Clinical and epidemiological performance of WHO Ebola case definitions: a systematic review and meta-analysis

Grazia Caleo [1], Foivi Theocharaki [2], Kamalini Lokuge [3], Helen A Weiss [4], Leena Inamdar [5], Francesco Grandesso [6], Kostas Danis [7], Biagio Pedalino [8], Gary Kobinger [9], Armand Sprecher [10], Jane Greig [11], Gian Luca Di Tanna [12]

Affiliations collapse

Affiliations

[1]Manson Unit, Médecins sans Frontières, London, UK; MRC Tropical Epidemiology Group, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK. Electronic address: grazia.caleo@london.msf.org.
[2]Centre for Longitudinal Studies, University College London, London, UK.
[3]Manson Unit, Médecins sans Frontières, London, UK; National Centre for Epidemiology and Population Health, Research School of Population Health, Australian National University, Canberra, ACT, Australia.
[4]MRC Tropical Epidemiology Group, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.
[5]Public Health England, Leeds, UK.
[6]Epicentre, Paris, France.
[7]Santé Publique France, Saint-Maurice, France.
[8]TEPHINET, Task Force for Global Health, Decatur, GA, USA; Istituto Superiore Di Sanita, Rome, Italy.
[9]Département de Microbiologie-Infectiologie et d'Immunologie, Université Laval, Quebec City, QC, Canada.
[10]Médecins sans Frontières, Brussels, Belgium.
[11]Manson Unit, Médecins sans Frontières, London, UK.
[12]The George Institute for Global Health, University of New South Wales, Sydney, NSW, Australia.

Summary

Background

Ebola virus disease case definition is a crucial surveillance tool to detect suspected cases for referral and as a screening tool for clinicians to support admission and laboratory testing decisions at Ebola health facilities. We aimed to assess the performance of the WHO Ebola virus disease case definitions and other screening scores.

Methods

In this systematic review and meta-analysis, we searched PubMed, Scopus, Embase, and Web of Science for studies published in English between June 13, 1978, and Jan 14, 2020. We included studies that estimated the sensitivity and specificity of WHO Ebola virus disease case definitions, clinical and epidemiological characteristics (symptoms at admission and contact history), and predictive risk scores against the reference standard (laboratory-confirmed Ebola virus disease). Summary estimates of sensitivity and specificity were calculated using bivariate and hierarchical summary receiver operating characteristic (when four or more studies provided data) or random-effects meta-analysis (fewer than four studies provided data).

Findings

We identified 2493 publications, of which 14 studies from four countries (Sierra Leone, Guinea, Liberia, and Angola) were included in the analysis. 12021 people with suspected disease were included, of whom 4874 were confirmed as positive for Ebola virus infection. Six studies explored the performance of WHO case definitions in non-paediatric populations, and in all of these studies, suspected and probable cases were combined and could not be disaggregated for analysis. The pooled sensitivity of the WHO Ebola virus disease case definitions from these studies was 81·5% (95% CI 74·1–87·2) and pooled specificity was 35·7% (28·5–43·6). History of contact or epidemiological link was a key predictor for the WHO case definitions (seven studies) and for risk scores (six studies). The most sensitive symptom was intense fatigue (79·0% [95% CI 74·4–83·0]), assessed in seven studies, and the least sensitive symptom was pain behind the eyes (1·0% [0·0–7·0]), assessed in three studies. The performance of fever as a symptom varied depending on the cut-off used to define fever.

Interpretation

WHO Ebola virus disease case definitions perform sub-optimally to identify cases at both community level and during triage at Ebola health facilities. Inclusion of intense fatigue as a key symptom and contact history could improve the performance of case definitions, but implementation of these changes will require effective collaboration with, and trust of, affected communities.

Funding

Médecins Sans Frontières (MSF).

Introduction

Ebola virus disease case definition is a crucial surveillance tool to detect suspected cases for referral and as a screening tool for clinicians to support admission and laboratory testing decisions at Ebola health facilities. However, there have been long-standing concerns about the poor performance of the WHO Ebola virus disease case definitions, including the inability to distinguish Ebola virus disease from common diseases such as malaria and typhoid fever [1-3].

The scale of the 2014–16 west African Ebola epidemic further challenged the operational use and validity of the WHO case definitions in detecting suspected cases at the community level and allocating patients appropriately to high-risk or low-risk wards for testing at specialised isolation centres[4]. Consequently, during and since this epidemic, organisations involved in the Ebola virus disease response have estimated the sensitivity and specificity of the WHO case definitions and its constituent symptoms and signs, and developed alternative definitions and risk scores to identify clinical and epidemiological factors that could predict infection under outbreak conditions [5, 6]. Discordance on the use of WHO Ebola virus disease case definitions with consequent delay on outbreak control and community disengagement have been reported in west Africa and, in the current outbreak, in the Democratic Republic of the Congo along with its bordering countries [7-9].

However, the operational use and performance of those definitions and risk scores has not been rigorously evaluated. Such an evaluation is needed to guide communities and public health practitioners to improve the effectiveness and efficiency of identification and management of suspected cases during Ebola virus disease responses.

We aimed to assess the performance of the WHO Ebola virus disease case definitions and other clinical and epidemiological characteristics, such as symptoms and signs at admission and contact history, as the index test or test under assessment, against the reference standard of laboratory-confirmed Ebola virus infection.

Methods

Search strategy and selection criteria

For this systematic review and meta-analysis, we searched PubMed, Scopus, Embase, and Web of Science, without regional restrictions, for studies in English published between June 13, 1978 (when the first Ebola virus disease outbreaks were reported on), and Jan 14, 2020 [10,11] We also endeavoured

to capture data on the current outbreak of Ebola virus disease in the Democratic Republic of the Congo by contacting relevant people involved in the response.

The search terms included “Ebola”, “EVD infection”, “case definition”, “admission symptoms”, “sensitivity”, “specificity”, “likelihood”, “score”, “classification”, “validity” and “performance” (Appendix, pp 5-6).

We included observational retrospective studies that estimated the sensitivity and specificity of WHO Ebola virus disease case definitions and other clinical and epidemiological characteristics (symptoms and signs at admission and contact history) against the reference standard (laboratory confirmation of Ebola virus infection), and studies that developed, or externally validated, predictive risk scores (based on a combination of symptoms and signs, and epidemiological information) to predict the risk of being positive for Ebola virus.

We also included studies looking at sensitivity and specificity of WHO case definitions for Ebola or Marburg virus infections because they belong to the same family of viruses (Filoviridae) and share the same case definitions, and the reference standard is laboratory confirmation of infection [(12)].

We excluded studies on the sensitivity and specificity of diagnostic tests, animal and vaccine studies, studies of survivors of Ebola virus disease, and studies on predictors of outcomes or severity of Ebola virus disease, community surveillance, and outbreak and clinical management. Studies specifically on frequency of symptoms at admission were also excluded as a previous review exists [(13).]

Two reviewers (GC and FT) independently screened all titles and abstracts to identify those meeting the selection criteria, and a third author (LI) arbitrated for studies without consensus. A full-text review was then done for these articles, and their bibliographies were assessed for other eligible studies. We extracted data on author, year of publication, country, virus, period of data collection, study design, study objective, outcomes measured, setting in which data were collected (eg, Ebola treatment centres), age of population included in the study, study size including number of patients who were negative and positive for Ebola virus, diagnostic method, limitation of individual studies, and performance of the WHO Ebola virus disease case definitions, and individual symptoms and signs, and epidemiological links or contact history with known patients with Ebola virus disease.

Performance data extracted included sensitivity, specificity, predictive values and risk score, and area under the receiver operating characteristic (ROC) curve (AUC). We developed a spreadsheet to compile extracted data based on the Cochrane data tool [(14)]. The primary data extracted from each article were checked by a second researcher (FT). No protocol was developed for this study.

WHO Ebola virus disease case definitions were used to define suspected, probable, and confirmed cases, which varied by context and period of outbreak. In 2014 in Sierra Leone, WHO included miscarriage as an additional symptom (eg, abdominal pain) or sign (eg, vaginal bleeding) to the existing

definitions [12,15] For paediatric populations, the modified WHO case definition used in Sierra Leone was evaluated (figure 1) [15].

Data analysis

We derived the numbers of true positive, false negative, true negative, and false positive cases in each study using data provided in each article for each symptom and sign, and WHO Ebola virus disease case definition. Sensitivity and specificity are correlated, and univariate measures of heterogeneity, such as I^2, are not suitable to report heterogeneity in diagnostic test accuracy reviews [16]. We used bivariate and hierarchical summary ROC (HSROC) models for meta-analysis [17,18].

The bivariate model provides estimation of a summary of sensitivity and specificity, whereas the HSROC model provides the estimation of a summary curve from studies that have used different thresholds, the 95% confidence region for the summary point, and the 95% prediction region. The prediction region graphically illustrates between-study heterogeneity as well as the bivariate relationship between sensitivity and specificity [19]. Only studies that used comparable thresholds, symptoms and signs, or definitions were combined using these methods.

Given that HSROC models cannot be fitted when there are data from fewer than four studies, for some symptoms and signs we did a random-effects meta-analysis to calculate pooled estimates for sensitivity and specificity [20] .Compared with bivariate and hierarchical models, pooled estimation from random-effects meta-analysis could slightly overestimate point estimation, so estimates from the random-effects model are provided for completeness.

We summarised, without any further re-analysis, studies that developed or externally validated risk scores for predicting Ebola virus infection. Scores were used to identify individuals with a higher or lower risk of Ebola virus infection during screening at Ebola health facilities. To obtain the risk scores, these studies used the regression coefficients of independent risks obtained by multivariable logistic regression against Ebola virus infection and then converted regression coefficients into an integer-based point-scoring system. Reviewed studies assigned positive and negative risk scores with calculated AUC to epidemiological, demographic, and clinical characteristics. Positive values indicated higher risk of Ebola virus infection and negative values indicated higher risk of another infection such as malaria or typhoid.

Values assigned to the risk score varied by study; therefore, a meta-analysis of risk scores was not done, but instead evidence was systematically reviewed. For comparability, we reclassified the risk scores reported in the included studies into categories, from very low risk to very high risk (appendix p 7). STATA 15 was used for statistical analysis.

PRISMA guidelines for Diagnostic Test Accuracy Studies (PRISMA-DTA) were followed (appendix pp 2–4) [21].

Role of the funding source

GC, KL, AS, and JG were employed by the funder, and participated in planning the study, carrying out the research, and writing the report. The funder of the study had no further role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Of the 2493 studies initially screened using the article title, 143 were deemed to be potentially eligible on the basis of the abstract, and their full-text articles were assessed. Of these studies, 18 met the inclusion criteria, but three were excluded because data on sensitivity and specificity could not be extrapolated (appendix p 8). One was excluded because it is yet unpublished (FG). Of the 14 included studies, 11 were full manuscripts [5,6,22,24,25,27,29–33], one a letter [28] , one an oral plenary abstract [26] , and one a conference poster [23] (the author of this poster was also contacted and they provided an abstract with additional data [Kuehne A, Epicentre, Paris, France, personal communication]; table 1). 13 studies were published between 2015 and 2019 and assessed Ebola virus disease in the west Africa outbreak (seven in Sierra Leone [5,6,25,26,30,32,33], four in Guinea [24,27,28,31] , and two in Liberia [23,29]). The remaining article was published in May, 2010, assessing Marburg virus in Angola.[22] .

Overall, 12 021 people with suspected disease were included, of whom 4874 were confirmed as positive for Ebola virus infection. Study populations varied from 75 to about 2847 (table 1). All studies, apart from the national surveillance study, included patients who presented alive to health facilities for assessment. The national surveillance study included all cases (suspected, probable, and confirmed), including patients both alive and deceased, identified in both the community and health facilities. Eight studies' data were from single Ebola treatment centres [23,27–33], with the remaining using a national surveillance list (24), three from Ebola holding units [5,25,26], and two from hospitals screening patients for Ebola virus disease while still functioning as general health facilities [6,22]. All studies covered distinct patient groups from different periods and geographical areas, except for two studies from Guinea [24,27]. Although these two studies covered overlapping patient groups, they reported on different clinical and epidemiological characteristics (WHO case definition performance vs symptom performance) [24,27].

All selected manuscripts analysed all ages combined, except one author who assessed, in two different studies, the sensitivity and specificity of 2014 WHO Ebola case definitions and also developed a risk score specifically for the paediatric population (younger than 13 years) [25,26].

Six studies explored the performance of a WHO case definition in non-paediatric populations [5,22,24,29–31] . In all of these studies, suspected and probable cases were combined and could not be disaggregated for analysis. The following results therefore apply to this combined group of suspected and probable

cases. The pooled sensitivity was 81·5% (95% CI 74·1–87·2) and pooled specificity was 35·7% (28·5–43·6; figure 2). One study assessed WHO 2014 case definitions for a paediatric population (younger than 13 years old); the sensitivity was 98·0% (95% CI 95·0–99·0) and specificity was 5·0% (3·0–7·0) [25].

When WHO subdefinitions were assessed, history of contact and symptoms had high specificity compared with clinical symptoms alone, ranging from 62·3% (95% CI 49·8–73·5) to 94·4% (95% CI not provided in original paper; table 2). The highest sensitivity (100·0%) was documented for the WHO subdefinitions in which fever was not mandatory. Among studies using clinical symptoms and signs alone, the definition including three or more symptoms (intense fatigue, confusion, conjunctivitis, hiccups, diarrhoea, and vomiting) had the highest specificity (79·1% [95% CI not provided in original paper]). Unexplained death had high specificity (92·8% [95% CI not provided in original paper]) but the lowest sensitivity (14·2% [95% CI not provided in original paper]; table 2).

For children, the highest specificity (97·0% [95% CI not provided in original paper]) was with a case definition of contact, fever, and conjunctivitis, or contact, fever, anorexia, and two of abdominal pain, diarrhoea, or male sex (older than 2 years; table 2) [26].

Seven articles developed a risk score [22,23,25,29,31–33] and among those five [25,29,31–33] did an internal validation (using bootstrap or test and training methods) and one assessed a risk score according to outbreak prevalence in a paediatric population (25). An eighth study [28] externally validated the score developed by Oza and colleagues [33] without developing an alternative score. Of the 44 potential predictors of Ebola virus infection included across the seven studies that developed risk scores, 20 were found to be positive or negative predictors (figure 3). The score system ranged from very low to very high risk, with intermediate categories varying across studies (appendix p 7).

One study created a malaria sensitive score aiming to discriminate between Ebola virus infection and malaria infection, which indicated a predictor power of 89·6% (95% CI 86–93) to discriminate Ebola virus positive versus negative, reaching a discrimination power of 98·5% (95% CI not provided in original paper) during the malaria season [32]. The same study obtained similar results (AUC 76·8% [95% CI not provided in original paper] vs 75·0% [70·0–80·0]), when externally validating the scores developed by Levine and colleagues [29,32].

The study validating Oza and colleagues' algorithm found poorer performance in their cohort (AUC 58% [95% CI 56–61] vs 83·0% [79–86]) [28,33].

The highest performing score was developed by Hartley and colleagues [32], a key difference being referral time (figure 3). For the adult population (six studies [22,23,29,31–33]), a positive risk score for infection was associated in more than one study with each of the following five characteristics: epidemiological link

(eg, history of contact), diarrhoea, conjunctivitis, unexplained bleeding, difficulty swallowing (also called dysphagia; figure 3).

Fever was assessed at different thresholds (>38·0°C or ≥38·5°C), and inclusion of fever in the final predictive score was only reported by two studies [31,32] (figure 3). Discordant values were assigned across studies (either positive or negative) for anorexia or loss of appetite, muscle pain (also called myalgia), and abdominal pain.

For the paediatric population (one study [25]), positive predictors were age (2 years or older), sex (male), epid-emiological link, diarrhoea, conjunctivitis, fever (>38·0°C), anorexia or loss of appetite, and abdominal pain. Negative predictors were difficulty swallowing, rash, headache, and difficulty breathing (also called dyspnoea; figure 3). The same study compared two different time periods over the Ebola virus disease 2014–16 outbreak in Sierra Leone (high prevalence in October, 2014 [77% of suspected cases testing positive], and low prevalence in March, 2015 [4% of suspect cases testing positive]): a low cutoff for the risk score (with high sensitivity) performed better at periods of high prevalence transmission, and a high cutoff with high specificity performed better during low prevalence [25]. Similarly, the positive predictive value decreased from 93% to 31%, and the negative predictive value increased from 23% to 90% when comparing high (early) to low (late) transmission periods in the Ebola virus disease outbreak in another study in Liberia in an all ages population [23].

Eight studies measured sensitivity and specificity of individual symptoms at admission, assessing a total of 35 symptoms [5,22–24,27,29–31]. The pooled sensitivity per symptom ranged from 79·0% (95% CI 74·4–83·0) for intense fatigue (seven studies) to 1·0% (0·0–7·0) for pain behind the eyes (three studies). By contrast, the pooled specificity ranged from 98·0% (95% CI 91·0–100·0) for pain behind the eyes to 32·3% (95% CI 25·8–39·4) for intense fatigue (appendix p 9).

Haemorrhagic symptoms and signs were the most specific indicator of infection. Other symptoms and signs with high specificity included confusion, coma, hiccups, rash, and sore throat with specificity ranging from 92·0% (95% CI 91·0–94·0) for hiccups to 97·8% (95% CI 95·2–99·0) for rash (appendix p 9). Performance of fever was assessed by seven studies, but each one used a different definition of fever [5,22,23,27,29–31]. The optimal performance (definition that achieved best balance between maximising sensitivity vs maximising specificity) for fever was a threshold at ≥38·5°C (sensitivity 80·2% [95% CI 69·2–88·2]; specificity 82·6% [71·2–90·3]; table 3).[31] In the random- effects analysis, a threshold at greater than 38·0°C (three studies 22,27,29) gave a pooled sensitivity of 80·0% (95% CI 69·0–90·0) and specificity of 25·0% (17·0–33·0; table 3).

Seven studies assessed sensitivity and specificity of an epidemiological link [5,22–24,29–31]. Across these studies, the sensitivity of an epidemiological link ranged from 21·6% (95% 17·9–25·6) to 100·0% and specificity ranged from 29·0% (95% CI 19·0–41·3) to 86·0% (95% CI 74·0–94·0). The most sensitive definition was history of contact with a person with confirmed Ebola virus infection (100·0%; table 3).

The most specific definition was direct contact with an individual potentially infected with Marburg virus or his or her body fluids, or direct contact during funeral practices [22].

Discussion

Our results indicate that, for all ages combined, the WHO case definitions have a sensitivity of 81·5% and a specificity of 35·7%. The sensitivity is not high enough to achieve acceptable false negative rates, particularly in low-prevalence settings, the primary requirement for community-based screening. The low specificity results in high numbers of false positives and thus potentially unnecessary admissions to Ebola treatment centres, with associated risk of nosocomial transmission and costs of managing suspected cases [1]. As a consequence, a large number of people who do not have Ebola virus disease will experience unnecessary invasive procedures, risk of being infected with Ebola virus, isolation from family, fear of being stigmatised, and delay to appropriate care, and community mistrust in response activities will increase.

In our meta-analysis, fever had low specificity (25·0%), except for when defined as a threshold at 38·5°C or more (82·6%), and the WHO case subdefinition had 100% sensitivity only when fever was not a mandatory criterion. In the risk score systematic review, the association of fever with Ebola virus infection was not consistent across studies, with only two studies including it in the final predictive score. Presence of fever is likely to be related to the stage of infection at admission, with previous studies reporting absence of fever in a large proportion of suspected cases at admission [34]. This finding is consistent with a recent Ebola seminar reporting that fever was absent in at least 10% of the cases in the west Africa outbreak [35].

Therefore, exclusion of fever from the case definition at the community level is likely to increase the sensitivity of the case definition. Intense fatigue was the most sensitive symptom (79·0%) that could be used at the community level to facilitate early referral of suspected cases and prevent community transmission.

The meta-analysis did not identify any individual symptom or sign having an optimal trade-off between sensitivity and specificity. Conjunctivitis, unexplained bleeding, difficulty swallowing, and diarrhoea were individual symptoms and signs with the best discriminatory accuracy in the studies that explored risk score for the all-age population and with the exception of diarrhoea all had high specificity (>80%) in the studies that explored their performance. However, these symptoms and signs could also be a proxy for late-stage disease when the virus infects endothelial cells, compromising vascular integrity, with massive tissue injury resulting in disseminated intravascular coagulopathy with risk of thrombosis, bleeding, and damage to the adrenal glands and gastrointestinal system [36–38]. These symptoms and

signs could enable health practitioners to prioritise patients for admission to an Ebola treatment centre when resources are scarce but are less useful at the community level because they appear at a late stage of the disease when transmission risk is the highest.

None of the studies assessed miscarriage, despite it being included in the December, 2014, WHO case definition [15]. History of miscarriage and other associated pregnancy complications (eg, stillbirth) could help to identify cases that can be a major source of nosocomial transmission in general health facilities [39].

Although only one study focused on a paediatric population, this study used data from 11 Ebola holding units and included a large population of children (1006), providing useful guidance for this age group [26].

The WHO paediatric definition had very high sensitivity (98·0%) but very poor specificity (5·0%). When the same authors assessed a WHO subdefinition (including contact, fever, and conjunctivitis, or contact, fever, anorexia, and two of abdominal pain, diarrhoea, or male sex [older than 2 years]), the sensitivity dropped markedly to 23·0% but the specificity improved to 97·0%. The optimal fever temperature cutoff for the paediatric population was not explored. However, in another study of a paediatric population of patients with confirmed Ebola virus disease admitted to one Ebola treatment centre in Sierra Leone, 25% of children aged 5 years and younger were afebrile [40]. This difference might be due to several factors: how fever was assessed (either reported in their history or measured at admission), age groups included (younger than 13 years vs younger than 5 years), period of data collection (August–March, 2015, vs June–Dec, 2014) when seasonality of other febrile illnesses could have influenced fever prevalence, background Ebola virus transmission rates, and viraemia at admission and time since onset of symptoms.

The paediatric analysis did not explore sensitivity and specificity of individual symptoms and signs at admission for children. Alongside the fact that they might have different clinical presentations compared with adults, children are more likely to experience adverse outcomes from Ebola virus disease and are less able to report symptoms and history of contact.

Similarly, pregnant women with non-Ebola virus disease-related complications usually present with symptoms (such as bleeding and abdominal pain) that mimic Ebola virus infection [39]. As suggested elsewhere, the paediatric and pregnant women populations might require adaptation of case definitions that take into account their specific characteristics [41–43]. None of the selected manuscripts explored the performance of WHO Ebola case definitions among pregnant women. Therefore, further evidence specifically applicable to children and pregnant women is required to develop appropriate tools for screening for Ebola virus disease in these populations.

Reported history of contact was a strong predictor for paediatric and adult populations, often performing better than many of the clinical symptoms included in accepted case definitions, as also reported by other studies [44]. However, it is likely that this is an underestimate of the potential performance of actual contact history in screening for Ebola virus disease.

Levels of disclosure of self-reported clinical information and contact history depend on community engagement with intervention strategies, including trust in the health-care provider. Therefore, to improve WHO case definition performance, effective and trusted collaboration with communities is essential to ensure reliable understanding and reporting of such crucial epidemiological information. Equally, it is the responsibility of response agencies to understand the underlying pattern of Ebola virus transmission, local traditions, coping mechanisms, and family dynamics in order to identify people at risk of infection. Genetic sequencing has also been put forward as a tool for identifying chains of transmission when contact history is unknown [45].

One of the limitations in interpreting the results of this meta-analysis is that all the evidence reviewed, apart from the national surveillance study, came from patients triaged at health facilities or Ebola isolation centres. Thus, this meta-analysis might represent only cases with severe symptoms, limiting generalisability to the performance of these screening criteria at the community level and in early stages of disease. Second, there was significant heterogeneity between selected studies, and considerable variation in the quality of data on clinical symptoms and recollection of patients' history, with different variables and thresholds used in each study, and limited data on co-infection. For example, fever is a key symptom in the WHO case definitions, but different temperatures were used to define fever, which could explain the between-study heterogeneity. Inconsistency on thresholds for fever and the decision to include fever or not have been reported in the Democratic Republic of the Congo and in four neighbouring countries [9].

For the two studies with overlapping patient populations, performance of WHO case definitions was assessed only using national surveillance data, with Ebola treatment centre data for these patients being assessed for only individual symptoms or WHO subdefinitions. These two studies were therefore not included together in pooled estimations, so the cohort overlap would not have affected results. Individual studies mentioned small sample size and poor quality of data as part of their limitations.

A range of contextual factors related to study setting will affect the performance of Ebola virus disease case definitions, including seasonally occurring diseases such as malaria and Lassa fever, which have a similar clinical presentation to Ebola virus disease. Such factors will affect the generalisability of our findings to other settings. In addition, only two of the recommended risk scores were externally validated [28,32] , limiting the generalisability of those scores because performance appears to vary across outbreak periods and populations.

Finally, there is potential for publication language bias because we considered only studies in English. However, for Guinea, a French-speaking country, we included data from national surveillance and two major Ebola treatment centres; therefore, we consider that bias due to language restrictions was minimised in our results. We included peer-reviewed abstract and poster data to capture data on paediatric populations and additional evidence for all age cohorts, and we sought unpublished evidence from French- speaking countries.

This systematic review is relevant to inform public health practitioners in the current Ebola virus disease outbreak in the Democratic Republic of the Congo, in which only 8% of suspected cases isolated are confirmed, possibly because of inconsistent use of WHO case definition at community and health facility levels [46] .

In conclusion, this first systematic review and meta-analysis of the strengths and limitations of the WHO Ebola virus disease case definitions highlights the need for further studies to assess consistent thresholds for fever, to explore viraemia and symptoms and signs at admission, and to externally validate risk scores for Ebola virus infection. The sensitivity and specificity of WHO Ebola case definitions could be improved by excluding fever and instead including both intense fatigue and history of contact. However, reliable disclosure of reported symptoms and history of contact requires effective collaboration with, and the trust of, affected communities. To achieve this trust and collaboration, responding organisations must recognise the paramount role of communities in controlling transmission and ending outbreaks. We also identified important gaps related to the paediatric and pregnant population, which must be addressed through future research.

Contributors

GC, KL, and HAW conceived the idea of this study. GC and FT undertook the literature review and extracted the data with help from LI. GC wrote the first and final drafts of the manuscript. GC, KL, HAW, and GLDT contributed to the analysis and interpretation of the data. KL, HAW, FG, KD, BP, GK, AS, JG, and GLDT reviewed early and late drafts of the manuscript, and all authors have given signed or electronic approval to be authors on the manuscript.

Declaration of interests

We declare no competing interests.

Acknowledgments

This work was supported by funding from Médecins sans Frontières, Operational Centre Amsterdam (Netherlands). We thank Holly Baker (Médecins sans Frontières, London, UK) for supporting the final stages of manuscript preparation.

Research in context

Evidence before this study

There have been long-standing concerns about the poor performance of WHO case definitions for Ebola virus disease, including their inability to distinguish Ebola virus infection from common tropical diseases. We did a systematic search of the scientific literature using PubMed, Scopus, Embase, and Web of Science, without regional restrictions, for research articles published in English between June 13, 1978, and Jan 14, 2020. We used the search terms “Ebola”, “EVD infection”, “case definition”, “admission symptoms”, “sensitivity”, “specificity”, “likelihood”, “score”, “classification”, “validity” and “performance”. We also contacted relevant experts. We found that different organisations have attempted to assess the performance of WHO Ebola case definitions and developed alternative definitions and risk scores. However, there has been no systematic and rigorous evaluation of those studies. Such an evaluation is needed to guide communities and public health practitioners to improve the effectiveness and efficiency of identification and management of suspected cases during an Ebola virus disease outbreak.

Added value of this study

To our knowledge, this study is the first systematic review and meta-analysis that assesses the performance of the WHO Ebola virus disease case definitions, and other clinical and epidemiological characteristics such as symptoms and signs at admission and contact history, against the reference standard (laboratory confirmation of Ebola virus infection). Our analysis provides the most comprehensive evidence on the limitations of WHO case definitions and its constituent symptoms and signs, and predictive risk scores. We show that the WHO case definitions perform sub optimally to identify cases at both the community level and during triage at general and specialist health facilities. The performance of fever as a symptom varied depending on the cut off used to define fever. The most sensitive symptom was intense fatigue. History of contact was a key predictor for the WHO case definitions and for risk scores. This study identifies important gaps related to the paediatric and pregnant population and highlights the need to use consistent thresholds (e.g., for fever) to explore viraemia and symptoms at admission, and to externally validate risk scores for Ebola virus infection.

Implications of all the available evidence

Inclusion of intense fatigue as a key symptom could improve the sensitivity, the primary requirement for community-based screening, of WHO and alternative case definitions. Inclusion of contact history will improve specificity, resulting in a lower number of false positives and thus a lower number of unnecessary admissions to Ebola health facilities. These improvements will contribute to reduced isolation from family, fear of being stigmatised, delay to appropriate care, and community mistrust in response activities.

Figure 1. WHO Ebola virus disease case definitions for all ages and the paediatric population

	WHO case definitions (August, 2014) all ages[12]	**WHO case definition (December, 2014) all ages in Sierra Leone[15]**	**Late 2014 WHO case definition for paediatric population in Sierra Leone[15]**
Suspected	Any person, alive or dead, suffering or having suffered from sudden onset of high fever and having had contact: • a suspect, probable, or confirmed Ebola virus disease case • with a dead or sick animal (for Ebola) • a mine (for Marburg); OR any person with sudden onset of high fever and at least three of the following symptoms: • headaches • lethargy • anorexia or loss of appetite • aching muscles or joints • stomach pain • difficulty swallowing • vomiting • difficulty breathing • diarrhoea • hiccups; OR any person with inexplicable bleeding; OR any sudden, inexplicable death	Any person having had contact with a clinical case and presenting with acute fever (>38°C); OR having had contact with a clinical case (suspected, probable, or confirmed) and presenting with three or more of the symptoms below; OR presenting with acute fever and presenting with three or more of the symptoms below: • headache • nausea or vomiting • loss of appetite • diarrhoea • intense fatigue • abdominal pain • generalised or articular pain • difficulty in swallowing • difficulty in breathing • hiccups • miscarriage; OR any person with unexplained bleeding or miscarriage; OR any unexplained death	Any child with fever and either one symptom (in children younger than 5 years), two symptoms (in children aged 5–12 years), or more than three symptoms (in children older than 12 years); for children younger than 1 years old, maternal history is very important
Confirmed	Any suspected or probable cases with a positive laboratory result; laboratory-confirmed cases must test positive for the virus antigen, either by detection of virus RNA by RT-PCR, or by detection of IgM antibodies directed against Marburg or Ebola	Any person with a positive PCR test for Ebola or Marburg virus	Any person with a positive PCR test for Ebola or Marburg virus
Probable	Any suspected case evaluated by a clinician; OR any deceased suspected case (where it has not been possible to collect specimens for laboratory confirmation) having an epidemiological link with a confirmed case	A suspect case that is known to have had contact with a known case (suspected, probable, or confirmed); OR any person who is, on clinical or epidemiological grounds, very likely to have Ebola or Marburg	Not further specified

Table 1. Overview of articles included in the systematic review and meta-analysis

	Country	Virus	Period of data collection	Design	Objective	Outcomes	Setting of data collection	Age of study population	Patients positive for Ebola virus/ total patients	Method (reference standard) and timing of Ebola virus confirmatory testing	Limitations
Roddy et al (2010)[22]	Angola	Marburg	March–July, 2005	Observational retrospective study of data at admission	Evaluate the diagnostic validity of individual patient clinical and epidemiological characteristics and WHO-recommended case definitions for Marburg haemorrhagic fever, and develop a data-derived diagnostic algorithm for Marburg haemorrhagic fever that improves the WHO-recommended definitions	Sensitivity and specificity of WHO case definition, WHO case subdefinitions, symptoms at admission, and epidemiological link; and risk score	Screening at one hospital	All ages	41/102	Quantitative PCR on admission	Small sample; only saw patients at admission; data only captured Marburg haemorrhagic fever; hospital-based data collection; detailed data not available for all Marburg haemorrhagic fever cases; only presenting symptoms were recorded; highlights the necessity of collecting high-quality clinical and epidemiological data during outbreaks; over-representation of individuals with more serious symptoms that required hospital admission; no reported validation (external or internal)
Kuehne et al (2015)[23]	Liberia	Ebola	August, 2014–March, 2015	Observational retrospective study of data at admission and clinical results	Study the discriminative accuracy (sensitivity, attributable frequency, diagnostic test odds ratio, area under the receiver operating characteristic curve) of clinical signs, contact history, and combinations thereof	Sensitivity and specificity of WHO case subdefinitions, symptoms at admission, and epidemiological link; and risk score	One Ebola treatment centre	All ages	1235/1832	Quantitative PCR on admission	Reporting bias; poor data quality; conference poster and abstract data (Kuehne A, Epicentre, Paris, France, personal communication); no reported validation (external or internal)
Levine et al (2015)[29]	Liberia	Ebola	September, 2014–January, 2015	Observational retrospective study of data at admission	Develop a clinical prediction model that can help to guide care for patients with suspected Ebola virus disease,	Sensitivity and specificity of WHO case definition, symptoms at	One Ebola treatment centre	All ages	160/382	Quantitative PCR on admission	Data collected only at admission, different stages of disease process; data might not be representative of all patients

					provide specific parameters for isolation and admission to treatment centres, and maximise resource use	admission, and epidemiological link; and risk score					with Ebola virus disease; poor data quality; small sample; patients pre-screened by Ebola treatment units (ambulance travel); only assessed 14 variables; no reported external validation, only internal validation
	Country	**Virus**	**Period of data collection**	**Design**	**Objective**	**Outcomes**	**Setting of data collection**	**Age of study population**	**Patients positive for Ebola virus/total patients**	**Method (reference standard) and timing of Ebola virus confirmatory testing**	**Limitations**
Lado et al (2015)[s]	Sierra Leone	Ebola	May, 2014–December, 2014	Observational retrospective study of data at admission	Identify clinical characteristics that were predictive of Ebola virus disease diagnosis and assess the accuracy of suspected Ebola virus disease case definitions	Sensitivity and specificity of WHO case definition, WHO case subdefinition, symptoms at admission, and epidemiological link	One Ebola holding unit	All ages	464/724	Quantitative PCR on admission	Small sample; poor accuracy on reporting of symptoms and history; no access to patients who chose not to present to hospital or did not have access; no reported validation (external or internal)
Arranz et al (2016)[30]	Sierra Leone	Ebola	December, 2014–March, 2015	Observational retrospective study of data at admission	Compare the clinical characteristics of confirmed cases (patients with Ebola virus disease) and non-confirmed cases (patients without Ebola virus disease), assess the diagnostic validity of initial symptoms used in WHO case definition to diagnose Ebola virus disease in a low-incidence situation	Sensitivity and specificity of WHO case definition, WHO case subdefinition, symptoms at admission, and epidemiological link	One Ebola treatment centre	All ages	31/75	Quantitative PCR on admission	Only data at admission; poor data quality; retrospective design; small sample; no reported validation (external or internal)
Loubet et al (2016)[31]	Guinea	Ebola	December, 2014–February, 2015	Observational retrospective study of data at admission	Identify epidemiological, sociodemographic, and clinical variables associated with Ebola virus disease diagnosis and to create, based on these variables, a predictive score for identification of confirmed Ebola virus disease	Sensitivity and specificity of WHO case definition, WHO case subdefinition, symptoms at admission, and epidemiological link; and risk score	One Ebola treatment centre	All ages	76/145	Quantitative PCR on admission	Data collected only at admission; poor data quality; retrospective design; patients included might have been reluctant to come to the Ebola treatment centre, and thus were more likely to present severe clinical presentation with late symptoms; temperature taking might be affected by several factors; small sample size; anorexia and temperature (the factors that in that study

										were associated with an increased likelihood of Ebola virus disease) are not easy to measure and interpret; no reported external validation, only internal validation
	Country	**Virus**	**Period of data collection**	**Design**	**Objective**	**Outcomes**	**Setting of data collection**	**Age of study population**	**Patients positive for Ebola virus/total patients**	**Method (reference standard) and timing of Ebola virus confirmatory testing**	**Limitations**
Hartley et al (2017)[32]	Sierra Leone	Ebola	December, 2014–November, 2015	Observational retrospective study of data at admission	Construct an easy-to-use and highly accurate triage scoring system that discriminates Ebola virus infection risk in a malaria-sensitive manner and improve the predictive accuracy for Ebola virus disease and malaria	Risk score	One Ebola virus treatment centre	All ages	158/566	Quantitative PCR on admission; rapid diagnostic malaria test (histidine-rich protein-II antigen rapid diagnostic kits were used)	Only the most prevalent symptoms at admission were included in the score; poor data quality; did not fully cover all the malaria season because the Ebola treatment centre was opened from December to June; recall bias
Fitzgerald et al (2017)[26]	Sierra Leone	Ebola	August, 2014–March, 2015	Observational retrospective study of data at admission	Refine the case definition and describe outcomes of admitted children	Sensitivity and specificity of WHO case subdefinitions	11 Ebola holding units	Paediatric population (younger than 13 years)	309/1006	Quantitative PCR on admission	Only included children younger than 13 years; oral plenary abstract; no reported external validation, only internal validation
Ingelbeen et al (2017)[27]	Guinea	Ebola	March, 2014– September, 2015	Observational retrospective study of data at admission	Describe the burden of non-cases in relation to the phase of the outbreak; determine the duration of their stay at the Ebola treatment centre and (potential) subsequent nosocomial infections; compare characteristics, outcome, and risk factors for death in confirmed cases and non-cases to improve the selection, diagnosis, and care of people with suspected Ebola virus disease	Sensitivity and specificity of WHO case subdefinitions and symptoms on admission	One Ebola treatment centre	All ages	822/2362	Quantitative PCR on admission; Xpert Ebola Assay (Cepheid GeneXpert, Sunnyvale, CA USA) on admission	The Ebola treatment centre for part of the outbreak was located within one hospital but then moved to another area in July; could not assess possible drivers for the large proportion of non- cases; no reported validation (internal or external)
Oza et al (2017)[33]	Sierra Leone	Ebola	November, 2014–March, 2015	Observational retrospective study of data at admission	Develop two Ebola risk scores to supplement the broad WHO case definition by further separating triaged patients based on their likelihood of being positive for Ebola virus	Risk score	One Ebola treatment centre	All ages	252/424	Quantitative PCR on admission; biochemistry laboratory tests with the Piccolo Xpress (Abaxis, Union City, CA, USA) and i-STAT (Abbott Point of Care, Princeton, NJ, USA) device	Only one treatment centre; investigated 14 commonly recorded symptoms; small amount and poor quality of patient data; excluded exposure as a potential predictor because of large amount of missing data or poor data quality; patients might not be representative of the overall population of suspect Ebola cases; no reported external validation, only internal validation

	Country	Virus	Period of data collection	Design	Objective	Outcomes	Setting of data collection	Age of study population	Patients positive for Ebola virus/total patients	Method (reference standard) and timing of Ebola virus confirmatory testing	Limitations
Hsu et al (2018)[24]	Guinea	Ebola	March– October, 2014	Observational retrospective study of surveillance data	Assess the diagnostic performance of the WHO suspected case definition by using epidemiological surveillance and diagnostic test	Sensitivity and specificity of WHO case definition, WHO case subdefinition, symptoms at admission, and epidemiological link	National surveillance line list	All ages	1304/2847	Quantitative PCR (on admission and for deceased patients at the community level)	Unknown how representative the database was for all patients with Ebola virus disease; only 1412 patients had complete data to assess and analyse the WHO case definition; possible overestimation of performance of WHO definition because only common symptoms were recorded in the early stage of the outbreak; poor data quality; no reported validation (internal or external)
Fitzgerald et al (2018)[25]	Sierra Leone	Ebola	August, 2014– March, 2015	Observational retrospective study of data at admission	Develop a predictive score that could be used to tailor the paediatric case definition for suspected Ebola virus disease according to the clinical and epidemiological setting	Sensitivity and specificity of WHO case definition and risk score	11 Ebola holding units	Paediatric population (younger than 13 years)	309/1006	Quantitative PCR on admission	Only included children younger than 13 years; poor data quality; no data on the true Ebola status of people who did not meet the WHO case definition and were not admitted; no reported validation, only internal validation
Ingelbeen et al (2018)[28]	Guinea	Ebola	March, 2014– September, 2015	Observational retrospective study of data at admission	Validate risk score by Oza and colleagues[13]	Risk score	One Ebola treatment centre	All ages	805/2311	Quantitative PCR on admission; Xpert Ebola Assay (Cepheid GeneXpert) on admission	Did not propose another algorithm; letter; no reported external validation, only internal validation
Huizenga et al (2019)[6]	Sierra Leone	Ebola	September, 2014– November, 2015	Observational retrospective study of data at admission	Evaluate the pre-existing health-care infrastructure during the Ebola virus disease outbreak, and assess the provided health care and safeguard functionality of a health-care system for all patients not suspected to have or diagnosed with Ebola virus disease	Sensitivity and specificity of WHO case subdefinitions	Screening at one hospital	All ages	22/1556	Quantitative PCR on admission	Scant description of data; poor data quality; no reported validation (external or internal)

Figure 2. HSROC summary of sensitivity and specificity

HSROC=hierarchical summary receiver operating characteristic.

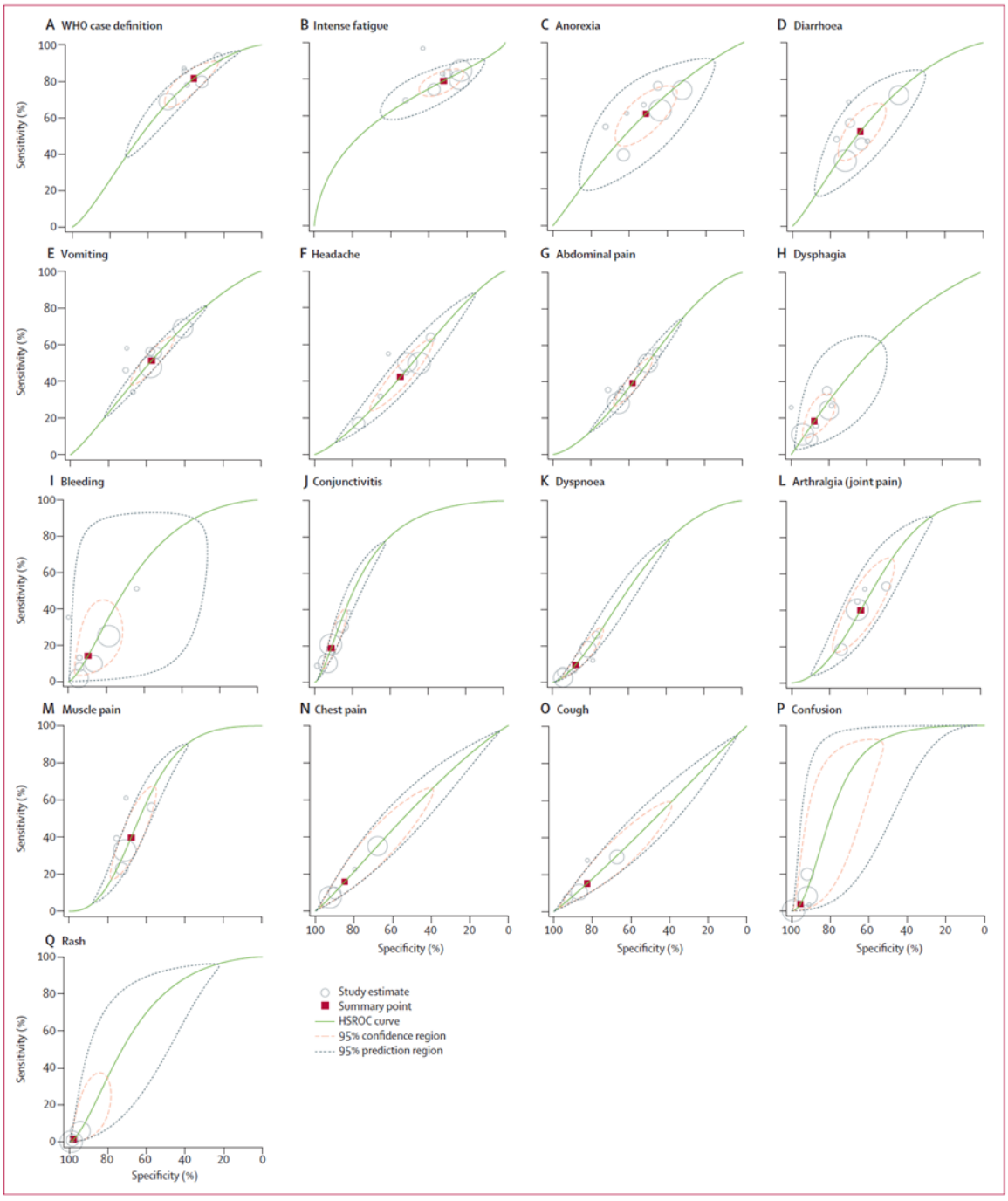


Table 2. Sensitivity and specificity of WHO Ebola virus disease subdefinitions against reference standard of laboratory-confirmed Ebola virus infection, in decreasing order of sensitivity

	WHO subdefinition	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
Huizenga et al (2019)[6]	WHO definition, with the difference that fever with sudden onset is not a mandatory criterion	100·0%	42·5%*	2·4%*	100·0%
Fitzgerald et al (2017)[26]	Contact alone, fever (in children older than 2 years) OR fever and conjunctivitis (in children younger than 2 years)	94·0%*	35·0%*	Not provided	Not provided
Roddy et al (2010)[22]	Epidemiological link or a combination of myalgia or arthralgiaand any haemorrhage	79·0% (64·0–91·0)	73·0% (60·0–84·0)	Not provided	Not provided
Loubet et al (2016)[31]	WHO subdefinition 2 (temperature ≥37·5°C plus risk factor†)	75·0% (63·5–83·9)	62·3% (49·8–73·5)	Not provided	Not provided
Roddy et al (2010)[22]	WHO case definition (clinical criteria only‡)	73·0% (57·0–86·0)	43·0% (30·0–56·0)	Not provided	Not provided
Roddy et al (2010)[22]	Fever plus three or more symptoms§	68·0% (52·0–82·0)	46·0% (33·0–59·0)	Not provided	Not provided
Loubet et al (2016)[31]	Temperature ≥38·5°C plus risk factor†	68·4% (56·6–78·3)	82·6% (71·2–90·3)	Not provided	Not provided
Arranz et al (2016)[30]	Contact and three symptoms§	67·7% (51·3–84·2)	81·8% (70·4–93·2)	72·4% (56·1–88·7)	78·3% (66·3–90·2)
Loubet et al (2016)[31]	WHO subdefinition 3 (temperature ≥37·5°C plus clinical symptoms§)	67·1% (55·2–77·2)	76·8% (64·8–85·8)	Not provided	Not provided
Loubet et al (2016)[31]	WHO subdefinition 1 (risk factor plus clinical symptoms§)	63·2% (51·3–73·7)	66·7% (54·2–77·3)	Not provided	Not provided
Lado et al (2015)[5]	Three or more major symptomsfj	57·8% (52·1–61·4)	70·8% (64·7–76·4)	77·9% (73·1–82·3)	47·5% (42·3–52·7)
Arranz et al (2016)[30]	Fever and three symptoms§	58·1% (40·7–75·4)	50·0% (35·2–64·8)	45·0% (29·6–60·4)	62·9% (46·8–78·9)
Hsu et al (2018)[24]	Clinical criteria§	57·2%*	62·0%*	66·4%*	52·5%*
Ingelbeen et al (2017)[27]	WHO case definition (clinical criteria only‖)	56·9%*	46·4%*	36·3%*	66·8%*
Roddy et al (2010)[22]	Epidemiological link and two or more general symptoms§	54·0% (37·0–70·0)	91·0% (80·0–97·0)	Not provided	Not provided
Roddy et al (2010)[22]	Epidemiological link and three or more general symptoms§	54·0% (37·0–70·0)	93·0% (83·0–98·0)	Not provided	Not provided
Arranz et al (2016)[30]	Contact plus fever	48·4% (30·8–66·0)	77·3% (64·9–89·7)	60·0% (40·8–79·2)	68·0% (55·1–80·9)
Roddy et al (2010)[22]	Fever plus haemorrhage	44·0% (28·0–60·0)	72·0% (59·0–83·0)	Not provided	Not provided
Ingelbeen et al (2017)[27]	Three major signs**	27·7%*	79·1%*	41·5%*	67·2%*
Fitzgerald et al (2017)[26]	Contact, fever, and conjunctivitis OR contact, fever, anorexia, and two of abdominal pain, diarrhoea, or male sex (older than 2 years)	23·0%*	97·0%*	Not provided	Not provided
Kuehne et al (2015)[23]	History of contact, gastrointestinal symptoms†† and illness duration of >3 days	20·0%*	94·4%*	Not provided	Not provided
Hsu et al (2018)[24]	Unexplained death	14·2%*	92·8%*	72·0%*	45·2%*

*95% CI not provided in the original paper. †For example, being a health worker, have attended a funeral, and having contact with a relative suspect of having Ebola virus.

‡Fever plus three other symptoms or fever and haemorrhage. §Symptoms or criteria not specifed in original paper. fjThree or more symptoms among the following: intense fatigue, confusion, conjunctivitis, hiccups, diarrhoea, or vomiting. ‖Acute fever and presenting three or more of the following: headache, anorexia or lack of appetite, lethargy, muscle or joint pain, breathing difficulties, vomiting, diarrhoea, stomach ache, difficulty swallowing, and hiccups; or any person with unexplained bleeding.

**As proposed by Lado and colleagues.[5] ††Diarrhoea, vomiting, and anorexia or loss of appetite.

Figure 3. Overview of risk score by symptoms and epidemiological characteristics

	AUC (95% CI) on own study database	AUC (95% CI) of Levine et al algorithm[29] on Hartley et al[32] database	AUC (95% CI) of Oza et al algorithm[33] on Ingelbeen et al[28] database	Epidemiological link	Referral (4–9 days)	Days since first symptom	Duration of illness >3 days	Gastrointestinal symptoms*	Male sex	Age (≥2 years)	Age (<2 years)	Diarrhoea	Conjunctivitis	Fever (>38·0°C)	Unexplained bleeding	Nausea or vomiting	Fever (≥38·5°C)
Hartley et al (2017)[32]	89% (86–93)	NA	NA	6	3	Y	NA	NA	NA	NA	NA	3	4	1	2	Y	NA
Oza et al (2017)[33]	83% (79–86)†	NA	58% (56–61)	NA	NA	NA	NA	NA	NA	NA	NA	2	2	Y	Y	1	NA
Loubet et al (2016)[31]	82% (77–87)	NA	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA	Y	NA	3
Fitzgerald et al (2018; paediatric population)[25]	80%‡	NA	NA	2	NA	Y	NA	NA	1	2	Y	1	2	1	Y	Y	NA
Levine et al (2015)[29]	75% (70–80)	76%‡	NA	2	NA	NA	NA	NA	NA	NA	NA	1·5	NA	Y	Y	Y	NA
Kuehne et al (2015)[23]	53–59‡	NA	NA	+	NA	NA	+	+	NA	NA	NA	NA	NA	NA	NA	NA	NA
Roddy et al (2010)[22]	‡	NA	NA	+	NA	NA	NA	NA	NA	NA	NA	Y	Y	Y	+	Y	NA

(Figure 3 continues on next page)

	Joint pain	Anorexia or loss of appetite	Muscle pain	Difficulty swallowing	Abdominal pain	Rash	Headache	Difficulty breathing	Fatigue, weakness, or asthenia	Hiccups	Cough	Diarrhoea or vomiting	Epigastralgia	Anuria	Haematuria	Disorientation	Hepatomegaly	Haemoptysis	Malaria infection	ORL haemorrhage	Dehydration	Haematochezia	Joint or muscle pain	Bleeding at injection site	Bloody gingivitis	Jaundice	Non-menstrual vaginal bleeding	Bloody diarrhoea	Haematemesis	Epistaxis
Hartley et al (2017)[32]	NA	Y	–2	2	Y	Y	Y	Y	Y	Y	NA	NA	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA
Oza et al(2017)[33]	NA	–1	NA	Y	Y	Y	–1	–1	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	Y	NA	NA	NA	NA	NA	NA	NA
Loubet et al (2016)[31]	Y	2	Y	Y	Y	NA	Y	Y	Y	NA	Y	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Fitzgerald et al (2018; paediatric population)[25]	Y	1	Y	–1	1	–2	–1	–1	Y	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Levine et al (2015)[29]	Y	1	1	1	–1	NA	Y	Y	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Kuehne et al (2015)[23]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Roddy et al (2010)[22]	+	Y	+	Y	Y	NA	Y	Y	Y	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	Y	Y	Y	Y	Y	Y	Y

Predictive scores (numeric or + symbol) are shown in shaded cells (blue indicates positive scores and light pink indicates negative scores). Y indicates that the characteristic was assessed, but not used. AUC=area under the receiver operating characteristic curve. NA=not assessed. ORL=otorhinolaryngology. *Diarrhoea, vomiting, or anorexia or loss of appetite. †95% CI is taken from Ingelbeen et al (2018)[28] because, although Oza and colleages do not report 95% CIs in their manuscript, Ingelbeen and colleagues have externally validated Oza and colleagues' score and they do report the 95% CI. ‡95% CI, AUC, or both AUC and 95% CI not given in original paper.

Table 3. Sensitivity and specificity of fever, epidemiological link, or contact history, ordered by optimal performance

	Variable	Sensitivity (95% CI)	Specificity (95% CI)
Fever cutoff			
Loubet et al (2016)[31]	≥38·5°C	80·2% (69·2–88·2)	82·6% (71·2–90·3)
Loubet et al (2016)[31]	≥38·0°C	88·2% (78·2–94·1)	72·5% (60·2–82·2)
Loubet et al (2016)[31]	≥37·5°C	93·4% (84·7–97·5)	50·7% (38·5–62·9)
Kuehne et al (2015)[23]	History of fever	85·3%*	26·4%*
Lado et al (2015)[5]	≥37·5°C or referred	85·9% (82·4–89·0)	16·4% (12·0–21·6)
Arranz et al (2016)[30]	≥38·0°C or referred	61·3% (44·1–78·4)	29·5% (16·1–43·0)
Roddy et al (2010)[22]	>38·0°C	85·0% (71·0–94·0)	20·0% (11·0–32·0)
Levine et al (2015)[29]	>38·0°C	85·0% (79·0–91·0)	21·0% (16·0–27·0)
Ingelbeen et al (2017)[27]	>38·0°C	71·5%*	30·5%*
Pooled analysis†	>38·0°C	80·0% (69·0–90·0)	25·0% (17·0–33·0)
Epidemiological link			
Hsu et al (2018)[24]	Contact with infected persons or body fluid, handling of bushmeat, attending the funeral of a patient with Ebola virus disease	74·7%*	67·1%*
Roddy et al (2010)[22]	Epidemiological link‡	67·0% (50·0–81·0)	86·0% (74·0–94·0)
Arranz et al (2016)[30]	History of contact with a person with confirmed Ebola virus disease	100·0%	59·0% (43·5–74·4)
Levine et al (2015)[29]	Sick contact§	65·0% (58·0–73·0)	61·0% (54·0–67·0)
Loubet et al (2016)[31]	Health worker or having had contact with a person with suspected Ebola virus disease or having attended funerals	81·5% (44·0–60·7)	29·0% (19·0–41·3)
Kuehne et al (2015)[23]	Contact to case	47·3%*	71·2%*
Lado et al (2015)[5]	Travel to an Ebola virus disease hotspot area, health-care work, funeral attendance, or contact with an ill family member or friend¶	21·6% (17·9–25·6)	84·6% (79·6–88·8)

Optimal performance is the definition that achieved best balance between maximising sensitivity versus maximising specificity. *95% CI not provided in original paper.

†The pooled analysis was used for the studies that had the same cut-off for fever (>38°C).[22,27,29] ‡Epidemiological link was defined as direct contact with an individual potentially infected with Marburg haemorrhagic fever or his or her body fluids or direct contact during funeral practices. §Direct or indirect contact with a patient with suspected or confirmed Ebola virus disease in the previous 21 days, including living in the same household or providing direct care for the patient. ¶A contact is any person who comes into contact with a case or suspected case by sleeping in the same household within the past month; direct physical contact with the case (dead or alive); touching his or her linens or body fluid; or attendance at a funeral of a person with confirmed or suspected Ebola virus disease.

4.5.6 Reference

1 Zachariah R, Harries AD. The WHO clinical case definition for suspected cases of Ebola virus disease arriving at Ebola holding units: reason to worry? Lancet Infect Dis 2015; 15: 989–90.

2 Pittalis S, Fusco FM, Lanini S, et al. Case definition for Ebola and Marburg haemorrhagic fevers: a complex challenge for epidemiologists and clinicians. New Microbiol 2009; 32: 359–67.

3 Kunkel A, Keita M, Diallo B, et al. Assessment of a health facility based active case finding system for Ebola virus disease in Mbandaka, Democratic Republic of the Congo, June–July 2018. BMC Infect Dis 2019; 19: 981.

4 Vogt F, Fitzpatrick G, Patten G, et al. Assessment of the MSF triage system, separating patients into different wards pending Ebola virus laboratory confirmation, Kailahun, Sierra Leone, July to September 2014. Euro Surveill 2015; 20: 30097.

5 Lado M, Walker NF, Baker P, et al. Clinical features of patients isolated for suspected Ebola virus disease at Connaught Hospital, Freetown, Sierra Leone: a retrospective cohort study. Lancet Infect Dis 2015; 15: 1024–33.

6 Huizenga E, van der Ende J, Zwinkels N, et al. A modified case definition to facilitate essential hospital care during Ebola outbreaks. Clin Infect Dis 2019; 68: 1763–68.

7 Desclaux A, Malan MS, Egrot M, Sow K, Akindès F. Surveillance in the field: over-identification of Ebola suspect cases and its contributing factors in West African at-risk contexts.
Glob Public Health 2019; 14: 709–21.

8 Biedron C, Lyman M, Stuckey MJ, et al. Evaluation of infection prevention and control readiness at frontline health care facilities in high-risk districts bordering Ebola virus disease-affected areas in the Democratic Republic of the Congo—Uganda, 2018.
MMWR Morb Mortal Wkly Rep 2019; 68: 851–54.

9 Medley AM, Mavila O, Makumbi I, et al. Case definitions used during the first 6 months of the 10th Ebola virus disease outbreak in the Democratic Republic of the Congo–four neighboring countries, August 2018-February 2019. MMWR Morb Mortal Wkly Rep 2020;
69: 14–19.

10 Report of an International Commission. Ebola haemorrhagic fever in Zaire, 1976. Bull World Health Organ 1978; 56: 271–93.

11 Report of a WHO/International Study Team. Ebola haemorrhagic fever in Sudan, 1976. Bull World Health Organ 1978; 56: 247–70.

12 WHO. Case definition recommendations for Ebola or Marburg virus diseases. August, 2014. https://www.who.int/csr/resources/ publications/ebola/case-definition/en/ (accessed May 29, 2020).

13 Petti S, Messano GA, Vingolo EM, Marsella LT, Scully C. The face of Ebola: changing frequency of haemorrhage in the West African compared with Eastern-Central African outbreaks. BMC Infect Dis 2015; 15: 564.

14 Higgins JP, Green S (eds). Cochrane handbook for systematic reviews of interventions, version 5.1. March, 2011. http://handbook-5-1. cochrane.org/ (accessed May 29, 2020).

15 WHO. Clinical management of patients in the Ebola treatment centres and other care centres in Sierra Leone: a pocket guide. Geneva: World Health Organization, 2014.

16 McGrath TA, Alabousi M, Skidmore B, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. Syst Rev 2017; 6: 194.

17 Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001; 20: 2865–84.

18 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; 58: 982–90.

19 Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. Syst Rev 2013; 2: 82.

20 Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. Arch Public Health 2014; 72: 39.

21 Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009; 339: b2700.

22 Roddy P, Thomas SL, Jeffs B, et al. Factors associated with Marburg hemorrhagic fever: analysis of patient data from Uige, Angola.
J Infect Dis 2010; 201: 1909–18.
23 Kuehne A, Gergonne B, Bawo L, et al. Differentiating high and low suspect Ebola cases based on clinical presentation and history of contact. 2015. https://issuu.com/aminataepicentre/docs/2015_ epicentre_scientific_day_poste (accessed May 29, 2020).
24 Hsu CH, Champaloux SW, Keïta S, et al. Sensitivity and specificity of suspected case definition used during West Africa Ebola epidemic. Emerg Infect Dis 2018; 24: 9–14.
25 Fitzgerald F, Wing K, Naveed A, et al. Development of a pediatric Ebola predictive score, Sierra Leone. Emerg Infect Dis 2018;
24: 311–19.
26 Fitzgerald F, Wing K, Naveed A, et al. Refining the paediatric Ebola case definition: a study of children in Sierra Leone with suspected Ebola virus disease. Lancet 2017; 389: S19.
27 Ingelbeen B, Bah EI, Decroo T, et al. Mortality among PCR negative admitted Ebola suspects during the 2014/15 outbreak in Conakry, Guinea: a retrospective cohort study. PLoS One 2017; 12: e0180070.
28 Ingelbeen B, De Weggheleire A, Van Herp M, van Griensven J. Symptom-based Ebola risk score for Ebola virus disease, Conakry, Guinea. Emerg Infect Dis 2018; 24: 1162.
29 Levine AC, Shetty PP, Burbach R, et al. Derivation and internal validation of the Ebola prediction score for risk stratification of patients with suspected Ebola virus disease. Ann Emerg Med 2015; 66: 285–293.e1.
30 Arranz J, Lundeby KM, Hassan S, et al. Clinical features of suspected Ebola cases referred to the Moyamba ETC, Sierra Leone: challenges in the later stages of the 2014 outbreak. BMC Infect Dis 2016; 16: 308.
31 Loubet P, Palich R, Kojan R, et al. Development of a prediction model for Ebola virus disease: a retrospective study in Nzérékoré Ebola Treatment Center, Guinea. Am J Trop Med Hyg 2016;
95: 1362–67.
32 Hartley M-A, Young A, Tran A-M, et al. Predicting Ebola infection: a malaria-sensitive triage score for Ebola virus disease. PLoS Negl Trop Dis 2017; 11: e0005356.
33 Oza S, Sesay AA, Russell NJ, et al. Symptom- and laboratory-based Ebola risk scores to differentiate likely Ebola infections. Emerg Infect Dis 2017; 23: 1792–99.
34 Haaskjold YL, Bolkan HA, Krogh KØ, et al. Clinical features of and risk factors for fatal Ebola virus disease, Moyamba District, Sierra Leone, December 2014–February 2015. Emerg Infect Dis 2016; 22: 1537–44.
35 Malvy D, McElroy AK, de Clerck H, Günther S, van Griensven J. Ebola virus disease. Lancet 2019; 393: 936–48.
36 Ansari AA. Clinical features and pathobiology of Ebolavirus infection. J Autoimmun 2014; 55: 1–9.
37 Sullivan N, Yang Z-Y, Nabel GJ. Ebola virus pathogenesis: implications for vaccines and therapies. J Virol 2003; 77: 9733–37.
38 Wong G, Kobinger GP, Qiu X. Characterization of host immune responses in Ebola virus infections. Expert Rev Clin Immunol 2014; 10: 781–90.
39 Bower H, Grass JE, Veltus E, et al. Delivery of an Ebola virus- positive stillborn infant in a rural community health center, Sierra Leone, 2015. Am J Trop Med Hyg 2016; 94: 417–19.
40 Shah T, Greig J, van der Plas LM, et al. Inpatient signs and symptoms and factors associated with death in children aged
5 years and younger admitted to two Ebola management centres in Sierra Leone, 2014: a retrospective cohort study. Lancet Glob Health 2016; 4: e495–501.
41 Kangbai JB, Heumann C, Hoelscher M, Sahr F, Froeschl G. Epidemiological characteristics, clinical manifestations, and treatment outcome of 139 paediatric Ebola patients treated at a Sierra Leone Ebola treatment center. BMC Infect Dis 2019; 19: 81.
42 Mpofu JJ, Soud F, Lyman M, et al. Clinical presentation of pregnant women in isolation units for Ebola virus disease in Sierra Leone, 2014. Int J Gynaecol Obstet 2019; 145: 76–82.

43 Garde DL, Kahn RJ, Mesman AW, Koroma AP, Marsh RH. Care of Pregnant Women: Experience from a Maternity-Specific Ebola Isolation Unit in Sierra Leone. J Midwifery Womens Health 2019; 64: 493–99.
44 Miglietta A, Solimini A, Djeunang Dongho GB, et al. The Ebola virus disease outbreak in Tonkolili district, Sierra Leone:
a retrospective analysis of the Viral Haemorrhagic Fever surveillance system, July 2014–June 2015. Epidemiol Infect 2019; 147: e103.
45 Pini A, Zomahoun D, Duraffour S, et al. Field investigation with real-time virus genetic characterisation support of a cluster of Ebola virus disease cases in Dubréka, Guinea, April to June 2015. Euro Surveill 2018; 23: 1–7.
46 Epicentre. EpiSitrep maladie à virus Ebola en Nord Kivu et Ituri, May to June 2019. Epicentre; Paris, France; 2019.

Supplementary material

THE LANCET
Infectious Diseases

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Caleo G, Theocharaki F, Lokuge K, et al. Clinical and epidemiological performance of WHO Ebola case definitions: a systematic review and meta-analysis. *Lancet Infect Dis* 2020; published online June 25. https://doi.org/10.1016/S1473-3099(20)30193-6.

Contents

A. Methods Appendix

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	2-3
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	3, Table 1
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	Table 1
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	2-3
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	3

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	3;7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	3;7
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	7;9-10, Table 1
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	Table 1

Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	10-12, Figure 2 Figure 3 Table 2-3
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	10-12, Figure 2
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	Appendix page 7 and 9
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	12-14
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	13-14, Table 1
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	12-14
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	7

Adapted From: McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.

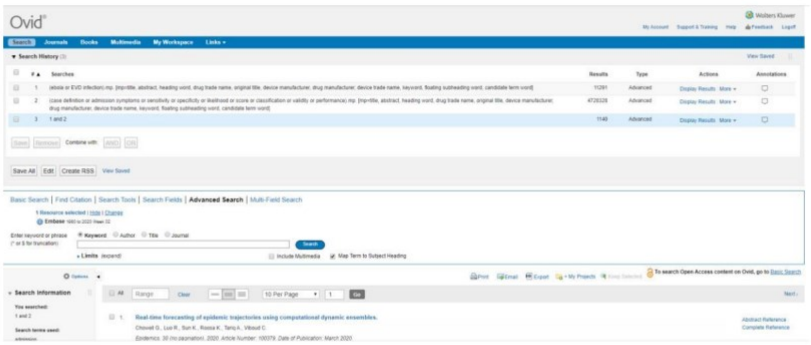
For more information, visit: www.prisma-statement.org.

A2. Systematic review search terms

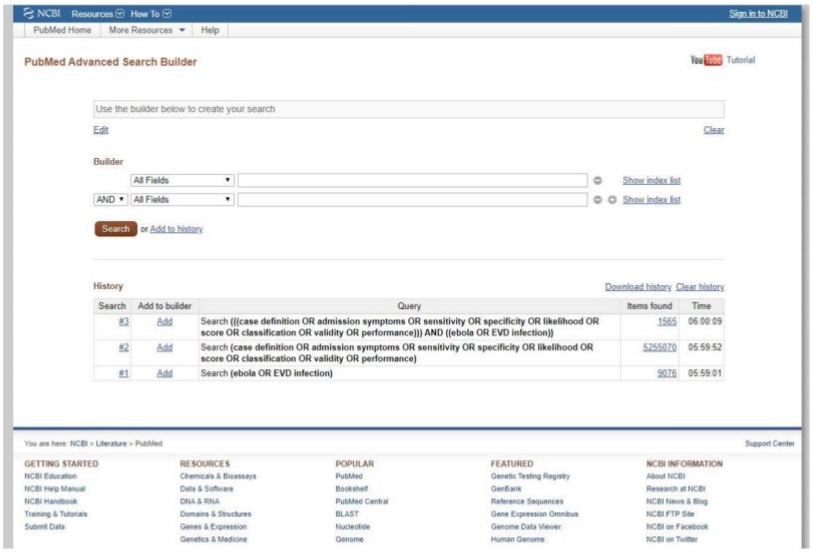
To identify eligible studies, we searched PubMed, Scopus, EMBASE, Web of Science, without regional restrictions. We included articles in English published from 13 June 1978 (when the first EVD outbreaks were reported on) to 14 January 2020.

The search terms used in our systematic review in order to identify the studies needed are: (ebola OR EVD infection) AND (case definition OR admission symptoms OR sensitivity OR specificity OR likelihood OR score OR classification OR validity OR performance)

Embase Search Criteria



PubMed Search Criteria



A3. Risk score

We summarised, without any further re-analysis, studies that developed or externally validated risk scores for predicting EV infection. Scores were used to identify individuals with a higher or lower risk of EV infection during screening at ETCs.

For comparability, we re-classified the risk scores reported in the included studies into categories, from very low risk to very high risk (Table S1)

Supplementary Table 1. Classification of the risk score for Ebola virus disease (EVD) infection across selected studies

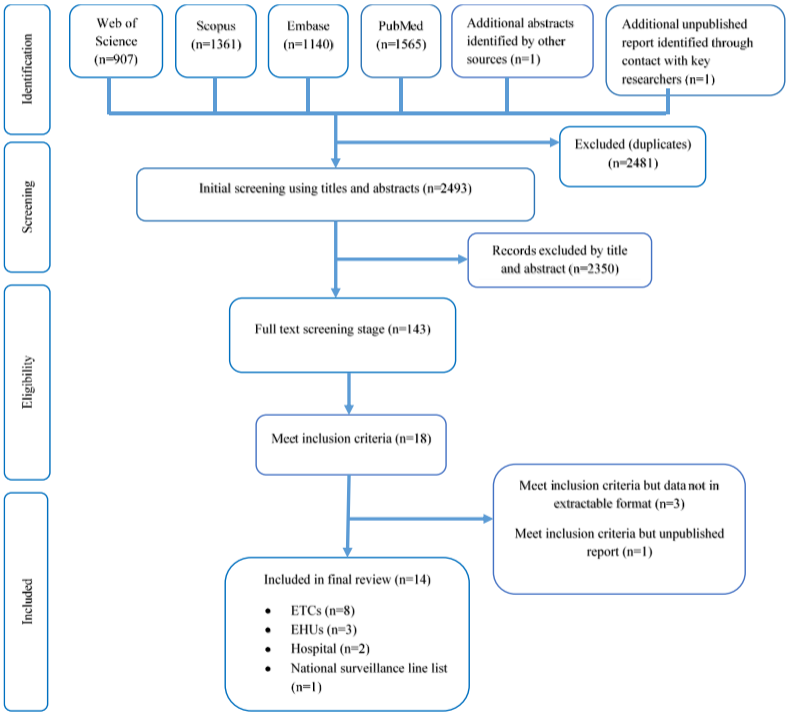
Classification	Levine et al., 2015[29]	Loubet et al., 2016[31]	Hartley et al., 2017[32]	Oza et al., 2017 and Ingelbeen et al., 2018 [33, 28]	Fitzgerald et al., 2018*[25]
Very low risk	≤0	-*	≤1	-*	
Low	1	0-2	2-4	- (lowest score: -3)	0-3
Moderate	2	2-4	5-7	0	-*
High risk	3	≥4 (4-6)	8-11	+ (highest score: +5)	
Very high risk	≥4	-*	≥12	-*	7-10

*In high prevalence, score ≥7 indicates high risk for a child being a case (0-13 years old).

B. Results Appendix

B1. Systematic review flowchart

Supplementary Figure 2: PRISMA flow diagram depicting the number of articles at each of the identification, screening and eligibility stages



B2. Pooled sensitivity and specificity of individual symptoms

Supplementary Table 2. Performance of symptoms at admission ordered by number of studies

Symptom	Number of studies	Model	Sensitivity (95% CI)	Specificity (95% CI)
intense fatigue	7	Bivariate and HSROC	79·0 (74·4-83·0)	32·3 (25·8-39·4)
anorexia	7		61·0 (50·8-70·2)	51·3 (41·7-60·8)
diarrhoea	7		51·4 (42·2-60·6)	64·4 (56·7-71·3)
vomiting	7		51·2 (43·4-58·9)	57·4 (50·6-63·8)
headache	7		42·2 (30·9-54·2)	55·0 (45·0-64·5)
abdominal pain	7		39·3 (31·7-47·5)	58·0 (51·9-63·8)
dysphagia/difficulties to swallow	7		18·2 (12·3-26·2)	87·6 (82·0-91·7)
bleeding	7		14·3 (6·2-29·4)	89·5 (80·7-94·6)
conjunctivitis	6		19·0 (11·8-29·2)	91·1 (87·4-93·9)
hiccups	6	Random effects*	10·0 (9·0-12·0), I2=0· 00%	92·0 (91·0-94·0), I2=9·96%
dyspnoea	6	Bivariate and HSROC	9·8 (4·9-18·3)	87·8 (81·2-92·3)
arthralgia/joint pain	5		39·9 (28·0-53·1)	63·5 (55·6-70·7)
muscle pain	5		39·8 (28·3-52·5)	67·8 (62·1-72·9)
chest pain	4		15·8 (8·1–28·4)	84·7 (73·6-91·7)
cough	4		15·2 (8·4–25·8)	82·6 (71·5-90·0)
confusion	4		3·7 (0·6-19·6)	95·4 (89·1-98·1)
rash	4		1·5 (0·4-4·6)	97·8 (95·2-99·0)
'muscle or joint' pain	3	Random effects*	58·0 (47·0-69·0), I2=62·7%	53·0 (42·0-65·0), I2=73·1%
sore throat	3		8·0 (0·0-26·0), I2=99·3% 3·0	93·0 (81·0-99·0), I2=98·4%
jaundice	3		(0·0–9·0), I2=93·6%	97·0 (93·0-99·0), I2=82·7%
pain behind the eyes	3		1·0 (0·0–7·0), I2=98·1%	98·0 (91·0-100·0), I2=98·1%
haematemesis	2		20·0 (17·0–24·0), I2= 0·0%	97·0 (95·0-99·0), I2=0·0%
melena	2		4·0 (3·0-6·0), I2=0·0% 2·0	96·0 (94·0-98·0), I2=0·0%
coma	2		(2·0-3·0), I2=0·0% 83·9	95·0 (94·0-96·0), I2=0·0%
digestive symptoms	1		(70·9-96·8)	45·5 (30.7-60·2)
non-menstrual vaginal bleed	1		20·0 (8·0-39·0)	91·0 (75·0-98·0)
bloody gingivitis	1		17·0 (7·0-32·0)	93·0 (84·0-98·0)
bleed from injection site	1		12·0 (4·0-26·0)	97·0 (89·0-100·0)
back pain	1		6·5 (0·0-15·1)	93·2 (85·7-100·0)
epistaxis	1		5·0 (1·0-17·0)	98·0 (91·0-100·0)
epigastralgia	1		4·0 (1·0-12·0)	98·5 (91·0-99·9)
Neurological symptoms	1		3·2 (0·0-9·4)	90·9 (82·4-99·4)
loss of consciousness	1		3·1 (1·7-5·1)	97·6 (94·8-99·1)
haemoptysis	1		0·0 (0·0-9·0)	98·0 (91·0-100·0)

* I2 are reported where a random-effects models were performed (no. of studies <4). For hiccups the HSROC did not achieve convergence