
A.M. Rojek¹,*, A. Salam¹,², R.J. Ragotte¹, E. Liddiard¹, A. Elhussain¹, A. Carlqvist¹, M. Butler¹, N. Kayem¹, L. Castle¹, L."O. Odondi¹, K. Stepniewska¹,³, P.W. Horby¹
¹Epidemic Diseases Research Group, University of Oxford, Oxford, UK
²United Kingdom Public Health Rapid Support Team, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, UK
³WorldWide Antimalarial Resistance Network, Oxford, UK

Abstract

Background—Over 28 000 individuals were infected with Ebola virus during the West Africa (2013–2016) epidemic, yet there has been criticism of the lack of robust clinical descriptions of Ebola virus disease (EVD) illness from that outbreak.

Objectives—To perform a meta-analysis of published data from the epidemic to describe the clinical presentation, evolution of disease, and predictors of mortality in individuals with EVD. To assess the quality and utility of published data for clinical and public health decision-making.


Eligibility—Studies that sequentially enrolled individuals hospitalized for EVD and that reported acute clinical outcomes.

Methods—We performed meta-analyses using random-effect models and assessed heterogeneity using the $I^2$ method. We assessed data representativeness by comparing meta-analysis estimates with WHO aggregate data. We examined data utility by examining the availability and compatibility of data sets.

Results—In all, 3653 articles were screened and 34 articles were included, representing 16 independent cohorts of patients (18 overlapping cohorts) and at least 6168 individuals. The pooled estimate for case fatality rate was 51% (95% CI 46%–56%). However, pooling of estimates for clinical presentation, progression, and predictors of mortality in individuals with EVD were hampered by significant heterogeneity, and inadequate data on clinical progression. Our
assessment of data quality found that heterogeneity was largely unexplained, and data availability and compatibility were poor.

Conclusions—We have quantified a missed opportunity to generate reliable estimates of the clinical manifestations of EVD during the West Africa epidemic. Clinical data standards and data capture platforms are urgently needed.

Keywords
Ebola; Ebola virus disease; emerging infection; epidemic; outbreak; viral hemorrhagic

Introduction
The Ebola virus disease (EVD) epidemic in West Africa (2013–2016) was unprecedented, with at least 28 000 individuals infected [1]. Over the course of the epidemic there were gains in clinical understanding and clinical care despite the challenging circumstances—the estimated case fatality rate in the most affected countries fell from 70% to 39% [2] and was 18.5% in individuals evacuated to high-income countries [3]. Case reports detailed previously unseen or unrecognized clinical manifestations ranging from disease relapse [4], through to survival following in utero infection [5], and sustained virus sequestration in immune privileged sites [4,6,7].

It is critical to translate the personal expertise gained by clinicians during the outbreak into an evidence base to inform the treatment of individuals in future outbreaks, and to provide reliable data on clinical presentation and natural history with which to design clinical trials. Before the West Africa outbreak there were variations in clinical data estimates (including case fatality rates) due to case ascertainment biases, inconsistencies in data collection methods, and the temporal and spatial distribution of outbreaks. Although it was hoped that clear and actionable data would become available due to the scale and duration of the West Africa outbreak, when the WHO released their most recent guidelines for clinical care they concluded that their recommendations were still limited by the scarcity of robust clinical descriptions [8]. This is despite over 2800 articles being published during the outbreak [9]. Why this considerable academic output did not translate into robust data to inform clinical care has not been evaluated.

The first objective of this work is to determine the extent to which the literature can be synthesized by meta-analysis to describe the evolution of patient signs and symptoms during the course of EVD, and to identify clinical predictors of mortality with greater confidence. The second aim is to determine the extent to which the literature constitutes ‘actionable data’ [10]. This is the extent to which the findings are accessible, objective and sufficient to inform clinical, public health and research decision-making for future outbreaks [10]. This work is the first systematic analysis of clinical data accumulated during the outbreak.

Methods
This systematic review was conducted in accordance with PRISMA guidelines [11] and prospectively registered on the PROS-PERO database (registration CRD42017070150).
Data sources and searches
We conducted a systematic search of PubMed for articles published between 1 January 2014 and 31 May 2017, updating a search published previously [9]. The search term was ‘Ebola’. We searched for epidemiological data (case number and mortality data) on the WHO website.

Study selection and criteria
We included articles describing hospitalized individuals with laboratory-confirmed EVD who were treated in the three most affected countries (Liberia, Sierra Leone, Guinea) during the epidemic in West Africa (2013–2016).

We included articles that described one or more clinical aspects of disease (defined as a symptom, examination finding, or laboratory investigation finding) and an acute patient outcome (at a minimum, death up to or within 28 days of discharge). Hospitalization was defined as admission to an Ebola Treatment Unit/Centre (ETC) or a pre-existing hospital. A post hoc addition was made to include several sites that were originally termed Ebola holding centres but admitted patients and provided supportive medical therapy (e.g. intravenous fluids) and that were retrospectively retermed as ETCs by national governments. To identify representative clinical data sets, articles must have included all EVD patients in their facility for the defined period and so articles that reported a subset of patients (usually on the basis of age or pregnancy status) were excluded, as were articles that included additional enrolment criteria (usually clinical trials).

We excluded articles that described fewer than ten patients. Language was restricted to articles with English and French abstracts.

Data extraction and quality assessment
We reviewed manuscript titles, abstracts and full text in duplicate based on pre-agreed selection criteria. For eligible articles, data were extracted in duplicate into an electronic data form, with two independent reviewers following the study data management standard operating procedure. The senior author made final determination of article selection or data extraction where there was disagreement. Quality assessment was performed only by the senior author.

We extracted study characteristics including the number of patients included, selection criteria, dates of patient enrolment and dates of publication. Where hospitalized and non-hospitalized patients were included, we extracted only data on hospitalized patients.

We extracted patient information including demographics, temporal characteristics of disease, clinical manifestations of disease (based on the WHO case definition used during the outbreak [12]), patient co-morbidities, vital signs, predictors of mortality, mortality and patient treatment. Data extraction was recorded by survival status of the patient where that information was available. We recorded symptom data at three time-points (on admission (defined as ≤24 h), during hospitalization (subsequent days) or at any time (for when an article did not describe the time of data collection, or this could not be deducted). Vital sign
data were abstracted on admission only because there was little longitudinal data on pilot testing.

For all variables, we extracted numerator and denominator data for dichotomous variables, and summary statistics for continuous variables (mean or median, and standard deviation, or interquartile range, or range, or proportion high or low). For predictors of survival we extracted unadjusted and adjusted odds or risk ratios with 95% confidence intervals. Data were reported as missing for each variable when it was not available in the manuscript. Manuscript authors were not contacted for missing data.

Quality assessment

An existing observational data quality assessment scale [13] was piloted but found to be unsuitable so we devised a study-specific scale where assessment criteria were based on parameters independently identified as critical for outbreak evaluation [14]. These include that the data source is defined; the timeframe of data collection is defined; and there is transparency regarding the proportion of cases for which an outcome is known. Clinical interpretability of data was assessed based on whether each article provided an estimate of clinical descriptors that were known or suspected co-variates of outcome at the time the West Africa outbreak commenced (age, pregnancy status, viral load and time from symptom onset to patient admission) [15–17].

Treatment of duplicate data

Duplicate or overlapping data sets were first identified on the basis of replication of authors, facilities, enrolment periods, sample size or statement of previous publication. Because some facility names and affiliated organizations changed during the outbreak, articles were then manually mapped by the senior author who was familiar with ETC operations during the outbreak.

When duplicate or overlapping data sets occurred, they were ranked from that which included the most patients, to that with the fewest number of patients. The first ranked (largest data set) was always used for reporting patient variables when data were available. However, because more detailed clinical description could be included in one or more lower ranked data sets, these were not excluded outright. Instead, a variable of interest was searched in a step-wise manner down the ranking scale and then reported for the highest ranked article where it was contained. We refer to the first ranked data sets collectively as the ‘primary’ data set.

Statistical analysis

Descriptive statistics are presented as frequencies for categorical variables, means and standard deviations for normally distributed data, and median with range for other continuous variables. Relative risks were calculated from the number of events and participants in each group. All p-values are calculated from two-tailed tests of statistical significance with a type I error rate of 5%.
Meta-analysis

Individual proportion or relative risk estimates are displayed graphically with forest plots, and on summary forest plots when variables are clinically related. Meta-analysis was performed using a random effects model. A binomial-specific method was used for meta-analysis of proportions [18]. Heterogeneity for all meta-analysis was assessed using the $\chi^2$ test for assessment of heterogeneity and quantified with the $I^2$ statistic. Estimates of publication bias in meta-analysis (such as Egger’s test) were not included because of their limited utility when there are a small number of publications [19]. Meta-analysis was only performed when three or more articles reported data for the variable.

Actionable data analysis

Three criteria were used to perform an actionable data analysis: the representativeness of the data to the underlying population, the ‘transparency’ of reporting [10], and the ability for the data to be quantitatively synthesized (compatibility) [10].

We assessed data representativeness by comparison with WHO data sets. For symptom data, this was aggregate data from the first year of the outbreak (data stratified by hospitalization status was not available) [20]. There are no aggregated estimates available for vital sign data. The comparative estimate for case fatality rate (CFR) was based on individual patient data (IPD) on hospitalized individuals during the entire outbreak [2].

We assessed data transparency by reporting whether anonymized IPD were available for an article, and whether publication of the same cohort in another manuscript was acknowledged when applicable.

We assessed data compatibility for synthesis by comparing the consistency of data reporting for two key co-variates of outcome that could be used in meta-regression (patient age and estimate of viral load) and by mapping predictors of patient outcomes.

We performed all analyses with STATA/MP version 15.0.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report.

Results

A total of 3653 articles were screened (Fig. 1), and 34 articles were eligible for inclusion (see Supplementary material, Appendix S1) [21–54]. Out of these, 16 articles represent the largest, or only, representation of patients (termed the primary data set) and 18 articles duplicate or overlap with these data and are only used when data are missing from the primary data set. The total number of patients included in the primary data set is 6168. For this data set the median number of patients per article was 207 (interquartile range 108–607).
Quality of evidence

Appendix S2 (see Supplementary material) details the quality of included evidence. Articles reported the data source and time of collection in 47% (16/34) of articles, the exact dates of the study in 79% (27/34) of articles, and provided specific details of laboratory confirmation in 59% (20/34) of articles. Most articles (53%, 18/34) described if patients lost to follow up were excluded from analysis, and detailed the proportion of cases for whom the clinical outcome was known (97%, 33/34). Fewer articles (35%, 12/34) detailed the extent of missing data for other variables, or described statistical management of missing data. Reporting of covariates for estimating mortality varied by category—88% (30/34) reported age, 68% (23/34) time from symptom onset to hospitalization, 59% (20/34) viral load (or cycle threshold), and 9% (3/34) pregnancy status.

Case fatality rate

The pooled CFR of the primary data set was 51% (95% CI 46%–56%); however, heterogeneity was significant ($I^2 = 92.8\%$, $p < 0.01$). The forest plot for CFR is provided in the Supplementary material (Appendix S3).

Symptoms of EVD infection

In pooled analysis of the primary data set, the most common presenting symptoms were fever (76%, 95% CI 66%–85%), fatigue (71%, 95% CI 64%–74%) and anorexia (64%, 95% CI 51%–76%) (Fig. 2). There was high (defined as $I^2$ value > 75% [55]) heterogeneity for symptoms of haemorrhagic manifestations, hiccups, difficulty breathing, conjunctivitis, abdominal pain, myalgia, vomiting, headache, anorexia and fever. There was low to moderate heterogeneity for all other presented symptoms based on $I^2$ estimate, although the $\chi^2$ for heterogeneity was still significant for diarrhoea ($p < 0.01$) and lethargy ($p < 0.01$) (see Supplementary material, Appendix S4).

The pooled relative risk of death was significantly >1 in patients presenting with the symptoms of haemorrhage, hiccups, diarrhoea, breathing difficulty, conjunctivitis and difficulty swallowing compared with those presenting without the symptom (Fig. 3). Heterogeneity was low to moderate for all estimates except for confusion and vomiting, although the $\chi^2$ test for heterogeneity was significant for headache ($p = 0.046$) (see Supplementary material, Appendix S5).

Systematic review was not feasible for symptoms during hospitalization. Only one article systematically reported patient symptoms in the period following admission [31], although sporadic reporting of some other symptoms occurred in a further two articles.

Meta-analysis of vital sign data was not feasible, because only two articles systematically reported data [38,52], with some detail in one further article [42].

Supportive care for EVD infection

Only one article differentiated between medications that were prescribed (as part of a protocol) and those that were actually administered to patients. Five articles detailed the exact number of patients receiving antibiotics and antimalarials (and one article detailed
route of admission) and six articles described intravenous fluid use (with an estimate of fluid volume provided in half of these manuscripts).

**Actionable data analysis**

**Data representativeness**— Fig. 4 depicts the CFR reported in the published articles compared with reference data from WHO (CFR from all hospitalized patients during the outbreak) [2]. Our estimate of CFR was 51%, the WHO estimate was 54.5%.

Fig. 2 shows the meta-analysis estimates for the proportion of patients presenting with a symptom compared with the reference data (WHO data for all confirmed and probable patients presenting during the first year of the outbreak).

Fig. 3 depicts a comparison of our meta-analysis estimates for the relative risk of death for a patient presenting with a symptom at the time of admission, in comparison with reference data (as outlined above).

**Data transparency**

Eighty-five per cent (29/34) of articles made no statement on IPD availability, and data were only available for two articles where a statement was made. Eighty-two per cent (28/34) of articles reported duplicate or overlapping data (although this may be an underestimation, as we did not systematically review articles that did not meet our inclusion criteria). In only one of these articles was there acknowledgement of duplicate presentation (see Supplementary material, Appendix S6).

**Data compatibility**

Of the 15 manuscripts presenting data according to age categories, the categories selected for stratification were the same for only two articles (see Supplementary material, Appendix S7). Between the 15 manuscripts that stratified by viral load for analysis, ten different categorization systems were used (see Supplementary material, Appendix S7). The diversity of variables found to be predictive of case fatality are shown in Fig. 5.

**Discussion**

In this analysis of the clinical presentation of EVD, we report relatively comprehensive data for symptoms on admission (and their association with CFR). We planned to describe the evolution of patient symptoms and predictors for death because this information is not available in epidemiology data sets and is of significant utility for clinical decision-making. However, we found the literature inadequate to perform such analyses. In some ETCs, the lack of data reflects the clinical reality of working in a humanitarian setting, where at times, careful clinical observation was not feasible. However, in many other ETCs, especially those reporting data later during the outbreak, some data are expected to exist.

Meta-analysis indicated substantial heterogeneity for analysis of the proportion of patients presenting with a symptom, and for analysis of CFR. We attempted to limit clinical diversity by only including articles that enrolled all patients in a sequential manner presenting for treatment. However, populations may have differed between studies due to ETCs using...
different admission routes. Moreover, selection bias may have occurred because early during the outbreak some overwhelmed ETCs turned away patients and in other cases fears or rumours regarding ETCs may have altered admission patterns. The representativeness of estimates is further diminished by limited reporting of covariates for mortality (including time to hospitalization, cycle threshold and pregnancy status).

When we compared with WHO estimates, the limitations of clinical decision-making based on the literature became evident. For example, our estimates of the proportion of patients presenting with vomiting (50%, 95% CI 41%–58%) and diarrhoea (51%, 95% CI 45%–57%) are substantially lower than WHO estimates (66%, 95% CI 64%–67%; and 64%, 95% CI 62%–65%, respectively) with nonoverlapping confidence intervals, and our point estimates of the relative risk of death when presenting with each of these symptoms are higher than WHO estimates—diarrhoea 1.4 (95% CI 1.2–1.6) compared with 1.1 (95% CI 1.0–1.2); vomiting 1.7 (95% CI 0.9–3.0) compared with 1.1 (95% CI 1.0–1.1). Whether gastrointestinal manifestations are a common manifestation of EVD among all patients, or is relatively rarer but more discriminative of patients at risk of death, is unclear, which limits the utility of this information for clinical triage. The meta-analysis estimate of CFR is lower than for the WHO estimate in all hospitalized individuals with known outcomes. This might represent publication bias, where ETCs with high survival rates are more likely to publish their findings.

Two issues identified from the actionable data analysis have particular clinical relevance. First, the lack of available IPD represents a significant impediment to pooling or comparison of data, especially given the lack of consistency in reporting of summary measures. Second, overlapping cohorts of patients often only became evident due to author experience with the outbreak, and only one article declared previous publication (although we recognize that authors of some manuscripts may not be aware of duplicate publication). The consequence is that those relying on the literature as an evidence base may see consensus in the literature where it does not exist.

Despite poorly standardized reporting of longitudinal clinical data, there was near universal consistency in the reporting of symptoms on admission, no doubt facilitated by use of WHO case reporting forms. Therefore, a solution is in adoption of such best practice standardized case reporting forms for clinical data. These provide clarity regarding objective data (e.g. discriminating between lack of a symptom, with lack of collection of data on a symptom), provide data dictionaries (that provide clear agreed definitions for subjective terms such as ‘fatigue’) and can provide options to allow for different complexities of data collection depending on resources (ranging from symptom reporting only through to comprehensive laboratory and treatment reporting). These forms are increasingly being produced for reporting the natural history of other emerging infectious diseases, such as the WHO protocol for Zika virus infection [56]; the findings of this work further support their use.

Our findings demonstrate that clinical data standards and patient registries, akin to the value of cancer registries, are needed. The aggregated epidemiological data sets curated by WHO are clearly valuable (although not without data collection and analysis constraints themselves) [2]. Due to these, confidence in estimates of the epidemiology of EVD have

Clin Microbiol Infect. Author manuscript; available in PMC 2020 December 11.
improved markedly and have been instrumental in directing public health control. An evidence base of this calibre is just as important for informing the clinical care of patients but does not exist. Such a data set is feasible given that sufficient data for publication of a clinical data set exist for at least 6168 patients. This approach has been used previously for aggre-gation of clinical trial data for malaria, and in this instance has allowed IPD pooled meta-analysis to inform treatment guidelines with greater statistical power than would otherwise be possible [57]. Centralized patient registries have also been suggested for clinical trials of other tropical diseases [58], rare diseases [10] and resistant pathogens [59], where fragmented data reporting also occurs. Importantly, in outbreak settings, centralized databases should also be used to facilitate and advocate for ethical data collection and storage practices. Some of this work must be undertaken in advance of outbreaks, to test and streamline data collection tools that are feasible to use in the early stages of outbreaks when resources are stretched. Therefore, to guide these patient registries, we suggest that a systematic assessment of clinical knowledge gaps for epidemic-prone infections should be prioritized. In the first instance, this should be undertaken for WHO research and development blueprint priority pathogens.

There are several limitations to our analysis. We do not wish for our comparison with WHO data to be misconstrued as comparison with reference standard data, but rather that it constitutes the best available comparison to approximate the representativeness of data included. In particular, our comparisons for symptom data are limited as the WHO data include non-hospitalized cases. For pragmatic reasons we limited analysis to clinical data available in standardized formats (e.g. mean, median, categorical data) although clinically meaningful data may exist in other formats. Planned subgroup analysis by admission route was not possible due poor consistency of reporting. Subgroup analysis for CFR stratified by date was not possible because the distribution of date of admissions within each study was not available.

Conclusions

This systematic review summarizes clinical data for 6168 individuals with EVD. Synthesis of data largely failed due to lack of standardization and transparency in reporting of clinical data. Given the volume of clinical data that was collected, more profound gains in our knowledge regarding the natural history of EVD and factors influencing patient outcomes could have been achieved if the data had been more thoughtfully defined, assembled and shared. The most appropriate method of achieving this is through data-standardization initiatives, patient registries and structured assessment of data priorities. Although not diminishing the practical difficulties of doing so, we advocate strongly for a patient-centred research response to be embedded in the clinical response to EVD outbreaks as a means to improve evidence-based care and patient survival.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.
Acknowledgements

This report is independent research funded and undertaken by the UK Public Health Rapid Support Team, which is funded by the UK Government. The views expressed in this publication are those of the author(s) and not necessarily those of the National Health Service, the National Institute for Health Research or the Department of Health. This work was supported by the Wellcome Trust of Great Britain (grant numbers 107834/Z/15/Z and 106491/Z/14/Z). AR was funded by a Rhodes Scholarship. The funders had no role in the study design, the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the paper for publication.

References


Fig. 1.
Summary of article selection process.
Fig. 2.
Meta-analysis of proportion of individuals with Ebola virus disease (EVD) presenting with a symptom compared with reference data. Blue represents a meta-analysis estimate with low or moderate heterogeneity for the pooled estimate, red represents a meta-analysis estimate with high heterogeneity. Grey is WHO reference data. All estimates are shown with their 95% CI.
Fig. 3.
Meta-analysis of the relative risk of death in individuals with Ebola virus disease (EVD) presenting with a symptom on admission, compared with symptom absence, and showing comparison reference data. Blue represents a meta-analysis estimate with low or moderate heterogeneity for the pooled estimate, red represents a meta-analysis estimate with high heterogeneity. Grey is WHO reference data. All estimates are shown with their 95% CI.
Fig. 4.
Funnel plot of case fatality rate (CFR) by country. The horizontal red line is a reference value that depicts the overall CFR for hospitalized patients during the outbreak reported by WHO, with 95% and 99.8% CI.
**Fig. 5.**
Reporting of predictors of mortality for patients with laboratory confirmed Ebola virus disease (EVD). Cells with a red tab reflect a complex relationship (such as statistical significance for only a subset of data). Asterisk indicates the it includes if association on univariate analysis is not supported by multivariate analysis. VL, viral load; Ct, cycle threshold; Rx, treatment.