

Sputum Gram Stain for Bacterial Pathogen Diagnosis in Community-acquired Pneumonia: A Systematic Review and Bayesian Meta-analysis of Diagnostic Accuracy and Yield

Hiroaki Ogawa,¹ Georgios D. Kitsios,² Mitsunaga Iwata,¹ and Teruhiko Terasawa,^{1,2}

¹Department of Emergency and General Internal Medicine, Fujita Health University School of Medicine, Toyoake, Aichi, Japan; and ²Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh Medical Center and University of Pittsburgh School of Medicine, Pennsylvania, United States of America

(See the Editorial Commentary by Meyer Sauter on pages 514–6.)

Background. The clinical role of sputum Gram stain (SGS) in community-acquired pneumonia (CAP) diagnosis remains controversial. A 1996 meta-analysis of the diagnostic accuracy of SGS reported heterogeneous results. To update the available evidence, we performed a systematic review and a Bayesian standard and latent-class model meta-analysis.

Methods. We searched Medline, Embase, and Cochrane Central by 23 August 2018 to identify studies reporting on the diagnostic accuracy, yield (percentage of patients with any pathogen[s] correctly identified by SGS), and clinical outcomes of SGS in adult patients with CAP. Two reviewers extracted the data. We quantitatively synthesized the diagnostic accuracy and yield, and descriptively analyzed other outcomes.

Results. Twenty-four studies with 4533 patients were included. The methodological and reporting quality of the included studies was limited. When good-quality sputum specimens were selected, SGS had a summary sensitivity of 0.69 (95% credible interval [CrI], .56–.80) and specificity of 0.91 (CrI, .83–.96) for detecting *Streptococcus pneumoniae*, and a sensitivity of 0.76 (CrI, .60–.87) and specificity of 0.97 (CrI, .91–.99) for *Haemophilus influenzae*. Adjusted analyses accounting for imperfect reference standards provided higher-specificity estimates than the unadjusted analyses. Bacterial pathogens were identified in 73% (CrI, 26%–96%) of good-quality specimens, and 36% (CrI, 22%–53%) of all specimens regardless of quality. Evidence on other bacteria was sparse.

Conclusions. SGS was highly specific to diagnose *S. pneumoniae* and *H. influenzae* infections in patients with CAP. With good-quality specimens, SGS can provide clinically actionable information for pathogen-directed antibiotic therapies.

Keywords. community-acquired pneumonia; diagnosis; meta-analysis; sensitivity and specificity; sputum Gram stain.

Community-acquired pneumonia (CAP) develops in previously healthy individuals or those with limited contact with medical institutions or settings; they are a leading infectious cause of death, with many lives lost globally [1]. The mainstay of pneumonia therapy is an appropriate antimicrobial treatment, covering causative agents while avoiding antimicrobial overuse and development of resistance. Thus, accurate and timely microbiological diagnosis is critical for pneumonia management.

Despite improvements in microbiological diagnostic procedures, the causative pathogen is not detected in about half of CAP cases [2, 3]. Among cases with identified etiology, stand-alone bacteria or viruses and coinfection thereof are the most common

pathogens [2, 3]. Conventional bacterial CAP diagnosis requires growth and isolation of the culprit pathogen(s) in blood or respiratory specimen cultures following incubation on appropriate media, identification of the isolated bacterial species, and antibiotic susceptibility testing, steps that collectively require 2–3 days before actionable results are available. Additionally, the antibiotic therapies prior to culture specimen acquisition further reduce the sensitivity of microbiological cultures. Therefore, along with the apparent success of empirical treatments and the impetus to administer antibiotics early [4] as well as the possibility of bacterial coinfection, even in the cases with apparently stand-alone viral infection, current guidelines recommend empirical treatments with broad-spectrum antimicrobials for patients with CAP [5–7].

Gram stain (GS) of expectorated sputum is an inexpensive, noninvasive, readily available test that can promptly identify causative bacteria if performed by an experienced observer in a qualified laboratory on good-quality specimens [8]. Results of sputum GS can further facilitate interpretation the results of the sputum culture. Currently, for the rapid detection of bacterial pathogens, sputum GS, along with urine antigen tests for

Received 19 December 2018; editorial decision 16 July 2019; accepted 2 September 2019; published online September 5, 2019.

Correspondence: T. Terasawa, Section of General Internal Medicine, Department of Emergency and General Internal Medicine, Fujita Health University School of Medicine, 1–98 Dengakugakubo, Kutsukakecho, Toyoake, Aichi 470–1192, Japan (terasawa@fujita-hu.ac.jp).

Clinical Infectious Diseases® 2020;71(3):499–513

© The Author(s) 2019. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com.
DOI: 10.1093/cid/ciz876

Streptococcus pneumoniae and *Legionella pneumophila* are the most commonly used rapid point-of-care tests. Moreover, rapid multiplex polymerase chain reaction (PCR) tests targeting the nucleic acids of viruses are widely used. However, the detection of a plausible viral culprit using PCR does not exclude the possibility of bacterial superinfection and the subsequent need for targeted antibiotic therapies. Additionally, despite the recent emergence of PCR-based tests for the syndromic testing of bacterial pathogens in respiratory specimens with rapid turnaround times [9, 10], the clinical adoption of such tests varies and sputum GS remains the frontline diagnostic tool in most institutions.

A previous meta-analysis assessing the diagnostic accuracy of sputum GS in CAP published in 1996 [11] reported heterogeneous results with limited conclusions. The meta-analysis assumed that the various culture-based imperfect reference standards in the primary studies were perfectly accurate [11]. However, it is now well-appreciated that the naively estimated test accuracy is biased when the reference standard is imperfect [12]. In addition, the meta-analysis assessed *S. pneumoniae* only, whereas other clinical outcomes, such as diagnostic yield, were not considered. Several primary studies have assessed this topic since 1996, thus providing available data for an updated and methodologically appropriate analysis. Therefore, we conducted a systematic review of the literature on sputum GS and performed a meta-analysis.

METHODS

This systematic review and meta-analysis followed the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) extension for diagnostic test accuracy statement [13], and was exempted from ethics review. The protocol has been published elsewhere [14].

Literature Search

We searched Medline, Embase, and Cochrane Central from the beginning until 23 August 2018, and used “sputum,” “Gram stain,” “pneumonia,” and their synonyms as search terms (Supplementary Materials). The references of eligible studies and review articles were also examined.

Inclusion Criteria

Two investigators (H. O., T. T.) independently screened abstracts and examined full-text articles for eligibility. Prospective or retrospective studies that used sputum GS in ≥ 10 adult (≥ 18 years of age) patients with CAP were included. The pneumonia cases developed in nursing home residents were also encompassed [6]. Our primary outcome of interest was the diagnostic accuracy for specific bacterial etiologies. The secondary outcomes included the diagnostic yield, defined as the percentage of patients with any pathogen(s) correctly identified by sputum GS, effect of sputum GS on diagnostic and therapeutic management, and other patient-relevant clinical

outcomes. Discrepancies were resolved via consensus. Details are provided in the Supplementary Materials.

Data Extraction

One investigator (H. O.) extracted descriptive data, which were confirmed by another investigator (T. T.). The extracted descriptive data included the study, patient, and test characteristics (Supplementary Materials).

Two reviewers (H. O., T. T.) independently extracted the numerical data. Discrepant extractions were resolved by consensus. When the data were not extractable, we contacted the authors for additional data. See the Supplementary Materials for operational definitions.

Risk of Bias

Two reviewers (H. O., T. T.) independently assessed the risk of bias and concerns about applicability based on the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool [15]. Discrepancies were resolved by consensus (Supplementary Table 1).

Data Synthesis

For all clinical outcomes, individual patients were considered as the unit of analysis. For diagnostic accuracy, the sensitivity and specificity were calculated as summary measure. A hierarchical Bayesian latent class model (LCM) was used to estimate both unadjusted and adjusted summary receiver operating characteristic (ROC) curves and the corresponding sensitivity and specificity values with 95% credible intervals (CrIs) [12]. The Bayesian LCM meta-analysis represents a recently developed extension of the standard meta-analysis of diagnostic accuracy and accounts for multiple imperfect reference standards under the LCM assumption that the true disease status is unobservable to calculate adjusted sensitivity and specificity. Summary positive and negative likelihood ratios (PLRs and NLRs, respectively) were calculated from the summary sensitivity and specificity estimates. For diagnostic yield, we used the hierarchical Bayesian random-effects meta-analysis of proportions [16]. For complete details regarding the methodology, model fitting, choice of prior distributions for the parameters, and operational definitions used in the sensitivity analysis, see the Supplementary Materials.

We visually assessed between-study heterogeneity by plotting the accuracy estimates in the ROC space. Alternative models were compared by using the deviance information criterion (DIC) and considering the differences in DIC scores > 5 as important. The lack of data resulted in subgroup analyses performed only on the year of publication and study location. Funnel plot asymmetry was not examined, because the required tests do not allow a valid assessment of the extent and impact of publication and selective reporting bias in studies of diagnostic accuracy [13]. All analyses were conducted by using OpenBUGS

3.2.3 (OpenBUGS Project Management Group, www.openbugs.net) and Stata SE software 14.1 (StataCorp, College Station, Texas). *P* values for all comparisons were 2-tailed, and statistical significance was defined as *P* < .05.

RESULTS

Literature Search and Eligible Studies

After abstract screening, 142 potentially eligible full publications were reviewed (Figure 1). After exclusions (see the Supplementary Materials for details), 24 independent studies [17–40] (22 on diagnostic accuracy, 4 on diagnostic yield, and 1 on changes in patient management) were included in this review.

Study and Clinical Characteristics

The eligible 24 studies (9 from the United States [17–21, 24, 25, 29, 32], 8 from Europe [22, 26, 27, 30, 31, 34, 35, 40], 4 from Japan [28, 33, 36, 38], and 1 each from Australia [39], Bangladesh [37],

and China [23]) provided relevant data from 4533 patients (Table 1). Fifteen of the 24 (63%) studies were prospective [19, 22, 25–27, 30–34, 36–40], and 4 (17%) were retrospective [21, 28, 29, 35]. The other 5 studies (21%) did not provide adequate information to classify the exact study design [17, 18, 20, 23, 24]. One study [38] assessed patients with healthcare-associated pneumonia (HCAP) jointly with those with CAP. The separate data on CAP were provided by the author through personal communication.

Fourteen of the 24 (58%) studies assessed hospitalized patients on admission [17, 18, 20, 24–28, 33–35, 38–40], whereas another 4 (17%) considered the patients treated in the emergency department [23, 29, 30, 32]. The other 6 (25%) studies did not report on clinical context (inpatients vs outpatients) [19, 21, 22, 31, 36, 37]. All studies defined CAP as a combination of acute symptoms of lower respiratory tract infection with a new radiographic infiltrate.

The median sample size was 131 patients (range, 16–533) with an average age of 31–77 years (Supplementary Table 2). Nine

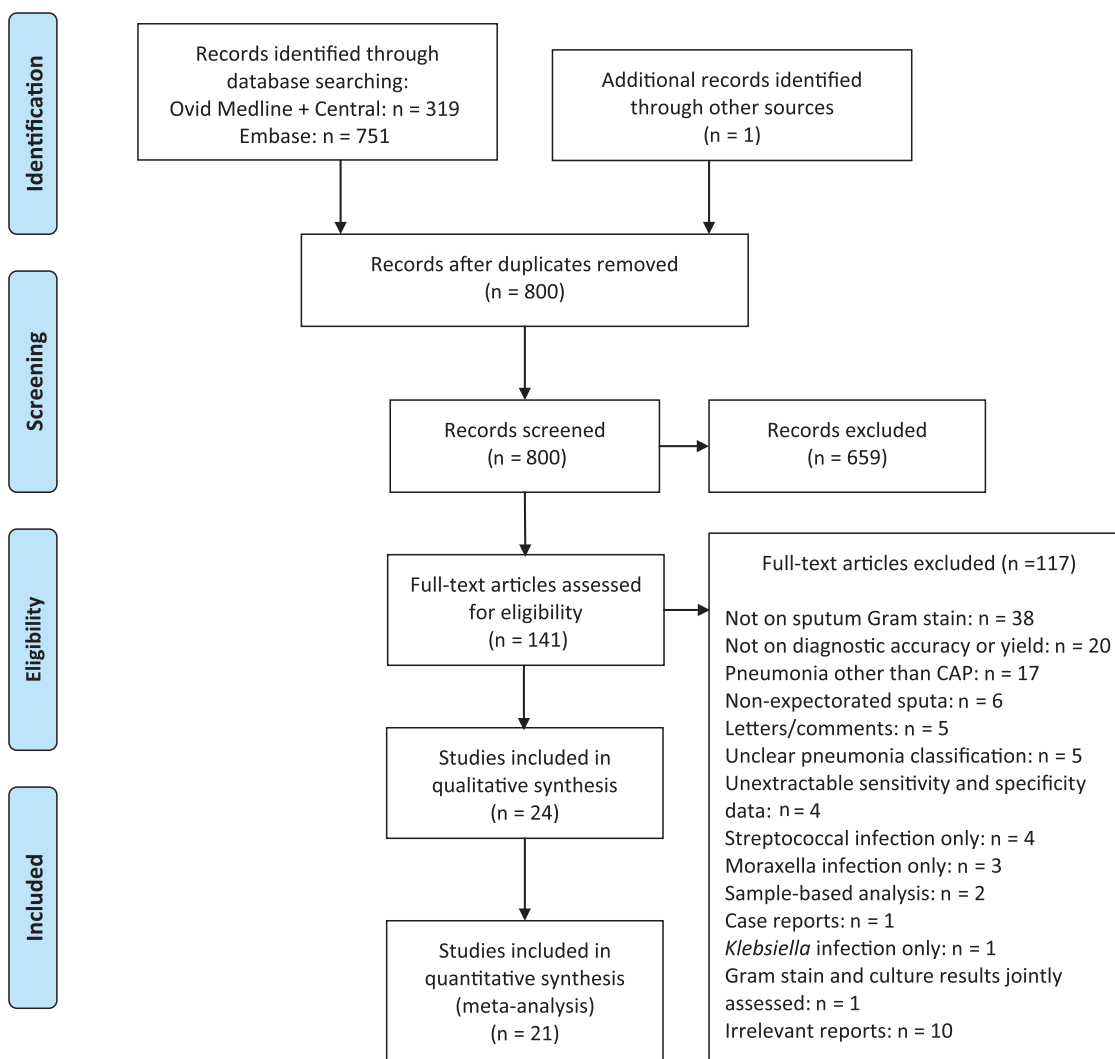


Figure 1. Preferred Reporting Items for Systematic Review and Meta-Analyses flow diagram. Abbreviation: CAP, community-acquired pneumonia.

Table 1. Study Characteristics

Study ID (Location)	Study Year	Design (No. of Centers)	Enrollment	Clinical Context	Definition of Pneumonia	Inclusion Criteria	Exclusion Criteria	Comparator Tests ^a	Reference Standard for Positive Diagnosis
Merrill 1973 [17] (Charlottesville, VA, US)	ND	ND (1)	ND	"Acute" pneumonia on admission ^b	Acute change in health, fever >37.8°C, cough, new X-ray pulmonary infiltrates	Adults, "acute" pneumonia that required hospitalization, no previous antibiotic therapy	ND	Sputum antigen for <i>Streptococcus pneumoniae</i> by Quellung reaction	(1) Sputum culture only, (2) CRS (APR for cultures of multiple specimens)
Thorsteinsson 1975 [18] (Houston, TX, US)	ND	ND (1)	ND	"Acute" pneumonia on admission ^b	Symptoms/signs of acute pneumonia, X-ray pulmonary infiltrates	"Acute" pneumonia that required hospitalization, no previous antibiotic therapy	ND	None	(1) Sputum culture only, (2) transtracheal aspirate culture only, (3) bronchial aspirate culture only, (4) CRS (APR by cultures of multiple specimens)
Rein 1978 [19] (Charlottesville, VA, US)	ND	Prospective (1)	Consecutive	CAP	Acute productive cough, new X-ray pulmonary infiltrates	CAP	ND	Sputum culture, mouse inoculation of sputum, sputum antigen for <i>S. pneumoniae</i> by Quellung reaction	(1) Sputum culture only, (2) CRS (APR for cultures of multiple specimens, mouse inoculation of sputum, sputum antigen for <i>S. pneumoniae</i>)
Boerner 1982 [20] (Durham, NC, US)	ND ^c	ND (1)	ND	CAP on admission	Symptoms/signs of acute RTI, X-ray pulmonary infiltrates or consolidations	CAP that required hospitalization	ND	None	CRS (APR by cultures of multiple specimens)
Dans 1984 [21] (Baltimore, MD, US)	1971–1972, 1979–1980	Retrospective (1)	Inconsecutive	CAP	Fever, X-ray pulmonary infiltrates, treating physicians' clinical diagnosis	CAP that required hospitalization	CAP as the secondary diagnosis, incomplete data	None	Sputum culture only
BTS 1987 [22] (nationwide, UK)	Nov 1982–Dec 1983	Prospective (25)	Inconsecutive	CAP	Acute symptoms, new segmental or lobar X-ray pulmonary infiltrates	Adults (15–74 y), CAP that required hospitalization	Pneumonia not the main reason of admission, pneumonia as the terminal event, pulmonary TB	None	CRS (APR by cultures of multiple specimens or sputum antigen for <i>S. pneumoniae</i>)
Zhang 1988 [23] (Shanghai, China)	Dec 1986–Feb 1987	ND (1)	Inconsecutive	CAP in ED	History/symptoms of acute LRTI, X-ray pulmonary infiltrations	CAP	No sputum collected	None	CRS (APR by cultures of multiple specimens, urine antigen for <i>S. pneumoniae</i> , or serology for <i>S. pneumoniae</i>)
Gleckman 1988 [24] (Worcester, MA, US)	Jan 1982–July 1987	ND (1)	Consecutive	CAP with bacteremia on admission	Symptoms/signs of acute RTI, new X-ray pulmonary infiltrates	Adults, CAP with isolation of a bacterium from blood that required hospitalization	Any coexistent infection	None	(1) Sputum culture only, (2) blood culture only, (3) CRS (APR by cultures of sputum or blood)

Table 1. Continued

Study ID (Location)	Study Year	Design (No. of Centers)	Enrollment	Clinical Context	Definition of Pneumonia	Inclusion Criteria	Exclusion Criteria	Comparator Tests ^a	Reference Standard for Positive Diagnosis
Fine 1991 [25] (Pittsburgh, PA, US)	Jul 1986–Mar 1987	Prospective (2)	Inconsecutive	CAP or HCAP ^d on admission	Symptoms/signs of LRTI, new X-ray pulmonary infiltrates	> 16 y, CAP or HCAP ^d that required hospitalization	No sputum collected or missing results	None	CRS (APR by cultures of multiple specimens)
Bohte 1996 [26] (Leiden, Netherlands)	Jan 1991–Apr 1993	Prospective (6)	Inconsecutive	CAP on admission	New X-ray pulmonary infiltrates	≥ 18 y, CAP that required hospitalization	HCAP; hospitalization ≤ 1 wk, failures to obtain serologic tests, concomitant infection	None	CRS (APR for cultures of multiple specimens)
Roson 2000 [27] (Barcelona, Spain)	Feb 1995–May 1997	Prospective (1)	Consecutive	CAP on admission	≥ 1 signs/symptoms of LRTI, new X-ray pulmonary infiltrates	CAP that required hospitalization	Neutropenia, AIDS, transplantation, pneumonia of “unknown origin”	None	CRS (APR for cultures of multiple specimens and PCR of needle aspirate for <i>S. pneumoniae</i>)
Sato 2002 [28] ^e (Tokyo, Japan)	Jan 1997–Dec 2000	Retrospective (1)	ND	CAP on admission	Acute signs/symptoms of LRTI, new X-ray pulmonary infiltrates	CAP that required hospitalization	Aspiration pneumonia, patients requiring ventilator, HCAP ^d	None	Sputum culture only
Butler 2003 [29] (Atlanta, GA, US)	Jan 1997–Mar 1998	Retrospective (1)	Inconsecutive	CAP in ED	≥ 1 signs/symptoms of LRTI, X-ray pulmonary infiltrates	≥ 18 years, CAP that required hospitalization	Use of antimicrobials ≤ 7 d, no timely informed consent, HIV infection, anuria due to AKI/CKD, use of urinary catheter for > 24 h, bleeding diathesis, abnormality/alteration of the upper respiratory tract	Urinary antigen for <i>S. pneumoniae</i> , sputum PCR	CRS (APR for cultures of multiple specimens)
Garcia 2004 [30] (Barcelona, Spain)	Oct 1996–Apr 2002	Prospective (1)	Consecutive	CAP in ED	Signs/symptoms of LRTI, new X-ray pulmonary infiltrates	> 14 y, CAP	Neutropenia, HIV infection, TB, fungal infection, patients treated with immunosuppressive drugs, disease duration ≥ 2 wk	None	CRS (APR for cultures of multiple specimens)
Roson 2004 [31] (Barcelona, Spain)	Jun 2000–Apr 2002	Prospective (1)	Consecutive	CAP	Acute respiratory illness, new X-ray pulmonary infiltrates	Adult, non-severely immunosuppressed, CAP	Neutropenia, AIDS, transplant recipients, pneumococcal vaccination ≤ 1 wk	Urinary antigen for <i>S. pneumoniae</i>	CRS (APR for cultures of multiple specimens)
Yang 2005 [32] (Baltimore, MA, US)	Oct 2001–May 2003	Prospective (1)	Consecutive	CAP in ED	Acute signs/symptoms of LRTI, leukocytosis, new X-ray pulmonary infiltrates	> 17 y, CAP, excess of sputum samples available, no missing data on reference standard	Failures to receive a reference standard	None	CRS (APR for cultures of multiple specimens, sputum or BAL fluid antigen for <i>S. pneumoniae</i>)
Miyashita 2008 [33] (Kurashiki, Japan)	Jan 2004–Jul 2007	Prospective (1)	ND	CAP on admission	Signs/symptoms of LRTI, new X-ray pulmonary infiltrates	CAP that required hospitalization	HAP, HIV infection, use of immunosuppressive therapy or steroids, HAP	None	Sputum culture only
Anevlavis 2009 [34] (Athens and Alexandroupolis, Greece)	Jan 2002–Jun 2008	Prospective (2)	Inconsecutive	CAP on admission	Signs/symptoms of LRTI, increased PMNs, X-ray pulmonary infiltrates	Selected “bacterial” CAP, no antimicrobial therapy < 2 wk, same organism identified from both blood and sputum	ND	None	Both blood and sputum cultures positive

Table 1. Continued

Study ID (Location)	Study Year	Design (No. of Centers)	Enrollment	Clinical Context	Definition of Pneumonia	Inclusion Criteria	Exclusion Criteria	Comparator Tests ^a	Reference Standard for Positive Diagnosis
Ferre 2011 [35] (Barcelona, Spain)	Oct 2005–Nov 2007	Retrospective (1)	Consecutive	Hospitalized CAP from ED on admission	Signs/symptoms of LRTI, X-ray pulmonary infiltrate	CAP that required hospitalization from ED	Pediatric or gynecology patients, cases requiring ICU care, empyema, immunosuppressed patients, HIV infection, patients on HD	None	CRS (APR for cultures of multiple specimens, urinary antigen for <i>S. pneumoniae</i>)
Fukushima 2013 [36] (nationwide, Japan)	Mar 2006–Mar 2007	Prospective (14)	ND	CAP	ND	≥16 y, CAP	ND	None	Sputum culture only
Akter 2014 [37] (Dhaka, Bangladesh)	Jul 2011–Jun 2012	Prospective (1)	Consecutive	CAP	Fever, signs/symptoms of LRTI, new or progressing X-ray pulmonary infiltrates	> 18 y, CAP	TB, BA, CHD, AKI/CKD, foreign body aspiration, current use of or recently completed antibiotic therapy	None	Sputum PCR for <i>S. pneumoniae</i> and <i>H. influenzae</i>
Fukuyama 2014 [38] ^b (Uruma, Japan)	Aug 2010–Jul 2012	Prospective (1)	Consecutive	Hospitalized CAP or HCAP ^c from ED on admission	Signs/symptoms of LRTI, new X-ray pulmonary infiltrate	CAP or HCAP ^c that required hospitalization from ED	Nonpneumonia causes identified later through clinical follow-up	None	CRS (APR for cultures of multiple specimens, urinary antigen for <i>S. pneumoniae</i>)
Studies that assessed diagnostic yield									
Lim 1989 [39] (Adelaide, Australia)	Apr 1987–Mar 1988	Prospective (1)	Consecutive	CAP on admission	Symptoms/signs of acute pneumonia, new X-ray pulmonary infiltrates	CAP that required hospitalization	Patients with immunosuppressive disorders, treated with immunosuppressive drugs, or with disorders that affects consciousness	Sputum culture, blood culture, viral culture	An operational diagnostic algorithm consisting of definitive and presumptive etiologies
van der Eerden 2005 [40] (Alkmaar, Netherlands)	Dec 1998–Nov 2000	Prospective (1)	ND	CAP on admission	Symptoms/signs of acute pneumonia, new X-ray pulmonary consolidations	≥18 y, CAP that required hospitalization	Severe immunosuppression, malignancy, pregnancy, lactation, severe allergy to antibiotics, obstruction pneumonia, <8 d after hospital discharge	Sputum culture, sputum and urine antigen for <i>S. pneumoniae</i> , urine antigen for <i>Legionella pneumophila</i> , and serological tests	An operational diagnostic algorithm consisting of definitive and presumptive etiologies

Abbreviations: AKI, acute kidney injury; APR, any tests positive rule (ie, at least 1 positive result for any of the multiple reference standard tests performed was deemed as composite reference standard positive); BA, bronchial asthma; BAL, bronchoalveolar lavage; BTS, British Thoracic Society; CAP, community-acquired pneumonia; CHD, congenital heart disease; CKD, chronic kidney disease; CRS, composite reference standard; ED, emergency department; GA, Georgia; HD, hemodialysis; HAP, hospital-acquired pneumonia; HCAP, healthcare-associated pneumonia; HIV, human immunodeficiency virus; ICU, intensive care unit; LRTI, lower respiratory tract infection; MA, Massachusetts; MD, Maryland; NC, North Carolina; ND, no data; PA, Pennsylvania; PCR, polymerase chain reaction; PMN, polymorphonuclear; RTI, respiratory tract infection; TB, tuberculosis; TX, Texas; UK, United Kingdom; US, United States; VA, Virginia.

^aOnly tests that were clearly defined and analyzed, the results of which were reported in comparison with the Gram stain results, were considered.

^bNot specifically referred to as CAP.

^cOne-year period.

^dPatients from nursing home. These patients were included (see text).

^eThese studies also assessed diagnostic yield.

^fAccording to the American Thoracic Society 2005 criteria. These patients were excluded from analysis.

studies reported on the proportion of patients with chronic obstructive pulmonary disease (min–max, 20%–45%) [22, 25, 27, 29, 31, 33, 34, 38, 40], and 6 reported on patients with immunosuppressive conditions (min–max, 0%–6%) [27, 30, 35, 38–40]. Five studies reported aspiration pneumonia as the cause in $\leq 37\%$ [27, 29, 31, 32, 38]. The Pneumonia Severity Index scores were available in 6 studies [27, 31, 33, 35, 38, 40] in which approximately 50% were at low risk. Pretest antibiotic use was reported in 14 studies [17, 18, 20, 22–24, 26, 29–31, 33, 34, 37, 38]. Six studies included antibiotic-naive patients only, whereas 8 studies also included patients previously treated with antibiotics. A uniform assessment of the prevalence of specific pathogens across 24 studies was impossible owing to the wide range of pretest antibiotic use (0%–51%) and inconsistency of performed standard reference tests.

Test and Reference Standard Characteristics

Methods on the collection and process of the sputum specimens were poorly described (Supplementary Table 3). Ten studies reported on the test performers and interpreters [17, 19–21, 25, 26, 28, 33, 34, 38]. Supervised resident physicians [17, 20, 25, 28, 38] and/or medical students [17, 19, 21] implemented sputum GS in 7 studies. Similar protocols assessed the quality of sputum samples and defined a good-quality specimen as one containing ≥ 25 leukocytes and < 10 squamous epithelial cells per low-power field. Overall, 15 of 24 (63%) studies reported variable proportions of good-quality specimens (median 61%; range, 36%–100%) in patients with expectorated sputum samples [20, 21, 23–25, 27, 30–38].

Studies have adopted similar visual interpretation criteria for microorganisms screened using GS (Supplementary Table 3). For instance, gram-positive diplococci, typically assessed as the predominant morphology, were defined as the positive criterion for *S. pneumoniae*, whereas other studies have reported a similar positive criterion for *Haemophilus influenzae* as small-sized gram-negative coccobacilli.

The variously defined composite reference standards, based on cultures, antigen tests, and molecular tests on multiple different specimens, were used (Tables 1 and 2; Supplementary Tables 4 and 5). The culture of expectorated sputum specimens was the most commonly adopted component test.

Study Validity

Use of imperfect reference standards in all studies suggested a high risk of bias and limited generalizability of the naively calculated accuracy estimates (Figure 2). Spectrum bias was suspected in 2 studies that excluded patients with specific GS results [20, 24] and another that included patients with only culture-documented bacteremic pneumonia [34]. These 3 studies were excluded from the quantitative synthesis.

Diagnostic Accuracy

Data on good-quality specimens to diagnose *S. pneumoniae* were available from 11 studies (1794 patients, 611 with

S. pneumoniae infection; median prevalence, 37%; min–max, 19%–81%) [19, 23, 25, 27, 29, 30, 32, 33, 35, 37, 38]. The data points were relatively collected together in sensitivity (range, 0.49–0.89) and specificity (range, 0.72–0.99) (Figure 3 and Supplementary Figure 1). The summary estimates for sensitivity and specificity were 0.69 (95% CrI, .56–.80) and 0.91 (95% CrI, .83–.96), respectively, and the summary estimates for PLR and NLR were 7.6 (95% CrI, 4.3–15.1) and 0.34 (95% CrI, .24–.47), respectively (Figures 4 and 5).

Six studies (873 patients, 157 with *H. influenzae* infection; median prevalence, 17%; min–max, 10%–28%) assessed good-quality specimens to diagnose *H. influenzae* [25, 27, 33, 35, 37, 38]. The sensitivity (range, 0.64–0.88) and specificity (range, 0.88–1.0) plots were closely clustered (Figure 3 and Supplementary Figure 2). The summary estimates for sensitivity and specificity were 0.76 (95% CrI, .60–.87) and 0.97 (95% CrI, .91–.99), respectively, and the summary estimates for PLR and NLR were 24.7 (95% CrI, 8.7–123.3) and 0.25 (95% CrI, .14–.41), respectively (Figures 4 and 5).

Only 1 study reported high specificity for *S. aureus* and other gram-negative organisms based on good-quality specimens [38], which precluded quantitative synthesis (Figure 3).

Regarding the studies that evaluated all specimens regardless of sputum quality, the ROC plots for *S. pneumoniae* (11 cohorts from 10 studies; 2162 patients) showed a right-diagonal curvilinear relationship, suggesting a trade-off between sensitivity and specificity [17, 18, 21, 22, 26–28, 31, 35, 36]. Data were sparse for *H. influenzae* (3 studies, 734 patients) [27, 28, 35]. Therefore, we constructed summary ROC curves only for these studies (Supplementary Figures 3–5).

Adjusted Diagnostic Accuracy

Expectedly, the model-adjusted accuracy estimates of the reference standards were < 1 , and adjusted values were typically lower for sensitivity than specificity in both the models of conditional dependence and independence (Supplementary Figure 6). Overall, the adjusted analyses generated extremely high summary estimates in specificity and PLRs; these were higher than the estimates in the unadjusted analysis for both *S. pneumoniae* and *H. influenzae* (Figures 4 and 5). Although the conditional independence model yielded the highest summary estimates, the conditional dependence model showed the lowest DICs among the 3 alternative models, suggesting that it was the best-fit model. However, wide CrIs for the summary estimates substantially overlapped across the alternative model.

Subgroup and Sensitivity Analyses

For both *S. pneumoniae* and *H. influenzae*, the summary estimates in the subgroup analyses were consistent with those in the main analysis (Figures 4 and 5). The stability analyses calculated similar results when the clinical reference standards

Table 2. Reference Standard Tests^a

Target Pathogen	RS Group per Pathogen	Culture							Inoculation			Antigen Test			PCR		Frequency of Cohorts [Reference]
		Sputum	Tracheal Aspirates	Bronchial Aspirates	Pleural Effusion	BAL Fluid	Blood	Urine	Sputum	Urine	Blood	Sputum	Urine	Sputum	Urine	Specimen From the Lesion	
Studies that reported the data on any-quality specimens																	
SP	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5 [17, 21, 28, 36]
	2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2 [26, 31]
	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [35] ^b
	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [22]
	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [27]
	6	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [18]
HI	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [28]
	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [35]
	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [27]
Studies that reported the data on good-quality specimens only																	
SP	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [37]
	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [29] ^c
	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2 [33, 35] ^b
	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [23]
	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [19]
	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [25]
	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [27]
	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [38]
	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [32]
	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [30]
HI	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [37]
	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [33]
	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [35]
	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2 [25, 27]
	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [38]
KP	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [38]
MC	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [38]
PA	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [38]
SA	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 [38]

Abbreviations: BAL, bronchoalveolar lavage; HI, *Hemophilus influenzae*; KP, *Klebsiella pneumoniae*; MC, *Moraxella catarrhalis*; PA, *Pseudomonas aeruginosa*; PCR, polymerase chain reaction; RS, reference standard; SA, *Staphylococcus aureus*; SP, *Streptococcus pneumoniae*.

^aTests adopted in each composite RS group (shown in a row) are coded as "1" and those not adopted as "0".

^bData on good-quality specimens for detecting SP were extractable for RS based on sputum culture only.

^cData were based on the composite RS excluding sputum Gram stain.

QUADAS-2	Risk of bias				Concerns about applicability		
	D1: Patient Selection	D2: Index Test	D3: Reference Standard	D4: Flow and Timing	D1: Patient Selection	D2: Index Test	D3: Reference Standard
Studies that assessed diagnostic accuracy							
Merrill 1973 [17]	?	☹	☹	?	?	☹	☹
Thorsteinsson 1975 [18]	?	?	☹	?	?	?	☹
Rein 1978 [19]	☺	☺	☹	☺	☺	☺	☹
Boerner 1982 [20]	☹	☹	☹	☹	☹	☹	☹
Dans 1984 [21]	☹	☹	☹	?	☺	☹	☹
BTS 1987 [22]	☹	?	☹	☹	☺	?	☹
Zhang 1988 [23]	☹	☺	☹	?	☺	?	☹
Gleckman 1988 [24]	☹	☺	☹	☹	☹	?	☹
Fine 1991 [25]	☹	☺	☹	☹	☺	☺	☹
Bohte 1996 [26]	☹	☺	☹	☹	☺	?	☹
Roson 2000 [27]	☺	☺	☹	☺	☺	☺	☹
Sato 2002 [28]	☹	?	☹	☹	☺	☺	☹
Butler 2003 [29]	☹	☺	☹	☹	☺	?	☹
Garcia 2004 [30]	☺	☺	☹	☺	☺	?	☹
Roson 2004 [31]	☺	☺	☹	☺	☺	?	☹
Yang 2005 [32]	☺	☺	☹	☺	☺	?	☹
Miyashita 2008 [33]	☺	☺	☹	☺	☺	☺	☹
Anevlavis 2009 [34]	☹	☺	☹	☹	☹	☺	☹
Ferre 2011 [35]	☹	☺	☹	?	☺	☺	☹
Fukushima 2013 [36]	?	?	☹	?	☺	?	☹
Akter 2014 [37]	☺	?	☹	☺	☺	?	☹
Fukuyama 2014 [38]	☺	☹	☹	☺	☺	☹	☹
Studies that assessed diagnostic yield							
Lim 1989 [39]	☺	?	☹	☺	☺	?	☹
van der Eerden 2005 [40]	?	?	☹	☺	☺	?	☹

Key

☺	Low risk of bias or concern about applicability
☹	High risk of bias or concern about applicability
?	Unclear risk of bias or concern about applicability

Figure 2. Risk of bias and concerns regarding applicability of included studies. Abbreviations: BTS, British Thoracic Society; ID, identification number; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies 2.

adopted in each study were categorized into the operationalized groups and the informative priors were additionally used (Supplementary Table 6). The results were also similar

when alternative prior distributions for the between-study random-effects parameters were utilized (Supplementary Table 6).

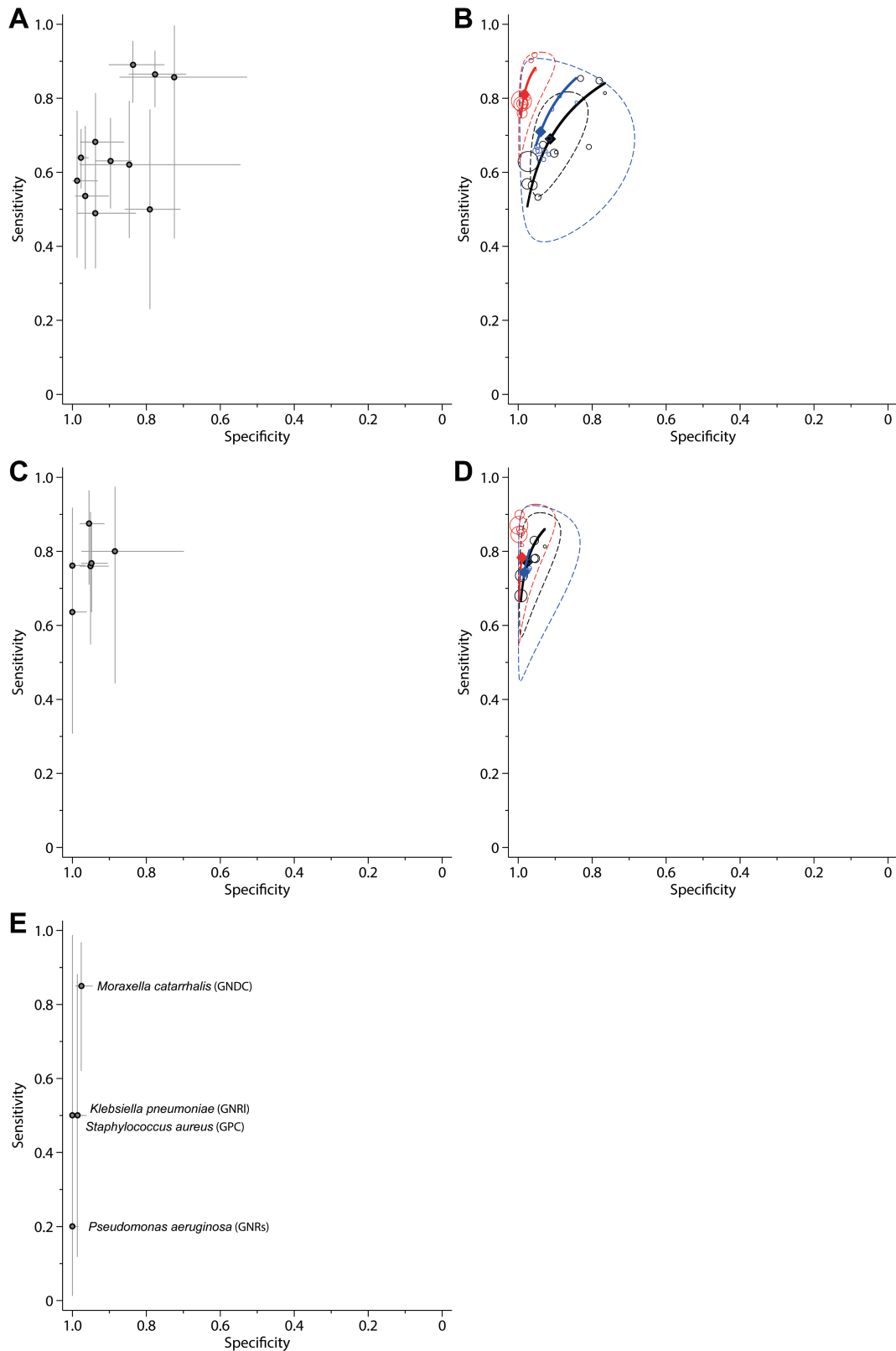


Figure 3. Diagnostic accuracy of good-quality sputum specimens for the diagnosis of *Streptococcus pneumoniae* (A and B), *Haemophilus influenzae* (C and D), and other bacteria (E). Cross-hair receiver operating characteristic (ROC) plots (A, C, and E) show reported prior-point estimates (shown as circles), and confidence intervals (shown as extended lines). ROC plots and hierarchical summary ROC curves (B and D) show individual study posterior-point estimates (the size of each circle is proportional to the sample size for each study). The dashed elliptical boundary represents the 95% credible region for the summary estimates (closed diamond). The standard (black) and latent-class model analyses based on the conditional dependence model (blue) and the conditional independence model (red) are presented. E, Causative bacteria and their diagnostic criteria of visual assessment (in parentheses). Abbreviations: GNDC, gram-negative diplococci; GNRI, large-sized gram-negative rods; GNRs, small-sized gram-negative rods; GPC, gram-positive cocci in clusters.

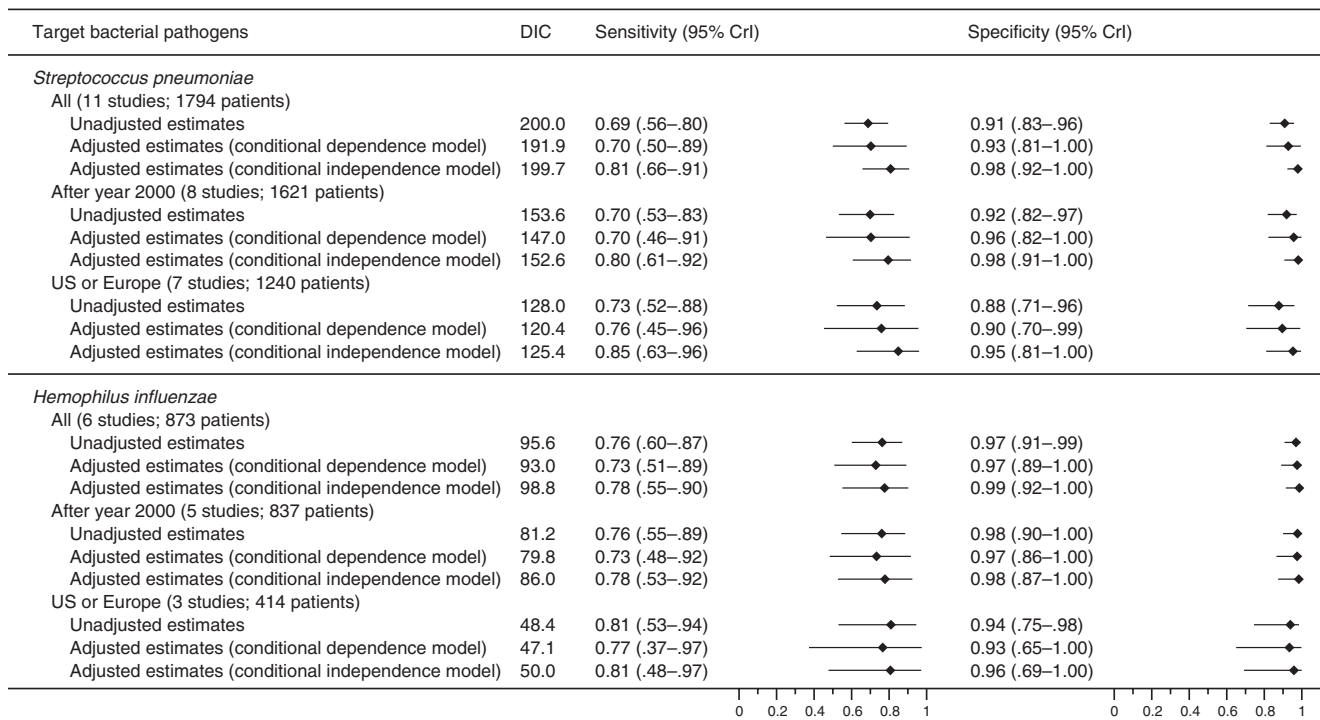


Figure 4. Summary sensitivity and specificity of sputum Gram stain in community-acquired pneumonia. Diamonds represent point estimates. Extending lines represent the 95% credible interval of each estimate. Abbreviations: CrI, credible interval; DIC, deviance information criteria; US, United States.

Diagnostic Yield and Effect on Management Decisions and Clinical Outcomes

Four studies (2 from Japan [28, 38], 1 from the Netherlands [40], and 1 from Australia [39]) contributed data of diagnostic yield. On average, sputum GS was able to successfully identify bacterial pathogens in 27% (95% CrI, 13%–48%) of all patients with CAP (regardless of whether sputum specimens were successfully obtained) and 36% (95% CrI, 22%–53%) of those with successfully expectorated sputum samples (regardless of the quality of specimens) (Figure 6). Limited heterogeneous data from 2 studies [38, 40] suggested that good-quality specimens exhibited a higher average yield of 73% (95% CrI, 26%–96%).

No study provided data on the effects on changes in patient management. Only 1 study reported the association of sputum GS-directed treatments with clinical outcomes [38] (Supplementary Table 7). However, the results were based on joint data on both CAP and HCAP.

DISCUSSION

Summary of Evidence

This meta-analysis determined that GS of good-quality sputum samples showed high specificity and PLRs to diagnose *S. pneumoniae* and *H. influenzae* as the etiology of CAP. However, the test was not very sensitive and NLR was not sufficiently low for the exclusion of these bacteria. These findings

were consistent across the subgroup, sensitivity, and adjusted analyses. Studies that have failed to select good-quality specimens reported heterogeneous accuracy estimates, consistent to those found in a previous meta-analysis [11]. Data on other bacteria were limited.

When results of both tests are not correlated, the sensitivity and specificity of an imperfect reference standard test would, in theory, underestimate the specificity and sensitivity of the index test, respectively [41]. Judging by the model-corrected low sensitivity and high specificity of the reference standards in our adjusted analysis, the observed moderate and large improvement in sensitivity and specificity, respectively, of sputum GS based on the conditional independent model are consistent with the underlying theory. Moreover, our slightly improved, adjusted estimates in the conditional dependence model are consistent with the underlying theory that naively calculated test accuracy is overestimated when the index and the reference standard tests concur to a degree greater than that expected by chance; thus, the model corrects not only underestimation (resulting from imperfect reference standards) but also overestimation (resulting from between-test correlations) [41].

Our meta-analysis of diagnostic yield found that sputum GS could diagnose bacterial pathogens in approximately one-third of patients when samples were successfully collected. Selecting good-quality specimens could increase this yield, although data

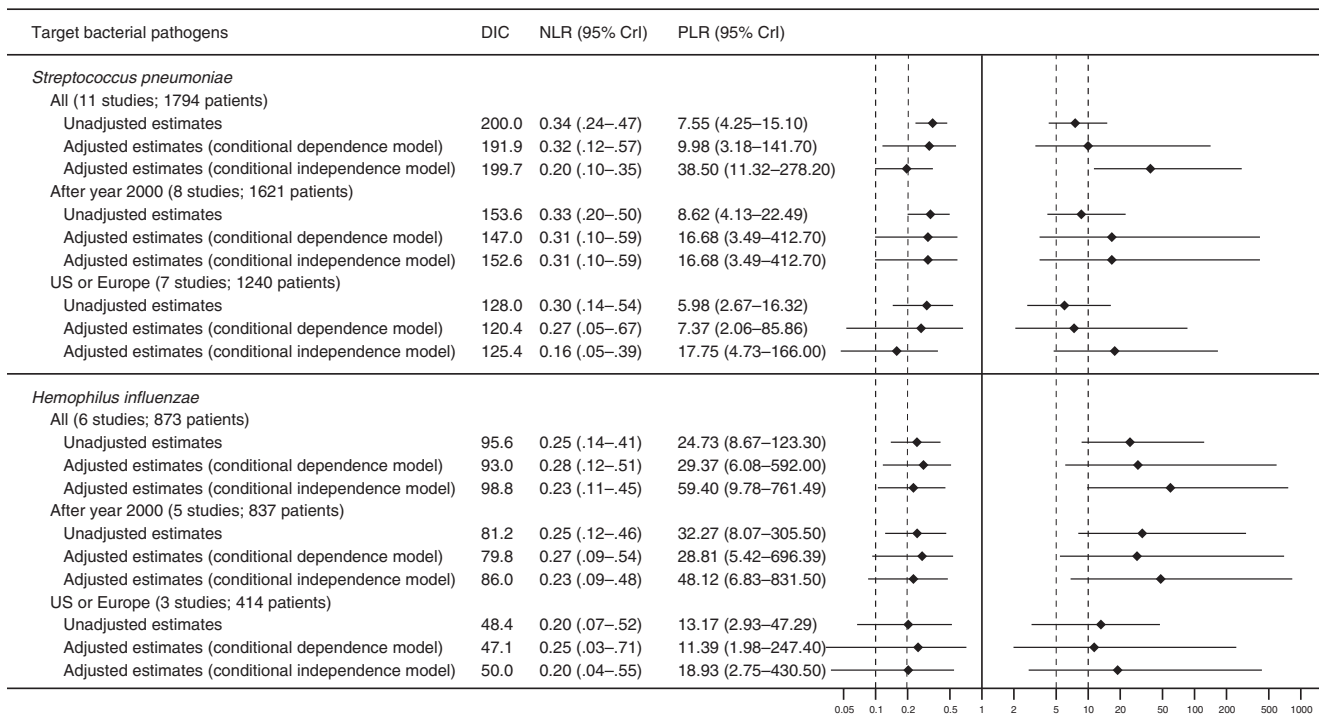


Figure 5. Summary positive and negative likelihood ratios of sputum Gram stain in community-acquired pneumonia. Diamonds represent point estimates. Extending lines represent the 95% credible interval of each estimate. Bold (drawn at 10 and 0.1) and thin (drawn at 5 and 0.2) dashed vertical lines represent clinically important thresholds that would make large and moderate shifts, respectively, in probability. Abbreviations: CrI, credible interval; DIC, deviance information criteria; NLR, negative likelihood ratio; PLR, positive likelihood ratio; US, United States.

supporting this are limited. However, findings from non-US countries were better than the yields based on both sputum GS and cultures reported from the United States [42]. The yields could depend on the prevalence of pneumonia etiologies, and thus should be carefully interpreted in countries, such as the United States, where the prevalence of pneumococcal disease is low [43].

The high PLR in our meta-analysis suggests that a positive result in GS from good-quality specimens can assist in the rapid diagnosis of *S. pneumoniae* and *H. influenzae* infection (Supplementary Figure 7). This supports the current guidelines recommending pretreatment sputum GS and culture if good-quality specimens can be obtained and quality performance measures are met [5, 7]. In addition to its low cost and wide availability, sputum GS offers the advantages of performing a wide screening for causative pathogens [5], which has the potential to optimize the initial antibiotic selection with its consistently high specificity for several particular organisms [38]. Emerging multiplex PCR tests applied directly on respiratory specimens may allow for a rapid detection of an expansive list of bacterial pathogens with ultrarapid turnaround times (1–2 hours) [9, 10]. On the other hand, up to 40% of patients are unable to produce sputum in a timely manner [4, 30], thereby rendering urinary antigen tests

more attractive as an alternative to detect *S. pneumoniae* and *L. pneumophila* serogroup 1. Urinary tests are advantageous in their rapidity, simplicity, and ability to detect antigens even days after the initiation of empiric antibiotic therapies. Nonetheless, the lack of organism isolates for in vitro susceptibility testing with reliance on multiplex PCRs or urinary antigens can limit the ability to narrow initial empiric antibiotic therapies.

There are limitations to this study. Variations in the pretest disease duration, sample collection methods, transport, and processing as well as in the experience of the interpreters explain the variable accuracy of previously reported values [5]. We were not able to specify important characteristics owing to a lack of data.

Antibiotic therapies prior to culture specimen acquisition can increase the number of false-negative results [44], which will affect the accuracy of sputum GS testing. Sparse data precluded the assessment of bias in the results due to pretest administration of antibiotics.

Our LCM meta-analysis assumed a common “average” accuracy for the composite reference standard for all participants in each study. However, the test selection should be individualized depending on, for example, the severity of pneumonia, because sicker patients might be more likely to receive more invasive

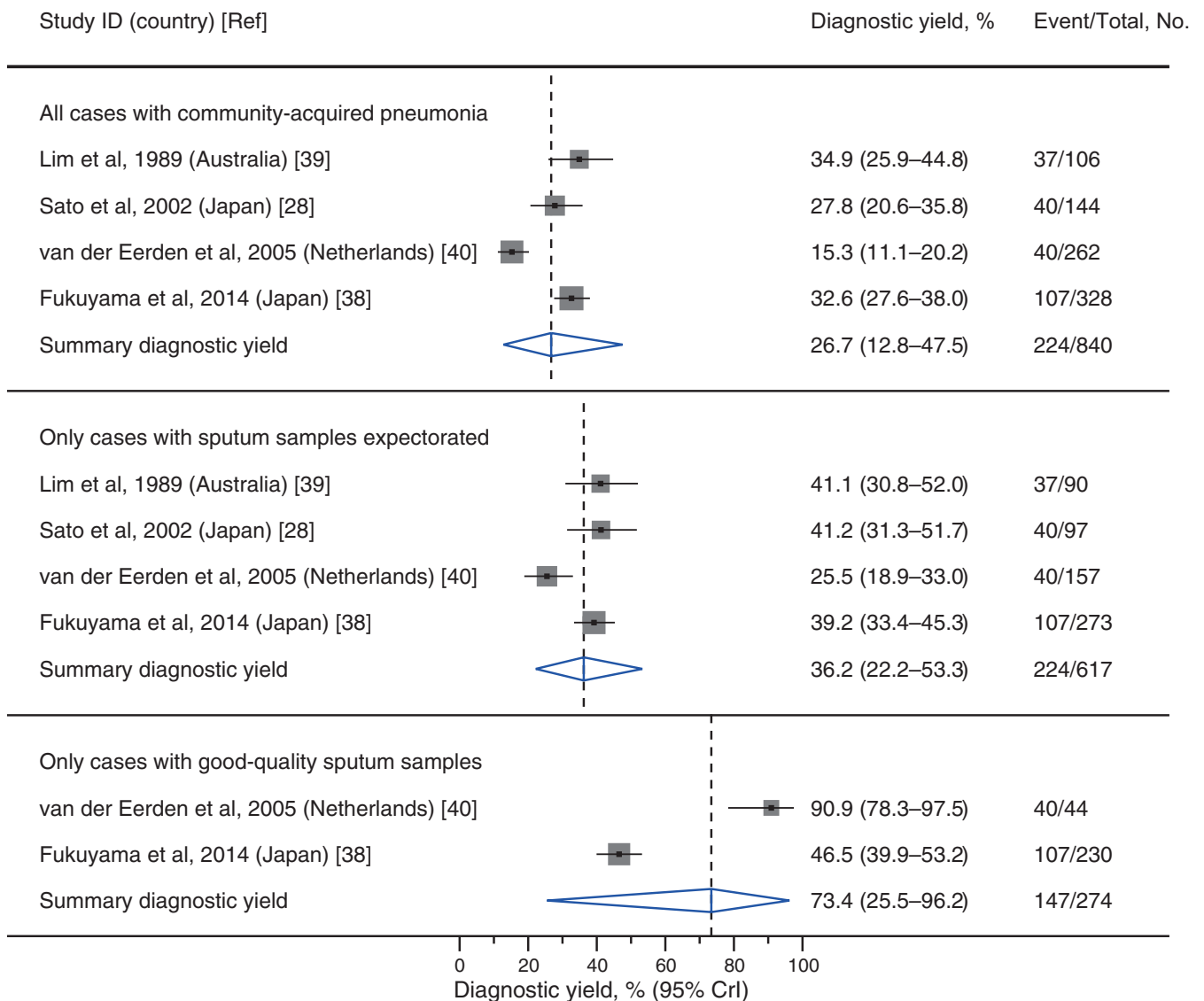


Figure 6. Summary diagnostic yield of sputum Gram stain in community-acquired pneumonia. Closed squares represent reported point estimates. Extended lines represent the 95% confidence interval of each estimate. Diamonds depict the meta-analytic result, with the center of the diamond and dashed vertical line representing the summary estimate and the width of the diamond representing the 95% CrI. Abbreviations: CrI, credible interval; ID, identification number.

and accurate tests, causing variations in the accuracy of the individual reference standards. This was not addressable without access to the individual-level data.

CONCLUSIONS

To diagnose *S. pneumoniae* and *H. influenzae* infections in patients with CAP, sputum GS of good-quality specimens is a highly specific test that can provide actionable information for pathogen-directed therapies. Yet, evidence on improved outcomes beyond accuracy or yield is sparse. Given the anticipated wider introduction of highly accurate molecular-based point-of-care tests into clinical practice [45], future research should empirically examine the effect of sputum GS, along with other point-of-care tests, on the management decisions,

patient-relevant clinical outcomes, and other broader epidemiological measures of antibiotic resistance.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. The authors thank Drs Hajime Fukuyama and Samuel Yang for providing unpublished data from their group's original work and Drs Nandini Dendukuri and Ian Schiller for providing their original codes in R to depict adjusted summary receiver operating characteristic curves and credible regions for adjusted summary sensitivity and specificity, which the authors modified and ported to Stata.

Disclaimer. The funding sources had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Financial support. This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology of Japan (grant numbers 26461518 and 26460755 to T. T.) and the National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (grant number K23 HL139981 to G. D. K.).

Potential conflicts of interest. G. D. K. has received grants from the NHLBI and Karius, Inc. M. I. has received personal fees from Otsuka Pharmaceutical Co, Ltd, Astellas Pharma Inc, Eisai Co, Ltd, and Daiichi Sankyo Co, Ltd. All other authors report no potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **2017**; 390:1151–210.
2. Jain S, Self WH, Wunderink RG, et al; CDC EPIC Study Team. Community-acquired pneumonia requiring hospitalization among U.S. adults. *N Engl J Med* **2015**; 373:415–27.
3. Musher DM, Roig IL, Cazares G, Stager CE, Logan N, Safar H. Can an etiologic agent be identified in adults who are hospitalized for community-acquired pneumonia: results of a one-year study. *J Infect* **2013**; 67:11–8.
4. van der Eerden MM, Vlasplolder F, de Graaff CS, et al. Comparison between pathogen directed antibiotic treatment and empirical broad spectrum antibiotic treatment in patients with community acquired pneumonia: a prospective randomised study. *Thorax* **2005**; 60:672–8.
5. Mandell LA, Wunderink RG, Anzueto A, et al; Infectious Diseases Society of America/American Thoracic Society. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis* **2007**; 44(Suppl 2):S27–72.
6. Eccles S, Pincus C, Higgins B, Woodhead M; Guideline Development Group. Diagnosis and management of community and hospital acquired pneumonia in adults: summary of NICE guidance. *BMJ* **2014**; 349:g6722.
7. Woodhead M, Blasi F, Ewig S, et al; Joint Taskforce of the European Respiratory Society and European Society for Clinical Microbiology and Infectious Diseases. Guidelines for the management of adult lower respiratory tract infections—full version. *Clin Microbiol Infect* **2011**; 17(Suppl 6):E1–59.
8. Skerrett SJ. Diagnostic testing for community-acquired pneumonia. *Clin Chest Med* **1999**; 20:531–48.
9. Leber AL, Everhart K, Daly JA, et al. Multicenter evaluation of Biofire FilmArray respiratory panel 2 for detection of viruses and bacteria in nasopharyngeal swab samples. *J Clin Microbiol* **2018**; 56:e01945–17.
10. Gadsby NJ, McHugh MP, Russell CD, et al. Development of two real-time multiplex PCR assays for the detection and quantification of eight key bacterial pathogens in lower respiratory tract infections. *Clin Microbiol Infect* **2015**; 21:788.e1–13.
11. Reed WW, Byrd GS, Gates RH Jr, Howard RS, Weaver MJ. Sputum Gram's stain in community-acquired pneumococcal pneumonia. A meta-analysis. *West J Med* **1996**; 165:197–204.
12. Dendukuri N, Schiller I, Joseph L, Pai M. Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics* **2012**; 68:1285–93.
13. McInnes MDF, Moher D, Thoms BD, et al; the PRISMA-DTA Group. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* **2018**; 319:388–96.
14. Ogawa H, Kitsios G, Iwata M, Terasawa T. Sputum Gram stain for diagnosing causative bacterial pathogens and guiding antimicrobial therapies in community-acquired pneumonia: a systematic review and meta-analysis protocol. *Fujita Med J* **2018**. Available at: https://www.jstage.jst.go.jp/article/fmj/advpub/0/advpub_2018-019/_article/-char/ja/. Accessed 3 June 2019.
15. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* **2011**; 155:529–36.
16. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med* **2010**; 29:3046–67.
17. Merrill CW, Gwaltney JM Jr, Hendley JW, Sande MA. Rapid identification of pneumococci. Gram stain vs. the Quellung reaction. *N Engl J Med* **1973**; 288:510–2.
18. Thorsteinsson SB, Musher DM, Fagan T. The diagnostic value of sputum culture in acute pneumonia. *JAMA* **1975**; 233:894–5.
19. Rein MF, Gwaltney JM Jr, O'Brien WM, Jennings RH, Mandell GL. Accuracy of Gram's stain in identifying pneumococci in sputum. *JAMA* **1978**; 239:2671–3.
20. Boerner DF, Zwadyk P. The value of the sputum Gram's stain in community-acquired pneumonia. *JAMA* **1982**; 247:642–5.
21. Dans PE, Charache P, Fahey M, Otter SE. Management of pneumonia in the prospective payment era. A need for more clinician and support service interaction. *Arch Intern Med* **1984**; 144:1392–7.
22. Community-acquired pneumonia in adults in British hospitals in 1982–1983: a survey of aetiology, mortality, prognostic factors and outcome. The British Thoracic Society and the Public Health Laboratory Service. *Q J Med* **1987**; 62:195–220.
23. Zhang XP, Deng KE, Ye YQ, Luo WT. Rapid detection of pneumococcal antigens in sputa in patients with community-acquired pneumonia by coagglutination. *Med Microbiol Immunol* **1988**; 177:333–8.
24. Gleckman R, DeVita J, Hibert D, Pelletier C, Martin R. Sputum Gram stain assessment in community-acquired bacteremic pneumonia. *J Clin Microbiol* **1988**; 26:846–9.
25. Fine MJ, Orloff JJ, Rihs JD, et al. Evaluation of housestaff physicians' preparation and interpretation of sputum Gram stains for community-acquired pneumonia. *J Gen Intern Med* **1991**; 6:189–98.
26. Bohte R, Hermans J, van den Broek PJ. Early recognition of *Streptococcus pneumoniae* in patients with community-acquired pneumonia. *Eur J Clin Microbiol Infect Dis* **1996**; 15:201–5.
27. Rosón B, Carratalà J, Verdagué R, Dorca J, Manresa F, Gudiol F. Prospective study of the usefulness of sputum Gram stain in the initial approach to community-acquired pneumonia requiring hospitalization. *Clin Infect Dis* **2000**; 31:869–74.
28. Sato T, Aoshima M, Ohmagari N, Tada H, Chohnabayashi N. Usefulness of sputum Gram staining in community-acquired pneumonia. *Nihon Kokyuki Gakkai Zasshi* **2002**; 40:558–63.
29. Butler JC, Bosshardt SC, Phelan M, et al. Classical and latent class analysis evaluation of sputum polymerase chain reaction and urine antigen testing for diagnosis of pneumococcal pneumonia in adults. *J Infect Dis* **2003**; 187:1416–23.
30. García-Vázquez E, Marcos MA, Mensa J, et al. Assessment of the usefulness of sputum culture for diagnosis of community-acquired pneumonia using the PORT predictive scoring system. *Arch Intern Med* **2004**; 164:1807–11.
31. Rosón B, Fernández-Sabé N, Carratalà J, et al. Contribution of a urinary antigen assay (Binax NOW) to the early diagnosis of pneumococcal pneumonia. *Clin Infect Dis* **2004**; 38:222–6.
32. Yang S, Lin S, Khalil A, et al. Quantitative PCR assay using sputum samples for rapid diagnosis of pneumococcal pneumonia in adult emergency department patients. *J Clin Microbiol* **2005**; 43:3221–6.
33. Miyashita N, Shimizu H, Ouchi K, et al. Assessment of the usefulness of sputum Gram stain and culture for diagnosis of community-acquired pneumonia requiring hospitalization. *Med Sci Monit* **2008**; 14:CR171–6.
34. Anevlavis S, Petrolou N, Tzavaras A, et al. A prospective study of the diagnostic utility of sputum Gram stain in pneumonia. *J Infect* **2009**; 59:83–9.
35. Ferré LC, Llopis RF, Jacob J, et al. Is sputum Gram staining useful in the emergency department's management of pneumonia? *Emergencias* **2011**; 23:108–11.
36. Fukushima K, Nakamura S, Takayanagi N, et al. Usefulness of a sputum antigen detection kit for *Streptococcus pneumoniae* in the rapid diagnosis of community-acquired pneumonia: comparison with urinary antigen detection test and sputum Gram stain [in Japanese]. *Nihon Kokyuki Gakkai* **2013**; 2:343–8.
37. Akter S, Shamsuzzaman SM, Jahan F. Community acquired bacterial pneumonia: aetiology, laboratory detection and antibiotic susceptibility pattern. *Malays J Pathol* **2014**; 36:97–103.
38. Fukuyama H, Yamashiro S, Kinjo K, Tamaki H, Kishaba T. Validation of sputum Gram stain for treatment of community-acquired pneumonia and healthcare-associated pneumonia: a prospective observational study. *BMC Infect Dis* **2014**; 14:534.
39. Lim I, Shaw DR, Stanley DP, Lumb R, McLennan G. A prospective hospital study of the aetiology of community-acquired pneumonia. *Med J Aust* **1989**; 151:87–91.
40. van der Eerden MM, Vlasplolder F, de Graaff CS, Groot T, Jansen HM, Boersma WG. Value of intensive diagnostic microbiological investigation in low- and high-risk patients with community-acquired pneumonia. *Eur J Clin Microbiol Infect Dis* **2005**; 24:241–9.
41. Trikalinos TA, Balion CM. Chapter 9: options for summarizing medical test performance in the absence of a "gold standard." *J Gen Intern Med* **2012**; 27(Suppl 1):S67–75.

42. Ewig S, Schlochtermeier M, Göke N, Niederman MS. Applying sputum as a diagnostic tool in pneumonia: limited yield, minimal impact on treatment decisions. *Chest* **2002**; 121:1486–92.
43. Musher DM, Abers MS, Bartlett JG. Evolving understanding of the causes of pneumonia in adults, with special attention to the role of pneumococcus. *Clin Infect Dis* **2017**; 65:1736–44.
44. Musher DM, Montoya R, Wanahita A. Diagnostic value of microscopic examination of Gram-stained sputum and sputum cultures in patients with bacteremic pneumococcal pneumonia. *Clin Infect Dis* **2004**; 39:165–9.
45. Zumla A, Al-Tawfiq JA, Enne VI, et al. Rapid point of care diagnostic tests for viral and bacterial respiratory tract infections—needs, advances, and future prospects. *Lancet Infect Dis* **2014**; 14:1123–35.