Rapid influenza diagnostic tests:

a meta-analysis of 127 studies

Caroline Chartrand

Department of Epidemiology, Biostatistics & Occupational Health

McGill University, Montreal, Canada

May 2011

A thesis submitted to McGill University in partial fulfillment of the

requirements of the degree of Master of Science

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABBREVIATIONS AND ACRONYMS

AUC – Area under the curve

C - Celsius

CDC – US Centers for Disease Control and Prevention

CI – Confidence interval

CLIA – Clinical laboratory improvement amendments

DOR – Diagnostic odds ratio

FDA – Food and drug administration

FN – False negative

FP – False positive

H – Hemagglutinin

HSROC – Hierarchical summary receiver operating characteristics

IDSA – Infectious Disease Society of America

IF - Immunofluorescence

ILI – Influenza-like illness

LR+ - Positive likelihood ratio

LR - - Negative likelihood ratio

MCDK – Madin-Darby canine kidney

N – Neuraminidase

NA – Nasal aspirate

NS – Nasal swab

NW – Nasal wash

NPA – Nasopharyngeal aspirate

NPS – Nasopharyngeal swab

NPW – Nasopharyngeal wash

PMK – Primary rhesus monkey

POC – Point-of-care

RDOR – Relative diagnostic odds ratio

RIDT – Rapid influenza diagnostic test

RNA- Ribonucleic acid

RT-PCR – Reverse-transcriptase polymerase chain reaction

SROC – Summary receiver operating characteristics

TAT – Turn-around-time

TN – True negative

TP- True positive

TS – Throat swab

WHO – World Health Organization

ABSTRACT (ENGLISH)

BACKGROUND: Timely diagnosis of influenza is important to administer appropriate antiviral therapy, institute proper infection control measures, and decrease ancillary test usage. While viral culture or reverse-transcriptase polymerase chain reaction (RT-PCR) are considered the most accurate diagnostic tests, a vast array of rapid influenza diagnostic tests (RIDTs) is available and could potentially impact patient management at the point-of-care.

OBJECTIVE: To summarize, using meta-analysis, available evidence on the diagnostic accuracy of RIDTs compared to a reference standard in adults and children with influenza-like illness and to evaluate patient and test factors associated with higher accuracy.

METHODS: Four databases (MEDLINE, EMBASE, Biosis, Web of Science) were searched up to and including September 2010 for studies on RIDTs' accuracy compared to a reference standard of either RT-PCR (first choice) or viral culture. Sensitivity and specificity were pooled using a bivariate random effects regression model and investigation of heterogeneity was done using subgroup analyses and meta-regression using an extension of the summary receiver operating characteristic curve (SROC) model.

RESULTS: A total of 100 articles, comprising 127 studies were identified. The pooled sensitivity of all RIDTs was 64.5% (95% CI: 60.6, 68.6), while the pooled specificity was 98.1% (95% CI: 97.3, 98.6). Sensitivity estimates were highly heterogeneous. Some of this heterogeneity was explained by significantly higher sensitivity in children (71.1%, 95% CI: 65.6, 76.1) than in adults (51.6%, 95% CI: 43.9, 59.1). Virus type also accounted for some of the heterogeneity in sensitivity (68.0%, 95% CI: 62.3, 73.1, for influenza A versus

51.8%, 95% CI: 42.8, 60.6, for influenza B) as well as the circulating strain of influenza A (56.9%, 95% CI: 50.9, 62.6, for influenza A/H1N1/2009 versus 72.8%, 95% CI: 65.9-78.8, for other seasonal influenza A strains). Finally, RIDTs performed better when compared against culture as the reference standard (sensitivity of 71.0%, 95% CI: 65.9, 75.6) than when compared against RT-PCR (sensitivity of 56.0%, 95% CI: 49.7, 62.1). Few studies reported duration of symptoms before testing, but studies that did showed a trend toward better accuracy at 24-48h with a rapid decline thereafter. When a meta-regression was conducted including several study-level covariates, only age remained significant with a relative diagnostic odds ratio (RDOR) of 2.67 (95% CI: 1.17, 6.11) for children versus adults.

CONCLUSION: RIDTs have modest sensitivity and high specificity, but heterogeneity in sensitivity is a concern. While they are more accurate in children than adults, and for influenza A compared to influenza B, these factors do not completely explain the heterogeneity in sensitivity. Because of their high specificity, RIDTs may be useful to rule in influenza. However, a negative test cannot be used to rule out influenza and should be confirmed by one of the reference standard tests. Further work is needed to summarize the clinical impact of RIDTs on patient management and patient-important outcomes.

ABSTRACT (FRENCH)

INTRODUCTION: Poser rapidement le diagnostic d'influenza permet d'administrer une thérapie antivirale appropriée, de débuter en temps opportun des mesures de prévention des infections et de diminuer le recours à d'autres tests diagnostiques. Bien que la culture virale et le RT-PCR demeurent les outils diagnostiques les plus fiables, il existe une vaste gamme de tests de diagnostic rapide de l'influenza (TDRI) pouvant potentiellement avoir un impact sur la prise en charge des patients.

OBJECTIFS: Résumer, par le biais d'une méta-analyse, l'ensemble des données disponibles sur la sensibilité et la spécificité des TDRIs comparés à un test de référence, chez les adultes et les enfants souffrant d'un syndrome d'allure grippal, ainsi qu'évaluer les facteurs liés au test ou au patient qui sont associés à une plus grande fiabilité.

MÉTHODES: Nous avons cherché à travers quatre bases de données (PubMed, EMBASE, Biosis, Web of Science), jusqu'en septembre 2010, pour des études sur la fiabilité des TDRIs comparés au RT-PCR (1er choix) ou à la culture virale. Nous avons méta-analysé la sensibilité et spécificité des TDRIs au moyen d'un *bivariate random effect regression model* et tenté d'expliquer l'hétérogénéité des résultats au moyen d'analyses de sous-groupes et d'une méta-régression, via une extension du modèle SROC (*summary receiver operating characteristic curve*).

RÉSULTATS: Nous avons identifiés 100 articles, comprenant 127 études. La sensibilité globale des TDRIs était de 64.5% (95% CI: 60.6, 68.6), alors que leur spécificité globale était de 98.1% (95% CI: 97.3, 98.6). Par contre, on a retrouvé une grande hétérogénéité au niveau de la sensibilité. Une partie de cette hétérogénéité pourrait être expliquée par une sensibilité

significativement plus élevée lorsque le test est utilisé chez les enfants (71.1%, 95% CI: 65.6, 76.1) plutôt que chez les adultes (51.6%, 95% CI: 43.9, 59.1). La sensibilité des TDRIs variait également en fonction du type de virus (68.0%, 95% CI: 62.3, 73.1, pour l'influenza A versus 51.8%, 95% CI: 42.8, 60.6, pour l'influenza B) ainsi que de la souche d'influenza A en circulation (56.9%, 95% CI: 50.9, 62.6, pour l'influenza A/H1N1/2009 versus 72.8%, 95% CI: 65.9-78.8, pour les autres souches saisonnières d'influenza A). Finalement, les TDRIs affichaient une meilleure performance lorsque comparés à la culture virale (sensibilité: 71.0%, 95% CI: 65.9, 75.6) plutôt qu'au RT-PCR (sensibilité: 56.0%, 95% CI: 49.7, 62.1). Peu d'études ont évalué l'effet de la durée des symptômes sur la fiabilité des TDRIs, mais les quelques études qui se sont penchées sur le sujet tendaient à démontrer une meilleure sensibilité 24-48h après le début des symptômes suivi d'un déclin rapide. Lorsque plusieurs de ces variables furent analysées en même temps, au moyen d'une méta-régression, seulement l'âge est demeuré significativement associé à la fiabilité des TDRIs, avec un rapport de cotes diagnostiques de 2.67 (95% CI: 1.17, 6.11) pour les enfants versus les adultes.

CONCLUSION: Les TDRIs ont une sensibilité modeste et une bonne spécificité, mais une grande hétérogénéité au niveau de la sensibilité demeure une préoccupation. Bien que les TDRIs soient plus fiables chez les enfants que chez les adultes et pour détecter l'influenza A versus l'influenza B, ces facteurs ne suffisent pas à expliquer l'hétérogénéité notée au niveau de la sensibilité. Puisqu'ils sont très spécifiques, les TDRIs sont utiles pour confirmer le diagnostic d'influenza. Cependant, un TDRI négatif n'est pas suffisant pour infirmer le diagnostic d'influenza et devrait être confirmé au moyen d'un des tests de référence. D'autres études sont nécessaires pour résumer l'impact clinique des TDRIs sur la prise en charge des patients.

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the work of Lorie Kloda, librarian at the McGill University Life Sciences library, who helped me devise the electronic search for this systematic review. I would also like to thank Jessica Minion for her role as the second reviewer as well as her contribution to the protocol development and her helpful input and feedback during the analysis and writing of this manuscript. Thank you to Timothy Brewer, thesis committee member, for his helpful advice.

Finally, I would like to thank Madhukar Pai, my thesis supervisor, for his invaluable help and support through every steps of this systematic review, from protocol development to data analysis and manuscript writing.

INTRODUCTION AND REVIEW OF THE LITERATURE

About the virus

Influenza is an enveloped RNA virus belonging to the Orthomyxoviridae family. There are three types of influenza virus (A, B and C), but only the first two are responsible for the seasonal influenza epidemics and the majority of human cases. Influenza C is a minor agent causing sporadic mild illness mostly in children and is not included in current influenza vaccine preparations, nor is it targeted by available diagnostic tools(1). Influenza A can be further classified into subtypes based on the hemagglutinin (H) and neuraminidase (N) glycoproteins present on its membrane. Currently, the most common circulating subtypes are H1N1, H1N2 and H3N2(1). Each of these subtypes encompasses many different strains. The different strains of influenza A and B are the products of minor antigenic variations called *antigenic drift*, a continuous process that gives rise to the annual influenza epidemics. Influenza A, however, is the only type that can cause global pandemics because of *antigenic shift*, major changes in the hemagglutinin, or neuraminidase glycoproteins, producing novel infectious subtypes for which individuals may have no prior immunological protection(1, 2).

Epidemiology of flu

Yearly epidemics of influenza occur worldwide, usually in autumn or winter in temperate climates. The global impact of these epidemics is difficult to ascertain with precision since many of those infected do not seek medical care or are not tested for influenza. Rough estimates based on physician visits for respiratory illness shows that 5 to 20% of United States residents contract influenza annually(3). Although a majority of infected individuals exhibit a self-limiting, acute febrile illness or are asymptomatic, severe disease and complications may occur, especially in the elderly, young children and individuals with underlying medicals conditions such as

pulmonary or cardiac disease, diabetes or immunosuppression(4). In children, influenza accounts for an average of 6 to 15 outpatient visits per 100 children per year and 3 to 9 courses of antibiotics per 100 children per year (5). Worldwide, three to five million individuals develop severe illness each year and 250 000 to 500 000 die of influenza-related causes(6). Even in developed countries like the United States, influenza is responsible for over 200 000 hospitalizations annually and between 3 000 and 49 000 deaths(7, 8). Moreover, as illustrated by the recent 2009 H1N1 pandemic that affected 214 countries and resulted in at least 18 000 deaths(9), influenza has the potential to rapidly spread globally. As pandemics go, however, the 2009 one was relatively mild compared to some of the previous pandemics, such as the Spanish influenza of 1918 which killed between 20-100 million people or the Hong Kong influenza of 1968 which was responsible for one million deaths(10).

Clinical Aspects

Early identification of influenza as the cause of an acute febrile respiratory illness is important for optimal patient management, allowing for more appropriate use of outpatient services and ancillary tests, prompt institution of infection control measures, and timely administration of antiviral treatments in high-risk patients. Studies in children with fever and/or respiratory illness during the influenza season have shown that a physician's knowledge of a patient's influenza status at the time of the patient visit can decrease additional testing (like complete blood count, chest X-ray, blood culture or urinanalysis)(11-15), reduce antibiotics prescriptions(12, 13, 15, 16), and potentially decrease time spent in the emergency department(12).

Successful management of institutional outbreaks or prevention of nosocomial spread of influenza, as well as proper treatment of infected patients also depends on prompt identification of influenza as the etiology(17). Current antiviral treatments for influenza, adamantanes and

neuraminidases inhibitors, have been shown to reduce the duration of clinical illness by approximately one day, reduce the severity of illness in hospitalized patients, and decrease influenza-related complications (e.g. bacterial or viral pneumonia or exacerbation of chronic diseases)(4, 18). As such, they are recommended in patients who are hospitalized, have severe, complicated or progressive illness, or are at high risk for influenza complications. However, they must be instituted within 48 hours of symptom onset to produce maximum benefits(3, 4). Since most patients are likely to present to medical attention towards the end of this time frame, it leaves clinicians with a very narrow window of opportunity to make the correct diagnosis and start the antiviral treatment.

Successful prevention and control of nosocomial transmission of influenza also depends on early identification of the virus in infected patients, thus enabling appropriate isolation and cohorting measures. Studies conducted in nursing homes have shown that access to rapid diagnosis for influenza serves to reduce the severity of institutional outbreaks through decrease of the duration or the final attack rate(19, 20).

Diagnosis

Influenza presents clinically as a respiratory illness with abrupt onset of fever, chills, headache, myalgia, sore throat, and/or dry cough(1, 2). However, the clinical diagnosis of influenza is complicated by the frequent co-circulation of other viruses (respiratory syncitial virus, adenovirus, rhinovirus, parainfluenza) and bacterial agents (*Chlamydophila pneumoniae, Mycoplasma pneumoniae, Streptococcus pneumoniae* and *Legionella* sp.) that can produce similar symptoms. The case definition of influenza-like illness (ILI), defined by the US Centers for Disease Control and Prevention (CDC) and World Health Organization (WHO) as fever (temperature above 37.8 ^{o}C) and cough or sore throat(21, 22), has been shown to have a sensitivity of 64-65% and a specificity of 67%(23, 24), which makes it only slightly better than

a coin toss for the identification of influenza in a given patient. Moreover, young children are more likely to present with atypical signs and symptoms like abdominal pain or vomiting, and the above case definition has been shown to perform even worse in studies of older individuals(4) making it even more challenging to diagnose influenza based solely on clinical presentation in the population at higher risk of complications. For these reasons, physicians often turn to laboratory diagnostics when the diagnosis of influenza is likely to impact patient management.

There are three main types of diagnostic tests for influenza: virus isolation in cell culture, detection of influenza-specific RNA, and direct influenza antigen detection in respiratory specimens. Viral culture mainly relies on the identification of a cytopathic effect in a cell culture (usually primary rhesus monkey (PMK) or Madin-Darby canine kidney (MDCK) cells) inoculated with a respiratory specimen containing influenza viruses. It is the time-honored gold standard for influenza diagnosis and provides essential virologic data on strain characteristics, such as antiviral susceptibility and relatedness to vaccine strains while monitoring for the emergence of novel influenza strains. Although viral culture is vital for global influenza surveillance, it has several limitations that lessen its usefulness in the clinical setting. It requires viable viruses (and therefore may be falsely negative late in the course of illness), and the characteristic cytopathic effect may be absent even in the presence of influenza or may be induced by other viruses. More importantly, turn-around-time (TAT) for results may take 3-10 days, making it of little use for patient management. Shell vial culture is a modification of the method, in which specimens are centrifuged onto a monolayer of cells, and can produce results much faster, in 48 hours, with comparable accuracy(3, 25).

Reverse transcriptase polymerase chain reaction (RT-PCR), which detects relatively stable matrix protein genes in the influenza RNA, is considered the most sensitive and specific test for influenza, with a 2-13% higher detection

rate over culture (26). It can detect viable and non-viable viruses and can even identify specific strains, with the addition of primers targeting genes for surface antigens. It is also the most expensive and least widely available diagnostic test for influenza because of the specialized equipment and expertise required and, since samples are usually run in batches, takes in practice longer than the theoretical turn-around-time of 4-6 hours. As with all PCR-based tests, it is also prone to cross-contamination and, since it detects specific RNA sequences, available primers may fail to identify novel influenza strains resulting from major antigenic changes(3, 4, 25).

The last of the traditional influenza diagnostic tests is immunofluorescence (IF), which detects influenza antigens in respiratory specimens through fluorescent antibody staining and microscopic visualization. It has a sensitivity of 70-100% and a specificity of 80-100% compared to culture(25) and can identify viable and non-viable viruses as well as simultaneously detect other respiratory viruses. Although results can be available in as little as 2-4 hours, this test depends heavily on specimen quality and laboratory expertise and is usually not conducted outside regular laboratory hours, limiting its usefulness in a 24/7 healthcare system(3, 4).

Rapid influenza diagnostic tests

Despite their good accuracy, none of the above-mentioned conventional influenza diagnostic tests is able to provide a timely answer to the physician's question "Does my patient have influenza?" This has spurred the development of an array of commercially developed rapid influenza diagnostic tests (RIDTs) that are simple to use, give results in 15 to 30 minutes and, in some cases, can be used at the point-of-care (POC) in a routine clinical setting. These tests are usually immunochromatographic assays that detect specific influenza viral antigens in respiratory specimens, though one test (ZstatFlu test) detects viral neuraminidase activity through its catalysis of a chemical reaction(26). The targeted antigens are usually

nucleoproteins, the internal structural proteins of the influenza virus, that differ between influenza A and B viruses but tend to be highly conserved among A or B strains.

As shown in Figure 1(27), RIDTs are made of a nitrocellulose strip that can be presented as is (dipstick), or housed in plastic (cassette), or bound to thick paper (card)(28). On one end of the strip, dye-labeled free antibodies, specific for the targeted influenza antigen(s), are placed. They are then mixed with the respiratory specimen and are drawn by capillary action across the strip. At the other end are two or more lines, the test line(s), with bound antibodies specific for the same influenza antigen(s), and the control line, to which are bound antibodies targeted to the dye-labeled free antibodies, to ensure that the specimen has migrated completely across the strip. If influenza antigens are present in the respiratory specimen, some of the dye-labeled free antibodies will be trapped on the test line(s), producing a color change(28).

In general, RIDTs can either detect only influenza A, or detect influenza A and B but not distinguish between the two, or detect influenza A and B and distinguish between them. The test is positive if a color change is detected at the influenza A or B position and the control position, and the test is negative if the color change appears only at the control position. The test should be considered invalid if no color change is detected at the control position, regardless of any color change at the influenza A or B position. Most manufacturers recommend that even very faint color change should be read as positive, which may lend an element of subjectivity to these tests that usually produce dichotomous (yes or no) results.

A number of different RIDTs are currently available on the market, with different characteristics, complexity, and manufacturer-stated accuracy. Table 1(27) lists the most frequently used or researched RIDTs. Some of

these tests are waived under the 1988 Clinical Laboratory Improvement Amendments (CLIA), which means they are deemed simple, safe, and accurate enough to be licensed by the FDA for use outside of a laboratory setting, making them amenable to point-of-care use. The cost of RIDTs (around $15-20 USD per test for kit and reagents(29)) also compares favorably with laboratory-based influenza tests, such as RT-PCR or IF.

The strength of RIDTs lies in their quick turn-around time, simplicity and potential for POC use, which makes them the most likely test to directly impact patient management. However, current RIDTs seem to suffer from inconsistent accuracy with reported sensitivity ranging from as low as 10% to as high as 80%(3, 4, 26, 30), with specificity usually exceeding 90%. Even so, the Infectious Disease Society of America (IDSA), the CDC and the WHO still consider them an integral part of their diagnostic guidelines, recognizing their usefulness in patient and outbreak management, especially when other tests such as RT-PCR or IF are not readily available while cautioning against potential misdiagnosis associated with their use(3, 26, 30). Therefore, it is crucial to better understand factors that influence the accuracy of RIDTs, as this may enable care providers to more accurately judge their potential impact and appropriately tailor their use.

Review of the existing literature

To our knowledge, only two other systematic reviews have been conducted on the accuracy of RIDTs. The first one was done by Uyeki (31)in 2002 and included only pediatric studies, making it impossible to judge whether RIDTs perform differently in children versus adults. Many studies on the accuracy of RIDTs have since been published, particularly in the wake of the 2009 H1N1 pandemic, which could also dramatically change the estimated RIDTs' accuracy. The other review was published recently by Petrozzino *et al*(24) sponsored by Quidel Corporation, and therefore only addressed QuickVue, the manufacturer's test. This review was also conducted before the

emergence of the influenza A/H1N1/2009 strain, which reportedly decreased the accuracy of available RIDTs(32). Thus, a new and up-to-date systematic review of all commercial RIDTs, including studies published after the emergence of influenza A/H1N1/2009, is both justified and timely.

Objectives

The objective of this systematic review is to summarize available evidence on the diagnostic accuracy of RIDTs compared to a reference standard (either viral culture or RT-PCR) in patients with an influenza-like illness and to evaluate patient and test factors that might account for heterogeneity in reported accuracy (e.g. patient age, type of specimen, duration of illness prior to testing, etc.).

METHODS

Before this systematic review was conducted, a protocol was prepared using standard guidelines for the systematic review of diagnostic studies [Appendix 1 et 2](33, 34).

Search Strategy

To identify relevant citations, the following 4 electronic databases were searched: PubMed, EMBASE (1980-2010), BIOSIS (1969-2010) and Web of Science. The search strategy was designed with the help of an experienced librarian and contained search terms for the influenza disease or virus combined with search terms for rapid diagnostic immunoassays, including brand names for the most common commercial RIDTs. The complete electronic search strategy for PubMed is presented in Appendix 3. Studies published in either English or French were considered. The databases were searched in March 2010 and an update of the search, in PubMed only, was conducted in September 2010. Additionally, the bibliography of included studies, recent narrative reviews on RIDTs, and guidelines on influenza were hand searched for additional relevant studies. Diagnostic manufacturers were also contacted for help in identifying additional or unpublished studies.

Eligibility Criteria

Studies were considered for inclusion if they assessed the accuracy of an RIDT against one of the accepted reference standards. For the purpose of this systematic review, RIDTs were defined as any commercially available assay that identified influenza viral antigens or neuraminidase activity in respiratory specimens in 30 minutes or less, through simple immunochromatographic formats. In-house tests and pre-commercial versions were excluded since they are not available to most healthcare providers assessing patients with possible influenza, may not be

standardized, and may result in inconsistent accuracy estimates. Acceptable reference standards included viral culture or RT-PCR. If both were available, RT-PCR was chosen because of its superior sensitivity and specificity which, in turn, reduces the risk of outcome misclassification. Studies were excluded if they compared RIDTs to immunofluorescence or ELISA (since those are not widely acknowledged gold standards for influenza diagnosis), if they used the result of the RIDTs as part of a composite reference standard (incorporation bias) or if they performed the reference standard only on samples with negative RIDTs (partial verification bias). RIDTs could be tested alone against the reference standard or in parallel with other RIDTs. If a publication included more than one RIDT, each test comparison was included as a separate 'study'. Sufficient data for the construction of a two-by-two table had to be reported for inclusion in the final analysis.

We excluded conference abstracts since they usually contained insufficient information on many of the data items relevant to our analysis, such as type of respiratory specimens collected, setting, participant characteristics, POC usage, etc., as well as on issues essential for quality assessment, such as blinding, selection of patients, reporting of indeterminate results, etc. We also excluded case-control type studies (testing with the RIDT of known positive and/or negative samples) which, by creating an extreme contrast, can overestimate a test's accuracy (18 such studies were excluded) (35).

Study Selection

Following the electronic database search, titles and abstracts were screened for relevance by one reviewer (CC). Full text articles were obtained for relevant citations and assessed for eligibility by one reviewer (CC). When doubt existed about eligibility, it was discussed with a second reviewer (MP) until a consensus was reached.

Data Extraction

A data extraction form was created and piloted on a subset of 5 included articles by two reviewers (CC and JM) before being finalized. The final version of the data extraction form is shown in Appendix 4. One reviewer (CC) extracted the data from all the articles included in the final review. For validation, a second reviewer (JM) extracted data from a randomly chosen sample of 22 articles. The numbers in the extracted two-by-two tables matched exactly in 20 of the 22 articles, with minor differences for the other two articles. Attempts were made to contact the authors if information was lacking to construct the main two-by-two table or for one of the pre-specified subgroups (see below). Out of the 25 authors contacted via email, thirteen provided new data.

Data items

The following assumptions or simplifications made during data extraction deserve special mention. When two or more RIDTs were assessed in parallel, each RIDT was considered to represent a separate study, within the same published article. Hence, the total number of articles and the total number of studies are not equal and the latter will be considered to be the denominator when assessing the proportion of studies presenting one characteristics or another. For the reference standards, both traditional viral culture and shell vial culture were considered together, regardless of the cell line used or variation in techniques. Similarly, RT-PCR was considered as a whole, independently of the platform or specific primers used. Children were defined as individuals younger than 18 years. The study population was considered to be mostly pediatric or mostly adults if 85% of individuals were below or above that cut-off, respectively. However, in mixed-study populations with separate results for children and adults, the cut-off used by the authors, which usually varied from 15 to 21 years old, was used.

Point-of-care testing was defined as a test conducted at the patient bedside (or in a routine clinic setting), immediately upon respiratory specimen acquisition. When studies failed to mention when and where the RIDT was done, it was presumed not to have been done at the point-of-care. Blinding of either the reference standard or the RIDT was acknowledged if it was specifically stated in the article, or if the RIDTs were performed at the point-of-care, since their results were available immediately and blinding to the reference standard results was assumed. Finally, the subtypes of influenza A circulating when the study was conducted were either determined from the article itself or from the information available in the CDC or WHO databases pertaining to past influenza seasons(36, 37).

Assessment of Study Quality

The methodological quality of the included studies was assessed using Quality Assessment of Diagnostic Accuracy Studies (QUADAS), a validated tool to evaluate the presence of bias and variation in diagnostic accuracy studies (38). The 14 item QUADAS checklist was included within the data extraction form in Appendix 4. For this review, the spectrum of patients was considered adequate if the study took place during either the influenza season or a documented outbreak of influenza. Selection was considered appropriate if patients were recruited consecutively based on clearly defined clinical signs and symptoms or if all specimens received for influenza testing by the laboratory during a defined period were included. As previously stated, blinding was considered to have occurred if it was clearly stated in the article or, for the index test, if the RIDT was done at the point-of-care. Handling of indeterminate RIDTs' results (no control line, positive for both influenza A and B, etc.) had to be clearly described by either reporting the number of invalid tests or stating if samples were retested or discarded. Quality assessment was used to present an overall picture of the methodological quality of the included studies, as well as to conduct certain

subgroup analyses to evaluate the impact of the methodological quality on the reported accuracy (see below).

Summary measures

Data were extracted to construct two-by-two tables containing true positive (TP), false positive (FP), false negative (FN) and true negative (TN) values for each study. True positives were defined as patient specimens positive for influenza according to the RIDT under investigation that were also found to be positive for influenza according to the reference standard (either viral culture or RT-PCR). False positives were patient specimens positive according to the RIDT, but for which the reference standard was negative. True negatives were defined as specimens that were negative for influenza according to the RIDT and the reference standard, while false negatives were specimens that were negative according to the RIDT but positive according to the reference standard.

Those numbers were used to calculate the sensitivity (i.e. the proportion of specimens positive for influenza detected by the RIDT) and the specificity (i.e. the proportion of specimens negative for influenza correctly identified by the RIDT) of RIDTs compared to viral culture or RT-PCR. Positive likelihood ratio (LR+), which measure how much more often a positive test occurs in diseased versus non-diseased individuals, and negative likelihood ratio (LR-), which measure how less likely a negative test result is in diseased versus non-diseased individuals, were also calculated since they are of particular interest to clinicians. Table 2 provides detailed definitions and formulae for the various diagnostic effect measures.

Analysis

Data were analyzed using STATA/IC 11.0 (Stata Corp. Texas, USA). Forest plots were used to visually summarize the sensitivity and specificity estimates for each study, along with their 95% confidence intervals (CIs),

using the MetaDiSc software(39). Meta-analysis of data from diagnostic studies represents a particular challenge since diagnostic accuracy is usually summarized by two measures (sensitivity and specificity), rather than a single effect measure. Since these two measures are generally inversely correlated and depend on cut-off values used, the analysis must account for these issues. Two models, the hierarchical summary receiver operating characteristic (HSROC) curves and the bivariate random effects regression models (different parametrizations of the same model in absence of covariates), were used to pool accuracy measures across studies, using the user-written program "metandi" in STATA(40). Both are examples of random effects models that assume that studies included in the meta-analysis are estimating different true effects that vary around some overall central value. Random effects models thus allow for two source of variability, random error and between-studies variability, and are preferred when the data are heterogeneous, since they lead to more conservative estimates with wider confidence intervals(34).

Since sensitivity and specificity should be treated jointly and vary according to thresholds (i.e. explicit or implicit cut-off values determining positive versus negative results), HSROC curves were analyzed to explore the influence of those thresholds(41). HSROC curves use a hierarchical model to account for both within-study variation, with parameters for accuracy and asymmetry for each study underlying ROC curve, and between-study variations in accuracy and positive threshold, according to a normal distribution(40). Each study is visually represented by a point in the ROC space (with a total area of 100%), according to its sensitivity (y axis) and 1-specificity estimate (x axis), along with the summary curve of the HSROC model, obtained by fitting a regression curve through these points. The closer the curve is to the upper left hand corner of the ROC plot (sensitivity and specificity are both 100%), the better the performance of the test. The area under the HSROC curve (AUC) represents the overall accuracy of the

test, with an AUC of 50% indicating no discriminatory ability while AUC of 100% would indicate perfect discrimination between diseased and non-diseased individuals.

Sensitivity and specificity were pooled using bivariate random effects regression models. The bivariate model takes into consideration the potential trade-off between sensitivity and specificity by explicitly incorporating this negative correlation in the analysis(40), thus overcoming the problems associated with simple weighted averages of the sensitivity and specificity estimates. These models also use a random effects approach by assuming that sensitivity (and specificity) values from individual studies, after logit transformation, are approximately normally distributed around a mean value(42). This variation in underlying accuracy estimates between studies can be related to undetected differences in study population, differences in implicit threshold (cut-off), or unnoticed variations in test protocol. A bivariate normal distribution is obtained by combining those two normally distributed outcome measures (sensitivity and specificity), while acknowledging the possible correlation between the two.

Investigation of heterogeneity

Heterogeneity is usually a concern with meta-analyses and refers to a high degree of variability in accuracy estimates across studies. Heterogeneity could be due to variability in thresholds, prevalence of disease, populations studied, variations in assay methods and reference standard tests, and differences in study quality. When significant heterogeneity is present, summary estimates from meta-analyses are not meaningful. Heterogeneity of the accuracy estimates was assessed using the I^2 statistic and explored through subgroup analyses and meta-regression (see below). The I^2 statistic measures the proportion of the variability in study estimates that is attributable to heterogeneity rather than random chance(43). Values above

75% indicate high heterogeneity and make a summary measure hard to interpret.

Contrary to popular belief, test accuracy is not a fixed property of a test, and it can vary between patient subgroups, depending on factors such as spectrum of disease, type of specimen used, test interpreter, etc.(33) This heterogeneity in meta-analysis is analogous to "effect measure modification" in epidemiologic studies. Since the reported accuracy of RIDTs is known to vary greatly, with reported sensitivity ranging from 10 to 80%(30), a large between-studies variation in accuracy was expected and the following subgroup analysis were planned a priori (pre-specified) to tease out the potential sources of heterogeneity. These included:

- Population age (children versus adults). Since children have been shown to have higher and longer viral shedding than adults(4), RIDTs are expected to perform better in the pediatric population.
- Virus type (influenza A versus influenza B and subtypes of influenza A). Influenza B tends to cause a milder illness than influenza A(2), which could contribute to a lower accuracy of the RIDTs. Moreover, the antibodies used in the manufacturing of the test might have different affinity for the different types, subtypes and strains of influenza.
- Reference standard used (viral culture or RT-PCR). RT-PCR has been shown to be both more sensitive and specific than culture(3). This could affect the accuracy of RIDTs since a more accurate gold standard will reduce outcome misclassification.
- Brand of RIDT. Since different RIDTs may use different reagents and have different intrinsic cut-off points for determining positive and negative results, they might perform differently.
- Type of specimen. Although most manufacturers' package inserts list a wide variety of acceptable respiratory specimens, different types of specimens are expected to yield different amounts of respiratory

epithelial cells, and thus quantities of virus. Nasopharyngeal aspirates (NPA) or swabs (NPS) are expected to have a higher yield than nasal specimens, which in turn are expected to be better than throat swabs (TS)(3).

- Duration of symptoms before testing. Viral shedding peaks at 24 to 48 hours from illness onset and declines rapidly thereafter, becoming undetectable after 5 to 10 days(4). The accuracy of RIDTs is expected to decrease as the interval from start of illness to testing increases.
- Point-of-care testing versus laboratory testing. Experienced laboratory personnel are expected to obtain higher accuracy from these tests than clinician or nurses performing them infrequently.
- Methodological quality (such as lack of blinding, clear definition of ILI, etc.). Studies of higher quality are expected to have more consistent results.

Provided a sufficient number of studies are identified for each subgroup (at least 4 studies are necessary for the HSROC model used to converge), separate pooled sensitivity and specificity estimates along with separate HSROC curves were generated and compared among different subgroups.

It is, however, possible that the differences in accuracy found between subgroups do not reflect actual variations due to these subgroups' characteristics, but rather the influence of another variable intricately linked to the first one (i.e. confounding). For example, if RIDTs perform differently in children versus adults, it might be because of biological differences in amount of viral shedding, but it can also be due to preferential infection with more or less virulent types or subtypes of influenza. In case of the latter, difference in accuracy would be due to virus type rather than age. Multivariable meta-regression was carried out to take into account the possible interrelations between the different subgroups and adjust for confounding.

The meta-regression analysis was done using an extension of the SROC model(44), a similar model to the HSROC without the hierarchical component that allows for between-studies variations (i.e. a fixed effect model). This linear regression model, in which individual studies are the unit of analysis, allows for the inclusion of covariates to explain the variability in the dependent variable, the diagnostic odds ratio (DOR)(45). The DOR (the odds of a positive result in diseased individuals compared to the odds of a positive result in non-diseased individuals) is a single indicator of diagnostic accuracy that varies from 0 to infinity, where a high value indicates a good test performance and a value of one (null) indicates inability to discriminate between individuals with and without disease. After antilogarithmic transformation, the coefficient of each covariate can be interpreted as a relative DOR (RDOR)(45). An RDOR is the ratio of two DORs and indicates the diagnostic accuracy of a test in studies with a particular characteristic (covariate=1) relative to the accuracy of the test in studies without that characteristic (covariate=0). Values above 1.0 indicate higher diagnostic accuracy in studies with a certain characteristics than in those without it, while the reverse holds true for value below 1.0. The covariates included in the meta-regression model were chosen from the previously mentioned subgroups. However, since the resulting estimates tend to become less precise as the number of covariates included in the model increase, only the covariates showing between-subgroups variation (statistically significant or not) were included in the final model.

Reporting of the review

The PRISMA statement for the reporting of systematic reviews and meta-analyses (46) was used as the template for preparing this report.

Design issues to minimize bias in the systematic review

We used several methods to minimize bias in our systematic review. First, we used a standard, pre-specified protocol for doing our review. Second, we carried out a comprehensive search for published and unpublished studies, with the support of an experienced librarian. Third, we included studies published in two languages (English and French), and aimed to limit publication bias by searching for unpublished studies by contacting test manufacturers. Fourth, to ensure reproducibility, two reviewers were involved in data extraction. Fifth, we used rigorous methods for data analysis, including bivariate random effects models, HSROC analyses and methods for exploring heterogeneity, including meta-regression. Lastly, we used standardized tools, such as QUADAS and PRISMA, to ensure that our review meets the best standards for evidence synthesis.

RESULTS

Study selection

The selection of included studies is shown in Figure 2. The initial search yielded 2497 citations, after exclusion of duplicates. After screening titles and abstracts, 113 were eligible for full-text review. Of these, 19 articles focusing on the clinical impact of RIDTs on patient management were put aside (though 3 were later re-included because they also provided sufficient information on test accuracy). Sixteen articles were further excluded because of case-control design (3), use of an inappropriate reference standard (6), testing of known positive samples only (4), and presentation of data published in full elsewhere (3) (Table 3 provides reasons for exclusion). Finally, five articles were excluded because of insufficient information to construct a two-by-two table, despite contacting the authors. An additional 27 articles were identified through hand searches and updating of the search in PubMed, for a total of 100 included articles. Since, as previously stated, some of these articles evaluated more than one RIDT, the final analysis included 127 studies.

Characteristics of included studies

Table 3 describes the 127 included studies while Table 4 summarizes their main study-level characteristics. Most studies (50%) included a population composed of both adults and children, though 35% and 14% included only children and adults, respectively. Less than a third of the studies (30%) defined the basis on which patients or specimens were recruited and even fewer (13%) gave any information on duration of patient clinical symptoms before testing. The recent H1N1/2009 pandemic spurred a renewed interest in RIDTs, and 21% of the included studies were conducted during that pandemic.

The included studies evaluated 20 different RIDTs. Of these, the most frequently studied tests were the Binax tests (Binax NOW Flu A and Flu B (6 studies) and Binax NOW Influenza A&B (17 studies), Inverness Medical, Portland, ME, USA), the Directigen tests (Directigen Flu A (11 studies) and Directigen Flu A+B (23 studies), Becton Dickinson, Franklin Lakes, NJ, USA) and the QuickVue tests (QuickVue Influenza Test (18 studies) and QuickVue Influenza A+B (18 studies), Quidel Corp, San Diego, CA, USA). Both reference standards were used with almost equal frequency, with 53 studies (42%) using RT-PCR and 67 studies (53%) using viral culture as a reference standard (7 studies used both).

Quality of included studies

Figure 3 presents an overview of the quality of included studies, using the QUADAS criteria for quality assessment of studies of diagnostic accuracy(38). Only 1 study was conducted outside the influenza season or a declared influenza outbreak (spectrum criteria), though 13 studies (10%) did not provide a clear indication on the timing of the study. Only a third (37%) of the included studies gave a clear rationale for patient or specimen inclusion (selection criteria). While, because of our inclusion criteria, most studies were free of a partial verification, differential verification, and incorporation bias and used an appropriate gold standard; only 38% reported blinding of the evaluation of the RIDTs' result (mostly because they were evaluated at the point-of-care). Even fewer (11%) reported blinding of the evaluation of the reference standard. As well, very few studies (18%) mentioned how they handled uninterpretable RIDT results, and only 7 studies provided actual numbers of indeterminate tests.

Overall accuracy of RIDTs

Figures 4 and 5 display the forest plots of sensitivity and specificity estimates from each study. Specificity appears to be more consistent across studies than sensitivity, with sensitivity estimates ranging from 10.2% to 100% and

specificity estimates ranging from 50.5% to 100%. The forest plot of sensitivity suggests a very high degree of heterogeneity.

Overall, for RIDTs (N=127 studies) compared to one of the two acceptable reference standards, the pooled sensitivity from bivariate random effects regression was 64.5% (95% CI: 60.6, 68.6) and the pooled specificity was 98.1% (95% CI: 97.3, 98.6), with an overall LR+ of 33.2 (95% CI: 24.6, 44.9) and an overall LR- of 0.36 (95% CI: 0.32, 0.40). Figure 6 shows the overall HSROC curve, along with the accuracy estimates for each study and the overall summary point. The HSROC plot shows greater variation in sensitivity than in specificity, with only 16 studies (12.6%) reporting specificity estimates below 85%. As expected, the I^2 statistic showed high heterogeneity, with a value of 95.9%.

Subgroup analyses

To investigate this heterogeneity, subgroup analyses were conducted according to the pre-specified subgroups. The number of studies in each subgroup may differ from those reported in Table 3 since many studies reported data on particular subgroups, in addition to the overall results. For example, although some studies recruited only adults or children, many that recruited a mixed population stratified their results according to age.

Table 5 presents the accuracy estimates for the different subgroup analyses. Based on non-overlapping confidence intervals, RIDTs showed a significantly higher pooled sensitivity in children (71.1%, 95% CI: 65.6, 76.1) compared to adults (51.6%, 95% CI: 43.9, 59.1). Figure 7 shows that both HSROC curves follow a similar pattern, but the one for adults is shifted downward compared to the one for children.

Virus types and subtypes also had a significant effect on the sensitivity of RIDTs. RIDTs had a significantly increased pooled sensitivity for detecting

influenza A (68.0%, 95% CI: 62.3, 73.1) compared to influenza B (51.8%, 95% CI: 42.8, 60.6), based on non-overlapping confidence intervals (HSROC curves shown in Figure 8). As well, RIDTs performed significantly worse during the recent outbreak of pandemic influenza A/H1N1/2009, with a pooled sensitivity of 56.9% (95% CI: 50.9, 62.6) compared to a pooled sensitivity of 72.8% (95% CI: 65.9, 78.8) for influenza A in other influenza seasons (HSROC curves shown in Figure 9). No significant difference was found between seasonal subtypes of influenza A (H1N1 versus H3N2), though those estimates are based on the most prevalent circulating subtypes during the study period and don't represent stratification according to individual specimen subtypes within studies.

There was no significant difference between the different RIDTs in term of sensitivity and specificity, with considerable overlap between the different accuracy estimates (Table 5). Directigen Flu A had the highest pooled sensitivity (79.4%, 95% CI: 67.2, 87.9), followed by ZstatFlu (75.0%, 95% CI: 55.5, 87.8), and QuickVue Influenza Test (70.7%, 95% CI: 58.8, 80.4). Specificity was consistent among most RIDTs, with the exception of Flu OIA (86.1%, 95% CI: 79.9, 90.6) and ZstatFlu (89.7%, 95% CI: 82.5, 94.1), which had markedly decreased pooled specificity compared to the others. The difference was statistically significant in the case of Flu OIA.

As expected, RIDTs appeared to perform better when compared to viral culture rather than RT-PCR (pooled sensitivity of 71.0% (95% CI: 65.9,75.6) versus 56.0% (95% CI: 49.7, 62.1)), owing to the increased accuracy of the latter (HSROC plots are shown in Figure 10).

Neither the type of specimen collected from patients nor whether or not the RIDT was performed at the point-of-care had a noticeable impact on their accuracy. As well, none of the quality issues investigated (spectrum of disease, patient selection, blinding, and handling of uninterpretable results)

had a significant effect on pooled accuracy estimates. Studies sponsored by the industry tended to show higher sensitivity for the RIDTs (pooled sensitivity of 71.7% (95% CI: 64.1, 78.2) compared to 62.7% (95% CI: 58.0, 67.2) for studies not sponsored by the industry), but the difference was not statistically significant.

Effect of symptom duration on test accuracy

Only 6 studies gave information on duration of symptoms before testing, but because the time stratification that was used varied across these studies, meta-analysis was not attempted. As shown in Table 6, in the 6 studies that conducted internal subgroup comparison between patients with different time elapsed since onset of symptoms, there seemed to be a trend towards lower accuracy on the first day of symptoms, with highest sensitivity on days 2 and 3 and a rapid decline thereafter.

Meta-regression

To further investigate interrelations between the different subgroups and residual heterogeneity within subgroups, meta-regression analysis was performed by extending the SROC model. The covariates used in meta-regression have to be study-level variables (i.e. variables like influenza A versus influenza B could not be included because they arise from subgroups within each study, but studies could be classified according to whether or not they were conducted during the influenza A/H1N1/2009 pandemic). The number of covariates that can be included in the meta-regression analysis is also limited by the available number of studies, with the need for at least 10 studies for each covariate (approximate rule of thumb). For this reason, covariates like RIDTs used were not included since they resulted in too many strata with insufficient studies in each stratum. The choice of the included covariates was based on the pre-specified subgroups and those variables that showed meaningful variations in the subgroup analyses.

Table 7 presents the RDOR estimates from the meta-regression analyses. Studies conducted in children showed a nearly threefold higher DOR (RDOR 2.67, 95% CI: 1.17, 6.11) compared to studies conducted in adults. None of the other covariates included in the model (reference standard used, study conducted during the H1N1/2009 pandemic, RIDTs performed at the point-of-care, selection of patients based on a clear clinical definition, and sponsoring by the industry) reached statistical significance. Interestingly, contrarily to what was shown in the subgroup analyses, studies conducted during the 2009 H1N1 pandemic tended towards a higher diagnostic accuracy (RDOR 1.72, 95% CI: 0.83, 3.55) compared to those conducted during other influenza seasons, though the difference was not statistically significant.

DISCUSSION

Early and rapid diagnosis of influenza is important from a clinical as well as public health perspective. A large number of rapid diagnostic tests are now available, and some of them can be potentially used at the point-of-care to improve clinical decision-making. To our knowledge, this is the most comprehensive systematic review and meta-analysis on the accuracy of RIDTs and provides the most up to date synthesis of 127 studies.

Major findings

Overall, RIDTs were found to have high specificity (pooled estimate 98.1% (95% CI: 97.3, 98.6)), with modest and highly variable sensitivity (pooled estimate 64.5% (95% CI: 60.6, 68.6), with estimates ranging from 10.2% to 100%). For the clinician, this means that although false negatives are frequent, a positive test is unlikely to be a false positive result and can be acted upon immediately (a LR+ of 33 reinforces this conclusion as a large LR+ will result in a substantial increase in post-test probability which can rule in influenza based on a positive result). In the presence of a positive RIDT, a clinician can confidently make the diagnosis of influenza and start appropriate infection control measures and antiviral therapy if indicated, while forgoing unnecessary additional diagnostic testing and antibiotic prescription. However, a negative RIDT result has a high likelihood of being a false negative and needs to be confirmed by other laboratory diagnostic tests if the result is likely to impact patient management (a modest LR- of 0.36 confirms this observation).

A very important finding is that RIDTs perform significantly better in children than in adults, with a difference of almost 20% in sensitivity between the two groups. This is plausible since young children have been shown to have higher viral loads, probably due to immature immune

defenses, and longer viral shedding (up to 10 days compared to 5 days for adults)(4), both factors contributing to more virus-rich specimens in children. This increased sensitivity could also be due to the acquisition of better, but more "invasive", specimens in children, such as nasopharyngeal aspirates and swabs, than in adults who are more likely to voice their objection to such uncomfortable approaches and make clinicians opt for specimens such as throat and nasal swabs instead.

After adjusting for other factors such as reference standard used, RIDTs performed at the point-of-care, and selection of patients based on a clear clinical definition through meta-regression, RIDTs still showed significantly increased accuracy in children compared to adults. Thus, confounding is unlikely to be the sole explanation for the observed difference in test sensitivity between children and adults.

We also found that the type of influenza virus in circulation had an impact on RIDTs' accuracy. RIDTs had a significantly higher sensitivity for detecting influenza A compared to influenza B. Studies have shown that infection with influenza A/H3N2 (the most common circulating subtype of influenza A in North America in the past decades) leads to more severe disease and higher annual rates of influenza-associated hospitalization and death than infection with influenza B. Conversely, influenza A/H1N1 has been shown to have the lowest severity index and the lowest annual rates of morbidity and mortality(7, 8, 47). More severe disease usually means higher viral load and thus higher likelihood of isolating the virus from respiratory specimens. Although we did not find significant differences in the accuracy of RIDTs for detecting influenza A/H3N2 versus influenza A/H1N1, this could be explained by the fact that those subgroups were based on the most prevalent circulating subtype of influenza A during the year(s) when the study was conducted, and as such does not exclude different degrees of mixing between

the two types that could blur any real differences in the performance of RIDTs.

During the H1N1/2009 pandemic, many reports were made of even lower sensitivity of RIDTs for this new pandemic strain, compared to published accuracy estimates. Indeed, comparing the accuracy of RIDTs for influenza A in studies conducted during the pandemic versus the previous seasons, we found a 15% reduction in sensitivity. This could be explained by the nature of the virus itself since, as previously stated, H1N1 induces a milder disease than both influenza A/H3N2 and influenza B. It can also be due to the nature of a pandemic itself, where heightened awareness and panic might lead to indiscriminate testing for the virus without rigorous selection of who should be tested. Finally, although past studies have been conducted using both viral culture and RT-PCR as reference standards, all the studies conducted during the H1N1/2009 pandemic, with one exception, used RT-PCR as a reference standard. Since RIDTs will appear to perform less well when compared to this more accurate reference standard, this could have created a spurious difference in accuracy between seasonal and H1N1/2009 influenza A. Interestingly, the difference was no longer significant when adjusted for the reference standard and other variables through meta-regression, even suggesting a higher diagnostic accuracy for H1N1/2009.

Overall, no single commercial brand of RIDTs seemed to perform markedly better than the others. Directigen Flu A had the highest sensitivity, but this moderately complex laboratory-based test detects only influenza A, making it useless in seasons where influenza B represents a significant proportion of circulating virus. Flu OIA and ZstatFlu had decreased specificity compared to the other tests, perhaps making them a less desirable option since high specificity is the best asset of RIDTs.

Also, no difference in accuracy was found between the different kinds of respiratory specimens, though these analyses were limited by the absence of stratification by specimen type in most studies. Though common practice guidelines have held nasopharyngeal specimens as the best specimen type(3, 4), followed by nasal specimens and throat swabs, other studies have failed to show a difference between them(48-50). This equivalence in accuracy across respiratory specimen types is a potential argument in favor of less uncomfortable (nasal and throat swabs) and more patient-accepted modes of specimen acquisition.

Point-of-care testing also showed no impact on the accuracy of RIDTs. Thus, administering of the RIDTs by personnel other than a trained laboratory technician doesn't seem to adversely impact these tests' performance. This is good news since it is likely that it is when they are used as first-line tests, outside of the laboratory setting, that these tests find their most useful application in the diagnostic algorithm for influenza. However, none of the studies directly compared accuracy between RIDTs performed at the point-of-care versus in the laboratory setting.

Although very few studies examined the effect of duration of symptoms before testing on RIDTs accuracy, those that did seemed to indicate a better accuracy 2 to 3 days after symptoms onset, with a rapid decline thereafter. This is likely explained by the fact that viral shedding in influenza has been shown to peak at 24-48 hours of illness, with a dramatic decrease after 3 to 5 days(4). Duration of a patient symptoms is thus likely to have an important impact on the accuracy of RIDTs (and other diagnostic test for influenza) and testing within 3 to 5 days of disease onset is recommended by the CDC and IDSA(3, 4).

Strengths and limitations of the systematic review

Our systematic review has several strengths. First, we used a standard protocol, including a comprehensive search strategy with various overlapping approaches, which enabled us to retrieve a large number of studies. By contacting several authors, we were able to gather information missing from the original publications. We used rigorous methods of data analysis, including bivariate random effects regression models and HSROC analyses, both random effects models that have been shown to yield more conservative pooled estimates. Lastly, subgroups analyses, according to predefined subgroups, and meta-regression were used in an attempt to explain some of the observed heterogeneity in accuracy estimates.

However, our systematic review had some limitations. Over the years, RT-PCR has gradually replaced viral culture as the preferred reference standard for influenza diagnosis. Since RT-PCR has better sensitivity and specificity than viral culture, there is a significant difference in accuracy of the RIDTs depending on the reference standard used. Though we acknowledged the superiority of RT-PCR by preferring it every time accuracy estimates were reported for both viral culture and RT-PCR, both are currently held as accepted reference standards, and choosing only RT-PCR would have biased our search to only recently conducted studies.

Considerable heterogeneity was found in the pooled estimates, which was not unexpected given the different products, specimen types, settings, and populations studied. In the presence of significant heterogeneity, summary measures can be difficult to interpret or even misleading. Despite conducting subgroup analyses and meta-regression, substantial heterogeneity remained unexplained. Many factors, possibly contributing to this residual heterogeneity, could not be assessed because they were not reported in most studies.

For example, duration of clinical symptoms before testing is likely to have an important effect on test performance since viral shedding decreases dramatically after 3 to 5 days of illness(4). This information was only mentioned in 13% of the included studies, with only 6 studies providing stratification by time since onset. Many studies failed to stratify by specimen type, making it difficult to assess the contribution of each to overall RIDTs' accuracy. Also, some subgroups, such as children, were by necessity broad and could encompass very different groups, from neonates to adolescents, for example. Similarly, the elderly were included within the adult subgroups because of insufficient data to analyze them separately. Finally, other variables, such as flu vaccination coverage of the population under study, inclusion or exclusion of individuals with co-morbidities, type of swab used (nylon versus cotton), transport medium used, time elapsed before specimen processing, and storage of specimens before testing (room temperature, refrigeration, freezing, etc.), were reported so infrequently that their impact was impossible to assess.

Limitations of the literature reviewed

The included studies had some shortcomings in their methodological qualities. Few studies (37%) gave a clear rationale for the selection of the included specimens or patients. Less than half (41%) reported blinded assessment of the RIDTs or reference test results. Only 18% mentioned how they dealt with uninterpretable results, of particular interest since very faint lines observed during reading may be an important source of false positive results(51). However, when assessed using subgroup analyses, none of these quality issues had a significant impact on RIDTs' accuracy.

Language and publication bias

Although we searched several different sources, it is possible that we may have missed some eligible studies, or that a large number of unpublished studies exist. Furthermore, we only extracted data on studies in English and

French, which could have biased our results. We are aware of at least 8 Japanese language studies on RIDTs, though they pertained to different tests than the main ones studied here and might not have met our inclusion criteria. Also, it was not possible to formally assess publication bias since there is no good method to do so when dealing with diagnostic studies.

Unlike meta-analyses of randomized trials where funnel plots and regression asymmetry tests (e.g. Egger test) are used to explore potential publication bias, these techniques are misleading when applied to diagnostic meta-analyses and are discouraged by the Cochrane Diagnostic Test Accuracy Methods Group (33). In fact, funnel plot-based tests have proven to be seriously misleading for diagnostic studies, and other alternatives suffer from poor power(33). It is thus safer to just assume that publication bias is likely present, with small studies showing poor accuracy likely being unreported.

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

Conclusion

The most important advantage of RIDTs is their rapid turn-around time, providing clinicians with an answer concerning their patient's influenza status in mere minutes. Although they undoubtedly have higher accuracy, RT-PCR, viral culture and immunofluroscence take hours or even days to give results, even discounting transportation time to the nearest laboratory. Thus, RIDTs fill a void at the patient bedside that none of the other tests is likely to fill in the near future, as a first-line test to be confirmed at a later date by more time-consuming, definitive testing.

As long as clinicians understand the limitations of RIDTs, namely that a negative result is unreliable and should be confirmed using culture or RT-PCR, RIDTs could enable clinicians to institute prompt infection-control measures during an influenza outbreak, start antiviral treatment in high-risk populations while they are still likely to impact the course of illness, and make informed decisions concerning further diagnostic investigations and the empirical start of antibiotics in unwell patients with fever. Although far from perfect, RIDTs seem to be the best options that clinicians have to guide them at the point-of-care, while waiting on more definitive results.

Clinical implications

The focus of our meta-analysis was test accuracy. Test sensitivity and specificity are considered surrogates of patient-important outcomes. Patient-important outcomes require more sophisticated and often more resource-intensive research (e.g. randomized controlled trials), wherein a study shows that implementing a diagnostic test in a given situation results in clinically relevant improvements in patient care and/or patient outcomes.

Indeed, a few studies have examined the clinical impact of RIDTs on patient management and outcomes. When compared to patients with negative RIDTs, patients with positive RIDTs had a decreased number of additional tests ordered during their emergency visit (13, 52, 53), though the exact tests responsible for this decrease (e.g. chest X-ray, blood tests, etc.) varied across studies. They also received a decreased number of antibiotic prescriptions and an increased number of antiviral prescriptions (13, 52, 54-56), of particular interest in an era where antibiotic resistance is a growing concern. However, this difference could be spurious and reflect a decrease in the total amount of additional testing and antibiotic prescriptions when clinicians have access to RIDTs and the results of the test is positive for influenza, or an increase in the number of tests performed and antibiotics prescribed when the test is negative, above what the physician would have done had he only relied on clinical judgment.

To circumvent this problem, studies have looked at the differences in terms of the ordering of additional tests and the prescribing of antibiotics and antiviral medication in populations of influenza positive patients. When the positive influenza status was known during the emergency department visit through RIDTs' use, there was a decrease in the use of ancillary tests (blood tests, blood culture, urinalysis and culture, and chest X-ray)(12) in the time spent in the emergency department(12) and in antibiotic prescriptions(12, 57-59) with an increase use of antiviral medication(12, 55, 57-59). This finding however, does not account for the potential inverse effect of a negative RIDT result.

Studies looking at differences in practice between clinicians given access to RIDTs, independent of the RIDTs' results, versus those that only had access to traditional reference standards, yielded conflicting results. Some reported a decrease in the overall number of additional tests(14), while others reported no difference or even an increase in the number of chest X-rays(16), mainly

in RIDTs' negative patient. As well, some reported a decrease in antibiotic prescription(60) while other reported no difference(14) or even an increase(16), attributable to RIDTs' negative patients. When available in nursing homes, RIDTs have been shown to help control influenza outbreaks, by decreasing their duration or the final attack rate(19, 20).

Needs for future research

Though not conclusively proven, RIDTs appear to have a positive impact on patient-important outcomes in certain circumstances. Further studies are needed to confirm these benefits and see if a better understanding by clinicians of RIDTs' strengths and weaknesses (for example the fact that a negative RIDT is still compatible with a clinical diagnostic of influenza) might lead to an even greater reduction in ancillary testing and unnecessary antibiotics use. As well, cost-effectiveness studies are essential to see if these benefits offset the added cost of performing RIDTs.

REFERENCES

1. AAP. Influenza. In: Pickering L, Baker C, Kimberlin D, Long S, editors. Red Book: 2009 Report of the Committee on Infectious Diseases. 28th ed. Elk Grove Village, IL: American Academy of Pediatrics; 2009. p. 400-12.
2. Cox NJ, Subbarao K. Influenza. Lancet. 1999 Oct 9;354(9186):1277-82.
3. Harper SA, Bradley JS, Englund JA, File TM, Gravenstein S, Hayden FG, et al. Seasonal influenza in adults and children--diagnosis, treatment, chemoprophylaxis, and institutional outbreak management: clinical practice guidelines of the Infectious Diseases Society of America. Clin Infect Dis. 2009 Apr 15;48(8):1003-32.
4. Fiore AE, Fry A, Shay D, Gubareva L, Bresee JS, Uyeki TM, et al. Antiviral agents for the treatment and chemoprophylaxis of influenza --- recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep. 2011 Jan 21;60(1):1-24.
5. Neuzil KM, Mellen BG, Wright PF, Mitchel EF, Jr., Griffin MR. The effect of influenza on hospitalizations, outpatient visits, and courses of antibiotics in children. N Engl J Med. 2000 Jan 27;342(4):225-31.
6. WHO. Influenza (Seasonal). [Internet] Geneva: World Health Organization; [cited March 8, 2011]; Available from: http://www.who.int/mediacentre/factsheets/fs211/en/index.html.
7. Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, et al. Influenza-associated hospitalizations in the United States. JAMA. 2004 Sep 15;292(11):1333-40.
8. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA. 2003 Jan 8;289(2):179-86.
9. WHO. Pandemic (H1N1) 2009 - update 112. [Internet] Geneva: World Health Organization; [updated August 6, 2010; cited March 8, 2011]; Available from: http://www.who.int/csr/don/2010_08_06/en/index.html.
10. Burkardt HJ. Pandemic H1N1 2009 ('swine flu'): diagnostic and other challenges. Expert Rev Mol Diagn. 2011 Jan;11(1):35-40.
11. Abanses JC, Dowd MD, Simon SD, Sharma V. Impact of rapid influenza testing at triage on management of febrile infants and young children. Pediatr Emerg Care. 2006 Mar;22(3):145-9.
12. Bonner AB, Monroe KW, Talley LI, Klasner AE, Kimberlin DW. Impact of the rapid diagnosis of influenza on physician decision-making and patient management in the pediatric emergency department: results of a randomized, prospective, controlled trial. Pediatrics. 2003 Aug;112(2):363-7.
13. Esposito S, Marchisio P, Morelli P, Crovari P, Principi N. Effect of a rapid influenza diagnosis. Arch Dis Child. [Article]. 2003 Jun;88(6):525-6.
14. Poehling KA, Zhu Y, Tang Y-W, Edwards K. Accuracy and impact of a point-of-care rapid influenza test in young children with respiratory

illnesses. Archives of Pediatrics & Adolescent Medicine. 2006 Jul;160(7):713-8.
15. Sharma V, Dowd MD, Slaughter AJ, Simon SD. Effect of rapid diagnosis of influenza virus type a on the emergency department management of febrile infants and toddlers. Arch Pediatr Adolesc Med. 2002 Jan;156(1):41-3.
16. Cohen R, Thollot F, Lecuyer A, Koskas M, Touitou R, Boucherat M, et al. [Impact of the rapid diagnosis downtown in the assumption of responsibility of the children in period of influenza]. Arch Pediatr. 2007 Jul;14(7):926-31.
17. Salgado CD, Farr BM, Hall KK, Hayden FG. Influenza in the acute hospital setting. Lancet Infect Dis. 2002 Mar;2(3):145-55.
18. Jefferson T, Jones M, Doshi P, Del Mar C, Dooley L, Foxlee R. Neuraminidase inhibitors for preventing and treating influenza in healthy adults. Cochrane Database Syst Rev. 2010(2):CD001265.
19. Church DL, Davies HD, Mitton C, Semeniuk H, Logue M, Maxwell C, et al. Clinical and economic evaluation of rapid influenza a virus testing in nursing homes in calgary, Canada. Clin Infect Dis. 2002 Mar 15;34(6):790-5.
20. Leonardi GP, Leib H, Birkhead GS, Smith C, Costello P, Conron W. Comparison of rapid detection methods for influenza A virus and their value in health-care management of institutionalized geriatric patients. J Clin Microbiol. 1994 Jan;32(1):70-4.
21. CDC. U.S. influenza sentinel provider surveillance network. [Internet] Atlanta: Center for Disease Control and Prevention; 2006 [cited April 13, 2011]; Available from: http://www.doh.state.fl.us/disease_ctrl/epi/htopics/flu/FSPISN/RecruitmentCDCsys2006.pdf.
22. WHO. WHO recommended surveillance standards. [Internet] Geneva: World Health Organization; 1999 [cited April 13, 2011]; Available from: http://www.who.int/csr/resources/publications/surveillance/whocdscsrisr992.pdf.
23. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? JAMA. 2005 Feb 23;293(8):987-97.
24. Petrozzino JJ, Smith C, Atkinson MJ. Rapid diagnostic testing for seasonal influenza: an evidence-based review and comparison with unaided clinical diagnosis. J Emerg Med. 2010 Oct;39(4):476-90 e1.
25. Gavin PJ, Thomson RB. Review of rapid diagnostic tests for influenza. Clinical and Applied Immunology Reviews. 2003;4:151-72.
26. WHO. WHO recommendations on the use of rapid testing for influenza diagnosis. [Internet] Geneva: World Health Organization; 2005 [cited July 22, 2010]; Available from: http://www.who.int/csr/disease/avian_influenza/guidelines/rapid_testing/en/index.html.
27. Chartrand C, Minion J, Pai M. Rapid diagnostics for influenza: what are the options? Future Microbiol. 2010 Oct;5:1451-5.
28. WHO. Use of influenza rapid diagnostic tests. Geneva: World Health Organization; 2010.

29. Rapid diagnostic tests for influenza. Med Lett Drugs Ther. 1999 Dec 17;41(1068):121-2.
30. CDC. Guidance for clinicians on the use of rapid influenza diagnostic tests for the 2010-2011 influenza season. [Internet] Atlanta: Centers for Disease Control and Prevention; [updated December 22, 2010; cited March 8,2011]; Available from: http://www.cdc.gov/flu/professionals/diagnosis/clinician_guidance_ridt.htm.
31. Uyeki TM. Influenza diagnosis and treatment in children: a review of studies on clinically useful tests and antiviral treatment for influenza. Pediatr Infect Dis J. 2003 Feb;22(2):164-77.
32. Centers for Disease C, Prevention. Evaluation of rapid influenza diagnostic tests for detection of novel influenza A (H1N1) Virus - United States, 2009. MMWR Morb Mortal Wkly Rep. 2009 Aug 7;58(30):826-9.
33. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working G. Systematic reviews of diagnostic test accuracy. Ann Intern Med. 2008 Dec 16;149(12):889-97.
34. Pai M, McCulloch M, Enanoria W, Colford JM, Jr. Systematic reviews of diagnostic test evaluations: What's behind the scenes? ACP J Club. 2004 Jul-Aug;141(1):A11-3.
35. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006 Feb 14;174(4):469-76.
36. CDC. Past weekly surveillance reports. [Internet] Atlanta: Centers for Disease Control and Prevention; [updated January 21, 2011; cited January 27, 2011]; Available from: http://www.cdc.gov/flu/weekly/pastreports.htm.
37. WHO. FluNet. [Internet] Geneva: World Health Organization; 2011 [cited January 27, 2011]; Available from: http://www.who.int/csr/disease/influenza/influenzanetwork/flunet/en/.
38. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003 Nov 10;3:25.
39. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol. 2006;6:31.
40. Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. The Stata Journal. 2009;9(2):211-29.
41. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001 Oct 15;20(19):2865-84.
42. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005 Oct;58(10):982-90.

43. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 6;327(7414):557-60.
44. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Med Decis Making. 1993 Oct-Dec;13(4):313-21.
45. Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. Stat Med. 2002 Jun 15;21(11):1525-37.
46. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009 Jul 21;6(7):e1000097.
47. Simonsen L, Clarke MJ, Williamson GD, Stroup DF, Arden NH, Schonberger LB. The impact of influenza epidemics on mortality: introducing a severity index. Am J Public Health. 1997 Dec;87(12):1944-50.
48. Heikkinen T, Marttila J, Salmi AA, Ruuskanen O. Nasal swab versus nasopharyngeal aspirate for isolation of respiratory viruses. J Clin Microbiol. 2002 Nov;40(11):4337-9.
49. Lambert SB, Whiley DM, O'Neill NT, Andrews EC, Canavan FM, Bletchly C, et al. Comparing nose-throat swabs and nasopharyngeal aspirates collected from children with symptoms for respiratory virus identification using real-time polymerase chain reaction. Pediatrics. 2008 Sep;122(3):e615-20.
50. Sung RY, Chan PK, Choi KC, Yeung AC, Li AM, Tang JW, et al. Comparative study of nasopharyngeal aspirate and nasal swab specimens for diagnosis of acute viral respiratory infection. J Clin Microbiol. 2008 Sep;46(9):3073-6.
51. Quach C, Newby D, Daoust G, Rubin E, McDonald J. QuickVue influenza test for rapid detection of influenza A and B viruses in a pediatric population. Clin Diagn Lab Immunol. 2002 Jul;9(4):925-6.
52. de La Rocque F, Lecuyer A, Wollner C, d'Athis P, Pecking M, Thollot F, et al. Impact of influenza rapid diagnostic tests (IRDT) on the diagnosis of influenza and on the management of influenza in children in ambulatory pediatric setting. [French]. Arch Pediatr. 2009 March;16(3):288-93.
53. Pierron S, Haas H, Berlioz M, Ollier L, Albertini M. Impact of rapid influenza test during influenza epidemic in all febrile children less than 6 years old in a pediatric emergency department. Arch Pediatr. [Article]. 2008 Aug;15(8):1283-8.
54. Bhavnani D, Phatinawin L, Chantra S, Olsen SJ, Simmerman JM. The influence of rapid influenza diagnostic testing on antibiotic prescribing patterns in rural Thailand. Int J Infect Dis. 2007 Jul;11(4):355-9.
55. Noyola DE, Demmler GJ. Effect of rapid diagnosis on management of influenza A infections. Pediatr Infect Dis J. 2000 Apr;19(4):303-7.
56. Theocharis G, Vouloumanou EK, Rafailidis PI, Spiropoulos T, Barbas SG, Falagas ME. Evaluation of a direct test for seasonal influenza in outpatients. Eur J Intern Med. Oct;21(5):434-8.

57. D'Heilly SJ, Janoff EN, Nichol P, Nichol KL. Rapid diagnosis of influenza infection in older adults: influence on clinical care in a routine clinical setting. J Clin Virol. 2008 Jun;42(2):124-8.
58. Falsey AR, Murata Y, Walsh EE. Impact of rapid diagnosis on management of adults hospitalized with influenza. Arch Intern Med. 2007 Feb 26;167(4):354-60.
59. Jennings LC, Skopnik H, Burckhardt I, Hribar I, Del Piero L, Deichmann KA. Effect of rapid influenza testing on the clinical management of paediatric influenza. Influenza Other Respi Viruses. 2009 May;3(3):91-8.
60. Ozkaya E, Cambaz N, Coskun Y, Mete F, Geyik M, Samanci N. The effect of rapid diagnostic testing for influenza on the reduction of antibiotic use in paediatric emergency department. Acta Paediatr. 2009 Oct;98(10):1589-92.
61. Agoritsas K, Mack K, Bonsu BK, Goodman D, Salamon D, Marcon MJ. Evaluation of the Quidel QuickVue test for detection of influenza A and B viruses in the pediatric emergency medicine setting by use of three specimen collection methods. Journal of Clinical Microbiology. 2006 Jul;44(7):2638-41.
62. Alexander R, Hurt AC, Lamb D, Wong FY, Hampson AW, Barr IG. A comparison of a rapid test for influenza with laboratory-based diagnosis in a paediatric population. Commun Dis Intell. 2005;29(3):272-6.
63. Arsene S, Vabret A, Dina J, Tecu C, Brouard J, Eckard P, et al. Comparison of the quick view influenza test (Quidel) to an immunofluorescence assay for the detection of influenza virus infections. Roum Arch Microbiol Immunol. 2004 Jul-Dec;63(3-4):235-43.
64. Bellei N, Benfica D, Perosa AH, Carlucci R, Barros M, Granato C. Evaluation of a rapid test (QuickVue) compared with the shell vial assay for detection of influenza virus clearance after antiviral treatment. J Virol Methods. 2003 Apr;109(1):85-8.
65. Bellmann-Weiler R, Beikircher B, Kurz K, Theurl I, Weiss G. Accuracy of bedside antigen tests in the diagnosis of new influenza A/H1N1v infection. Clin Microbiol Infect. 2010 Apr 6.
66. Biggs C, Walsh P, Overmyer CL, Gonzalez D, Feola M, Mordechai E, et al. Performance of influenza rapid antigen testing in influenza in emergency department patients. Emergency Medicine Journal. 2010 January;27(1):5-7.
67. Boivin G, Hardy I, Kress A. Evaluation of a rapid optical immunoassay for influenza viruses (FLU OIA test) in comparison with cell culture and reverse transcription-PCR. J Clin Microbiol. 2001 Feb;39(2):730-2.
68. Boivin G, Cote S, Dery P, De Serres G, Bergeron MG. Multiplex real-time PCR assay for detection of influenza and human respiratory syncytial viruses. J Clin Microbiol. 2004 Jan;42(1):45-51.
69. Boon AC, French AM, Fleming DM, Zambon MC. Detection of influenza a subtypes in community-based surveillance. J Med Virol. 2001 Sep;65(1):163-70.
70. Booth S, Baleriola C, Rawlinson WD. Comparison of two rapid influenza A/B test kits with reference methods showing high specificity and sensitivity for influenza A infection. J Med Virol. 2006 May;78(5):619-22.

71. Cazacu AC, Greer J, Taherivand M, Demmler GJ. Comparison of lateral-flow immunoassay and enzyme immunoassay with viral culture for rapid detection of influenza virus in nasal wash specimens from children. J Clin Microbiol. 2003 May;41(5):2132-4.
72. Cazacu AC, Chung SE, Greer J, Demmler GJ. Comparison of the directigen flu A+B membrane enzyme immunoassay with viral culture for rapid detection of influenza A and B viruses in respiratory specimens. J Clin Microbiol. 2004 Aug;42(8):3707-10.
73. Cazacu AC, Demmler GJ, Neuman MA, Forbes BA, Chung S, Greer J, et al. Comparison of a new lateral-flow chromatographic membrane immunoassay to viral culture for rapid detection and differentiation of influenza A and B viruses in respiratory specimens. J Clin Microbiol. 2004 Aug;42(8):3661-4.
74. Chan KH, Maldeis N, Pope W, Yup A, Ozinskas A, Gill J, et al. Evaluation of the Directigen FluA+B test for rapid diagnosis of influenza virus type A and B infections. J Clin Microbiol. 2002 May;40(5):1675-80.
75. Chen Y, Xu F, Gui X, Yang K, Wu X, Zheng Q, et al. A rapid test for the detection of influenza A virus including pandemic influenza A/H1N1 2009. J Virol Methods. 2010 Jul;167(1):100-2.
76. Cheng CK, Cowling BJ, Chan KH, Fang VJ, Seto WH, Yung R, et al. Factors affecting QuickVue Influenza A + B rapid test performance in the community setting. Diagn Microbiol Infect Dis. 2009 Sep;65(1):35-41.
77. Choi YJ, Kim HJ, Park JS, Oh MH, Nam HS, Kim YB, et al. Evaluation of new rapid antigen test for detection of pandemic influenza A/H1N1 2009 virus. J Clin Microbiol. 2010 Jun;48(6):2260-2.
78. Covalciuc KA, Webb KH, Carlson CA. Comparison of four clinical specimen types for detection of influenza A and B viruses by optical immunoassay (FLU OIA test) and cell culture methods. J Clin Microbiol. 1999 Dec;37(12):3971-4.
79. Crum-Cianflone NF, Blair PJ, Faix D, Arnold J, Echols S, Sherman SS, et al. Clinical and Epidemiologic Characteristics of an Outbreak of Novel H1N1 (Swine Origin) Influenza A Virus among United States Military Beneficiaries. Clinical Infectious Diseases. [Article]. 2009 Dec;49(12):1801-10.
80. Cruz AT, Cazacu AC, McBride LJ, Greer JM, Demmler GJ. Performance characteristics of a rapid immunochromatographic assay for detection of influenza virus in children during the 2003 to 2004 influenza season. Annals of Emergency Medicine. 2006 Mar;47(3):250-4.
81. Cruz AT, Cazacu AC, Greer JM, Demmler GJ. Rapid assays for the diagnosis of influenza A and B viruses in patients evaluated at a large tertiary care children's hospital during two consecutive winter seasons. J Clin Virol. 2008 Feb;41(2):143-7.
82. Cruz AT, Demmler-Harrison GJ, Caviness AC, Buffone GJ, Revell PA. Performance of a rapid influenza test in children during the H1N1 2009 influenza a outbreak. Pediatrics. 2010 Mar;125(3):e645-50.

83. Dale SE, Mayer C, Mayer MC, Menegus MA. Analytical and clinical sensitivity of the 3M rapid detection influenza A+B assay. J Clin Microbiol. 2008 Nov;46(11):3804-7.
84. de la Tabla VO, Antequera P, Masia M, Ros P, Martin C, Gazquez G, et al. Clinical evaluation of rapid point-of-care testing for detection of novel influenza A (H1N1) virus in a population-based study in Spain. Clin Microbiol Infect. 2010 Dec 29.
85. Diederen BM, Veenendaal D, Jansen R, Herpers BL, Ligtvoet EE, Ijzerman EP. Rapid antigen test for pandemic (H1N1) 2009 virus. Emerg Infect Dis. 2010 May;16(5):897-8; author reply 8.
86. Dominguez EA, Taber LH, Couch RB. Comparison of rapid diagnostic techniques for respiratory syncytial and influenza A virus respiratory infections in young children. J Clin Microbiol. 1993 Sep;31(9):2286-90.
87. Drinka PJ. Experience with the rapid Directigen test for influenza. J Am Med Dir Assoc. 2006 Jan;7(1):37-9.
88. Dunn JJ, Gordon C, Kelley C, Carroll KC. Comparison of the Denka-Seiken INFLU A.B-Quick and BD Directigen Flu A+B kits with direct fluorescent-antibody staining and shell vial culture methods for rapid detection of influenza viruses. J Clin Microbiol. 2003 May;41(5):2180-3.
89. Effler PV, Ieong MC, Tom T, Nakata M. Enhancing public health surveillance for influenza virus by incorporating newly available rapid diagnostic tests. Emerg Infect Dis. 2002 Jan;8(1):23-8.
90. Fader RC. Comparison of the Binax NOW Flu A enzyme immunochromatographic assay and R-Mix shell vial culture for the 2003-2004 influenza season. Journal of Clinical Microbiology. 2005 Dec;43(12):6133-5.
91. Faix DJ, Sherman SS, Waterman SH. Rapid-test sensitivity for novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med. 2009 Aug 13;361(7):728-9.
92. Fernandez C, Cataletto M, Lee P, Feuerman M, Krilov L. Rapid influenza A testing for novel H1N1: point-of-care performance. Postgrad Med. 2010 Jan;122(1):28-33.
93. Foo H, Blyth CC, van Hal S, McPhie K, Ratnamohan M, Fennell M, et al. Laboratory test performance in young adults during influenza outbreaks at World Youth Day 2008. J Clin Virol. 2009 Dec;46(4):384-6.
94. Fuenzalida L, Blanco S, Prat C, Vivancos M, Dominguez MJ, Modol JM, et al. Utility of the rapid antigen detection BinaxNOW A&B test for detection of novel influenza A (H1N1) virus. Clin Microbiol Infect. 2010 Dec 29.
95. Ganzenmueller T, Kluba J, Hilfrich B, Puppe W, Verhagen W, Heim A, et al. Comparison of the performance of direct fluorescent antibody staining, a point-of-care rapid antigen test and virus isolation with that of RT-PCR for the detection of novel 2009 influenza A (H1N1) virus in respiratory specimens. J Med Microbiol. 2010 Jun;59(Pt 6):713-7.
96. Ghebremedhin B, Engelmann I, Konig W, Konig B. Comparison of the performance of the rapid antigen detection actim Influenza A&B test and RT-

PCR in different respiratory specimens. J Med Microbiol. 2009 Mar;58(Pt 3):365-70.
97. Gooskens J, Swaan CM, Claas EC, Kroes AC. Rapid molecular detection of influenza outbreaks in nursing homes. J Clin Virol. 2008 Jan;41(1):7-12.
98. Gordon A, Videa E, Saborio S, Lopez R, Kuan G, Reingold A, et al. Performance of an influenza rapid test in children in a primary healthcare setting in Nicaragua. PLoS One. 2009;4(11)(e7907).
99. Gordon A, Videa E, Saborio S, Lopez R, Kuan G, Balmaseda A, et al. Diagnostic accuracy of a rapid influenza test for pandemic influenza A H1N1. PLoS One. 2010;5(4):e10364.
100. Grijalva CG, Poehling KA, Edwards KM, Weinberg GA, Staat MA, Iwane MK, et al. Accuracy and interpretation of rapid influenza tests in children. Pediatrics. 2007 Jan;119(1):e6-11.
101. Grondahl B, Puppe W, Weigl J, Schmitt HJ. Comparison of the BD Directigen Flu A+B Kit and the Abbott TestPack RSV with a multiplex RT-PCR ELISA for rapid detection of influenza viruses and respiratory syncytial virus. Clinical Microbiology and Infection. 2005 Oct;11(10):848-50.
102. Hamilton MS, Abel DM, Ballam YJ, Otto MK, Nickell AF, Pence LM, et al. Clinical evaluation of the ZstatFlu-II test: a chemiluminescent rapid diagnostic test for influenza virus. J Clin Microbiol. 2002 Jul;40(7):2331-4.
103. Hara M, Takao S, Fukuda S, Shimazu Y, Miyazaki K. Evaluation of three immunochromatographic kits for rapid detection of influenza virus A and B. Labmedicine. [Article]. 2008 Oct;39(10):603-6.
104. Harnden A, Brueggemann A, Shepperd S, White J, Hayward AC, Zambon M, et al. Near patient testing for influenza in children in primary care: comparison with laboratory test. BMJ. 2003 Mar 1;326(7387):480.
105. Hawkes M, Richardson SE, Ipp M, Schuh S, Adachi D, Tran D. Sensitivity of rapid influenza diagnostic testing for swine-origin 2009 a (H1N1) influenza virus in children. Pediatrics. 2010 Mar;125(3):e639-44.
106. Herrmann B, Larsson C, Zweygberg BW. Simultaneous detection and typing of influenza viruses A and B by a nested reverse transcription-PCR: comparison to virus isolation and antigen detection by immunofluorescence and optical immunoassay (FLU OIA). J Clin Microbiol. 2001 Jan;39(1):134-8.
107. Hindiyeh M, Goulding C, Morgan H, Kenyon B, Langer J, Fox L, et al. Evaluation of BioStar FLU OIA assay for rapid detection of influenza A and B viruses in respiratory specimens. J Clin Virol. 2000 Aug;17(2):119-26.
108. Hulson TD, Mold JW, Scheid D, Aaron M, Aspy CB, Ballard NL, et al. Diagnosing influenza: the value of clinical clues and laboratory tests. J Fam Pract. 2001 Dec;50(12):1051-6.
109. Hurt AC, Alexander R, Hibbert J, Deed N, Barr IG. Performance of six influenza rapid tests in detecting human influenza in clinical specimens. J Clin Virol. 2007 Jun;39(2):132-5.
110. Johnston SL, Bloy H. Evaluation of a rapid enzyme immunoassay for detection of influenza A virus. J Clin Microbiol. 1993 Jan;31(1):142-3.
111. Karre T, Maguire HF, Butcher D, Graepler A, Weed D, Wilson ML. Comparison of Becton Dickinson Directigen EZ Flu A+B test against the CDC

real-time PCR assay for detection of 2009 pandemic influenza A/H1N1 virus. J Clin Microbiol. 2010 Jan;48(1):343-4.
112. Kim YK, Uh Y, Chun JK, Kim C, Kim HY. Evaluation of new hemagglutinin-based rapid antigen test for influenza A pandemic (H1N1) 2009. J Clin Virol. 2010 Sep;49(1):69-72.
113. Kok J, Blyth CC, Foo H, Patterson J, Taylor J, McPhie K, et al. Comparison of a rapid antigen test with nucleic acid testing during cocirculation of pandemic influenza A/H1N1 2009 and seasonal influenza A/H3N2. J Clin Microbiol. 2010 Jan;48(1):290-1.
114. Landry ML, Cohen S, Ferguson D. Real-time PCR compared to Binax NOW and cytospin-immunofluorescence for detection of influenza in hospitalized patients. J Clin Virol. 2008 Oct;43(2):148-51.
115. Leonardi GP, Mitrache I, Pigal A, Freedman L. Public hospital-based laboratory experience during an outbreak of pandemic influenza A (H1N1) virus infections. J Clin Microbiol. 2010 Apr;48(4):1189-94.
116. Liao RS, Tomalty LL, Majury A, Zoutman DE. Comparison of viral isolation and multiplex real-time reverse transcription-PCR for confirmation of respiratory syncytial virus and influenza virus detection by antigen immunoassays. J Clin Microbiol. 2009 Mar;47(3):527-32.
117. Likitnukul S, Boonsiri K, Tangsuksant Y. Evaluation of sensitivity and specificity of rapid influenza diagnostic tests for novel swine-origin influenza A (H1N1) virus. Pediatr Infect Dis J. 2009 Nov;28(11):1038-9.
118. Louie JK, Guevara H, Boston E, Dahlke M, Nevarez M, Kong T, et al. Rapid influenza antigen test for diagnosis of pandemic (H1N1) 2009. Emerg Infect Dis. 2010 May;16(5):824-6.
119. Marcante R, Chiumento F, Palu G, Cavedon G. Rapid diagnosis of influenza type A infection: comparison of shell-vial culture, directigen flu-A and enzyme-linked immunosorbent assay. New Microbiol. 1996 Apr;19(2):141-7.
120. Mee Lee H, Ki Park H, Hwang HS, Chun MY, Pai HJ, Hee Oh S, et al. Diagnostic value of the rapid influenza antigen test for novel influenza A (H1N1). Scand J Infect Dis. 2010 Aug 25.
121. Monto AS, Rotthoff J, Teich E, Herlocher ML, Truscon R, Yen HL, et al. Detection and control of influenza outbreaks in well-vaccinated nursing home populations. Clin Infect Dis. 2004 Aug 15;39(4):459-64.
122. Newton DW, Mellen CF, Baxter BD, Atmar RL, Menegus MA. Practical and sensitive screening strategy for detection of influenza virus. Journal of Clinical Microbiology. [Article]. 2002 Nov;40(11):4353-6.
123. Nilsson AC, Alemo B, Bjorkman P, Dillner L, Melhus A, Nilsson B, et al. Around-the-clock, rapid diagnosis of influenza by means of membrane chromatography antigen testing confirmed by polymerase chain reaction. Infect Control Hosp Epidemiol. 2008 Feb;29(2):177-9.
124. Nougairede A, Ninove L, Zandotti C, de Lamballerie X, Gazin C, Drancourt M, et al. Point of Care Strategy for Rapid Diagnosis of Novel A/H1N1 Influenza Virus. PLoS One. [Article]. 2010 Feb;5(2).

125. Nougairede A, Ninove L, Zandotti C, Thiberville SD, Gazin C, La Scola B, et al. Interim report on the A/H1N1 influenza virus pandemic in Marseille, France, April-November 2009. Clinical Microbiology and Infection. [Article]. 2010 Apr;16(4):322-5.
126. Noyola DE, Clark B, O'Donnell FT, Atmar RL, Greer J, Demmler GJ. Comparison of a new neuraminidase detection assay with an enzyme immunoassay, immunofluorescence, and culture for rapid detection of influenza A and B viruses in nasal wash specimens. J Clin Microbiol. 2000 Mar;38(3):1161-5.
127. Noyola DE, Paredes AJ, Clark B, Demmler GJ. Evaluation of a neuraminidase detection assay for the rapid detection of influenza A and B virus in children. Pediatr Dev Pathol. 2000 Mar-Apr;3(2):162-7.
128. Poehling KA, Griffin MR, Dittus RS, Tang YW, Holland K, Li H, et al. Bedside diagnosis of influenzavirus infections in hospitalized children. Pediatrics. 2002 Jul;110(1 Pt 1):83-8.
129. Pregliasco F, Puzelli S, Mensi C, Anselmi G, Marinello R, Luisa Tanzi M, et al. Influenza Virological Surveillance in Children: The Use of the QuickVue Rapid Diagnostic Test. Journal of Medical Virology. 2004 Jun;73(2):269-73.
130. Rahman M, Kieke BA, Vandermause MF, Mitchell PD, Greenlee RT, Belongia EA. Performance of Directigen flu A plus B enzyme immunoassay and direct fluorescent assay for detection of influenza infection during the 2004-2005 season. Diagnostic Microbiology & Infectious Disease. 2007 Aug;58(4):413-8.
131. Rahman M, Vandennause MF, Kieke BA, Belongia EA. Performance of Binax NOW Flu A and B and direct fluorescent assay in comparison with a composite of viral culture or reverse transcription polymerase chain reaction for detection of influenza infection during the 2006 to 2007 season. Diagnostic Microbiology & Infectious Disease. 2008 Oct;62(2):162-6.
132. Rashid H, Shafi S, Haworth E, El Bashir H, Ali KA, Memish ZA, et al. Value of rapid testing for influenza among Hajj pilgrims. Travel Med Infect Dis. 2007 Sep;5(5):310-3.
133. Rawlinson WD, Waliuzzaman ZM, Fennell M, Appleman JR, Shimasaki CD, Carter IW. New point of care test is highly specific but less sensitive for influenza virus A and B in children and adults. J Med Virol. 2004 Sep;74(1):127-31.
134. Reina J, Munar M, Blanco I. Evaluation of a direct immunofluorescence assay, dot-blot enzyme immunoassay, and shell vial culture in the diagnosis of lower respiratory tract infections caused by influenza A virus. Diagn Microbiol Infect Dis. 1996 Jul;25(3):143-5.
135. Reina J, Padilla E, Alonso F, de Gopegui ER, Munar M, Mari M. Evaluation of a new dot blot enzyme immunoassay (Directigen flu A+B) for simultaneous and differential detection of influenza A and B virus antigens from respiratory samples. Journal of Clinical Microbiology. [Article]. 2002 Sep;40(9):3515-7.

136. Rodriguez WJ, Schwartz RH, Thorne MM. Evaluation of diagnostic tests for influenza in a pediatric practice. Pediatr Infect Dis J. 2002 Mar;21(3):193-6.
137. Rouleau I, Charest H, Douville-Fradet M, Skowronski DM, De Serres G. Field performance of a rapid diagnostic test for influenza in an ambulatory setting. J Clin Microbiol. 2009 Sep;47(9):2699-703.
138. Ruest A, Michaud S, Deslandes S, Frost EH. Comparison of the Directigen flu A+B test, the QuickVue influenza test, and clinical case definition to viral culture and reverse transcription-PCR for rapid diagnosis of influenza virus infection. J Clin Microbiol. 2003 Aug;41(8):3487-93.
139. Sambol AR, Abdalhamid B, Lyden ER, Aden TA, Noel RK, Hinrichs SH. Use of rapid influenza diagnostic tests under field conditions as a screening tool during an outbreak of the 2009 novel influenza virus: Practical considerations. J Clin Virol. [Article]. 2010 Mar;47(3):229-33.
140. Sandora TJ, Smole SC, Lee GM, Chung S, Williams L, McAdam AJ. Test Characteristics of Commercial Influenza Assays for Detecting Pandemic Influenza a (H1n1) in Children. Pediatr Infect Dis J. [Article]. 2010 Mar;29(3):261-2.
141. Scansen KA, Bonsu BK, Stoner E, Mack K, Salamon D, Leber A, et al. Comparison of Polyurethane Foam to Nylon Flocked Swabs for Collection of Secretions from the Anterior Nares in Performance of a Rapid Influenza Virus Antigen Test in a Pediatric Emergency Department. Journal of Clinical Microbiology. [Article]. 2010 Mar;48(3):852-6.
142. Schultze D, Thomas Y, Wunderli W. Evaluation of an optical immunoassay for the rapid detection of influenza A and B viral antigens. Eur J Clin Microbiol Infect Dis. 2001 Apr;20(4):280-3.
143. Simmerman JM, Chittaganpitch M, Erdman D, Sawatwong P, Uyeki TM, Dowell SF. Field performance and new uses of rapid influenza testing in Thailand. Int J Infect Dis. 2007 Mar;11(2):166-71.
144. Smit M, Beynon KA, Murdoch DR, Jennings LC. Comparison of the NOW Influenza A & B, NOW Flu A, NOW Flu B, and Directigen Flu A+B assays, and immunofluorescence with viral culture for the detection of influenza A and B viruses. Diagn Microbiol Infect Dis. 2007 Jan;57(1):67-70.
145. Steed LL, Salmon VC, Overall Jr JC. Identification of influenza A virus by shell vial culture and two commercially available antigen detection methods. Clinical and Diagnostic Virology. 1994;2(4-5):261-9.
146. Stein J, Louie J, Flanders S, Maselli J, Hacker JK, Drew WL, et al. Performance characteristics of clinical diagnosis, a clinical decision rule, and a rapid influenza test in the detection of influenza infection in a community sample of adults. Ann Emerg Med. 2005 Nov;46(5):412-9.
147. Stripeli F, Sakkou Z, Papadopoulos N, Georgiou V, Gratsia P, Christodoulou I, et al. Performance of rapid influenza testing in hospitalized children. Eur J Clin Microbiol Infect Dis. 2010 Jun;29(6):683-8.
148. Suntarattiwong P, Jarman RG, Levy J, Baggett HC, Gibbons RV, Chotpitayasunondh T, et al. Clinical performance of a rapid influenza test and comparison of nasal versus throat swabs to detect 2009 pandemic influenza

A (H1N1) infection in Thai children. Pediatr Infect Dis J. 2010 Apr;29(4):366-7.
149. Talbot HK, Williams JV, Zhu Y, Poehling KA, Griffin MR, Edwards KM. Failure of routine diagnostic methods to detect influenza in hospitalized older adults. Infect Control Hosp Epidemiol. 2010 Jul;31(7):683-8.
150. Uyeki TM, Prasad R, Vukotich C, Stebbins S, Rinaldo CR, Ferng Y-h, et al. Low Sensitivity of Rapid Diagnostic Test for Influenza. Clinical Infectious Diseases. 2009 May 1;48(9):E89-E92.
151. Velasco JM, Montesa-Develos ML, Jarman RG, Lopez MN, Gibbons RV, Valderama MT, et al. Evaluation of QuickVue influenza A+B rapid test for detection of pandemic influenza A/H1N1 2009. J Clin Virol. 2010 Jun;48(2):120-2.
152. Waner JL, Todd SJ, Shalaby H, Murphy P, Wall LV. Comparison of Directigen FLU-A with viral isolation and direct immunofluorescence for the rapid detection and identification of influenza A virus. J Clin Microbiol. 1991 Mar;29(3):479-82.
153. Watcharananan S, Kiertiburanakul S, Chantratita W. Rapid influenza diagnostic test during the outbreak of the novel influenza A/H1N1 2009 in Thailand: an experience with better test performance in resource limited setting. J Infect. 2010 Jan;60(1):86-7.
154. Weinberg A, Mettenbrink CJ, Ye D, Yang CF. Sensitivity of diagnostic tests for influenza varies with the circulating strains. J Clin Virol. 2005 Jun;33(2):172-5.
155. Weitzel T, Schnabel E, Dieckmann S, Borner U, Schweiger B. Evaluation of a new point-of-care test for influenza A and B virus in travellers with influenza-like symptoms. Clin Microbiol Infect. 2007 Jul;13(7):665-9.
156. Yoo Y, Jang WS, Dae WP, Jeong YK, Hye KS, Lee Y, et al. Clinical evaluation of the SD bioline influenza virus antigen test for rapid detection of influenza viruses A and B in children and adults during the influenza season. Clinical and Vaccine Immunology. 2007 August;14(8):1050-2.
157. Atmar RL, Baxter BD, Dominguez EA, Taber LH. Comparison of reverse transcription PCR with tissue culture and other rapid diagnostic assays for detection of type A influenza virus. Journal of Clinical Microbiology. [Article]. 1996 Oct;34(10):2604-6.
158. Cheng PK, Wong KK, Mak GC, Wong AH, Ng AY, Chow SY, et al. Performance of laboratory diagnostics for the detection of influenza A(H1N1)v virus as correlated with the time after symptom onset and viral load. J Clin Virol. 2010 Feb;47(2):182-5.
159. Chomel JJ, Remilleux MF, Marchand P, Aymard M. Rapid diagnosis of influenza A. Comparison with ELISA immunocapture and culture. J Virol Methods. 1992 Jun;37(3):337-43.
160. Drinka PJ, Krause P, Nest L, Miller J, Gauerke C. Experience with a rapid diagnostic test for influenza. Infection Control and Hospital Epidemiology. [Letter]. 2002 01;23(10):561.
161. Ginocchio CC, Lotlikar M, Falk L, Arora S, Kowerska M, Bornfreund M, et al. Clinical performance of the 3M Rapid Detection Flu A+B Test compared

to R-Mix culture, DFA and BinaxNOW Influenza A&B Test. J Clin Virol. 2009 Jun;45(2):146-9.
162. Landry ML, Cohen S, Ferguson D. Impact of sample type on rapid detection of influenza virus A by cytospin-enhanced immunofluorescence and membrane enzyme-linked immunosorbent assay. Journal of Clinical Microbiology. [Article]. 2000 Jan;38(1):429-30.
163. Landry ML, Ferguson D. Suboptimal detection of influenza virus in adults by the Directigen Flu A+B enzyme immunoassay and correlation of results with the number of antigen-positive cells detected by cytospin immunofluorescence. J Clin Microbiol. 2003 Jul;41(7):3407-9.
164. Landry ML, Cohen S, Ferguson D. Comparison of Binax NOW and Directigen for rapid detection of influenza A and B. J Clin Virol. 2004 Oct;31(2):113-5.
165. Magauran CE, Yen-Lieberman B, Dhar S, Schindler S, Starkey C, Gordon SM. Lessons learned: Moving from a rapid immunoassay to a molecular platform for respiratory viral testing during influenza season. Infectious Diseases in Clinical Practice. 2007 Jan;15(1):22-5.
166. Mehlmann M, Bonner AB, Williams JV, Dankbar DM, Moore CL, Kuchta RD, et al. Comparison of the MChip to viral culture, reverse transcription-PCR, and the QuickVue influenza A+B test for rapid diagnosis of influenza. J Clin Microbiol. 2007 Apr;45(4):1234-7.
167. Ryan-Poirier KA, Katz JM, Webster RG, Kawaoka Y. Application of Directigen FLU-A for the detection of influenza A virus in human and nonhuman specimens. J Clin Microbiol. 1992 May;30(5):1072-5.
168. Steininger C, Redlberger M, Graninger W, Kundi M, Popow-kraupp T. Near-patient assays for diagnosis of influenza virus infection in adult patients. Clinical Microbiology and Infection. 2009;15(3):267-73.
169. van Hal SJ, Foo H, Blyth CC, McPhie K, Armstrong P, Sintchenko V, et al. Influenza outbreak during Sydney World Youth Day 2008: the utility of laboratory testing and case definitions on mass gathering outbreak containment. PLoS One. 2009;4(9):e6620.
170. Vasoo S, Stevens J, Singh K. Rapid antigen tests for diagnosis of pandemic (Swine) influenza A/H1N1. Clin Infect Dis. 2009 Oct 1;49(7):1090-3.
171. Weinberg A, Waker ML. Evaluation of three immunoassay kits for rapid detection of influenza virus A and B. Clinical and Diagnostic Laboratory Immunology. 2005 Mar;12(3):367-70.
172. Apisarnthanarak A, Mundy LM. Rapid testing for pandemic influenza A (H1N1): diagnostic test utility and specimen source. Infect Control Hosp Epidemiol. 2010 Jun;31(6):663-4.
173. Ginocchio CC, Zhang F, Manji R, Arora S, Bornfreund M, Falk L, et al. Evaluation of multiple test methods for the detection of the novel 2009 influenza A (H1N1) during the New York City outbreak. J Clin Virol. 2009 Jul;45(3):191-5.

174. Nougairede A, Ninove L, Zandotti C, Salez N, Mantey K, Resseguier N, et al. Novel Virus Influenza A (H1N1sw) in South-Eastern France, April-August 2009. PLoS One. [Article]. 2010 Feb;5(2).
175. Watanabe M, Nakagawa N, Ito M, Ihara T. Sensitivity of rapid immunoassay for influenza A and B in the early phase of the disease. Pediatr Int. 2009 Apr;51(2):211-5.

Figure 1 Mode of action of RIDTs

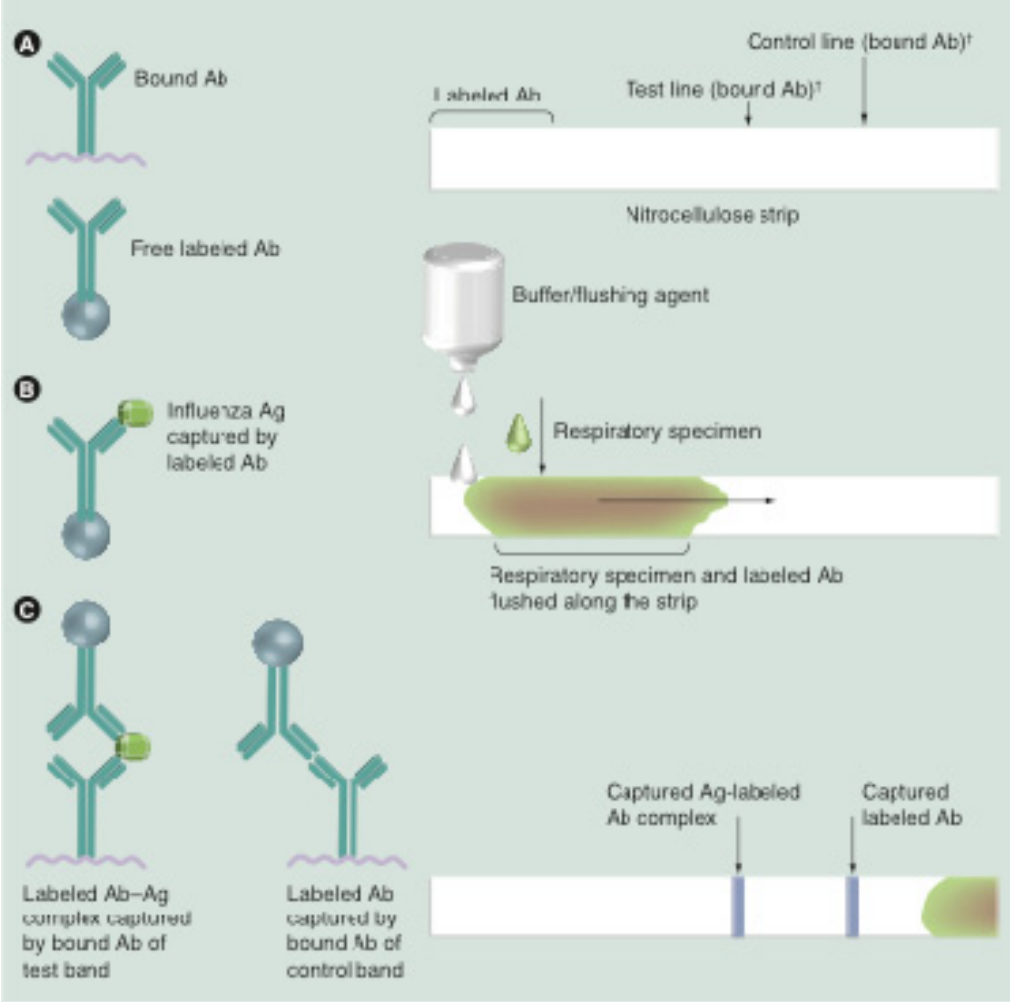


Figure 1. Mode of action of influenza rapid diagnostic tests. (A) Dye-labeled free Abs, specific for the target antigen, are present on the lower end of the nitrocellulose strip. Abs, also specific for the target antigen, are bound to the strip in a thin test line and Abs specific for the dye-labeled Abs are bound to the control line. **(B)** Respiratory specimen and buffer are added to the strip and mix with the dye-labeled free Abs. They are drawn up the strip, across the two lines of bound Abs. **(C)** If influenza antigen are present, some of the dye-labeled Abs will be trapped on the test line. Other dye-labeled Abs are trapped on the control line.
Ab: Antibody; Ag: Antigen.
†Not normally visible.
Reproduced with permission from [5].

From: Chartrand C, Minion J, Pai M. Rapid diagnostics for influenza: what are the options? Future Microbiol. 2010 Oct; 5:1451-5.

Reproduced with permission from: WHO. Use of influenza rapid diagnostic tests. Geneva: World Health Organization; 2010.

Figure 2 Study selection

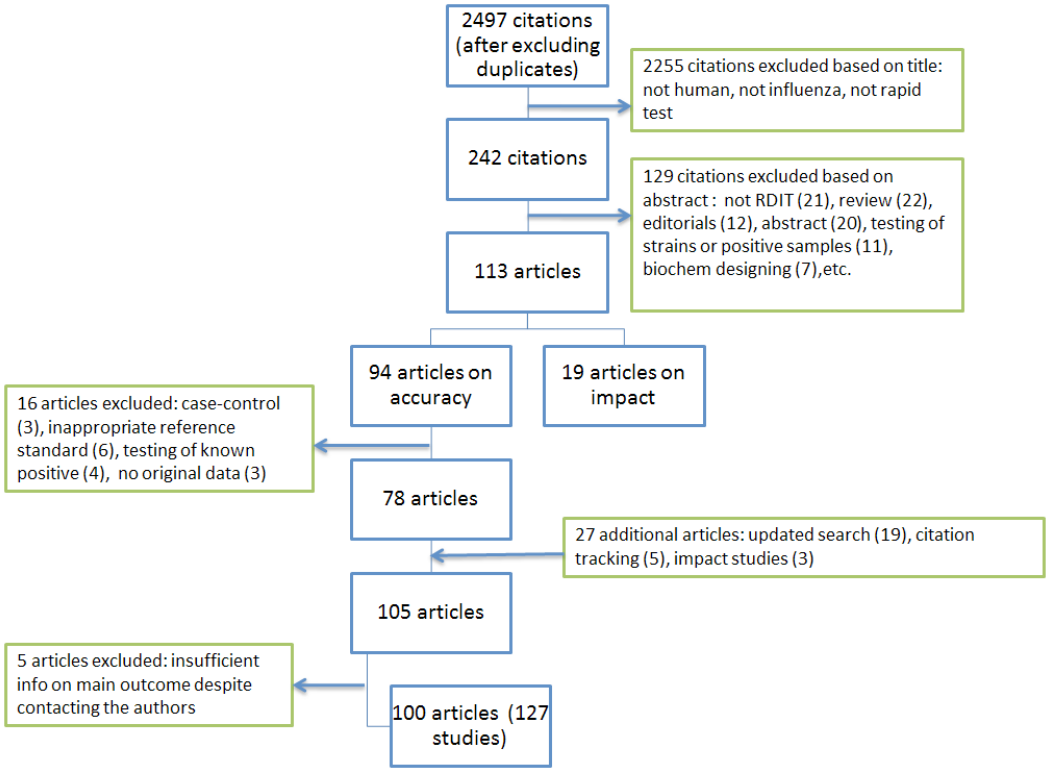
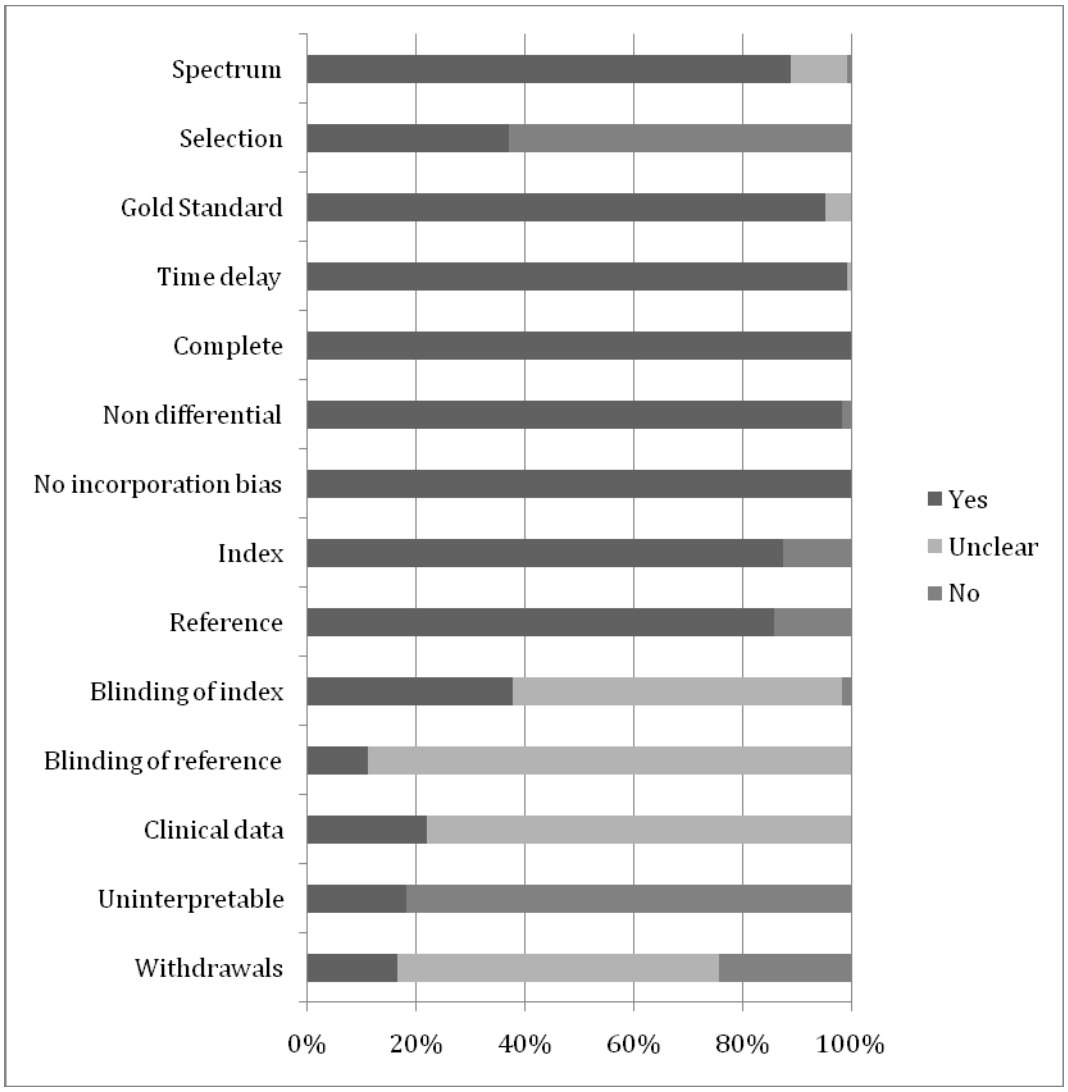


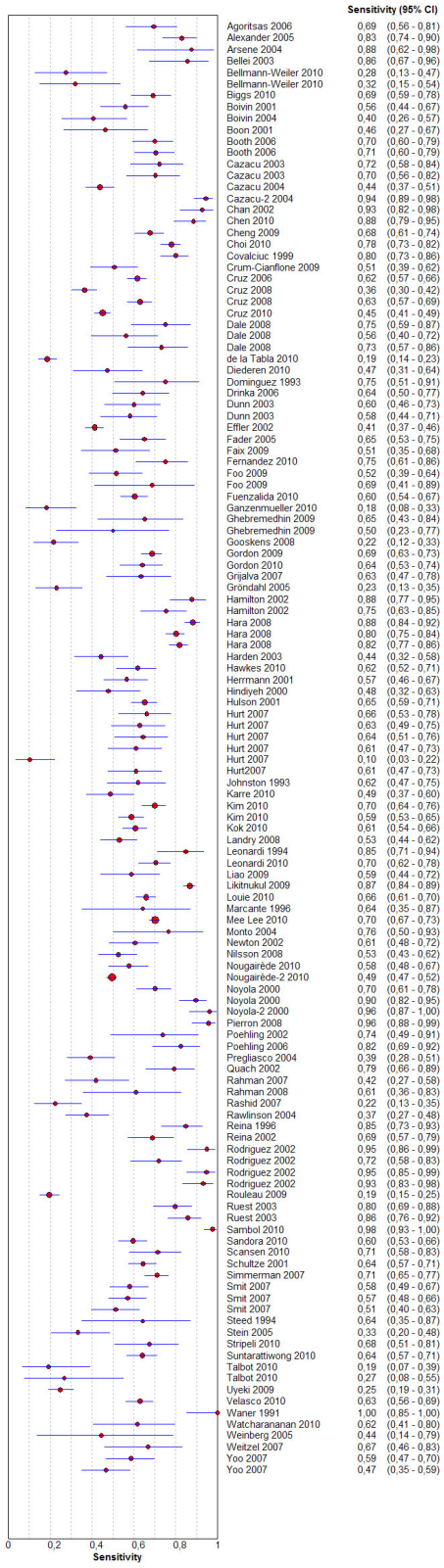
Figure 3 Quality of included studies



Definition of the QUADAS elements
Spectrum: was the spectrum of patients representative of the patients who will receive the test in practice?
Selection: were selection criteria clearly described?
Gold Standard: is the reference standard likely to correctly classify the target condition?
Time delay: is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
Complete: did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?
Non differential: did the patients receive the same reference standard regardless of the index test result?
No incorporation bias: was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
Index: was the execution of the index test described in sufficient detail to permit replication of the test?

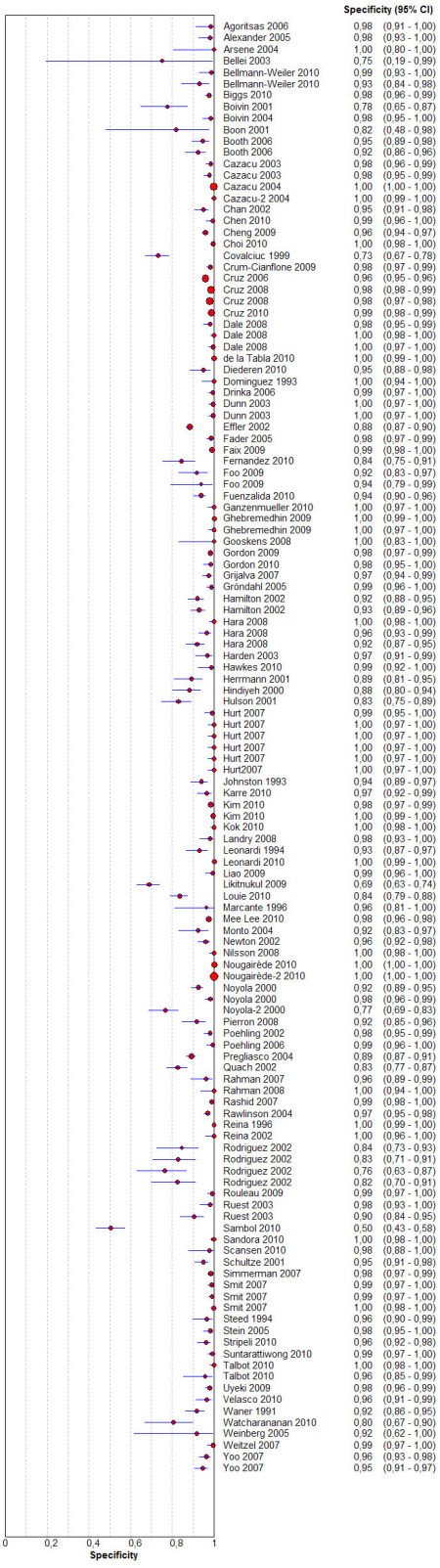
Reference: was the execution of the reference standard described in sufficient detail to permit replication of the test?
Blinding of index: were the index test results interpreted without knowledge of the results of the reference standard?
Blinding of reference: were the reference standard results interpreted without knowledge of the results of the index test?
Clinical data: were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
Uninterpretable: were uninterpretable / intermediate test results reported?
Withdrawals: were withdrawals from the study explained?

Figure 4 Forest plot of sensitivity estimates



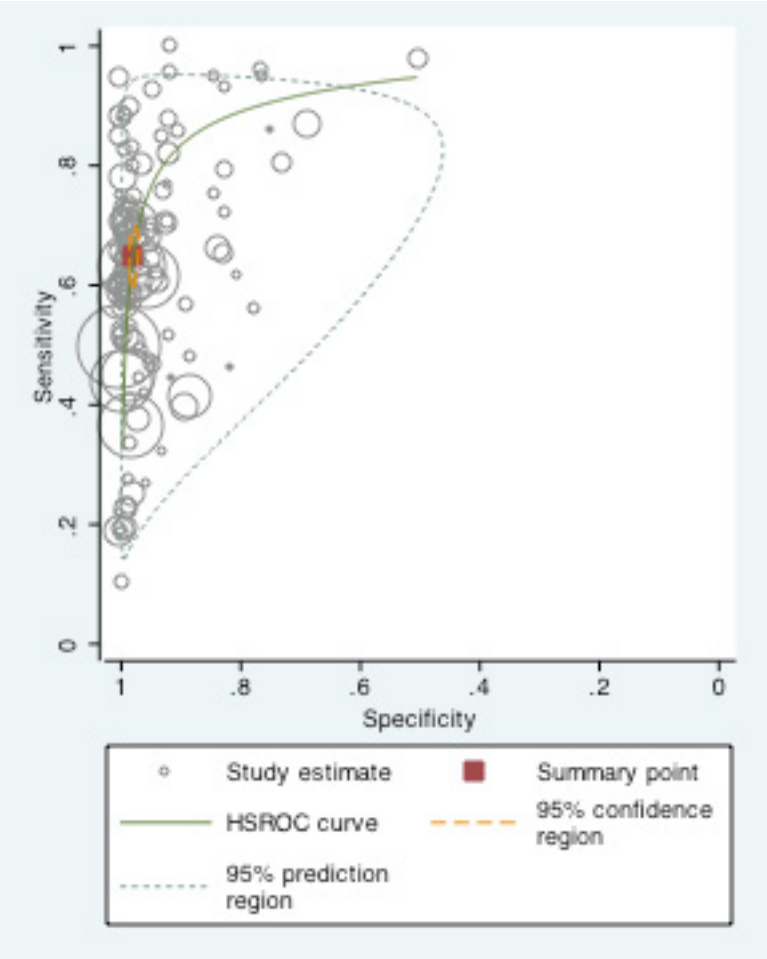
Point estimates of sensitivity from each study are shown as closed circles. Size of the circles is proportionate to the size of the study. Solid lines represent 95% confidence intervals.

Figure 5 Forest plot of specificity estimates



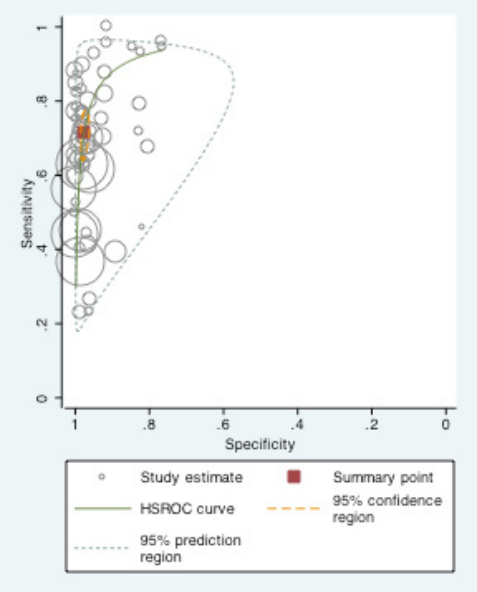
Point estimates of specificity from each study are shown as closed circles. Size of the circles is proportionate to the size of the study. Solid lines represent 95% confidence intervals.

Figure 6 HSROC plot of RIDTs studies (n=127)

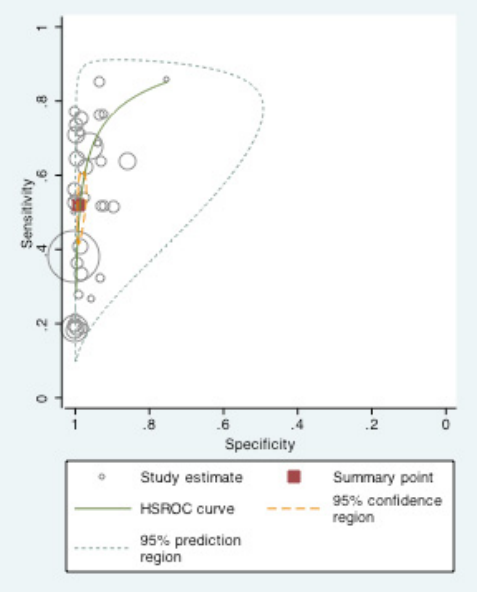


Individual studies are shown as open circles whose size is proportionate to the size of the study. Summary point is shown as a closed square, representing sensitivity estimates pooled using bivariate random effects regression model. The HSROC curve is shown as a full line and is truncated outside the area for which data exist.

Figure 7 HSROC plot of RIDTs in children (n=61) versus in adults (n=36)



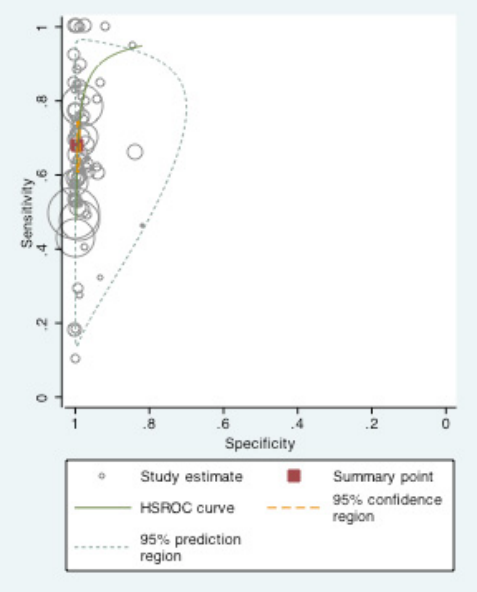
Children



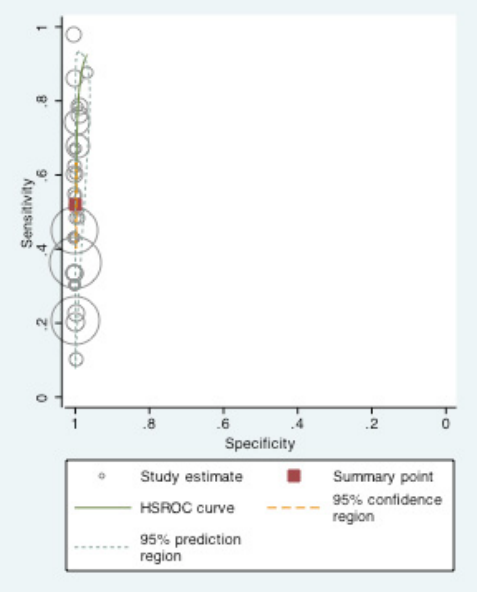
Adults

Individual studies are shown as open circles whose size is proportionate to the size of the study. Summary point is shown as a closed square, representing sensitivity estimates pooled using bivariate random effects regression model. The HSROC curve is shown as a full line and is truncated outside the area for which data exist.

Figure 8 HSROC plot of RIDTs in influenza A (n=74) versus in influenza B (n=37)



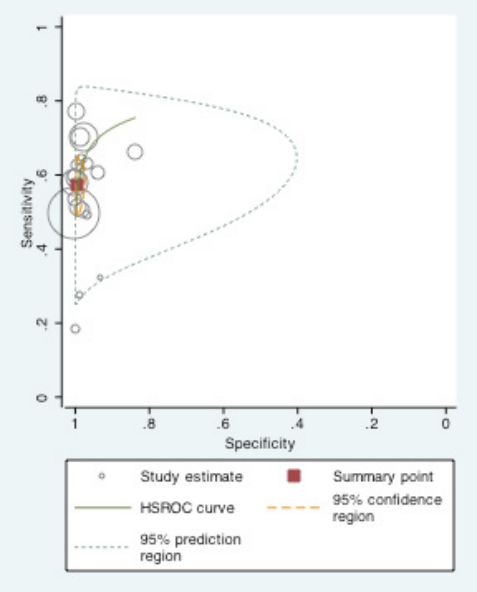
Influenza A



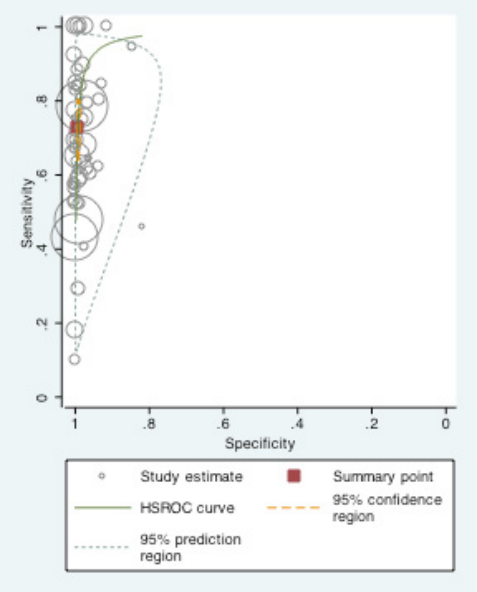
Influenza B

Individual studies are shown as open circles whose size is proportionate to the size of the study. Summary point is shown as a closed square, representing sensitivity estimates pooled using bivariate random effects regression model. The HSROC curve is shown as a full line and is truncated outside the area for which data exist.

Figure 9 HSROC plot of RIDTs in influenza A/H1N1/2009 (n=20) versus in seasonal influenza A (n=54)



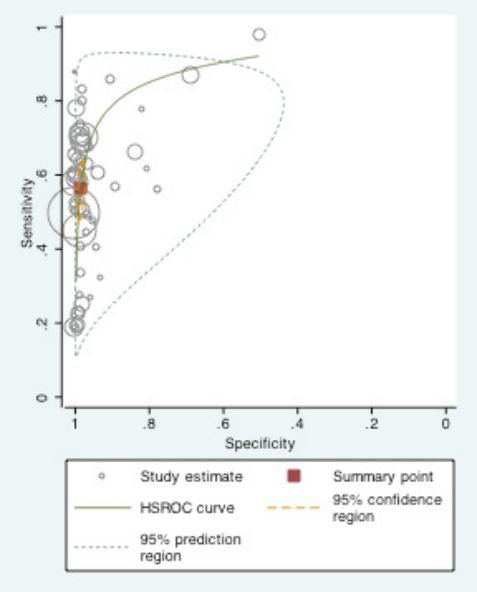
Influenza A/H1N1/2009



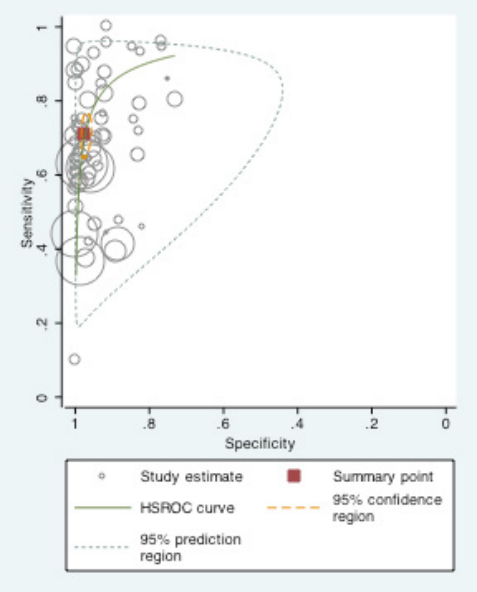
Seasonal influenza A

Individual studies are shown as open circles whose size is proportionate to the size of the study. Summary point is shown as a closed square, representing sensitivity estimates pooled using bivariate random effects regression model. The HSROC curve is shown as a full line and is truncated outside the area for which data exist.

Figure 10 HSROC plot of RIDTs accuracy compared to RT-PCR (n=56) versus culture (n=67)



RT-PCR



Culture

Individual studies are shown as open circles whose size is proportionate to the size of the study. Summary point is shown as a closed square, representing sensitivity estimates pooled using bivariate random effects regression model. The HSROC curve is shown as a full line and is truncated outside the area for which data exist.

Table 1 Commercially available RIDTs

Table 1. Profile of commonly used influenza rapid tests.

Manufacturer	Test	CLIA status[†]	Distinguish between influenza A and B?	Acceptable specimens	Assay time	Sensitivity	Specificity	Ref.
Inverness Medical (ME, USA)	BinaxNOW®, Influenza	CLIA waived	Yes	NS, NA, NW, NPS	15 min	A: 78–82% B: 58–71%	A: 92–94% B: 97%	[102]
	BinaxNOW®, Influenza A & B	CLIA waived	Yes	NS, NA, NW, NPS	15 min	A: 77–83% B: 50–59%	A: 96–99% B: 100%	[7,102]
Becton Dickinson (NJ, USA)	Directigen™ Flu A	CLIA moderate	Detects only influenza A	NPS, NPA, NPW, TS	15 min	67–96%	88–100%	[102]
	Directigen™ Flu A + B	CLIA moderate	Yes	NS, NPS, NPA, NPW, TS, BAL	15 min	A: 77–96% B: 71–88%	A: 90–99.6% B: 97–100%	[8,102]
Biostar Inc. (CO, USA)	Flu OIA®	CLIA waived	No	NA, NPS, TS, sputum	20 min	62–88%	52–80%	[102]
	Flu OIA® A/B	CLIA waived	Yes	NA, NPS, TS, sputum	20 min			
Cuidel (CA, USA)	QuickVue™ Influenza Test	CLIA waived	No	NS, NA, NW, NPS	10 min	73–81%	96–99%	[103]
	QuickVue™ Influenza A+B	CLIA waived	Yes	NS, NA, NW, NPS	10 min	A: 72–94% B: 62–82%	A: 90–99% B: 96–99%	[9,102]
Standard Diagnostic Inc. (South Korea)	SD BIOLINE influenza Ag Test		Yes	NS, NA, NPS, NPA, TS	10 min	91.8%	98.9%	[10]
Remel Inc. (KS, USA)	XPECT Flu® A&B	CLIA moderate	Yes	NW, NPS, TS, sputum, tracheal aspirate, BAL	15 min	A: 89–100% B: 93–100%	A: 100% B: 100%	[11,102]
Zyme Tx Inc. (OK, USA)	ZstatFlu-II test	CLIA waived	No	TS	30 min	65–96%	77–98%	[12,102]

Sensitivity and specificity estimates were derived from package inserts.
[†]Clinical Laboratory Improvement Amendment of 1988. CLIA waived tests can be performed in a clinical or office setting.
BAL: Broncho-alveolar lavage; CLIA: Clinical Laboratory Improvement Amendment; NA: Nasal aspirate; NPA: Nasopharyngeal aspirate; NPS: Nasopharyngeal swab; NPW: Nasopharyngeal wash; NS: Nasal swab; NW: Nasal wash; TS: Throat swab.

From: Chartrand C, Minion J, Pai M. Rapid diagnostics for influenza: what are the options? Future Microbiol. 2010 Oct; 5: 1451-5.

Table 2 Definitions of accuracy outcomes

Sensitivity	Sensitivity refers to the proportion of people with disease who have a positive test result (formula shown below)
Specificity	Specificity refers to the proportion of people without disease who have a negative test result (formula shown below)
Positive predictive value	Positive predictive value is the proportion of patients with positive test results who are correctly diagnosed (formula shown below)
Negative predictive value	Negative predictive value is the proportion of patients with negative test results who are correctly diagnosed (formula shown below)
Positive likelihood ratio (LR+)	LR+ is the ratio of the likelihood of a positive test result when the disease is present compared with when it is absent (formula shown below)
Negative likelihood ratio (LR-)	LR- is the ratio of the likelihood of a positive test result when the disease is absent compared with when it is present (formula shown below)
Indeterminate results	Proportion of results that meet the specific criteria for being classified as "indeterminate"

		Target Disorder		Totals
		Present	Absent	
Diagnostic Test Result	Positive	a	b	a+b
	Negative	c	d	c+d
	Totals	a+c	b+d	a+b+c+d

Sensitivity = a/(a+c)

Positive Predictive Value = a/(a+b)

Specificity = d/(b+d)

Negative Predictive Value = d/(c+d)

Likelihood Ratio for a positive test result = LR+ = sens/(1-spec)

Likelihood Ratio for a negative test result = LR- = (1-sens)/spec

Diagnostic odds ratio (DOR) = (a * d) / (b * c)

Table 3 Study characteristics

Author, Year	Population	Specimen type	RIDT	Reference	Total number of specimens (Ref+/Ref-)	Sensitivity (%) with 95% CI	Specificity (%) with 95% CI
Agoritsas, 2006(61)	Children	NPW	QuickVue	Culture and RT-PCR	59 / 63	69.5 (56.1-80.8)	98.4 (91.5-100)
Alexander, 2005(62)	Children	NPA, NS, TS, BAL	Directigen A+B	RT-PCR	94 / 97	83.0 (73.8-89.9)	97.9 (92.7-99.7)
Arsene, 2004(63)	Children	NS, NA[2]	QuickVue	RT-PCR	16 / 17	87.5 (61.7-98.4)	100 (80.5-100)
Bellei, 2003(64)	Adults	NPS	QuickVue	Culture	28 / 4	85.7 (67.3-96.0)	75.0 (19.4-99.4)
Bellmann-Weiler, 2010[1](65)	Adults	NS, TS	Binax NOW II	RT-PCR	29 / 78	27.6 (12.7-47.2)	98.7 (93.1-100)
			Directigen A+B	RT-PCR	25 / 71	32.0 (14.9-53.5)	93.0 (84.3-97.7)
Biggs, 2010(66)	Mixed	NPS	Binax NOW II	RT-PCR	97 / 469	69.1 (58.9-78.1)	97.7 (95.8-98.8)
Boivin, 2001(67)	Mixed	TS	Flu OIA	RT-PCR	77 / 58	55.8 (44.1-67.2)	77.6 (64.7-87.5)
Boivin, 2004(68)	Children	NPA	Directigen A+B	RT-PCR	42 / 130	40.5 (25.6-56.7)	98.5 (94.6-99.8)
Boon, 2001(69)	Children	NS, TS	Directigen A	Culture	26 / 11	46.2 (26.6-66.6)	81.8 (48.2-97.7)
Booth, 2006(70)	Mixed	NPA, NS, TS	Binax NOW	Culture	93 / 131	69.9 (59.5-79.0)	94.7 (89.3-97.8)
			ImmunoCard	Culture	95 / 129	70.5 (60.3-79.4)	92.2 (86.2-96.2)
Cazacu, 2003(71)	Children	NW	Directigen A+B	Culture	54 / 302	72.2 (58.4-83.5)	98.3 (96.2-99.5)
			QuickVue	Culture	54 / 302	70.4 (56.4-82.0)	97.7 (95.3-99.1)
Cazacu, 2004(72)	Children	NPS, NW, T, BAL, S	Directigen A+B	Culture	219 / 3873	43.8 (37.2-50.7)	99.7 (99.5-99.9)
Cazacu, 2004(73)	Mixed	NPS, NW, TS, T, BAL, S	XPECT Flu	Culture	125 / 275	94.4 (88.8-97.7)	100 (98.7-100)
Chan, 2002(74)	Children	NPA	Directigen A+B	Culture	54 / 196	92.6 (82.1-97.9)	94.9 (90.8-97.5)
Chen, 2010(75)	NR	NS, TS	Flu A Dot-ELISA	Culture	78 / 147	88.5 (79.2-94.6)	99.3 (96.3-100)
Cheng, 2009(76)	Adults	NS, TS	QuickVue A+B	Culture	186 / 812	67.7 (60.5-74.4)	95.8 (94.2-97.1)
Choi, 2010[1](77)	Mixed	NPS	SD Bioline H1N1	RT-PCR	313 / 446	78.0 (72.9-82.4)	99.6 (98.4-99.9)
Covalciuc, 1999(78)	Mixed	NPS, NA, TS, S	Flu OIA	Culture	151 / 253	80.1 (72.9-86.2)	73.1 (67.2-78.5)

Crum-Cianflone, 2009[1](79)	Mixed	NPS	QuickVue A+B	RT-PCR	79 / 492	50.6 (39.1-62.1)	98.2 (96.6-99.2)
Cruz, 2006(80)	Children	NW, NS, T, BAL, S	Binax NOW	Culture	437 / 3946	61.6 (56.8-66.1)	95.8 (95.1-96.4)
Cruz, 2008(81)	Children	NPS, NW, T, BAL, S	XPECT Flu	Culture	259 / 4112	36.3 (30.4-42.5)	98.4 (98.0-98.8)
			Binax NOW II	Culture	283 / 4532	62.9 (57.0-68.5)	97.9 (97.4-98.3)
Cruz, 2010[1](82)	Children	NPS, NW, NS, T, BAL, S	Binax NOW II	RT-PCR	689 / 2341	45.0 (41.2-48.8)	98.6 (98.1-99.1)
Dale, 2008(83)	Adults	NPS, NS	3M	Culture	40 / 202	75.0 (58.8-87.3)	98.0 (95.0-99.5)
			Binax NOW II	Culture	41 / 208	56.1 (39.7-71.5)	100 (98.2-100)
			QuickVue A+B	Culture	41 / 208	73.2 (57.1-85.8)	99.5 (97.4-100)
De la Tabla, 2010[1](84)	Adults	NPS, TS	Clearview	RT-PCR	297 / 698	18.5 (14.3-23.4)	100 (99.5-100)
Diederen, 2010[1](85)	Mixed	NPA	Binax NOW II	RT-PCR	38/ 97	47.4 (31.0-64.2)	94.8 (88.4-98.3)
Dominguez, 1993(86)	Children	NW, NS, TS	Directigen A	Culture	20 / 61	75.0 (50.9-91.3)	100 (94.1-100)
Drinka, 2006(87)	Adults	NPS	Directigen A+B	Culture	53 / 274	64.2 (49.8-76.9)	99.3 (97.4-99.9)
Dunn, 2003(88)	NR	NPS, NW, TS, T, BAL, S	INFLU A+B	Culture	55 / 200	60.0 (45.9-73.0)	99.5 (97.2-100)
			Directigen A+B	Culture	55 / 200	58.2 (44.1-71.3)	99.5 (97.2-100)
Effler, 2002(89)	NR	NPS, TS	Flu OIA	Culture	473 / 1596	41.2 (36.8-45.8)	88.2 (86.5-89.8)
Fader, 2005(90)	Mixed	NA, NW	Binax NOW	Culture	77 / 378	64.9 (53.2-75.5)	98.4 (96.6-99.4)
Faix, 2009[1](91)	NR	NR	QuickVue A+B	RT-PCR	39 /728	51.3 (34.8-67.6)	99.0 (98.0-99.6)
Fernandez, 2010[1](92)	Mixed	NPS, NS	QuickVue A+B	Culture[3]	52 / 95	75.0 (61.1-86.0)	84.2 (75.3-90.9)
Foo, 2009(93)	Adults	NS, TS	QuickVue A+B	Culture and RT-PCR	64 / 73	51.6 (38.7-64.2)	91.8 (83.0-96.9)
			Binax NOW II	Culture and RT-PCR	16 / 32	68.8 (41.3-89.0)	93.8 (79.2-99.2)

Fuenzalida, 2010[1](94)	Mixed	NPA	Binax NOW II	RT-PCR	227 / 285	60.4 (53.7-66.8)	93.7 (90.2-96.2)
Ganzenmueller, 2010[1](95)	Mixed	NPS, TS, BAL	QuickVue A+B	RT-PCR	44 / 128	18.2 (8.2-32.7)	100 (97.2-100)
Ghebremedhin, 2009(96)	Children	NPA, NPS, T, BAL	Actim Influenza	RT-PCR	23 / 450	65.2 (42.7-83.6)	100 (99.2-100)
			Binax NOW II	RT-PCR	14 / 135	50.0 (23.0-77.0)	100 (97.3-100)
Cooskens, 2008(97)	Adults	NPS, NPS, TS	Directigen A+B	RT-PCR	65 / 20	21.5 (12.3-33.5)	100 (83.2-100)
Gordon, 2009(98)	Children	NS	QuickVue A+B	RT-PCR	359/798	68.5(63.4-73.3)	98.1 (96.9-98.9)
Gordon, 2010[1](99)	Children	NS, TS[2]	QuickVue A+B	RT-PCR	92 / 173	64.1 (53.5-73.9)	98.3 (95.0-99.6)
Grijalva, 2007(100)	Children	NS, TS	Mixed tests[4]	Culture and RT-PCR	41 / 229	63.4 (46.9-77.9)	97.4 (94.4-99.0)
Gröndahl, 2005(101)	Children	NPA	Directigen A+B	RT-PCR	61 / 238	23.0 (13.2-35.5)	98.7 (96.4-99.7)
Hamilton, 2002(102)	Children	NA	ZstatFlu	Culture	65 / 235	87.7 (77.2-94.5)	91.9 (87.7-95.1)
			Directigen A+B	Culture	65 / 235	75.4 (63.1-85.2)	92.8 (88.7-95.7)
Hara, 2008(103)	Children	NPA	ESPLINE	Culture	323 / 171	88.2 (84.2-91.5)	100 (97.9-100)
			Directigen A+B	Culture	323 / 171	80.2 (75.4-84.4)	96.5 (92.5-98.7)
			Binax NOW II	Culture	323 / 171	81.7 (77.1-85.8)	91.8 (86.6-95.5)
Harden, 2003(104)	Children	NPA, NS[2]	QuickVue	RT-PCR	61 / 96	44.3 (31.5-57.6)	96.9 (91.1-99.4)
Hawkes, 2010[1](105)	Children	NPS, NS	Binax NOW II	RT-PCR	107 / 71	61.7 (51.8-70.9)	98.6 (92.4-100)
Herrmann, 2001(106)	Mixed	NPA, NPS	Flu OIA	RT-PCR	92 / 92	56.5 (45.8-66.8)	89.1 (80.9-94.7)
HIndiyeh, 2000(107)	NR	NPS, NW, TS, BAL, S	Flu OIA	Culture	44 / 101	47.7 (32.5-63.3)	88.1 (80.2-93.7)
Hulson, 2001(108)	Mixed	TS	ZstatFlu	Culture	241 / 117	65.1 (58.8-71.1)	82.9 (74.8-89.2)
Hurt, 2007(109)	Mixed	NPA, NS, TS, BAL, S	Binax NOW II	Culture	59/ 118	66.1 (52.6-77.9)	99.2 (95.4-100)
			Directigen A+B	Culture	59 / 118	62.7 (49.1-75.0)	100 (96.9-100)
			INFLU A+B	Culture	59 / 118	64.4 (50.9-76.4)	100 (96.9-100)

			ESPLINE	Culture	59 / 118	61.0 (47.4-73.5)	100 (96.9-100)
			Rockeby Influenza	Culture	49 / 128	10.2 (3.4-22.2)	100 (97.2-100)
			QickVue A+B	Culture	59 / 118	61.0 (47.4-73.5)	100 (96.9-100)
Johnston, 1993(110)	NR	NPS, TS	Directigen A	Culture	50 / 161	62.0 (47.2-75.3)	93.8 (88.9-97.0)
Karre, 2010[1](111)	NR	NPW	Directigen A+B	RT-PCR	80 / 145	48.8 (37.4-60.2)	96.6 (92.1-98.9)
Kim, 2010[1](112)	Mixed	NPS	SD Bioline H1N1	RT-PCR	260 / 688	70.0 (64.0-75.5)	98.4 (97.2-99.2)
			SD Bioline	RT-PCR	260 / 688	58.8 (52.6-64.9)	99.6 (98.7-99.9)
Kok, 2010(113)	NR	NS, TS	QuickVue A+B	RT-PCR	269 / 231	60.6 (54.5-66.5)	100 (98.4-100)
Landry, 2008(114)	Mixed	NPS	Binax NOW II	RT-PCR	132 / 105	53.0 (44.2-61.8)	98.1 (93.3-99.8)
Leonardi, 1994(20)	Adults	NPS, TS	Directigen A	Culture	46 / 114	84.8 (71.1-93.7)	93.0 (86.6-96.9)
Leonardi, 2010[1](115)	Mixed	NPS	Directigen A+B	Culture	145 / 468	70.3 (62.2-77.6)	100 (99.2-100)
Liao, 2009(116)	Mixed	NPA, NPS	Directigen A+B	Culture and RT-PCR	51 / 129	58.8 (44.2-72.4)	99.2 (95.8-100)
Likitnukul, 2009[1](117)	Mixed	NS	Mixed tests[5]	RT-PCR	569 / 272	86.8 (83.8-89.5)	68.8 (62.9-74.2)
Louie, 2010[1](118)	Mixed	NPS, NS, TS	QuickVue	RT-PCR	404 / 299	65.8 (61.0-70.5)	83.6 (78.9-87.6)
Marcante, 1996(119)	Mixed	NPA	Directigen A	Culture	14 / 27	64.3 (35.1-87.2)	96.3 (81.0-99.9)
Mee Lee, 2010[1](120)	Mixed	NPS	SD Bioline	RT-PCR	1225 / 929	70.0 (67.4-72.6)	97.5 (96.3-98.4)
Monto, 2004(121)	Adults	TS	Directigen A+B	Culture	17 / 65	76.5 (50.1-93.2)	92.3 (83.0-97.5)
Newton, 2002(122)	Mixed	NW, NS, TS, S	Directigen A	Culture	71 / 219	60.6 (48.3-72.0)	95.6 (92.3-98.1)
Nilsson, 2008(123)	Adults	NPA	Binax NOW II	RT-PCR	120 / 155	52.5 (43.2-61.7)	100 (97.6-100)
Nougairède, 2010[1](124)	Mixed	NS	Directigen A+B	RT-PCR	111 / 1863	57.7 (47.9-67.0)	100 (99.8-100)
Nougairède, 2010[1](125)	Mixed	NS	Directigen A+B	RT-PCR	1615 / 5844	49.4 (46.9-51.9)	100 (99.9-100)
Noyola, 2000(126)	Children	NA, NW	ZstatFlu	Culture	124 / 355	70.2 (61.3-78.0)	92.4 (89.1-94.9)
			Directigen A	Culture	97 / 320	89.7 (81.9-94.9)	98.1 (96.0-99.3)

Noyola, 2000(127)	Children	NA, NW, TS	ZstatFlu	Culture	51 / 145	96.1 (86.5-99.5)	76.6 (68.8-83.2)
Pierron, 2008(53)	Children	NPA	QuickVue	Culture	69 / 108	95.7 (87.8-99.1)	91.7 (84.8-96.1)
Poehling, 2002(128)	Children	NS	QuickVue	RT-PCR	19 / 214	73.7 (48.8-90.9)	98.1 (95.3-99.5)
Poehling, 2006(14)	Children	NS, TS[2]	QuickVue	Culture and RT-PCR	51 / 154	82.4 (69.1-91.6)	99.4 (96.4-100)
Pregliasco, 2004(129)	Children	NS, TS	QuickVue	Culture	74 / 770	39.2 (28.0-51.2)	89.1 (86.7-91.2)
Quach, 2002(51)	Children	NPA	QuickVue	Culture	53 / 247	79.2 (65.9-89.2)	82.6 (77.3-87.1)
Rahman, 2007(130)	Mixed	NPS	Directigen A+B	Culture	43 / 75	41.9 (27.0-57.9)	96.0 (88.8-99.2)
Rahman, 2008(131)	Mixed	NPS	Binax NOW	RT-PCR	18 / 55	61.1 (35.7-82.7)	100 (93.5-100)
Rashid, 2007(132)	Mixed	NS	QuickVue	RT-PCR	58 / 497	22.4 (12.5-35.3)	99.0 (97.7-99.7)
Rawlinson, 2004(133)	Mixed	NPA, TS	ZstatFlu	Culture	91 / 495	37.4(27.4-48.1)	97.0 (95.1-98.3)
Reina, 1996(134)	Children	NPA	Directigen A	Culture	59 / 318	84.7 (73.0-92.8)	100 (98.8-100)
Reina, 2002(135)	Mixed	NPA, TS	Directigen A+B	Culture	74 / 86	68.9 (57.1-79.2)	100 (95.8-100)
Rodriguez, 2002(136)	Children	NW, NS, TS	Directigen A	Culture[3]	58 / 58	94.8 (85.6-98.9)	84.5 (72.6-92.7)
			ZstatFlu	Culture[3]	57 / 58	71.9 (58.5-83.0)	82.8 (70.6-91.4)
			QuickVue	Culture[3]	57 / 55	94.7 (85.4-98.9)	76.4 (63.0-86.8)
			Flu OIA	Culture[3]	58 / 57	93.1 (83.3-98.1)	82.5 (70.1-91.3)
Rouleau, 2009(137)	Mixed	NPA	QuickVue A+B	RT-PCR	267 / 221	19.5 (14.9-24.7)	99.1 (96.8-99.9)
Ruest, 2003(138)	Mixed	NPA	Directigen A+B	RT-PCR	79 / 105	79.7 (69.2-88.0)	98.1 (93.3-99.8)
			QuickVue	RT-PCR	84 / 115	85.7 (76.4-92.4)	90.4 (83.5-95.1)
Sambol, 2010[1](139)	Mixed	NPS, NW, NS	Mixed tests[6]	RT-PCR	130 / 206	97.7 (93.4-99.5)	50.5 (43.5-57.5)
Sandora, 2010[1](140)	Children	NPS	Binax NOW II	RT-PCR	208 / 333	59.6 (52.6-66.3)	99.7 (98.3-100)
Scansen, 2010(141)	Children	NPS, NS[2]	QuickVue A+B	RT-PCR	56 / 44	71.4 (57.8-82.7)	97.7 (88.0-99.9)
Schultze, 2001(142)	Mixed	NPS, NA, TS, S	Flu OIA	Culture[3]	205 / 195	64.4 (57.4-70.9)	94.9 (90.8-97.5)
Simmerman, 2007(143)	Mixed	NPS, NS[2]	QuickVue	RT-PCR	252 / 840	71.0 (65.0-76.6)	98.5 (97.4-99.2)
Smit, 2007(144)	Mixed	NPS, NW, TS	Binax NOW II	Culture	119 / 402	58.0 (48.6-67.0)	98.8 (97.1-99.6)

			Binax NOW	Culture	119 / 402	57.1 (47.7-66.2)	99.0 (97.5-99.7)
			Directigen A+B	Culture	78 / 331	51.3 (39.7-62.8)	99.7 (98.3-100)
Steed, 1994(145)	Mixed	NPA, NPW, NPS, NA, NW, NS, TS	Directigen A	Culture	14 / 83	64.3 (35.1-87.2)	96.4 (89.8-99.2)
Stein, 2005(146)	Adults	NPW	QuickVue	RT-PCR	48 / 169	33.3 (20.4-48.4)	98.2 (94.9-99.6)
Stripeli, 2010(147)	Children	NPA, NS[2]	QuickVue	RT-PCR	40 / 177	67.5 (50.9-81.4)	96.0 (92.0-98.4)
Suntarattiwong, 2010[1](148)	Children	NS, TS	QuickVue A+B	RT-PCR	181 / 237	64.1 (56.6-71.1)	99.2 (97.0-99.9)
Talbot, 2010(149)	Adults	NS, TS	QuickVue A+B	RT-PCR	26 / 201	19.2 (6.6-39.4)	100 (98.2-100)
			Binax NOW II	RT-PCR	15 / 46	26.7 (7.8-55.1)	95.7 (85.2-99.5)
Uyeki, 2009(150)	Mixed	NS, TS	QuickVue A+B	RT-PCR	210 / 447	24.8 (19.1-31.2)	97.8 (95.9-98.9)
Velasco, 2010[1](151)	Mixed	NS	QuickVue A+B	RT-PCR	226 / 114	62.8 (56.2-69.1)	96.5 (91.3-99.0)
Waner, 1991(152)	Children	NPW, NPS, TS, T, BAL, S	Directigen A	Culture	23 / 167	100 (85.2-100)	91.6 (86.3-95.3)
Watcharananan, 2010[1](153)	Mixed	NPS	QuickVue A+B	RT-PCR	26 / 51	61.5 (40.6-79.8)	80.4 (66.9-90.2)
Weinberg, 2005(154)	NR	NPS, NW, TS, T, BAL, S	Binax NOW	Culture	9 / 12	44.4 (13.7-78.8)	91.7 (61.5-99.8)
Weitzel, 2007(155)	Mixed	NS	ImmunoCard	Culture and RT-PCR	27 / 176	66.7 (46.0-83.5)	99.4 (96.9-100)
Yoo, 2007(156)	Mixed	NPA, NS, TS[2]	SD Bioline	Culture	75 / 220	58.7 (46.7-69.9)	96.4 (93.0-98.4)
			QuickVue	Culture	75 / 220	46.7 (35.1-58.6)	94.5 (90.7-97.2)

Excluded studies	**Reasons for exclusion**
Atmar, 1996(157)	RIDT portion already reported by Dominguez, 1993
CDC, 2009(32)	Testing of known positive samples
Cheng, 2010(158)	Testing of known positive samples
Chomel, 1992(159)	ELISA used as reference standard
Drinka, 2002(160)	No original data provided (letter to the editor)
Ginocchio, 2009(161)	Composite reference standard that included the RIDT
Landry, 2000(162)	Immunofluorescence used as reference standard
Landry, 2003(163)	Immunofluorescence used as reference standard

Landry, 2004(164)	Case-control design
Magauran, 2007(165)	Reference standard performed only for RIDT negative specimen
Mehlmann, 2007(166)	Case-control design
Ryan-Poirier, 1992(167)	Testing of known positive samples
Steininger, 2009(168)	Testing of known positive samples
Van Hal, 2009(169)	Modeling using previously published data
Vasoo, 2009(170)	Case-control design
Weinberg, 2005(171)	Composite reference standard that included the RIDT
Studies excluded because of insufficient information to construct the main two-by-two table despite contacting the authors Apisarnthanarak, 2010(172) D'Heilly, 2008(57) Ginocchio, 2009(173) Nougairède, 2010(174) Watanabe, 2009(175)	

[1]Studies conducted during the pandemic of influenza A/H1N1/2009
[2]Differences in the type of specimen used for the RIDT and for the reference standard
[3]Reference standard was culture and/or immunofluorescence positive
[4]Mixed tests: Directigen A+B, QuickVue A+B, Directigen A, Binax NOW
[5]Mixed tests: QuickVue A+B, SD Bioline
[6]Mixed tests: Binax NOW II, Tru Flu, XPECT Flu, QuickVue A+B
NPA: nasopharyngeal aspirate, NPS: nasopharyngeal swab, NPW: nasopharyngeal wash, NA: nasal aspirate, NS: nasal swab, NW: nasal wash, TS: throat swab, T: tracheal specimen, BAL: broncho-alveolar lavage, S: sputum
Binax NOW: Binax NOW Flu A and Flu B, Binax NOW II: Binax NOW Influenza A&B, Directigen A: Directigen Flu A, Directigen A+B: Directigen Flu A+B, QuickVue: QuickVue Influenza Test, QuickVue A+B: QuickVue Influenza A+B, ZstatFlu: ZstatLFlu test, SD Bioline: SD Bioline influenza Ag test, SD Bioline H1N1: SD Bioline Influenza Ag A/B/A(H1N1)Pandemic, XPECT Flu: XPECT Flu A&B, ImmunoCard: ImmunoCard STAT! Flu A and B, 3M: 3M rapid detection influenza A+B, INFLU A+B: INFLU A.B – Quick, Actim influenza: Actim influenza A&B test, ESPLINE: ESPLINE influenza A and B – N, Rockeby influenza: Rockeby Influenza A Antigen Test, Clearview: Clearview exact influenza A&B test

Table 4 Characteristics of included studies (n=127 studies)

	Numbers of studies (%)
Population	
Children	45 (35.4%)
Adults	18 (14.2%)
Mixed	64 (50.4%)
Setting	
Emergency room	8 (6.3%)
Outpatient clinics	30 (23.6%)
Hospital wards	11 (8.7%)
Nursing home	4 (3.1%)
Mixed	31 (24.4%)
Clear definition of ILI[1]	
Yes	38 (29.9%)
Study conducted during the H1N1 pandemic	
Yes	27 (21.3%)
Commercial RDITs	
Binax NOW Flu A and Flu B	6 (4.7%)
Binax NOW Influenza A&B	17 (13.4%)
Directigen Flu A	11 (8.7%)
Directigen Flu A + B	23 (18.1%)
FLU OIA	7 (5.5%)
QuickVue Influenza Test	18 (14.2%)
QuickVue Influenza A+B	18 (14.2%)
ZstatFlu test	6 (4.7%)
Mixed tests[2]	3 (2.4%)
Others: SD Bioline influenza Ag test (3), XPECT Flu A&B (2), ImmunoCard STAT! Flu A and B Plus (2), 3M rapid detection influenza A+B (1), INFLU A B-Quick (2), Actim Influenza A&B test (1), ESPLINE Influenza A&B (2), Rockeby Influenza A Antigen tests (1), Flu A Dot-ELISA (1), SD Bioline Influenza Ag A/B/H1N1 Pandemic (2), Clearview Exact Influenza A&B (1)	
Reference standard	
RT-PCR	53 (41.7%)
Culture	67 (52.8%)
Both	7 (5.5%)
Type of specimen	
Throat swab	3 (2.4%)
Nasal swab	8 (6.3%)
Nasopharyngeal swab	14 (11.0%)
Nasopharyngeal aspirate	16 (12.6%)
Nasal swab and throat swab	17 (13.4%)
Mixed	61 (48.0%)

Others: Nasal aspirate (2), Nasal wash (2), Nasopharyngeal wash (3)	
Duration of symptoms before testing	
Any information	17 (13.4%)
Point-of-care testing	
Yes	25 (19.7%)

[1]Article provided a clear definition of the clinical symptoms on the basis of which patients were recruited for the study.
[2]More than one RIDT was used concomitantly without separate data on the results of each test.

Table 5 Accuracy estimates from subgroup analyses

Study characteristics (# of studies)	Pooled Sensitivity (%) with 95% CI	Pooled Specificity (%) with 95% CI
Population		
Children (n=61)	71.1 (65.6-76.1)	97.6 (96.6-98.4)
Adults (n=36)	51.6 (43.9-59.1)	98.7 (97.5-99.3)
Virus type		
Influenza A (n=74)	68.0 (62.3-73.1)	99.3 (98.9-99.5)
Influenza B (n=37)	51.8 (42.8-60.6)	99.8 (99.7-99.9)
Pandemic A/H1N1/2009(n=20)	56.9 (50.9-62.6)	99.3 (98.1-99.7)
Seasonal influenza A (n=54)	72.8 (65.9-78.8)	99.2 (98.7-99.5)
Seasonal A H3N2 (n=34)	73.1 (63.2-81.2)	99.2 (98.5-99.5)
Seasonal A H1N1 (n=11)	66.7 (52.1-78.7)	99.6 (98.9-99.9)
Index test		
Binax NOW Flu A and B (n=6)	61.3 (57.0-65.4)	97.6 (95.3-98.8)
Binax NOW Influenza A&B (n=17)	57.6 (51.4-63.6)	98.4 (97.2-99.1)
Directigen Flu A (n=11)	79.4 (67.2-87.9)	96.3 (92.5-98.3)
Directigen Flu A+B	60.4 (51.6-68.4)	99.2 (98.3-99.7)
QuickVue Influenza Test (n=18)	70.7 (58.8-80.4)	95.5 (92.4-97.4)
QuickVue Influenza A+B (n=18)	53.5 (43.7-63.1)	98.5 (96.8-99.3)
Flu OIA (n=7)	65.0 (49.4-77.9)	86.1 (79.9-90.6)
ZstatFlu (n=6)	75.0 (55.5-87.8)	89.7 (82.5-94.1)
Reference standard		
RT-PCR (n=56)	56.0 (49.7-62.1)	98.9 (97.6-99.2)
Culture (n=67)	71.0 (65.9-75.6)	97.3 (96.0-98.2)
Type of specimen		
Nasopharyngeal aspirate (n=19)	68.9 (56.2-79.3)	97.4 (95.3-98.6)
Nasopharyngeal swab (n=17)	67.9 (62.0-73.4)	98.1 (95.7-99.2)
Nasal swab (n=9)	64.4 (51.3-75.5)	99.3 (95.5-99.9)
Throat swab (n=8)	64.5 (42.4-81.9)	92.2 (81.0-97)
Testing at the point-of-care		
POCT (n=27)	62.3 (51.8-71.7)	96.7 (94.9-98.0)
Not POCT (n=103)	64.9 (60.5-69.1)	98.3 (97.5-98.8)
Setting		
Emergency room (n=9)	71.4 (61.1-79.8)	97.6 (93.6-99.1)
Hospitalized patients (n=11)	60.7 (41.4-77.1)	98.0 (96.9-98.7)
Outpatient clinics (n=32)	66.6 (58.7-73.7)	96.0 (93.6-97.6)
Nursing home (n=4)	64.6 (35.5-85.9)	98.7 (89.1-99.6)
QUALITY ISSUES		
Spectrum of disease		

During influenza season (n=114)	64.2 (59.8-68.4)	98.0 (97.2-98.6)
Outside influenza season (n=13)	68.5 (56.8-78.3)	98.4 (96.9-99.2)
Patient selection		
ILI defined[1] (n=38)	61.5 (53.2-69.2)	97.4 (95.6-98.4)
ILI not defined (n=89)	66.0 (61.3-70.4)	98.3 (97.5-98.8)
Blinding		
Any blinding reported (n=52)	68.7 (61.7-75.0)	96.5 (94.8-97.7)
No blinding reported (n=77)	61.0 (56.0-65.8)	98.7 (98.0-99.2)
Handling of indeterminate results		
Reported (n=23)	67.5 (58.7-75.2)	98.2 (96.6-99.1)
Not reported (n=104)	64.0 (59.4-68.4)	98.0 (97.2-98.6)
Industry sponsoring		
Sponsored (n=26)	71.7 (64.1-78.2)	97.3 (95.5-98.5)
Not sponsored (n=101)	62.7 (58.0-67.2)	98.2 (97.4-98.8)

[1]Article provided a clear definition of the clinical symptoms on the basis of which patients were recruited for the study.

Table 6 Studies with subgroups for duration of symptoms

Study	Duration[1]	Sensitivity (%) with 95% CI	Specificity (%) with 95% CI
Gordon, 2009	Day 1	51.9 (40.3-63.3)	98.4 (95.3-99.7)
	Day 2	75.1 (68.3-81.1)	97.9 (96.0-99.1)
	Day 3	74.2 (62.0-84.2)	97.9 (94.1-99.6)
	Day 4	57.9 (33.5-79.7)	98.6 (94.2-100)
Gordon, 2010	< 24h	41.7 (22.1-63.4)	97.9 (88.9-99.9)
	≥ 24h	72.1 (59.9-82.3)	98.4 (94.3-99.8)
Nilsson, 2008	1-3 days	71.4 (58.7-82.1)	100 (95.1-100)
	1-5 days	62.8 (51.7-73.0)	100 (96.7-100)
	> 5 days	13.8 (3.9-31.7)	100 (90.0-100)
Phoeling, 2002	< 4 days	100 (63.1-100)	96.6 (90.4-99.3)
	≥ 4days	54.5 (23.4-83.3)	98.4 (94.4-99.8)
Stein, 2005	< 48h	58.3 (27.7-84.8)	96.2 (80.4-99.9)
	> 48h	25.0 (12.1-42.2)	98.6 (95.0-99.8)
Stripeli, 2010	< 48h	75.0 (42.8-94.5)	100 (92.1-100)
	≥ 48h	65.4 (44.3-84.8)	94.2 (88.4-97.6)

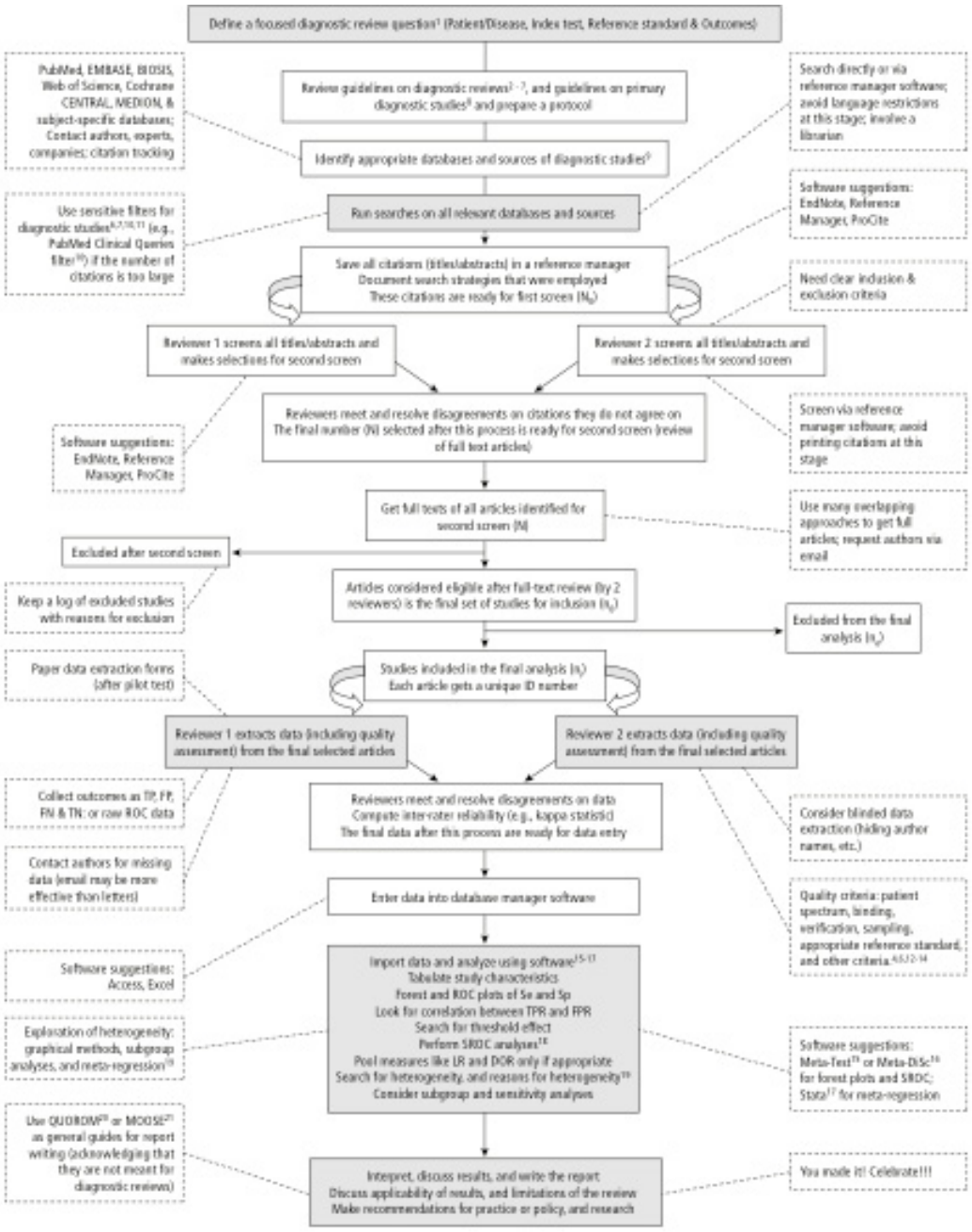
[1]Duration of clinical symptoms at the time of testing by the RIDT

Table 7 Results from the meta-regression analysis

Covariates	Relative DOR (95% CI)
Population	
Children (1) versus Adults (0)	2.67 (1.17-6.11)
Mixed (2) versus Adults (0)	1.27 (0.58-2.82)
Reference standard	
Culture (1) versus RT-PCR (0)	1.81 (0.94-3.50)
Pandemic Influenza A/H1N1/2009	
A/H1N1/2009 (1) versus seasonal influenza (0)	1.72 (0.83-3.55)
Point-of-care	
RIDT performed at POC (1) versus elsewhere (0)	0.81 (0.39-1.66)
Selection of patients	
Clear definition of ILI (1) versus none (0)	0.79 (0.39-1.57)
Industry	
Industry sponsored (1) versus not (0)	1.44 (0.72-2.72)

Meta-regression was performed using an extension of the SROC model, a linear regression model, in which individual studies are the unit of analysis, that allows for the inclusion of covariates to explain the variability in the dependant variable, the diagnostic odds ratio. The coefficient of each covariate, after antilogarithmic transformation, can be interpreted as a relative DOR. A RDOR is the ratio of two DORs and indicates the diagnostic accuracy of a test in studies with a particular characteristic (covariate=1) relative to the accuracy of the test in studies without that characteristic (covariate=0). Values above 1.0 indicate higher diagnostic accuracy in studies with a certain characteristics than in those without it, while the reverse holds true for value below 1.0.

Appendix 1 Protocol for a Diagnostic Meta-Analysis



From: Pai M, McCulloch M, Enanoria W, Colford JM, Jr. Systematic reviews of diagnostic test evaluations: What's behind the scenes? ACP J Club. 2004 Jul-Aug;141(1):A11-3.

Appendix 2 Review protocol

PROTOCOL

Rapid diagnostic tests for the diagnosis of influenza: a systematic review

Caroline Chartrand[1], Jessica Minion[1], Charles Frenette[1],Thimothy Brewer[1], Keertan Dheda[2], Francine Ducharme[3], Madhukar Pai[1]

[1]McGill University, Montreal, Canada

[2]Unviversity of Cape Town, South Africa

[3]University of Montreal, Montreal, Canada

Contact:
Madhukar Pai, MD, PhD
Assistant Professor & CHIR Investigator
Dept, of Epidemiology, Biostatistics & Occupationnal Health
McGill University
1020 Pine Avenue West
Montreal, Canada H3A 1A2
Tel: 514-398-5422
Fax: 514-398-4503
Email: madhukar.pai@mcgill.ca

BACKGROUND

Influenza occurs in yearly epidemics throughout the world. Most human cases of influenza are due to either influenza A or influenza B viruses. Influenza A is classified into subtypes according to the hemagglutinin and neuraminidase glyproteins present on its membrane. Common circulating subtypes currently include H1N1, H2N2 and H3N2. These subtypes encompass many different strains including the recent pandemic-causing 2009 strain of H1N1. It is estimated that 5 to 20% of US residents contract influenza annually (1). Although the majority of those infected with influenza exhibit a self-limited, acute febrile illness or are asymptomatic, some may develop severe disease or complications leading to hospitalization or death. For example, in the United States alone, more than 200 000 individuals are hospitalized each year for complications related to influenza and 36 000 die of influenza-related causes annually (1). The recent pandemic of 2009 H1N1influenza affected more than 211 countries worldwide with at least 15174 deaths in laboratory confirmed cases, which likely underestimate the actual number of death related to H1N1 as many were never tested for influenza (2).

Early identification of influenza as the cause of an acute febrile respiratory illness might lead to earlier antiviral treatment aimed at decreasing duration, severity and complications of the disease and more appropriate use of outpatient services and antibiotics treatment. A clinical diagnosis of influenza, based on the presence of fever and cough or sore throat, is however difficult because many other viruses such as respiratory syncytial virus (RSV), adenoviruses, rhinoviruses, parainfluenza, and bacterial agents including *Chlamydophilaa pneumoniae*, *Mycoplasma pneumoniae*, *Streptococcus pneumoniae* and *Legionnella* species can produce similar symptoms. The sensitivity and specificity of the case definition of influenza-like illness has been estimated to be 64% and 67%, respectively (3). Thus, when the diagnosis of influenza impacts clinical management, as in the decision to start antiviral treatment in a patient belonging to a high-risk group, clinicians must rely on laboratory support to confirm or refute their clinical suspicion.

Although high sensitivity and specificity is claimed, the gold-standard viral culture and the relatively newly developed real-time polymerase chain reaction (RT-PCR) do not provide clinicians with a time-appropriate answer to the question: “Does this patient have influenza?” and are not available in all settings. Thus, an array of rapid diagnostic tests for influenza has been developed in recent years that can give results within 15-30 min and so could potentially have an impact on clinical decision-making if used at the point-of-care (POC). However, the reported sensitivity and specificity of these tests vary greatly, ranging from 40-90% and 75-100%, respectively (1,4). There

are also limited data on whether these rapid tests actually impact clinical decisions or improve patient-important outcomes.

The avian flu and swine flu epidemics of the past few years have brought renewed interest in influenza. Rapid diagnostic tests are part of the arsenal of tests recommended by the US Centers for Diseases Control (CDC) and Infectious Disease Society of America (IDSA) (1,4). Given their widespread use, it is important to systematically review the existing evidence on the accuracy and clinical impact of these tests for influenza diagnosis. To our knowledge, there has been no systematic review on this topic.

OBJECTIVES

Primary objective

To summarize available evidence on the diagnostic accuracy of rapid diagnostic tests for influenza compared to a reference standard (viral culture or RT-PCR) in adults and children with an influenza-like illness and to evaluate patient and test factors associated with higher accuracy (e.g. patient age, type of specimen, duration of illness prior to testing, etc.).

Secondary objective

To summarize available evidence on clinical impact of rapid diagnostic tests in adults and children (i.e. whether these tests actually change clinical decisions, alter treatment plan, reduce morbidity and mortality, etc.).

METHODS

This systematic review will be done using standard methods used for the review of diagnostic test accuracy (5, 6).

Criteria for considering studies for this review

Types of studies

For the primary objective, studies assessing the diagnostic accuracy of rapid diagnostic tests for influenza against a reference standard (viral culture or RT-PCR) will be included. Studies using a case-control design (testing of known positive and negative stored samples) and studies evaluating only sensitivity, through the used of known positive stored samples, will be excluded. The different rapid diagnostic tests may be tested alone against a reference test or in parallel with some other rapid tests. Sufficient data for the construction of a two-by-two table must be reported for inclusion of the studies in the final analysis.

For the secondary objective, studies that assess the clinical impact of rapid diagnostic tests will be included, including those that evaluate impact of rapid diagnostic tests on clinical decisions, on treatment choices and on patient outcomes (e.g. duration of symptoms, mortality).

Participants

Studies will be included if they evaluate patients presenting for an influenza-like illness (defined by the CDC and WHO as fever higher than 37.8°C (100°F) plus either cough or sore throat (1,2)), during the influenza season. No exclusion will be made based on the age of the patients, co-morbidities of the patients, setting of the study or circulating strains of influenza.
We will also include purely laboratory based studies if they include all specimen sent for influenza testing to their laboratory, since those will be assumed to come from patients in which the diagnostic of influenza is entertained. Studies using patients with voluntary infection will be excluded since they don't represent real life use of the rapid diagnostic tests.

Diagnostic tests for influenza

Diagnostic tests for influenza can be broadly classified into the following three categories (Appendix II):

1) Reference tests. This category includes viral culture and RT-PCR
2) Rapid diagnostic tests. These are commercially developed assays for which results are available in less than 30 minutes. Some rapid tests can be performed at the point-of-care without any laboratory requirements.
3) Other laboratory tests. This category includes immunofluorescence and direct fluorescent antibody staining, enzyme immunoassay (EIA) and serology.

This review addresses rapid diagnostic tests, which we define as any commercially available assay capable of giving result in 30 minutes or less. A table summarizing the different diagnostic tests for influenza is included in appendix II.

Index tests

For the purpose of this review, we define rapid diagnostic tests as tests that detect influenza A or A/B within 30 minutes by identifying influenza viral antigen or viral neuraminidase activity in a clinical respiratory specimen. This review will be limited to commercially developed assays regardless of whether they are FDA-approved or not. In-house tests and pre-commercial versions will be excluded from this analysis since they are not routinely available to most healthcare providers assessing patients with possible influenza, may not be standardized and may result in inconsistent accuracy

estimates. Commercial assays included in the review include, but are not limited to:

- BinaxNOW Influenza A&B (Inverness Medical Innovations Inc., MA, USA)
- Directigen Flu A and A+B (Becton-Dickinson, NJ, USA)
- Flu OIA and Flu OIA A/B (Thermo BioStar, CO, USA)
- QuickVue Influenza Test, Influenza A&B Test (Quidel Corporation, CA, USA)
- Rapid Detection Flu A+B Test (3M, MN, USA)
- SAS Influenza A Test and Influenza B Test (SA Scientific, Ltd., TX, USA)
- TRU FLU (Meridian Bioscience, Inc., OH, USA)
- XPECT Flu A&B (Remel, KS, USA)
- ZStat Flu Test (ZymeTx, Inc., OK, USA)

Some of these tests are Clinical Laboratory Improvement Amendments (CLIA)-waived, which mean they can be used as point-of-care tests performed by doctors or nurses in a clinical setting, while the rest are only licensed to be performed by trained laboratory personnel in a diagnostic laboratory setting, which may influence their accuracy.

These tests usually report results in a dichotomous yes/no manner as lines in a results window. If a line only appears at the control position, the test is negative. If a line appears at the control position and at the influenza A or B positions, the test is positive for influenza A or B respectively. Even faint lines appearing at the influenza A or B positon are considered to be positive. The test is considered invalid if no line appears at the control position even if one appears at the influenza A or B positions. It is also recommended to redo the test if the result is positive for influenza A and B.

Specimen for testing (index test)

Rapid diagnostic tests can be use with different types of upper respiratory specimens that might affect their accuracy. The type of specimens includes nasopharyngeal aspirates, swabs or washes, nasal aspirates, swabs or washes and throat swabs. Nasopharyngeal aspirates are expected to have the highest sensitivity and throat swab the lowest.

Reference standards

The following will be considered adequate reference standards for the diagnosis of influenza:

-Positive viral culture (or shell vial culture) for influenza A or B
-Positive RT-PCR for influenza A or B

Viral culture relies on the identification of a cytopathic effect in a cell culture inoculated with influenza. Results are available within 3 to 10 days though shell vial culture, which consist of enhancing viral infectivity of the cell by centrifugation, might provide results in as little as two days. Since it as long been considered the gold standard for influenza diagnosis, there is no formal evaluation of its accuracy though it has many limitations. It requires viable viruses (and therefore may be falsely negative it test is performed late in the clinical course), is effective only if the cell line used is sensitive to the inoculated virus strain, and the characteristic cytopathic effect of influenza might not be observed in infected cell line or might be caused by other respiratory viruses (1,4).

The RT-PCR, based on detection of influenza viral RNA in a clinical specimen, is the most sensitive and specific test for diagnosis of influenza (sensitivity between 86% and 100%, according to the CDC (1)). It uses primers complementary to the relatively stable gene 7, encoding the conserved matrix proteins and can detect viable and non-viable viruses. Results may be available in as little as 2-4h, though because samples are usually run in batches it might actually take longer than that for the clinician to receive the test results. Since it requires special equipment, it is not available in all settings and is more expensive than the other diagnostic tests (1,4). All PCR-based tests are prone to cross-contamination and therefore quality assurance is necessary to minimize false-positive results. Each RT-PCR assay has its limit of detection (for example 50 copies/μl). Those are clearly defined for commercially available kits, but might not be for in-house assays.

Search methods for identification of studies

Electronic searches

The following 4 electronic databases will be searched with the support of an experienced librarian:

MEDLINE, EMBASE, Web of Science, Biosis

The following search terms will be used:

#1: “Influenza, Human” [Mesh] “Influenza A virus” [Mesh] “Influenza B virus” [Mesh] OR “influenza” OR “flu” OR “grippe”[tiab]
#2: “rapid test*” OR “rapid diagnos*” OR “rapid diagnostic test*”OR “point-of-care test*” OR “antigen detection test*” OR “immunoassay*” OR “immunochromatographic test*” OR “rapid antigen test*” OR “antigen detection” OR “BinaxNOW” OR “Directigen Flu” OR “Flu OIA” OR “OSOM Influenza” OR “QuickVue Influenza” OR “Rapid Detection Flu” OR “ SAS Influenza” OR “ TRU FLU” OR “XPECT flu” OR “ZStat flu”
#1 AND #2

Searching other resources

To identify additional studies we will scan bibliographies of included studies and recent review articles and guidelines on influenza, contact experts and test manufacturers, enter the relevant identified studies into PubMed and use the Related Articles feature and search the Science Citation Index to identify articles that cite the relevant identified studies.

Language will be restricted to French and English. The time period will be from 1980 to February 2010.

Data collection and analysis

The first selection of studies, based on title and abstract, will be done by one reviewer (CC). Articles on influenza diagnosis by means other than rapid diagnostic test will be excluded at this point. Full text articles for the remaining studies will then be independently assessed by two reviewers (CC and JM). Disagreement will be resolved by consensus.

Data extraction and management

Data will be extracted on:

Author, journal and year of publication
Study design
Study population
Index test and how it was performed
Reference standard and how it was performed
Type of specimen used
Duration of symptoms before testing
Setting test is performed in; type of personnel performing test
QUADAS checklist items for study quality
Data for a two-by-two table on diagnostic accuracy (i.e. sensitivity and specificity)
Data on impact of the test on clinical outcomes, such as treatment decision, management decision, morbidity, mortality.

The data extraction form (see Appendix I) will be piloted by the two reviewers before being finalized.

Assessment of methodological quality

For quality assessment, the QUADAS checklist will be used, using "yes", "no" or "unclear" to score each item (7). Interpretation for each item is given below.

1. Was the spectrum of patients representative of the patients who will receive the test in practice?
The study population will be considered representative if patients being evaluated have clinical symptoms compatible with an influenza-like illness, defined by the CDC as the acute onset of fever (>37.8^{o}C) and cough and/or sore throat (1). Moreover, the study population should be evaluated during influenza season, which spans from November to May in the Northern hemisphere, or a clear rationale for the study time period (ex: local epidemic) should be provided. If known influenza patients are compared with healthy controls (i.e. case-control design) or if subjects with voluntary infection are used, then the spectrum of patients will be considered not representative. In case of laboratory-based studies, the spectrum of patients will be considered adequate if only slides sent for influenza testing were included.

2. Were selection criteria clearly described?
It should be clear if the patients were selected randomly, consecutively or neither. If patients with ILI were recruited in a clinical setting, a clear definition of what is considered to be an ILI should be given. For laboratory based studies, it should be clear how specimen were chosen for inclusion (every specimen during a selected time period, randomly chosen subset, etc.).

3. Is the reference standard likely to correctly classify the target condition?
The reference standard will be considered appropriate if it consists of either viral culture, shell vial culture or RT-PCR done using a well described standardized method by a trained laboratory technician.

4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
Since viral shedding in influenza patients vary greatly from day to day, peaking at 24-48h of illness then rapidly declining, the time period will be considered adequate if the specimen for the index and reference test were collected simultaneously, or at least on the same day, and stored properly pending their analysis.

5. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?
Reference standard should be performed on every patient, even if healthy controls are used, because influenza may be asymptomatic in some patients and is very prevalent in the general population during the influenza season.

6. Did patients receive the same reference standard regardless of the index test result?
The same reference standard should be applied to every patient in the study.

7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?

It is not expected that the index test (commercial rapid diagnostic tests) will form part of the reference standard (viral culture or RT-PCR).

8. Was the execution of the index test described in sufficient detail to permit replication of the test?
Since these are commercially available tests, a reference to the package insert with a statement that it was rigorously followed will be considered adequate as long as the type of specimen used is clearly stated since most of these tests have specific instructions depending on the type of specimen collected.

9. Was the execution of the reference standard described in sufficient detail to permit its replication?
The methods used for the viral culture or the RT-PCR must be clearly described or referenced and conform to accepted standards.

10.Were the index test results interpreted without knowledge of the results of the reference standard?
It should be stated that the individual performing and interpreting the index test had no knowledge of the results of the reference test or it should be explicitly said that the index test was performed before the reference standard. The only exception will be if the index test is done at the point-of-care. It will then be assumed that the results were available before the laboratory-based reference standard.

11.Were the reference standard results interpreted without knowledge of the results of the index test?
It should be stated that the technician performing and interpreting the reference standard was blinded to the index test result.

12. Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice?
The test should give a straight yes, no or indeterminate answer and their interpretation should be independent of any clinical information.

13. Were uninterpretable/indeterminate test results reported?
Results from all specimen collected should be reported, including those that were read as invalid or uninterpretable. Alternatively, redoing the test on specimens that gave invalid results initially will be considered acceptable.

14. Were withdrawals from the study explained?
A clear description of how many patients were approached enrolled and tested should be given for studies conducted in a clinical setting. For laboratory based studies, it should be stated whether or not all available specimens were included and the reasons for inclusion.

Statistical analysis and data synthesis

The data extracted in the two-by-two tables will be used to calculate sensitivity, specificity, positive and negative likelihood ratios and diagnostic odds ratio for each study (definitions are provided in Appendix III). Data from studies evaluating the same commercial test will be combined to evaluate overall performance of the test.

Studies evaluating the clinical impact of rapid diagnostic tests for influenza are likely to report very different clinical outcomes, such as the impact on the decision to prescribe antiviral treatment or antibiotic treatment, the impact on the decision to hospitalize or discharge the patient, the impact on duration of symptoms, occurrence of complications or mortality, etc. Data from different studies are thus expected to be too heterogeneous to be pooled together and a narrative summary of the extracted data will be provided.

Investigation of heterogeneity

If adequate information is available, studies will be evaluated separately according to:

1) Children or adults. Since children have higher and longer viral shedding than adults, it is expected that the sensitivity of the rapid diagnostic tests will be higher in the pediatric population.
2) Duration of symptoms before testing. Viral shedding peaks at 24 to 48h of illness and then rapidly declines. Respiratory specimen for testing should be collected within 5 days of symptoms onset. The sensitivity of the rapid diagnostic tests for influenza is expected to decrease as the length of the duration of symptoms before testing increases.
3) Type of specimen used for testing. Nasopharyngeal aspirates are expected to have the highest sensitivity and throat swab, although more convenient, is expected to have the lowest because of the amount of viral particles present in each specimen. The other types of specimens are expected to fall somewhere in between.
4) Personnel performing test. Experienced laboratory personnel are expected to obtain higher accuracy from these tests than clinicians performing them on an infrequent basis.
5) Type of assay (each commercial assay will be analyzed separately).
6) Virus type (influenza A or B and subtype of influenza A).
7) Recent (2005 or after) versus older studies. The rapid diagnostic tests for influenza might have different sensitivity and specificity in regard to newer strains of influenza like the avian H5N1 and the 2009 strain of H1N1. This could impact the accuracy of these tests in recent studies compared to older ones.
8) Quality of studies (assessed using QUADAS)

Meta-analysis

Forrest plots visually displaying sensitivity and specificity estimates and their 95% confidence intervals (CIs) from each study will be constructed using MetaDiSc software (8). Since these measures tend to be correlated and vary according to thresholds (either explicit or implicit cut-off values determining positive versus negative results), hierarchical summary receiver operating characteristic (HSROC) curves will be analyzed to explore the influence of those thresholds (9). Accuracy measures will be pooled, if appropriate, using bivariate random effects regression models (10), using the user-written program "metandi" in STATA (11). Heterogeneity of accuracy estimates will be assessed using the I^2 statistic (12).

Subgroup analysis

Since heterogeneity in accuracy estimates is expected, we will performed separate meta-analyses within pre-specified subgroups, provided sufficient studies are identified in each subgroup. The subgroups are:

- Children or adults
- Duration of symptoms before testing
- Type of specimen used for testing
- Type of commercial assay
- Virus type (influenza A or B and subtype of influenza A)
- Quality of studies

Separate HSROC curves will be generated for each subgroup, with separate pooled sensitivity and specificity estimate for each subgroup.

Source of support:

European Commission, EU-FP7 TBSusgent Grant

REFERENCES

1. Centers for Disease Control and Prevention. Information for Health Professionals. August 1st, 2009. Available at: http://www.cdc.gov/flu/professionals/index.htm. Accessed February 11, 2010.
2. World Health Organisation. Influenza. Available at: http://www.who.int/topics/influenza/en/ Accessed February 11, 2010.
3. Call SA, Vollenweider MA, Hornung CA et al. Does this patient have influenza? *JAMA* 2005; 293: 987-997.
4. Harper SA et al. Seasonal Influenza in Adults and Children—Diagnosis, Treatment, Chemoprophylaxis, and Institutional Outbreak Management: Clinical Practice Guidelines of the Infectious Diseases Society of America. *Clin Infect Dis 2009*;48:1003–32.
5. Pai M, McCulloch M, Enanoria W, Colford JM Jr. Systematic reviews of diagnostic test evaluations: What's behind the scenes? *ACP J Club* 2004; 141: A11-3.
6. Leeflang MMG, Deeks JJ, Gatsonis C et al. Systematic Review of Diagnostic Test Accuracy. *Ann Intern Med* 2008; 149: 889-897.
7. Whiting P, Rutjes AW, Reitsma JB, Bossut PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3:25
8. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.
9. Rutte CM, Gatsonis CA. A hierchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med.* 2001; 20: 2865-84.
10. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderma AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58: 982-90.
11. Harbord R, Whiting P. Metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. Stata Journal 2009; 9(2).
12. Higgins JP, Thompson SG, Deeks JJ, Altman Dg. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327 : 557-60.

Appendix 3 Search string used in PubMed

Search #1: “Influenza, Human” [Mesh] OR “Influenza A virus” [Mesh] OR “Influenza B virus” [Mesh] OR “influenza” OR “flu” OR “grippe”

Search #2: “rapid test*” OR “rapid diagnos*” OR “rapid diagnostic test*” OR “point-of-care test*” OR “antigen detection test*” OR “antigen detection” OR “rapid antigen test*” OR “immunoassay*” OR “immunochromatographic test*” OR “Binax NOW” OR “Directigen Flu” OR “Flu OIA” OR “QuickVue Influenza” OR “Rapid Detection Flu” OR “SAS Influenza” OR “ TRU FLU” OR “ XPECT flu” OR “Zstat flu”

Search #3: #1 AND #2

Restricted to Humans and English and French

Appendix 4 Data extraction form

Study ID : Reviewer :

APPENDIX I
Rapid diagnostic tests for the diagnosis of influenza
EXTRACTION FORM

Author : Journal : Year :
Country : Language : Sponsor:
Main author email:

Index test :
Manufacturer :
Performed at: □Point of care □Laboratory □NR

Specimen for testing: □Nasopharyngeal: □aspirate □wash □swab
□Nasal: □aspirate □wash □swab
□Throat: □swab
□NR
□ Frozen stored specimens

Reference test: □viral culture □shell vial culture □RT-PCR □both culture and PCR
□ If RT-PCR: name of the assay □ in house
limit of detection: copies/µL □ unknown

Quality control reported: Index test: Y/N Reference test: Y/N

Patient population:
□ adults only □ pediatric only □ both □NR
Age range:
Setting: □ outpatient clinic □ER dept □ hospitalized/institutionalized
□ research setting □other: □NR
Co-morbidities: □ included □ excluded □NR
Type: □ >65 years old □ cardiovascular disease □ lung disease
□ diabetes □ immunosuppressed □ others:
Influenza vaccination: □ included □ excluded □NR
Definition of ILI: □ fever
□ respiratory symptoms: □ cough □ sore throat □ rhinitis
□ headache □ myalgia □ malaise
□ other: □ NR
□Laboratory specimens (no description of patients)
Time since onset of sx: □ ≤2 days □ ≤3 days □ ≤4 days □ ≤5 days □ >5 days
□ NR
Influenza season: □ yes □ no □ NR date:
Predominant circulating strains:

Study Design:

Design: □ case-control □ cross-sectional □ NR
□ prospective □ retrospective □ NR
Selection: □ consecutive □ random □ other □NR
Blinding: □ index □ reference □NR
Validation: □ complete □ partial □ differential □NR
Intermediate results reported: Y/N

Strata	Index	Reference	IND	TP	FP	FN	TN	Total

Notes:

QUADAS CHECKLIST

1. Was the spectrum of patients representative of the patients who will receive the test in practice?	YES	NO	UNCLEAR
2.Were selection criteria clearly described?	YES	NO	UNCLEAR
3. Is the reference standard likely to correctly classify the target condition?	YES	NO	UNCLEAR
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	YES	NO	UNCLEAR
5. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?	YES	NO	UNCLEAR
6. Did the patients receive the same reference standard regardless of the index test result?	YES	NO	UNCLEAR
7.Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	YES	NO	UNCLEAR
8. Was the execution of the index test described in sufficient detail to permit replication of the test?	YES	NO	UNCLEAR
9. Was the execution of the reference standard described in sufficient detail to permit its replication?	YES	NO	UNCLEAR
10. Were the index test results interpreted without knowledge of the results of the reference standard?	YES	NO	UNCLEAR
11.Were the reference standard results interpreted without knowledge of the results of the index test?	YES	NO	UNCLEAR
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	YES	NO	UNCLEAR
13. Were uninterpretable / intermediate test results reported?	YES	NO	UNCLEAR
14. Were withdrawals from the study explained?	YES	NO	UNCLEAR