systems and support health worker adherence to IMCI guidelines, 2) the effect of IMCI over time for the two training approaches, 3) the overall effect of IMCI training, and 4) the absolute level of healthcare quality delivered to ill children after IMCI training.

METHODS

In preparing this report, Preferred Reporting Items for Systematic Reviews and Meta-Analyses [Moher et al., 2009] guidelines were followed. No formal protocol was prepared, although a short guidance document was written that described the methods.

Definitions

IMCI training was defined as in-service training that used IMCI materials, lasted ≥5 days, and covered enough of the standardized IMCI training content so that one could reasonably expect that health workers would be able to follow all IMCI guidelines as adapted to the country in which the study was conducted. Courses less than 5 days were considered too short to be of practical value, and the shortest courses that countries typically offer are 5 days long [Goga et al., 2009]. "Standard training" was defined as an in-service IMCI course with a duration of 11 days or more, and "short training" was an in-service IMCI course with a duration between 5 and 10 days. Definitions of the adequacy of study designs and several analysis-related terms are shown in Box 1. Study identification numbers (Study IDs) are labels representing all reports for a given study (see reference list organized by Study ID and Annex 1A).

Inclusion criteria

We included studies that: 1) investigated the effectiveness of the health facility component of the IMCI strategy by comparing standard versus short training (direct comparison studies) or compared IMCI-trained health workers versus health workers without IMCI training (indirect comparison studies); 2) reported quantitative results on measures of health worker practices related to managing ill children less than five years old, in either public or private health facilities; and 3) were conducted in a low- or middle-income country [World Bank 2005]. Published and unpublished studies were eligible for inclusion. No studies were excluded based on adequacy of statistical analysis or data collection method. Restrictions on the timing of studies and the language of study reports depended on the source (Table 1).

As outcomes measured on extremely small samples might be unreliable, we excluded outcome measures for a study group at a particular time point if they were based on <15 ill child consultations; and thus we excluded any study in which all outcomes had a measure based on <15 consultations. Studies of data collected from IMCI follow-up visits were excluded, as the IMCI trainers or supervisors who conducted the follow-up visits were present during data collection and presumably were actively trying to improve health worker practices. Post-only studies (i.e., performance only measured after IMCI implementation) without controls were excluded, as effect sizes cannot be directly estimated with this design.

Sources and search strategies

We searched five sources to identify relevant reports. The sources and search strategies are listed in Table 1. Although the searches were conducted in 2006 and 2007, several included reports were published after these years because our search had identified the study reports while still in draft form, and we followed-up with investigators to obtain final versions.

Data collection methods

Data on study outcomes and attributes of the IMCI courses were collected using slightly different methods. Regarding study outcomes, we focused on direct measures of health worker behavior (e.g., tasks related to treatment or counseling) and patient knowledge of how to administer treatments at home. Health outcomes (e.g., mortality rates) were not considered, as few studies reported them and it was too difficult to attribute changes in health outcomes to IMCI training. Outcome data for most studies were imported from a pre-existing database on medicine use that is supported by WHO and the International Network for Rational Use of Drugs (INRUD) (source 3 in Table 1) [WHO, 2009] and coordinated by an investigator of this review (KAH). For this database, one investigator (VI) abstracted information from study reports and entered it into a database (Microsoft Access, Microsoft, Inc., Redmond, Washington), and another investigator (KAH) reviewed the abstraction for accuracy. Before data from the WHO/INRUD database were imported, one investigator (SYR) checked the data against the original reports. For a small number of reports, discrepancies were identified; and in these cases, discrepancies were resolved through a consultative process.

For studies in the WHO/INRUD database, after the consultative process described above, results from the WHO/INRUD database were used, with five exceptions [Study IDs 5, 9, 11, 29, 30]. In four of these exceptions [Study IDs 5, 9, 29, 30], study groups, study areas, or outcomes were defined differently from what this review required because the purpose of our review was different from that of the WHO/INRUD database. In the remaining exception [Study ID 11], the measures for one outcome were slightly different between our review and the WHO/INRUD database because different publications were used. For studies not in the WHO/INRUD database, one investigator (SYR) abstracted the outcomes from study reports; and results from these studies were added to the WHO/INRUD database. Outcome definitions varied by study; however, most studies defined their outcomes according to WHO standard IMCI indicators [WHO, 2003b]. Details on the outcomes are available in Annex 1A.

Regarding the attributes of IMCI training courses, we collected information on: training quality; course duration; administrative level of country at which trainings occurred; types of health workers trained; proportion of children managed by an IMCI-trained health worker in geographic areas where IMCI was implemented[1]; time between IMCI training and evaluation; sample sizes of health facilities, health workers, and patients involved in the evaluation; and additional interventions that were used to strengthen IMCI implementation (e.g., extra supervision or job aids). Whenever possible, we collected sample size information on the data used to calculate effect sizes for each outcome of health worker performance. If outcome-specific sample sizes were not provided, we collected the study's overall sample size. If a study measured several outcomes, we collected information on the maximum sample size used to calculate effect sizes. Also, for studies on short training, we collected information on how the training was shortened (i.e., major changes in content or educational methods). A data manager abstracted the information from reports and entered it into a database (Microsoft Excel, Microsoft, Inc., Redmond, Washington), and an investigator (SYR) reviewed the abstraction for accuracy. If details on training attributes were not in a report, authors of the report were

[1] For example, in a district where IMCI is implemented, not all health workers might have received IMCI training. Thus, perhaps, only 80% of ill children were seen by IMCI-trained health workers.

contacted via E-mail using a standardized form letter. If authors did not respond to the initial E-mail in 2 days, they were contacted twice a week for 2 weeks (minimum of four attempts in all).

Analysis

Our primary analysis focused on studies with at least one effect size based on ≥20 consultations per study group and time point. This sample size criterion is the same as that used in another review [Rowe et al., 2009b].

The ideal design for a study to answer our question would involve a direct comparison of the performance of health workers who received short training (experimental group) versus standard training (comparison group). Unfortunately, only five studies with this design were identified [Quality Assurance Project, 2006; Surjono et al., 1998; and Study IDs 17, 20, 21]. One of these studies [Surjono et al., 1998] was excluded because it lacked quantitative measures of treatment quality, counseling by health workers, or patient knowledge on administering treatments; and another study [Quality Assurance Project, 2006] was excluded because it used simulated, rather than real, patients in its assessment of health workers' clinical skills. However, as there were other studies that compared IMCI-trained health workers to non-IMCI-trained health workers, we also performed an indirect comparison of standard versus short training by contrasting the effects of "standard training versus no IMCI" in a group of studies, and the effects of "short training versus no IMCI" in a different group of studies.

To perform both the direct and indirect comparisons, we faced several choices. First, we needed to decide which study designs to include. The two choices were to include only first-tier study designs, or include both first- and second-tier designs (Box 1). Studies with a second-tier design without a comparison (post-only study without controls) were not used. As illustrated in Figure 1, the advantage of only including first-tier designs is that results would be less susceptible to bias; but the disadvantage is that the analysis would have fewer studies, which means that results of the pooled analysis might be less representative of IMCI in the "target population" (i.e., all developing countries that are implementing or might implement IMCI) and might be more influenced by studies with atypically large or small effects. In contrast, if first- or second-tier designs are included, results might be more susceptible to bias, but potentially more representative and less influenced by outliers.

Second, as outcomes of health worker performance varied among studies, we needed a common metric that could be used across as many studies as possible. We developed two such "summary measures" (see Box 1 and Annex 1A for details). The first was the median effect size (MES), which was the median of effect sizes for all outcomes for a given comparison from a given study. MES reflected a study's "middle" effect size, as shown in Figure 2. MES was available for all studies, and has been used in other reviews [Jamtvedt et al., 2006; Rowe et al., 2009b]. The second summary measure was the percent of patients needing an oral antimicrobial or oral rehydration solution (ORS) who received these treatments according to IMCI guidelines ("patient treated according to IMCI guidelines", or PTIG), which was not available for all studies. PTIG was chosen in part because, as an indicator of correct treatment among children with a potentially life-threatening illness, it had clear clinical and public health relevance. Similar to the discussion above, each summary measure has advantages and disadvantages (Figure 1).

As shown in Figure 1, we decided to perform a main analysis of both MES and PTIG for studies with first-tier designs, and a sensitivity analysis of both summary measures for studies with first-tier or second-tier designs. Altogether, we performed four sets of analyses, each with advantages and disadvantages.

For each summary measure, we calculated effect sizes defined as the percentage-point (%-point) "difference of differences" (equation 1) [Ross-Degnan et al., 1997; WHO, 2001].

Equation 1: $\text{Effect size} = (\text{follow-up} - \text{baseline})_{\text{intervention}} - (\text{follow-up} - \text{baseline})_{\text{control}}$

Effect sizes were calculated such that a value greater than zero indicated an improvement in case management quality. For follow-up measurements, we had to consider the possibility that health worker performance might change over time after IMCI training—in particular, that performance might deteriorate. However, time between training and evaluation (i.e., "time since training") was a complex factor to account for because: 1) it generally varied among studies; 2) some studies measured outcomes at more than one time point since training; and 3) some studies had training that occurred relatively slowly such that at any given time point, when a cross-sectional survey was conducted, time since training varied substantially among enrolled health workers (i.e., there was no single time since training that represented all health workers well). To address these complexities, for most analyses, we used the one time point from each study that was furthest from IMCI training. In a set of secondary analyses, we analyzed time since training directly by using as many time points as possible from each study. For the analyses of time since training, for one study [Study ID 24] that used cross-sectional surveys in which no single time since training value existed that represented all health workers well, we accepted results from statistical models provided by investigators that estimated IMCI training effect sizes at several time points after training.

All analyses were performed with SAS version 9.1 (SAS Institute Inc., Cary, North Carolina). An exact, nonparametric test was used to test differences between medians, and linear regression modeling with the REG procedure was used to test for trends. As regression models based on extremely small samples might be unreliable, we conducted modeling only when analyses involved $\geq$10 effect sizes. Hypothesis testing was done at an alpha level of 0.05 and assumed each effect size had no uncertainty. Ideally, statistical tests should account for the uncertainty (standard error) of each effect size, which depends on sample size and correlation between outcomes of patients managed by the same health worker or at the same health facility ("clustering"). However, in this review, the degree of clustering (design effect or intraclass correlation coefficient) was not known for most studies. We considered p-values from 0.05–0.10 to be of "borderline" significance.

Effect sizes were stratified into three groups, which reflected the type of comparison: 1) effect sizes from studies that directly compared standard versus short IMCI training, 2) effects from studies of standard IMCI training versus no IMCI, and 3) effects from studies of short IMCI training versus no IMCI. Groups 2 and 3 were both needed for the indirect comparison described above. To identify and control for confounding, we further stratified effect sizes by time since training, and whether or not other interventions besides IMCI training were implemented to support health workers' adherence to IMCI guidelines (Table 2). In addition, as

standard and short training are a dichotomization of training duration, we analyzed MES and PTIG as a function of training duration as a continuous variable.

As part of the sensitivity analysis, to simultaneously account for potential confounders, multivariable linear regression modeling was performed on MES and PTIG using the REG procedure. For MES, six models were run (Tables 7a and 7b). Model 1 included training approach (standard or short), whether other interventions were implemented (yes or no), time since training (in months), study design (first-or second-tier), a "time since training x training approach" interaction term, and the "median baseline outcome measure" for each study. "Median baseline outcome measure" was defined as the median of all outcome measures at baseline for the intervention group, or in the case of no baseline values (i.e., post-only studies), the median of all outcome measures for the control group. We included this variable in the model because we hypothesized that if the same intervention were evaluated in two settings with different baseline performance levels, the study with the lower baseline performance level might find a larger effect size. Model 2 included the same factors as Model 1, except for the interaction term; and Model 3 included the same factors as Model 2, except for time since training, as this factor was not available for all studies. Models 4–6 were the same as Models 1–3, except that "training approach" (a dichotomous variable for training duration) was replaced by training duration (a continuous variable). For PTIG, sixteen models were run (Tables 8a and 8b). Six models were identical to the MES models; and ten additional models were run containing one, two, or three factors, as there were fewer studies with PTIG values and we were concerned that larger models would be over-specified.

RESULTS

Literature search and data abstraction

Three of our sources searched databases for eligible studies: source 1 retrieved 126 titles, source 3 retrieved 7,824 titles from the INRUD bibliography and hundreds more from a hand-search of many thousands of documents in the WHO archives, and source 4 retrieved 39,806 titles (Table 1). These titles were screened by staff working on the projects described in Table 1 (some of whom are also authors of this review). From this screening process, 232 reports were assessed for eligibility: 59 were included, 169 were excluded because they did not meet inclusion criteria, and four were excluded because they related to studies in which IMCI training lasted <5 days[2]. The 59 included reports presented results from 31 distinct studies; some studies had results in more than one report, and one report presented results for more than one study. Of these, our analysis focused on the 29 studies with at least one effect size based on $\geq$20 consultations per study group and time point (i.e., our primary analysis). One alternative analysis (described below), which involved effect sizes based on $\geq$15 consultations per study group and time point, included all 31 studies.

General descriptive results

The 29 studies were from 23 countries in six WHO regions (Figure 3, Annex 1A). Sample size information is provided in Annex 1B. Sixteen (55%) studies were published or dated in the past five years, and almost half (14/29, or 48%) were published in scientific journals (Figure 4). Five (17%) studies had a first-tier study design, and 24 (83%) had a second-tier design (Table 3).

Among the 29 studies, there were 31 comparisons (Box 1): 27 studies each had one comparison (one effect size, based on a comparison between two study groups, such as standard training versus no IMCI), and two studies each had two comparisons (three study groups, see footnote of Table 3).

Table 4a shows the number of studies and comparisons used in the main analysis and sensitivity analysis. Five studies had a first-tier study design, including two direct comparison studies (standard versus short IMCI training) and three indirect comparison studies (three studies comparing standard IMCI training to no IMCI). Twenty-four studies had a second-tier design, including one direct comparison and 23 indirect comparisons (16 studies of standard training, and seven studies of short training). Only half (15/29, or 52%) of studies used the PTIG outcome (Table 4b).

[2] These studies evaluated training courses that lasted four days (Uzochukwu BS, Onwujekwe OE, Ezeilo EA, Nwobi E, Ndu AC, Onoka C. Integrated management of childhood illness in Nigeria: does short-term training of health workers improve their performance? Public Health. 2008;122(4):367–70), three days (Tawfik Y, Nsungwa-Sabitii J, Greer G, Owor J, Kesande R, Prysor-Jones S. Negotiating improved case management of childhood illness with formal and informal private practitioners in Uganda. Tropical Medicine & International Health 2006;11:967–973), two days (Chakraborty S, D'Souza SA, Northrup RS. Improving private practitioner care of sick children: testing new approaches in rural Bihar. Health Policy and Planning 2000;15:400–407), and one day (Luby S, Zaidi N, Rehman S, Northrup R. Improving private practitioner sick-child case management in two urban communities in Pakistan. Tropical Medicine & International Health 2002;7:210–219).

Although training duration is categorized as either standard or short, there was considerable variation in actual course lengths (Figure 5). The median duration for standard and short training was 11 days and 7 days, respectively. In half of studies (15/29, or 52%), training occurred at the district level or lower, with training in other studies occurring at the national, state, or province level (Figure 6). Studies rarely reported information on the IMCI training quality criteria. However, among the 10 studies that provided any training quality data, all reported completion of all IMCI modules; and among 9 studies that reported the time spent on clinical practice, 7 reported the minimum of 30%, one reported daily practice, and one reported less clinical instruction than the 11-day course (Annex 2). Nurses and physicians were the most common health worker types included in studies (Table 5). Among the 18 studies for which a follow-up period could be determined (i.e., the maximum value of time since training), the median was 22.5 months (range: 1.5, 81.0).

Regarding how standard IMCI training was modified to shorten its duration, 8 of the 10 studies of short training mentioned changing at least one of the training methods, usually shortening it (Table 6); although in one study, computer training was added [Study ID 17]. One study reported changing the content of the training modules (as opposed to simply shortening the standard training modules). Investigators in China added materials related to early child development [Study ID 31].

In 26 studies, the training coverage in IMCI areas was available. The data for the large majority of these studies (23/26, or 88%) were analyzed such that the outcomes for ill children managed by IMCI-trained health workers were compared to those of ill children managed by non-IMCI trained health workers. Three (12%) of the studies [Study IDs 4, 6, 28] provided data stratified by IMCI area, but did not provide data stratified by IMCI training status. Thus, for these three studies, the analysis compared outcomes of ill children managed by a mix of IMCI-trained and non-IMCI trained health workers (in the IMCI area) to those of ill children managed by non-IMCI trained health workers (in the no-IMCI area) (Figure 7), which would probably underestimate IMCI's effect.

Nineteen (66%) of the 29 included studies used the same basic health facility survey methodology to assess health care quality. As recommended by WHO [WHO, 2003b], a sample of health facilities is selected and consultations of ill children are sampled within selected health facilities. Consultations are silently observed. Then, caretakers (usually the child's parent) are interviewed, and the child is re-examined to obtain a "gold standard" determination of the child's IMCI illness classifications. At the end of the day, health workers are interviewed and a health facility assessment is conducted to collect information on the availability of equipment and drugs. In one of these 19 studies [Study ID 28], the follow-up survey used all the above-mentioned steps of the standard WHO health facility survey tool, while the baseline survey did not (it only observed consultations and conducted a chart review and health facility assessment). The 10 studies that did not use the basic health facility survey methodology lacked several of the data collection components: 2 studies [Study IDs 27, 29] observed consultations and conducted caretaker and health worker interviews and health facility assessments (no re-examination), 4 studies [Study IDs 15, 22, 25, 26] observed consultations and conducted caretaker interviews only (no re-examination, health worker interviews, or health facility assessments), 3 studies [Study IDs 10, 12, 17] observed consultations only, and 1 study [Study ID 23] conducted a chart

review. Among the studies that reported sample size information, the majority of the studies involved >100 patients from >10 health facilities (Annex 1B).

The 29 studies reported on 42 distinct outcomes. Nearly all studies reported results for more than one outcome, and many studies used the same standard WHO IMCI indicators [WHO, 2003b] (Annex 1A). For these 42 outcomes, we abstracted 280 outcome measures (Box 1), and these outcome measures were used to calculate 117 effect sizes. The large number of outcome measures reflects the fact that outcomes (usually several per study) were measured in multiple studies for multiple study groups at multiple time points. The number of effect sizes was considerably smaller than the number of outcome measures because several (typically 2 or 4) outcome measures were used to calculate a single effect size (see equation 1). The number of outcomes included in the MES calculation varied considerably among the 29 included studies, ranging from 1–8 outcomes per study (median = 3). Figure 2 illustrates the variation in effect sizes that can be masked by summarizing a study with a single MES.

An analysis of the 26 indirect comparison studies showed that, before IMCI training, performance levels were very poor (Figures 8–9). Figure 9 shows that typically only one in five children needing an antimicrobial or ORS received these treatments according to IMCI guidelines. IMCI training (standard and short combined) generally had a moderate effect (median MES increase of 19 %-points [Figure 10a], and median PTIG increase of 27 %-points [Figure 11a]). Performance levels after training from direct and indirect comparison studies usually revealed considerable room for improvement. For all outcomes combined, the median of study-specific medians of post-training outcome measures was 75% (range 41, 97) (Figure 12a), and the median of post-training PTIG measures was 66% (range 11, 100) (Figure 13a). The latter result for PTIG means that, in general, even after IMCI training (and sometimes other supports), 34% of ill children needing an antimicrobial or ORS did not receive these treatments according to IMCI guidelines.

Stratifying effect sizes by whether or not other interventions besides IMCI training were implemented revealed no important differences. IMCI's effect on MES was slightly lower when other interventions were present (Figure 10b), and the effect on PTIG was slightly higher when other interventions were present (Figure 11b). Furthermore, median post-training measures for all outcomes combined (Figure 12b) or for PTIG (Figure 13b) were somewhat or slightly lower, respectively, when other interventions were present.

Studies that directly compared standard and short training

Three studies directly compared standard and short training. None were published in a scientific journal. The first was a pre-post study with a randomized comparison group from Uganda [Study ID 17]; although, as health worker performance was not measured before training[3], for the purpose of this review, the design was a post-only study with a randomized comparison group (a first-tier design). Standard 11-day IMCI training was compared to a 9-day course that included computer-based training, in addition to the usual discussions, video viewing, and clinical practice. Clinical officers and nurses were enrolled. The follow-up period (i.e., the time between training and evaluation) was relatively short (3–4 months). The second study was a post-only study with a randomized comparison group (a first-tier design) from Zambia [Study ID

[3] Instead of a performance outcome, health worker knowledge was measured before (as well as after) training.

20]. Standard 11-day IMCI training was compared to a 6-day course, which was developed by WHO's African Regional Office for training physicians. In this study, however, nurses, midwives, clinical officers, environmental technicians, and health information officers were enrolled (i.e., no physicians). The follow-up period was 4–6 months. The third study was a post-only study with a non-randomized comparison group (a second-tier design) from Kosovo [Study ID 21]. Standard 11-day IMCI training was compared to an 8-day course. Only physicians were enrolled. The follow-up period was 2–3 years.

Main analysis of the MES outcome (first-tier study designs only)

The following results should be interpreted with caution, as they are based on a small number of effect sizes ($\leq$6) from only five studies. The direct comparison of standard versus short training, which included only two effect sizes from two studies (Figure 14, left-hand column of data points), showed that standard training was slightly better than short training (by about 5 %-points; no statistical test performed because of small sample size).

No indirect comparison (i.e., "standard training versus no IMCI" compared to "short training versus no IMCI") was performed because no studies of first-tier design compared short training to no IMCI. For studies comparing standard training to no IMCI, an analysis of training duration as a continuous variable suggested an increasing effect size with increasing training duration (Figure 15). However, as this analysis included only four effect sizes from three studies, no statistical testing was performed. We then stratified the effect sizes by the presence of other interventions (in addition to IMCI training) to support IMCI implementation (Figure 16). We found that standard training with other interventions seemed much more effective than without other interventions (median MES with other interventions = 55.0 %-points, median MES without other interventions = 19.9 %-points — a difference of about 35 %-points).

An analysis of IMCI's effect over time revealed no evidence of deteriorating performance over time after standard training; although this observation was based on 9 effect sizes from only three studies (Figure 17a). In one of the three studies in this analysis [Study ID 24, with 2 comparisons], IMCI's effect seemed to be unchanged or decreasing over time (Figure 17b).

Main analysis of the PTIG outcome (first-tier study designs only)

The following results should be interpreted with caution, as they are based on a very small number of effect sizes ($\leq$5) from only four studies. The direct comparison of standard versus short training, which included only two effect sizes from two studies, showed that the effects of standard and short training were very similar (standard was better by about 3 %-points; no statistical test) (Figure 18).

An indirect comparison was not performed because no studies of first-tier design compared short training to no IMCI. For studies comparing standard training to no IMCI, an analysis of training duration as a continuous variable suggested an increasing effect size with increasing training duration (Figure 19). However, as this analysis included only three effect sizes from two studies, no statistical testing was performed. When we stratified the three effect sizes by the presence of other interventions, we found that standard training with other interventions seemed much more effective than without other interventions (median MES with other interventions = 61.1 %-points, median MES without other interventions = 19.1 %-points —

a difference of about 42 %-points) (Figure 20). An analysis of IMCI's effect over time after standard training revealed no trends; although this observation was based on 8 effect sizes from only two studies (Figures 21a and 21b).

Sensitivity analysis of the MES outcome

The direct comparison of standard versus short training, which included three effect sizes from three studies, showed that standard training was slightly better than short training (by about 2 %-points; no statistical test) (Figure 22).

The indirect comparison, which included 28 effect sizes from 26 studies, suggested that standard training was somewhat better than short training (by about 9 %-points; exact median test p = 1.0) (Figure 22). With this same group of studies, an analysis of training duration as a continuous variable revealed a non-significant trend of increasing effect size with increasing training duration (1.9 %-point increase per additional day of training, p = 0.15) (Figure 23). We then stratified the 28 effect sizes by the presence of other interventions and found, regardless of whether other interventions had been implemented, standard training appeared somewhat better than short training by about 6 %-points (Figure 24; 20.9 versus 15.0 %-points when no other interventions present; 16.6 versus 10.7 %-points when other interventions present). For either standard or short training, the effect of other interventions was about –4 %-points (16.6 versus 20.9 %-points for standard training; 10.7 versus 15.0 %-points for short training).

An analysis of IMCI's effect over time after training revealed no trends for either standard training (p = 0.30) or short training (only 4 effect sizes, no statistical testing) (i.e., neither trend appeared different from a slope of zero; Figures 25a and 25b). Statistical modeling (Table 7a, Model 1, described below) found no difference between the two slopes (interaction p = 0.43).

In contrast to the isolated results from the above descriptive and univariate analyses, results of multivariable linear regression modeling revealed a more complex network of associations. In Tables 7a and 7b, Models 1–6 examined the effects of training duration, other interventions, study design adequacy, baseline performance levels, and time since training. Models 1–3 analyzed training duration as a dichotomous variable, and Models 4–6 analyzed training duration as a continuous variable. Both showed generally similar trends; but as it seemed logical to analyze training duration as a continuous variable, Models 4–6 are discussed in detail.

Model 4 results suggested that the relationship between IMCI's effect size and training duration does not depend on time since training (the "time since training x training duration" interaction term was not statistically significant, p = 0.48). Model 5, which was identical to Model 4, except the interaction term was dropped (and thus the effect of training duration was averaged across all values of "time since training"), suggested that IMCI's effect increased with longer training duration (2.5 %-point increase per extra day of training, p = 0.11). Model 6 showed no relation between IMCI's effect and training duration (p = 0.51); this model was different from Models 4 and 5 because it excluded time since training and included four additional studies for which time since training could not be analyzed. Models 4, 5, and 6 suggested that IMCI's effect decreased as baseline performance levels increased, although this finding was not statistically significant (effect size decreased 0.3 to 0.4 %-points per additional 1

%-point increase in baseline performance level; p-values ranged from 0.11 to 0.13). Furthermore, Models 4–6 showed that there was no significant difference between training with interventions versus training without other interventions (–2 to 8 %-point difference, p-values ranged from 0.23 to 0.80).

Sensitivity analysis of the PTIG outcome

The direct comparison of standard versus short training, which included only three effect sizes from three studies, showed that the effects of standard and short training were very similar (short was better by about 3 %-points; no statistical test) (Figure 26).

The indirect comparison, which included 14 effect sizes from 12 studies, suggested that standard training was much better than short training (by about 23 %-points); although the result was not statistically significant (exact median test p = 0.19) (Figure 26). With this same group of studies, an analysis of training duration as a continuous variable revealed a borderline significant trend of increasing effect size with increasing training duration (5.6 %-point increase per extra day of training, p = 0.09) (Figure 27). Stratification by the presence of other interventions showed that, regardless of whether other interventions had been implemented, standard training appeared substantially better than short training (Figure 28; by 30 %-points [36.2 versus 6.0 %-points] when no interventions present; by 23 %-points [38.0 versus 15.3 %-points] when other interventions present). For standard training, the effect of other interventions was about 2 %-points (38.0 versus 36.2 %-points); and for short training, the effect of other interventions was about 9 %-points (15.3 versus 6.0 %-points).

An analysis of IMCI's effect over time suggested a trend of increasing effect after standard training (p = 0.17) and short training (only three effect sizes, no statistical testing; also, effect sizes were relatively small, ranging from –10 %-points to +22 %-points) (Figures 29a and 29b).

Attempts to perform multivariable modeling were limited by the small number of effect sizes and studies. Thus, models with more than two variables should probably be considered unreliable (i.e., Models 1–5 and 9–13). In Table 8a, Models 6–8 suggested that standard training was better than short training (by about 24 to 33 %-points), although at best these associations were only of borderline statistical significance ($p \geq 0.08$). Results from Models 14–16 were similar, showing a borderline significant trend that IMCI's effect increased with training duration (4.9 to 6.3 %-point increase per extra day of training, p-values ranged from 0.09 to 0.11) (Table 8b). Additionally, Models 6 and 14 showed that IMCI's effect decreased as baseline performance levels increased, although at best this finding was borderline significant (effect size decreased 0.8 to 0.9 %-points per additional 1 %-point increase in baseline performance level; p-values ranged from 0.09 to 0.15). Furthermore, Models 7 and 15 showed that there was no significant difference between training with interventions versus training without other interventions (–9 to –5 %-point difference, p-values ranged from 0.58 to 0.74).

Alternative analyses

We investigated the impact of tightening and relaxing the sample size inclusion criteria in two alternative sensitivity analyses. In the first analysis, we excluded one outlier [Study ID 25] because it involved only two health facilities and unusually high effect sizes (Annexes 1A, 1B).

In the second analysis, effect sizes based on ≥15 consultations per study group and time point were included, which allowed two studies to be added [Study IDs 32, 33]. Results from both analyses were similar to those from the primary analysis (primary analysis summary on Table 9, alternate analyses results in Annexes 3 and 4 of the report).

Cost of training approaches

Few studies reported training costs, although the direct comparisons studies did. In Zambia, standard training was US$828 per trainee, and short training was US$450 per trainee (46% lower) (Mwinga et al., 2006). In Uganda, standard training was US$472 per trainee, and short training was US$335 per trainee (29% lower, assuming no costs for computers) or US$410 per trainee (13% lower, assuming computers were rented) (Tavrow et al., 2002). In Kosovo, standard training was US$430 per trainee, and short training was US$240 per trainee (44% lower) (Syla, 2003).

From other studies, costs of standard training per trainee were as follows: US$291 in Bolivia (personal communication, T. Lambrechts, WHO, November 6, 2007), US$793 in Kenya (Quality Assurance Project, 2006), between US$730 and US$811 in Morocco (Naimoli 2001), and US$850 in Benin (Rowe et al., 2009a).

DISCUSSION

Our objective was to answer the question: does shortening the in-service IMCI training course reduce its effectiveness? The review was complicated by the fact that only two ideal studies were identified (i.e., studies with a first-tier design that directly compared standard versus short training). However, we did find one study that directly compared standard versus short training with a second-tier design, as well as numerous studies comparing IMCI to no IMCI, which permitted an indirect comparison of the two training approaches. Although we could have been purists and included only ideal studies, we decided to examine a broader range of studies. In so doing, to the best of our knowledge, this review represents the most comprehensive examination of IMCI effectiveness on health worker performance to date. We performed a series of layered analyses to show results based on more and less restrictive inclusion criteria. Additionally, as studies used a variety of outcomes, we estimated effects for two summary measures, each with advantages and disadvantages.

Based on the available evidence, standard in-service IMCI training seemed more effective than short training, although the magnitude of the difference is unclear and might be small. Differences (i.e., standard training minus short training) from the four analyses of direct comparisons ranged from –3 to +5 %-points[4] (median = 2 %-points) across all analyses of the two summary measures, and differences from two analyses of indirect comparisons ranged from 9 to 23 %-points[5] (median = 16 %-points). Indirect comparisons that examined training duration as a continuous variable suggested that the effect of IMCI training increases by 2 to 6 %-points[6] per additional day of training, although at best these associations were of borderline statistical significance. To capture the complexity of these results, our best estimate of the difference between standard and short training was a range bounded by the median values of the direct and indirect comparisons—i.e., a range of 2 to 16 %-points.

Are these modest differences relevant? In even a small country (e.g., population 12 million, with 2 million children <5 years old) with a plausible disease burden (3 illness episodes per child per year) and low access to health facility-based services (e.g., 25% of ill children are brought to health facilities), a difference of 2–16 %-points in health worker performance translates into an additional 30,000–240,000 children treated correctly per year (i.e., 2 million children x 3 illnesses per child per year x 25% access x 2–16% improved treatment). Even if only 5% of these illnesses are life-threatening and the protective efficacy of correct treatment to prevent death is only 78% [Lindblade et al., 2007], a 2–16 %-point difference could prevent 1000–9000 child deaths per year—a considerable impact[7].

---

[4] The four differences are: –2.7 %-points (sensitivity analysis of PTIG), 2.1 %-points (sensitivity analysis of MES), 2.7 %-points (main analysis of PTIG), and 5.0 %-points (main analysis of MES).

[5] The two differences are: 8.7 %-points (sensitivity analysis of MES), 22.7 %-points (sensitivity analysis of PTIG).

[6] The slopes (and their p-values) from univariate linear regression models with training as a continuous variable are: 1.9 (p = 0.15) (sensitivity analysis of MES) and 5.6 (p = 0.09) (sensitivity analysis of PTIG).

[7] This calculation is not meant to imply that a policy of standard training will save more lives than a policy of short training because it does not account for the lower cost of short training. For example, while health worker performance might be somewhat lower with short training, the lower cost of short training could allow more health workers to be trained and therefore more ill children to be managed by IMCI-trained health workers and more lives saved.

Sensitivity analyses of indirect comparisons revealed an advantage of the standard training over the short training for both summary measures, but especially for PTIG. This suggests the standard training might be more effective than short training in improving behaviors that are harder to change. For example, prescribing all necessary drugs with the correct dose, as in PTIG, might be harder to change than other behaviors.

Although assessing the effectiveness of other interventions intended to support IMCI was not our primary focus, the analyses did provide some interesting insights. First, in several analyses, IMCI's effect with other interventions was greater than without other interventions; however the median effects of other interventions from different analyses varied substantially, ranging from –9 to +42 %-points. Second, an examination of individual studies with other interventions revealed two with particularly large effect sizes; both had first-tier study designs. The study with the largest PTIG effect was from Bangladesh [Study ID 1], and the other interventions involved purchasing essential IMCI drugs and supplies and monthly supervision by IMCI-trained medical officers (Table 2). The study with the second largest PTIG effect was from Benin [Study ID 24], where investigators implemented a package of relatively inexpensive post-IMCI-training supports that included increased supervision, supervision of supervisors, job aids, and non-financial incentives. Over a 3-year period, among consultations performed by IMCI-trained health workers with the study supports, the proportion of ill children who received IMCI-recommended treatments was 27 %-points higher than for IMCI-trained health workers in a comparison area with "usual" supports.

This review also provided an opportunity to examine baseline and post-IMCI-implementation values for our two outcomes. At baseline (before IMCI training), median outcome measures and PTIG measures were generally very low. Perhaps these results should not be surprising because health workers had not yet been trained. However, if IMCI guidelines reflect a minimum level of care recommended by WHO in low-resource settings, then it should be quite concerning that, in the absence of IMCI, health systems in many countries seemed to be typically providing grossly inadequate care for ill children. Regarding post-IMCI performance levels, it is also concerning that, in general, even after IMCI training (and sometimes other supports), 34% of ill children needing an antimicrobial or ORS were not receiving these treatments according to IMCI guidelines. This result illustrates the challenge of getting health workers to follow clinical guidelines, and it underscores the importance of providing ongoing support for health workers after training.

Limitations

In summarizing results of this review, we tried to make the best use of all existing data, which required a variety of analytic approaches, all of which had important limitations (Figure 1). The heterogeneity in IMCI implementation, study designs, and outcomes precluded a meta-analysis, which would have been preferable to our "quantitative summary". Specific limitations are as follows. First, as previously mentioned, we identified only two ideal studies; and one of these (the Uganda study [Tavrow et al., 2002]) shortened training by only 2 days. Second, the main analysis, which was based on all studies with first-tier designs, included very few studies. Third, the indirect comparisons were susceptible to bias caused by differences between standard and short training studies, besides training duration. For example, some studies might have implemented training that did not meet all of the quality training criteria, which might have

caused effect sizes to be lower. Data on training quality were scarce, and it is unclear to what degree the trainings were comparable. Additionally, on average, one group of studies might have been from settings with relatively better infrastructure, which might have caused effect sizes to be greater. Also, one group of studies might have been from settings with well-performing health workers at baseline, which might have caused effect sizes to be smaller. Reliable measurement and quantitative adjustment of non-training factors was challenging. We attempted to account for two such factors (the presence of other interventions to support IMCI and baseline performance levels); however, our methods were simplistic. Fourth, the summary measure MES included different outcomes in different studies. Thus, comparisons based on the MES might have been susceptible to bias because some indicators of performance might be easier to improve than others.

A fifth limitation was that when comparing training approaches, we often did not test for statistical significance because of the small number of effect sizes, or we used tests that did not account for the variance of each effect size. An ideal analysis would have estimated effect size differences between standard and short training (e.g., with a random-effects model) and would have estimated a confidence interval around the differences. However, such an approach would have required the actual study datasets, which would have been too difficult to obtain, and thus was simply impractical. Sixth, when we did perform statistical testing, we found that almost none of the results were significant; and the few results that were significant close to the 0.05 level appeared in the course of numerous comparisons. If one applied even a modest reduction of the threshold for determining significance (e.g., $p < 0.01$) to adjust for multiple comparisons, most significant findings would disappear. Thus, it seems that there was little convincing statistical evidence of a difference between training approaches. With that said, one must bear in mind that each data point represented an entire study, with most studies involving over 100 patients. Seventh, only about two-thirds of studies used the "gold standard" determination of the child's IMCI illness classifications. Without such a standard, the effect of training might be overestimated depending on the correctness of the health worker's ability to classify illness. Finally, in the analysis of IMCI's effect over time, there was insufficient evidence to conclude that the effect increased or decreased over time since training for either standard or short training. However, "time since training" values were missing or might have been inaccurate for some studies. For example, if the training was rolled out over a long time period such that groups of health workers received the training at different times, the outcomes measured at one point in time after training would reflect a variety of "times since training". Furthermore, in this analysis, many studies reported results for just one time point. Such results can only be meaningfully combined if one assumes study settings are similar (i.e., no confounding by site-specific characteristics), which might be a weak assumption.

Summary

- All direct comparisons showed little difference between standard and short training, with median effects from different analyses ranging from –3 to +5 %-points (median = 2 %-points) (Table 9).
- In all indirect comparisons, effect sizes for standard training versus no IMCI were greater than short training versus no IMCI, with differences ranging from 9 to 23 %-points (median = 16 %-points).

- When we analyzed training duration as a continuous variable, we generally found that IMCI's effect increased with longer training duration (by 2–6 %-points per additional day of training).
- Our best estimate was a 2–16 %-point advantage for standard training.
- There was insufficient evidence to conclude that IMCI's effect increased or decreased over time since training for either standard or short training.
- In the three studies that compared the costs of the two training approaches, the cost of short training was 13–46% lower than standard training.
- The influence of other interventions on IMCI's effect varied substantially, with median effects of the other interventions ranging from –9 to +42 %-points.
- Post-IMCI-implementation outcome values show that, even after IMCI training (and sometimes other supports), 34% of ill children needing an antimicrobial or ORS did not receive these treatments according to IMCI guidelines.

Conclusions

Four broad conclusions can be drawn from this analysis. First, there were too few direct comparisons of standard and short training with first-tier study designs to conclude firmly whether shortening IMCI training reduces its effectiveness (and if so, to what degree effectiveness is reduced). Additional direct comparisons with first-tier designs would be needed to answer the question definitively. While such studies would be welcome, they might be difficult to realize because countries usually do not implement two different types of IMCI training at the same time. Moreover, program implementers and donors often resist the use of first-tier study designs.

Second, standard in-service IMCI training seemed more effective than short training, although the magnitude of the difference is unclear, ranging from –3 to +23 %-points. Our best estimate was a difference of 2–16 %-points.

Third, all three direct comparison studies found that shortening IMCI training reduces costs. Direct training costs for short training were 13–46% lower than standard training, health workers were away from their clinics for a shorter time, and presumably IMCI course facilitators spent less time on training.

Fourth, post-IMCI-implementation outcome values show that even after IMCI training, considerable room for improvement exists. Although this review focused on training duration, an equal (or greater) consideration should be given to designing and implementing other interventions to support health workers after IMCI training, regardless of training duration. Results from this and other reviews [Ross-Degnan et al., 1997; Rowe et al., 2005; Rowe et al., 2009b; WHO, 2001; WHO, 2009] clearly show that not all such interventions have the same effect, and therefore we recommend strongly that additional research be conducted to identify effective and affordable interventions to improve and maintain health worker performance in low-resource settings. Similarly, research is needed on the effectiveness of pre-service training on IMCI guidelines, as such an approach might render moot concerns about the difference between shorter versus longer in-service training duration.

Policy implications

Based on limited evidence, standard training seemed more effective than short training, although the difference might be small. Where possible, standard training is recommended. However, we acknowledge that in some settings only shortened training is feasible. When countries need to shorten training, WHO recommends that "core competencies" (i.e., those addressing major causes of deaths for which effective interventions exist) should be retained [WHO, 2007]. In all circumstances, it is critical to implement strategies to strengthen health systems in addition to training in IMCI.

ACKNOWLEDGMENTS

We thank Karen Herman and Chunying Xie for their excellent assistance with the literature review. We also thank the Ministries of Health of Cambodia, Ecuador, Eritrea, Ethiopia, Niger, Peru, South Africa, Uzbekistan, Vietnam, and Zambia for giving us permission to include data from unpublished health facility survey reports in our database. We greatly appreciated the thoughtful comments provided by Rajiv Bahl, Jose Martines, and the participants of the *Technical consultation on IMCI training approaches and pre-service IMCI*, which was convened by the WHO, Geneva, Switzerland, November 19–23, 2007.

SOURCE OF FUNDING

S.Y. Rowe and K. Herman were partly supported by WHO's Department of Child and Adolescent Health and Development.

REFERENCES

Adam T, Manzi F, Armstrong Schellenberg J, et al. Does the Integrated Management of Childhood Illness cost more than routine care? Results from the United Republic of Tanzania. Bulletin of the World Health Organization 2005;83:369–77.

Adam T, Edwards SJ, Amorim DG, et al. Cost implications of improving the quality of child care using integrated clinical algorithms: Evidence from Northeast Brazil. Health Policy 2009;89(1):97–106

Amaral J. Manejo de casos da AIDPI nas unidades de saúde no Brasil — 2002. 2002. Unpublished report.

Amaral J, Gouws E, Bryce J, et al. Effect of Integrated Management of Childhood Illness (IMCI) on health worker performance in Northeast-Brazil. Cadernos de saude publica/Ministerio da Saude, Fundacao Oswaldo Cruz, Escola Nacional de Saude Publica 2004;20 Suppl 2:S209–19.

Amaral JJ, Victora CG. The effect of training in Integrated Management of Childhood Illness (IMCI) on the performance and healthcare quality of pediatric healthcare workers: a systematic review. Revista Brasileira de Saúde Materno Infantil 2008;8(2):151–61.

Anonymous. Implementación de la estrategia de Atención Integrada a las Enfermedades Prevalentes de la Infancia, Reporte Final, Linea Base, Provincia de Imbabura, Republica del Ecuador. 1997. Unpublished report.

Anonymous. Informe Preliminar: Evaluación AIEPI a Servicios de Salud, Primera Prueba Mundial, Bolivia, 12–30 de Abril 1999. Unpublished report.

Anonymous. Evaluación del proceso de implementacion de la estrategia de Atención Integrada a las Enfermedades Prevalentes en la Infancia - AIEPI - en ambitos de las Direcciones de Salud Callao y Lima Norte (Ventanilla y Puente Piedra), 2004. Unpublished report.

Arifeen SE, Bryce J, Gouws E, et al. Quality of care for under-fives in first-level health facilities in one district of Bangladesh. Bulletin of the World Health Organization 2005;83:260–7; and personal communication from S. El Arifeen, International Centre for Diarrhoeal Diseases Research, Bangladesh, June 13, 2007.

Arifeen SE, Hoque DME, Akter T, et al. Effect of the Integrated Management of Childhood Illness strategy on childhood mortality and nutrition in a rural area in Bangladesh: a cluster randomised trial. Lancet 2009;374:393–403.

(Armstrong Schellenberg, 2004a) Armstrong Schellenberg J, Bryce J, de Savigny D, et al. The effect of Integrated Management of Childhood Illness on observed quality of care of under-fives in rural Tanzania. Health Policy and Planning 2004; 19(1):1–10.

(Armstrong Schellenberg, 2004b) Armstrong Schellenberg JR, Adam T, Mshinda H, et al. Effectiveness and cost of facility-based Integrated Management of Childhood Illness (IMCI) in Tanzania. Lancet 2004;364:1583–94.

Assimadi J, Agbèrè AD, Alihonou F, et al. [Management of childhood illness in children under 5 years of age before implementation of an integrated strategy in Togo] Archives de pédiatrie, 1999;6(10):1135–6.

Assimadi JK, Lawson-Evi K, Atakouma DY, et al. [Evaluation of payment of health personnel in an integrated management of childhood diseases program during the first year of its functioning] Archives de pédiatrie, 2003;10(2):158–9.

Atakouma Dzayissé Y, Tsolenyanu E, Gbadoe A, et al. [Usefulness of integrated management of childhood illness for caring children less than 5 years in peripheral health centres at Lome (Togo)] Archives de pédiatrie, 2006;13(12):1552–3.

Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet, 2010; 375: 1969–87.

Briggs J, Gabra M, Ickx P. Senegal Assessment: Drug Management for Childhood Illness. Rational Pharmaceutical Management Plus Program, Management Sciences for Health. June 2002.

Bryce J, Gouws E, Adam T, et al. Improving quality and efficiency of facility-based child health care through Integrated Management of Childhood Illness in Tanzania. Health Policy and Planning 2005;20 Suppl 1:i69–i76.

Burnham GM. Evaluation of Integrated Management of Childhood Illness (IMCI) Performance in Urban Health Centers, Lusaka, Zambia, August 10–23, 1997. BASICS, 1997.

Centers for Disease Control and Prevention. Health worker performance after training in integrated management of childhood illness—Western Province, Kenya, 1996–1997. Morbidity Mortality Weekly Report 1998;47(46):998–1001.

Choi M, Salgado R, Mehari M, Wendo D. Assessment of Child Health Services in Eritrea After Three Years of Implementation of the Integrated Management of Childhood Illness (IMCI) Strategy. John Snow International, Inc., 2002. Unpublished report.

Chopra M, Patel S, Cloete K, et al. Effect of an IMCI intervention on quality of care across four districts in Cape Town, South Africa. Archives of Disease in Childhood 2005;90:397–401.

Cordero D, Salgado R, Drasbeck C. An analysis of the IMCI implementation process in four countries of Latin America. Edited by Tina G. Sanghvi and Mireille Mather. Arlington, Virginia: BASICS II for the United States Agency for International Development and the Pan American Health Organization. 2004.

Degbey H. Enquête dans les formations sanitaires mettant en oeuvre la PCIME au Niger (15 mars - 15 avril 2005). Lettre de l'OMS Niger [Internet] 2005 [cited September 15, 2010]; 40: 6-7. Available from: http://www.santetropicale.com/alire/omsnig0605.pdf

El Arifeen S. Section 4: Health Facility Survey. 2000; pp. xxv–xxxiv. Unpublished report.

El Arifeen S, Blum LS, Hoque DM, et al. Integrated Management of Childhood Illness (IMCI) in Bangladesh: early findings from a cluster-randomised study. Lancet 2004;364:1595–602.

Federal Ministry of Health of Sudan. Health Facility Survey on Quality of Outpatient Child Health Services, IMCI Health Facility Survey, Sudan, March–April 2003. Eastern Mediterranean Regional Office of the World Health Organization, 2004.
Document number: WHO-EM/CAH/013/E/G/02.04/500. Available at: http://www.emro.who.int/cah/pdf/imcisurvey/sud/report03.pdf, accessed October 29, 2008.

Gilroy K, Winch PJ, Diawara A, et al. Impact of IMCI training and language used by provider on quality of counseling provided to parents of sick children in Bougouni District, Mali. Patient Education and Counseling 2004;54:35–44.

Goga A. IMCI health facility survey: Eastern Cape and North West Provinces, 2003. Final draft. September 20, 2003. South Africa National Department of Health (plus personal communication from A. Goga, Medical Research Council of South Africa, on July 13, 2009, which included supplementary tables).

Goga AE, Muhe LM, Forsyth K, et al. Results of a multi-country exploratory survey of approaches and methods for IMCI case management training. Health Research Policy and Systems 2009;7:18.

Gouws E, Bryce J, Habicht JP, et al. Improving antimicrobial use among health workers in first-level facilities: results from the multi-country evaluation of the Integrated Management of Childhood Illness strategy. Bulletin of the World Health Organization 2004;82:509–15.

Gove S for the WHO Working Group on Guidelines for Integrated Management of the Sick Child. Integrated management of childhood illness by outpatient health workers: technical basis and overview. Bulletin of the World Health Organization 1997;75 (supplement 1):7–24.

Health Care Provider Performance (HCPP) Review Study Group. Protocol for a systematic review of the effectiveness and costs of interventions to improve health care provider performance and related health outcomes in low- and middle-income countries. October 16, 2008. Available from Alexander Rowe, U.S. Centers for Disease Control and Prevention, Atlanta, Georgia (E-mail: axr9@cdc.gov).

Huicho L, Davila M, Campos M, et al. Scaling up integrated management of childhood illness to the national level: achievements and challenges in Peru. Health policy and planning 2005;20:14–24.

Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. 2006. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD000259.

Kelley, E., et al., The impact of QA methods on compliance with the Integrated Management of Childhood Illness algorithm in Niger. Operations Research Results 2000;1(2): p. 16.

Kelley E, Geslin C, Djibrina S, et al. Improving performance with clinical standards: the impact of feedback on compliance with the integrated management of childhood illness algorithm in Niger, West Africa. The International journal of health planning and management 2001;16:195–205.

Kelley, E., et al., Compliance, workload, and the cost of using the Integrated Management of Childhood Illness algorithm in Niger. Operations Research Results 2002;2(11): p. 15.

Lambrechts T, Bryce J, Orinda V. Integrated Management of Childhood Illness: a summary of first experiences. Bulletin of the World Health Organization 1999;77:582–594.

Lindblade KA, Hamel MJ, Feikin DR, Odhiambo F, Adazu K, Williamson J, Vulule JM, Slutsker L. Mortality of sick children after outpatient treatment at first-level health facilites in rural western Kenya. Tropical Medicine and International Health 2007;12:1258–1268.

Lee L, Odhacha A, Hamel MJ, et al. Effectiveness of in-service training and follow-up supervision on correct integrated management of childhood illness (IMCI) in Western Kenya. Unpublished manuscript.

Mbuya C, Mgalula L, Kasale H, et al. IMCI Implementation: A Report on Experiences in Morogoro and Rufiji Districts in Tanzania, Preliminary Report. 2003.

Mehari M, Wendo D, Ghebrat Y, Salgado R. Preparedness of Eritrean Health Services for the Integrated Management of Childhood Illness (IMCI) Strategy. 2000. Unpublished report.

Mgalula, L. Activities and Methods, Substudy 3: Health Facility Survey. 2000; pp. 5–19. Unpublished report.

Ministry of Health of Ecuador. Evaluación de servicios de salud sobre "La Atención Integrada a las Enfermedades Prevalentes de la Infancia AIEPI". 2000. Unpublished report.

Ministry of Health of Niger. Enquête dans les formations sanitaires mettant en œuvre la PCIME. 2005. Unpublished report.

Ministry of Health of Perú. Evaluación de Servicios de Salud. Segunda Prueba Mundial, Perú, 13–24 Octubre 1999. Ministerio de Salud, Dirección General de las Personas, Comisión Coordinadora AIEPI – MINSA Perú, Programa de Salud Básica Para Todos, Proyecto 2000, OPS/OMS. 2000. Unpublished report.

Ministry of Health of Vietnam, National Institute of Hygiene and Epidemiology of Vietnam. Official Travel Report, Annex 1: IMCI health facility survey in Vietnam; Planning Workshop, Hanoi, June 17–21, 2002. World Health Organization, Western Pacific Regional Office, Reference number MR/2002/0289 (plus personal communication from T. Lambrechts, WHO, on July 1, 2009 and December 9, 2009, which included data summary tables).

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Medicine 2009;6:e1000097.

Mwinga K, Siziya S, Kasungami D, et al. Research Study Report: Comparative study of Integrated Management of Childhood Illness (IMCI) case management skills of primary health care workers trained in eleven-day standard course and six-day abridged course for physicians. Unpublished report from the Zambia Ministry of Health, 2006

Naimoli, JF. The Integrated Management of Childhood Illness (IMCI) pilot project in Morocco, Preliminary results from an evaluation of the management of childhood illness in public sector IMCI and non-IMCI facilities in four Moroccan provinces, April 2000. 2000. Unpublished report.

Naimoli JF. Theoretical and empirical advances in research on the implementation of an integrated approach to managing childhood illness in outpatient facilities in developing countries [ScD Dissertation]. [Cambridge (MA)]: Harvard University; 2001. 347 p.

Naimoli JF, Rowe AK, Lyaghfouri A, et al. Effect of the Integrated Management of Childhood Illness strategy on health care quality in Morocco. International Journal for Quality in Health Care 2006;18:134–44.

Ortiz JL, Chumacero LL, Dávila M. Impacto de la aplicación de la estrategia AIEPI en el Uso de Antibióticos para el tratamiento de las Infecciones Respiratorias Agudas y Enfermedades Diarreicas, en Centros de Salud de Lambayeque – Perú. 1998. Unpublished report.

Osterholt DM, Onikpo F, Lama M, et al. Improving pneumonia case-management in Benin: a randomized trial of a multi-faceted intervention to support health worker adherence to Integrated Management of Childhood Illness guidelines. Human Resources for Health 2009;7:77.

Pariyo GW, Gouws E, Bryce J, et al. Improving facility-based care for sick children in Uganda: training is not enough. Health Policy and Planning 2005;20 Suppl 1:i58–i68.

Quality Assurance Project. Evaluation of an IMCI Computer-based Training Course in Kenya. Operations Research Results. Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project (QAP), Bethesda, MD. August 2006.

Rathmony, H and Ministry of Health of Cambodia. Cambodia Health Facility Survey, Summary Data Tables. 2006.

Rathmony, H and Ministry of Health of Cambodia. Cambodia Health Facility Survey, PowerPoint presentation. 2006.

Rehlis, N. The Ministry of Health of Uzbekistan and The World Bank Project, Final Report, Evaluation of The Project Health-1 (Loan 4396 - UZ). 2003. Unpublished report.

Rehlis N. Health Facility Survey in Cambodia. Principal Indicators. June 2007. Unpublished report.

Ross-Degnan D, Laing R, Santoso B, Ofori-Adjei, D, Lamoureux C, Hogerzeil H. Improving pharmaceutical use in primary care in developing counties: a critical review of experience and lack of experience. Presented at the International Conference on Improving Use of Medicines, Chiang Mai, Thailand, April 1997. Available from Dr. Ross-Degnan (E-mail address: Dennis_Ross-Degnan@hms.harvard.edu)

Rowe AK, de Savigny D, Lanata CF, Victora CG. How can we achieve and maintain high-quality performance of health workers in low-resource settings? Lancet 2005;366:1026–35.

(Rowe 2009a) Rowe AK, Onikpo F, Lama M, et al. A multifaceted intervention to improve health worker adherence to Integrated Management of Childhood Illness guidelines in Benin. American Journal of Public Health;99:837–46 (and personal communication from A. Rowe, Centers for Disease Control and Prevention, November 10, 2007, August 9, 2010 and October 3, 2010).

(Rowe 2009b) Rowe AK, Rowe SY, Peters DH, Holloway KA, Chalker J, Ross-Degnan D. Improving health care provider performance in low and middle income countries. Chapter in: *Implementing health services strategies in low and middle income countries*. Washington, DC: The World Bank.

Rowe AK, Onikpo F, Lama M, Osterholt DM, Deming MS. Impact of a malaria control project in Benin that included the Integrated Management of Childhood Illness strategy. American Journal of Public Health. In press.

Salgado R, Baruda G, Choi M, Weinstock A. Health Facility Survey of Child Health Services, Southern Nations, Nationalities and Peoples Region (SNNPR) Ethiopia, March 2002. John Snow International, Inc., 2002. Unpublished report.

South Africa National Department of Health. Report of IMCI health facility survey in South Africa. March 18, 2005. South Africa.

Surjono A, Lestari LED, Soenarto Y, et al. Implementation of Integrated Management of Childhood Illness (IMCI) in East Nusa Tenggara Province: Randomized controlled trial, comparison of the 11-days training to the 5-days training program. October 1998. (unpublished report)

Syla, S. A comparison of standard training (11-days) with adapted 8 days in the Integrated Management of Childhood Illness (IMCI) in Kosova [Masters Thesis]. [Copenhagen]: Institute of Public Health, University of Copenhagen; 2003. (and personal communication from S. Syla, Kosovo WHO Country Office, Prishtina, November 14 and 15, 2007)

Tavrow P, Kekitiinwa Rukyalekere A, Maganda A, et al. A comparison of computer-based and standard training in the Integrated Management of Childhood Illness in Uganda. Quality Assurance Project, 2002.

World Bank. *Income categories of countries*. 2005. Available online at: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20421402~menuPK:64133156~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html, accessed October 14, 2005.

World Health Organization. Interventions and strategies to improve the use of antimicrobials in developing countries. Drug Management Program, World Health Organization, Geneva. 2001. Document number: WHO/CDS/CSR/DRS/2001.9. (Plus a personal communication from J. Chalker, Management Sciences for Health, on May 17, 2004, which included revised results tables.)

(World Health Organization 2003a). World Health Organization. *The analytic review of the Integrated Management of Childhood Illness strategy. Final Report.* 2003a. Available online at: http://www.who.int/child_adolescent_health/documents/9241591730/en/index.html, accessed September 17, 2004.

(World Health Organization 2003b). World Health Organization, Department of Child and Adolescent Health and Development. Health facility survey: tool to evaluate the quality of care delivered to sick children attending outpatients facilities (using the Integrated Management of Childhood Illness clinical guidelines as best practices). Indicators listed in Annex D. Geneva, Switzerland: World Health Organization, 2003. Available at: http://www.who.int/child_adolescent_health/documents/9241545860/en/index.html, accessed October 17, 2008. The list of indicators is in Annex D.

World Health Organization. Regional Office for the Eastern Mediterranean. Implementation of IMCI in Sudan. Systematic approach to IMCI implementation at district level: key steps and tools. Available at: http://www.emro.who.int/cah/IMCI-implementation-sud.htm#Section2.4, accessed October 6, 2008.

World Health Organization. Handbook: IMCI integrated management of childhood illness. Geneva, Switzerland: World Health Organization, 2005. Available at: http://whqlibdoc.who.int/publications/2005/9241546441.pdf, accessed June 3, 2008.

World Health Organization. Report of technical consultation on IMCI training approaches and pre service IMCI. Geneva, Switzerland, 19–23 November, 2007. Available online at: http://www.who.int/child_adolescent_health/documents/imci/en/index.html, accessed December 19, 2008.

World Health Organization. Medicines use in primary care in developing and transitional countries. Fact book summarizing results from studies reported between 1990 and 2006. World Health Organization: Geneva, 2009, document number WHO/EMP/MAR/2009.3, available at: http://www.who.int/medicines/publications/who_emp_2009.3/en/index.html, accessed December 23, 2009.

World Health Organization. Child and adolescent health and development: progress report 2009: highlights. Geneva: World Health Organization, 2010.

Zamora ADG, Cordero DV, Mejia MS. Evaluación de la estrategia "Atención integrada a las enfermedades prevalentes de la infancia" (AIEPI) en Servicios de Salud, Bolivia 1999. Revista Chilena de Pediatría, 2002;73(2):184–191.

Zhang Y, Dai Y, Zhang S. Impact of implementation of Integrated Management of Childhood Illness on improvement of health system in China. Journal of Paediatrics and Child Health. 2007;43(10):681–5.

REFERENCES (organized by Study ID)

Study ID	References
1	1. Arifeen, et al. WHO Bulletin 2005;83:260–7; and personal communication from S. El Arifeen, International Centre for Diarrhoeal Diseases Research, Bangladesh, June 13, 2007. 2. El Arifeen, et al. Lancet 2004;364:1595–602. 3. El Arifeen. Section 4: Health Facility Survey, 2000, pp. xxv–xxxiv. 4. Arifeen, et al. Lancet 2009;374:393–403.
3	1. Amaral, et al. Cadernos de saude publica 2004;20 Suppl 2:S209–19. 2. Amaral. Manejo de casos da AIDPI nas unidades de saúde no Brasil -- 2002. 2002. 3. Gouws, et al. WHO Bulletin 2004;82:509–15.
4	1. Rathmony and Ministry of Health. Cambodia Health Facility Survey, Summary Data Tables. 2006. 2. Rathmony and Ministry of Health. Cambodia Health Facility Survey, PowerPoint presentation. 2006. 3. Rehlis N. Health Facility Survey in Cambodia. Principal Indicators. June 2007.
5	1. Anonymous. Implementación de la estrategia de Atención Integrada a las Enfermedades Prevalentes de la Infancia, Reporte Final, Linea Base, Provincia de Imbabura, Republica del Ecuador, 1997. 2. Ministry of Health of Ecuador. Evaluación de servicios de salud sobre "La Atención Integrada a las Enfermedades Prevalentes de la Infancia AIEPI", 2000.
6	1. Choi, et al. Assessment of child health services in Eritrea after three years of implementation of the Integrated Management of Childhood Illness (IMCI) strategy, 2002. 2. Mehari, et al. Preparedness of Eritrean Health Services for the Integrated Management of Childhood Illness (IMCI) Strategy, 2000.
7	Salgado, et al. Health Facility Survey of Child Health Services, Southern Nations, Nationalities and Peoples Region (SNNPR) Ethiopia, March 2002.
9	1. Centers for Disease Control and Prevention. MMWR 1998; 47(46):998–1001. 2. Lee, et al. unpublished manuscript
10	Gilroy, et al. Patient Educ Couns 2004;54:35–44.
11	1. Naimoli. The Integrated Management of Childhood Illness (IMCI) pilot project in Morocco, Preliminary results from an evaluation of the management of childhood illness in public sector IMCI and non-IMCI facilities in four Moroccan provinces, April 2000. 2. Naimoli JF. Theoretical and empirical advances in research on the implementation of an integrated approach to managing childhood illness in outpatient facilities in developing countries [ScD Dissertation]. [Cambridge (MA)]: Harvard University; 2001. 347 p. 3. Naimoli, et al. Int J Qual Health Care 2006;18:134–44
12	1. Kelley, et al. Operations Research Results 2000;1:16. 2. Kelley, et al. Int J Health Plan Management 2001;16:195–205. 3. Kelley, et al. Operations Research Results 2002;2:15.

Study ID	References
13	1. Ministry of Health of Niger. Enquête dans les formations sanitaires mettant en œuvre la PCIME. 2005. 2. Degbey H. Enquête dans les formations sanitaires mettant en oeuvre la PCIME au Niger (15 mars- 15 avril 2005) Lettre de l'OMS Niger 2005; 40: 6-7. Available at: http://www.santetropicale.com/alire/omsnig0605.pdf, accessed September 15, 2010.
14	1. Huicho, et al. Health Policy Plan 2005;20:14–24. 2. Ministry of Health of Perú. Evaluación de Servicios de Salud. Segunda Prueba Mundial, Perú, 13–24 Octubre 1999. 2000.
15	Chopra, et al. Arch Dis Child 2005;90:397–401.
16	1. Armstrong Schellenberg, et al. Lancet 2004;364:1583–94. 2. Armstrong Schellenberg, et al. Health Policy Plan 2004;19:1–10. 3. Bryce, et al. Health Policy Plan 2005;20 Suppl 1:i69–i76. 4. Gouws, et al. WHO Bulletin 2004;82:509–15. 5. Mbuya, et al. IMCI Implementation: A Report on Experiences in Morogoro and Rufiji Districts in Tanzania, Preliminary Report. 2003. 6. Mgalula. Activities and Methods, Substudy 3: Health Facility Survey. 2000; pp. 5–19.
17	Tavrow, et al. Quality Assurance Project, 2002.
18	1. Gouws, et al. WHO Bulletin 2004;82:509–15. 2. Pariyo, et al. Health Policy Plan 2005;20 Suppl 1:i58–i68.
19	Rehlis. The Ministry of Health of Uzbekistan and The World Bank Project, Final Report, Evaluation of The Project Health-1 (Loan 4396 - UZ). 2003.
20	Mwinga, et al. Research Study Report: Comparative study of Integrated Management of Childhood Illness (IMCI) case management skills of primary health care workers trained in eleven-day standard course and six-day abridged course for physicians. 2006.
21	Syla S. A comparison of standard training (11-days) with adapted 8 days in the Integrated Management of Childhood Illness (IMCI) in Kosova. 2003. (and personal communication from S. Syla, Kosovo WHO Country Office, Prishtina, November 14 and 15, 2007)
22	Anonymous. Evaluación del proceso de implementacion de la estrategia de Atención Integrada a las Enfermedades Prevalentes en la Infancia - AIEPI - en ambitos de las Direcciones de Salud Callao y Lima Norte (Ventanilla y Puente Piedra), 2004.
23	Ortiz, et al. Impacto de la aplicación de la estrategia AIEPI en el Uso de Antibióticos para el tratamiento de las Infecciones Respiratorias Agudas y Enfermedades Diarreicas, en Centros de Salud de Lambayeque - Perú, 1998.
24	1. Osterholt, et al. Hum Resour Health 2009;7:77. 2. Rowe, et al. Am J Public Health 2009;99:837–46 (and personal communication from A. Rowe, Centers for Disease Control and Prevention, November 10, 2007, August 9, 2010 and October 3, 2010).
25	Atakouma Dzayisse, et al. Arch Pediatr 2006;13:1552–3.

Study ID	References
26	1. Assimadi, et al. Arch Pediatr 1999;6:1135–6. 2. Assimadi, et al. Arch Pediatr 2003;10:158–9.
27	Burnham GM. Evaluation of Integrated Management of Childhood Illness (IMCI) Performance in Urban Health Centers, Lusaka, Zambia, August 10–23, 1997. BASICS, 1997.
28	1. Anonymous. Informe Preliminar: Evaluación AIEPI a Servicios de Salud, Primera Prueba Mundial, Bolivia, 12–30 de Abril 1999. 2. Zamora, et al. Rev Chil Pediatr, 2002;73(2):184–191. 3. Cordero D, Salgado R, Drasbeck C. An analysis of the IMCI implementation process in four countries of Latin America. Edited by Tina G. Sanghvi and Mireille Mather. Arlington, Virginia: BASICS II for the United States Agency for International Development and the Pan American Health Organization. 2004.
29	Briggs, et al. Senegal Assessment: Drug Management for Childhood Illness. 2002.
30	1. Federal Ministry of Health of Sudan. Health Facility Survey on Quality of Outpatient Child Health Services, IMCI Health facility survey, Sudan, March–April 2003. 2004. 2. World Health Organization. Regional Office for the Eastern Mediterranean. Implementation of IMCI in Sudan. Systematic approach to IMCI implementation at district level: key steps and tools. Available at: http://www.emro.who.int/cah/IMCI-implementation-sud.htm#Section2.4, accessed October 6, 2008.
31	Zhang, et al. J Paediatr Child Health. 2007 Oct;43(10):681–5.
32	1. Goga A. IMCI health facility survey: Eastern Cape and North West Provinces, 2003 (plus personal communication from A. Goga, Medical Research Council of South Africa, July 13, 2009). 2. South Africa National Department of Health. Report of IMCI health facility survey in South Africa, 2005.
33	Ministry of Health of Vietnam, National Institute of Hygiene and Epidemiology of Vietnam. Official Travel Report, Annex 1: IMCI health facility survey in Vietnam; Planning Workshop, Hanoi, June 17–21, 2002. WHO, Western Pacific Regional Office (plus personal communication from T. Lambrechts, WHO, on July 1, 2009 and December 9, 2009).

Box 1. Definitions used in the review

Adequacy of study designs

- "First-tier" study designs are:
  - Pre-post study (i.e., performance measured before and after Integrated Management of Childhood Illness [IMCI] implementation) with randomized controls
  - Pre-post study with non-randomized controls
  - Post-only study with randomized controls
  - Interrupted time series with at least three data points before and after IMCI training

- "Second-tier" study designs are:
  - Interrupted time series with fewer than three data points before and after IMCI training (e.g., pre-post study without controls)
  - Post-only study with non-randomized controls
  - Case-control study (i.e., stratify patients assessed in a cross-sectional health facility survey according to whether the consultation was performed by an IMCI-trained or non-IMCI-trained health worker; similar to a post-only study with non-randomized controls)

Note 2. The classification scheme does not imply that "second-tier" designs are without value; indeed, many seem quite robust for demonstrating whether IMCI had a positive effect. The classification is intended to identify studies that are relatively less susceptible to bias.

Analysis-related terms

- An outcome measure is a numerical value for an outcome for a particular study group at a particular time point.

- A comparison is a contrast between two study groups. For example, a study with three study groups (intervention 1, intervention 2, and controls) could have up to three comparisons (intervention 1 versus controls, intervention 2 versus controls, and intervention 1 versus intervention 2). Multiple effect sizes could relate to a single comparison if measurements are made over time (e.g., baseline measure, follow-up measure 1, and follow-up measure 2).

- A summary measure is an outcome derived from a group of related outcomes that are not identical. Summary measures are used to analyze results from studies that use different outcomes. The two used in this review are:
  - Median effect size (MES)—i.e., the median of effect sizes for all outcomes for a given comparison from a given study (see Annex 1A for details). For example, as shown in Figure 2, one study [Study ID 20] had seven effect sizes; the MES is the median of the seven effect sizes, which was 7.0 percentage-points.
  - Percent of patients needing an oral antimicrobial (antibiotic or antimalarial) or oral rehydration solution who received these treatments according to IMCI guidelines ("patient treated according to IMCI guidelines", or PTIG) (see Annex 1 for details). In addition to selecting appropriate medicines, treatment almost always involves correct dosing and treatment duration.

Table 1. Sources and search strategies used to identify studies included in the review

Source (date of search)	Search strategy
1. WHO/CAH literature search (January 2007)	Searched OVID "MEDLINE R In process" and other non-indexed citations for reports with key word "IMCI" (personal communication, A. Goga, May 10, 2007)
2. WHO/CAH reports (May 2007)	Searched reports from CAH and WHO regional offices (mainly unpublished reports) (person communication, T. Lambrechts, May–June 2007)
3. WHO/INRUD database (December 2006)	Searched WHO/INRUD database[a] for IMCI intervention studies (personal communication, K. Holloway, June 27, 2007)
4. HCPP Review Study Group (May 2006)	Searched HCPP database[b] with key words "Integrated Management of Childhood Illness" and "IMCI" (personal communication, Samantha Rowe, May 2006)
5. Investigators of IMCI evaluations (July 2007)	Unpublished reports from Bangladesh (personal communication, S. Arifeen, August 14, 2007), Benin (personal communication, A. Rowe, November 10, 2007), South Africa (personal communication, A. Goga, July 13, 2009), and Vietnam (personal communication, T. Lambrechts, July 1 and December 9, 2009).

Footnotes for Table 1.

CAH = WHO's Department of Child and Adolescent Health and Development; HCPP = Health Care Provider Performance; IMCI = Integrated Management of Childhood Illness; INRUD = International Network for Rational Use of Medicines; WHO = World Health Organization.

[a] This database is the result of a systematic review on medicine use. The project is supported by WHO and INRUD. Included studies were published from 1990–2006 (as found in searches conducted in December 2006), written in English, French, Spanish, Portuguese, or Russian, and obtained as full-text articles (i.e., complete text of studies was required; studies were excluded if only an abstract was available). See reference WHO, 2009 for detailed search strategy.

[b] This database is the result of a systematic review on improving health care provider performance. The HCPP Review Study Group is jointly supported by staff from the Centers for Disease Control and Prevention, Harvard University, WHO, Management Sciences for Health, the World Bank, and Johns Hopkins University. The detailed search strategy can be found in the HCPP Review protocol (HCPP Review Study Group 2008) and in Rowe *et al*. 2009b. Studies published from 1951–2006 (as found in searches of electronic databases conducted in May 2006) were included; there were no language restrictions.

Table 2. Interventions to support the Integrated Management of Childhood Illness (IMCI) strategy that complement IMCI in-service training

Country (Study ID)	Interventions
Bangladesh [Study ID 1]	Essential IMCI drugs and supplies were purchased, monthly supervision by IMCI-trained medical officers, standard government health information system forms were modified
Benin [Study ID 24]	Increased IMCI-related supervision, supervision of supervisors, training on and distribution of job aids, and non-financial incentives
Bolivia [Study ID 28]	Ministry of Health already had started to integrate child health programs (the acute diarrheal disease and acute respiratory infections programs had merged); IMCI implemented with a supervision and monitoring system in which supervisors visited health workers 1–2 times per year
Brazil [Study ID 3]	IMCI implemented in the context of a Family Health Program supported by the World Bank and Ministry of Health
Cambodia [Study ID 4]	IMCI implemented in the context of extra health system supports (e.g., contracting, support from non-governmental organizations) and additional supervision of health workers
China [Study ID 31]	Increased supervision of drug use by health authorities; in some areas, local governments provided IMCI-trained health workers with subsidies to compensate for the reduction of income due to the rational use of drugs
Eritrea [Study ID 6]	Donor-funded IMCI implementation resulted in “steadfast and continuous support” that allowed the strategy to have “steady implementation without any major disruptions”
Ethiopia [Study ID 7]	Some IMCI-trained health workers also attended an Integrated Refresher Training course
Niger [Study ID 12]	IMCI implemented in an area with a pre-existing audit and feedback system
Peru [Study ID 22]	IMCI implemented in a health system that integrates health promotion, prevention, and care (*Modelo de Atención Integral de Salud*) with a comprehensive health insurance program for patients (*Seguro Integral de Salud*)

Table 2 continued on next page.

Table 2, continued. Interventions to support the Integrated Management of Childhood Illness (IMCI) strategy that complement IMCI in-service training

Country (Study ID)	Interventions
Sudan [Study ID 30]	IMCI implemented with a supervision system that involved training supervisors on IMCI-related supervision and the intent of quarterly visits to all health facilities
Tanzania [Study ID 16]	Multi-faceted health systems intervention including local governments having increased control over health budgets and plans, donor-supported "basket" funding from the health-sector-wide approach, sentinel burden-of-disease information tool, and a district health budget mapping tool
Uganda [Study ID 18]	For nurse assistants, IMCI training given as part of a 3-month in-service training course
Uzbekistan [Study ID 19]	IMCI implemented as part of "Project Health", which included a 10-month in-service training program and supplying health facilities with medical equipment, furniture, and drugs
Zambia [Study ID 27]	Drugs were available through donor funding at the central medical store; emergency drug funds were available for districts to locally purchase drugs; improved supervision and recognition of quality assurance at an institutional level (e.g., health facility managers and clinicians); donors provided family planning commodities, equipment, drugs, and training in family planning and communication for behavior change

Table 3. Study designs and number of comparisons among the 29 studies included in the primary analysis

Study design	Adequacy of the study design	No. of studies (%)	No. of comparisons
Pre-post study with randomized controls	First tier	1 (3)	1
Pre-post study with non-randomized controls	First tier	1 (3)	2[a]
Post-only study with randomized controls	First tier	3 (10)	3
Pre-post study with no controls	Second tier	10 (35)	10
Post-only study with non-randomized controls	Second tier	8 (28)	9[a]
Case-control study	Second tier	6 (21)	6
Total		29 (100)	31

Footnotes for Table 3.

IMCI = Integrated Management of Childhood Illness.

[a] In this row, one study had three study groups (IMCI with extra interventions, IMCI without extra interventions, and no IMCI) and two comparisons (IMCI with extra interventions versus no IMCI, and IMCI without extra interventions versus no IMCI).

Table 4a. Study design adequacy and number of comparisons among the 29 studies included in the primary analysis

Study design adequacy	Comparison			Total
	Standard versus short training (references[a])	Standard training versus no IMCI (references[a])	Short training versus no IMCI (references[a])	
First tier (for main analysis)	2 studies with 2 comparisons (Study IDs: 17, 20)	3 studies with 4 comparisons (Study IDs: **1**, 10, **24**[b])	0 studies with 0 comparisons	5 studies with 6 comparisons
Second tier	1 study with 1 comparison (Study ID: 21)	16 studies with 17 comparisons (Study IDs: **4**[b], **6**, **7**, 9, 11, **12**, 15, **16**, **18**, **19**, 25, 26, **27**, **28**, 29, **30**)	7 studies with 7 comparisons (Study IDs: **3**, 5, 13, 14, **22**, 23, **31**)	24 studies with 25 comparisons
All study designs (for sensitivity analysis)	3 studies with 3 comparisons	19 studies with 21 comparisons	7 studies with 7 comparisons	29 studies with 31 comparisons

Footnotes for Table 4a.

IMCI = Integrated Management of Childhood Illness.

[a] Reference numbers in bold font indicate studies in which IMCI was implemented with complementary interventions to support IMCI (see Table 2).

[b] Two studies [Study IDs 4 and 24] each had two comparisons: IMCI with extra interventions versus no IMCI, and IMCI without extra interventions versus no IMCI.

Table 4b. Study design adequacy and number of comparisons among the 15 studies included in the primary analysis in which the "patient treated according to IMCI guidelines" (PTIG) outcome was measured

Study design adequacy	Comparison			Total
	Standard versus short training (references[a])	Standard training versus no IMCI (references[a])	Short training versus no IMCI (references[a])	
First tier (for main analysis)	2 studies with 2 comparisons (Study IDs: 17, 20)	2 studies with 3 comparisons (Study IDs: **1**, **24**[b])	0 studies with 0 comparisons	4 studies with 5 comparisons
Second tier	1 study with 1 comparison (Study ID: 21)	7 studies with 8 comparisons (Study IDs: **4**[b], **6**, 11, **16**, **18**, 25, **30**)	3 studies with 3 comparisons (Study IDs: **3**, 13, 14)	11 studies with 12 comparisons
All study designs (for sensitivity analysis)	3 studies with 3 comparisons	9 studies with 11 comparisons	3 studies with 3 comparisons	15 studies with 17 comparisons

Footnotes for Table 4b.

IMCI = Integrated Management of Childhood Illness.

[a] Reference numbers in bold font indicate studies in which IMCI was implemented with complementary interventions to support IMCI (see Table 2).

[b] Two studies [Study IDs 4 and 24] each had two comparisons: IMCI with extra interventions versus no IMCI, and IMCI without extra interventions versus no IMCI.

Table 5. Types of health workers trained in the 29 studies included in the primary analysis

Health worker type[a]	No. of studies (%)
Nurse	24 (83)
Physician	19 (66)
Nursing aide or clinical assistant	7 (24)
Midwife	5 (17)
Clinical Officer	4 (14)
Medical assistant	4 (14)
Paramedic or non-physician	2 (7)
Physician assistant	1 (3)
Medical officer	1 (3)
Other[b]	8 (28)
Don't know	1 (3)

Footnotes for Table 5.

[a] No standard definitions were used; categories were based on the terminology used in the report.

[b] Includes: assistant clinical officer, environmental technician, health information officer, "*assistant sociaux*", health assistant, nutritionist, dentist, pre-graduate medical and nursing students, public health officer, and "other health worker".

Table 6. Modifications to reduce course duration in 10 studies of short training included in the primary analysis

Modification	No. of studies (%)
Modification of content[a]	1 (10)
Modification of at least one training method	8 (80)
Clinical practice	5 (63[b])
Non-interactive sessions (e.g., participants reading text)	7 (88)
Interactive sessions	3 (38)
Computer-based training added	1 (13)
Other	1 (13)

Footnotes for Table 6.

[a] For example, adding materials related to early child development.

[b] That is, 63% of studies that changed at least one training method modified something about clinical practice (usually shortening it).

Table 7a. Results of a sensitivity analysis in the primary analysis with the median effect size (MES) summary measure: multivariable linear regression models[a] (training duration analyzed as a dichotomous variable)

Variable	Model 1[b]		Model 2		Model 3	
	Coeff.[c]	p-value	Coeff.[c]	p-value	Coeff.[c]	p-value
Intercept	46.27	0.07	33.68	0.09	34.28	**0.01**
Training approach (1 = standard, 0 = short)	–1.95	0.93	13.18	0.26	1.33	0.86
Other interventions implemented (1 = yes, 0 = no)	5.72	0.42	6.92	0.32	–1.35	0.83
First-tier study design (1 = yes, 0 = no)	–9.80	0.36	–10.76	0.31	7.96	0.44
Median baseline or control value	–0.39	0.14	–0.42	0.11	–0.29	0.10
Time since training (months)	–0.27	0.45	–0.03	0.88	—	—
"Time since training x training approach" interaction	0.33	0.43	—	—	—	—

Footnotes for Table 7a.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness.

[a] Models 1 and 2 involved 25 effect sizes from 17 comparisons from 15 studies. Model 3 involved 28 effect sizes (one per comparison) from 26 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] As an example, the model was:
$MES = \beta_0 + (\beta_1 \times [\text{Training approach}]) + (\beta_2 \times [\text{Other interventions}]) + (\beta_3 \times [\text{First-tier study design}]) + (\beta_4 \times [\text{Median baseline or control value}]) + (\beta_5 \times [\text{Time since training}]) + (\beta_6 \times [\text{Time since training}] \times [\text{Training approach}]) + \varepsilon$,
where β_0 is the parameter estimate for the intercept, β_1 is the parameter estimate for the "Training approach" variable, β_2 is the parameter estimate for the "Other interventions" variable, and so on; and ε is the error term.

[c] Coefficients are percentage-point differences. For example, in row 2 of Model 3, the coefficient means that the effect size for standard training is 1.33 percentage-points higher than for short training, after controlling for all other factors in the model (although the p-value of 0.86 indicates that there is insufficient evidence to conclude that this difference is different from zero).

Table 7b. Results of a sensitivity analysis in the primary analysis with the median effect size summary measure: multivariable linear regression models[a] (training duration analyzed as a continuous variable)

Variable	Model 4 [b]		Model 5		Model 6	
	Coeff.[c]	p-value	Coeff.[c]	p-value	Coeff.[c]	p-value
Intercept	29.76	0.35	15.40	0.52	25.28	0.17
Training duration (days)	1.26	0.58	2.48	0.11	0.89	0.51
Other interventions implemented (1 = yes, 0 = no)	6.35	0.38	8.14	0.23	–1.55	0.80
First-tier study design (1 = yes, 0 = no)	–10.81	0.30	–9.53	0.35	7.49	0.47
Median baseline or control value	–0.39	0.13	–0.40	0.11	–0.27	0.12
Time since training (months)	–0.42	0.49	–0.01	0.95	—	—
"Time since training x training duration" interaction	0.04	0.48	—	—	—	—

Footnotes for Table 7b.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness.

[a] Models 4 and 5 involved 25 effect sizes from 17 comparisons from 15 studies. Model 6 involved 28 effect sizes (one per comparison) from 26 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] For an example of the model, see footnote "b" of Table 7a.

[c] Coefficients are percentage-point differences. See footnote "c" of Table 7a.

Table 8a. Results of a sensitivity analysis in the primary analysis with the PTIG summary measure: multivariable linear regression models[a] (training duration analyzed as a dichotomous variable)

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8	
Variable	Coeff.[b]	P	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p
Intercept	–20.95	0.58	29.89	0.27	34.51	0.20	34.93	0.16	10.87	0.53	34.83	0.14	10.79	0.52	9.10	0.55
Training approach (1=standard, 0=short)	104.77	**0.02**	41.42	**0.02**	23.27	0.27	23.49	0.23	30.85	0.14	24.25	0.17	33.20	0.09	31.66	0.08
Other interventions implemented (1=yes, 0=no)	23.57	**0.03**	22.41	0.05	1.69	0.92	1.85	0.90	–5.31	0.74	—	—	–5.06	0.74	—	—
First-tier study design (1=yes, 0=no)	–34.07	**0.02**	–29.00	**0.04**	1.49	0.94	—	—	8.89	0.65	—	—	—	—	—	—
Baseline or control value of PTIG	–1.56	**0.01**	–1.33	**0.02**	–0.83	0.23	–0.85	0.18	—	—	–0.83	0.15	—	—	—	—
Time since training (months)	1.08	0.08	0.20	0.46	—	—	—	—	—	—	—	—	—	—	—	—
"Time since training x training approach" interaction	–1.08	0.11	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Footnotes for Table 8a.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness; p = p-value for test that coefficient is different from zero; PTIG = percent of patients treated according to IMCI guidelines.

[a] Models 1 and 2 involved 15 effect sizes from 10 comparisons from 8 studies. Models 3–8 involved 14 effect sizes (one per comparison) from 12 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] Coefficients are percentage-point differences. For example, in row 2 of Model 7, the coefficient means that the effect size for standard training is 33.2 percentage-points higher than for short training, after controlling for the other factor in the model (and the p-value of 0.09 indicates that there is insufficient evidence to conclude that this difference is different from zero).

Table 8b. Results of a sensitivity analysis in the primary analysis with the PTIG summary measure: multivariable linear regression models[a] (training duration analyzed as a continuous variable)

	Model 9		Model 10		Model 11		Model 12		Model 13		Model 14		Model 15		Model 16	
Variable	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p
Intercept	–52.59	0.49	–8.68	0.82	3.84	0.92	3.87	0.92	–24.80	0.51	4.73	0.89	–27.68	0.44	–25.25	0.46
Training duration (days)	11.51	0.11	7.30	**0.02**	5.04	0.18	4.98	0.15	5.84	0.14	4.86	0.11	6.27	0.09	5.57	0.09
Other interventions implemented (1=yes, 0=no)	18.17	0.12	18.22	0.10	–1.26	0.94	–1.37	0.93	–8.75	0.60	—	—	–8.85	0.58	—	—
First-tier study design (1=yes, 0=no)	–31.30	**0.04**	–30.43	**0.04**	–1.52	0.94	—	—	7.63	0.70	—	—	—	—	—	—
Baseline or control value of PTIG	–1.37	**0.03**	–1.31	**0.03**	–0.93	0.16	–0.92	0.12	—	—	–0.93	0.09	—	—	—	—
Time since training (months)	0.91	0.45	0.11	0.68	—	—	—	—	—	—	—	—	—	—	—	—
"Time since training x training duration" interaction	–0.08	0.50	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Footnotes for Table 8b.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness; p = p-value for test that coefficient is different from zero; PTIG = percent of patients treated according to IMCI guidelines.

[a] Models 9 and 10 involved 15 effect sizes from 10 comparisons from 8 studies. Models 11–16 involved 14 effect sizes (one per comparison) from 12 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] Coefficients are percentage-point differences. See footnote "b" of Table 8a.

Table 9. Summary of all analyses: primary analysis

Study design adequacy and summary measure	Analysis[a]					
	Direct comparison (standard compared to short training)	Indirect comparison (standard training vs. no IMCI compared to short training vs. no IMCI)	Training duration analyzed as a continuous variable	Indirect comparison stratified by presence of other interventions (besides IMCI training)	Impact of presence of other interventions (besides IMCI training)	Analysis of time since training
Main analysis (first-tier study designs only)						
Outcome = MES Note: Few studies (≤6 effect sizes[b] from 5 studies)	Standard was slightly better (5 %-point difference, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Suggested increasing effect size with longer training duration (only 4 effect sizes, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Standard training with other interventions seemed much more effective than without other interventions (35 %-point difference)	No apparent trends over time after standard training (no statistical testing); no analysis for trends after short training (no studies of short vs. no IMCI)
Outcome = PTIG Note: Very few studies (≤5 effect[b] sizes from 4 studies)	Standard and short were very similar (3 %-point difference, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Suggested increasing effect size with longer training duration (only 3 effect sizes, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Standard training with other interventions seemed much more effective than without other interventions (42 %-point difference)	No apparent trends over time after standard training (no statistical testing); no analysis for trends after short training (no studies of short vs. no IMCI)
Sensitivity analysis (first- and second-tier study designs)						
Outcome = MES ≤31 effect sizes[b] from 29 studies	Standard and short were very similar (2 %-point difference, no statistical testing)	Standard was somewhat better than short (9 %-point difference, p=1.0)	Trend of increasing effect with longer training duration (univariate p=0.15 [1.9 %-point increase per extra day of training]; multivariable p-value range: 0.11–0.51)	Standard was somewhat better than short (by 6 %-points) regardless of other interventions	*Univariate analyses*: Any training with other interventions seemed slightly less effective than without other interventions (–4 %-point difference). *Multivariable analyses*: Impact of other interventions was mixed (–2 to 8 %-points, p-value range: 0.23–0.80).	No apparent trends over time after either standard (p=0.30) or short training (only 4 effect sizes, no statistical testing)

Table 9 continued on next page.

Table 9, continued. Summary of all analyses: primary analysis

Study design adequacy and summary measure	Analysis[a]					
	Direct comparison (standard compared to short training)	Indirect comparison (standard training vs. no IMCI compared to short training vs. no IMCI)	Training duration analyzed as a continuous variable	Indirect comparison stratified by presence of other interventions (besides IMCI training)	Impact of presence of other interventions (besides IMCI training)	Analysis of time since training
Sensitivity analysis (first- and second-tier study designs)						
Outcome = PTIG ≤17 effect sizes[b] from 15 studies	Standard and short were very similar (–3 %-point difference, no statistical testing)	Standard was much better than short (23 %-point difference; p=0.19)	Trend of increasing effect with longer training duration (univariate p=0.09 [5.6 %-point increase per extra day of training]; multivariable p-value range: 0.09–0.11)	Standard was much better than short (by 23 %-points with other interventions, by 30 %-points without other interventions), regardless of other interventions	*Univariate analyses*: Standard training with other interventions seemed as effective as without other interventions (2 %-point difference). Short training with other interventions seemed somewhat more effective than without other interventions (9 %-point difference). *Multivariable analyses*: No significant difference between training with other interventions and training without other interventions (–9 to –5 %-point difference, p-value range: 0.58–0.74)[c]	Suggested that PTIG increases over time after standard training (p=0.17); after short training, a possible increase over time (only 3 effect sizes, no statistical testing)

Footnotes for Table 9.

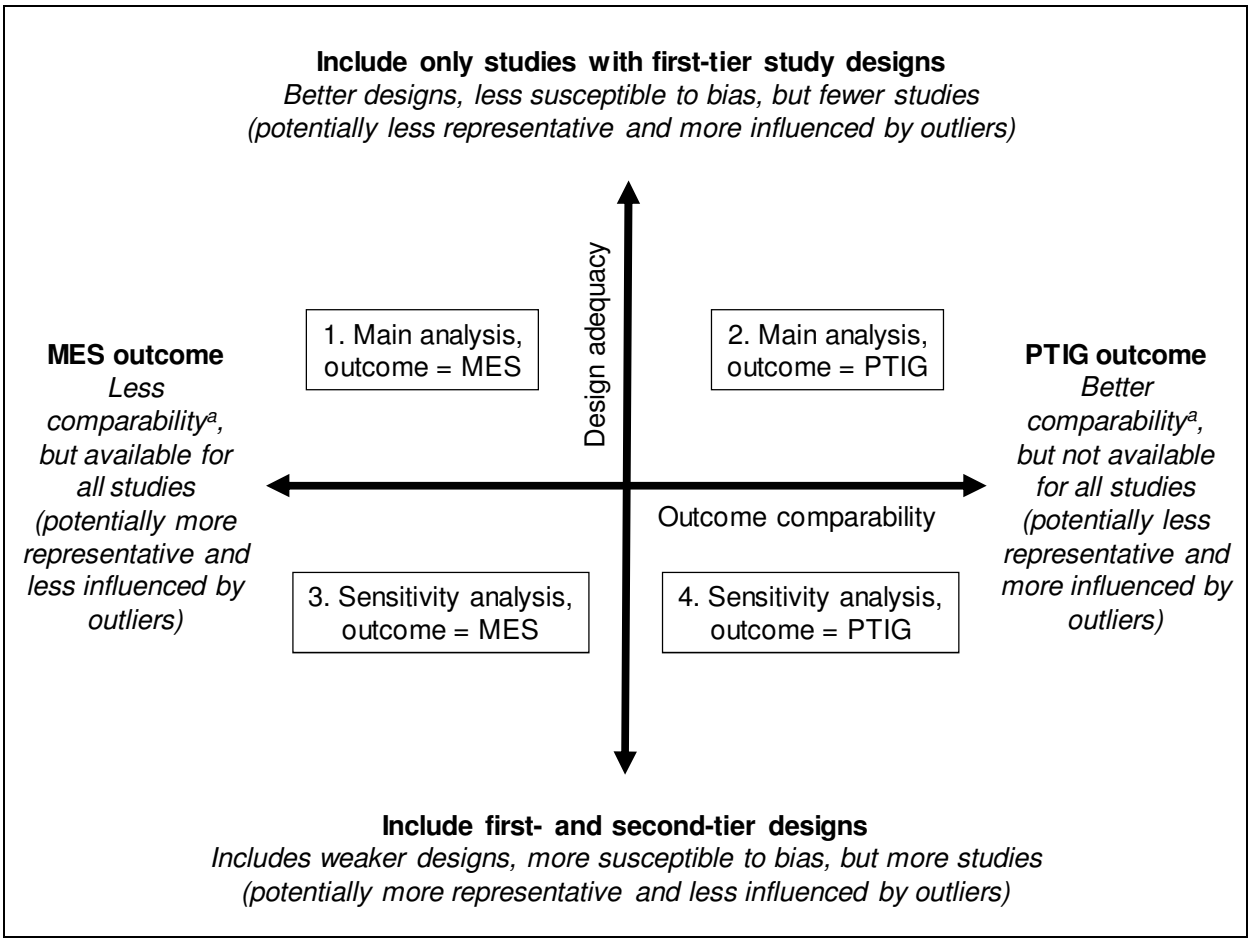
IMCI = Integrated Management of Childhood Illness; MES = median effect size; PTIG = patient (needing oral antimicrobial or rehydration therapy) treated according to IMCI guidelines; %-point = percentage-point.

[a] All differences are all in terms of "effect of standard training" minus "effect of short training" (i.e., positive differences indicate standard training led to greater improvements in healthcare quality than short training).

[b] Excludes the effect sizes from Study ID 24 for specific months since training that used measures approximated by regression models (see Methods for details).

[c] Multivariable p-values were from models with 2 independent variables, since models with >2 independent variables might be unreliable, given the small number of effect sizes.

Figure 1. Four analytic approaches, contrasted by study design adequacy and outcome comparability

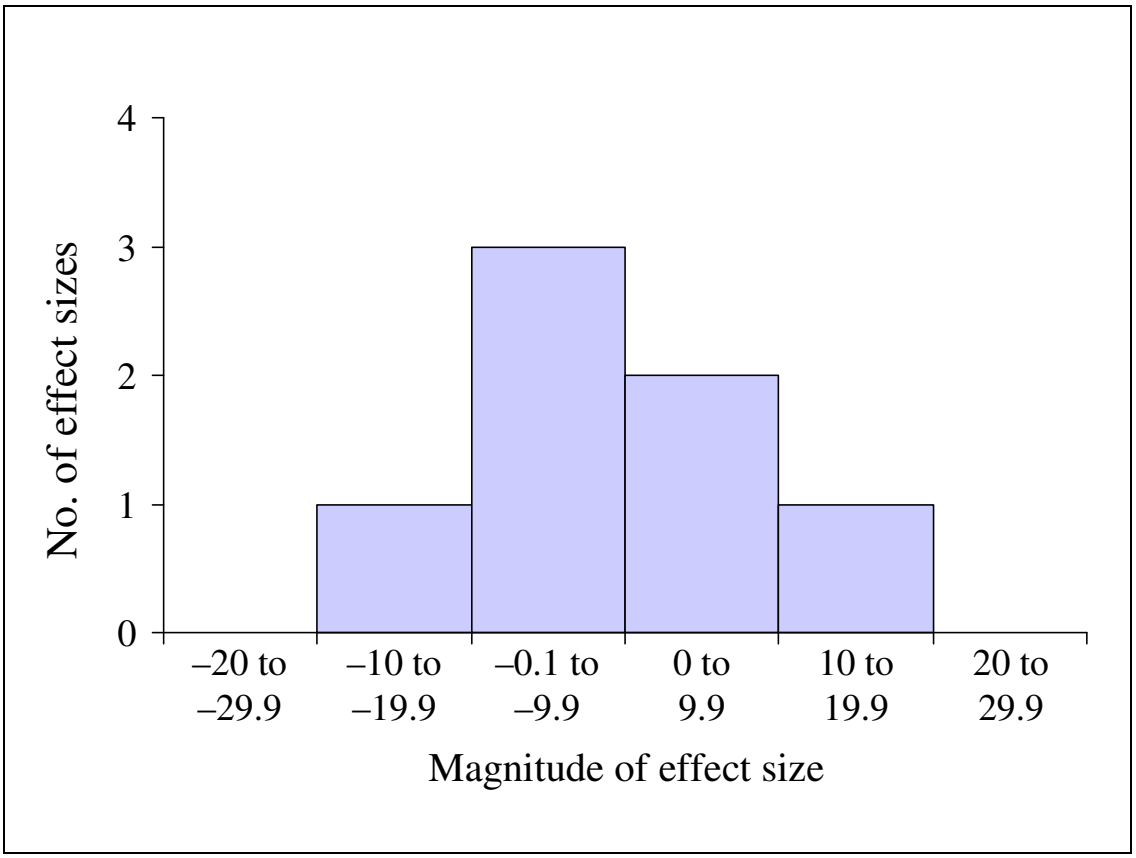


Footnotes for Figure 1.

MES = Median effect size for all outcomes of health worker performance (see Methods); PTIG = the percent of patients treated according to Integrated Management of Childhood Illness guidelines.

[a] Comparability is the degree to which outcomes from different studies have similar (or identical) definitions. Better comparability facilitates the interpretation of comparisons across studies, with perfect comparability being identical outcomes for all studies.

Figure 2. Distribution of seven outcomes (i.e., seven effect sizes) in one study [Study ID 20] that contributed to its median effect size[a]



Footnote for Figure 2.

[a] The median effect size for this study is 7.0 percentage-points (minimum = –17.0, maximum = 10.3). Note that the comparison is standard training performance minus short training performance.

Figure 3. Distribution of 23 countries where the 29 studies included in the primary analysis were conducted, by World Health Organization region

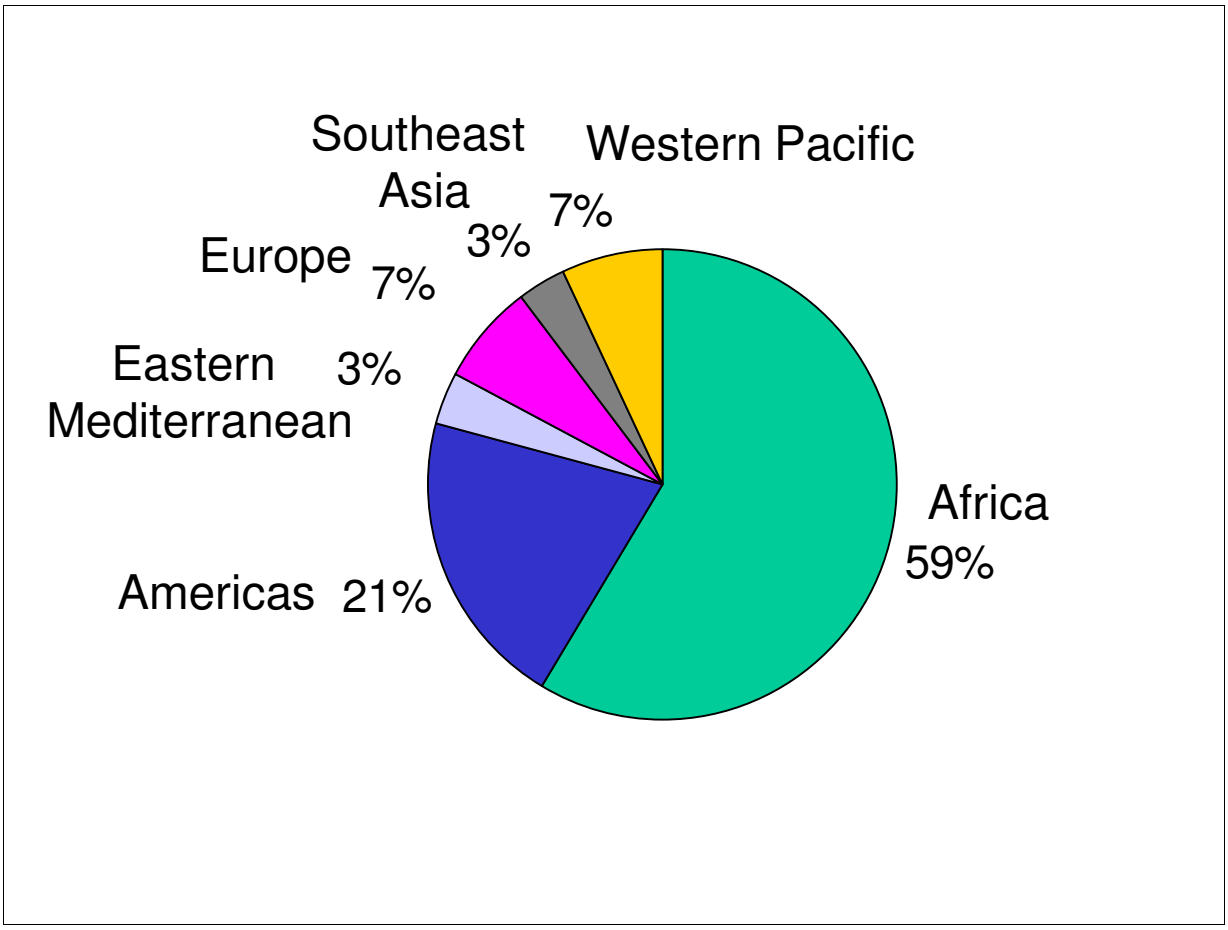


Figure 4. Distribution of the 29 studies included in the primary analysis, by year of publication or date of report (for unpublished studies)

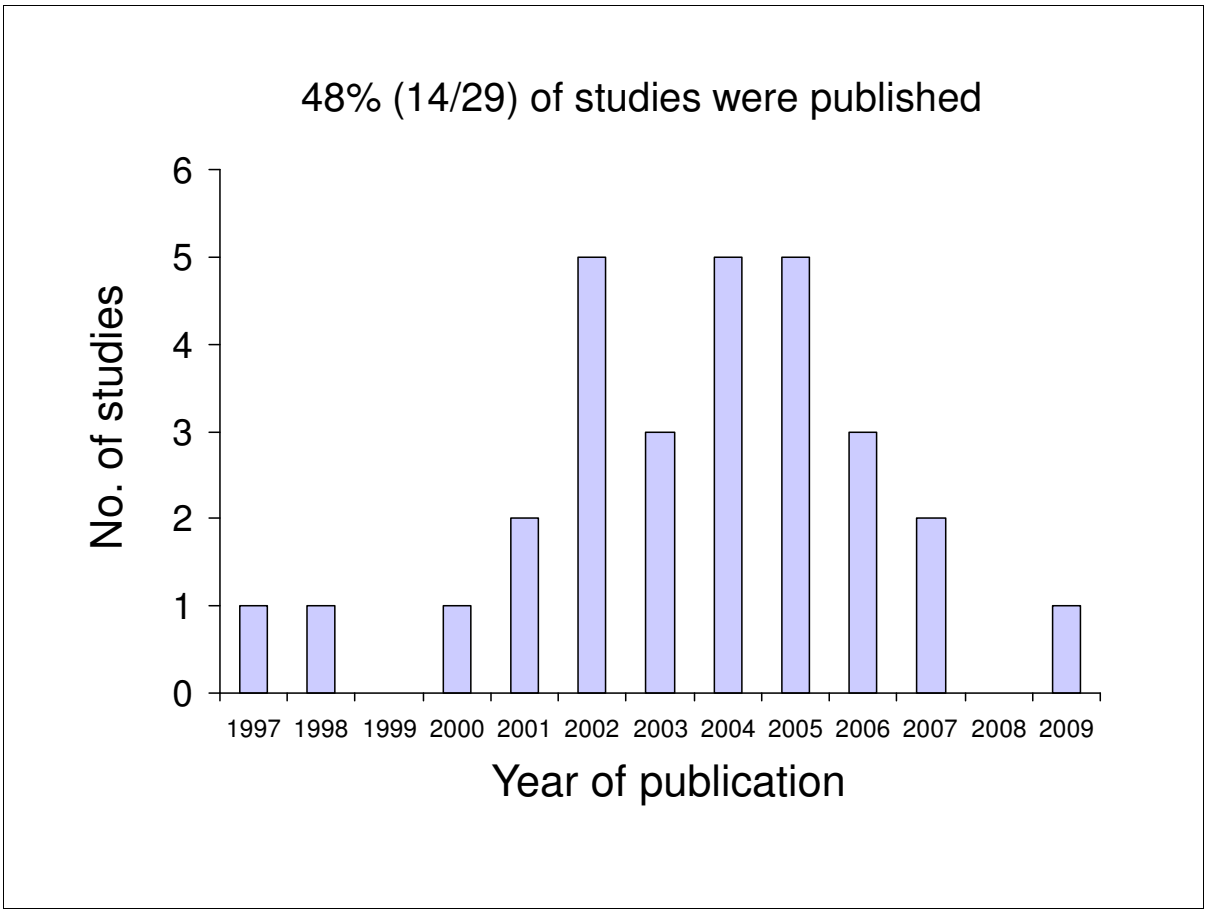
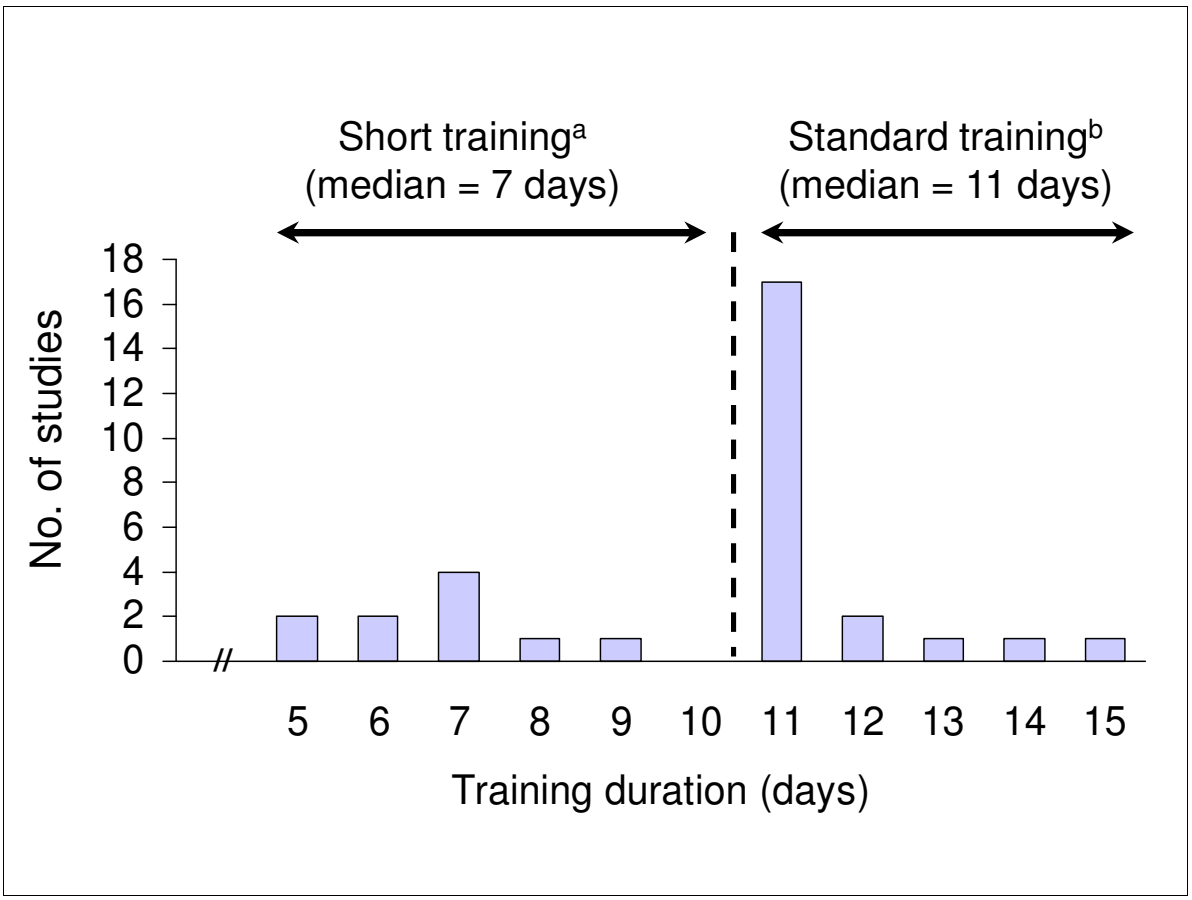


Figure 5. Distribution of the 29 studies included in the primary analysis, by duration of training



Footnotes for Figure 5.

IMCI = Integrated Management of Childhood Illness.

[a] Includes standard versus short training (N = 3 studies, median duration of short training = 8 days) and short training versus no IMCI (N = 7 studies, median training duration = 7 days).

[b] Includes standard versus short training (N = 3 studies, median duration of standard training = 11 days) and standard training versus no IMCI (N = 19 studies, median training duration = 11 days).

Figure 6. Distribution of the 29 studies included in the primary analysis, by administrative level of country at which training was done

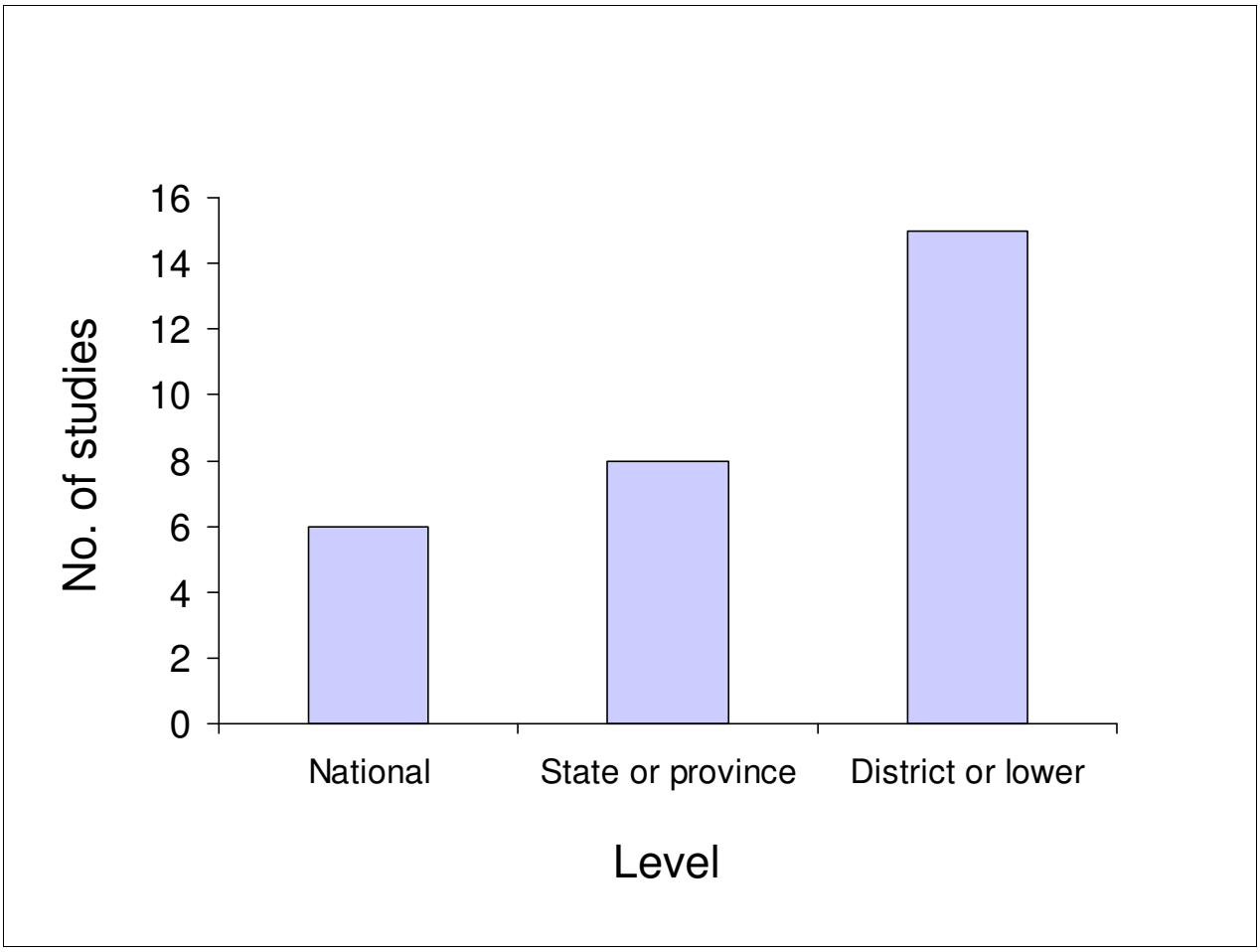
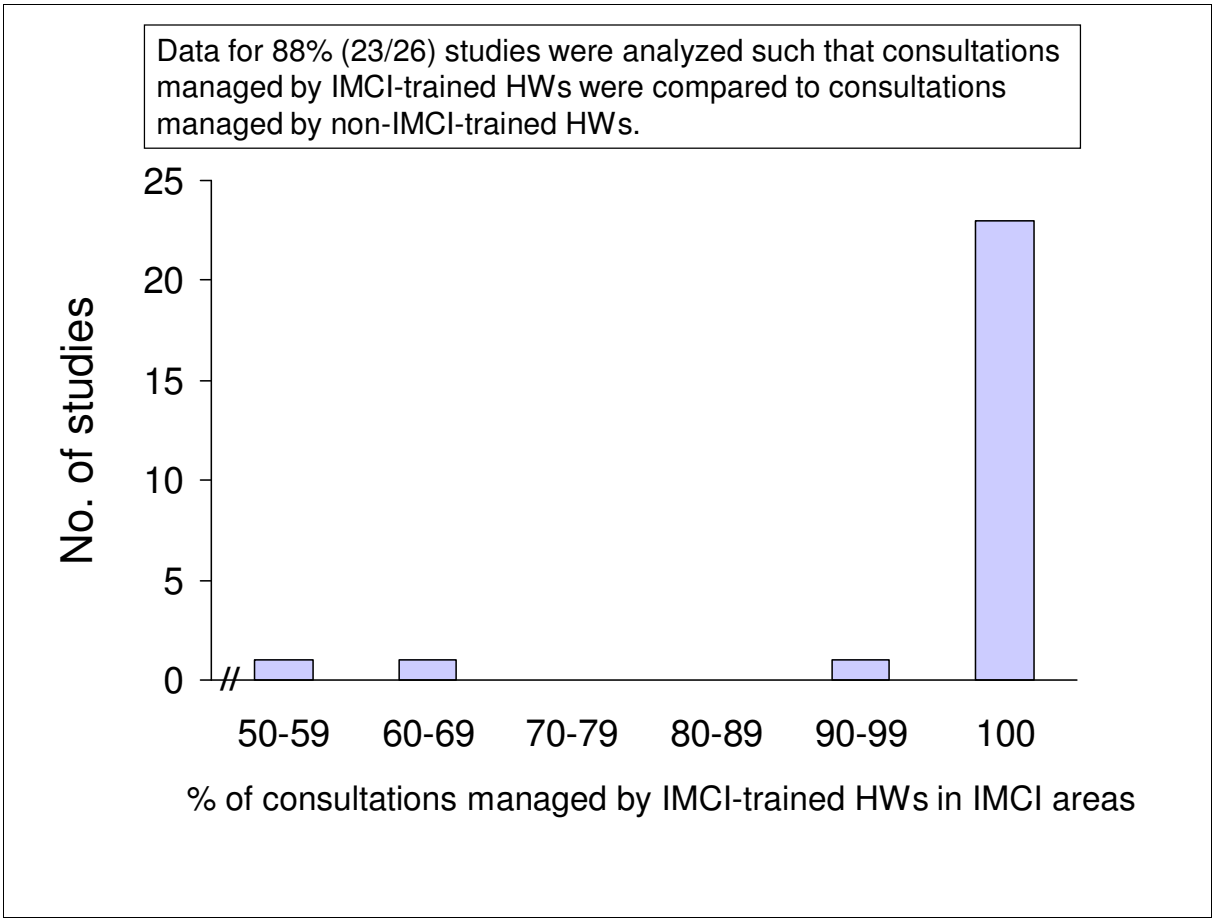


Figure 7. Distribution of 26 studies included in the primary analysis with available training coverage data, by proportion of consultations performed by IMCI-trained health workers in areas where IMCI was implemented

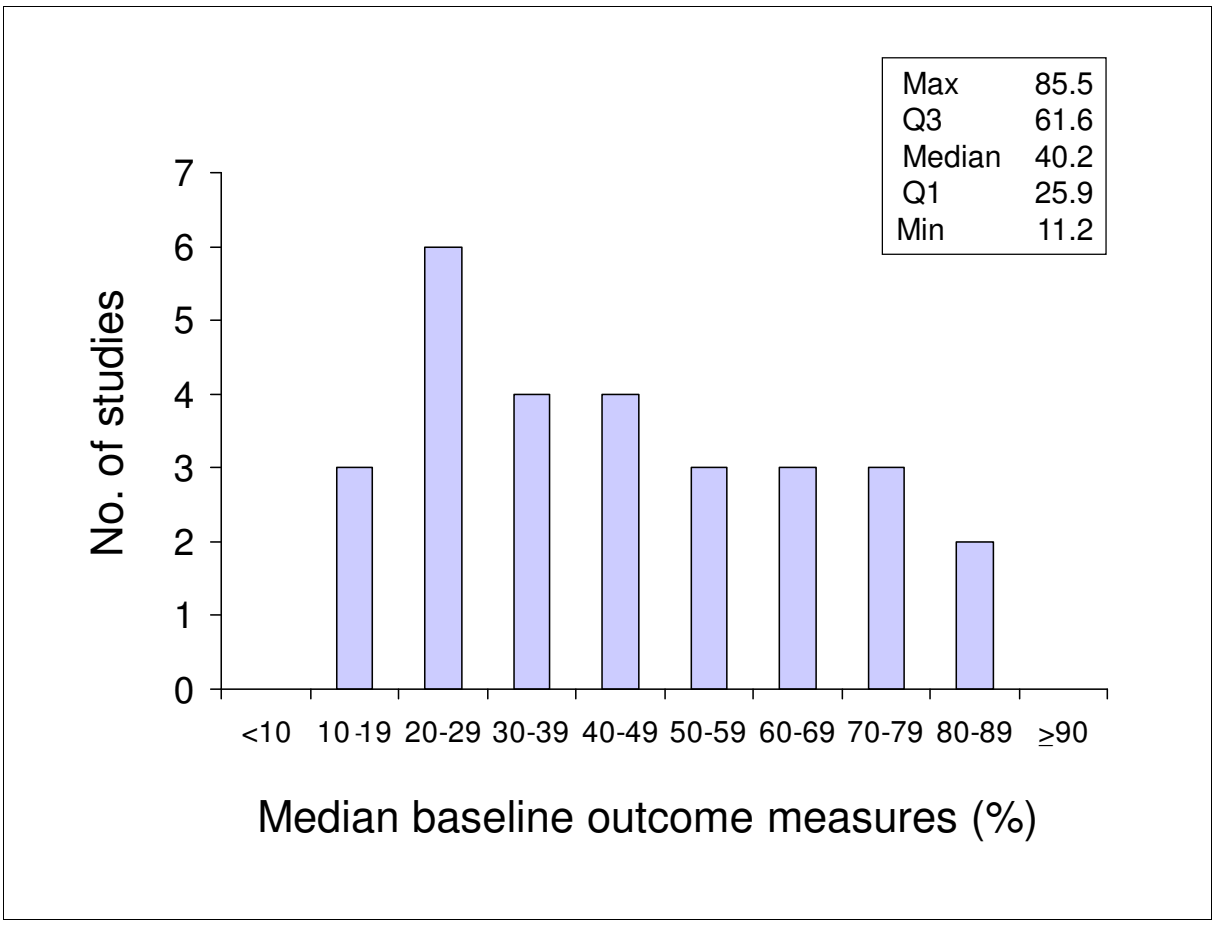


Footnotes for Figure 7.

HW = Health worker; IMCI = Integrated Management of Childhood Illness.

Note. In 26 of the 29 included studies, the training coverage in IMCI areas was available. Data for 23 of these 26 studies were analyzed such that outcomes for ill children managed by an IMCI-trained health worker were compared to those of ill children managed by a non-IMCI-trained health worker. The remaining 3 of the 29 studies did not provide data stratified by IMCI-training status, and thus were analyzed such that outcomes for ill children in an IMCI area (containing a mixture of IMCI-trained and non-IMCI trained health workers) were compared to those of ill children in a non-IMCI area (containing only non-IMCI trained health workers).

Figure 8. Distribution of the study-specific medians of baseline outcome measures[a]

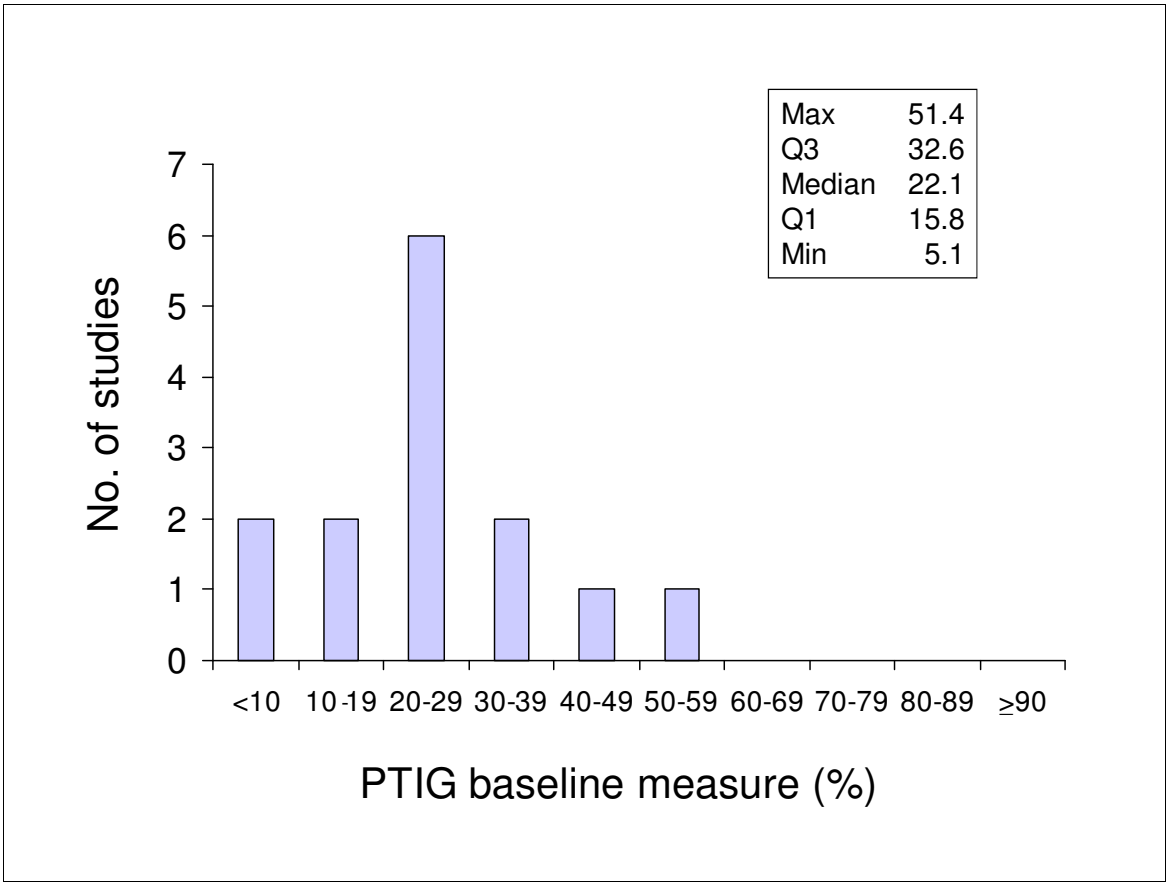


Footnotes for Figure 8.

IMCI = Integrated Management of Childhood Illness.

[a] This graph shows 28 values (one per comparison) from 26 studies (first-tier design or second-tier design) that compared IMCI training (standard or short) to no IMCI in the primary analysis. The three studies that directly compared standard training to short training were excluded because no baseline outcome measures of health worker performance were available.

Figure 9. Distribution of baseline PTIG measures[a]



Footnotes for Figure 9.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

[a] This graph shows 14 values (one per comparison) from 12 studies (first-tier design or second-tier design) that compared IMCI training (standard or short) to no IMCI in the primary analysis. The three studies that directly compared standard training to short training were excluded because no baseline outcome measures of health worker performance were available.

Figure 10a. Distribution of median effect sizes in the primary analysis (N=28 effect sizes [one per comparison] from 26 studies[a])

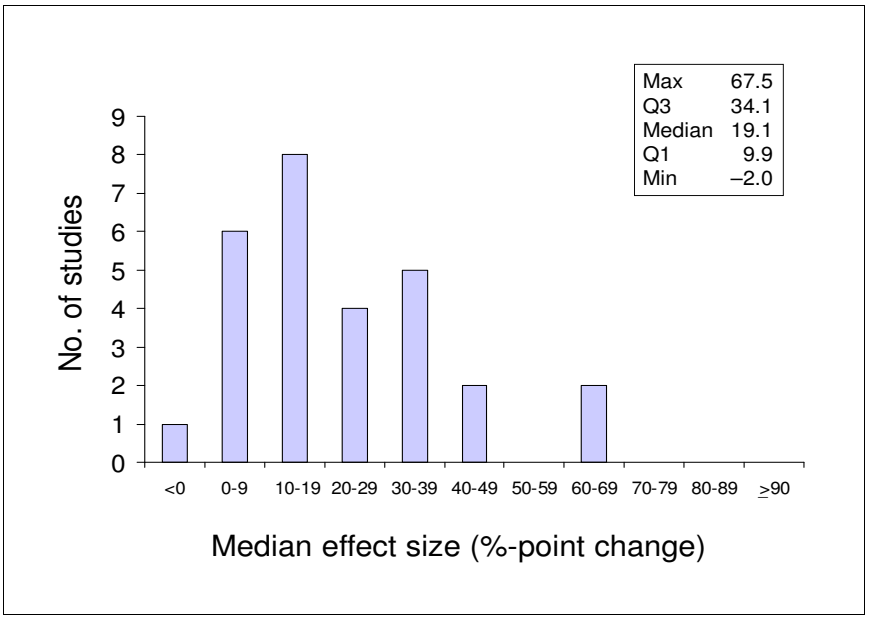
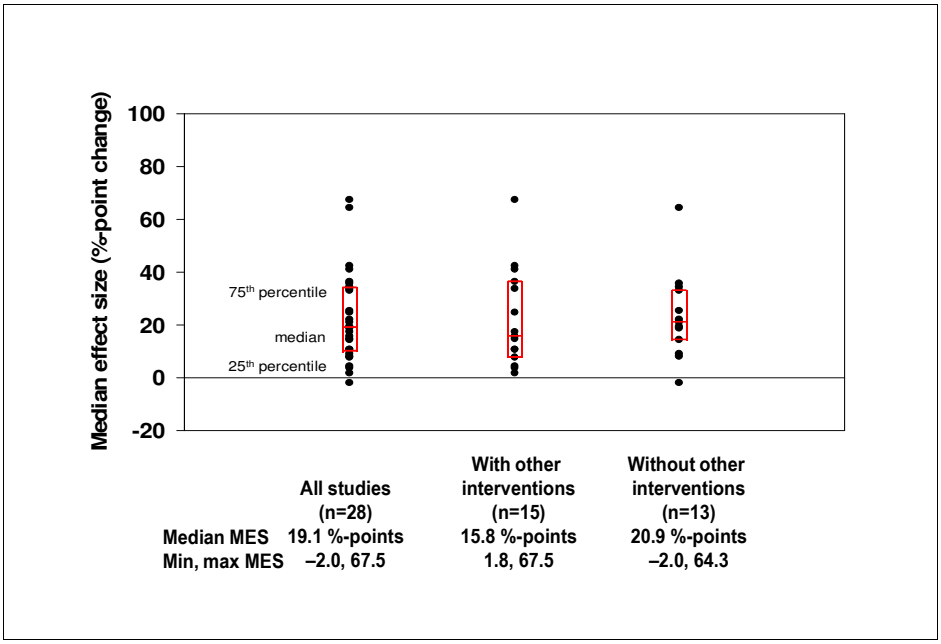


Figure 10b. Median effect sizes (MES) in the primary analysis stratified by presence of other interventions (N=28 effect sizes [one per comparison] from 26 studies[a])



Footnotes for Figures 10a–10b.

IMCI = Integrated Management of Childhood Illness.

[a] Studies comparing standard training versus no IMCI or short training versus no IMCI. Baseline values shown in Figure 8.

Figure 11a. Distribution of PTIG effect sizes in primary analysis (N=14 effect sizes [one per comparison] from 12 studies[a])

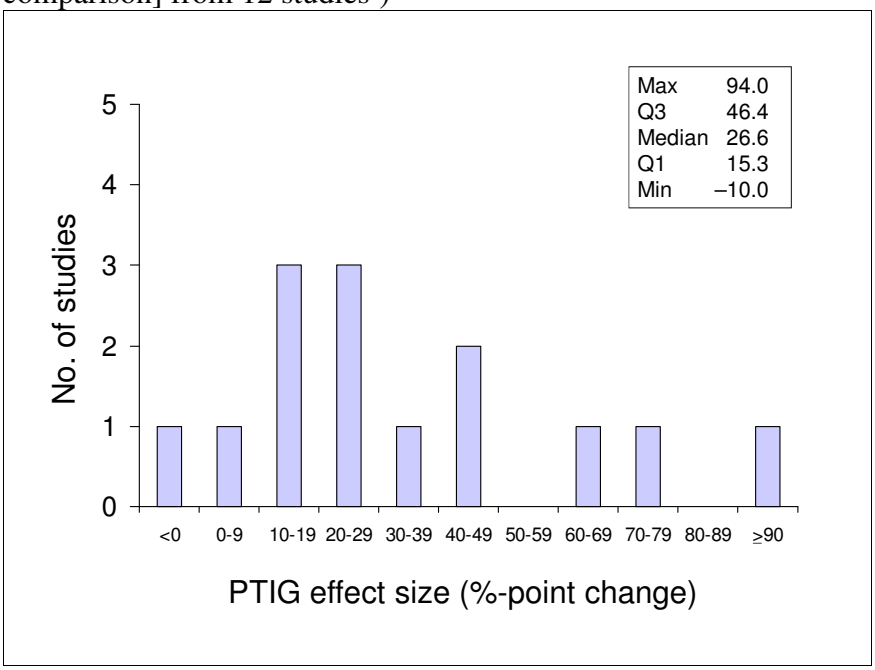
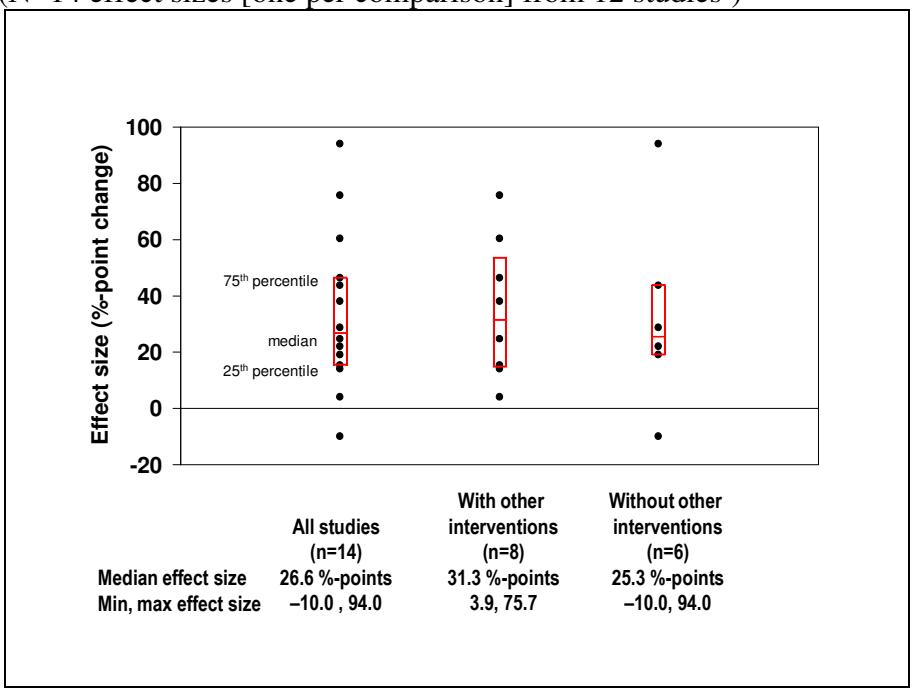


Figure 11b. PTIG effect sizes in the primary analysis stratified by presence of other interventions (N=14 effect sizes [one per comparison] from 12 studies[a])



Footnotes for Figures 11a–11b.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

[a] Studies comparing standard vs. no IMCI or short vs. no IMCI. Baseline values in Figure 9.

Figure 12a. Distribution of 34 median post-intervention values from 29 studies included in the primary analysis (28 values from 26 indirect comparison studies and 6 values from 3 direct comparison studies)

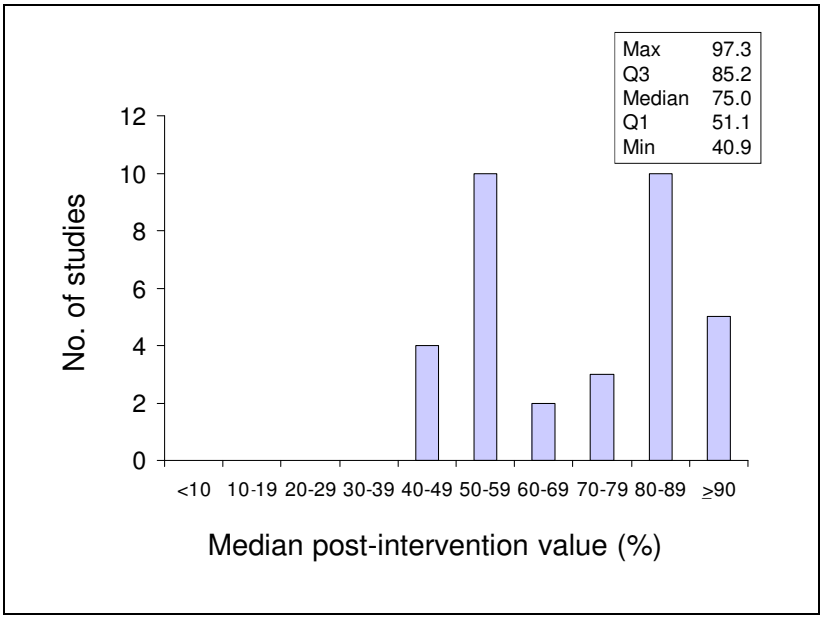
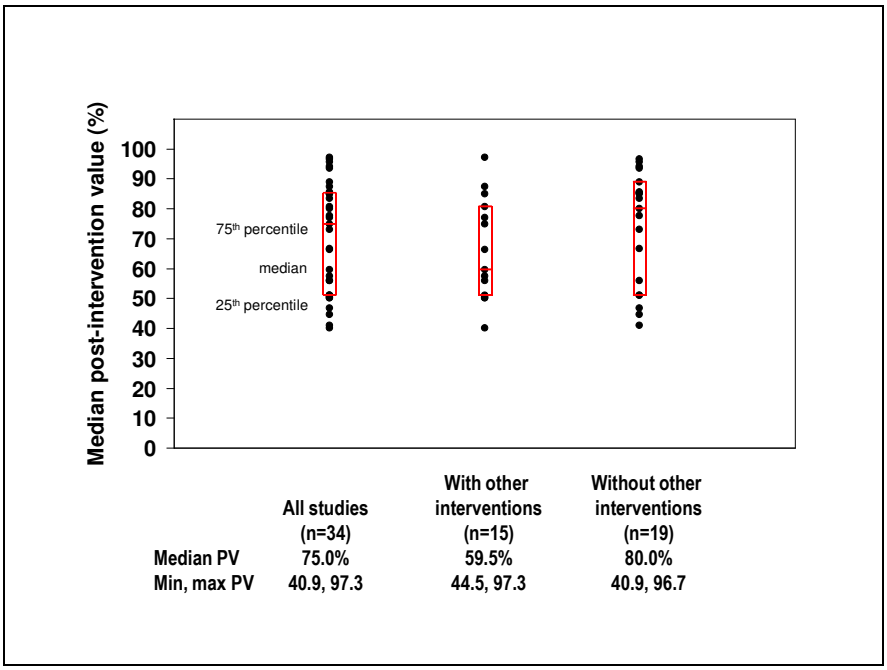


Figure 12b. Median post-intervention values in the primary analysis stratified by the presence of other interventions (28 values from 26 indirect comparison studies and 6 values from 3 direct comparison studies)



Footnotes for Figures 12a–12b.

IMCI = Integrated Management of Childhood Illness; PV = post-intervention value.

Figure 13a. Distribution of 20 post-intervention PTIG values from 15 studies included the primary analysis (14 values from 12 indirect comparison studies and 6 values from 3 direct comparison studies)

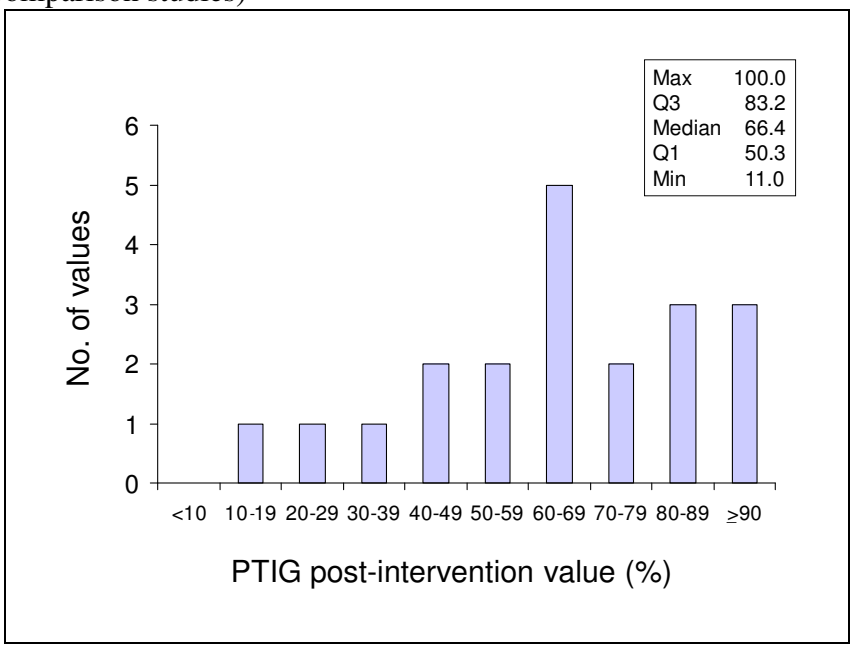
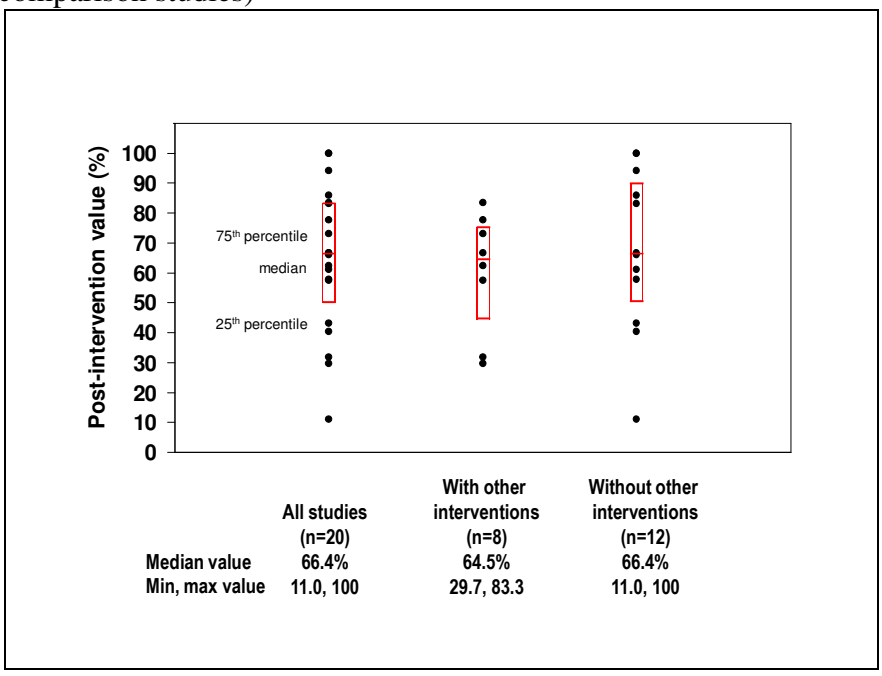


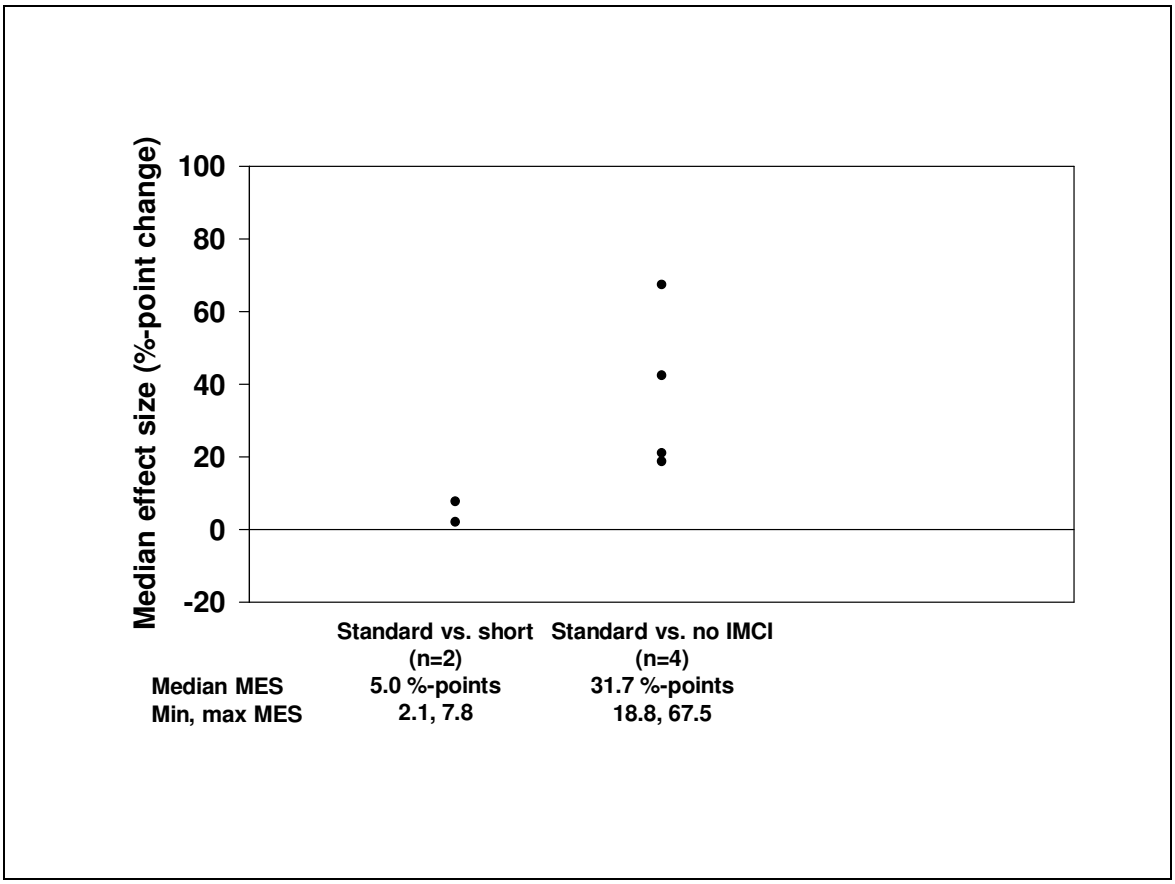
Figure 13b. Post-intervention PTIG values in the primary analysis stratified by the presence of other interventions (14 values from 12 indirect comparison studies and 6 values from 3 direct comparison studies)



Footnotes for Figures 13a–13b.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

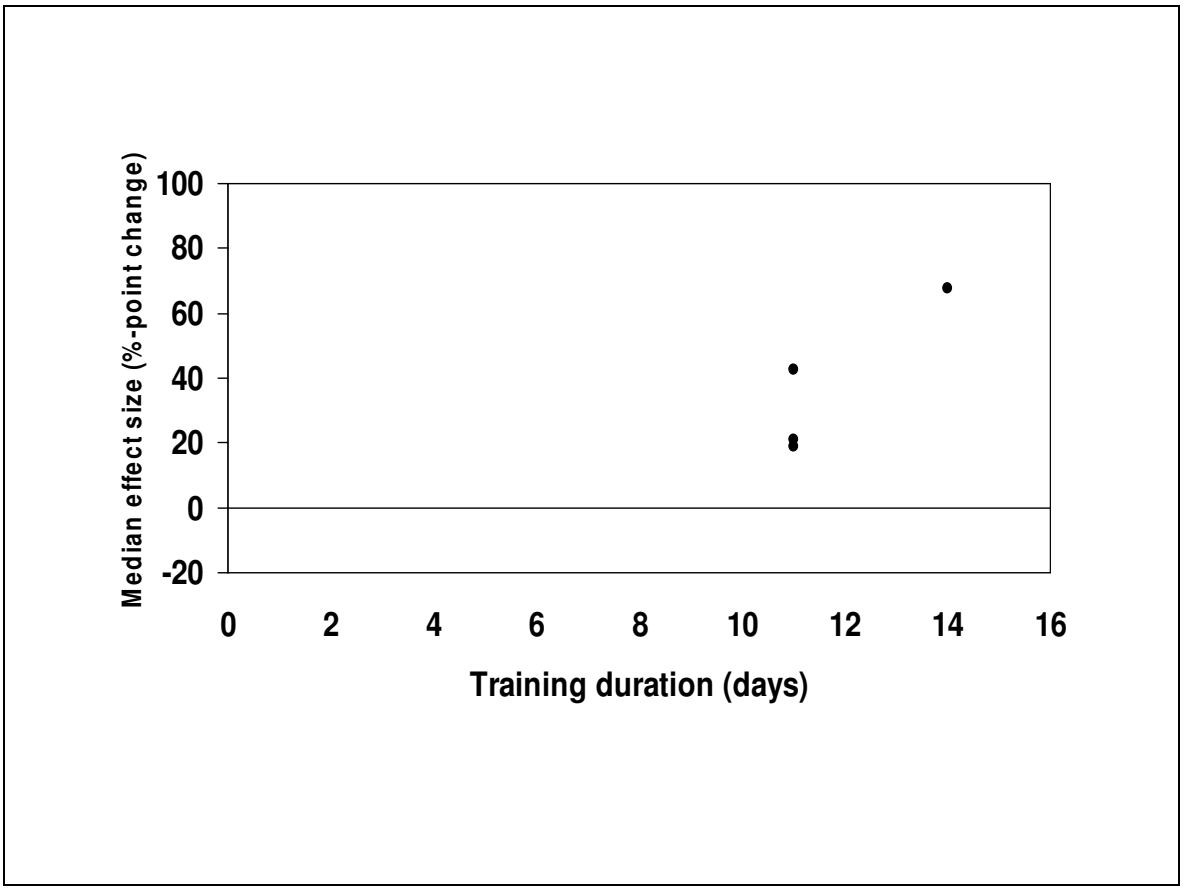
Figure 14. Main analysis: Median effect size stratified by comparison type (N=6 effect sizes [one per comparison] from 5 studies included in the primary analysis)



Footnotes for Figure 14.

IMCI = Integrated Management of Childhood Illness; MES = median effect size.

Figure 15. Main analysis: Median effect size versus training duration (N=4 effect sizes [one per comparison] from 3 studies[a] included in the primary analysis)



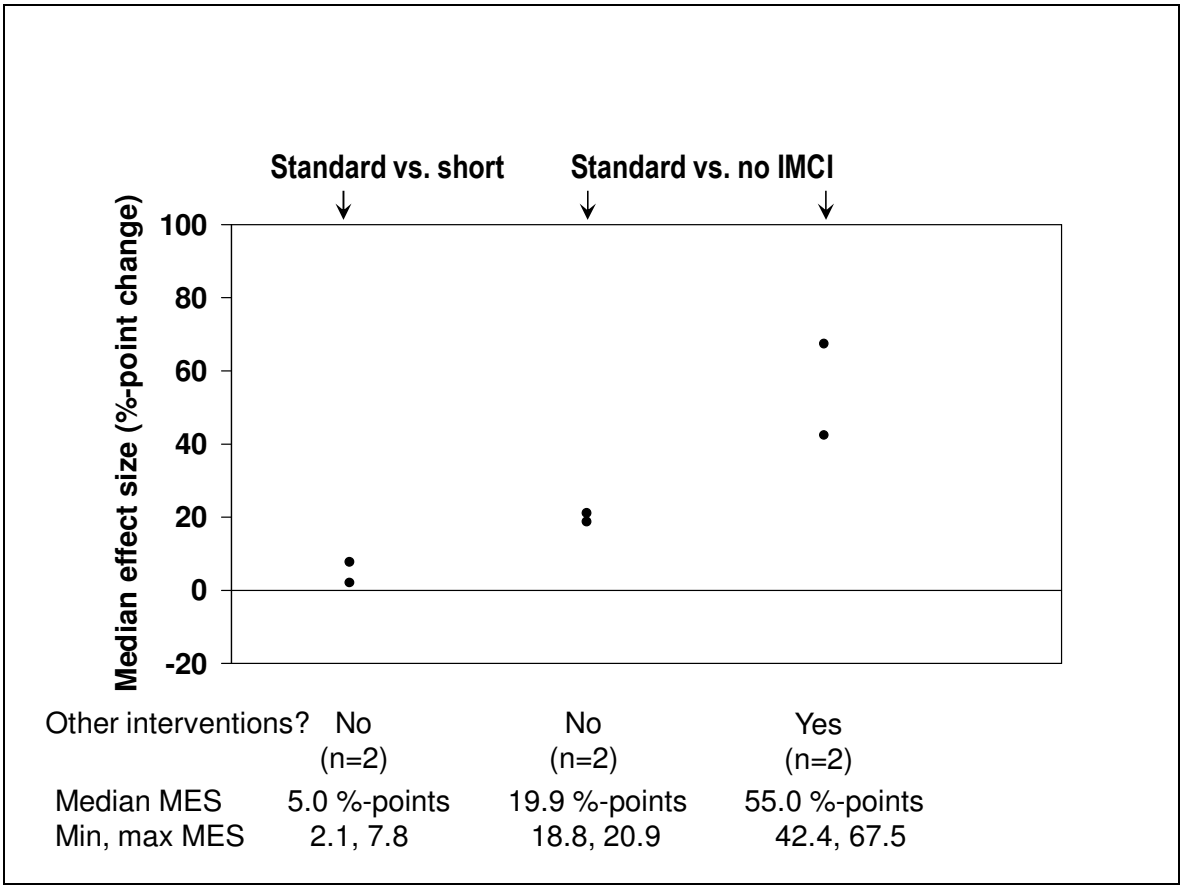
Footnotes for Figure 15.

IMCI = Integrated Management of Childhood Illness.

[a] Studies comparing standard versus no IMCI or short versus no IMCI.

Note. Slope not determined because there were only 4 effect sizes.

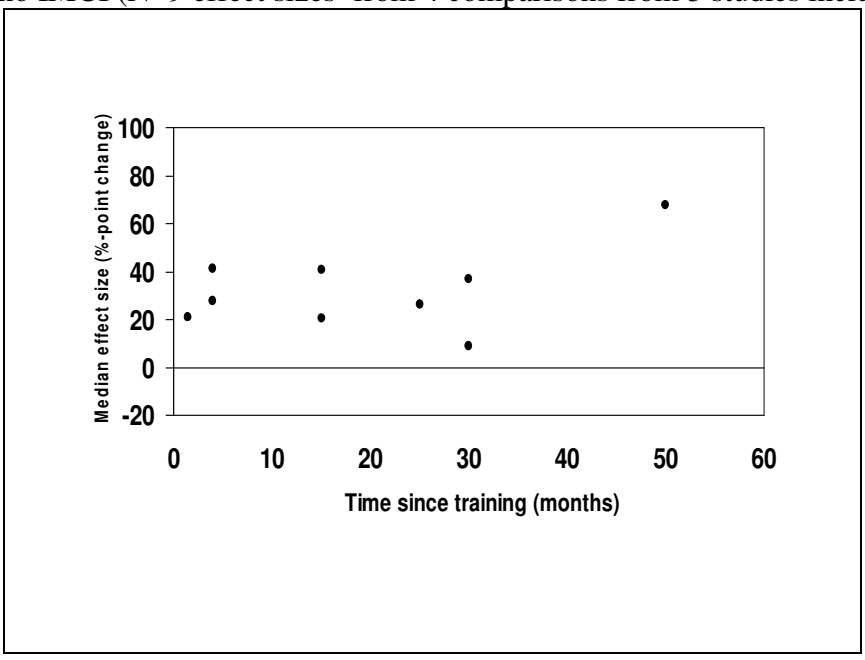
Figure 16. Main analysis: Median effect size stratified by comparison type and presence of other interventions (N=6 effect sizes [one per comparison] from 5 studies included in the primary analysis)



Footnotes for Figure 16.

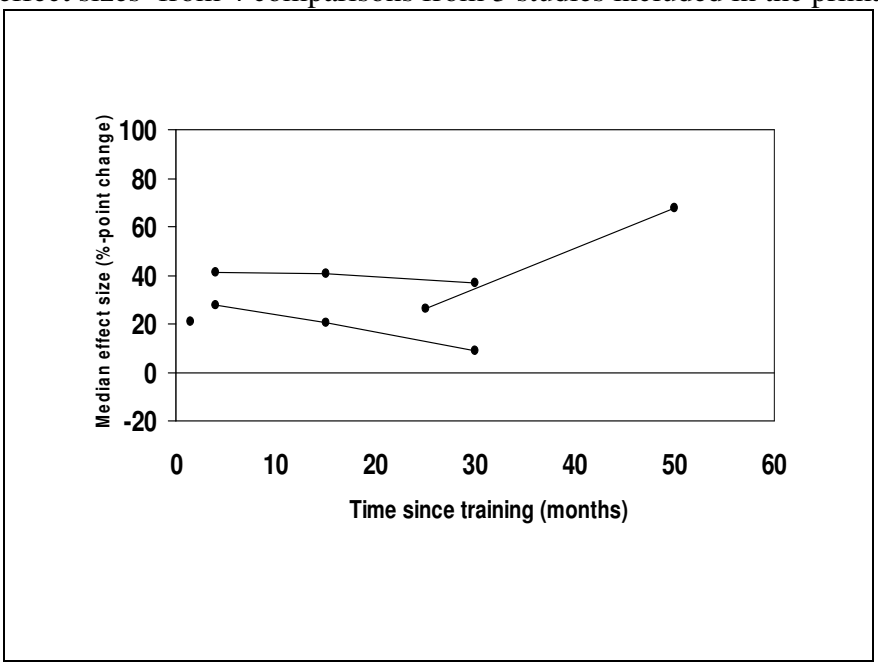
IMCI = Integrated Management of Childhood Illness; MES = median effect size.

Figure 17a. Main analysis: Median effect size versus time since initial training, standard versus no IMCI (N=9 effect sizes[a] from 4 comparisons from 3 studies included in the primary analysis)



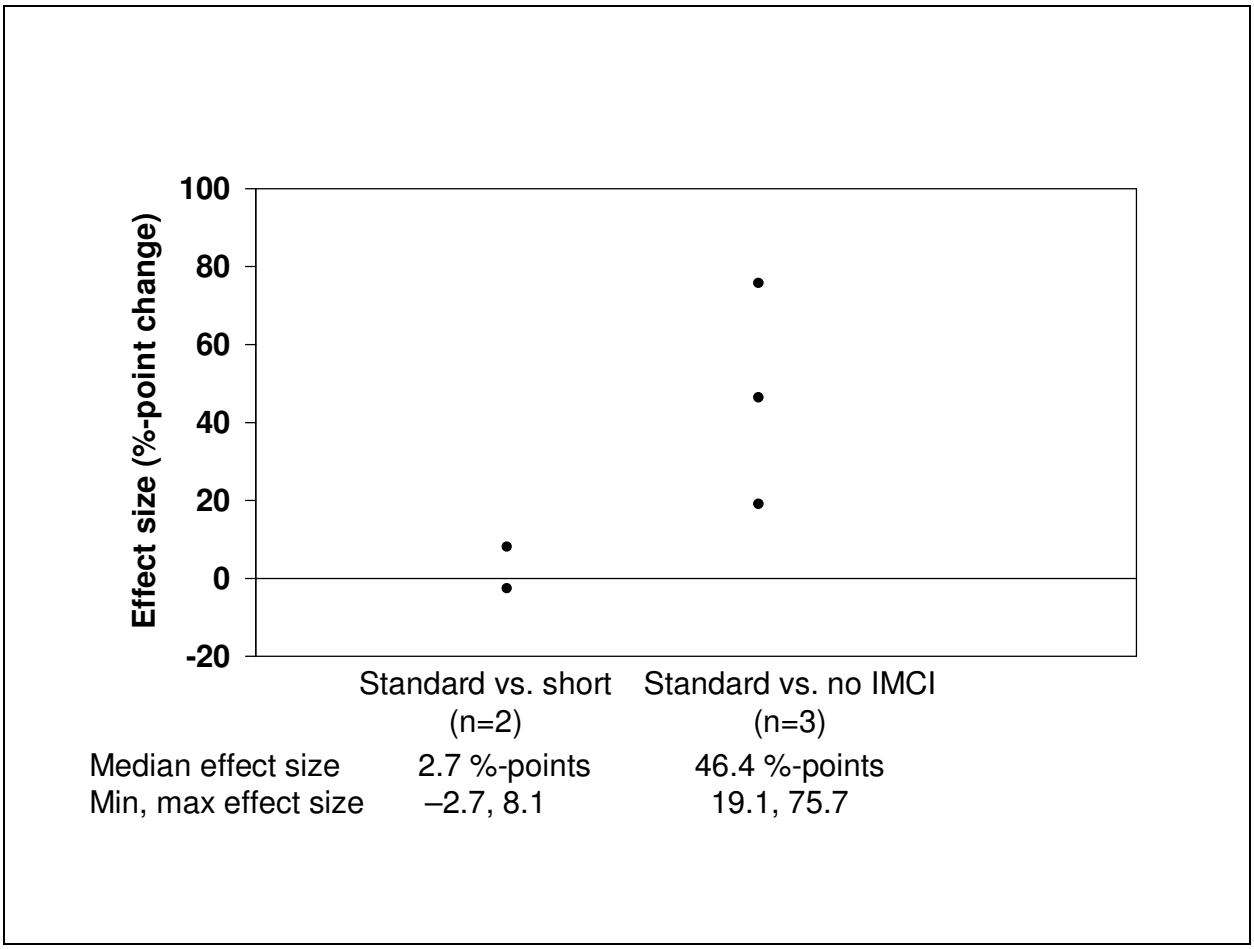
Note. Slope not determined because there were only 9 effect sizes.

Figure 17b. Main analysis: Median effect size versus time since initial training, standard versus no IMCI (multiple time points from the same comparison per study are connected by lines) (N=9 effect sizes[a] from 4 comparisons from 3 studies included in the primary analysis)



Footnotes for Figures 17a–17b.
IMCI = Integrated Management of Childhood Illness.
[a] Includes 6 effect sizes for 2 comparisons from Study ID 24 for 4, 15, and 30 months since training that used measures approximated by regression models (See Methods for details.)

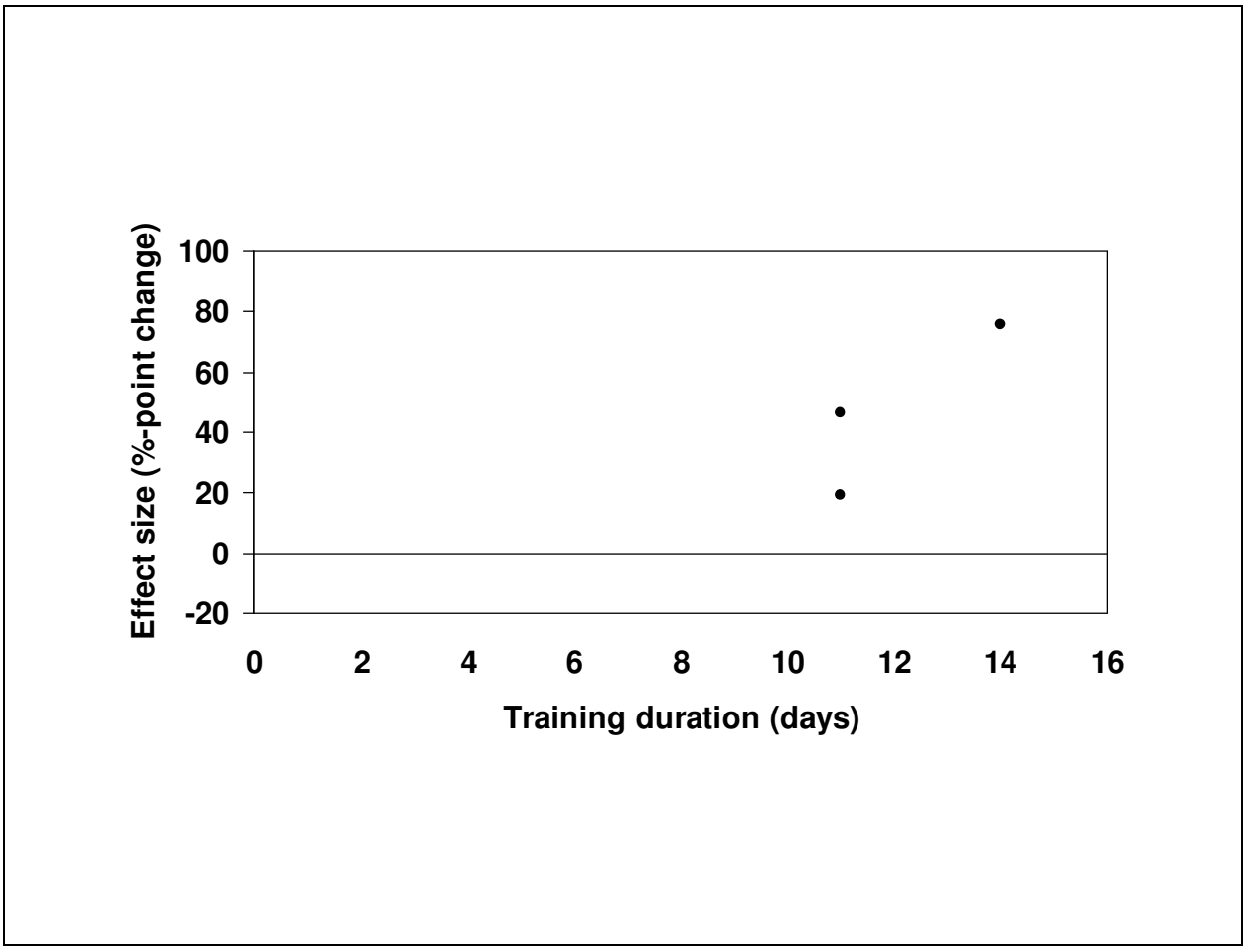
Figure 18. Main analysis: PTIG effect size stratified by comparison type (N=5 effect sizes [one per comparison] from 4 studies included in the primary analysis)



Footnotes for Figure 18.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

Figure 19. Main analysis: PTIG effect size versus training duration (N=3 effect sizes [one per comparison] from 2 studies[a] included in the primary analysis)



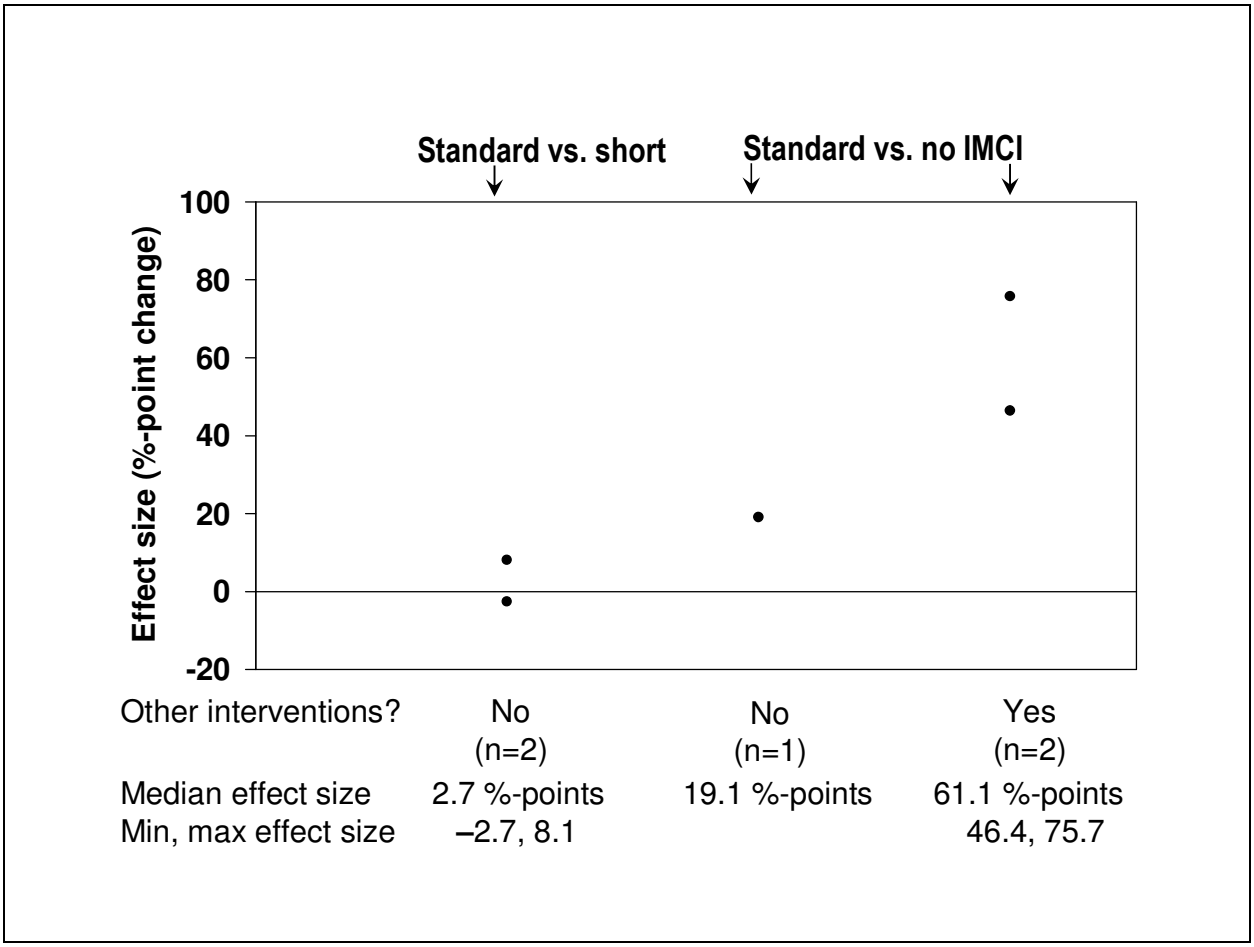
Footnotes for Figure 19.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

[a] Studies comparing standard versus no IMCI or short versus no IMCI.

Note. Slope not determined because there were only 3 effect sizes.

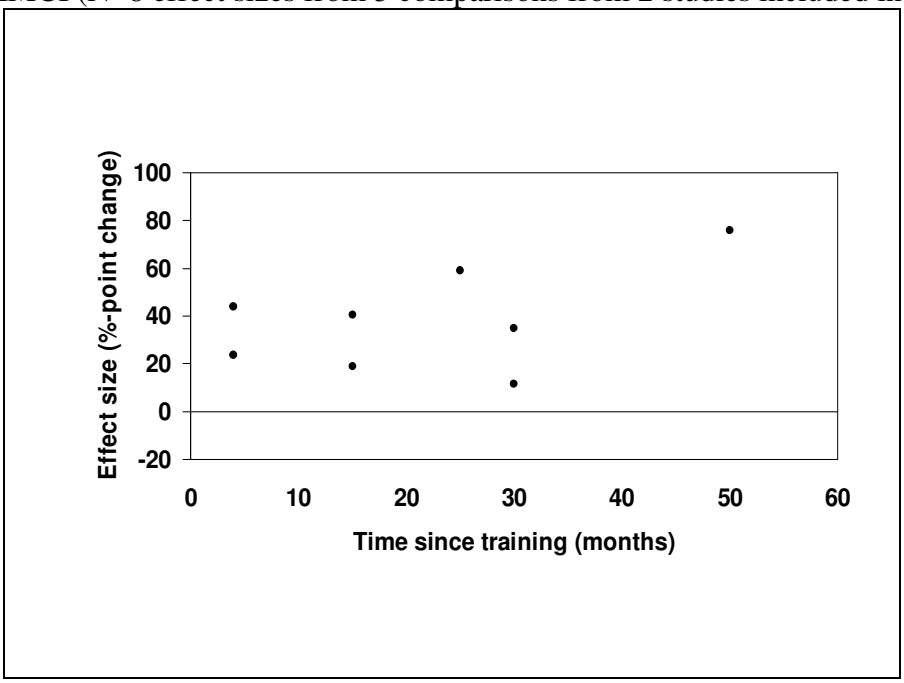
Figure 20. Main analysis: PTIG effect size, stratified by comparison type and presence of other interventions (5 effect sizes [one per comparison] from 4 studies included in the primary analysis)



Footnotes for Figure 20.

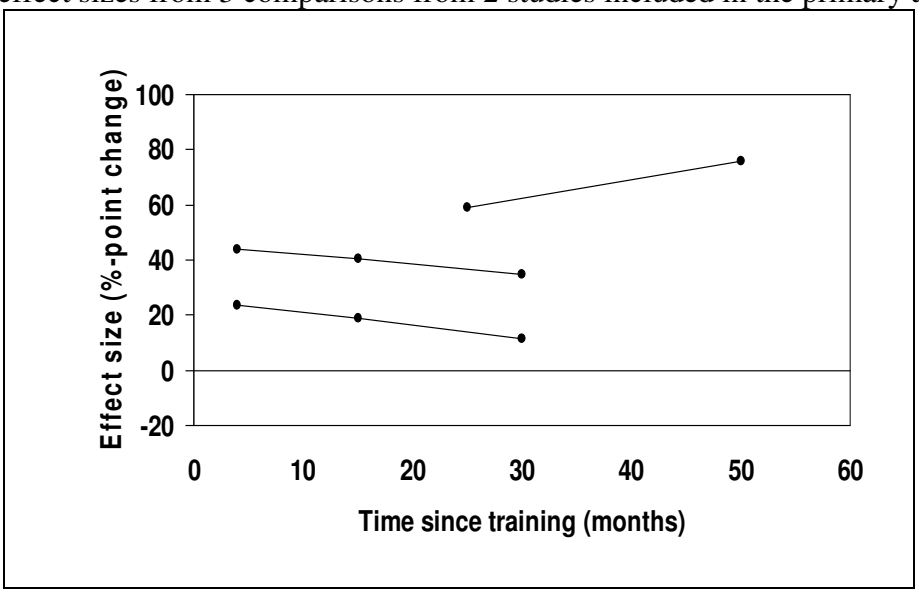
IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

Figure 21a. Main analysis: PTIG effect size versus time since initial training, standard versus no IMCI (N=8 effect sizes from 3 comparisons from 2 studies included in the primary analysis)



Note. Slope not determined because there were only 8 effect sizes.

Figure 21b. Main analysis: PTIG effect size versus time since initial training, standard versus no IMCI (multiple time points from the same comparison per study are connected by lines) (N=8 effect sizes from 3 comparisons from 2 studies included in the primary analysis)

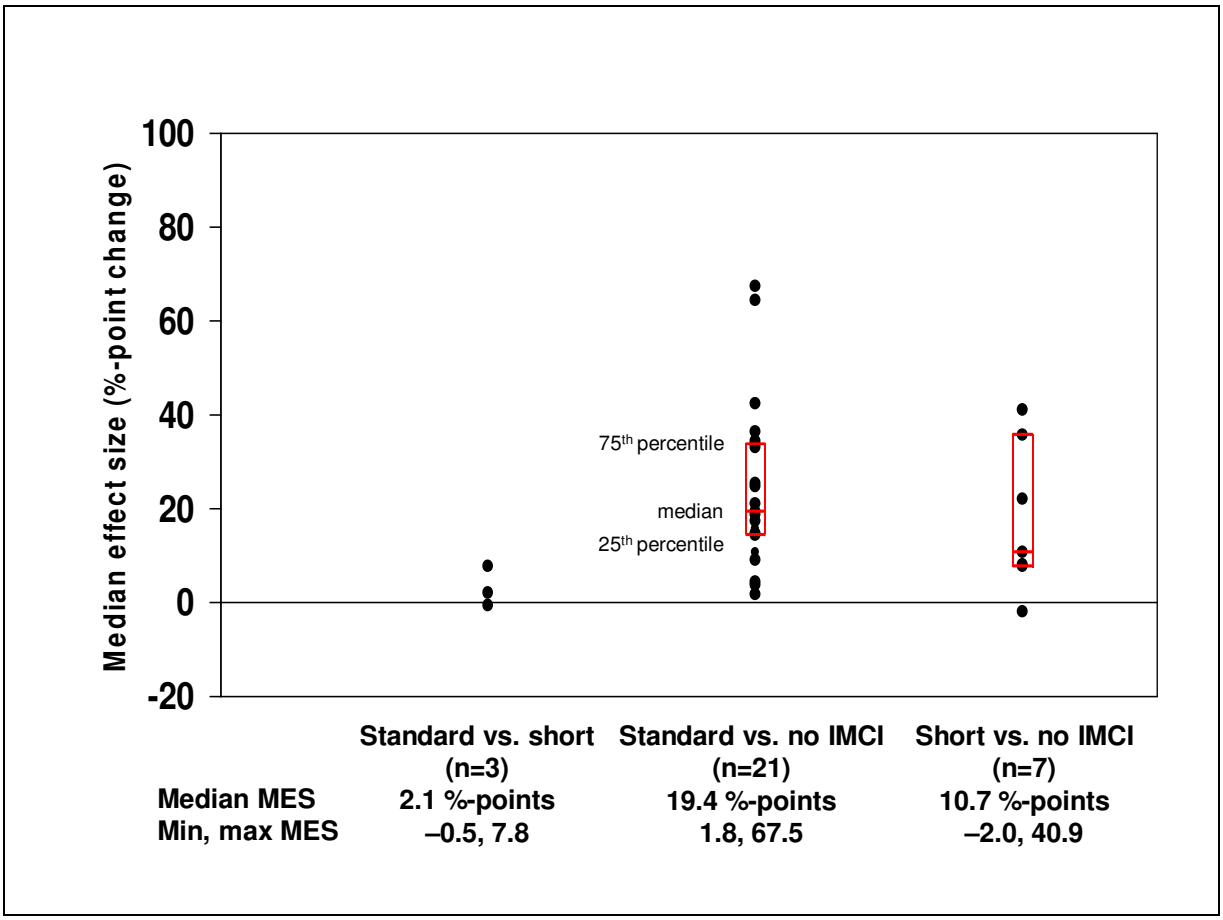


Footnotes for Figures 21a–21b.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

[a] Includes 6 effect sizes for 2 comparisons from Study ID 24 for 4, 15, and 30 months since training that used measures approximated by regression models (See Methods for details.)

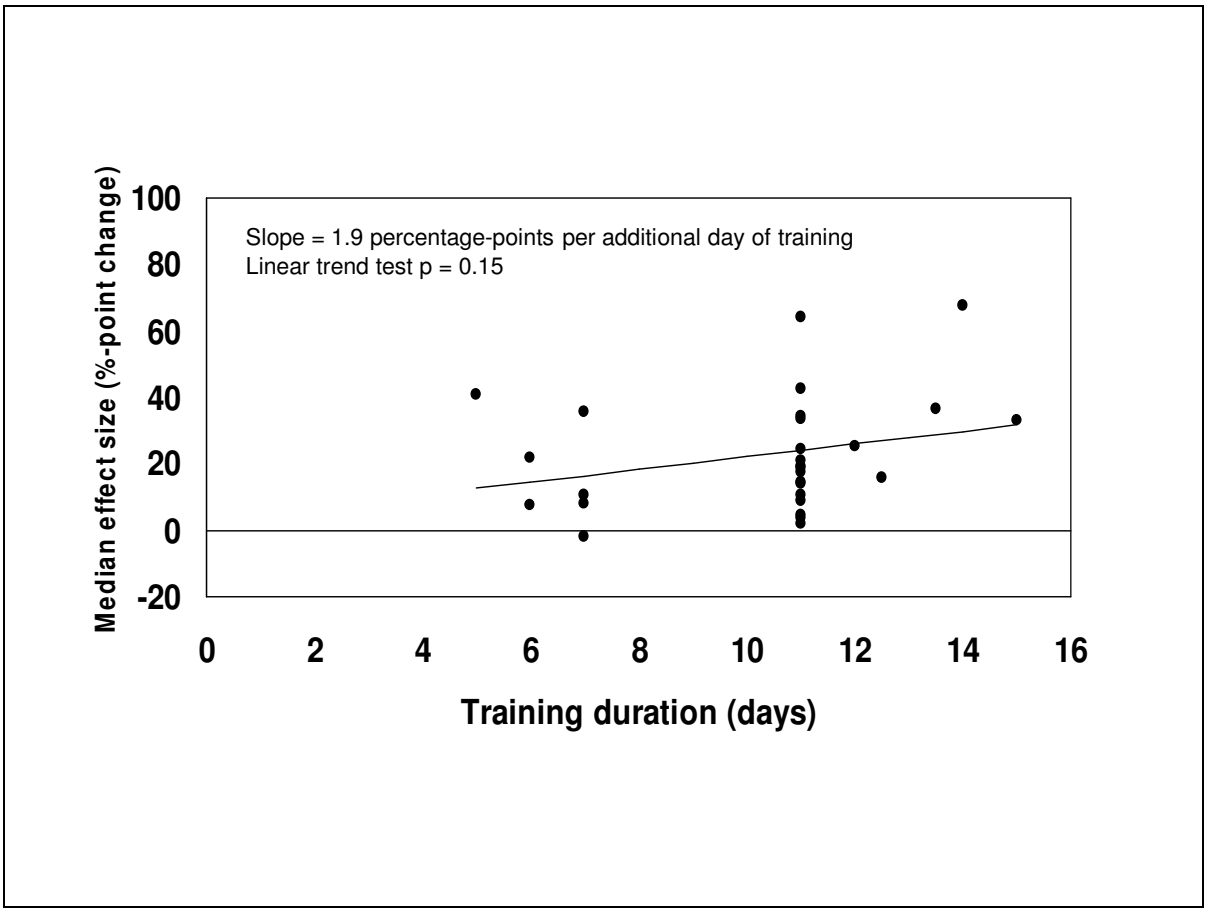
Figure 22. Sensitivity analysis: Median effect size stratified by comparison type (N=31 effect sizes [one per comparison] from 29 studies included in the primary analysis)



Footnotes for Figure 22.

IMCI = Integrated Management of Childhood Illness; MES = median effect size.

Figure 23. Sensitivity analysis: Median effect size versus training duration (N=28 effect sizes [one per comparison] from 26 studies[a] included in the primary analysis)

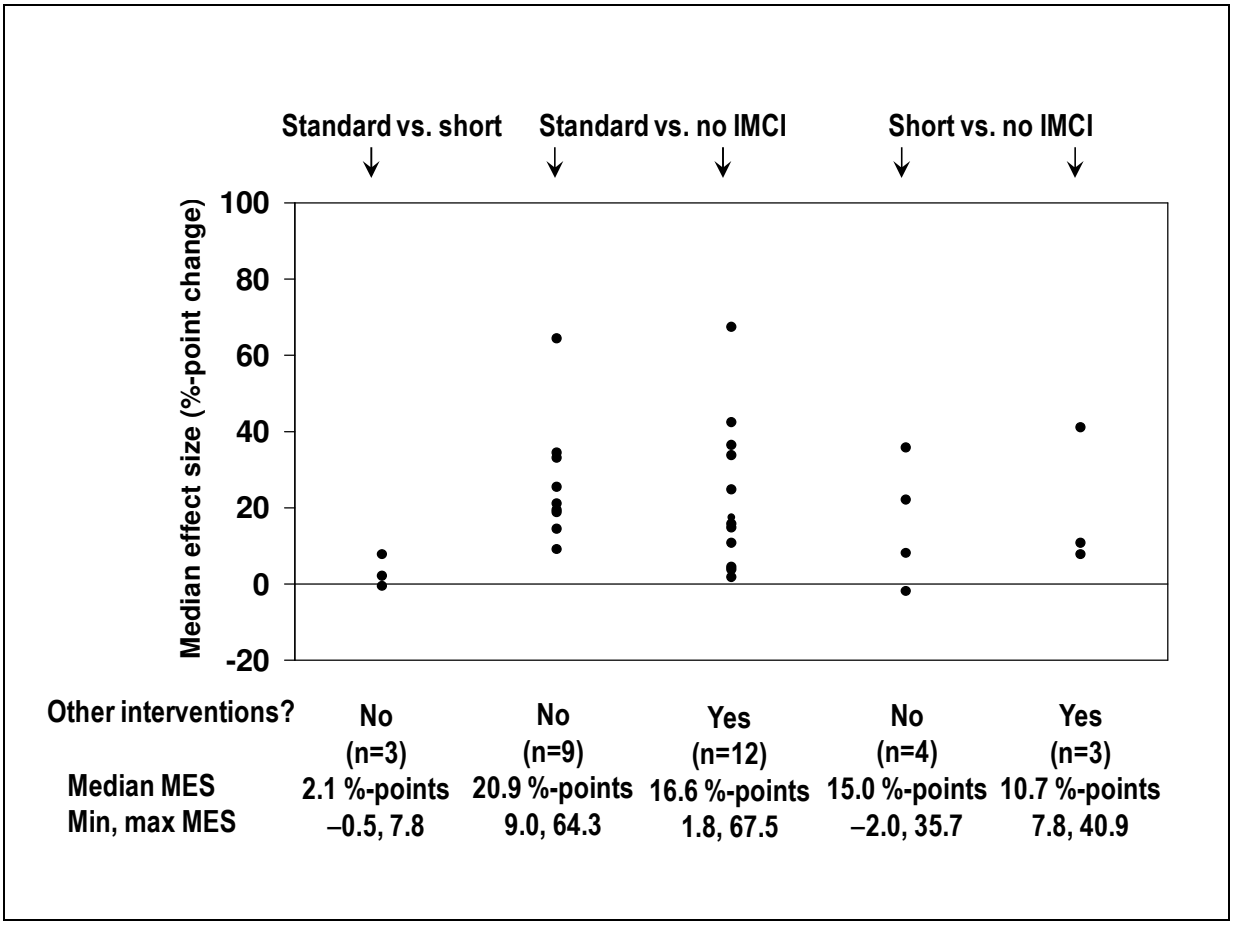


Footnotes for Figure 23.

IMCI = Integrated Management of Childhood Illness; MES = median effect size.

[a] Studies comparing standard versus no IMCI or short versus no IMCI.

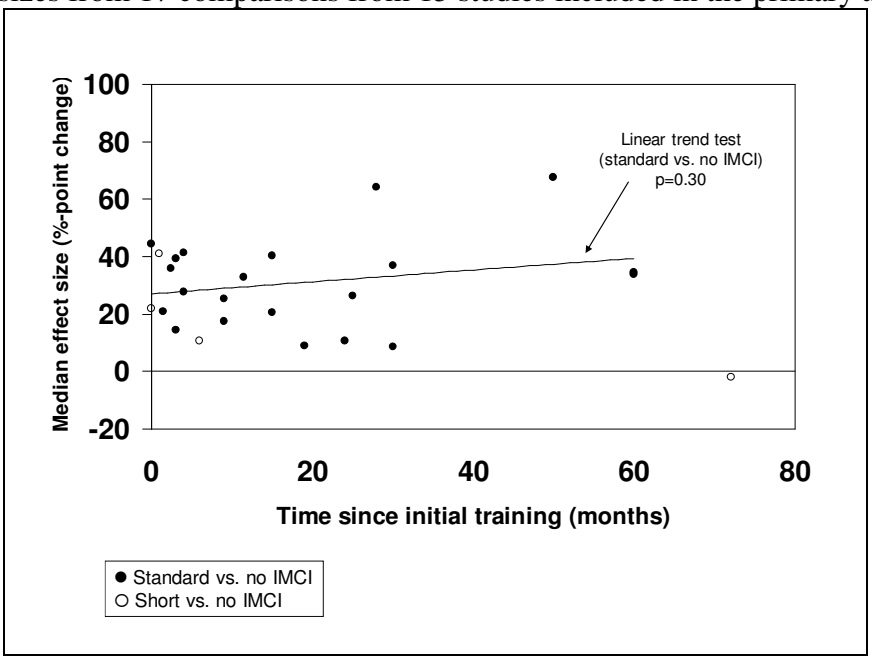
Figure 24. Sensitivity analysis: Median effect size stratified by comparison type and presence of other interventions (N=31 effect sizes [one per comparison] from 29 studies included in the primary analysis)



Footnotes for Figure 24.

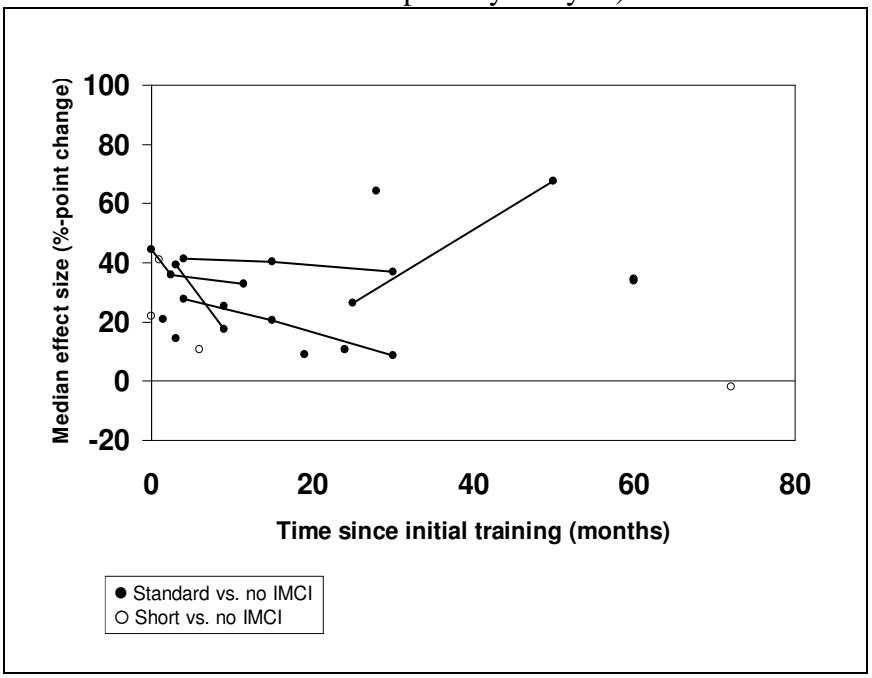
IMCI = Integrated Management of Childhood Illness; MES = median effect size.

Figure 25a. Sensitivity analysis: Median effect size versus time since initial training (N=25 effect sizes from 17 comparisons from 15 studies included in the primary analysis)



Note. Slope for short vs. no IMCI not determined because there were only 4 effect sizes.

Figure 25b. Sensitivity analysis: MES vs. time since initial training (multiple time points from the same comparison per study are connected by lines) (N=25 effect sizes from 17 comparisons from 15 studies included in the primary analysis)

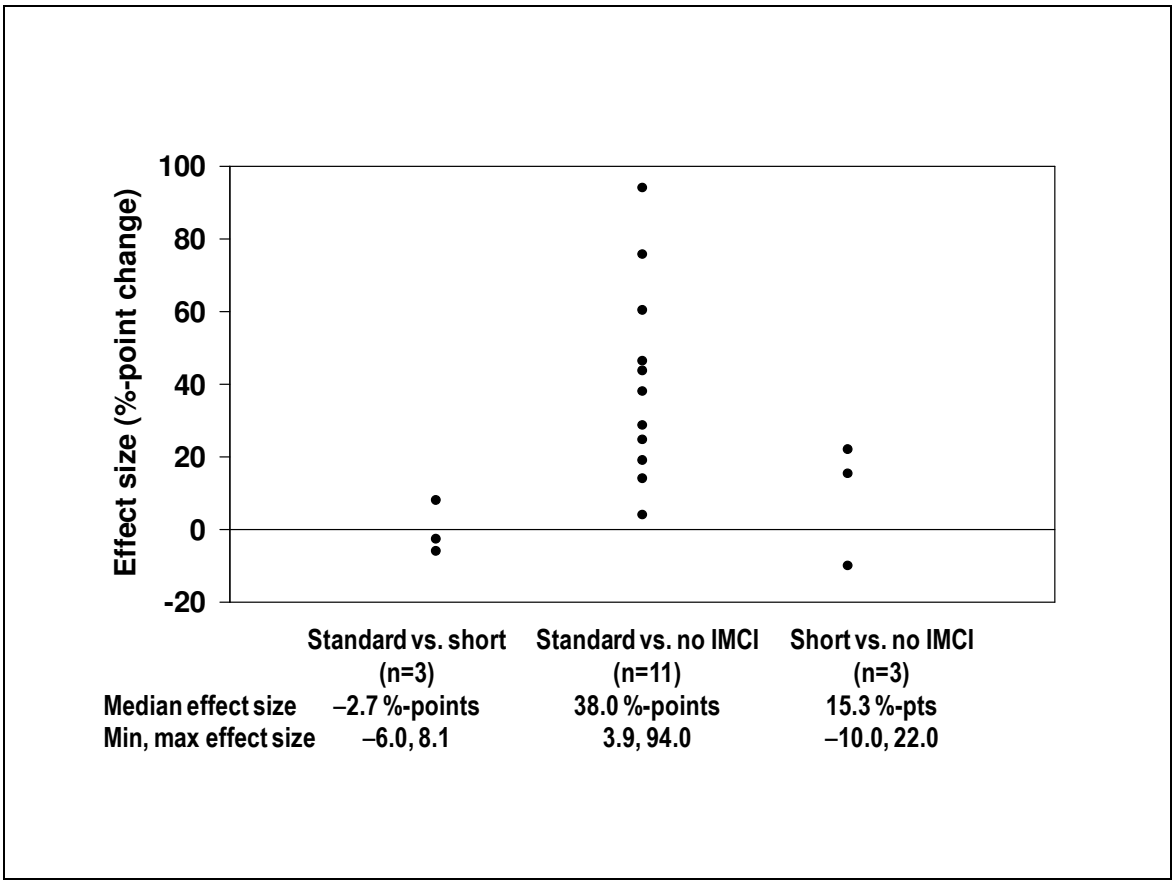


Footnotes for Figure 25a–25b.

IMCI = Integrated Management of Childhood Illness; MES = median effect size.

[a] Includes 6 effect sizes for 2 comparisons from Study ID 24 for 4, 15, and 30 months since training that used measures approximated by regression models (See Methods for details.)

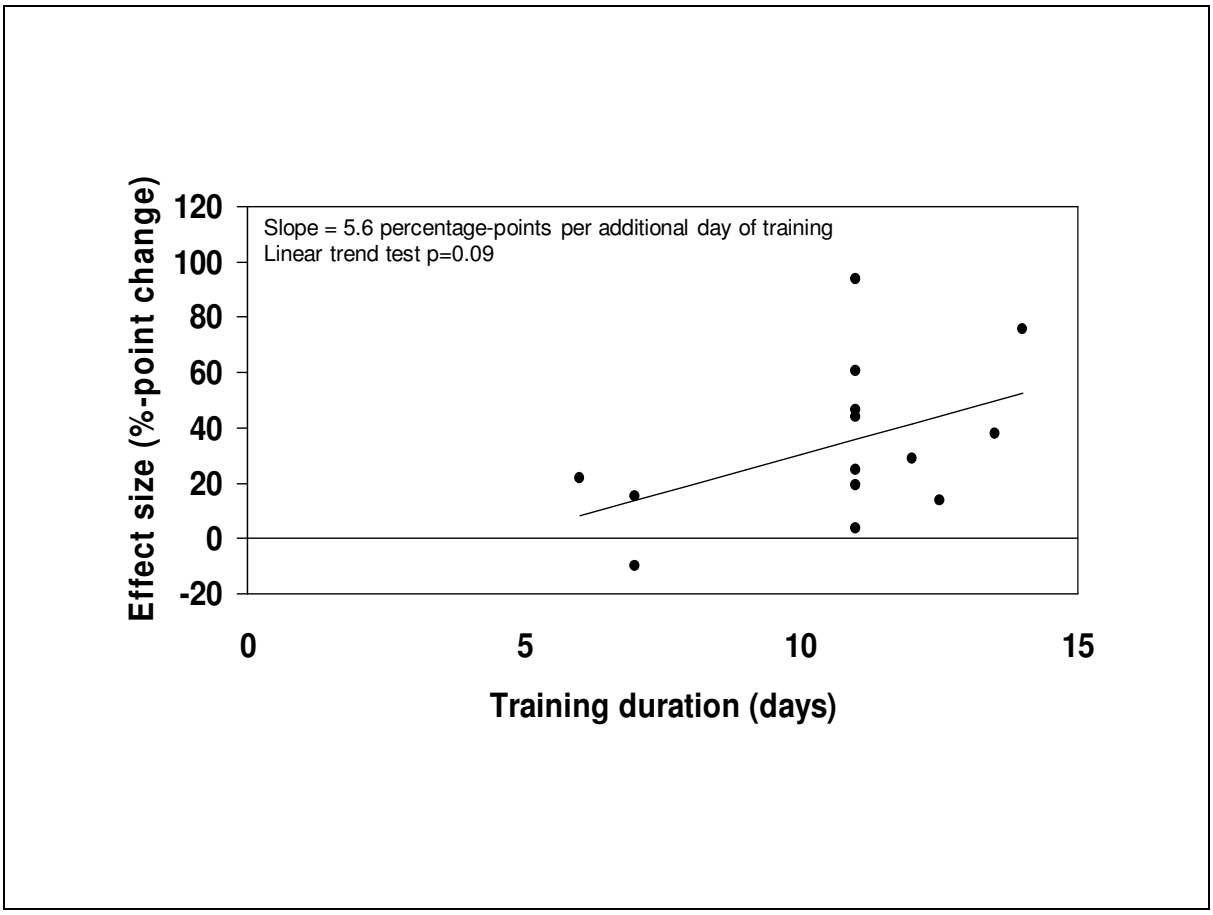
Figure 26. Sensitivity analysis: PTIG effect size stratified by comparison type (N=17 effect sizes [one per comparison] from 15 studies included in the primary analysis)



Footnotes for Figure 26.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

Figure 27. Sensitivity analysis: PTIG effect size versus training duration (N=14 effect sizes [one per comparison] from 12 studies[a] included in the primary analysis

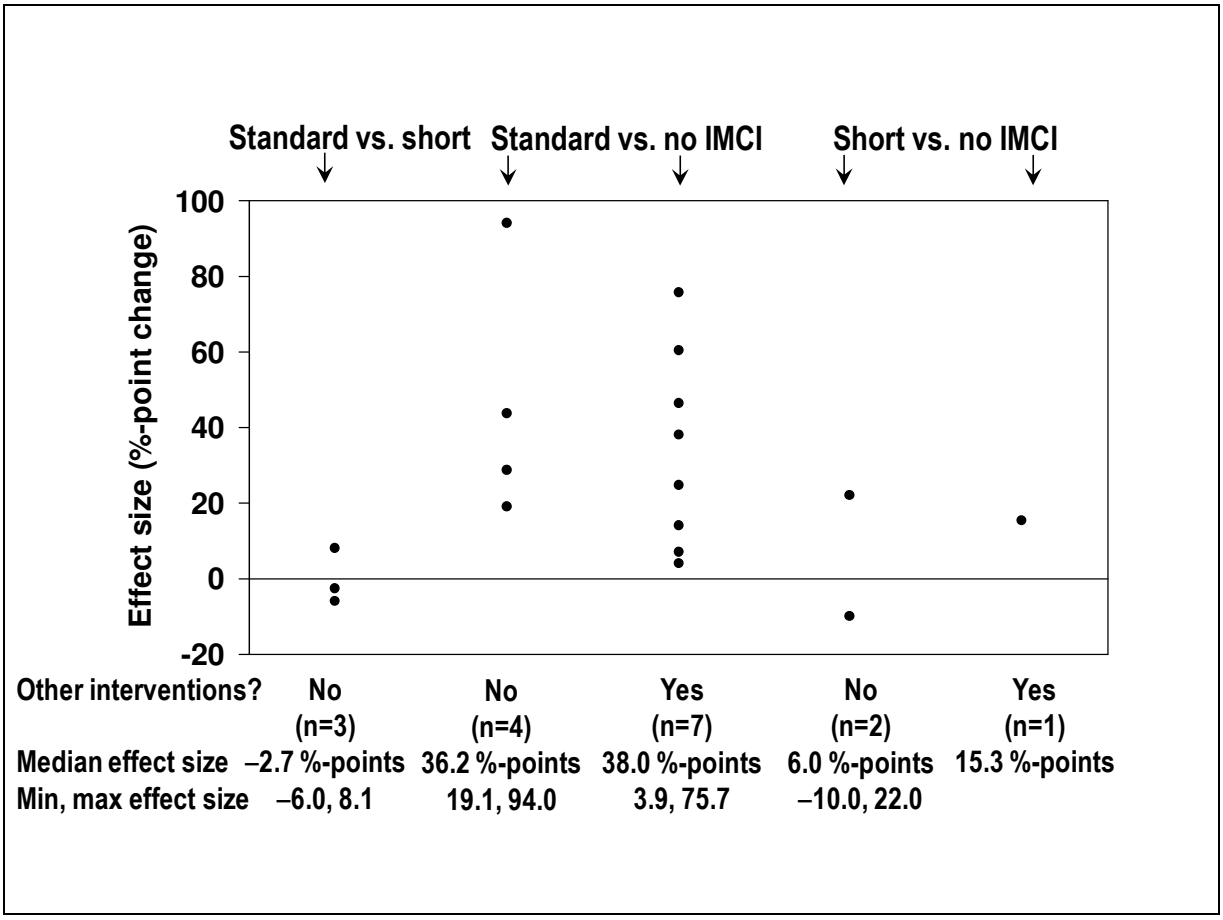


Footnotes for Figure 27.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

[a] Studies comparing standard versus no IMCI or short versus no IMCI.

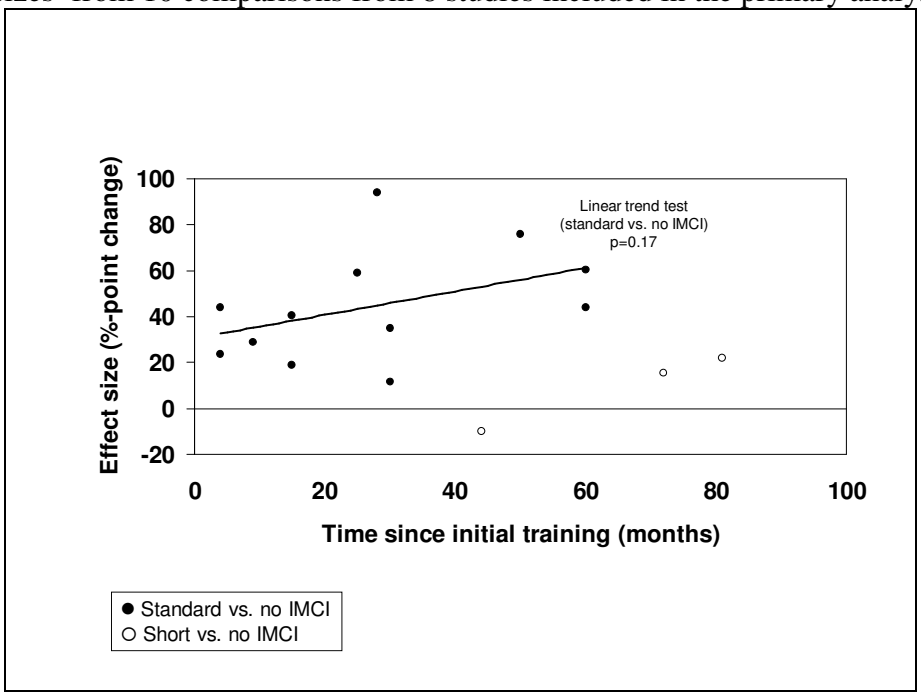
Figure 28. Sensitivity analysis: PTIG effect size stratified by comparison type and presence of other interventions (N=17 effect sizes [one per comparison] from 15 studies included in the primary analysis)



Footnotes for Figure 28.

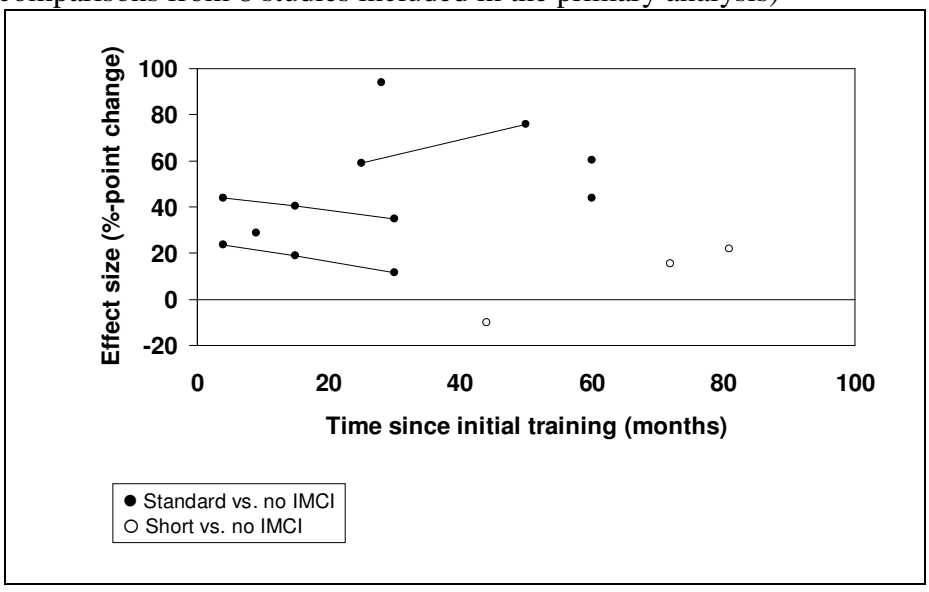
IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

Figure 29a. Sensitivity analysis: PTIG effect size versus time since initial training (N=15 effect sizes[a] from 10 comparisons from 8 studies included in the primary analysis)



Note. Slope for short vs. no IMCI not determined because there were only 3 effect sizes.

Figure 29b. Sensitivity analysis: PTIG effect size versus time since initial training (multiple time points from the same comparison per study are connected by lines) (N=15 effect sizes[a] from 10 comparisons from 8 studies included in the primary analysis)



Footnotes for Figures 29a–29b.

IMCI = Integrated Management of Childhood Illness; PTIG = patient treated according to IMCI guidelines.

[a] Includes 6 effect sizes for 2 comparisons from Study ID 24 for 4, 15, and 30 months since training that used measures approximated by regression models ((See Methods for details.)

Annex 1. Characteristics, summary measures, and sample size data for studies

A. Selected characteristics for all studies included in the review and details on summary measures for effect sizes based on ≥20 consultations per study group and time point

B. Sample size data for all studies included in the review

Annex 1A. Selected characteristics for all studies included in the review and details on summary measures for effect sizes based on ≥20 consultations per study group and time point

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
1	1. Arifeen et al. WHO Bulletin 2005;83:260–7; and personal communication from S. El Arifeen, June 13, 2007. 2. El Arifeen. Section 4: Health Facility Survey, 2000, pp. xxv–xxxiv. 3. El Arifeen et al. Lancet 2004;364:1595-602. 4. Arifeen et al. Lancet 2009; 374: 393–403.	Bangladesh	Pre-post with randomized controls	yes	yes	14	yes	Median of all outcomes: At 25 months: 14.6 At 50 months: 13.5 PTIG: 10.8	WHO IMCI PI-7[e]	At 25 months: 59.2 At 50 months: 75.7
									WHO IMCI SI-6	At 50 months: 76.8
									WHO IMCI SI-13	At 25 months: 26.7 At 50 months: 67.5
									WHO IMCI PI-8	At 25 months: 26.0 At 50 months: 33.6
									WHO IMCI PI-11	At 25 months: 1.6 At 50 months: 57.9
									MES	At 25 months: 26.4 At 50 months: 67.5
3	1. Amaral et al. Cadernos de saude publica 2004;20 Suppl. 2: S209–19. 2. Amaral. Manejo de casos da AIDPI nas unidades de saúde no Brasil -- 2002. 3. Gouws et al. WHO Bulletin 2004; 82: 509–15.	Brazil	Post-only with non-randomized controls	no	yes	6–8	yes	Median of all outcomes: 64.4 PTIG: 51.4	WHO IMCI PI-7[e]	15.3
									WHO IMCI PI-8	6.0
									WHO IMCI PI-11	–11.7
									WHO IMCI SI-13	32.9
									MES	10.7

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
4	1. Rathmony and Ministry of Health of Cambodia. Cambodia Health Facility Survey, Summary Data Tables. 2006. 2. Rathmony and Ministry of Health of Cambodia. Cambodia Health Facility Survey, PowerPoint presentation. 2006. 3. Rehlis N. Health Facility Survey in Cambodia. Principal Indicators. June 2007.	Cambodia	Post-only with non-randomized controls	no	yes	11	no	Median of all outcomes: 22.9 PTIG: 22.9	WHO IMCI PI-7[e]	43.8
									WHO IMCI PI-8	34.4
									WHO IMCI PI-11	3.7
									MES	34.4
4	1. Rathmony and Ministry of Health of Cambodia. Cambodia Health Facility Survey, Summary Data Tables. 2006. 2. Rathmony and Ministry of Health of Cambodia. Cambodia Health Facility Survey, PowerPoint presentation. 2006. 3. Rehlis N. Health Facility Survey in Cambodia. Principal Indicators. June 2007.	Cambodia	Post-only with non-randomized controls	no	yes	11	yes	Median of all outcomes: 22.9 PTIG: 22.9	WHO IMCI PI-7[e]	60.4
									WHO IMCI PI-8	33.7
									WHO IMCI PI-11	18.5
									MES	33.7
5	1. Anonymous. Implementación de la estrategia de Atención Integrada a las Enfermedades Prevalentes de la Infancia, Reporte Final, Linea Base, Provincia de Imbabura, Republica del Ecuador, 1997. 2. Ministry of Health of Ecuador. Evaluación de servicios de salud sobre "La Atención Integrada a las Enfermedades Prevalentes de la Infancia AIEPI", 2000.	Ecuador	Pre-post with no controls	no	yes	7	no	Median of all outcomes: 50.0	WHO IMCI PI-11	35.7

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
6	1. Choi et al. Assessment of child health services in Eritrea after three years of implementation of the Integrated Management of Childhood Illness (IMCI) strategy, 2002. 2. Mehari et al. Preparedness of Eritrean Health Services for the Integrated Management of Childhood Illness (IMCI) Strategy, 2000.	Eritrea	Pre-post with no controls	no	yes	11	yes	Median of all outcomes: 40.4 PTIG: 27.8	WHO IMCI PI-7[e]	3.9
									WHO IMCI PI-8	5.0
									WHO IMCI PI-11	–16.3
									WHO IMCI SI-6	–1.0
									WHO IMCI SI-8	7.0
									WHO IMCI SI-13	25.0
									MES	4.5
7	Salgado et al. Health Facility Survey of Child Health Services, Southern Nations, Nationalities and Peoples Region (SNNPR) Ethiopia, March 2002.	Ethiopia	Case-control	no	yes	11	yes	Median of all outcomes: 40.7	WHO IMCI PI-11	–1.4
									WHO IMCI SI-13	8.9
									MES	3.8
9	1. Centers for Disease Control and Prevention. MMWR 1998; 47(46): 998–1001. 2. Lee et al. unpublished manuscript	Kenya	Pre-post with no controls	no	yes	15	no	Median of all outcomes: 40.0	WHO IMCI SI-7	At 11.5 months: 33.0
									% of malaria patients correctly treated (by drug name only)	At <1 month: 19.0 At 2.5 montns: 23.0 At 11.5 months: 0
									% of pneumonia patients correctly treated (by drug name only)	At <1 month: 70.0 At 2.5 montns: 49.0 At 11.5 months: 46.0
									MES	At <1 month: 44.5 At 2.5 montns: 36.0 At 11.5 months: 33.0
10	Gilroy et al. Patient Educ Couns 2004; 54: 35–44.	Mali	Post-only with randomized controls	yes	no	11	no	Median of all outcomes: 30.2	% of patients for whom HW said how many days to give drug	20.9

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
11	1. Naimoli. The Integrated Management of Childhood Illness (IMCI) pilot project in Morocco, Preliminary results from an evaluation of the management of childhood illness in public sector IMCI and non-IMCI facilities in four Moroccan provinces, April 2000. 2. Naimoli JF. Theoretical and empirical advances in research on the implementation of an integrated approach to managing childhood illness in outpatient facilities in developing countries [ScD Dissertation]. [Cambridge (MA)]: Harvard University; 2001. 347 p. 3. Naimoli et al. Int J Qual Health Care 2006; 18: 134–44.	Morocco	Post-only with non-randomized controls	no	yes	12	no	Median of all outcomes: 46.8 PTIG: 32.6	% of patients needing antibiotic who were prescribed first or second line antibiotic (by drug name only)[e]	28.6
									WHO IMCI SI-13	43.6
									100% – % of patients with oral antibiotic prescribed without IMCI indication	20.2
									% of pneumonia patients with no pharyngitis who were prescribed first or second line antibiotic (by drug name only)	22.0
									MES	25.3
12	1. Kelley et al. Operations Research Results 2000; 1: 16. 2. Kelley et al. Int J Health Plan Management 2001; 16: 195–205. 3. Kelley et al. Operations Research Results 2002; 2: 15.	Niger	Pre-post with no controls	no	no	11	yes	Median of all outcomes: 35.5	% of all tasks done correctly per child	2.0
									% of 17 assessment and counseling tasks done correctly per child	27.1
									MES	14.6

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
13	1. Ministry of Health of Niger. Enquête dans les formations sanitaires mettant en œuvre la PCIME. 2005. 2. Degbey H. Enquête dans les formations sanitaires mettant en oeuvre la PCIME au Niger (15 mars - 15 avril 2005). Lettre de l'OMS Niger [Internet] 2005 [cited September 15, 2010]; 40: 6-7. Available from: http://www.santetropicale.com/alire/omsnig0605.pdf	Niger	Case-control	no	yes	6	no	Median of all outcomes: 25.5 PTIG: 21.0	WHO IMCI PI-7[e]	22.0
									WHO IMCI SI-6	–1.0
									WHO IMCI SI-8	16.0
									WHO IMCI PI-8	37.5
									WHO IMCI SI-13	35.7
									MES	22.0
14	1. Huicho et al. Health Policy Plan 2005; 20: 14–24. 2. Ministry of Health of Perú. Evaluación de Servicios de Salud. Segunda Prueba Mundial, Perú, 13–24 Octubre 1999.	Peru	Post-only with non-randomized controls	no	yes	7	no	Median of all outcomes: 21.0 PTIG: 21.0	WHO IMCI PI-7[e]	–10.0
									WHO IMCI PI-8	–2.0
									WHO IMCI PI-11	19.0
									WHO IMCI SI-8	–5.0
									WHO IMCI SI-13	31.0
									MES	–2.0
15	Chopra et al. Arch Dis Child 2005; 90: 397–401.	South Africa	Pre-post with no controls	no	no	11	no	Median of all outcomes: 71.0	WHO IMCI PI-8	22.0
									WHO IMCI PI-11	–4.0
									MES	9.0

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
16	1. Armstrong Schellenberg et al. Health Policy Plan 2004; 19: 1–10. 2. Armstrong Schellenberg et al. Lancet 2004; 364: 1583–94. 3. Bryce et al. Health Policy Plan 2005; 20 Suppl. 1: i69–i76. 4. Gouws et al. Bull World Health Organ 2004; 82: 509–15. 5. Mbuya et al. IMCI Implementation: A Report on Experiences in Morogoro and Rufiji Districts in Tanzania, Preliminary Report. 2003. 6. Mgalula. Activities and Methods, Substudy 3: Health Facility Survey. 2000; pp. 5-19.	Tanzania	Post-only with non-randomized controls	no	yes	11–16	yes	Median of all outcomes: 37.5 PTIG: 35.0	WHO IMCI PI-7[e]	38.0
									WHO IMCI PI-8	29.0
									WHO IMCI PI-11	16.0
									WHO IMCI SI-6	35.0
									WHO IMCI SI-8	63.0
									WHO IMCI SI-13	78.0
									MES	36.5
17	Tavrow et al. Quality Assurance Project, 2002.	Uganda	Post-only with randomized controls (short vs. standard IMCI)	yes	no	Short: 9 Standard: 11	no	N/A	% of HWs who prescribed correct drugs in correct dosages for all classifications noted[e]	At 0 months: 8.5 At 3.5 months: 8.1
									% of HWs who gave all counseling messages in and checked caretaker's understanding	At 0 months: –4.2 At 3.5 months: –3.9
									MES	At 0 months: 2.2 At 3.5 months: 2.1
18	1. Gouws et al. Bull World Health Organ 2004; 82: 509–15. 2. Pariyo et al. Health Policy Plan 2005; 20 Suppl. 1: i58–i68.	Uganda	Post-only with non-randomized controls	no	yes	11–14	yes	Median of all outcomes: 28.2 PTIG: 43.6	WHO IMCI PI-7[e]	13.9
									WHO IMCI PI-11	15.8
									WHO IMCI SI-6	14.3
									WHO IMCI SI-8	22.5
									WHO IMCI SI-13	29.3
									MES	15.8

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
19	Rehlis. The Ministry of Health of Uzbekistan and The World Bank Project, Final Report, Evaluation of The Project Health-1 (Loan 4396 - UZ). 2003.	Uzbekistan	Case-control	no	yes	11	yes	Median of all outcomes: 57.8	Average of: % of patients who are prescribed antibiotics who know how to give antibiotics and % of patients who are prescribed other drugs who know how to give other drugs	3.5
									WHO IMCI SI-13	0
									MES	1.8
20	Mwinga, et al. Research Study Report: Comparative study of Integrated Management of Childhood Illness (IMCI) case management skills of primary health care workers trained in eleven-day standard course and six-day abridged course for physicians. 2006.	Zambia	Post-only with randomized controls (short vs. standard IMCI)	yes	yes	Short: 6 Standard: 11	no	N/A	WHO IMCI PI-7[e]	–2.7
									WHO IMCI PI-8	7.0
									WHO IMCI PI-11	8.6
									WHO IMCI SI-6	9.8
									WHO IMCI SI-8	–1.4
									WHO IMCI SI-13	10.3
									MES	7.8
21	Syla. A comparison of standard training (11-days) with adapted 8 days in the Integrated Management of Childhood Illness (IMCI) in Kosova. 2003. (and personal communication from S. Syla, Kosovo WHO Country Office, Prishtina, November 14 and 15, 2007)	Kosovo	Post-only with non-randomized controls (short vs. standard IMCI)	no	yes	Short: 8 Standard: 11	no	N/A	% of patients with pneumonia, streptococcal sore throat, or acute ear infection who were prescribed antibiotics correctly[e]	–6.0
									WHO IMCI PI-11	–2.0
									WHO IMCI SI-13	2.0
									100% – % of patients who were prescribed unnecessary antibiotics	1.0
									MES	–0.5

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
22	Anonymous. Evaluación del proceso de implementacion de la estrategia de Atención Integrada a las Enfermedades Prevalentes en la Infancia - AIEPI - en ambitos de las Direcciones de Salud Callao y Lima Norte (Ventanilla y Puente Piedra), 2004.	Peru	Case-control	no	no	5–7	yes	Median of all outcomes: 85.5	WHO IMCI PI-11	7.8
									WHO IMCI SI-13	–2.5
									100% – % of patients who were prescribed unnecessary medicines	28.0
									MES	7.8
23	Ortiz, et al. Impacto de la aplicación de la estrategia AIEPI en el Uso de Antibióticos para el tratamiento de las Infecciones Respiratorias Agudas y Enfermedades Diarreicas, en Centros de Salud de Lambayeque - Perú, 1998.	Peru	Pre-post with no controls	no	no	7	no	Median of all outcomes: 81.0	WHO IMCI SI-6	0
									% of non-pneumonia patients not treated with antibiotics	15.0
									% of diarrhea patients without signs of dysentery not treated with antibiotics	8.0
									MES	8.0

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
24	1. Osterholt et al. Hum Resour Health 2009; 7: 77. 2. Rowe et al. Am J Public Health 2009; 99: 837–46 (plus personal communication from A. Rowe, CDC, November 10, 2009, August 9, 2010 and October 3, 2010).	Benin	Pre-post with non-randomized controls	yes	yes	11	no	Median of all outcomes: 19.9 PTIG: 21.3	% of patients who received IMCI-recommended treatment for all potentially life-threatening illnesses[e]	At 4 months (predicted): 23.5 At 15 months (predicted): 18.8 At 30 months (predicted): 11.4 Averaged over 3-year follow-up period: 19.1
									WHO IMCI SI-6	At 4 months (predicted): 31.6 At 15 months (predicted): 22.1 At 30 months (predicted): 5.9 Averaged over 3-year follow-up period: 18.5
									MES	At 4 months (predicted): 27.6 At 15 months (predicted): 20.5 At 30 months (predicted): 8.7 Averaged over 3-year follow-up period: 18.8

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
24	1. Osterholt et al. Hum Resour Health 2009; 7: 77. 2. Rowe et al. Am J Public Health 2009; 99: 837–46 (plus personal communication from A. Rowe, CDC, November 10, 2009, August 9, 2010 and October 3, 2010).	Benin	Pre-post with non-randomized controls	yes	yes	11	yes	Median of all outcomes: 11.2 PTIG: 15.8	% of patients who received IMCI-recommended treatment for all potentially life-threatening illnesses[e]	At 4 months (predicted): 44.0 At 15 months (predicted): 40.6 At 30 months (predicted): 35.0 Averaged over 3-year follow-up period: 46.4
									WHO IMCI SI-6	At 4 months (predicted): 38.7 At 15 months (predicted): 40.3 At 30 months (predicted): 38.9 Averaged over 3-year follow-up period: 38.4
									MES	At 4 months (predicted): 41.4 At 15 months (predicted): 40.5 At 30 months (predicted): 37.0 Averaged over 3-year follow-up period: 42.4
25	Atakouma Dzayisse et al. Arch Pediatr 2006; 13: 1552–3.	Togo	Post-only with non-randomized controls	no	no	11	no	Median of all outcomes: 31.4 PTIG: 6.0	WHO IMCI PI-7[e]	94.0
									WHO IMCI PI-11	34.6
									MES	64.3
26	1. Assimadi et al. Arch Pediatr 1999; 6: 1135–6. 2. Assimadi et al. Arch Pediatr 2003; 10: 158–9.	Togo	Pre-post with no controls	no	no	11	no	Median of all outcomes: 73.3	Baseline: % of children with fever correctly treated Follow-up: % of malaria patients correctly treated	19.4

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
27	Burnham. Evaluation of Integrated Management of Childhood Illness (IMCI) Performance in Urban Health Centers, Lusaka, Zambia, August 10–23, 1997. BASICS, 1997.	Zambia	Pre-post with no controls	no	no	11	yes	Median of all outcomes: 60.0	% of diarrhea cases who were given or recommended ORS or ORT	At 3 months: 34.1 At 9 months: 17.4
									average of: % of caretakers of patients prescribed ORS who know how to give ORS, % of caretakers of patients prescribed CQ who know how to give CQ, and % of caretakers of patients prescribed CTM who know how to give CTM	At 9 months: –30.7
									WHO IMCI SI-13	At 3 months: 44.2 At 9 months: 38.4
									MES	At 3 months: 39.2 At 9 months: 17.4
28	1. Anonymous. Informe Preliminar: Evaluación AIEPI a Servicios de Salud, Primera Prueba Mundial, Bolivia, 12–30 de Abril 1999. 2. Zamora et al. Rev Chil Pediatr, 2002; 73(2): 184–191. 3. Cordero D et al. BASICS II, 2004	Bolivia	Pre-post with no controls	no	Baseline: no Follow-up: yes	11	yes	Median of all outcomes: 76.5	100% – % of diarrhea patients without dehydration who received unnecessary antibiotics	10.8

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
29	Briggs et al. Senegal Assessment: Drug Management for Childhood Illness. Rational Pharmaceutical Management Plus Program, Management Sciences for Health. June 2002.	Senegal	Case-control	no	no	11	no	Median of all outcomes: 63.2	WHO IMCI PI-11	13.5
									WHO IMCI SI-13	15.9
									% of non-pneumonia patients not treated with antibiotics	50.1
									% of pneumonia patients given appropriate antibiotic (drug name only)	8.0
									% of diarrhea cases who were prescribed ORS	14.9
									% of diarrhea cases who were not prescribed antidiarrheals	9.4
									% of diarrhea patients without signs of dysentery not treated with antibiotics	–3.8
									% of malaria patients given appropriate antimalarial (drug name only)	26.0
									MES	14.2

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
30	1. Federal Ministry of Health of Sudan. Health Facility Survey on Quality of Outpatient Child Health Services, IMCI Health facility survey, Sudan, March-April 2003, 2004. 2. World Health Organization. Regional Office for the Eastern Mediterranean. Implementation of IMCI in Sudan. Systematic approach to IMCI implementation at district level: key steps and tools.	Sudan	Case-control	no	yes	11	yes	Median of all outcomes: 26.2 PTIG: 5.1	Average of: % of patients needing oral antibiotic prescribed antibiotic correctly, % of malaria patients prescribed recommended oral antimalarials correctly, and % of diarrhea patients prescribed ORS correctly[e]	24.6
									% of patients not needing antibiotics for an IMCI or non-IMCI reason who were prescribed no antibiotics	47.9
									% of diarrhea cases who were prescribed ORS	2.4
									Average of: % of caretakers of patients prescribed recommended oral antibiotics who know how to give antibiotics, % of caretakers of patients prescribed recommended oral antimalarials who know how to give antimalarials, and % of caretakers of patients prescribed ORS who know how to give ORS	22.7

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
30	(continued from previous page)	Sudan	Case-control	no	yes	11	yes	Median of all outcomes: 26.2 PTIG: 5.1	Average of: % of caretakers of patients prescribed recommended oral antibiotics who were given advice about dose, frequency, and duration of antibiotic treatment; % of caretakers of patients prescribed recommended oral antimalarials who were given advice about dose, frequency, and duration of antimalarial treatment; and % of caretakers of patients prescribed ORS who were given advice about dose, frequency, and duration of ORS treatment	35.1
									MES	24.6
31	Zhang et al. J Paediatr Child Health. 2007 Oct; 43(10): 681–5.	China	Pre-post with no controls	no	yes	5	yes	Median of all outcomes: 56.4	100% – % of patients with inappropriate use of injection or IV	40.9
									100% – % of patients with inappropriate use of antibiotics	53.1
									100% – % of patients with inappropriate use of hormones	4.9
									MES	40.9

Study ID	References	Country	Study design	Ade-quate study design[a]	All steps of the standard WHO health facility survey tool used[b]	Training duration (days)	Other interven-tions[c]	Baseline or control value used in models[d]	Outcomes	Effect size (percentage points)
32	1. Goga A. IMCI health facility survey: Eastern Cape and North West Provinces, 2003 (plus personal communication from A. Goga, Medical Research Council of South Africa, July 13, 2009). 2. World South Africa National Department of Health. Report of IMCI health facility survey in South Africa, 2005.	South Africa	Case-control	no	yes	11	no	See Annex 4A	See Annex 4A	See Annex 4A
33	Ministry of Health of Vietnam, National Institute of Hygiene and Epidemiology of Vietnam. Official Travel Report, Annex 1, 2002 (plus personal communication from T. Lambrechts, WHO, July 1, 2009 and December 9, 2009).	Vietnam	Case-control	no	yes	11	no	See Annex 4A	See Annex 4A	See Annex 4A

Footnotes for Annex 1A.
CQ = chloroquine; CTM = cotrimoxazole; HW = health worker; IMCI = Integrated Management of Childhood Illness; IV = intravenous therapy; MES = median effect size; N/A = not applicable; ORS = oral rehydration solution; ORT = oral rehydration therapy; PI = priority indicator; PTIG = percent of patients treated according to IMCI guidelines; SI = supplemental indicator; WHO = World Health Organization

For Study IDs 32 and 33, effect sizes were measured on <20 consultations per study group and time point. Thus, they were excluded from the primary analysis, but they were included in the alternative analysis that involved studies with effect sizes based on ≥15 consultations per study group and time point (see Annex 4A for results).

WHO IMCI PI-7 is the proportion of children who do not need urgent referral, who need an oral antibiotic or an antimalarial who are prescribed the drug(s) correctly (includes correct dosing of drugs). [WHO, 2003]

WHO IMCI PI-8 is the proportion of children not needing antibiotic who leaves facility without antibiotic. [WHO, 2003]

Footnotes for Annex 1A, continued:

WHO IMCI PI-11 is the proportion of children prescribed ORS, oral antibiotic, or oral antimalarial whose caretaker knows how to give treatment. [WHO, 2003]

WHO IMCI SI-6 is proportion of children with pneumonia correctly treated (includes correct dosing of drugs). [WHO, 2003]

WHO IMCI SI-7 is the proportion of children with dehydration correctly treated. [WHO, 2003]

WHO IMCI SI-8 is proportion of children with malaria correctly treated (includes correct dosing of drugs). [WHO, 2003]

WHO IMCI SI-13 is the proportion of children prescribed oral medication whose caretaker is advised on how to administer the treatment. [WHO, 2003]

[a] "First-tier" study designs include a pre-post study with either randomized or non-randomized controls and a post-only study with randomized controls.

[b] Steps of the standard WHO health facility survey tool include: silent observation of the consultation, caretaker interview, re-examination of the child to obtain a "gold standard" determination of the child's IMCI illness classifications, health worker interview, and a health facility assessment to collect information on the availability of equipment and drugs.

[c] "Other interventions" are strategies besides IMCI training that were implemented to support HWs' adherence to IMCI guidelines. (See Table 2 in report)

[d] The "models" refer to linear regression models of either MES or PTIG (see Tables 7a and 7b, 8a and 8b, and Annexes 3 and 4 in the report). Among pre-post studies with controls comparing IMCI to non-IMCI groups, the baseline measures of the IMCI group were used. Among studies comparing short and standard IMCI training, the "median baseline or control value" is not applicable.

[e] Outcome used for the summary measure PTIG.

Annex 1B. Sample size data for all studies included in the review

Country [Study identify-cation number[a]]	Sample size[b]		
	N_{HF}	N_{HW}	$N_{patient}$
Direct comparison studies with a first-tier design (2 comparisons[c] from 2 studies)			
Uganda [17]	NA	104[d]	NA
Zambia [20]	82	113	377
Direct comparison study with a second-tier design (1 comparison from 1 study)			
Kosovo [21]	30	56	351
Indirect comparison studies with a first-tier design (4 comparisons from 3 studies)			
Bangladesh [1]	20	586[f]	378[d]
Benin [24] comparison 1[c]	99	196	757[d]
Benin [24] comparison 2[c]	101	218	872[d]
Mali [10]	10	10	364
Indirect comparison studies with a second-tier design (24 comparisons from 23 studies)			
Bolivia [28]	36[d]	54[d]	102[d]
Brazil [3]	653	653	584
Cambodia [4] comparison 1[c]	80	NA	262[d]
Cambodia [4] comparison 2[c]	80	NA	216[d]
China [31]	419[d]	NA	696
Ecuador [5]	37[d]	47[d]	113
Eritrea [6]	74[d]	NA	360
Ethiopia [7]	43	43	102
Kenya [9]	36	NA	1043[d]
Morocco [11]	62	101	271
Niger [12]	NA	50[d]	NA
Niger [13]	44	NA	216
Peru [14]	90	202	372
Peru [22]	15	NA	58
Peru [23]	7	NA	670
Senegal [29]	41	NA	1217
South Africa [15]	21	42[d]	96
Sudan [30]	66	NA	254
Tanzania [16]	73	NA	404
Togo [25]	2	NA	300
Togo [26]	45[d]	221[d]	166

Annex 1B continued on next page.

Annex 1B, continued. Sample size data for all studies included in the review

Country [Study identify-cation number[a]]	Sample size[b]		
	N_{HF}	N_{HW}	$N_{patient}$
Uganda [18]	316 [d]	427 [d]	1265
Uzbekistan [19]	120	NA	170
Zambia [27]	8 [d]	NA	223 [d]
Indirect comparison studies with a second-tier design with <20 consultations per study group and time point (2 comparisons from 2 studies, used in one of the alternative analyses)			
South Africa [32]	32	NA	41
Vietnam [33]	70	NA	110

Footnotes to Annex 1B.

IMCI = Integrated Management of Childhood Illness, NA = not available, N_{HF} = health facility sample size, N_{HW} = health worker sample size, $N_{patient}$ = patient sample size

[a] Label that identifies a group of one or more reports with a single study (see reference list on page 31, which is organized by Study ID).

[b] In each study, N_{HF} and N_{HW} corresponded with the $N_{patient}$ whose outcomes were analyzed for our review, with some exceptions:

- N_{HF}, N_{HW}, and $N_{patient}$ were overall sample sizes for the study: Bolivia [28]
- N_{HF} and N_{HW} were overall sample sizes for the study: Bangladesh [1], Brazil [3], Ecuador [5], Ethiopia [7], Kosovo (21), Morocco [11], Peru [14], South Africa [15], Togo [26], Uganda [18], Zambia [20]
- N_{HF} and $N_{patient}$ were overall sample sizes for the study: Eritrea [6], Peru [22], Peru [23], Uzbekistan [19], Zambia [27]
- N_{HF} was the overall sample size of health facilities for the study: Cambodia [4], Kenya [9], Niger [13], Senegal [29], South Africa [32], Sudan [30], Vietnam [33], Tanzania [16]
- N_{HW} was the overall sample size of health workers for the study: Niger [12]

[c] See Box 1 for explanation of a comparison. Comparison 1 gives an effect size for IMCI training with other interventions, and comparison 2 gives an effect size for IMCI training without other interventions. Note that in Cambodia [4], the control group received new health systems support from contracting "Operational Districts". In Benin [24], the control group received nothing more than the pre-existing health system support.

[d] Details about sample sizes:

- Bangladesh [1]: N_{HW} is the sum of the number of health workers at baseline and the latest follow-up, which could be an over-estimate if some of the same health workers were observed at baseline and follow-up. $N_{patient}$ is the sum of the number of patients at baseline and the latest follow-up.

- Benin [24]: $N_{patient}$ for comparison 1 is the sum of the number of patients at baseline and all follow-up measures for the "IMCI with other interventions" group and the control group. $N_{patient}$ for comparison 2 is the sum of the number of patients at baseline and all follow-up measures for the "IMCI without other interventions" group and the control group.
- Bolivia [28]: N_{HF}, N_{HW}, and $N_{patient}$ are underestimates because they only include sample sizes at follow-up. Baseline sample sizes are not available.
- Cambodia [4]: $N_{patient}$ for comparison 1 is the sum of the number of patients in the "IMCI with other interventions" group and the control group. $N_{patient}$ for comparison 2 is the sum of the number of patients in the "IMCI without other interventions" group and the control group.
- China [31]: N_{HF} is the sum of the number of health facilities at baseline and follow-up, which could be an over-estimate if some of the same health facilities were observed at baseline and follow-up.
- Ecuador [5]: N_{HF} is the sum of the number of health facilities at baseline and follow-up, which could be an over-estimate if some of the same health facilities were observed at baseline and follow-up. N_{HW} is the sum of number of the health workers at baseline and follow-up, which could be an over-estimate if some of the same health workers were observed at baseline and follow-up.
- Eritrea [6]: N_{HF} is the sum of number of health facilities at baseline and follow-up, which could be an over-estimate if some of the same health facilities were observed at baseline and follow-up.
- Kenya [9]: $N_{patient}$ is the sum of the number of patients at baseline and at the latest follow-up.
- Niger [12]: The study involved 2 baseline and 1 post-IMCI measure. The N_{HF} at each measure was not available; however, the study covered 3 districts, each having 8–10 health facilities. N_{HW} is the sum of the number of health workers at the baseline measure closest to the time of IMCI training and that at follow-up, which could be an over-estimate if some of the same health workers were observed at baseline and follow-up. The $N_{patient}$ at each measure was not available; however, the study collected data on a total of 483 patients.
- South Africa [15]: N_{HW} is the sum of the number of the health workers at baseline and follow-up, which could be an over-estimate if some of the same health workers were observed at baseline and follow-up.
- Togo [26]: N_{HF} is an underestimate because it only includes the number of health facilities at baseline. The N_{HF} at follow-up was not available. N_{HW} is the sum of the number of health workers at baseline and follow-up, which could be an over-estimate if some of the same health workers were observed at baseline and follow-up.
- Uganda [17]: N_{HW} is the number of health workers at the latest follow-up.
- Uganda [18]: N_{HF} is the sum of the number of health facilities at all follow-ups, which could be an over-estimate if some of the same health facilities were observed at each follow-up. N_{HW} is the sum of the number of health workers at all follow-ups, which could be an over-estimate if some of the same health workers were observed at each follow-up.
- Zambia [27]: N_{HF} and $N_{patient}$ are underestimates because they only include sample sizes at the latest follow-up. Baseline sample sizes are not available.

Annex 2. Training quality among 10 studies that reported information about at least 1 of the quality criterion besides "completion of all training modules" (since all studies included in our review involved training that covered all IMCI modules).

Country (Study ID)	Trainee:facilitator ratio no more than 4:1	Completion of all modules	IMCI chart booklet to each trainee	Minimum 30% clinical practice	Minimum 20 sick children managed per trainee	No more than 24 trainees per training
Bangladesh (1)	N/A	Yes	Yes	N/A	N/A	Yes
Benin (24)	Yes	Yes	Yes	Yes	Yes	Yes
Kenya (9)	N/A	Yes	N/A	N/A ("daily practice" over 15-day training)	N/A	Yes
Kosovo (21)	11-day training: Yes 8-day training: No	Yes	N/A	Yes	N/A	N/A
Morocco (11)	Yes	Yes	Yes	Yes	N/A	Yes
Peru (14)	No	Yes	N/A	Yes	N/A	N/A
Peru (23)	N/A	Yes	N/A	Yes	N/A	N/A
Tanzania (16)	N/A	Yes	N/A	Yes	N/A	N/A
Uganda (18)	N/A ("low" ratio)	Yes	N/A	N/A ("less clinical instruction than 11-day course")	N/A	N/A
Zambia (20)	N/A	Yes	N/A	Yes	11-day training: Yes 6-day training: No	N/A

IMCI = Integrated Management of Childhood Illness, N/A = not available

Annex 3. Results of the alternative analysis #1 (exclude Togo [Study ID 25])

A. Sensitivity analysis with the median effect size (MES) summary measure: multivariable linear regression models (training duration analyzed as a dichotomous variable)

B. Sensitivity analysis with the median effect size (MES) summary measure: multivariable linear regression models (training duration analyzed as a continuous variable)

C. Sensitivity analysis with the PTIG summary measure: multivariable linear regression models (training duration analyzed as a dichotomous variable)

D. Sensitivity analysis with the PTIG summary measure: multivariable linear regression models (training duration analyzed as a continuous variable)

E. Summary of alternative analysis #1

Annex 3A. Results of a sensitivity analysis in the alternative analysis #1 (exclude Togo [Study ID 25]) with the median effect size (MES) summary measure: multivariable linear regression models[a] (training duration analyzed as a dichotomous variable)

Variable	Model 1[b]		Model 2		Model 3	
	Coeff.[c]	p-value	Coeff.[c]	p-value	Coeff.[c]	p-value
Intercept	40.47	0.09	30.14	0.09	28.92	**0.01**
Training approach (1 = standard, 0 = short)	–2.83	0.89	9.40	0.38	–1.23	0.86
Other interventions implemented (1 = yes, 0 = no)	8.11	0.22	9.14	0.15	2.44	0.67
First-tier study design (1 = yes, 0 = no)	–5.03	0.61	–5.68	0.56	12.71	0.18
Median baseline or control value	–0.32	0.19	–0.34	0.16	–0.23	0.15
Time since training (months)	–0.24	0.46	–0.05	0.78	—	—
"Time since training x training approach" interaction	0.27	0.48	—	—	—	—

Footnotes for Annex 3A.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness.

[a] Models 1 and 2 involved 24 effect sizes from 16 comparisons from 14 studies. Model 3 involved 27 effect sizes (one per comparison) from 25 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] As an example, the model was:
$MES = \beta_0 + (\beta_1 \times [\text{Training approach}]) + (\beta_2 \times [\text{Other interventions}]) + (\beta_3 \times [\text{First-tier study design}]) + (\beta_4 \times [\text{Median baseline or control value}]) + (\beta_5 \times [\text{Time since training}]) + (\beta_6 \times [\text{Time since training}] \times [\text{Training approach}]) + \varepsilon$,
where β_0 is the parameter estimate for the intercept, β_1 is the parameter estimate for the "Training approach" variable, β_2 is the parameter estimate for the "Other interventions" variable, and so on; and ε is the error term.

[c] Coefficients are percentage-point differences. For example, in row 2 of Model 3, the coefficient means that the effect size for standard training is 1.23 percentage-points lower than for short training, after controlling for all other factors in the model (although the p-value of 0.86 indicates that there is insufficient evidence to conclude that this difference is different from zero).

Annex 3B. Results of a sensitivity analysis in the alternative analysis #1 (exclude Togo [Study ID 25]) with the median effect size (MES) summary measure: multivariable linear regression models[a] (training duration analyzed as a continuous variable)

Variable	Model 4 [b]		Model 5		Model 6	
	Coeff.[c]	p-value	Coeff.[c]	p-value	Coeff.[c]	p-value
Intercept	10.12	0.73	6.49	0.76	19.88	0.22
Training duration (days)	2.22	0.29	2.52	0.06	0.71	0.55
Other interventions implemented (1 = yes, 0 = no)	10.05	0.13	10.52	0.08	1.97	0.73
First-tier study design (1 = yes, 0 = no)	–4.88	0.60	–4.48	0.61	11.87	0.20
Median baseline or control value	–0.30	0.19	–0.30	0.17	–0.20	0.20
Time since training (months)	–0.09	0.87	0.01	0.96	—	—
"Time since training x training duration" interaction	0.01	0.85	—	—	—	—

Footnotes for Annex 3B.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness.

[a] Models 4 and 5 involved 24 effect sizes from 16 comparisons from 14 studies. Model 6 involved 27 effect sizes (one per comparison) from 25 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] For an example of the model, see footnote "b" of Annex 3A.

[c] Coefficients are percentage-point differences. See footnote "c" of Annex 3A.

Annex 3C. Results of a sensitivity analysis in the alternative analysis #1 (exclude Togo [Study ID 25]) with the PTIG summary measure: multivariable linear regression models[a] (training duration analyzed as a dichotomous variable)

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8	
Variable	Coeff.[b]	P	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p
Intercept	–30.52	0.30	–1.75	0.94	16.59	0.47	21.86	0.31	6.36	0.64	22.69	0.28	6.57	0.63	9.10	0.47
Training approach (1=standard, 0=short)	81.63	**0.03**	40.02	**0.01**	16.18	0.36	18.70	0.27	18.23	0.27	23.13	0.14	23.55	0.15	26.34	0.08
Other interventions implemented (1=yes, 0=no)	23.24	**0.01**	22.49	**0.01**	9.93	0.49	10.39	0.46	8.21	0.54	—	—	7.60	0.57	—	—
First-tier study design (1=yes, 0=no)	–16.69	0.17	–10.82	0.35	13.21	0.45	—	—	17.00	0.28	—	—	—	—	—	—
Baseline or control value of PTIG	–0.89	0.08	–0.64	0.17	–0.35	0.56	–0.52	0.34	—	—	–0.44	0.40	—	—	—	—
Time since training (months)	0.91	0.06	0.36	0.11	—	—	—	—	—	—	—	—	—	—	—	—
"Time since training x training approach" interaction	–0.71	0.18	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Footnotes for Annex 3C.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness; p = p-value for test that coefficient is different from zero; PTIG = percent of patients treated according to IMCI guidelines.

[a] Models 1 and 2 involved 14 effect sizes from 9 comparisons from 7 studies. Models 3–8 involved 13 effect sizes (one per comparison) from 11 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] Coefficients are percentage-point differences. For example, in row 2 of Model 7, the coefficient means that the effect size for standard training is 23.55 percentage-points higher than for short training, after controlling for all other factors in the model (and the p-value of 0.15 indicates that there is insufficient evidence to conclude that this difference is different from zero).

Annex 3D. Results of a sensitivity analysis in the alternative analysis #1 (exclude Togo [Study ID 25]) with the PTIG summary measure: multivariable linear regression models[a] (training duration analyzed as a continuous variable)

	Model 9		Model 10		Model 11		Model 12		Model 13		Model 14		Model 15		Model 16	
Variable	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p
Intercept	–87.83	0.09	–47.51	0.10	–7.05	0.83	–6.34	0.84	–19.30	0.50	–9.76	0.75	–25.35	0.37	–26.50	0.32
Training duration (days)	11.38	**0.02**	7.49	**0.002**	3.79	0.22	4.29	0.14	3.92	0.20	4.92	0.06	4.90	0.10	5.26	**0.047**
Other interventions implemented (1=yes, 0=no)	17.86	**0.03**	17.91	**0.02**	7.24	0.61	7.16	0.60	5.55	0.68	—	—	4.40	0.74	—	—
First-tier study design (1=yes, 0=no)	–11.68	0.26	–10.74	0.29	10.52	0.54	—	—	15.58	0.31	—	—	—	—	—	—
Baseline or control value of PTIG	–0.59	0.16	–0.53	0.20	–0.42	0.46	–0.56	0.27	—	—	–0.51	0.29	—	—	—	—
Time since training (months)	1.03	0.19	0.29	0.12	—	—	—	—	—	—	—	—	—	—	—	—
"Time since training x training duration" interaction	–0.07	0.32	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Footnotes for Annex 3D.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness; p = p-value for test that coefficient is different from zero; PTIG = percent of patients treated according to IMCI guidelines.

[a] Models 9 and 10 involved 14 effect sizes from 9 comparisons from 7 studies. Models 11–16 involved 13 effect sizes (one per comparison) from 11 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] Coefficients are percentage-point differences. See footnote "b" of Annex 3C.

Annex 3E. Summary of alternative analysis #1 (exclude Togo [Study ID 25])

Study design adequacy and summary measure	Analysis[a]		
	Direct comparison (standard compared to short training)	Indirect comparison (standard training vs. no IMCI compared to short training vs. no IMCI)	Training duration analyzed as a continuous variable
Main analysis (first-tier study designs only)			
Outcome = MES Note: Few studies (≤6 effect sizes[b] from 5 studies)	Standard was slightly better (5 %-point difference, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Suggested increasing effect size with longer training duration (only 4 effect sizes; no statistical testing)
Outcome = PTIG Note: Very few studies (≤5 effect sizes[b] from 4 studies)	Standard and short were very similar (3 %-point difference, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Suggested increasing effect size with longer training duration (only 3 effect sizes; no statistical testing)
Sensitivity analysis (first- and second-tier study designs)			
Outcome = MES ≤30 effect sizes[b] from 28 studies	Standard and short were very similar (2 %-point difference, no statistical testing)	Standard was somewhat better than short (8 %-point difference, p=1.0)	Trend of increasing effect with longer training duration (univariate p=0.15 [1.7 %-point increase per extra day of training]; multivariable p-values ranged from 0.06–0.55)
Outcome = PTIG ≤16 effect sizes[b] from 14 studies	Standard and short were very similar (–3 %-point difference, no statistical testing)	Standard was much better than short (18 %-point difference; p=0.19)	Trend of increasing effect with longer training duration (univariate p=0.047 [5.3 %-point increase per extra day of training; multivariable p-values ranged from 0.06–0.10)[c]

Footnotes for Annex 3E.
IMCI = Integrated Management of Childhood Illness; MES = median effect size; PTIG = patient (needing oral antimicrobial or rehydration therapy) treated according to IMCI guidelines; %-point = percentage-point.

[a] All differences are all in terms of "effect of standard training" minus "effect of short training" (i.e., positive differences indicate standard training led to greater improvements in healthcare quality than short training).

[b] Excludes the effect sizes from Study ID 24 for specific months since training that used measures approximated by regression models. (See Methods for details.)

[c] Multivariable p-values were from models with 2 independent variables, since models with >2 independent variables might be unreliable, given the small sample size of effect sizes.

Annex 4. Results of the alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point)

A. Summary measures in alternative analysis #2 that were different from those in the primary analysis

B. Sensitivity analysis with the median effect size (MES) summary measure: multivariable linear regression models (training duration analyzed as a dichotomous variable)

C. Sensitivity analysis with the median effect size (MES) summary measure: multivariable linear regression models (training duration analyzed as a continuous variable)

D. Sensitivity analysis with the PTIG summary measure: multivariable linear regression models (training duration analyzed as a dichotomous variable)

E. Sensitivity analysis with the PTIG summary measure: multivariable linear regression models (training duration analyzed as a continuous variable)

F. Summary of alternative analysis #2

Annex 4A. Summary measures in alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point) that were different from those in the primary analysis (Annex 1A)

Study ID	References	Country	Baseline or control value used in models[a]	Outcomes	Effect size (percentage points)
1	1. Arifeen et al. WHO Bulletin 2005;83:260–7; and personal communication from S. El Arifeen, June 13, 2007. 2. El Arifeen. Section 4: Health Facility Survey, pp. xxv-xxxiv. 3. El Arifeen et al. Lancet 2004;364:1595-602. 4. Arifeen et al. Lancet 2009; 374: 393–403.	Bangladesh	Median of all outcomes: At 25 months: 13.5	WHO IMCI PI-7[b]	At 25 months: 59.2
				WHO IMCI SI-6	***At 25 months: 83.8***
				WHO IMCI SI-13	At 25 months: 26.7
				WHO IMCI PI-8	At 25 months: 26.0
				WHO IMCI PI-11	At 25 months: 1.6
				MES	At 25 months: 26.7
7	Salgado et al. Health Facility Survey of Child Health Services, Southern Nations, Nationalities and Peoples Region (SNNPR) Ethiopia, March 2002.	Ethiopia	Median of all outcomes: 43.0 PTIG: 43.0	***WHO IMCI PI-7[b]***	***7.0***
				WHO IMCI PI-11	–1.4
				WHO IMCI SI-13	8.9
				MES	7.0
15	Chopra et al. Arch Dis Child 2005; 90: 397–401.	South Africa	Median of all outcomes: 78.0	WHO IMCI PI-8	22.0
				WHO IMCI PI-11	–4.0
				WHO IMCI SI-6	***2.0***
				MES	2.0
21	Syla. A comparison of standard training (11-days) with adapted 8 days in the Integrated Management of Childhood Illness (IMCI) in Kosova. 2003. (and personal communication from S. Syla, Kosovo WHO Country Office, Prishtina, November 14 and 15, 2007)	Kosovo	N/A	% of patients with pneumonia, streptococcal sore throat, or acute ear infection who were prescribed antibiotics correctly[b]	–6.0
				WHO IMCI PI-11	–2.0
				WHO IMCI SI-13	2.0
				100% – % of patients who were prescribed unnecessary antibiotics	1.0
				% of patients who needed ORS who were given ORS	***–15.0***
				MES	–2.0

Annex 4A, continued. Summary measures in alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point) that were different from those in the primary analysis (Annex 1A)

Study ID	References	Country	Baseline or control value used in models[a]	Outcomes	Effect size (percentage points)
32	1. Goga A. IMCI health facility survey: Eastern Cape and North West Provinces, 2003 (plus personal communication from A. Goga, Medical Research Council of South Africa, July 13, 2009). 2. World South Africa National Department of Health. Report of IMCI health facility survey in South Africa, 2005.	South Africa	Median of all outcomes: 6.8 PTIG: 13.6	***WHO IMCI PI-7[b]***	***62.9***
				WHO IMCI SI-13	***47.1***
				MES	55.0
33	Ministry of Health of Vietnam, National Institute of Hygiene and Epidemiology of Vietnam. Official Travel Report, Annex 1, 2002 (plus personal communication from T. Lambrechts, WHO, July 1, 2009 and December 9, 2009).	Vietnam	Median of all outcomes: 31.7	***WHO IMCI PI-11***	***13.6***
				WHO IMCI SI-13	***50.3***
				MES	32.0

Footnotes for Annex 4A.

IMCI = Integrated Management of Childhood Illness; MES = median effect size; ORS = oral rehydration solution; PI = priority indicator; PTIG = percent of patients treated according to IMCI guidelines; SI = supplemental indicator; WHO = World Health Organization

Outcomes and effect sizes in bold italics were measured on 15–19 consultations per study group and time point. Thus, they were excluded from the primary analysis, but they were included in the alternative analysis #2.

WHO IMCI PI-7 is the proportion of children who do not need urgent referral, who need an oral antibiotic or an antimalarial who are prescribed the drug(s) correctly (includes correct dosing of drugs). [WHO, 2003]

WHO IMCI PI-8 is the proportion of children not needing antibiotic who leaves facility without antibiotic. [WHO, 2003]

WHO IMCI PI-11 is the proportion of children prescribed ORS, oral antibiotic, or oral antimalarial whose caretaker knows how to give treatment. [WHO, 2003]

WHO IMCI SI-6 is proportion of children with pneumonia correctly treated (includes correct dosing of drugs). [WHO, 2003]

Footnotes for Annex 4A, continued:

WHO IMCI SI-13 is the proportion of children prescribed oral medication whose caretaker is advised on how to administer the treatment. [WHO, 2003]

[a] The "models" refer to linear regression models of either MES or PTIG (see Tables 7a and 7b, 8a and 8b, and Annexes 3 and 4 in the report). Among pre-post studies with controls comparing IMCI to non-IMCI groups, the baseline measures of the IMCI group were used. Among studies comparing short and standard IMCI training, the "median baseline or control value" is not applicable.

[b] Outcome used for the summary measure PTIG.

Annex 4B. Results of a sensitivity analysis in the alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point) with the median effect size (MES) summary measure: multivariable linear regression models[a] (training duration analyzed as a dichotomous variable)

Variable	Model 1[b]		Model 2		Model 3	
	Coeff.[c]	p-value	Coeff.[c]	p-value	Coeff.[c]	p-value
Intercept	47.94	0.07	35.66	0.07	39.94	**0.001**
Training approach (1 = standard, 0 = short)	–1.94	0.93	12.65	0.28	1.72	0.82
Other interventions implemented (1 = yes, 0 = no)	6.21	0.38	7.33	0.29	–2.38	0.69
First-tier study design (1 = yes, 0 = no)	–10.76	0.31	–11.60	0.27	4.21	0.67
Median baseline or control value	–0.44	0.09	–0.46	0.07	–0.39	**0.02**
Time since training (months)	–0.27	0.45	–0.04	0.83	—	—
"Time since training x training approach" interaction	0.32	0.45	—	—	—	—

Footnotes for Annex 4B.
Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness.

[a] Models 1 and 2 involved 25 effect sizes from 17 comparisons from 15 studies. Model 3 involved 30 effect sizes (one per comparison) from 28 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] As an example, the model was:
$MES = \beta_0 + (\beta_1 \times [\text{Training approach}]) + (\beta_2 \times [\text{Other interventions}]) + (\beta_3 \times [\text{First-tier study design}]) + (\beta_4 \times [\text{Median baseline or control value}]) + (\beta_5 \times [\text{Time since training}]) + (\beta_6 \times [\text{Time since training}] \times [\text{Training approach}]) + \varepsilon$,
where β_0 is the parameter estimate for the intercept, β_1 is the parameter estimate for the "Training approach" variable, β_2 is the parameter estimate for the "Other interventions" variable, and so on; and ε is the error term.

[c] Coefficients are percentage-point differences. For example, in row 2 of Model 3, the coefficient means that the effect size for standard training is 1.72 percentage-points higher than for short training, after controlling for all other factors in the model (although the p-value of 0.82 indicates that there is insufficient evidence to conclude that this difference is different from zero).

Annex 4C. Results of a sensitivity analysis in the alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point) with the median effect size (MES) summary measure: multivariable linear regression models[a] (training duration analyzed as a continuous variable)

Variable	Model 4 [b]		Model 5		Model 6	
	Coeff.[c]	p-value	Coeff.[c]	p-value	Coeff.[c]	p-value
Intercept	30.28	0.35	16.83	0.48	32.21	0.07
Training duration (days)	1.33	0.57	2.46	0.11	0.82	0.54
Other interventions implemented (1 = yes, 0 = no)	6.91	0.34	8.54	0.21	–2.52	0.67
First-tier study design (1 = yes, 0 = no)	–11.58	0.27	–10.28	0.31	3.79	0.69
Median baseline or control value	–0.43	0.09	–0.44	0.08	–0.37	**0.02**
Time since training (months)	–0.40	0.52	–0.02	0.92	—	—
"Time since training x training duration" interaction	0.04	0.52	—	—	—	—

Footnotes for Annex 4C.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness.

[a] Models 4 and 5 involved 25 effect sizes from 17 comparisons from 15 studies. Model 6 involved 30 effect sizes (one per comparison) from 28 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] For an example of the model, see footnote "b" of Annex 4B.

[c] Coefficients are percentage-point differences. See footnote "c" of Annex 4B.

Annex 4D. Results of a sensitivity analysis in the alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point) with the PTIG summary measure: multivariable linear regression models[a] (training duration analyzed as a dichotomous variable)

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8	
Variable	Coeff.[b]	P	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p
Intercept	–20.95	0.58	29.89	0.27	40.28	0.09	40.34	0.07	13.05	0.44	40.48	0.06	12.92	0.43	9.10	0.56
Training approach (1=standard, 0=short)	104.77	**0.02**	41.42	**0.02**	23.06	0.22	23.08	0.20	31.77	0.11	22.68	0.16	34.00	0.07	30.77	0.09
Other interventions implemented (1=yes, 0=no)	23.57	**0.03**	22.41	**0.045**	–1.03	0.94	–0.99	0.94	–11.86	0.42	—	—	–11.45	0.42	—	—
First-tier study design (1=yes, 0=no)	–34.07	**0.02**	–29.00	**0.04**	0.23	0.99	—	—	10.15	0.58	—	—	—	—	—	—
Baseline or control value of PTIG	–1.56	**0.01**	–1.33	0.02	–0.99	0.11	–0.99	0.08	—	—	–1.01	0.045	—	—	—	—
Time since training (months)	1.08	0.08	0.20	0.46	—	—	—	—	—	—	—	—	—	—	—	—
"Time since training x training approach" interaction	–1.08	0.11	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Footnotes for Annex 4D.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness; p = p-value for test that coefficient is different from zero; PTIG = percent of patients treated according to IMCI guidelines.

[a] Models 1 and 2 involved 15 effect sizes from 10 comparisons from 8 studies. Models 3–8 involved 16 effect sizes (one per comparison) from 14 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] Coefficients are percentage-point differences. For example, in row 2 of Model 7, the coefficient means that the effect size for standard training is 34.0 percentage-points higher than for short training, after controlling for all other factors in the model (and the p-value of 0.07 indicates that there is insufficient evidence to conclude that this difference is different from zero).

Annex 4E. Results of a sensitivity analysis in the alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point) with the PTIG summary measure: multivariable linear regression models[a] (training duration analyzed as a continuous variable)

	Model 9		Model 10		Model 11		Model 12		Model 13		Model 14		Model 15		Model 16	
Variable	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p	Coeff.[b]	p
Intercept	–52.59	0.49	–8.68	0.82	6.22	0.86	6.31	0.86	–26.67	0.46	9.53	0.77	–29.60	0.39	–25.29	0.46
Training duration (days)	11.51	0.11	7.30	**0.02**	5.29	0.12	5.17	0.11	6.37	0.09	4.78	0.09	6.78	0.06	5.56	0.09
Other interventions implemented (1=yes, 0=no)	18.17	0.12	18.22	0.10	–4.06	0.78	–4.41	0.75	–15.52	0.30	—	—	–15.51	0.29	—	—
First-tier study design (1=yes, 0=no)	–31.30	**0.04**	–30.43	**0.04**	–3.20	0.86	—	—	7.66	0.68	—	—	—	—	—	—
Baseline or control value of PTIG	–1.37	**0.03**	–1.31	**0.03**	–1.05	0.08	–1.01	0.06	—	—	–1.08	**0.03**	—	—	—	—
Time since training (months)	0.91	0.45	0.11	0.68	—	—	—	—	—	—	—	—	—	—	—	—
"Time since training x training duration" interaction	–0.08	0.50	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Footnotes for Annex 4E.

Coeff. = Coefficient (i.e., parameter estimate from model); IMCI = Integrated Management of Childhood Illness; p = p-value for test that coefficient is different from zero; PTIG = percent of patients treated according to IMCI guidelines.

[a] Models 9 and 10 involved 15 effect sizes from 10 comparisons from 8 studies. Models 11–16 involved 16 effect sizes (one per comparison) from 14 studies. All studies compared either standard versus no IMCI or short versus no IMCI. Dashes indicate that a variable was not included in the model.

[b] Coefficients are percentage-point differences. See footnote "b" of Annex 4D.

Annex 4F. Summary of alternative analysis #2 (include studies that have effect sizes based on ≥15 consultations per study group and time point)

Study design adequacy and summary measure	Analysis[a]		
	Direct comparison (standard compared to short training)	Indirect comparison (standard training vs. no IMCI compared to short training vs. no IMCI)	Training duration analyzed as a continuous variable
Main analysis (first-tier study designs only)			
Outcome = MES Note: Few studies (≤6 effect sizes[b] from 5 studies)	Standard was slightly better (5 %-point difference, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Suggested increasing effect size with longer training duration (only 4 effect sizes; no statistical testing)
Outcome = PTIG Note: Very few studies (≤5 effect sizes[b] from 4 studies)	Standard and short were very similar (3 %-point difference, no statistical testing)	No comparison possible (no studies of short vs. no IMCI)	Suggested increasing effect size with longer training duration (only 3 effect sizes; no statistical testing)
Sensitivity analysis (first- and second-tier study designs)			
Outcome = MES ≤33 effect sizes[b] from 31 studies	Standard and short were very similar (2 %-point difference, no statistical testing)	Standard was somewhat better than short (10 %-point difference, p=1.0)	Trend of increasing effect with longer training duration (univariate p=0.14 [2.0 %-point increase per extra day of training]; multivariable p-values ranged from 0.11–0.54)
Outcome = PTIG ≤19 effect sizes[b] from 17 studies	Standard and short were very similar (–3 %-point difference, no statistical testing)	Standard was much better than short (23 %-point difference; p=0.20)	Trend of increasing effect with longer training duration (univariate p=0.09 [5.6 %-point increase per extra day of training; multivariable p-values ranged from 0.06–0.09)[c]

Footnotes for Annex 4F.

IMCI = Integrated Management of Childhood Illness; MES = median effect size; PTIG = patient (needing oral antimicrobial or rehydration therapy) treated according to IMCI guidelines; %-point = percentage-point.

[a] All differences are all in terms of "effect of standard training" minus "effect of short training" (i.e., positive differences indicate standard training led to greater improvements in healthcare quality than short training).

[b] Excludes the effect sizes from Study ID 24 for specific months since training that used measures approximated by regression models. (See Methods for details.)

[c] Multivariable p-values were from models with 2 independent variables, since models with >2 independent variables might be unreliable, given the small number of effect sizes.

ISBN 978 92 4 159721 0