Review

Evaluating the impact of healthcare provider training to improve tuberculosis management: a systematic review of methods and outcome indicators used

Shishi Wu a, Imara Roychowdhury b, Mishal Khan a,b,*

a Saw Swee Hock School of Public Health, National University of Singapore, Singapore
b Communicable Diseases Policy Research Group, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, UK

1. Introduction

Infectious diseases place a great burden on the health systems and economies of low- and middle-income countries (LMICs) and remain the leading causes of death and disability worldwide.1–3 According to the most recent World Health Organization (WHO) report, four of the top ten causes of death in LMICs each year are from infectious diseases, among which tuberculosis (TB) is the leading cause of death from a single infectious disease.4 It is widely recognized that developing human resources capacity is vital for TB control in LMICs.5 It is widely recognized that developing human resources capacity is vital for TB control in LMICs.5 and insufficient quality, quantity, and distribution of healthcare providers (HCPs) was highlighted as a major challenge in the WHO Global Plan to Stop TB 2006–2015.6,7 Inconsistent and inadequate quality of services provided by HCPs is documented as a prevalent problem that results in poor case detection and adherence to treatment, thereby hindering the progress on TB control.8–12 In line with the need for human resource capacity building, investments in TB healthcare provider (HCP) training programmes have increased; for example, in 2014, the Global Fund to Fight AIDS, Tuberculosis and Malaria provided 16 million person-episodes of training for HCPs, which was a ten-fold increase over the number trained in 2005.13

Evaluations are essential to determine whether public health programmes and interventions have been successful, and to inform decisions about future investments,14 but these are often challenging to design and conduct. Training evaluation refers to a systematic measurement of intended outcomes of training activities.15 It helps answer two main questions: whether training objectives were achieved, and whether the accomplishment of those objectives resulted in enhanced performance on the job.16 In addition, results of training evaluations help ensure training meets the needs of learners and organizations.

Despite growing investment in and focus on improving human resource capacity for TB control through training, no studies could...
be identified that have systematically reviewed the existing literature on evaluations of TB HCP training programmes. A systematic review was therefore conducted to investigate whether HCP training programmes to improve TB control are frequently evaluated, and to synthesize information on the methods and outcome indicators used in assessments.

2. Methods

This systematic scoping review was based on the methodological framework designed by Arksey and O’Malley. The following key steps were included in performing the review: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarizing, and reporting the results.

The population of interest for this review was TB HCPs, who were defined as doctors, nurses, healthcare workers, lay health workers, traditional health practitioners, clinical staff, and laboratory technicians. Teachers and other professionals delivering health services outside their routine work were not considered HCPs. The intervention of interest was any training or capacity building activity related to TB health service delivery or reporting.

A search for articles published after January 1, 2000 was conducted in three electronic databases on April 28, 2016: PubMed, Embase, and Cochrane Library. No language restrictions were applied at the search stage. In addition, the relevant grey literature was searched in Google Scholar (first 100 citations), as well as five major non-governmental organization (NGO) websites on July 18, 2016: WHO, Oxfam International, Save the Children, Community Health Workers Central (CHW Central), and Target TB, UK. The search terms used are summarized in Table 1.

All citations were imported into EndNote X7 and duplicate citations were removed manually. A two-stage screening process for eligibility was conducted. Articles were eligible for inclusion if the studies described evaluations of HCP training programmes for TB and contained descriptions of the training programme, method used to evaluate the programme, and outcomes assessed in the evaluation. Papers that did not report outcomes or results of the evaluation were excluded from analysis. The geographic areas of the studies and the time point of training (pre-employment or on the job) were not restricted. In the first stage screening, researchers reviewed only the titles and abstract of the citations to exclude irrelevant articles. Two researchers (SW, IR) conducted the screening independently. Results from both researchers were compared at the end of the first stage screening process. Titles for which an abstract was not available or for which either of the reviewers suggested inclusion were put forward for subsequent full-text review. In the second stage, a full-text review of articles included after title and abstract screening was conducted. If the studies did not meet the eligibility criteria, they were excluded at this stage. Articles that could not be obtained through online database and library searches at the National University of Singapore and London School of Hygiene and Tropical Medicine were also excluded from the final analysis.

Relevant information was then extracted from the articles included in the final analysis using a pre-designed standardized Microsoft Excel template. Table 2 summarizes data extracted and definitions used for categorizing data. For each study, the training outcomes evaluated were categorized into four levels (reaction, learning, behaviour, and results) based on the Kirkpatrick model; this was the first framework designed for training evaluation, and remains the most commonly used. When necessary, studies could be included in multiple categories. Descriptive statistics were used to summarize the data extracted.

### Table 1

<table>
<thead>
<tr>
<th>Search terms</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(healthcare workers OR healthcare providers OR healthcare professionals OR healthcare staff OR healthcare practitioners OR health workers OR health providers OR health professionals OR health staff OR health practitioners OR health-care workers OR health-care professionals OR health-care providers OR health-care practitioners OR health-care staff) AND (training) AND (evaluat* OR assess*) AND (tuberculosis OR TB)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Definitions of extracted data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data extracted</strong></td>
</tr>
<tr>
<td>Year of publication</td>
</tr>
<tr>
<td>Study location</td>
</tr>
<tr>
<td>Population trained</td>
</tr>
<tr>
<td>Training methods</td>
</tr>
<tr>
<td>Lecture</td>
</tr>
<tr>
<td>Small group based</td>
</tr>
<tr>
<td>Experiential hands-on training</td>
</tr>
<tr>
<td>Not specified</td>
</tr>
<tr>
<td>Evaluation methods</td>
</tr>
<tr>
<td>Pre- and post-training tests</td>
</tr>
<tr>
<td>Interviews</td>
</tr>
<tr>
<td>Review of patient records</td>
</tr>
<tr>
<td>Observation</td>
</tr>
<tr>
<td>Focus group discussion</td>
</tr>
<tr>
<td>Cost-effective analysis</td>
</tr>
<tr>
<td>Outcomes evaluated</td>
</tr>
<tr>
<td>Reaction</td>
</tr>
<tr>
<td>Learning</td>
</tr>
<tr>
<td>Behaviour</td>
</tr>
<tr>
<td>Results</td>
</tr>
</tbody>
</table>

HCP, healthcare provider.
3. Results

A total of 304 citations were retrieved from the three databases and 350 from the grey literature search. After removing duplicates, 499 unique publications were screened, of which 21 met the inclusion criteria and were included in the analysis. Figure 1 shows the screening and selection process of this study. A table providing information on the individual studies is given in the Supplementary Material.

All of the 21 TB HCP training evaluations were published after 2005, with a fairly even distribution of publications in the decade spanning 2005 to 2015 (Table 3). In terms of geographic distribution of studies, most (n = 16, 76%) took place in African countries, including the Democratic Republic of Congo, Ethiopia, Kenya, Malawi, Senegal, South Africa, Tanzania, and Uganda. South Africa was the country with the largest number of TB HCP training programme evaluations (n = 7, 33%). There was only one HCP training evaluation study conducted in North America (USA), one in South America (Ecuador), and three in Asia (India, Malaysia, and Taiwan).

3.1. Evaluation methods applied

A wide range of training evaluation methods was used. As shown in Figure 2, the most common method was reviewing patient records to assess diagnostic and treatment outcomes after HCPs had attended training sessions; eight evaluations (38%) used this method. The second most commonly used evaluation method was post-training interview with trainees, which included both semi-structured interviews and structured interviews; seven (33%)...
studies applied this approach. Pre- and post-training tests were used in six (29%) studies to assess trainees’ knowledge gain after training programmes. Approaches such as cost-effective analysis, post-training focus group discussion with trainees, and observation of on-the-job performance of trainees were not commonly used.

3.2. Training outcomes assessed

In terms of the outcomes evaluated in the studies, Figure 3 shows which of the four outcome levels based on the Kirkpatrick model – reaction, learning, behaviour, and results – were assessed. Among the 21 studies included, more than half \( (n = 12, 57\%) \) evaluated knowledge acquisition (learning) of trainees after training sessions. Nine (43\%) studies evaluated downstream results of the training programmes. Fewer studies \( (n = 4, 19\%) \) assessed whether trainees liked the programme or whether the programme was considered useful for trainees (reaction), and six (29\%) measured the behaviour change of trainees after they had finished the training and returned to their jobs.

4. Discussion

This scoping review of TB HCP training evaluations showed that there is very limited evidence available for decision-making, with only 21 studies identified on this important topic since 2000. Those papers that reported on evaluations of training programmes mainly applied the following methods: review of patient records, interview with trainees, and pre- and post-training tests. In terms of outcome indicators assessed, a particular gap was identified in robust evidence about behaviour change of trainees and their views on the usefulness of training.

Even though more funds have been invested in HCP training since 2000 in order to improve human resource capacity,\(^{24}\) the annual number of evaluations of TB HCP training has not increased since 2005. The present analysis of the geographic distribution of evaluation studies showed that most were conducted in Africa, with very limited evidence about training programme outcomes from Asia. Even though Asian countries account for a substantial proportion of global TB cases – among the top six countries with the largest number of incident cases, four of them are India, Indonesia, Pakistan, and China\(^{23}\) – only three studies were conducted in Asia.

It was found that most studies evaluated knowledge, using pre- and post-training tests, a method that is best suited for testing retention of factual knowledge.\(^{24}\) Improvements in service quality – such as standardized application of diagnostic and treatment guidelines and improved communication with TB patients – are important outcomes of training that were rarely evaluated. Here post-training tests have limitations, as they can only assess whether trainees know about the guidelines or best practices, but cannot determine whether trainees have actually applied these strategies in their jobs. The assessment of behavioural change of HCPs after attending training programmes is critical in determining whether the training will have an impact on patient care and TB control indicators. Currently, most evaluations of behavioural change have been conducted using qualitative research methods asking participants, for example, whether they have applied newly acquired skills.\(^{25}\) This approach largely relies on self-reported information, and the actual on-the-job performance of trainees is not objectively evaluated; there is a risk of bias in assessments using this method, as trainees may not be willing to reveal that they ignore or have forgotten the skills learned at the training sessions to evaluators, who are often directly involved in conducting or overseeing the interviews. Furthermore, it was found that many studies only conducted post training tests immediately after the training sessions; therefore, these evaluations do not provide evidence about whether the
training programmes have resulted in long-term knowledge gain.

In comparison to TB, twice as many training programme assessments have been published on HIV. The evaluation methods applied in assessing HIV HCP training programmes are also more diverse. For example, standardized patients, which refers to the use of actors trained to accurately portray a specific medical condition,19 have been used in evaluating on-the-job behaviour change of HIV HCPs after training programmes.20–22 Compared to self-reported changes in practice, this method is able to reduce bias because trainees do not know when a clinical encounter with a standardized patient will occur.20

Downstream programmatic results are commonly used to evaluate TB HCP training programmes. As part of these evaluations, researchers typically compared TB treatment outcomes before and after training programmes or between intervention and control groups by reviewing patient records. Manuals such as the WHO guidance for monitoring and evaluating national TB programmes are used by evaluators to define specific patient-level outcome indicators to assess the effectiveness of TB control programmes, such as the successful treatment rate, case detection, ‘appropriate care rate’, and proportion of patients successfully completing the full course of treatment.23,24 However, training programmes are often embedded in a broader national TB control strategy with other TB prevention and control activities ongoing in parallel. Therefore, changes in downstream programmatic or organizational results, such as improved TB case detection or the treatment success rate, may not directly result from the HCP training programmes, and the attribution of success is challenging using a simple before–after training evaluation approach. Other factors, such as improved supply of medical equipment or enhanced healthcare infrastructure, may also contribute to better patient outcomes.

This systematic review, which is the first to summarize evidence on TB HCP training evaluations, was strengthened by a broad search of the grey and published literature including three scientific databases, Google Scholar, and five NGO websites. However, it is recognized that some TB HCP training evaluations may not have been captured by this study if they were not accessible online. Additionally, as acknowledged in other scoping reviews, the quality of the studies included was not assessed, because the aim of this scoping review was to map the existing literature in terms of their volume, nature, and characteristics.17

The lack of rigorous TB HCP training evaluation studies identified through this review could reflect the limited experience, knowledge, and budget available to TB programme managers or researchers in LMICs to conduct training evaluations. It is believed that further investigation into the difficulties encountered in conducting training evaluations in LMICs would be useful in identifying strategies to increase the evidence base in this area. This could be done through a systematic review and analysis of the limitations mentioned in the existing literature and interviews with training programme managers.

In conclusion, as the number of TB HCP training interventions has been growing to address the shortage of human resources for healthcare delivery globally, rigorous training evaluations are critical to determine the effectiveness of interventions. However, through this review, it was found that evidence from evaluations of TB HCP training is extremely limited. Specific evidence gaps in TB HCP training include the following: results based on robust evaluation designs, assessments of on-the-job behaviour change in HCPs, and training evaluations in high TB burden Asian countries.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest: None.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijid.2016.11.421.

References


