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

Alexander K Rowe,1* Samantha Y Rowe,2 Kathleen A Holloway,3 Verica Ivanovska,3 Lulu Muhe4 and Thierry Lambrechts4

1Malaria Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA, 2Data Management Activity, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA, 3World Health Organization—formerly Headquarters, Geneva, Switzerland, now Regional Office Southeast Asia, Delhi, India and 4Department of Child and Adolescent Health and Development, World Health Organization, Geneva, Switzerland

*Corresponding author. Centers for Disease Control and Prevention, Mailstop F22, 4770 Buford Highway, Atlanta, GA 30341-3724, USA. Tel: +1–770–488 3588. Fax: +1–770–488 7761. E-mail: axr9@cdc.gov

Accepted 17 February 2011

Objective
Implementation of the Integrated Management of Childhood Illness (IMCI) strategy with an 11-day training course for health workers improves care for ill children in outpatient settings in developing countries. The 11-day course duration is recommended by the World Health Organization, which developed IMCI. Our aim was to determine if shortening the training (to reduce cost) reduces its effectiveness.

Methods
We conducted a systematic review to compare IMCI’s effectiveness with standard training (duration ≥11 days) versus shortened training (5–10 days). Studies were identified from a search of MEDLINE, two existing systematic reviews, and by contacting investigators. We included published or unpublished studies that evaluated IMCI’s effectiveness in developing countries and reported quantitative measures of health worker practices related to managing ill children under 5 years old in public or private health facilities. Summary measures were the median of effect sizes for all outcomes from a given study, and the percentage of patients needing oral antimicrobials or rehydration who were treated according to IMCI guidelines.

Findings
Twenty-nine studies were included. Direct comparisons from three studies showed little difference between standard and shortened training. Indirect comparisons from 26 studies revealed that effect sizes for standard training versus no IMCI were greater than shortened training versus no IMCI. Across all comparisons, differences ranged from -3 to +23 percentage-points, and our best estimate was a 2 to 16 percentage-point advantage for standard training. No result was statistically significant. After IMCI training (of any duration), 34% of ill children needing oral antimicrobials or rehydration were not receiving these treatments according to IMCI guidelines.

Conclusions
Based on limited evidence, standard IMCI training seemed more effective than shortened training, although the difference might be small. As sizable performance gaps often existed after IMCI training, countries should consider...
implementing other interventions to support health workers after training, regardless of training duration.

Keywords  Child health, developing country, quality improvement, systematic review, training

KEY MESSAGES

- Based on limited evidence, standard training (typically 11 days) for health workers on Integrated Management of Childhood Illness (IMCI) guidelines for managing ill children seemed more effective than short training (5–10 days), although the difference might be small.
- Where possible, standard IMCI training is recommended; however, in settings where only shortened training is feasible, core competencies should be retained (i.e. those addressing major causes of deaths for which effective interventions exist).
- Even after IMCI training (of any duration), typically one-third of ill children needing oral antimicrobials or rehydration were not receiving these treatments according to IMCI guidelines; thus, it is critical in all circumstances to implement strategies in addition to training to improve health worker adherence to guidelines.

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 health care quality at health facilities (Armstrong Schellenberg et al. 2004a; Amaral and Victoria 2008; Arifeen et al. 2009; Rove et al. 2009a), although its effect on mortality is uncertain (Armstrong Schellenberg et al. 2004b; Arifeen et al. 2009; Rove et al. in press).

To improve health care quality at outpatient health facilities, IMCI includes evidence-based guidelines (Gove et al. 1997; WHO 2005) for managing the leading causes of child deaths (pneumonia, diarrhoea and malaria) (Black et al. 2010). WHO recommends implementing the guidelines with an 11-day in-service training course for health workers, a follow-up visit to health workers’ facilities one month later to reinforce new practices, and job-aids (e.g. a chart booklet and wall chart of clinical algorithms, and a one-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 of ≤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 ≤24 participants (Lambrechts et al. 1999).

Despite favourable 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. 2009), concerns have been raised that the 11-day training is too expensive and that it takes health workers away from their clinics for too long (WHO 2003a; Goga et al. 2009). In response, many countries have shortened courses. 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 re-examination of 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 training course (or slightly longer courses) versus shortened training. Secondary objectives were to examine: (1) the effect of other interventions (besides IMCI training) to strengthen health systems and health worker adherence to IMCI guidelines; (2) the effect of IMCI over time; (3) the overall effect of IMCI training; and (4) the absolute level of health care quality delivered to ill children after IMCI training.

Methods

In preparing this review, 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 and lasted ≥5 days. Courses <5 days were considered too short to be of practical value, and the shortest courses that countries typically offer have a duration of 5 days (Goga et al. 2009). Standard and short training were defined as IMCI courses lasting ≥11 days and 5–10 days, respectively. The adequacy of study designs and analysis-related terms are
defined in Box 1. Study identification numbers (Study IDs) are labels representing all reports for a given study (see Table 2).

Sources, search strategies and inclusion criteria

We searched five sources to identify relevant reports (Table 1). 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 <5 years old in public or private facilities; and (3) were conducted in low- or middle-income countries (World Bank 2005). No study was excluded because of adequacy of statistical analysis or data collection method. Restrictions on the timing of studies and language of reports depended on the source (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.

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 IMCI follow-up visits were excluded, as these reflected practices when IMCI trainers were present and actively trying to improve adherence. Post-only studies (i.e. performance only measured after IMCI implementation) without controls were excluded, as effect sizes cannot be estimated with this design.

Data collection methods

Data on study outcomes and attributes of IMCI courses were collected using slightly different methods. Regarding study outcomes, we focused on direct measures of health worker behaviour (e.g. tasks related to treatment or counseling) and patient knowledge on administering treatments at home. Health outcomes (e.g. mortality) 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

---

**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.** 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 analyse results from studies that use different outcomes. The following two summary measures were used in this review.
  - Median effect size (MES) is the median of effect sizes for all outcomes for a given comparison from a given study. For example, Study ID 20 had seven effect sizes (-17.0, -2.7, -1.4, 7.0, 8.6, 9.8 and 10.3 percentage-points), and so the MES is the median of these seven values (7.0 percentage-points).
  - Patient treated according to IMCI guidelines (PTIG) is the percentage of patients needing an oral antimicrobial (antibiotic or antimalarial) or oral rehydration solution who received these treatments according to IMCI guidelines. In addition to selecting appropriate medicines, treatment almost always involved correct dosing and treatment duration.
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 co-ordinated 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. In a few cases, discrepancies were identified, and these 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. For 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 outcomes from study reports; and results were added to the WHO/INRUD database. Outcome definitions varied by study; however, most studies used WHO standard IMCI indicators (WHO 2003b). Details on the outcomes are available in Box 2 and the online annex.

Reports were also reviewed to determine: training quality; training duration; types of health workers trained; proportion of children managed by IMCI-trained health workers in geographic areas where IMCI was implemented; time between IMCI training and evaluation; sample sizes of health facilities, health workers and patients; and additional interventions (besides IMCI training) to strengthen IMCI implementation (e.g. extra supervision). 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. For studies on short training, we collected information on how the training was shortened. A data manager abstracted the information from reports and entered it into a database (Microsoft Excel, Microsoft, Inc., Redmond, WA, USA), and an investigator (SYR) reviewed the abstraction for accuracy. If needed details were not in a report, authors were contacted.

Analysis

We 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).

Ideally, we needed direct comparisons of health workers who received short training versus standard training. Only five such studies were identified (Surjono et al. 1998; Tavrow et al. 2002; Syla 2003; Mwinga et al. 2006; Quality Assurance Project 2006), two of which were excluded [one lacked quantitative measures of our outcomes (Surjono et al. 1998); and one used simulated, not real, patients (Quality Assurance Project 2006)]. However, as other studies compared IMCI-trained health workers with non-IMCI-trained health workers, we also performed an indirect comparison by contrasting effects of ‘standard training versus no IMCI’ in one group of studies and ‘short training versus no IMCI’ in a different group of studies.

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

<table>
<thead>
<tr>
<th>Source (date of search)</th>
<th>Search strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. WHO/CAH literature search (January 2007)</td>
<td>Searched OVID “MEDLINE R In process” and other non-indexed citations for reports with key word “IMCI” (personal communication, A Goga, 10 May 2007)</td>
</tr>
<tr>
<td>2. WHO/CAH reports (May 2007)</td>
<td>Searched reports from CAH and WHO regional offices (mainly unpublished reports) (person communication, T Lambrechts, May–June 2007)</td>
</tr>
<tr>
<td>5. Investigators of IMCI evaluations</td>
<td>Unpublished reports from Bangladesh (personal communication, S Arifeen, 14 August 2007), Benin (personal communication, A Rowe, 10 November 2006), South Africa (personal communication, A Goga, 13 July 2009), and Vietnam (personal communication, T Lambrechts, 1 July and 9 December 2009)</td>
</tr>
</tbody>
</table>

Notes:

*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 included if only an abstract was available). See WHO (2009) for detailed search strategy.

†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 1991–2006 (as found in searches conducted in May 2006) were included; there were no language restrictions.
For both direct and indirect comparisons, to balance strength of study design, risk of bias, representativeness among developing countries and influence of outliers (Figure 1, vertical dimension), we performed a main analysis of only studies with first-tier designs (Box 1) and a sensitivity analysis of studies with first- or second-tier designs. As outcomes of health worker performance varied among studies, we needed a common metric that could be used across as many studies as possible. Thus, for the two analyses, we used two summary measures (Figure 1, horizontal dimension): the median effect size (MES) and percentage of patients needing oral antimicrobials or oral rehydration solution who received these treatments according to IMCI guidelines (‘patient treated according to IMCI guidelines’, or PTIG) (Box 1). MES reflected a study’s ‘middle’ effect

Box 2 The six most commonly used outcomes (all outcomes available in the online annex)

- Proportion of children prescribed oral rehydration solution, oral antibiotic or oral antimalarial whose caretaker knows how to give the treatment [Integrated Management of Childhood Illness (IMCI) Priority Indicator 11 (WHO 2003b)]
- Proportion of children prescribed oral medication whose caretaker is advised on how to administer the treatment [IMCI Supplemental Indicator 13 (WHO 2003b)]
- Proportion of children who do not need urgent referral, who need an oral antibiotic or an antimalarial who are prescribed the drug(s) correctly, including correct dosing of drugs [IMCI Priority Indicator 7 (WHO 2003b)]
- Proportion of children with pneumonia correctly treated, including correct dosing of drugs [IMCI Supplemental Indicator 6 (WHO 2003b)]
- Proportion of children not needing antibiotic who leave facility without an antibiotic [IMCI Priority Indicator 8 (WHO 2003b)]
- Proportion of children with malaria correctly treated, including correct dosing of drugs ([IMCI Supplemental Indicator 8 (WHO 2003b)]

Figure 1 Four analytic approaches, contrasted by study design adequacy and outcome comparability Notes: MES = Median effect size for all outcomes of health worker performance; PTIG = the percentage of patients treated according to Integrated Management of Childhood Illness guidelines (see Methods). 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.
size, was available for all studies and has been used in other reviews (Jamtvedt et al. 2006; Rowe et al. 2009b). PTIG, although not available for all studies, had clear clinical and public health relevance and better comparability than MES. 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 (-%) ‘difference of differences’ (equation 1) (Ross-Degnan et al. 1997; WHO 2001).

\[
\text{Effect size} = \frac{(\text{follow-up} - \text{baseline})_{\text{intervention}} - (\text{follow-up} - \text{baseline})_{\text{control}}}{\text{baseline}}
\]  
(1)

Effect sizes were calculated such that values >0 indicated an improvement in case-management quality. For follow-up measurements, we had to consider the possibility that health worker performance might change (e.g. deteriorate) after IMCI training. However, time between training and evaluation (‘time since training’) varied among studies. For most analyses, we used the one time point from each study that was furthest from IMCI training. In additional analyses, we examined 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, NC, USA). An exact, nonparametric test was used to test differences between medians, and linear regression modelling with the REG procedure was used to test for trends and in multivariable modelling. As regression models based on extremely small samples might be unreliable, we conducted modelling only when analyses involved ≥10 effect sizes. Hypothesis testing was done with an alpha level of 0.05 and assumed each effect size had no uncertainty. Ideally, statistical tests should account for the uncertainty 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. However, in this review, the correlations (design effects or intraclass correlation coefficients) were not known for most studies. We considered \( P \)-values from 0.05 to 0.10 to be borderline significant.

To identify and control for confounding, we stratified effect sizes by time since training, and whether other interventions besides IMCI training were implemented. We also performed statistical modelling with the above factors, plus baseline values and a continuous variable for training duration.

Results

Literature search

Three of our sources searched databases for eligible studies: source 1 retrieved 126 titles, source 3 retrieved 7824 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 IMCI training lasted <5 days. The 59 included reports presented results from 31 distinct studies. Of these, our analysis focused on the 29 studies with at least one effect size based on ≥20 consultations per study group and time point. One additional analysis (described below), which allowed effect sizes based on ≥15 consultations per study group and time point, included all 31 studies.

General descriptive results

The 29 studies were from 23 countries (Table 2) in six WHO regions, and 14 (48%) were published in scientific journals. Five (17%) studies had first-tier study designs, and 24 (83%) had second-tier designs. The median duration for standard and short training was 11 and 7 days, respectively. 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 nine studies that reported the time spent on clinical practice, seven reported the minimum of 30%, one reported daily practice, and one reported less clinical instruction than the 11-day course (all details available in the online annex). Nurses and physicians were the most common health worker types studied. With short training, most courses were modified by shortening training methods (e.g. reducing the time participants spent reading text). 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).

Nineteen (66%) of the 29 studies used the health facility survey methodology recommended by WHO (observe consultations, interview caretakers, re-examine children, then interview health workers and assess the facility) (WHO 2003b). The remaining 10 studies used similar methods, but skipped at least one step (e.g. no re-examination). Among the studies that reported sample size information, the majority of the studies involved >100 patients from >10 health facilities.

The 29 studies reported on 42 distinct outcomes (Box 2), and most had multiple outcomes. Many studies used standard IMCI indicators (WHO 2003b). Only 15 (52%) studies used the PTIG outcome. For the 42 outcomes, we abstracted 280 outcome measures (Box 1); and with these measures, we calculated 117 effect sizes. The median number of outcomes per study included in the MES calculation was three (range: 1, 8). The 29 studies had 31 comparisons (Box 1): 27 studies each had one comparison (effect sizes 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).

In the 26 indirect comparison studies, the performance of health workers without IMCI training was very poor. The median PTIG value was 22.1% (range: 5.1, 51.4). IMCI training (standard and short) generally had a moderate effect [median MES increased 19.1%-points (range: -2.0, 67.5), and median PTIG increased 26.6%-points (range: -10.0, 94.0)]. Performance levels after training from direct and indirect comparison studies
<table>
<thead>
<tr>
<th>Country</th>
<th>Study identification number</th>
<th>Study design</th>
<th>Training duration (days)</th>
<th>Sample size</th>
<th>Median effect size (range of effect sizes)</th>
<th>PTTG effect size (range of effect sizes)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct comparison studies with a first-tier design (2 comparisons from 2 studies)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia [20]</td>
<td>PORC</td>
<td>11 vs. 6</td>
<td>82 113 377</td>
<td>No</td>
<td>7.8 (−2.7, 10.3)</td>
<td>−2.7</td>
<td>Mwinga et al. 2006</td>
</tr>
<tr>
<td>Uganda [17]</td>
<td>PORC</td>
<td>11 vs. 9</td>
<td>n.a. 104 n.a.</td>
<td>No</td>
<td>2.1 (−3.9, 8.1)</td>
<td>8.1</td>
<td>Tavrow et al. 2002</td>
</tr>
<tr>
<td>Direct comparison study with a second-tier design (1 comparison from 1 study)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kosovo [21]</td>
<td>PONRC</td>
<td>11 vs. 8</td>
<td>30 56 351</td>
<td>No</td>
<td>−0.5 (−6.0, 2.0)</td>
<td>−6.0</td>
<td>Syla 2003</td>
</tr>
<tr>
<td>Indirect comparison studies with a first-tier design (4 comparisons from 3 studies)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mali [10]</td>
<td>PORC</td>
<td>11</td>
<td>10 10 364</td>
<td>No</td>
<td>20.9 (only 1 effect size)</td>
<td>n.a.</td>
<td>Gillroy et al. 2004</td>
</tr>
<tr>
<td>Benin [24] comparison 1</td>
<td>PPNRC</td>
<td>11</td>
<td>99 196 757</td>
<td>Yes</td>
<td>42.4 (38.4, 46.4)</td>
<td>46.4</td>
<td>Osterholt et al. 2009; Rowe et al. 2009a</td>
</tr>
<tr>
<td>Indirect comparison studies with a second-tier design (24 comparisons from 23 studies)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China [31]</td>
<td>PPNC</td>
<td>5</td>
<td>419 n.a. 696</td>
<td>Yes</td>
<td>40.9 (4.9, 53.1)</td>
<td>n.a.</td>
<td>Zhang et al. 2007</td>
</tr>
<tr>
<td>Peru [22]</td>
<td>CACO and PONRC</td>
<td>5–7</td>
<td>15 n.a. 58</td>
<td>Yes</td>
<td>7.8 (−2.5, 28.0)</td>
<td>n.a.</td>
<td>Anonymous 2004</td>
</tr>
<tr>
<td>Niger [13]</td>
<td>CACO</td>
<td>6</td>
<td>44 n.a. 216</td>
<td>No</td>
<td>22.0 (−1.0, 37.5)</td>
<td>22.0</td>
<td>Degbey 2005; Ministry of Health of Niger 2005</td>
</tr>
<tr>
<td>Brazil [3]</td>
<td>PONRC</td>
<td>6–8</td>
<td>653 653 584</td>
<td>Yes</td>
<td>10.7 (−11.7, 32.9)</td>
<td>15.3</td>
<td>Amaral 2002; Amaral et al. 2004; Gouves et al. 2004</td>
</tr>
<tr>
<td>Ecuador [5]</td>
<td>PPNC</td>
<td>7</td>
<td>37 47 113</td>
<td>No</td>
<td>35.7 (only 1 effect size)</td>
<td>n.a.</td>
<td>Anonymous 1997; Ministry of Health of Ecuador 2000</td>
</tr>
<tr>
<td>Peru [14]</td>
<td>PONRC</td>
<td>7</td>
<td>90 202 372</td>
<td>No</td>
<td>−2.0 (−10.0, 31.0)</td>
<td>−10.0</td>
<td>Ministry of Health of Perú 2000; Huicho et al. 2005</td>
</tr>
<tr>
<td>Peru [23]</td>
<td>PPNC</td>
<td>7</td>
<td>7 n.a. 670</td>
<td>No</td>
<td>8.0 (0, 15.0)</td>
<td>n.a.</td>
<td>Ortiz et al. 1998</td>
</tr>
<tr>
<td>Bolivia [28]</td>
<td>PPNC</td>
<td>11</td>
<td>36 54 102</td>
<td>Yes</td>
<td>10.8 (only 1 effect size)</td>
<td>n.a.</td>
<td>Anonymous 1999; Zamora et al. 2002; Cordero et al. 2004</td>
</tr>
<tr>
<td>Cambodia [4] comparison 1</td>
<td>PONRC</td>
<td>11</td>
<td>80 n.a. 262</td>
<td>Yes</td>
<td>33.7 (18.5, 60.4)</td>
<td>60.4</td>
<td>Rathmony 2006a; Rathmony 2006b; Rehls 2007</td>
</tr>
<tr>
<td>Cambodia [4] comparison 2</td>
<td>PONRC</td>
<td>11</td>
<td>80 n.a. 216</td>
<td>No</td>
<td>34.4 (3.7, 43.8)</td>
<td>43.8</td>
<td>Rathmony 2006a; Rathmony 2006b; Rehls 2007</td>
</tr>
<tr>
<td>Ethiopia [7]</td>
<td>CACO</td>
<td>11</td>
<td>43 43 102</td>
<td>Yes</td>
<td>3.8 (−1.4, 8.9)</td>
<td>n.a.</td>
<td>Salgado et al. 2002</td>
</tr>
<tr>
<td>Senegal [29]</td>
<td>CACO</td>
<td>11</td>
<td>41 n.a. 1217</td>
<td>No</td>
<td>14.2 (−3.8, 50.1)</td>
<td>n.a.</td>
<td>Briggs et al. 2002</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Country [Study identification number]</th>
<th>Study design</th>
<th>Training duration (days)</th>
<th>Sample size[b] N_HF N_HW N_patient</th>
<th>Other interventions[c]</th>
<th>Median effect size[d] (range of effect sizes)</th>
<th>PTIG effect size[d]</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Africa [15]</td>
<td>PPNC</td>
<td>11</td>
<td>21 42 96</td>
<td>No</td>
<td>9.0 (−4.0, 22.0)</td>
<td>n.a.</td>
<td>Chopra et al. 2005</td>
</tr>
<tr>
<td>Sudan [30]</td>
<td>CACO</td>
<td>11</td>
<td>66 n.a. 254</td>
<td>Yes</td>
<td>24.6 (2.4, 47.9)</td>
<td>24.6</td>
<td>Federal Ministry of Health of Sudan 2004; WHO EMRO 2004</td>
</tr>
<tr>
<td>Uzbekistan [19]</td>
<td>CACO</td>
<td>11</td>
<td>120 n.a. 170</td>
<td>Yes</td>
<td>1.8 (0, 3.5)</td>
<td>n.a.</td>
<td>Rehlis 2003</td>
</tr>
<tr>
<td>Togo [25]</td>
<td>PONRC</td>
<td>11</td>
<td>2 n.a. 300</td>
<td>No</td>
<td>64.3 (34.6, 94.0)</td>
<td>94.0</td>
<td>Atakouma Dragiñé et al. 2006</td>
</tr>
<tr>
<td>Togo [26]</td>
<td>PPNC</td>
<td>11</td>
<td>45 221 166</td>
<td>No</td>
<td>19.4 (only 1 effect size)</td>
<td>n.a.</td>
<td>Assimadi et al. 1999; Assimadi et al. 2003</td>
</tr>
<tr>
<td>Zambia [27]</td>
<td>PPNC</td>
<td>11</td>
<td>8 n.a. 223</td>
<td>Yes</td>
<td>17.4 (−30.7, 38.4)</td>
<td>n.a.</td>
<td>Burnham 1997</td>
</tr>
<tr>
<td>Tanzania [16]</td>
<td>PONRC</td>
<td>11–16</td>
<td>73 n.a. 404</td>
<td>Yes</td>
<td>36.5 (16.0, 78.0)</td>
<td>38.0</td>
<td>Mgalula 2000; Mbuya C et al. 2003; Armstrong Schellenberg et al. 2004a; Armstrong Schellenberg et al. 2004b; Gouws et al. 2004; Bryce et al. 2005</td>
</tr>
<tr>
<td>Uganda [18]</td>
<td>PONRC</td>
<td>11–14</td>
<td>316 427 1265</td>
<td>Yes</td>
<td>15.8 (13.9, 29.3)</td>
<td>13.9</td>
<td>Gouws et al. 2004; Pariyo et al. 2005</td>
</tr>
<tr>
<td>Kenya [9]</td>
<td>PPNC</td>
<td>15</td>
<td>36 n.a. 1043</td>
<td>No</td>
<td>33.0 (0.0, 46.0)</td>
<td>n.a.</td>
<td>Centers for Disease Control and Prevention 1998; Lee et al. 2001</td>
</tr>
</tbody>
</table>

Indirect comparison studies with a second-tier design with <20 consultations per study group and time point (2 comparisons from 2 studies, used only in one of the alternative sensitivity analyses)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>South Africa</td>
<td>CACO</td>
<td>11</td>
<td>32 n.a. 41</td>
<td>No</td>
<td>55.0 (47.1, 62.9)</td>
<td>62.9</td>
<td>Goga 2003; South Africa National Department of Health 2005</td>
</tr>
<tr>
<td>Vietnam [33]</td>
<td>CACO</td>
<td>11</td>
<td>70 n.a. 110</td>
<td>No</td>
<td>32.0 (13.6, 50.3)</td>
<td>n.a.</td>
<td>Ministry of Health of Vietnam 2002</td>
</tr>
</tbody>
</table>

Notes:
- CACO = case-control; IMCI = Integrated Management of Childhood Illness; N_HF = health facility sample size; N_HW = health worker sample size; N_patient = patient sample size; PPNC = pre-post with randomized controls; PONRC = post-only with randomized controls; PPNC = pre-post with non-randomized controls; PONRC = post-only with non-randomized controls; PPNC = pre-post with no controls; PTIG = patient (needing oral antimicrobial or rehydration therapy) treated according to IMCI guidelines.
- Label that identifies a group of one or more reports with a single study.
- See online annex for details on sample sizes.
- Strategies besides IMCI training (e.g. extra supervision) implemented to support health workers’ adherence to IMCI guidelines.
- Percentage-point differences; positive effect size means standard training was better; negative effect size means shortened training was better.
- 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.
- Of the three outcomes analysed, one outcome was measured in a PONRC study design, and two outcomes were measured in a CACO design.
- PTIG was 7.0 percentage-points for this study; however, as the value was based on only 18 consultations in the IMCI group, it was not included in the primary analysis (it was only included in one of the alternative sensitivity analyses).
usually revealed considerable room for improvement. The median of study-specific medians of all post-training outcome measures was 75.0% (range: 40.9, 97.3), and the median of post-training PTIG measures was 66.4% (range: 11.0, 100.0). Thus, typically, even after IMCI training (and sometimes other supports), 34% of ill children needing oral antimicrobials or rehydration did not receive these treatments according to IMCI guidelines. Table 3 summarizes the various analyses performed, which are described in more detail below.

Direct comparison studies
A Zambian study compared standard 11-day training with a 6-day course developed for physicians (Mwinga et al. 2006). The study, however, enrolled nurse-level workers (no physicians). The follow-up period was 4–6 months. A Ugandan study compared 11-day training with a 9-day course that included computer-based training (Tavrow et al. 2002). Clinical officers and nurses were enrolled. Follow-up was at 3–4 months. A study from Kosovo compared 11-day training with an 8-day course (Syla 2003). Only physicians were enrolled. Follow-up was at 2–3 years.

MES main analysis
A direct comparison of standard versus short training (two effect sizes from two studies; Figure 2, top graph, filled circles in left column) suggested that standard was slightly better than short training [median = 5.0%-points (range: 2.1, 7.8); no statistical testing because of small sample]. An indirect comparison of standard versus short training could not be performed because no studies with a first-tier design compared short training to no IMCI.

PTIG main analysis
Direct comparison (two effect sizes from two studies; Figure 2, bottom graph) showed that standard and short training were very similar [standard was better by 2.7%-points (range: −2.7, 8.1); no statistical testing]. An indirect comparison could not be performed because no studies with a first-tier design compared short training to no IMCI.

MES sensitivity analysis
Direct comparison (three effect sizes from three studies; Figure 2, top graph, all circles in left column) showed that standard training was very similar [standard was better by 2.1%-points (range: −0.5, 7.8); no statistical testing].

Indirect comparison (28 effect sizes from 26 studies) suggested that standard training [median MES = 19.4%-points (range: 1.8, 67.5)] was somewhat better than short training [median MES = 10.7%-points (range: −2.0, 40.9)] (Figure 2, top graph, all circles in middle and right columns). This difference of about 9%-points was not statistically significant (P = 1.0). Training duration analysed as a continuous variable revealed a non-significant trend (P = 0.15) of increasing effect size with increasing training duration. Multivariable modelling also showed a non-significant trend (P ≥ 0.11) (details available in the online annex). Additionally, the models 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).

PTIG sensitivity analysis
Direct comparison (three effect sizes from three studies; Figure 2, bottom graph) showed that standard and short training were very similar [short was better by 2.7%-points (range: −6.0, 8.1); no statistical testing].

Indirect comparison (14 effect sizes from 12 studies) suggested that standard training [effect size = 38.0%-points (range: 3.9, 94.0)] was much better than short training [effect size = 15.3%-points (range: −10.0, 22.0)]. The 23%-point difference was not statistically significant (P = 0.19). Training duration analysed as a continuous variable revealed a borderline significant trend of increasing effect size with increasing training duration (5.6%-point increase per additional day of training; P = 0.09); and multivariable modelling supported this finding (P-values ranged from 0.09 to 0.11) (details available in the online annex). Additionally, the models showed that IMCI’s effect decreased as baseline performance levels increased, although this finding was borderline significant at best (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).

Other interventions and time since training
From the main analyses (Table 3, rows 1 and 2), standard training with other interventions seemed much more effective than without other interventions (35 to 42%-point difference). In contrast, the sensitivity analyses found, for any training duration, no significant difference between training with other interventions versus training without other interventions (−9 to 9%-points; P-values ranged from 0.23 to 0.80). None of the analyses revealed any evidence of deteriorating performance over time after IMCI training. Details on both analyses are available in the online annex.

Alternative analyses
We investigated the impact of relaxing and tightening the sample size inclusion criteria in two alternative sensitivity analyses. In the first 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). In the second analysis, we excluded one outlier (Study ID 25) because it had only two health facilities and unusually high effect sizes (Table 2). Results from both analyses were similar to those from the primary analysis (details available in the online annex).

Training costs
Few studies reported training costs, although the direct comparisons did. In Zambia, standard training was US$828 per trainee compared with US$450 per trainee for short training (46% lower) (Mwinga et al. 2006). In Uganda, standard training cost US$172 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
<table>
<thead>
<tr>
<th>Study design adequacy and summary measure</th>
<th>Analysis*</th>
<th>Indirect comparison (standard training vs. no IMCI compared with short training vs. no IMCI)</th>
<th>Training duration analysed as a continuous variable</th>
<th>Impact of other interventions (besides IMCI training)</th>
<th>Analysis of time since training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main analysis (first-tier study designs only)</td>
<td>Direct comparison (standard compared with short training)</td>
<td>Standard was slightly better (5%-point difference, no statistical testing)</td>
<td>No comparison possible (no studies of short vs. no IMCI)</td>
<td>Suggested increasing effect size with longer training duration (only 4 effect sizes, no statistical testing)</td>
<td>Standard training with other interventions seemed much more effective than without other interventions (35%-point difference)</td>
</tr>
<tr>
<td>Outcome= MES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>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)</td>
</tr>
<tr>
<td>Note: Few studies (&lt;6 effect sizes from 5 studies)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome= PTIG</td>
<td>Standard and short were very similar (3%-point difference, no statistical testing)</td>
<td>No comparison possible (no studies of short vs. no IMCI)</td>
<td>Suggested increasing effect size with longer training duration (only 3 effect sizes, no statistical testing)</td>
<td>Standard training with other interventions seemed much more effective than without other interventions (42%-point difference)</td>
<td>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)</td>
</tr>
<tr>
<td>Note: Very few studies (&lt;5 effect sizes from 4 studies)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis (first- and second-tier study designs)</td>
<td>Direct comparison (standard compared with short training)</td>
<td>Standard was somewhat better than short (9%-point difference, ( P = 1.0 ))</td>
<td>Trend of increasing effect size with longer training duration {univariate ( P = 0.15 ) (1.9%-point increase per extra day of training); multivariable ( P )-values ranged from 0.11–0.51}</td>
<td>For any training, no significant difference between training with other interventions versus training without other interventions ((-9) to (9)%-points, ( P )-values ranged from 0.23–0.80)</td>
<td>No apparent trends over time after either standard ((P=0.91)) or short training (only 4 effect sizes, no statistical testing)</td>
</tr>
<tr>
<td>Outcome= MES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(&lt;31) effect sizes from (29) studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome= PTIG</td>
<td>Standard and short were very similar (-3%-point difference, no statistical testing)</td>
<td>Standard was much better than short (23%-point difference; ( P = 0.19 ))</td>
<td>Trend of increasing effect size with longer training duration {univariate ( P = 0.09 ) (5.6%-point increase per extra day of training); multivariable ( P )-values ranged from 0.09–0.11}</td>
<td>For any training, no significant difference between training with other interventions versus training without other interventions ((-9) to (9)%-points, ( P )-values ranged from 0.58–0.74)</td>
<td>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)</td>
</tr>
<tr>
<td>(&lt;17) effect sizes from (15) studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- 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.
- All differences are in terms of 'effect of standard training' minus 'effect of short training' (i.e. positive differences indicate standard training led to greater improvements in health care quality than short training).
- *All differences are in terms of 'effect of standard training' minus 'effect of short training' (i.e. positive differences indicate standard training led to greater improvements in health care quality than short training).
we could have been purists and included only ideal studies, we
decided to broaden our scope. In so doing, to the best of our
knowledge, this review represents the most comprehensive
examination of IMCI effectiveness on health worker perform-
ance to date. We performed layerd 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 (standard training minus short training) from four
analyses of direct comparisons ranged from −3 to +5%-points
(median = 2%-points), and differences from two analyses of
indirect comparisons ranged from 9 to 23%-points (me-
dian = 16%-points). Indirect comparisons that examined train-
ing duration as a continuous variable suggested that the effect
of IMCI training increases by 2 to 6%-points 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 was a range bounded by the
median values of the direct and indirect comparisons, i.e. a 2 to
16%-point advantage for standard training.

Are these modest differences relevant? In even a small
country with low access to health facility-based services, an
improvement of 2−16%-points in health worker performance
could translate annually into an additional 30 000−240 000
children treated correctly and 1000−9000 child deaths pre-
vented—a considerable impact (details available in the online
annex).

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

Although the effectiveness of other interventions to support
IMCI was not our focus, the review did provide some
interesting insights. First, in several analyses, IMCI’s effect
with other interventions was greater than without other
interventions; however, results varied substantially from −9 to
+42%-points. Second, two studies with other interventions
had particularly large effect sizes; both had first-tier designs.
The largest PTIG effect was from Bangladesh (Study ID 1),
and the other interventions were purchasing essential IMCI
drugs and supplies and monthly supervision by IMCI-trained
medical officers. 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 3 years, among con-
sultations performed by IMCI-trained health workers with
study supports, the proportion of ill children receiving IMCI-
recommended treatments was 27%-points higher than for
IMCI-trained health workers in a comparison area with
‘usual’ supports.

Discussion
We addressed the question: does shortening IMCI in-service
training reduce its effectiveness? The review was complicated
by the fact that only two ideal studies were identified (direct
comparisons of standard versus short training with a first-tier
design). However, we did find one direct comparison with a
second-tier study design and numerous studies comparing IMCI
to no IMCI, which permitted indirect comparisons. Although

Figure 2 Median effect sizes (top graph) and PTIG effect sizes (bottom
graph) of IMCI training stratified by comparison type Notes:
IMCI = Integrated Management of Childhood Illness; MES = median
effect size; PTIG = patient treated according to IMCI guidelines. Filled
circles represent data from studies with first-tier designs, and open
circles represent data from studies with second-tier designs.

rented) (Tavrow et al. 2002). In Kosovo, costs were US$430 per
trainee for standard training and US$240 per trainee for short
training (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, 6 November 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).
We also examined baseline and post-IMCI-implementation values for our outcomes. At baseline (before IMCI training), median outcome measures and PTFG 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, studies in many countries identified 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 oral antimicrobials or rehydration were not receiving these treatments according to IMCI guidelines. This result illustrates the challenge of getting health workers to follow clinical guidelines and underscores the importance of providing ongoing support for workers after training.

Limitations
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 direct comparison studies with first-tier designs; and one of these (Tavrow et al. 2002) shortened training by only 2 days. Second, the main analysis included very few studies. Third, indirect comparisons were susceptible to bias caused by factors that were different between standard and short training studies, besides training duration. For example, data on training quality were scarce, and it is unclear to what degree the trainings were comparable. Additionally, reliable measurement and quantitative adjustment of non-training factors was challenging. We attempted to account for two such factors (other interventions and baseline performance levels); however, our methods were simplistic. Fourth, the summary measure MES included different outcomes in different studies. Thus, MES comparisons might have been biased because some performance indicators 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 small sample sizes) or we ignored the precision of effect sizes. The ideal analytic approach, however, would have required study datasets, which would have been too difficult to obtain, and thus this was impractical. Sixth, when we did perform statistical testing, we found that almost none of the results were significant; and the few significant results close to the 0.05 level appeared after numerous comparisons. 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 >100 patients. Seventh, only about two-thirds of studies used re-examinations to make a ‘gold standard’ determination of the child’s IMCI illness classifications. Without such a standard, training effects might be overestimated, depending on health workers’ ability to classify illnesses correctly. 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 for some studies, and many studies reported results for just one time point.

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, programme 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; WHO 2001; Rowe et al. 2005; Rowe et al. 2009b; 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 should be retained (i.e. those addressing major causes of deaths for which effective interventions exist) (WHO 2007). In all circumstances, it is critical to implement strategies to strengthen health systems in addition to training in IMCI.
**Funding**

SY Rowe and K Herman were partly supported by WHO’s Department of Child and Adolescent Health and Development.

**Acknowledgments**

Preliminary methods and results were presented at the WHO’s *Technical consultation on IMCI training approaches and review of pre-service training*, 19–23 November 2007, Geneva. 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 permission to include data from unpublished health facility survey reports. We greatly appreciated the thoughtful comments provided by Rajiv Bahl, Jose Martinez, and the participants of the above-mentioned WHO technical consultation.

**Conflict of interest**

None declared.

**Supplementary Data**

Supplementary data are available at *Health Policy and Planning* Online.

**References**


