

Original citation:

Auguste, Peter, Tsertsvadze, Alexander, Court, Rachel A. and Pink, Joshua. (2016) A systematic review of economic models used to assess the cost-effectiveness of strategies for identifying latent tuberculosis in high-risk groups. *Tuberculosis*, 99. pp. 81-91.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/79316>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

© 2016, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>

A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

A systematic review of economic models used to assess the cost-effectiveness of strategies for identifying latent tuberculosis in high-risk groups

Peter Auguste^{a*}, Alexander Tsertsvadze^b, Rachel Court^a, Joshua Pink^a

^a Warwick Evidence, Warwick Medical School, University of Warwick, Coventry, CV4 7AL

^b Evidence in Communicable Disease Epidemiology and Control, Health Sciences, Warwick Medical School, University of Warwick, Coventry, CV4 7AL

Correspondence to: Peter Auguste
Warwick Evidence
Warwick Medical School
University of Warwick
Coventry, CV4 7AL
Tel: +44 (0)2476 150648
Email: p.auguste@warwick.ac.uk

Declared competing interests of authors: none

This journal should be referenced as follows:

A systematic review of economic models used to assess the cost-effectiveness of strategies for identifying latent tuberculosis in high-risk groups

Abstract

Background

Timely diagnosis and treatment of latent tuberculosis infection (LTBI) through screening remains a key public health priority. Although globally it is recommended to screen people at high risk of developing TB, the economic evidence underpinning these recommendations is limited. This review critically appraised studies that had used a decision-analytical modelling framework to estimate the cost-effectiveness of interferon gamma release assays (IGRAs) compared to tuberculin skin test (TST) for detecting LTBI in high risk populations.

Methods

A comprehensive search of MEDLINE, EMBASE, NHS-EED was undertaken from 2009 up to June 2015. Studies were screened and extracted by independent reviewers. The study quality was assessed using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) and the Philips' checklist, respectively. A narrative synthesis of the included studies was undertaken.

Results

Ten of 8793 studies were considered relevant for inclusion. Two economic evaluations were conducted in a child population, six in an immunocompromised population and two in a recently arrived population. Most studies (n=7) used a decision tree structure with Markov nodes. In general, all models performed well in terms of reporting quality, but were subject to limitations to structure and model inputs. Models have not elaborated on their setting or the perspective of the studies was not consistent with their analyses. Other concerns were related to derivation of prevalence, test accuracy and transition probabilities.

Conclusion

Current methods available highlight limitations in the clinical effectiveness literature, model structures and assumptions, which impact on the robustness of the cost-effectiveness results. These models available are useful, but limited on the information that can be used to inform on future cost-effectiveness analysis. Until consideration is given on deriving the performance of tests used to identify LTBI that progresses to active TB, and the development of more comprehensive models, the economic benefit of LTBI testing with TST/IGRAs in high risk populations will remain unanswered.

Keywords: Systematic review, latent tuberculosis, decision-analytical modelling, cost-effectiveness

Background

Diagnosis and treatment of latent Tuberculosis infection (LTBI) through screening remains a key public health priority in the elimination of tuberculosis. For over a century, the tuberculin skin test (TST) has been used to diagnose LTBI, despite its many limitations. These include being neither very sensitive, due to anergy in an immunocompromised population, nor specific, due to cross-reactivity in people who are Bacilli Calmette-Guérin (BCG) vaccinated and those who are infected with non-tubercular mycobacteria (NTM).[1] Furthermore, TST requires people to return to have their results read, and there is the possibility of error when measuring the size of the induration of the skin reaction. This has led to the development of new *in vitro* interferon-gamma release assays (IGRAs) aimed at improving the diagnosis for LTBI.

Currently, two IGRAs are commercially available for the diagnosis of LTBI, QuantiFERON Gold In-tube (QFT-GIT) (Cellestis Ltd., Carnegie, Australia) and T-SPOT.TB (Oxford Immunotec Ltd, Oxford, UK). IGRAs do not boost responses due to repeated testing, and people are not required to make a second visit to have the results read.[2] These tests offer alternatives for the diagnosis of LTBI, but are more expensive. In the UK, current guidelines recommend the use of IGRAs and/or TST for the diagnosis of LTBI in high risk populations which include children, people who are immunocompromised or at risk of immunosuppression and people from countries with a high incidence of TB.[3] The health economic modelling which underpin these recommendations are based on ‘what-if’ analyses/scenarios rather than empirical screening evidence[4] and this offers little insight on which diagnostic strategy is the most cost-effective.

Decision makers, such as the National Institute for Health and Care Excellence (NICE), often rely on mathematical modelling to aid in decision making processes, as they are constantly faced with questions on what interventions should be funded. The purpose of modelling is to structure evidence on clinical and economic outcomes in a form that can be used to inform decisions on clinical practices and allocation of resources in order to achieve maximum benefits for health care.[5] Since the introduction of IGRAs, many studies have estimated the cost-effectiveness of various strategies for the diagnosis of LTBI using economic modelling in a decision analytical context. A previous clinical guideline [3] which included a systematic review highlighted that no published studies were identified in these high risk groups. Hence, in this review, the aim is to identify from recent literature the suitability of existing cost-effectiveness models that compared different diagnostic strategies for identifying LTBI in children, immunocompromised or at risk of immunosuppression and people from countries with a high incidence of TB.

Methods

Study eligibility criteria

Citations retrieved were screened by two reviewers (PA and AT) and included in the review if they met the following criteria: Children (immunocompetent), people who are immunocompromised or at risk from immunosuppression (e.g. transplant recipients or HIV) and recent arrivals from countries with a high incidence of TB (≥ 40 cases per 100,000), and comprising a formal economic evaluation involving direct comparison between IGRAs (QFT-G, QFT-GIT or T-SPOT.TB) and TST, and included a decision analytic model.

Search strategy

A search of the literature for published economic evaluations was performed for the purpose of identifying the suitability of existing cost-effectiveness models and their model design.

The cost-effectiveness search was developed and conducted as part of a wider systematic review that aimed to compare both the clinical effectiveness and cost-effectiveness of screening tests (IGRAs and TST) for LTBI in high risk groups.[6] Electronic databases were searched, applying the search strategy to the following databases: MEDLINE, MEDLINE In-Process Citations and Daily Update, Embase, NHS Economic Evaluation Database (NHS EED), Health Economics Evaluation Database (HEED), Science Citation Index, Research Papers in Economics (RePEC) and Cost-effectiveness analysis (CEA) Registry. The search was limited to English language and studies published between 2009 and June 2015. This time point was chosen because a clinical guideline [3] which included a systematic review[3] searched for studies published up to 2009, but did not identify any relevant economic modelling studies. Reference lists of potentially relevant articles were manually searched to identify additional studies. Details of search terms are presented in the Appendix.

Study selection

All citations retrieved were screened by two independent reviewers (PA and AT) at title/abstract level, of which potentially relevant publications were further examined for full text. Any disagreements between the reviewers were resolved by a consensus.

Data extraction

Data extraction was conducted by one reviewer (PA) and further cross-checked by a second reviewer (AT). Any disagreements were resolved by discussion or by recourse to a third party reviewer. Information was extracted on study details (title, author and year of study), baseline characteristics (population, intervention, comparator and outcomes), methods (study perspective, time horizon, discount rate, measure of effectiveness current, assumptions and analytical methods), results (study

parameters, base-case and sensitivity analysis results), discussion (study findings, limitations of the models and generalizability) and other (source of funding and conflicts of interests).

Quality assessment

The quality of the studies was assessed using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS)[7] and the Philips' checklist,[8] respectively. The CHEERS assessment tool comprises of six dimensions which include title and abstract, introduction, methods, results, discussion and other. The Philips' quality assessment tool comprises of two main dimensions, structure of the model and data used to parameterized the model. Study quality was assessed by one reviewer (PA) and cross-checked by a second reviewer (AT).

Data synthesis

Information extracted from the included studies were summarised and presented in Table 1. These findings were compared narratively, and recommendations for the future modelling of LTBI are discussed.

Results

The literature search identified 8793 records through electronic database searches and other sources. After removing duplicates, 4020 records were screened for inclusion. On the basis of title and abstract, 3995 records were excluded. The remaining 25 records were included for full-text screening. A further 15 articles were excluded at the full-text stage, and the reasons for exclusion are shown in Figure 1. There were no disagreements between the two reviewers, hence the third-party reviewer was not required. The literature search identified ten studies that estimated cost-effectiveness of IGRAs compared with TST in diagnosing LTBI in our three populations of interest, and included a decision analytical model.

Characteristics of included studies

The characteristics of these models are summarised in Table 1. Four[9-12] economic evaluations were conducted in Japan, three[13-15] in the USA and two[3, 4] in the UK, and one[16] in South Africa. Three studies[9-11] compared QFT-GIT with TST, two[13, 14] compared IGRA with TST, but have not suggested the type of IGRA being used, one[15] compared QFT-G with TST and four[3, 4, 12, 16] compared various testing strategies (TST, QFT, QFT-GIT, T-SPOT.TB, positive TST followed by QFT and positive TST followed by T-SPOT.TB). A clinical guideline which included an economic model[3] included a no testing strategy. Two[10, 16] economic evaluation were conducted in a child

population, six[9, 11-15] in the immunocompromised population and two[3, 4] in the recently arrived population.

Six[3, 9-13] studies reported results in terms of cost per quality-adjusted life year (QALY) gained, three studies[4, 15, 16] reported their results in terms of cost per life years saved (LYS), cost per false negative cases of LTBI avoided, cost per TB deaths avoided, cost per reactivation TB avoided or cost per TB avoided and one study[14] was based on number needed to screen to prevent one case of TB. From the base case results reported, IGRAs tended to be less costly and more effective than other strategies (e.g. TST) in identifying LTBI in these high-risk populations

All of the studies included a decision analytical model. The health states included in the models represented those that people would experience while being screened for LTBI. In the model with children, the health states included healthy, LTBI, TB and dead. There was some variation in the health states for the immunocompromised population, due to differences in underlying disease. In the models with recently arrived people, the health states included test results, treatment for LTBI and treatment for TB. One[4] of the model structures was illegible in this population.

Model time horizons ranged from one year to lifetime. In the models with children, the time horizon was lifetime (up to 80-years) with one-year cycle lengths. In the models with immunocompromised cohorts, the time horizons ranged from one-year to lifetime, with three-month or one-year cycle lengths and in the recently arrived cohort, the time horizons ranged from 15-years to 20-years, with annual cycle lengths. Authors suggested that their time horizons were long enough to measure the costs and benefits of these diagnostic strategies. All studies clearly stated and justified their time horizon, cycle lengths and discount rates, where appropriate.

Resource use and costs depended on the perspective taken. All studies clearly stated the perspective/viewpoint of their analyses. Six studies[3, 4, 12, 14, 15] conducted their analyses from the UK NHS or other national health payer perspective, and the remaining four studies[9-11, 13] conducted their analyses from the societal perspective. The six models that presented results from a health payer perspective included direct costs related to the health service (cost of diagnostic tests, chest x-ray and sputum examinations, treatment for LTBI/ TB and treatment for INH-induced hepatotoxicity). From the four models that presented results based on the societal perspective, three models[9-11] have not included any indirect costs.

Due to the uncertainty around model input parameters and assumptions made in the models, all authors conducted sensitivity analyses. Five studies[3, 4, 14-16] conducted deterministic (one- and two-way) sensitivity analyses alone. The remaining studies[9-13] conducted both deterministic and

PSAs. Sensitivity analyses were conducted around changing the prevalence of LTBI in these populations, test accuracies of diagnostic tests, cost of IGRAs, return rates for TST and varying the progression rate from LTBI to TB.

Quality assessment of the modelling methods

Structure

The structure of the models were generally of good quality. Studies clearly stated decision problems and objectives of the model, perspective of the analysis and presented model structures which represented the clinical pathway people would follow while being screened for LTBI. However, there were structural concerns identified; three studies[9-11] have undertaken their analyses from the societal perspective, but have not included indirect costs (e.g. productivity loss) in the analyses. In general, studies stated the location of the analyses, but not their setting, and this may impact on the generalisability of results. Clear, illustrative model structures were presented in majority of the studies except in the Pareek et al. study, where the illustrative structure was illegible.

All authors justified their choice of model structure, which represented the coherent theory of LTBI disease and its treatment. Six studies[9-13, 16] used decision tree structures with Markov nodes for their analyses, three studies[3, 4, 15] used decision tree structures alone and one study[14] used a Markov model alone. The guideline[3] which comprised of an economic evaluation included a proportion of people returning to have their TST result read. One study[15] included a proportion of people with indeterminate test results on an IGRA, and assumed that people received a second IGRA immediately, but this was not shown in the illustrative structure. All studies included chest x-ray and/or sputum examination to confirm initial active symptomatic TB. All studies included cost of treatment for LTBI/TB. As a result of adhering to treatment, all studies included a proportion of people developing Isoniazid (INH)-induced hepatotoxicity. Other adverse events were not considered. In the Markov models, similar health states were used to simulate the natural history of LTBI over time.

Data

Methods used to identify information to populate the models were satisfactory. Studies[3, 4, 9, 10, 12-16] conducted literature reviews, but have not specified the aim of the review. All [3, 4, 9-16] studies provided references for their model inputs, but were not clear on the choices between data sources or the quality of information used in the models. This might have been a result of a paucity of information in the literature.

Most models[9-16] used published sources to obtain or derive an estimate of the prevalence of LTBI, but some studies[9, 11, 13, 15] have not elaborated on what the prevalence represents (e.g. prevalence of LTBI in contact tracing, prevalence of LTBI based on occasional screening in the population of

interest or prevalence of LTBI that would develop to active TB). Additionally, studies[11-13, 15, 16] using multiple sources were not transparent on the methods to derive the prevalence of LTBI. Test performance for TST and IGRAs were required for the models. Most studies[11-13, 15, 16] conducted literature reviews, and have elaborated on the methods to derive an estimate of sensitivity and specificity. Methods included calculating an estimate based on an average of sensitivity/specificity obtained from the literature and obtaining estimates from meta-analyses. All costs required for the models have been referenced, and where applicable, inflated using the appropriate indices. Authors clearly stated the unit costs used in the models, but some[9-11, 13] authors have not elaborated on the resource used to estimate the unit costs, especially for the treatment of TB. The perspective of the analyses was stated, but in some studies[9-11], the costs did not reflect the viewpoint of the analyses. All authors[4, 9-14, 16], where necessary, discounted costs and benefits using the appropriate rates. Where results were reported in terms of QALYs, authors[3, 9-13] provided references used to obtain the utility weights, but have not elaborated on if the source of utility information was relevant to their population.

Uncertainty and assumptions

Uncertainty is unavoidable in economic modelling. Briggs and Gray 1999[17] and Philips et al.[8] have suggested methods to handle uncertainty. All models have undertaken univariate and multivariate sensitivity analysis on key model input parameters. Four studies[9-12] have also undertaken probabilistic sensitivity analysis (PSA) for joint parameter uncertainty. In order to have a workable model structure, most studies clearly stated their simplifying assumptions, except the model developed by Kowada 2014; these assumptions were unclear. In general, assumptions outlined appeared to be feasible, but strong in some studies[3, 4, 9]. In the NICE study,[3] authors assumed that people adhered to treatment of LTBI/TB, and it would not lead to any adverse events. Pareek et al.[4] assumed that testing with an IGRA would not lead to an indeterminate result. Kowada 2010[9] assumed that the chest x-ray is 100% sensitive and specific for diagnosing TB.

Summary of the general approaches to modelling LTBI

Children

Kowada 2012[10]

Kowada 2012 estimated the cost-effectiveness of QFT-GIT compared with TST or chest x-ray for the diagnosis of LTBI in children, using a decision tree structure with Markov nodes. The model started with a hypothetical cohort of children receiving one of three diagnostic strategies and continued with them occupying the LTBI/initial TB or no LTBI health state, characterised by the prevalence of the disease. On positive results, children received a chest x-ray to confirm TB. Children who received a negative result on the chest x-ray were treated for LTBI. Estimates of sensitivity and specificity of tests were obtained from a meta-analysis of developed-country studies. The analysis was conducted

from the societal perspective and base-case results were expressed as an incremental cost-effectiveness ratio (ICER) based on the outcome of cost per QALY gained. Kowada conducted one- and two-way sensitivity analyses and PSA. The base case results demonstrated that QFT-GIT alone was less costly and more effective than TST alone.

Mandalakas 2013[16]

Mandalakas and colleagues used a decision tree structure with Markov nodes to model young household contacts with an index case. The model started with children (< 5 years) who received one of five diagnostic strategies (no test, TST alone, IGRA alone, TST positive followed by IGRA and TST negative followed by IGRA). Children with positive test results were eligible for treatment for LTBI. Children entered the model at the LTBI health state, and could progress to no infection, initial infection, subsequent infection due to future exposures, pulmonary TB, disseminated TB, TB death or death from other causes. The analysis was conducted from the third-party payer and societal perspectives, and the main results were reported in terms of cost per life-year saved (LYS). Base case results indicated that for 0-2 year olds, the no testing strategy was the dominant strategy whilst for 3-5 year olds, an IGRA following a negative TST was the most effective strategy but not cost-effective compared to no testing.

Immunocompromised

Kowada 2010[9]

Kowada used a decision tree structure with Markov nodes to assess cost-effectiveness of QFT-GIT versus TST in people with rheumatoid arthritis, over a lifetime horizon, starting with a cohort aged 40 years. People with positive/negative results on the TST or positive QFT-GIT received a chest x-ray to diagnose TB, which was assumed to be 100% sensitive and specific. The author provided no comment/discussion on the sources of prevalence of LTBI in this population. Information on the sensitivity and specificity were obtained from a meta-analysis. The primary outcome measure of effectiveness was QALYs gained. The analysis was conducted from the societal perspective and results presented as cost per QALY gained. Kowada conducted one-way and two-way sensitivity analyses and PSA, but the distributions used were not presented. QFT-GIT alone was found to be the most cost-effective strategy, and the base-case results were robust to changes in model input parameters. Kowada suggested that results from the PSA showed that IGRA was the preferred option with 100% probability of being cost-effective compared to TST at a willingness-to-pay of US\$50,000 per QALY.

Kowada 2013[11]

Kowada used a decision tree structure with Markov nodes to assess QFT-GIT, TST or chest x-ray in people being screened before haemodialysis, over a lifetime horizon. People with positive results on

TST/QFT-GIT received a chest x-ray to detect TB, and were treated accordingly for TB/LTBI. The author assumed that chest x-ray was 100% sensitive and specific. The author conducted a review of the literature, but it was unclear on how the accuracy of the tests were derived. The primary outcome measure of effectiveness was QALYs gained, however, the author has not elaborated on the descriptive tools used to value these health states. The analysis was conducted from the societal perspective and results presented in terms of costs per QALY gained. Kowada conducted one- and two-way sensitivity analyses and PSA, but the distributions and the cost-effectiveness acceptability curve were not presented. The author demonstrated that QFT-GIT alone was the most cost-effective strategy for the diagnosis of LTBI.

Kowada 2014[12]

Kowada used a decision tree structure with Markov nodes and estimated the cost-effectiveness of IGRAs versus TST in HIV positive pregnant women in low incidence of TB countries. The model simulated the pathway of four cohorts (BCG-vaccinate during pregnancy, non-BCG vaccinated during pregnancy, BCG vaccinated postpartum period, and non-BCG vaccinated postpartum period), separately, and the cost-effectiveness was estimated over a thirty-year time horizon. The starting point of the model was women aged 20 years who received one of five (TST alone, QFT-G alone, T-SPOT alone, TST positive followed by QFT or TST positive followed by T-SPOT.TB) testing strategies. TST was considered positive if the induration was $\geq 5\text{mm}$ and $\geq 10\text{mm}$ in those who were non-BCG vaccinated and BCG-vaccinated, respectively. Women with positive TST, QFT-G or T-SPOT.TB strategies received a chest x-ray to diagnose TB. In the combination strategies, women who received a positive TST result then received QFT-G or T-SPOT.TB, and if positive, received a chest x-ray to diagnose TB. The analysis was conducted from the public health payer perspective and results presented in terms of cost per QALYs gained. Kowada conducted PSA, and one- and two-way sensitivity analyses. Base-case results showed that positive TST followed by QFT-G was the most cost-effective strategy for occasional screening of women who were non-BCG vaccinated during pregnancy. Results from the PSA showed that the TST followed by QFT-G strategy was the preferred option with 100% probability of being cost-effective at all willingness-to-pay values considered. The results from the sensitivity analyses showed that the base case results were sensitive to changes in the sensitivity of T-SPOT.TB, and the sensitivity of QFT-G in non-BCG vaccinated women.

Laskin et al., 2013[13]

Laskin and colleagues used a decision tree structure with Markov nodes to determine the most cost-effective screening strategy in children with new-onset idiopathic nephrotic syndrome. The model starts with children receiving TST/IGRA, and if positive children were eligible for LTBI treatment. The authors assumed that effective LTBI treatment provided long-term protection against LTBI/TB. The analyses were conducted from the societal perspective and included indirect costs on travel time

and loss of productivity. Base-case results showed that the no screen strategy was less costly and more effective than other strategies. Results from this study should be interpreted with caution because the discounted and undiscounted costs were similar despite the cost-effectiveness being measured over a lifetime horizon. Results were sensitive to changes in the prevalence of LTBI in this population, with the questionnaire followed by IGRA screening strategy to be the most cost-effective strategy at a prevalence of >4.9%. Results from the PSA showed that at a prevalence of 1.1%, no screening compared with IGRA was the preferred screening option, but the authors have not stated at what willingness-to-pay value.

Linac et al., 2011[14]

Linac and colleagues constructed a decision tree structure with Markov nodes and estimated the cost-effectiveness of using TST compared with IGRAs in various populations. The model began with a cohort receiving one of three diagnostic strategies (TST alone, IGRA alone or no screening), and continued with people characterised by their disease status (LTBI/no LTBI). People with positive IGRA/TST received treatment for LTBI. Costs related to a health service perspective were obtained from published sources. Utility values estimated were based on the SF-36 and EQ-5D descriptive systems. The primary outcome was cost per QALY gained over a lifetime horizon. Base-case results showed that in the HIV-infected cohort, screening with IGRA alone was marginally more costly and effective than the no screening option with an ICER of \$12,800. People who were on immunosuppressive medication, the reported ICER for TST compared with no screening was \$129,000. Sensitivity analysis results showed that increasing the mean age to 65 years, TST remained cost-effective in people living with HIV. Base-case results were sensitive to changes to the estimates on health-related quality of life for people who received treatment for TB. Screening with TST or IGRA resulted in ICERs greater than \$100,000 for people with diabetes or end-stage renal disease.

Swaminath et al., 2013[15]

Swaminath and colleagues used a decision tree structure and compared QFT-G with TST in people with inflammatory bowel disease. The model simulated people with moderate to severe active Crohn's disease being treated with immunosuppressive medication. On positive results, people received treatment for LTBI. Swaminath et al. suggested that people with indeterminate results on the QFT-G would immediately receive a second QFT-G test. However, this pathway was not shown in the decision tree structure. The prevalence of LTBI in this population was obtained from World Health Organization. Sensitivity and specificity of tests were derived based on information obtained from published sources, and not based on a systematic review. The analysis was conducted from the health payer perspective and results presented in terms of costs per false negative case avoided, TB reactivations and deaths avoided. The authors conducted one-way sensitivity analyses around input

parameters. Swaminath and colleagues concluded that QFT-G was less costly and more effective than the TST in this population.

Recently arrived

Pareek et al., 2013[4]

Pareek and colleagues used a decision tree structure and compared T-SPOT.TB, QFT-GIT, TST positive plus confirmatory T-SPOT.TB or TST positive plus confirmatory QFT-GIT for screening immigrants for LTBI. The illustrative model structure presented in the supplementary appendix was illegible. The authors suggested that immigrants who were symptomatic at screening or had a positive IGRA/TST result were referred for a chest x-ray and further clinically assessed. Immigrants with positive IGRA and/or positive TST result and normal chest x-ray without any symptoms of suggesting TB were considered to have LTBI. For a positive TST, cut-offs of ≥ 6 mm and ≥ 15 mm were used for BCG-unvaccinated and BCG-vaccinated participants, respectively. Additionally, the authors used a non-stratified cut-off of ≥ 10 mm to suggest a positive TST. Information required to populate the model were obtained from an observational study undertaken by the authors and from published sources. Study participants included recently arrived (≤ 5 years) immigrants to the UK aged ≥ 16 years (with symptoms of TB) or from a country with a high incidence of TB. Information on the prevalence of LTBI was derived from immigrants aged ≤ 35 years that had been tested with the three screening tests. The analysis was undertaken from the UK NHS perspective in a primary care setting. The outcome measures included in the analyses were the number of cases of TB avoided and the number of LTBI cases needed to be treated to prevent one case of TB over a 20-year time horizon. Base-case results showed that the screening strategy no port-of-entry chest x-ray and screening with QFT-GIT was cost-effective with an ICER of approximately £21,600 per case of TB avoided and the no port-of-entry chest x-ray and screening with one-step QFT-GIT was cost-effective, with an ICER of approximately £31,900 per case of TB avoided. These strategies were cost-effective in immigrants whose country of origin had an incidence of TB of 250 per 100,000 and 150 per 100,000, respectively. Sensitivity analyses results showed that increasing the prevalence and progression rate from LTBI to TB increased the cost-effectiveness of using the QFT-GIT. Reducing specificity resulted in the T-SPOT.TB becoming the most cost-effective strategy. Reducing the proportion of immigrants accepting and adhering to LTBI treatment lead to higher cost-effectiveness estimates.

CG117[3]

The authors of CG117 used a decision tree structure and compared four testing strategies: TST, IGRA, TST followed by IGRA for people with positive results and no test, in immigrants from countries with a high incidence of TB. In the TST/IGRA strategies, people who received a positive result were treated for LTBI. Conversely, people with negative results, a proportion were given BCG-vaccination. In the combination strategy, people who tested positive on the TST received a QFT test.

Immigrants with positive QFT results were treated for LTBI, and those with negative results, a proportion were given a BCG vaccination. The end-point of the model is people developing TB having received a BCG vaccination or treatment for LTBI. Sensitivity of tests were derived based on values obtained from two publications. Costs included in the model were those related to the UK NHS and Personal Social Services (PSS), and were presented in UK pounds sterling in 2008/09 prices. Costs obtained from published sources were inflated using the Hospital and Community Health Services Pay and Price Index. The results showed that positive TST followed by IGRA, and IGRA alone strategies were associated with ICERs below £30,000 per QALY compared with no testing strategy. Results from the sensitivity analyses showed that varying the cost of an IGRA (£50 to £60) changed the direction of the cost-effectiveness results.

Table 1. Summary characteristics of the models used to compare IGRAs and TST in identifying LTBI in children, immunocompromised and recently arrived immigrants

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting)	Intervention	Outcome(s)	Model type	Health states	Results (base case and sensitivity analysis)
Children							
Kowada 2012,[10] Japan	To assess the cost-effectiveness of school-based TB screening using QFT-GIT versus the TST and CXR	Cost-effectiveness analysis, societal perspective, setting not reported	QFT-GIT	Cost per QALY	Decision tree structure to model the short term events followed by a Markov modelling structure	Healthy, LTBI, TB and dead	QFT-GIT was less costly and more effective than TST strategy
Mandalakas 2013,[16] South Africa	To estimate the health and economic outcomes of five TB screening strategies	Cost-effectiveness analysis, third party payer and societal perspectives	IGRA (QFT, T-SPOT.TB)	Cost per LYS	Decision tree structure to model the short term events followed by a Markov modelling structure	LTBI health state, and could progress to no infection, initial infection, subsequent infection due to future exposures, pulmonary TB, disseminated TB, TB death and death from other causes	In the 0-2 cohort, no testing strategy dominated other strategies In the 0-3 cohort, the TST –ve followed by IGRA was the most - effective with a reported ICER of approximately US\$233 000 per LYS versus no testing
Immunocompromised							
Kowada 2010,[9] Japan	To assess the cost-effectiveness of QFT-GIT versus TST for TB screening of RA patients prior to	Cost-effectiveness analysis, societal perspective, setting not reported	QFT-GIT	Cost per QALY	Decision tree model with Markov nodes	No LTBI, LTBI, TB and death	QFT-GIT was less costly and more effective than TST strategy. At society's WTP per QALY, the probability of QFT-

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting)	Intervention	Outcome(s)	Model type	Health states	Results (base case and sensitivity analysis)
	initiation of TNF α antagonist therapy						GIT testing strategy has a 100% probability of being cost-effective compared to the TST
Kowada 2013,[11] Japan	To assess the cost-effectiveness of QFT-GIT compared with the TST and the CXR for TB screening of haemodialysis	Cost-effectiveness, societal perspective, setting not reported	QFT-GIT	Cost per QALY	Decision tree model with Markov nodes	Maintenance dialysis with no disorder, maintenance dialysis with LTBI, maintenance dialysis with TB and death	QFT-GIT was dominant compared to TST testing strategy. Results from the SA showed that the base-case results were sensitive to the BCG vaccination rate. At society's WTP per QALY, the probability of QFT-GIT testing strategy has a 100% probability of being cost-effective compared to the TST
Kowada 2014,[12] Japan	To assess the cost effectiveness for TB screening of high risk HIV positive pregnant women by using IGRAs compared to the TST in low	Cost-effectiveness analysis, health service perspective, low incidence of TB country, but setting not reported	1) TST alone, 2) QFT alone, 3) T-SPOT.TB, 4) TST followed by QFT and 5) TST followed by T-SPOT.TB	Cost per QALY	Decision tree model with Markov nodes	Non-LTBI and non-TB, LTBI, non MDR-TB, MDR-TB and dead	Base-case results showed that the T-SPOT.TB is less costly and was more effective compared to other strategies. SA showed that the cost-effectiveness

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting)	Intervention	Outcome(s)	Model type	Health states	Results (base case and sensitivity analysis)
	incidence countries						was sensitive to the sensitivity of T-SPOT.TB, the sensitivity of QFT, specificity of T-SPOT.TB and the specificity of QFT in close contacts
Laskin 2013,[13] USA	To determine the most cost-effective LTBI screening strategy before long-term steroid therapy in a child with new-onset idiopathic nephrotic syndrome	Cost-effectiveness analysis, societal perspective, setting not reported	IGRAs	Cost per QALY	Decision tree structure to model the short term events followed by a Markov modelling structure	Well, LTBI, TB, nephrotic relapse and dead) for the longer-term events	Base-case results showed that IGRA was less costly and produced moderately more QALYs compared to universal TST
Linan 2011,[14] USA	To estimate the cost-effectiveness of LTBI screening using the TST and IGRAs	Cost-effectiveness analysis, health service, setting not reported	IGRAs and TST	Number needed to screen to prevent one case of active TB, life expectancy, quality-adjusted life expectancy	Markov model	LTBI with INH, LTBI no INH, INH related hepatitis, < six months INH, 6-8 months INH, nine months INH, Active TB, post active TB and death	Base-case results showed that people who are taking immunosuppressive medications, TST screen was not likely to be cost-effective to the no screening strategy. Similar results were reported for people with ESRD
Swaminath 2013,[15] USA	To compare the performance of TST and QFT-G	Cost-effectiveness, health care payer,	QFT-G	Cost per false negative cases of LTBI avoided, cost	Decision tree model	True positive, true negative, false positive,	Base-case results showed that QFT-G dominated the TST

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting)	Intervention	Outcome(s)	Model type	Health states	Results (base case and sensitivity analysis)
	got screening LTBI among immunosuppressed IBD patients based on prevalence, mortality risk reactivation TB, and costs	setting not reported		per TB deaths avoided, cost per reactivation TB avoided (this can be derived from the information provided)		false negative, hepatitis, survive/death hepatitis	strategy. Additionally, the use of QFT-G would avoid 30 false-negative cases, 4.92 TB reactivations and 1.4 deaths compared with TST
Recently arrived							
CG117,[3] UK	To compare the cost and effects of four strategies of testing for people suspected with LTBI in England and Wales	Cost-effectiveness analysis, NHS and Personal Social Services (PSS)	1) TST, 2) IGRA, 3) TST followed by IGRA for people with positive TST and 4) no test (to inform and advise only)	Cost per QALY	Decision tree model	Test results, treatment for LTBI, treatment for TB	Results showed that TST +ve followed by IGRA and IGRA testing strategies were associated with ICERs below £30,000 per QALY compared with no testing. The results from the sensitivity analyses showed that varying the cost of an IGRA (£50 to £60) changes the direction of the cost-effectiveness results
Pareek 2013,[4] UK	To assess the cost-effectiveness of LTBI screening using different	Cost-effectiveness analysis, NHS,	1) T-SPOT.TB alone, 2) QFT-GIT alone, 3) TST plus	Cost per case of active TB avoided	Decision tree model	The illustrative modelling structure was presented in a	Results showed that screening of newly arrived immigrants from countries of

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting)	Intervention	Outcome(s)	Model type	Health states	Results (base case and sensitivity analysis)
	screening modalities at different incidence thresholds in a primary care setting, with and without CXR screening on arrival at port of entry	primary care setting	confirmatory T-SPOT.TB (if TST positive), and 4) TST plus confirmatory QFT-GIT (if TST positive)			supplementary web-appendix, but unfortunately, these structures were illegible	origin with moderate (not defined) TB incidence is likely to be cost-effective by the use of one-step IGRA testing compared to other screening strategies
BCG, Bacillus Calmette–Guérin; CXR, Chest x-ray, ESRD, End-stage renal disease; HIV, Human immunodeficiency virus; IGRA, Interferon-gamma release assay; INH, Isoniazid; LTBI, Latent tuberculosis infection; NHS, National Health Service; PSS, Personal Social Services; QALY, Quality adjusted life-years, QFT-G, QuantiFERON-Gold; QFT-GIT, QuantiFERON Gold-In-Tube; RA, Rheumatoid arthritis; SA, Sensitivity analysis; TB, Tuberculosis; TST, Tuberculin skin test; WTP, Willingness-to-pay							

Discussion

The evidence-base here offers some insight on the model structures which have been used to assess the cost-effectiveness of IGRAs compared with TST for the diagnosis of LTBI in high risk populations. We identified ten model-based economic evaluations, which mainly used decision tree structures with Markov nodes to simulate people being tested for LTBI, with majority of these models in the immunocompromised population. These results highlight that the evidence available for the other two populations is sparse.

We appraised models against frameworks on best practice for reporting an economic evaluation and economic modelling. In general, all models performed well in terms of reporting quality, and add to existing cost-effectiveness literature, but are subject to limitations. First, majority of the studies indicated the location of the study but have not stated the setting of the analysis and this may limit the generalisability of the results. Second, a majority of the studies used QALYs as their outcome measure and have referenced the source of their utility values, but have not provided commentary on the descriptive tools used to value these health states. When obtaining health state utility values from the literature it is important to consider the methods/tools used to generate these values and their relevance to the population to which they are going to be applied. Third, the perspective of the analysis was stated in all studies, however, some of the resource use and costs reported did not reflect studies' viewpoint. Fourth, studies were transparent about the methods to identify information to populate the models, but it was unclear on any assessment used on the quality of the information. Finally, all models have explored uncertainty around key model input parameters, but no attempt was made to explore methodological and structural uncertainty, or generalisability. Other concerns relate to the derivation of prevalence, test accuracy and transition probabilities; most studies have not elaborated on these statistical/pre-model analyses.

We identified one study[18] that provided a review of the cost-utility studies available up to 2014 on the use of IGRAs compared with TST for the diagnosis of LTBI. The results of all studies included in the review were presented in terms of 'cost per QALY.' Though useful, other forms of economic evaluation studies, more specifically cost-effectiveness analyses, may provide relevant information on prevalence of LTBI or resource use and costs, for example. Studies presenting results in terms of cost per QALY alone may suggest that QALY is likely to capture all the benefits of identifying people with LTBI. The authors concluded that screening with TST for LTBI in an HIV population is cost-effective, and screening with an IGRA within an immigrant population is moderately cost-effective. We identified a second systematic review[19] which focussed on the key model input parameters and the methodological differences in studies that assessed the cost-effectiveness of preventative treatment for TB in high risk populations, and not of LTBI diagnosis. In addition, economic models used to

assess the cost-effectiveness of strategies for identifying LTBI in a child population were not considered. These authors have outlined the limitations identified in the studies, but have not undertaken a formal quality appraisal of the economic models against the CHEERS [7] or Phillips et al. [8] checklists. Our current review identifies and appraises the economic models that have been used to inform on the diagnosis of LTBI in high risk populations.

For future advances in using economic models to aid in the decision making process for the most cost-effective strategy for identifying LTBI in high risk populations, analysts should consider the information required on prevalence of LTBI in these populations, diagnostic accuracy of test(s), and the illustrative model structure. Based on the studies identified, the methods used to derive prevalence, and sensitivity and specificity may not have provided the best estimates, and in some cases might have under/overestimated these input values. As no 'gold standard' test exists for LTBI diagnosis, estimates can be derived from meta-analysing studies that followed-up a cohort of people to the incidence of TB following testing with TST and/or IGRA. For this instance, best estimates would be based on the development of TB as a 'reference standard' for diagnosing LTBI. This method, as opposed to using exposure to TB and test agreement studies alone, may be more appropriate for use in decision analytical models. However, other points ought to be considered: serial testing, BCG vaccination history and anti-tuberculous treatment on testing positive, all of which can have an impact on evaluating test performance. Further discussion of these points are beyond the scope of this paper, but will be addressed in a subsequent manuscript.

The models available provide insight on the clinical pathway should screening for LTBI be undertaken, and which strategy is likely to be cost-effective in high-risk populations. In future models, it will be important to consider which diagnostic strategy is most likely to be cost effective to identify LTBI that progresses to active TB; and not sensitivity and specificity of diagnostic tests aimed at identifying LTBI in general. Such models would incorporate a decision tree structure and epidemiological model to estimate the cost-effectiveness. These models would also provide useful information on an estimate of the number of people who are treated/untreated for LTBI and further developed TB, and any new cases of LTBI.

We undertook a search of the literature to identify all relevant studies that compared TST and IGRAs for identifying LTBI in these three populations of interest. The main strength of this current review is the comprehensive search, reporting quality assessment and data extraction of the relevant information from these studies. Second, it provides a detailed overview and critique of the health economic models that have been used to estimate the cost-effectiveness of IGRAs compared with TST. In terms of limitations, some studies have not reported/presented information on model structure, how prevalence was derived; hence we could not provide a narrative for these studies.

Conclusion

This review highlights the health economic models available on the cost–effectiveness of diagnosing LTBI in high risk populations. The majority of the models were undertaken in an immunocompromised population, which suggests that there is a paucity of evidence available in a child population and recent arrivals population. In general, all models performed well in terms of defining the decision problem, including the study perspective, outlining the choice of comparators, presenting an illustrative model structure and providing a clear outline of the assumptions.

The evidence shows that the models available are based on identifying LTBI in general, and little is known about the cost-effectiveness of diagnostic tests that identify LTBI that progresses to active TB; which shows that research in this area is static. We propose that future pre-model analyses should consider deriving estimates based on the development of TB as a 'reference standard' for diagnosing LTBI in order to inform an economic model. However, the challenge/practicality is to identify prospective longitudinal studies with adequate sample size and a lengthy follow-up in people at high risk of developing TB.

Disclosure of interest

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf and declare that: no authors have support from any company for the submitted work; no authors have any relationships from any company that might have an interest in the submitted work in the previous three years; their spouses, partners, or children have no financial relationships that may be relevant to the submitted work; and no authors have any nonfinancial interests that may be relevant to the submitted work.

Contributors to authorship

PA, AT and JP designed and supervised the systematic review. The systematic searches were performed by RC with input from PA. PA, JP and RC prepared the manuscript as lead writers.

Funding

This systematic review was funded by the National Institute of Health Research Health Technology Assessment Programme (NIHR HTA) (13/178/01).

Acknowledgements

This project was funded by the by National Institute for Health Research Health Technology Assessment Programme and will be published in full in the Health Technology Assessment journal

series. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Department of Health.

References

1. Kim SH, Lee SO, Park JB, Park IA, Park SJ, Yun SC, Jung JH, Kim YH, Kim SC, Choi SH, Jeong JY, Kim YS, Woo JH, Park SK, Park JS, Han DJ. A prospective longitudinal study evaluating the usefulness of a T-cell-based assay for latent tuberculosis infection in kidney transplant recipients.[Erratum appears in *Am J Transplant*. 2011 Nov;11(11):2541]. *Am J Transplant* 2011; 11(9): 1927-1935.
2. Mahomed H, Hawkrigde T, Verver S, Abrahams D, Geiter L, Hatherill M, Ehrlich R, Hanekom WA, Hussey GD. The tuberculin skin test versus QuantiFERON TB Gold in predicting tuberculosis disease in an adolescent cohort study in South Africa.[Erratum appears in *PLoS One*. 2011;6(6). doi:10.1371/annotation/b371be66-de38-4ed6-b440-d27c9d7e552c]. *PLoS One* 2011; 6(3): e17984.
3. National Collaborating Centre for Chronic Conditions, Centre for Clinical Practice at NICE. Tuberculosis: clinical diagnosis and management of tuberculosis, and measures for its prevention and control. CG117. 2011 [cited; Available from: <http://www.ncbi.nlm.nih.gov/books/NBK97852/>]
4. Pareek M, Bond M, Shorey J, Seneviratne S, Guy M, White P, Lalvani A, Kon OM. Community-based evaluation of immigrant tuberculosis screening using interferon gamma release assays and tuberculin skin testing: observational study and economic analysis. *Thorax* 2013; 68(3): 230-239.
5. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce BR. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value Health* 2003; 6(1): 9-17.
6. HTA - 13/178/01: Accurate diagnosis of latent Tuberculosis in children, in people who are immunocompromised or at risk from immunosuppression, and recent arrivals from countries with a high incidence of Tuberculosis: systematic review and economic evaluation. 2014 [cited; Available from: <http://www.nets.nihr.ac.uk/projects/hta/1317801>]
7. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Eur J Health Econ* 2013; 14(3): 367-372.
8. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacoot N, Glanville J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004; 8(36): iii-iv, ix-xi, 1-158.
9. Kowada A. Cost effectiveness of interferon-gamma release assay for tuberculosis screening of rheumatoid arthritis patients prior to initiation of tumor necrosis factor-alpha antagonist therapy. *Mol Diagn Ther* 2010; 14(6): 367-373.
10. Kowada A. Cost effectiveness of interferon-gamma release assay for school-based tuberculosis screening. *Mol Diagn Ther* 2012; 16(3): 181-190.
11. Kowada A. Cost effectiveness of the interferon- release assay for tuberculosis screening of hemodialysis patients. *Nephrol Dial Transplant* 2013; 28(3): 682-688.

12. Kowada A. Cost effectiveness of interferon- release assay for TB screening of HIV positive pregnant women in low TB incidence countries. *J Infect* 2014; 68(1): 32-42.
13. Laskin BL, Goebel J, Starke JR, Schauer DP, Eckman MH. Cost-effectiveness of latent tuberculosis screening before steroid therapy for idiopathic nephrotic syndrome in children. *Am J Kidney Dis* 2013; 61(1): 22-32.
14. Linas BP, Wong AY, Freedberg KA, Horsburgh CR, Jr. Priorities for screening and treatment of latent tuberculosis infection in the United States. *Am J Respir Crit Care Med* 2011; 184(5): 590-601.
15. Swaminath A, Bhadelia N, Wang YC. Cost-effectiveness of QuantiFERON testing before initiation of biological therapy in inflammatory bowel disease. *Inflamm Bowel Dis* 2013; 19(11): 2444-2449.
16. Mandalakas AM, Hesseling AC, Gie RP, Schaaf HS, Marais BJ, Sinanovic E. Modelling the cost-effectiveness of strategies to prevent tuberculosis in child contacts in a high-burden setting. *Thorax* 2013; 68(3): 247-255.
17. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technol Assess* 1999; 3(2): 1-134.
18. Campbell JR, Sasitharan T, Marra F. A Systematic Review of Studies Evaluating the Cost Utility of Screening High-Risk Populations for Latent Tuberculosis Infection. *Applied health economics and health policy* 2015; 13(4): 325-340.
19. Diel R, Lampenius N, Nienhaus A. Cost Effectiveness of Preventive Treatment for Tuberculosis in Special High-Risk Populations. *Pharmacoeconomics* 2015; 33(8): 783-809.

Appendix

Example search strategy

This search was developed and conducted as part of a wider systematic review that aimed to compare both the clinical effectiveness and cost-effectiveness of screening tests (IGRAs and TST) for LTBI in high risk groups.[1] It was updated in Dec 2014 and June 2015

Ovid MEDLINE(R) 1946 to March Week 2 2014, searched on 21/03/2014

1	(laten* adj3 (tb* or tubercul*)).tw.	2701
2	ltb*.tw.	6939
3	tubercul*.tw.	158617
4	Tuberculosis/	51049
5	Latent Tuberculosis/	866
6	Tuberculosis, Pulmonary/	63874
7	Mycobacterium tuberculosis/	35401
8	1 or 2 or 3 or 4 or 5 or 6 or 7	195420
9	quantiferon*.tw.	819
10	QFT*.tw.	557
11	t spot*.tw.	261
12	exp Enzyme-Linked Immunosorbent Assay/	122317
13	Interferon-gamma Release Tests/	377
14	((interferon* or IFN*) adj3 gamma* adj3 (release* or test* or assay*)).tw.	3856
15	((y-interferon or interferon-y) adj3 (release* or test* or assay*)).tw.	7
16	IGRA*.tw.	448
17	9 or 10 or 11 or 12 or 13 or 14 or 15 or 16	126231
18	8 and 17	3837
19	Latent Tuberculosis/di	576
20	18 or 19	4058
21	Animals/ not Humans/	3812070
22	20 not 21	3477
23	limit 22 to english language	3011

24	limit 23 to ed=20091207-20140321	1285
----	----------------------------------	------