

University of Groningen

Making Informed Decisions

van der Pol, Simon

DOI:
[10.33612/diss.241727138](https://doi.org/10.33612/diss.241727138)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
van der Pol, S. (2022). *Making Informed Decisions: the Value of Testing Strategies in Healthcare*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.
<https://doi.org/10.33612/diss.241727138>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

CHAPTER

5

**Economic Analyses of Respiratory
Tract Infection Diagnostics**

A Systematic Review

Simon van der Pol
Paula Rojas García
Maarten J. Postma
Fernando Antoñanzas Villar
Antoinette D.I. van Asselt

Abstract

Background: Diagnostic testing for respiratory tract infections is a tool to manage the current COVID-19 pandemic, as well as the rising incidence of antimicrobial resistance (AMR). At the same time, new European regulations for market entry of in vitro diagnostics, in the form of the in vitro diagnostic regulation (IVDR), may lead to more clinical evidence supporting health-economic analyses.

Objective: To systematically review the methods used in economic evaluations of applied diagnostic techniques, for all patients seeking care for infectious diseases of the respiratory tract (such as pneumonia, pulmonary tuberculosis, influenza, sinusitis, pharyngitis, sore throats and general respiratory tract infections).

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, articles from three large databases of scientific literature were included (Scopus, Web of Science and PubMed) for the period January 2000 - May 2020.

Results: A total of 70 economic analyses are included, most of which use decision tree modelling for diagnostic testing for respiratory tract infections in the community-care setting. Many studies do not incorporate a generally comparable clinical outcome in their cost-effectiveness analysis: fewer than half the studies (33/70) used generalisable outcomes such as quality-adjusted life years (QALYs). Other papers consider outcomes related to the accuracy of the test or outcomes related to the prescribed treatment. The time horizons of the studies generally are limited.

Conclusion: The methods to economically assess diagnostic tests for respiratory tract infections vary and would benefit from clear recommendations from policy makers on the assessed time horizon and outcomes used.

Introduction

When diagnosing a patient with common symptoms with various possible pathologies, such as respiratory infections, clinicians historically had to either rely on clinical judgement and empirical therapy, or wait for the results of diagnostics performed in specialized laboratories¹²⁶. Recent developments have brought diagnostic tests to the point of care (POC): these novel diagnostics enable clinicians to rapidly and more accurately diagnose patients and to prescribe more appropriate treatment¹²⁷, which in turn improve understanding of the patient's condition and monitoring of the patient's clinical course¹²⁸.

In light of the COVID-19 pandemic, rapid diagnostic tests are regarded as a fundamental instrument in combating the spread of SARS-CoV-2^{129,130} and, consequently, have received a lot of attention. While societies worldwide are vaccinated at an unprecedented rate, rapid COVID-19 tests are expected to play an important role in reopening the economy. Since the start of the pandemic, many diagnostics have been developed and are entering the market¹³¹⁻¹³³.

The pandemic is a major risk to public health and economies worldwide, but infectious disease poses another threat as well: in 2018, the World Health Organisation (WHO) declared antimicrobial resistance (AMR) to be one of the ten greatest threats to public health¹³⁴. In Europe, there are estimated to be over 650,000 infections with resistant bacteria every year, causing over 30,000 attributable deaths¹³⁵. Innovative POC testing may have an important role in combating AMR, as it enables clinicians to prescribe antibiotics more accurately^{136,137}. Antibiotic prescriptions related to respiratory infections are especially relevant because this type of infection represents a third of total visits to primary care centres¹³⁸ and generate difficulties for medical professionals when diagnosing, as they tend to overestimate the proportion of patients presenting with bacterial infections and, consequently, overprescribe antibiotics¹³⁹.

In the field of pharmaceuticals, health technology assessment (HTA) plays an important role in assessing the value for money of new drugs¹⁴⁰, but in the case of diagnostic technologies, this is not as established. A complicating factor here is that, unlike pharmaceuticals, which directly influence a patient's health status, the impact of diagnostic technologies is indirect and only takes effect when diagnostic results change downstream clinical interventions¹⁴¹. Until now, the assessment of new diagnostic techniques often focused on technical capabilities, such as the test's sensitivity and specificity. However, starting in 2022, the European "in vitro diagnostic regulation" (IVDR) law will come into effect, making it mandatory for companies to prove the clinical effectiveness of new diagnostics before they enter the market with aligned data evidencing this¹⁴². These data will enable policy makers, payers and healthcare providers to better estimate the added clinical value of novel diagnostics and can be incorporated in HTA.

Considering the public health relevance of diagnostics of respiratory infections and the policy changes that may increase the focus on the HTA of diagnostics, we systematically reviewed the methods used in economic evaluations of applied diagnostic techniques, for all patients seeking care for infectious diseases of the respiratory tract (such as pneumonia, pulmonary tuberculosis (TB), influenza, bronchitis, bronchiolitis, sinusitis, pharyngitis, sore throats, group A beta-haemolytic streptococcal infections (GABHS) and general respiratory tract infections). Specifically, we report on the types of economic models used to assess current practices for implementing development diagnostic technologies so that the results generated would facilitate the identification of areas for improvement in economic evaluation of diagnostics. Finally, considering the evidence of increasing societal

costs of AMR¹³⁵ we evaluate how authors have modelled the influence of AMR and under which circumstances diagnostic tools would help reduce antibiotic prescription.

Methods

Search strategy

We conducted a systematic review of articles contained in PubMed, Scopus and Web of Science. The search syntax was constructed to include economic evaluations of diagnostic strategies of infectious diseases, see supplementary table 5.1 for the specific search syntax. The results were not limited to certain countries but, with the purpose of reflecting recent clinical practice, we only included articles published between January 2000 and May 2020.

Eligibility criteria

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines¹⁴³ were used for this study. Economic evaluations of diagnostics strategies for all respiratory tract infections were considered for inclusion. Articles meeting the eligibility criteria performed an economic analysis comparing both costs and effects, including Cost-Effectiveness Analyses (CEA), Cost-Utility Analyses (CUA) and Cost-Minimization Analyses (CMA) that incorporated clinical effects. Studies assessing only the test characteristics from a technical (laboratory) point of view were not considered in this review. Patient-relevant outcomes had to be included, such as Quality-Adjusted Life Years (QALYs), Disability-Adjusted Life Years (DALYs), life years gained or the proportion of correct diagnoses. An inclusion criterium was also utilised that at least two diagnostic strategies for respiratory tract infections were compared. Diagnostic strategies were defined as: “identifying the most likely cause of, and optionally optimal treatment for, a previously undetected disease in a clinically suspect patient who is seeking care”²⁶. Population screening, disease monitoring or genotyping of genetic material from patients were therefore explicitly not considered to be diagnostic strategies. Other exclusion criteria were studies focusing on animals, review articles, study protocols, comments on articles or individual case reports, and languages other than English, Spanish, Dutch, German or French.

Study selection

Phase one consisted of two reviewers independently screening the titles and abstracts (PRG & SvdP). In case of not reaching an agreement, a third person (ADIVa) was asked. The full-text screening phase was performed by the same two reviewers, applying the same criteria. This phase was also used to separate the diagnostic and screening strategies, as this distinction was often not clear from the abstract, and to assess whether the article concerned respiratory tract infections.

Data extraction

The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist⁴⁹, was used as a basis to create a standardised digital (Google) form to extract the relevant data from the articles. Various items relevant to diagnostics were added to this extraction tool. See supplementary file 5.1 for an overview of the items included. The data extraction was divided between three reviewers (PRG, SvdP and ADIVa); 10% of extractions was duplicated to check for consistency between authors. Furthermore, the reviewers reached consensus with continuous discussions on the extracted data and repeating the extraction if necessary. To analyse the reporting quality of the studies, a score was calculated based on the presence of the items as reported in CHEERS⁴⁹.

In the general and introduction parts of the extraction tool, the research question, specific

disease area and pathogens considered were included. In the methodology, we emphasised the type of model developed and its characteristics in terms of perspective, time horizon, setting, population included and incorporation of uncertainty analysis in parameter values (stochastic or deterministic). A section was included to assess whether the model included AMR and, if so, how. In the results sections, we paid attention to the incremental costs and outcomes and techniques for reporting uncertainty in the model. Finally, for the discussion, the focus was on the main findings, limitations, specific limitations in the assessment of diagnostics, and advantages/disadvantages of the modelling technique discussed by the authors.

Data analysis

The data extracted from the articles were analysed using R 3.6.3¹⁴⁴, categorising the data by the considered pathogens (influenza, streptococcus) and the type of models (decision tree, Markov). For data transformation and table creation, the packages Dplyr 1.0.0¹⁴⁵ and gt 0.2.2¹⁴⁶ were used. The code was made available on GitHub.

Results

Figure 5.1. shows the included and excluded studies in a PRISMA flow diagram. 70 papers were included in this review. Most studies were a CEA or CUA, comparing the standard of care, mostly consisting of empirical therapy, clinical judgement or traditional diagnostics (e.g. cultures or microscopy), to the use of rapid diagnostic tests. The most common

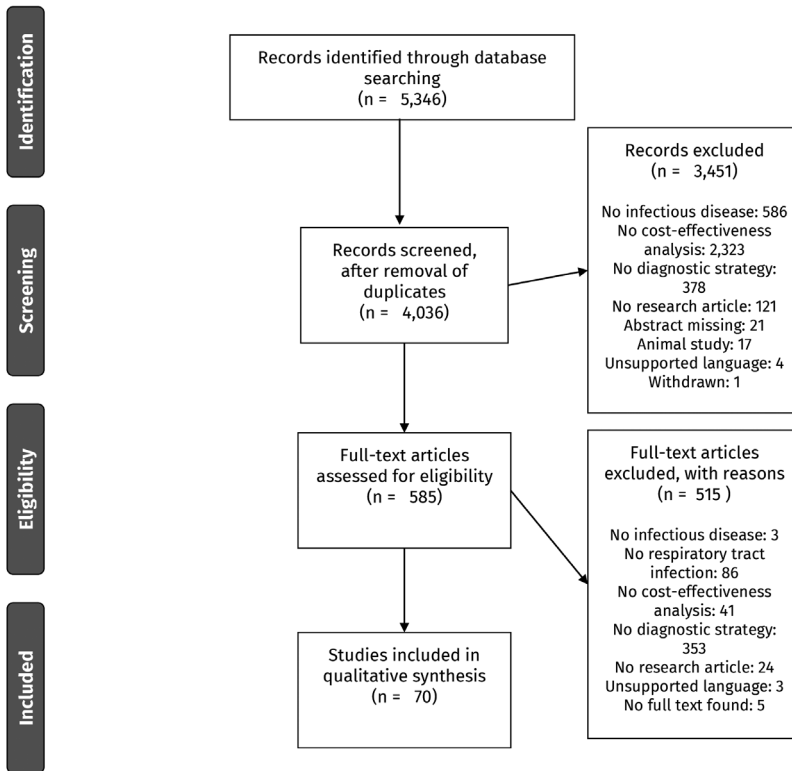


Figure 5.1. PRISMA flow diagram of inclusion and exclusion



rapid tests included are: Xpert for TB, influenza-specific POC tests, C-reactive protein (CRP) POC tests and procalcitonin (PCT) tests. Other diagnostics included are Polymerase Chain Reaction (PCR), microscopy, X-ray and clinical scoring algorithms. No tests for coronaviruses were included in this review. Most studies (46) included a decision tree model, two included a Markov model in addition to a decision tree, 12 were trial-based analyses, seven were categorised as a dynamic model, and three papers were categorised as 'other'. Table 5.1. provides an overview of the settings and tests assessed in the studies, while table 5.2. provides an overview of the methods used. Key characteristics and results of all studies can be found in supplementary tables 5.1 and 5.2.

The reporting quality, in the form of a CHEERS score, can also be found in the supplementary table 5.2. Most items were included by most studies. All articles reported comparators, model assumptions and incremental costs and outcomes of the model. Background and objectives were items included in all Markov model but only included in a 90% of trial-based analyses and 93% of decision tree models. The target population was included in all trial-based analyses, decision trees and Markov models but in 86% of all dynamic models.

Table 5.1. Overview of included studies

	Trial-based analysis (n = 12)¹⁴⁷⁻¹⁵⁸	Decision tree model (n = 46)^{161-182,184-193,193-207}	Markov model (n = 2)^{209,210}	Dynamic model (n = 7)²¹²⁻²¹⁸	Other (n = 3)²¹⁹⁻²²¹
Income²					
High income	6 (50%) ^{147-151,153}	32 (70%) 162,163,165, 170-177,179, 180,182,184- 186,188-193,195- 197,199-201,203- 205	1 (50%) ²¹⁰	2 (29%) ^{212,213}	1 (33%) ²²¹
Middle income	5 (42%) ^{152,155-158}	9 (20%) ^{161,166,178,181,187,194,198,202,207}	0 (0%)	2 (29%) ^{215,217}	2 (67%) ^{219,220}
Low income	1 (8%) ¹⁵⁸	5 (11%) 161,167-169,206	0 (0%)	4 (57%) ²¹⁵⁻²¹⁸	0 (0%)
Population					
Children and adolescents	1 (8%) ¹⁵⁸	11 (24%) 174,177,184-186, 189,194, 201-203,207	1 (50%) ²⁰⁹	1 (14%) ²¹³	2 (67%) ^{219,221}
Elderly	0 (0%)	4 (9%) 171,189,204,205	1 (50%) ²⁰⁹	0 (0%)	0 (0%)
Setting					
Primary care	5 (42%) ^{148-150,152,154}	26 (57%) 161,166- 169,175,176,178, 179,181,184- 186,189,190,193- 197,199,201-205	2 (100%) 209,210	1 (14%) ²¹⁶	1 (33%) ²¹⁹
Hospital	4 (33%) 147,153,155,158	18 (39%) 161-165, 169-172, 177, 180-182, 187,190,192, 206,207	0 (0%)	3 (43%) 212,213,215	3 (100%) ²¹⁹⁻²²¹
Emergency department	1 (8%) ¹⁵¹	5 (11%) 173,174, 177,181,198	0 (0%)	1 (14%) ²¹²	0 (0%)

Table 5.1. Overview of included studies (continued)

	Trial-based analysis (n = 12)¹⁴⁷⁻¹⁵⁸	Decision tree model (n = 46)^{161-182,184-193,193-207}	Markov model (n = 2)^{209,210}	Dynamic model (n = 7)²¹²⁻²¹⁸	Other (n = 3)²¹⁹⁻²²¹
Clinical indication					
Tuberculosis	6 (50%) 152,154-158	11 (24%) ¹⁶¹⁻¹⁷¹	0 (0%)	6 (86%) ²¹³⁻²¹⁸	2 (67%) ^{219,220}
Influenza	0 (0%)	14 (30%) ^{172-174,194-198,200-205}	0 (0%)	1 (14%) ²¹²	0 (0%)
Pneumonia	2 (17%) 151,153	3 (7%) 178,191,192	0 (0%)	0 (0%)	0 (0%)
Other ³	4 (33%) 147-150	18 (39%) ^{175-177, 179-182,184,185, 186,187-190, 193,199, 206,207}	2 (100%) 209,210	0 (0%)	1 (33%) ²²¹
Diagnostic strategies					
Rapid diagnostic test ⁴	3 (25%) ¹ 47-149	19 (41%) 172,175-177,179-182,194-198,200-203,205,207	1 (50%) ²¹⁰	0 (0%)	0 (0%)
Traditional diagnostic ⁵	6 (50%) 147,152, 154,156-158	15 (33%) 161-164, 166-169, 184-187,191,192,199	0 (0%)	3 (43%) 215,216,218	2 (67%) 219,220
Xpert ⁶	5 (42%) 152,154-157	8 (17%) ¹⁶²⁻¹⁶⁹	0 (0%)	5 (71%) 214-216, 217,218	1 (33%) ²¹⁹
Clinical rule	2 (17%) 153,158	6 (13%) 180,184,186, 189,194,198	0 (0%)	0 (0%)	1 (33%) ²²¹
Perspective					
Societal	0 (0%)	12 (26%) 173,177, 180,184,195,196, 198-201,203,205	1 (50%) ²⁰⁹	3 (43%) 212,213,215	1 (33%) ²²⁰
System ⁷	2 (17%) 149,156	18 (39%) 163,164,168-171, 175,176,178,179, 182,185-188, 196,197,202	1 (50%) ²¹⁰	4 (57%) 213,214, 217,218	2 (67%) ^{219,220}
Provider ⁸	8 (67%) ¹⁴⁷ ,148,150,151,153 ,154,157,158	18 (39%) 161-163,165, 167,172,174, 180,181,189-194, 204,206,207	0 (0%)	1 (14%) ²¹⁶	1 (33%) ²²¹

Table 5.1. Overview of included studies (continued)

Type of analysis	Trial-based analysis (n = 12) ¹⁴⁷⁻¹⁵⁸	Decision tree model (n = 46) ^{161-182,184-193,193-207}	Markov model (n = 2) ^{209,210}	Dynamic model (n = 7) ²¹²⁻²¹⁸	Other (n = 3) ²¹⁹⁻²²¹
Cost-utility analysis	3 (25%) ¹⁴⁷⁻¹⁴⁹	22 (48%) ^{163-165, 169-174,176,177, 179,184,187,197, 199-205}	2 (100%) ^{209,210}	6 (86%) ²¹³⁻²¹⁸	0 (0%)
Cost-effectiveness analysis	9 (75%) ¹⁵⁰⁻¹⁵⁸	20 (43%) ^{161,162,166-168,175, 178,180,185,186, 189-194, 196,198, 206,207}	0 (0%)	1 (14%) ²¹²	2 (67%) ^{219,221}
Cost-minimization analysis	0 (0%)	2 (4%) ^{177,188}	0 (0%)	0 (0%)	1 (33%) ²²⁰

Note that not all items are reported by all articles; hence not all columns sum to the total included articles

¹ Including a microsimulation and two database studies

² According to World Bank definitions

³ Including sinusitis, pharyngitis, sore throat, general respiratory infections

⁴ Includes rapid influenza tests, C-reactive protein tests and procalcitonin tests

⁵ Including microscopy and microbiological cultures

⁶ GeneXpert tuberculosis and rifampicin resistance test

⁷ Includes the healthcare system's and healthcare payer's perspective

⁸ Includes analyses from the perspective of a health centre, a laboratory or other provider of care

Table 5.2. Methods of included studies

	Trial-based analysis (n = 12)¹⁴⁷⁻¹⁵⁸	Decision tree model (n = 46) 161-182,184-193,193-207	Markov model (n = 2)^{209,210}	Dynamic model (n = 7)²¹²⁻²¹⁸	Other (n = 3)²¹⁹⁻²²¹
Time horizon					
Less than 1 year	4 (33%) ^{147-149,153}	16 (35%) 168,171,175, 176,179-182,190-193, 195,196, 198,200	1 (50%) ²⁰⁹	0 (0%)	1 (33%) ²²⁰
One year or more ^f	0 (0%)	7 (15%) ^{170,177,178,185,199,201,207}	1 (50%) ²¹⁰	6 (86%) ^{212-214,216-218}	0 (0%)
Lifetime	0 (0%)	7 (15%) 169,172-174, 200,203,204	0 (0%)	1 (14%) ²¹⁵	1 (33%) ²¹⁹
Unknown	8 (67%) ^{150-152,154-158}	16 (35%) 161-167,184, 186-189, 194,197,202, 206	0 (0%)	0 (0%)	1 (33%) ²²¹
Measurement of effectiveness					
Single-study based	10 (83%) ¹⁴⁹⁻¹⁵⁸	7 (15%) ^{161,162,169,170,174,175,177}	1 (50%) ²¹⁰	0 (0%)	2 (67%) 220,221
Synthesis	2 (17%) ^{147,148}	38 (83%) ^{164-168,171-173,176,178-182,184-207}	1 (50%) ²⁰⁹	7 (100%) 212-218	1 (33%) ²¹⁹
Clinical outcomes reported					
QALYs or DALYs	3 (25%) ¹⁴⁷⁻¹⁴⁹	22 (48%) 163-165,169-174, 176, 177,179,184, 187,197, 199-205	2 (100%) 209,210	6 (86%) ²¹³⁻²¹⁸	0 (0%)
Treatment-related ²	1 (8%) ¹⁵¹	7 (15%) ^{175,180,181,185,189,190,198}	0 (0%)	0 (0%)	0 (0%)
Based on diagnostic performance	5 (42%) ¹⁵⁴⁻¹⁵⁸	5 (11%) 161,162, 166,192,206	0 (0%)	0 (0%)	0 (0%)
Time-related ³	1 (8%) ¹⁵³	6 (13%) 167,191,194-196,207	0 (0%)	0 (0%)	2 (67%) ^{220,221}
Resistance included in analysis	4 (33%) ^{149,155-157}	9 (20%) 164,176,179-182,191,198,201	1 (50%) ²⁰⁹	4 (57%) 214,215,217,218	2 (67%) ^{219,220}

Table 5.2. Methods of included studies (continued)

	Trial-based analysis (n = 12)¹⁴⁷⁻¹⁵⁸	Decision tree model (n = 46)^{161-182,184-193,193-207}	Markov model (n = 2)^{209,210}	Dynamic model (n = 7)²¹²⁻²¹⁸	Other (n = 3)²¹⁹⁻²²¹
Sensitivity analyses					
Univariate	5 (42%) 147,149, 154,156,158	40 (87%) ^{161-168,170,172-174,176-178,180-182,184,186-193,195-206}	2 (100%) 209,210	6 (86%) ^{212-216,218}	3 (100%) ²¹⁹⁻²²¹
Multivariate	0 (0%)	8 (17%) ^{162,164,166,172,185,191,193,200}	0 (0%)	1 (14%) ²¹⁴	1 (33%) ²¹⁹
Probabilistic	5 (42%) 147-149,154,156	31 (67%) 162-166,168-176,179,180,184,187,191,192,194,196-204,207	1 (50%) ²¹⁰	6 (86%) 212,214-218	1 (33%) ²¹⁹

Note that not all items are reported by all articles; hence not all columns sum to the total included articles
 QALY: Quality-Adjusted Life Year; DALY: Disability-Adjusted Life Year; CEAC: Cost-Effectiveness Acceptability Curve

¹ Excluding lifetime horizons

² Includes number of correct diagnoses (for example true positives) and time to correct diagnosis

³ Includes time to correct diagnosis, hospital length-of-stay and disease duration

Regression models and trial-based analyses

When cost and effectiveness data obtained from a trial are directly used to analyse the costs and clinical effects of a new intervention using standard statistical methods, this is regarded as a regression model or trial-based analyses.

Several studies examined cost-effectiveness based upon (mostly) a single trial^{147–158}, without the use of a health-economic model. Four studies assessed diagnostics for general respiratory tract infections^{147–150}, two studies assessed diagnostics for pneumonia^{151,153} and three assessed diagnostics for tuberculosis (TB)^{152,154,155}.

Two studies by Oppong et al. assessed the cost-effectiveness of POC CRP testing^{148,149} and internet-based training for primary care clinicians¹⁴⁹ using regression models. The first study used data from an observational study in Norway and Sweden¹⁴⁸, while the second one used data from a multinational, cluster-randomised, factorial controlled trial in Belgium, the Netherlands, Poland, Spain and the United Kingdom (UK)¹⁴⁹. Both studies incorporated resource use, EQ-5D scores and antibiotic prescribing from the trials^{148,149}. Multilevel modelling was used to model the outcomes of interest: QALYs, antibiotic prescriptions and costs. Both studies found antibiotic prescribing to be less prevalent when a POC CRP test was performed and found no significant differences in health outcomes. The cost-effectiveness of POC CRP testing largely depended on the country, even when similar methods were used: while it was a dominant strategy in the Netherlands; in Spain, it was dominated by usual care^{148,149}. In the analysis where communication training was considered, this was considered to be the overall dominant strategy¹⁴⁹. Cost-effectiveness acceptability curves (CEAC) were included in both articles^{148,149}, using the net benefit regression framework¹⁵⁹. The reduction in antibiotic prescriptions was expressed in monetary terms as additional costs per patient prescription avoided¹⁴⁸ or additional costs per percentage reduction in antibiotic prescriptions¹⁴⁹.

Nicholson et al. conducted an RCT and health-economic evaluation of diagnostic tests for influenza, RSV and *Streptococcus pneumoniae* in adults hospitalised for chronic or acute cardiopulmonary illness in the UK¹⁴⁷. Three strategies were included in the analysis: POC tests, PCR and traditional, conventional laboratory diagnostic assessment. The clinical characteristics of the various diagnostic strategies were compared, and the cost-effectiveness was assessed using a bivariate model, mainly incorporating costs and QALYs of the trial. The authors performed a Bayesian analysis that used 50,000 replications of a Markov Chain Monte Carlo analysis. The probability of cost-effectiveness was calculated at the willingness to pay (WTP) threshold of £20,000/QALY. The authors concluded that there were no major differences between the different strategies, both in costs and QALYs, but that the PCR-based strategy was the most likely to be cost-effective.

Two studies assessed the cost-effectiveness of diagnosing community-acquired pneumonia (CAP). Böhmer et al. assessed a computer algorithm that aided the prescription of antibiotics in hospital settings compared to usual care¹⁵³. Two groups of 15 patients were followed in a single hospital in Germany and the clinical outcomes considered were: days with symptoms, days with antibiotics and hospital length-of-stay. Differences between trial arms were calculated using t tests. The algorithm was considered to be cost-saving, resulting in fewer costs and better treatment¹⁵³. Dinh et al. assessed the cost-effectiveness of a rapid pneumococcal antigen test for CAP patients in the Emergency Department (ED) setting¹⁵¹. 1,224 patients were included over three years; however, no control arm was used. In total, there were 51 positive test results, which led to a change in prescription for seven patients. The authors concluded that the costs of implementing the antigen test

(€8,748 annually) were too high compared to the benefits¹⁵¹.

Six trial-based analyses considered specific TB strategies: different culture-based methods in Kenya¹⁵⁸, Xpert testing in South Africa in laboratories¹⁵⁷ or at POC^{152,154}; automated microscopy in the South African laboratory setting¹⁵⁶; and a second Xpert test in China¹⁵⁵. Patient-level clinical and cost data were collected. Cost-effectiveness outcomes were mostly influenced by the number of patients correctly diagnosed, screened or treated^{152,154-158}, one study included TB morbidity measured with a numerical TB score¹⁵⁴. Differences between the various groups were compared using standard frequentist methods^{152,155}, logistic regression¹⁵², univariate analyses^{154,156,158} and Monte Carlo analyses^{154,156}. Two studies presented a cost-effectiveness frontier to compare the different strategies included in the analysis^{155,156}. The two studies assessing POC Xpert testing in South Africa had different conclusions compared to current care: cost-saving¹⁵² or cost-effective (albeit more expensive)¹⁵⁴, while the study in the South African laboratory setting concluded that the higher costs were not matched by the improvement in TB diagnostic efficacy¹⁵⁷. Wang et al. concluded that the price increase related to performing a second Xpert test is relatively high¹⁵⁵. Four studies questioned the affordability of the assessed TB diagnostic strategies in low-income countries¹⁵⁵⁻¹⁵⁸, even if the result of the analysis was that the assessed strategy was cost-effective.

There were some common gaps in the reporting quality of the trial-based and regression analyses. Mostly, this was a result of the relatively short time horizons: the percentage of trial-based studies that reported the time horizon was 20% and 30% reported the discount rate. 40% of the articles reported the measurement of effectiveness and the choice of model was reported in 20% of the articles. An example of a paper with a high reporting quality is written by Opong et al.¹⁴⁸ Resource use, clinical outcomes and statistical methods were clearly described, and the performed regression analysis provided detailed insight in the parameters relevant for the model.

Decision trees

Decision tree models are used to calculate costs and effectiveness outcomes of different clinical interventions, usually over a limited time period as time cannot be modelled explicitly¹⁶⁰. A combination of decisions and probability rates of occurrence are used to calculate the outcomes for various cohorts in the model.

A total of 46 articles compared diagnostic techniques using a decision tree model. 32 articles focused on the use of diagnostic tests that identify bacterial infections (11 TB-specific tests, seven CRP, four PCT, six group A beta-haemolytic *Streptococcus pyogenes* (GABHS) infection and four used other diagnostic techniques). 12 articles focused on the use of diagnostic tests to detect influenza (FLU OIA, QuickVue and ZstatFlu). Finally, two articles compared tests for both bacterial and viral infection.

The main diagnostic test to detect TB was the Xpert test¹⁶¹⁻¹⁶⁸, with other articles assessing the lateral flow urine lipoarabinomannan assay Alere Determine™ test¹⁶⁹, the IGRA (Interferon-Gamma Release Assay) test¹⁷⁰ and the T cells detection test¹⁷¹. The target population consisted of patients presenting with symptoms of active pulmonary TB disease^{161-166,170}, in some cases Human Immunodeficiency Virus (HIV) patients specifically^{167-169,171}. The comparators were sputum smear microscopy^{161-164,168}, culture¹⁶⁵⁻¹⁶⁷ and chest radiography^{164,166,168}. In the case of testing positive, the treatment was a routine TB regime. Time horizons considered were six months¹⁶⁸, two months¹⁷¹ or lifetime^{169,170}. The clinical outcomes used were QALYs^{163,165,170,171}, DALYs^{164,169}, TB cases detected^{161,162,169,170}, days free

from disease¹⁶⁷ and deaths averted¹⁶⁸. One analysis included multi-drug resistance into the model as they switched to a second-line treatment if rifampicin resistance was identified¹⁶⁴. One study concluded that the delay in diagnosing active TB caused the test strategy not to be cost-effective¹⁷¹. Other studies showed that the combination of Xpert with LF-LAM was cost-effective compared to the usual care^{163,165,169} and testing was the most cost-effective strategy while varying the incidence of TB (even when as low as 0.2%)¹⁷⁰. Uncertainty was included using deterministic^{161–168,170} and probabilistic sensitivity analyses^{162–166,168–171}.

Seven articles assessed a CRP test with a decision tree^{172–178}. Patients were adults in hospital care^{172,173} or children in the ED¹⁷⁴, both with suspected influenza symptoms, adults with symptoms of acute respiratory tract infection attending primary care^{175,176,178} and children visiting the paediatric ED of a hospital with meningeal signs¹⁷⁷. The most common treatment included in the models was amoxicillin, in case of a positive CRP result^{172–177}. In case of a negative result, antiviral therapy was prescribed^{172–176}. Other strategies compared were treatment without testing^{173,176}, no treatment¹⁷³ and intensified communication after targeted training for general practitioners (GPs)¹⁷⁵. Different time horizons were used: 28 days^{175,176}, 15 years¹⁷⁷ and lifetime^{173,174}. Nelson et al. followed patients from ED through the rest of their lives, with a life expectancy of 78.7 years¹⁷⁴. Most analyses used QALYs as the clinical outcome^{172–174,176,177} and one analysis used the number of antibiotic prescriptions safely saved¹⁷⁵. One analysis¹⁷⁶ included the cost of antibiotic resistance, by increasing the price of each prescription with the estimated cost of AMR, based on the annual cost of resistance in the United States. Most articles concluded CRP testing was cost-effective except for one¹⁷². Notably, this specific study estimated that performing the test was only profitable with an influenza prevalence of under 2.5%, whereas it was higher during the study period¹⁷². Uncertainty was included using deterministic sensitivity analysis^{172–174,176,177}, probabilistic sensitivity analyses^{172–174,176} and cost-effectiveness acceptability curves (CEACs)^{172–174,176}.

PCT testing was compared to usual care in four studies^{179–182}. Patients were adults and children with suspected acute respiratory infections presented in primary care¹⁷⁹, the intensive care unit¹⁸⁰ or the hospital in general^{181,182}. Antibiotic treatment was prescribed to all who tested positive. In the usual care arm, empirical antibiotics were prescribed, as judged by the physician. The time horizon was short in all cases: 30 days^{181,182} or one episode of acute respiratory infection^{179,180}. Antibiotics avoided was used in three articles as the clinical outcome, expressed as the number of prescriptions saved¹⁷⁹ or as a reduction of days of treatment^{180,181}. Also, QALYs¹⁷⁹ and number of infections averted were used¹⁸⁰. All articles included AMR: three studies assumed that the value of a safely avoided antibiotic prescription equalled the health system cost of resistant infections attributable to that prescription^{179,181,182}. In another article, the reduction in resistant infections was calculated, multiplying the correlation of reduction in days of antibiotics and rate of resistance (estimated in previous publications¹⁸³) by the difference in days of antibiotics between PCT and the usual care group¹⁸⁰. Three articles concluded that PCT was cost-effective^{179,181,182}. Uncertainty was included using a sensitivity analysis graph^{181,182}, CEAC¹⁷⁹ and tornado diagram¹⁸⁰.

Six decision-tree studies included a marker to detect GABHS infection^{184–189}. Children and adults with symptoms of pharyngitis presenting to primary care^{184–186} or hospitals¹⁸⁷, and in other cases with a sore throat^{188,189}, formed the population. The strategy of performing a rapid test was compared with treating all^{184–188}, treating none^{184,187}, using a clinical scoring measure to determine the treatment (triage the diagnosis and treat those with a high score

with antibiotics)^{184–186} or culture¹⁸⁷. If the diagnostic test was positive, penicillin was prescribed^{187–189} or two types of antibiotics depending on the severity of the infection. Clinical outcomes were expressed as quality-adjusted life-days (QALD)^{184,187}, proportion of patients cured without complications^{185,186} and rate of appropriate use of an antibiotic per patient treated¹⁸⁹. The use of the rapid diagnostic test to detect GABHS was cost-effective^{184–187,189} and cost-saving¹⁸⁸. Uncertainty was included using deterministic analyses^{186,188,189}, a tornado diagram^{184,186,187} or a two-way sensitivity analysis graph¹⁸⁵.

Three articles^{190–192} used other diagnostic techniques to detect pneumonia: plain chest x-ray and blood count¹⁹⁰, different cultures (BinaxNOW-SP and urinary antigen test) as an add-on to standard cultures¹⁹², bronchoalveolar lavage (BAL), mini-BAL, or endotracheal tube or bronchoscopy¹⁹¹. The included populations varied: Bertrán et al. considered a population of patients with community-acquired pneumonia, under age 65, without hospital admission criteria, and patients with acute exacerbation of chronic bronchitis due to respiratory infection¹⁹⁰. Ost et al., modelled a hypothetical cohort of immunocompetent patients in the ICU, intubated for 7 days, with evidence of ventilator-associated pneumonia¹⁹¹. The target population in Xie et al. was hospitalised patients with community-acquired pneumonia¹⁹². The three articles used a healthcare centre's perspective, included uncertainty using a deterministic sensitivity analysis and concluded that testing for pneumonia was cost-effective^{190–192}.

In another article, acute sinusitis was identified by ultrasound or radiographic evaluation¹⁹³. The target population consisted of patients presenting with acute sinusitis in primary care. The analysis took a healthcare centre's perspective, with a time horizon of seven days. The clinical outcome was the probability of being cured. As a result, authors suggested that the most rational clinical response to a suspected case of sinusitis was first “to wait and see for one week”, with a “selective strategy by means of structured clinical assessment” as the next step. Uncertainty was included with a deterministic sensitivity analysis and a two-way sensitivity analysis graph.

In 12 articles, the strategy of a diagnostic test that detects the presence of viruses was studied in a decision tree. The population of the studies were adults^{194–200}, children^{201–203} and the elderly^{204,205} with influenza-like illness during the influenza season^{196–198,200,203,205}, and, in two papers, only patients without prior influenza vaccination were included^{200,201}. The diagnostic test to detect influenza was included in all decision trees and was compared to clinical judgment^{194,204}, empirical treatment^{195,198–200,203}, treating none^{195–197,201,202}, targeted or universal rapid influenza testing¹⁹⁸, culture¹⁹⁹ as well as other rapid diagnostic tests such as FLU OIA, QuickVue, ZstatFlu^{200,203,205} or other tests¹⁹⁶. In the case of a positive influenza test, the treatment was oseltamivir^{196–198,200,201,203–205}. Amoxicillin^{195,198–200,203} and/or oseltamivir^{194–198,201–203,203–205} were included in the empirical treatment strategies. The time horizon was one year^{199,202} or a single episode of influenza-like illness^{195,196,198,200}. The clinical outcome most used was QALYs^{196,197,199–205}, quality-adjusted life expectancy (QALE)^{200,203,205}, days free from disease^{194,196} and antibiotic prescriptions saved¹⁹⁸. Two decision trees considered resistance: oseltamivir resistance was included as a percentage of H1N1 strains²⁰¹ and, as assumed by Michaelidis et al.¹⁷⁹, authors supposed that the relationship between antibiotic use and antibiotic resistance was one-to-one linear, in order to determine the impact of antibiotics used in treating influenza-like illness on overall societal antibiotic resistance¹⁹⁸. Four analyses showed that rapid testing for influenza was not cost-effective during the influenza season^{202,203}, testing was only a cost-effective approach early in a pandemic¹⁹⁷ and, in unvaccinated patients, antiviral therapy without testing was economical-

ly more reasonable¹⁹⁶. Other authors concluded that the optimal strategy depended on the patient's vaccination status and the risk of hospitalisation²⁰⁵. The findings showed that testing saved QALYs by reducing the rates of subsequent hospitalisation for influenza and mortality²⁰⁴. Uncertainty was included using deterministic^{196–205} and probabilistic sensitivity analyses^{196,198,200,203,205}.

Two articles included diagnostic strategies for both viral and bacterial infections^{206,207}. The populations were adults with suspected infections²⁰⁶ and children with symptoms of acute bronchiolitis²⁰⁷, both in a hospital setting. In one article, three testing strategies were compared: a comprehensive strategy in which all available diagnostic tests were requested simultaneously; a stepwise strategy in which a limited number of diagnostic tests could be requested, prioritising those for the most prevalent diseases; and a minimalist strategy in which a limited number of diagnostic tests, prioritising those with the highest sensitivity and specificity could be requested²⁰⁶. In the case of a positive result, the treatment was five days of ceftriaxone²⁰⁶. The stepwise strategy was the most cost-effective in terms of cost per correct diagnosis. Uncertainty was included using deterministic analysis. In the other article, several diagnostic techniques were available (blood count, CRP, PCT, chest X-ray and respiratory virus detection tests) and a clinician could order any of these based on clinical suspicion and adhering to clinical practice guideline²⁰⁷. This forms a branch of the tree (named good practice) and it was compared with another branch in which any test was performed or, if it is done, these tests were not the appropriate ones based on the symptoms that the patient (named lack of good practice). In the case of a positive result from the test, the treatment was antibiotics. The utilisation of good practice in the diagnosis and management of patients with bronchiolitis was associated with both fewer patients readmitted within 10 days post discharge and lower costs. Uncertainty was included using deterministic analysis and a tornado diagram.

Frequent gaps in the reporting of decision tree models are also related to the relatively short time horizons, as is the case with trial-based analyses. The time horizon was reported by 65% of papers and the discount rate was reported by 24%. 68% of the articles reported the measurement and valuation of preference-based outcomes, and 43% the currency, price date and conversion. Michaelidis et al. provide a high-ranking decision tree model, based on the CHEERS score¹⁷⁹. The authors described an easy-to-understand decision tree to model PCT-guided antibiotic therapy for outpatient respiratory-tract infections over a short time horizon (one episode). Extensive sensitivity analyses showed the uncertainty associated with several model assumptions.

Static Markov models

A static Markov model presents a set of mutually exclusive and exhaustive states that describe the progression of a disease for a cohort of patients. In contrast to decision trees, Markov models can be used to incorporate time in the health-economic model²⁰⁸.

Two studies used a static Markov model in combination with a decision tree to model the diagnosis and treatment of respiratory tract infection in the community setting^{209,210}. Balk et al. considered acute bacterial sinusitis in the United States and compared four strategies: not prescribing antibiotics, empirical amoxicillin treatment, amoxicillin treatment based on a set of clinical criteria and amoxicillin treatment based on an x-ray²⁰⁹. A combination of a decision tree and a Markov model was used to model a 14-day period using daily cycles: the decision tree was used to model the index consultation, including any tests performed and treatment decision, while the Markov model was used for the disease development using daily cycles. The Markov model incorporated disease complications,

antibiotic side effects and symptom improvements. AMR was considered, by reducing the efficacy of the antibiotic compared to placebo over the 14-day period. The prevalence of sinusitis was varied using deterministic sensitivity analyses. Balk et al. concluded that prescribing antibiotics based on clinical criteria was cost-effective for settings where most patients experience mild or moderate symptoms, however, they also concluded that empirical antibiotics were cost-effective if a sufficient proportion of the population experienced severe symptoms²⁰⁹. Hunter considered the implementation of POC CRP tests in the UK, comparing three strategies to the current standard of care, in which a GP decides to prescribe antibiotics based on the GP's views and the patient's expectations²¹⁰. The strategies considered are: a CRP test performed by the GP; a CRP test performed by a practice nurse and a CRP test performed by the GP in combination with communication training for the GP²¹⁰. A combination of a decision tree and Markov model was used, with a time horizon of 3 years using 28-day cycles after the index consultation²¹⁰. The decision tree was used to model the consultation and direct follow-up (up to 28 days), while the recurrence of respiratory tract infections following the initial disease episode was modelled using the Markov model with two states: healthy and respiratory tract infection. The prevalence was not varied in the model. The model was probabilistic and a CE plane and CEACs were reported. Hunter concluded that CRP implementation is cost-saving, with the strategy with communication training not being cost-effective²¹⁰. The reporting quality of the study by Hunter was high.

Dynamic models

Dynamic models are characterised by a changing rate of infection within the population, usually based on the number of infected individuals²¹¹. Dynamic models can be individual-based or cohort-based.

One paper considered influenza: Nshimyumukiza et al. compared an influenza rapid diagnostic test followed by antiviral treatment to empiric antiviral treatment in Quebec, Canada, by using an individual-based dynamic model²¹². The model consisted of two parts: a "susceptible, infected, recovered" (SIR) model and an economic model, considering a time horizon of 1 year. The compartmental SIR model consisted of three differential equations to model three states using single-day cycles. The economic analytical model was used to simulate infected persons who could decide to seek care if they were symptomatic. Patients who sought care within 48 hours received oseltamivir, reducing the probability of complications such as pneumonia and death. Two outcomes were reported, the saved costs and life-years per 100,000 person years, and uncertainty was included using deterministic and probabilistic sensitivity analyses. The authors concluded that the testing strategy was dominant (fewer deaths and fewer costs) compared to empirical antiviral treatment²¹².

Six studies assessed TB diagnostic strategies using a dynamic model²¹³⁻²¹⁸, of which five assessed one or more Xpert-based strategies compared to other interventions (e.g. standard care)²¹⁴⁻²¹⁸, one assessed a public-private mixed programme for TB diagnosis²¹⁵ and one a national TB strain service. Countries included were various African countries^{214,216,218}, India^{215,217} and the UK²¹³. The transmission models incorporated stages such as uninfected/susceptible, latent infection and active infection, in four cases including TB resistance^{214,215,217,218} or HIV status^{214,216,218}. Included were individual-level models^{215,217} (agent-based modelling) and compartmental models^{213,214,216,218}. Langley et al. also incorporated a discrete-event simulation to model the patient (presumptive TB cases, visiting a diagnostic centre) and sputum pathways (samples flowing from the diagnostic centre to the lab-

oratory and undergoing various diagnostic tests)²¹⁸. Clinical outcomes considered were DALYs^{214,216–218}, QALYs^{213,215} and life years²¹⁴, with all models considering a time horizon of 10 years or more. Five models incorporated a PSA^{214–218}, four included a CEAC^{214,215,217,218}, all incorporated a deterministic sensitivity analysis^{213–218}, and two a cost-effectiveness frontier^{215,218}. The conclusions of the papers vary and are dependent on various factors, such as the affordability^{214,215}, uncertainty of the input parameters²¹⁶ or procedural factors (e.g. number of referrals and cost-sharing)²¹⁷. Langley et al. identified three cost-effective strategies, including Xpert testing and two microscopy-based strategies for Tanzania²¹⁸. Mears et al., who described universal strain typing in the UK, concluded that this was unlikely to be a cost-effective strategy²¹³.

Due to the complexity of dynamic models, these often are accompanied by extensive supplementary material containing the details of the performed analysis. Areas for improvement regarding the reporting quality are currency conversion methods and details regarding the valuation of QALYs or DALYs. Other than that, there were no major differences between the reporting quality of the various papers.

Other models

Lee et al. assessed the cost-effectiveness of POC testing for TB, including rifampicin resistance, with a new PCR test for India's public sector (Truenat)²¹⁹. The compared strategies were smear microscopy, Xpert and the Truenat test in designated microscopy centres and Truenat at POC. A microsimulation model was used to model a cohort of patients seeking care with TB symptoms, over a lifetime horizon. TB prevalence was based on a previous implementation study. The cost-effectiveness measure was costs per life-year saved; a budget impact analysis was also performed, using time horizons of two and five years. Uncertainty was included using one-way and two-way deterministic sensitivity analyses, varying Truenat sensitivity and linkage to care, and a Tornado diagram, varying various key model parameters. The authors concluded that implementing Truenat at POC was cost-effective²¹⁹.

Two CMAs used accounting data to study optimal healthcare resource use. Bogdanova et al. compared two diagnostic algorithms for TB in Russia: a culture-based diagnostic algorithm and line probe assay to detect resistant TB, using costs collected from the government's accounting systems²²⁰. The reduction in the number of hospital days to the correct diagnosis and treatment initiation were considered clinical outcomes. Oostenbrink et al. used hospital data to assess cost-savings related to the implementation of a decision rule to diagnose and treat children with meningeal signs visiting a Dutch ED²²¹. The considered outcomes were safely avoided lumbar punctures and empirical antibiotic treatment. Both studies estimated and compared the resulting costs of each algorithm. One-way sensitivity analyses were included by changing the cost parameters of the model, and both studies concluded that the investigated strategy was cost-saving compared to current care^{220,221}.

Discussion

In this study, we reviewed 70 economic analyses of applied diagnostic techniques for infectious diseases of the respiratory tract, covering a broad range of illnesses for which individuals seek care including influenza, pneumonia, TB and GABHS. The diagnostic techniques assessed range from POC to laboratory testing in numerous different country settings.

Advantages and disadvantages of different modelling methods

12 studies assessed the cost-effectiveness of a new diagnostic strategy within the context of a single trial. Seven studies were performed in HICs^{147–151,153,155} and five in LMICs^{152,154,156–158}. The scope of trials of diagnostics is sometimes rather limited. Of those trials included in this review, trials have only a few patients¹⁵³, only one trial arm¹⁵¹ or a limited scope (e.g. the cost-effectiveness was only assessed from the perspective of the laboratory)^{156,157}. Most trial-based analyses resorted to outcomes related to the direct performance of the diagnostics^{151,154–158}. The generalisability of the studies was affected by all these factors, and the results may not be applicable outside the direct setting where the analysis was performed^{151–153,157}. These and other aspects make it difficult to assess the effects of the diagnostic method beyond the trial; no trial-based analysis reported a time horizon longer than a couple of weeks. In addition, only a few studies used generic clinical outcomes that can be compared between various studies and disease areas, such as QALYs^{147–149}.

The type of model most frequently found in the review was a decision tree (46 of 70 articles). One of the reported advantages of using a decision tree analysis is that the technique enables comparison of a large number of strategies and even the possibility of combining them sequentially, mostly a clinical scoring system to identify patients to be tested^{184–186,189}. A decision tree can easily be adapted to different health systems and settings. 29 studies were performed in HICs^{162,165,170,171,173–177,179,180,182,184–186,188–193,195–197,199–201,203,205} and 17 in LMICs^{161,163,164,166–169,172,178,181,187,194,198,202,204,206,207}. Although authors usually focus on one age group, one decision tree can also be applied to more than one group, as evidenced by some articles, which included both paediatric and adult populations in their analysis^{181,186,188,194}. Decision trees are straightforward to model and interpret for researchers and clinicians who may not be familiar with pharmacoeconomic methods. The (computational) simplicity of decision trees also enhances feasibility to include several sensitivity analyses, such as calculating the cost-effectiveness under various disease incidence values^{172–178}. A disadvantage of using decision trees can be that long-term outcomes are difficult to include, as time as a factor is not modelled explicitly. Therefore, many of these studies incorporate only a short time horizon or do not detail a time horizon at all. Yet, some overcame this disadvantage by estimating the life expectancy and applying that as time horizon^{173,174} or extended a decision tree with a Markov model to be able to model time explicitly^{209,210}. The main disadvantage found when using a decision tree was that several simplifying assumptions are needed^{164,176,199}, which makes it difficult to generalize the results. The testing strategy was cost-effective in 32 articles and cost-saving in six. Factors that most affected the results were the prevalence of infection and the patient's vaccination status. First, below and above certain prevalence percentages, the testing strategy is no longer cost-effective, either because there is not a sufficient number of cases or because empiric treatment dominates the other strategies^{162,167,172–174,185,187,192,194,199,201,206}. Also, the patient's vaccination status is a key aspect that affected the results, as the vaccination status influences the probability of a patient having the disease of interest: the cost-effectiveness of testing can be reduced in vaccinated populations^{177,195,196,201–203,205}. The test parameters affected the results of the analysis, but also the physician's judgment was found to be influential.

Seven studies used a dynamic model to assess the use of a novel diagnostic strategy^{212–218}. These models included transmission of influenza²¹² or TB^{213–218}, providing more flexibility compared to most other model types included in this review. For example, the TB models included time horizons ranging from ten years to a lifetime^{213–218}, and included either QALYs or DALYs^{213–218}. Two studies were performed in HICs^{212,213} and five in LMICs^{214–218}.

Dynamic models require more data than more straightforward models and this was mentioned as a disadvantage by some studies included^{212,214,216,217}. The authors of two papers mentioned that the time to treatment was an important aspect that was not modelled in their analysis^{213,216}. Langley et al. found a solution to this problem, as the authors not only modelled TB transmission, but also the operational process of transporting test samples to external diagnostic laboratories for different types of microscopy and Xpert in Tanzania²¹⁸. Using a discrete event simulation (DES), the authors modelled the time to start the correct treatment and loss-to-follow-up²¹⁸. Lee et al. did not include transmission in their TB microsimulation, instead, they used incidence data from a previous study. In this model, continuity of care after a visit to a health centre was improved when diagnosing at POC compared to external laboratories, resulting in better patient outcomes²¹⁹.

Three models were CMAs^{188,220,221} with a main focus on financial outcomes and not on clinical outcomes, which we believe to be a feasible approach when the new diagnostic strategy is at least as effective as current care. As the data requirements of the CMAs included in this review (unit-level accounting data or micro-costing) are larger than most other reviewed studies, this type of study may be less reproducible and more difficult to interpret for clinicians.

Inclusion of AMR

Twenty articles included resistance in some way^{149,155–157,164,176,179–182,191,198,201,209,214,215,217–220}. We identified three main methods for including AMR, which were used by more than one paper. Seven models incorporated AMR by adding a ‘societal cost’ to any antibiotics prescription^{149,176,179,181,181,191,198}, which was based on a paper by Oppong et al²²². Six models used a fixed percentage of resistant infections, in some cases varied in sensitivity analyses^{155–157,164,201,218}. Five studies modelled resistance dynamically, by changing the resistance rate based on the consumption of antibiotics^{180,214,215}, by modelling both a sensitive and resistant strain²¹⁷, or by decreasing the efficacy of antibiotics in the future²⁰⁹. Clearly, the consequences of increasing AMR can be incorporated in the numerator or denominator of the ICER (i.e., included as a cost or a clinical effect). Still, most studies did not include AMR in the analysis at all, even though this is highly relevant for patients, care providers and policy makers¹³⁵.

Limitations

This review focused on diagnostic strategies, as defined in the methods. The distinction between a diagnostic test and tests used for other purposes, such as screening, disease monitoring or (pharmaco-)genetic tests, can be difficult to make in certain cases. In many cases, the authors do not clearly specify the population to which the test is applied. We therefore tried to clearly define what we consider to be a diagnostic and made the distinction in the full-text screening round, so that we could decide from the methods, often the model specification, whether the paper should be included. We included five languages in this review, which limited the papers we included (seven papers were excluded based on language). Additionally, we only included papers from the year 2000 onwards, which allowed us to focus on current health-economic methods and diagnostic strategies.

Due to the many different diagnostic strategies that are included in this review, as well as the large number of healthcare systems (over 30 countries), we have not included any comparison of cost-effectiveness results. The quality of reporting was included using the CHEERS checklist as described above.

Opportunities for further research

Considering the number of articles included in this review, there is great interest in the cost-effectiveness of diagnostics for respiratory infections. However, there are some gaps where further research could be warranted. For example, in POC diagnostics in general practice: for many HICs we found no study on the cost-effectiveness of any test. Considering the significant reductions of antibiotic consumption linked to these tests¹⁴⁹, this may be an important opportunity to contain increasing AMR. Even more reductions in inappropriate antibiotic prescribing may be possible in LMICs²²³, but few studies considered the cost-effectiveness of diagnostics for other respiratory infections than TB^{181,187,194,198,202,206,207}.

In the hospital setting, multiplex PCR systems may increasingly play a role in quickly testing for a range of viruses and bacteria, which can provide valuable insight into the local transmission of respiratory pathogens²²⁴. Only two papers included an analysis of multiplex PCR; both papers considered patient-specific outcomes, but not the broader value of knowing the epidemiology of respiratory tract infections in the community^{147,174}. Although multiplex PCR systems may be major investments for hospitals, the collected data on the aetiology of ca-anti and AMR could be used to inform prescribing decisions by GPs in the community setting as well.

This review shows that there are many different methodological approaches used in literature to assess the cost-effectiveness of diagnostics for respiratory tract infections. While 33 studies used generalisable outcomes such as QALYs^{147–149,163,165,170,171,173,174,176,179,184,187,197,199–205,209,210,213,215,221} or DALYs^{164,169,214,216–218}, 10 studies used outcomes related to the accuracy of the diagnostic test^{154–158,161,162,166,192,206} (such as the percentage of correct diagnoses) and eight used an outcome related to the prescribed treatment (such as the number of antibiotics saved)^{151,175,180,181,185,189,190,198}. Therefore, many studies did not incorporate a generally comparable clinical outcome in their cost-effectiveness analysis. This also translated into the time horizon used, with varying durations, various studies not reporting a time horizon at all^{150–152,154–158,161–167,171,184–189,197,202,206,220,221} or using only a limited time horizon of less than one year^{147–149,153,168,175,176,179(p201),180–182,190–193,195,196,198,209}. Generalizable outcomes and sufficiently long time horizons are regarded as important principles when performing health-economic analyses^{50,225} and identified as important areas for improvement for economic analyses of diagnostics of respiratory tract infections.

Policy implications

None of the included articles assessed the diagnosis of a coronavirus. Due to the major economic impact of COVID-19, we expect any testing strategy here already to be worthwhile if this means that economies can function normally again. The recent public attention on rapid COVID-19 tests and the knowledge that respiratory-tract infections can be diagnosed rather precisely may result in a permanent change in treatment practice. There may be a shift for doctors, who have experienced the value of point-of-care testing during the COVID-19 pandemic, but also a shift for patients, who may demand to be informed regarding the cause of their respiratory complaints.

As the IVDR will come into effect soon, diagnostic companies will need to prove the clinical effectiveness of new products entering the European market¹⁴². Diagnostic test accuracy alone will not be sufficient to obtain market authorization and diagnostic companies will need to monitor patient outcomes associated with their tests over a longer term. We recommend these companies include quality-of-life measurements in their trials, enabling the calculation of QALYs or DALYs at a later stage if they want to draw any conclusions with regards to cost-effectiveness. Additionally, provided that sufficient clinical outcomes

are recorded, standard pharmacoeconomic methods can be used to extrapolate the trial results so that longer time horizons can be included. The increased availability of clinical data on the performance of diagnostics after the introduction of the IVDR will present decision makers with a more evidence-based method of assessing diagnostics¹⁴². We expect HTA will play an increasingly important role here, as it has with pharmaceuticals¹⁴⁰. Many European countries have developed guidelines for economic analyses, which are most often tailored to and used for pharmaceuticals¹⁴⁰. Health-economic guidelines for diagnostics are not as well developed, an issue which has previously been raised by Garfield et al., who looked at the assessment of molecular diagnostics from the perspective of various HTA agencies²²⁶. We would recommend decision makers to consider the application of the pharmacoeconomic guidelines to diagnostics and adapt these guidelines if needed.

Conclusion

This review shows that methods used to assess the cost-effectiveness of respiratory-tract infection diagnostics vary greatly. Main points for improvement in this field are the application of generalizable outcomes and the extrapolation of results beyond the time horizon of the trial.

Acknowledgments

We thank all members of work package 5 of VALUE-Dx for their valuable input.

