



HAL
open science

L'analyse Bayésienne décisionnelle pour le design d'essais cliniques dans le contexte de fièvres hémorragiques virales

Drifa Belhadi

► **To cite this version:**

Drifa Belhadi. L'analyse Bayésienne décisionnelle pour le design d'essais cliniques dans le contexte de fièvres hémorragiques virales. Bio-Informatique, Biologie Systémique [q-bio.QM]. Université Paris Cité, 2023. Français. NNT : 2023UNIP5211 . tel-04818566

HAL Id: tel-04818566

<https://theses.hal.science/tel-04818566v1>

Submitted on 4 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris Cité

École doctorale Pierre Louis de santé publique : épidémiologie et sciences de l'information
biomédicale (ED 393)

Infection, Anti-microbien, Modélisation, Évolution (UMR-S 1137)

L'Analyse Bayésienne Décisionnelle pour le design d'essais cliniques dans le contexte de fièvres hémorragiques virales

Par Drifa BELHADI

Thèse de doctorat de Biostatistique et Biomathématiques

Dirigée par France MENTRÉ et co-dirigée par Cédric LAOUÉNAN

Présentée et soutenue publiquement le 18/12/2023

Devant un jury composé de :

Laura RICHERT	PU-PH, Université de Bordeaux	Rapportrice
Olivier ROBINEAU	PU-PH, Université de Lille	Rapporteur
Marie JASPARD	PH, Sorbonne Université	Examinatrice
Jacques BÉNICHOU	PU-PH, Université de Rouen Normandie	Examineur
France MENTRÉ	PU-PH, Université Paris Cité	Directrice de thèse
Cédric LAOUÉNAN	PU-PH, Université Paris Cité	Co-directeur de thèse
Denis MALVY	PU-PH, Université de Bordeaux	Membre invité

Remerciements

Je tiens tout d'abord à remercier les membres du jury d'avoir accepté d'évaluer ce travail de thèse. Merci à Laura Richert et Olivier Robineau de m'avoir fait l'honneur d'être rapporteurs de cette thèse et pour vos rapports détaillés. Merci à Marie Jaspard et Jacques Bénichou de faire partie de ce jury et d'évaluer ce travail. Et merci à Denis Malvy pour ton aide tout au long de cette thèse, ta grande connaissance clinique ainsi que tes encouragements ont été précieux.

Merci à mes directeurs de thèse, France et Cédric. France, merci de m'avoir soutenu pendant ces 4 années de thèse. Ton aide et tes remarques toujours très justes et pointues ont été plus que précieuses. Tu m'as permis de tenir la barque, de ne pas m'éparpiller et de ne pas me décourager. Cédric, merci de m'avoir proposé cette thèse et de m'avoir fait confiance. Ton soutien, ton écoute, ton aide pendant les moments compliqués, nos nombreuses discussions et tes encouragements ont été tous ces éléments qui m'ont aidé à finir cette thèse !

Merci à toutes les personnes de IAME et notamment à l'équipe BIPID pour votre accueil. Merci à Jérémie pour ton accueil et ton soutien. Merci à Julie pour ta bonne humeur et ta spontanéité. Merci à Emmanuelle de m'avoir expliqué la beauté de l'Aïkido.

Marion, ça a commencé par une chute à St-Malo, et maintenant on se demande comment on pourrait vivre l'une sans l'autre. Tu es toujours là pour me soutenir, m'écouter, t'énerver contre des gens avec moi (sans même savoir pourquoi, juste par principe) et tu ne me juges jamais... pour tout ça merci. Paris, Bretagne, New York, Cambridge... je serais toujours là pour danser et crier sur ABBA avec toi. Emilie G, merci pour ton calme, ton écoute et pour me suivre dans mes sorties foireuses (même si cela veut dire aller voir un film bizarre sur un couple coincé dans une maison de la banlieue américaine avec un bébé reçu par la poste...). Jinju, I miss you !! Thank you so much for being so wise. Alex, pour ton chill et tes vaches (enfin poulets maintenant), et pour tous ces bons moments (le meilleur étant de très loin les danses traditionnelles à Newcastle). Guillaume, merci pour ton réconfort, ton humour et pour ce magnifique calendrier Reine des neiges toujours posé sur mon bureau. Aurélien, merci pour ta gentillesse, tu m'auras beaucoup fait rire (malgré toi parfois), même si je suis toujours là pour t'embrouiller quand tu fais des bêtises, ça aura été un plaisir de traverser cette fin de thèse avec toi. Selma, maintenant qu'on est voisines on va pouvoir venir très tôt au labo ensemble..... Antoine, a.k.a. la fusée, toujours prête pour un nouvel Ekiden là ! Mélanie, quand je pense que tu nous as fait courir à 1h du mat à St-Malo... Morgane, c'est quand tu veux pour refaire une soirée jeux de société ! Romain, pour m'avoir supporté dans le bureau pendant 4 ans, pour tous tes conseils et pour les schokobons ! Nadia, bon courage pour ta fin de thèse, ça va le faire !

Les nouveaux : Clarisse et Adrien, on aura partagé ce bureau pendant quelques mois seulement mais ça aura été un plaisir, vous verrez la thèse c'est beaucoup de bons moments aussi ! Arthur, Assil, Bach, Carlos, Davide, Gaëlle, Lobna, Lucie, Maxime, Nathanaël, Niels, Nicolas, Rémy, Tom, Zaïna merci pour la vie au labo !

Lionel, merci pour ta disponibilité et ta gentillesse. Houda, qu'est-ce qu'on ferait sans toi ! (et je n'écris pas ça parce que je sais que tu vas le lire haha) Merci de nous supporter et pour TOUT le travail que tu fais ! Les anciens : Antonio, Claire, François, Jérémy, Marc,

Mathieu, Nathan, Niccolo, merci pour m'avoir intégré si vite au début ! En même temps on a enchainé : meilleur St-Malo, Florence et escales post-COVID...

Et bien sûr merci à tous les gens de l'URC qui m'ont soutenu pendant ces années... Mes collègues de bureau Pauline et Toni, vous avez toujours été là pour me réconforter, pour pester avec moi, pour m'encourager... C'est vraiment le meilleur bureau du monde entier. Coralie, pour tous ces bons moments, toutes nos discussions, toutes nos sessions gym... Delphine, pour me motiver à aller courir, papoter avec toi en courant c'est le mieux ! Carine, pour te plaindre avec moi de tout ce qui ne va pas, ça fait du bien ! Marina, pour tous tes conseils précieux et ta gentillesse. Samira, pour ton soutien et tes encouragements.

Mes amis du dehors ! Baleina (tu te reconnaitras), tu es toujours là pour m'obliger à prendre soin de moi, pour me réconforter, pour parler pendant des heures alors qu'on n'a pas le temps... On se comprend et c'est ça qui est beau. Nina, ça fait 20 ans maintenant (même si toi tu m'avais oublié) ! Tu es mon roc, toujours là quand j'ai besoin de toi, même si tu n'as pas le temps ! Tu me suis toujours dans mes délires. Merci d'être là. Laura, depuis Londres toujours là ! A me faire rire, à écouter, à comprendre... C'est quand nos prochaines vacances ? Emilie C, tu as enfin quitté le sol britannique pour revenir parmi nous, maintenant on attend tes nouvelles péripéties ! Sophie, pour ta gentillesse et ta douceur. Vanessa, tu m'as appris à faire confiance en mes capacités.

Mes amis de la Croix-Rouge : Dom, merci d'être toujours là pour moi quand j'en ai besoin, tu es un ouragan sur lequel on peut compter, et pour tout ça merci tellement ! Micheline, pour me faire prendre du recul, parler 5 min avec toi me permet de me calmer, de réfléchir posément et de repartir ! Merci à vous deux ! Cécile et Thomas, merci pour votre soutien, votre aide, votre présence... c'est quand vous voulez la prochaine brocante ! Mélanie, merci pour ton soutien, promis c'est bon c'est fini la thèse ! Jorge, merci pour ton aide sur tous mes déménagements (promis j'arrête).

Merci à toute ma famille ! Ma maman, la plus forte, tu m'as montré que si je voulais quelque chose il ne fallait pas attendre que quelqu'un vienne me le donner. Tu es partie de rien et j'espère que tu es fière de tout ça. Djamila, tu m'as toujours soutenu, tu as toujours été fière de moi, tu as toujours voulu le meilleur pour moi... Et bien j'espère que je te fais ressentir la même chose. Kamel, on se dispute mais on est là en cas de besoin, merci pour ton soutien et ton aide. Malik, merci de toujours être là, de me faire rire et de me faire relativiser. Tahir, on a grandi ensemble, ça n'a pas toujours été facile, mais on peut arriver à avoir quelque chose de bien. Karim, mon beau-frère préféré (le seul haha), merci d'être là pour moi, ça m'aide vraiment. Wassila, merci d'être toujours aussi gentille et si drôle, c'est toujours un plaisir de papoter avec toi. Ouahiba, merci pour soutien, ta gentillesse et ton calme ! Et à tous mes neveux et nièces : Daril, c'est quand que tu arrêtes de grandir déjà ? Car c'est plus possible là ! Tamara, tu verras ma chérie, la prépa c'est dur mais ensuite ça va mieux et tu pourras en profiter ! Nayla et Manel, c'est un plaisir de vous voir grandir. Aksil, Amir, Ilian, Kaïs : vous êtes trop petits pour lire tout ça !

Martin, merci de me soutenir, de m'aider, de me faire rire, de me supporter quand ça ne va pas, de prendre soin de moi et de m'obliger à prendre soin de moi.

Je suis sûre que j'ai oublié des gens, donc merci tout le monde !

Et merci à moi-même (parce que quand même hein !).

*« If we're going to fight a disease, let's fight one of the most terrible diseases of all,
indifference.»*

*“- Quand faut y aller, faut y aller !
- Et quand on est con, on est con.”*

Résumé

L'Analyse Bayésienne Décisionnelle pour le design d'essais cliniques dans le contexte de fièvres hémorragiques virales

Les fièvres hémorragiques virales (FHV) sont des maladies souvent associées à des taux de mortalités élevés et des épidémies difficiles à prédire, ce qui entraîne des difficultés à conduire des essais cliniques méthodologiquement forts (tel qu'un essai contrôlé randomisé). Or il est nécessaire d'évaluer de nouveaux traitements contre les FHV. Du fait d'un seuil de significativité statistique très conservateur (seuil de 5%), les investigateurs doivent faire face à des nombres réduits de patients inclus impactant la puissance statistique de l'essai. Ainsi, dans le contexte de FHV, où les taux de mortalité sont élevés et peu de traitements efficaces disponibles, une erreur de type I de 5% peut ne pas être appropriée, voire même soulever des questions éthiques.

La "Bayesian Decision Analysis" (BDA) est une méthode permettant d'incorporer le fardeau de la maladie et les préférences patients lors du design des essais cliniques. En effet, cette méthode permet de prendre en compte le contexte de la maladie afin de calculer un nombre de sujets nécessaires et une erreur de type I optimaux à un essai clinique pour une maladie donnée. Nous avons voulu adapter la méthodologie de la BDA afin de pouvoir l'utiliser dans le contexte des FHV. Pour cela, étant donné que le taux de mortalité est le critère de jugement généralement le plus utilisé dans ce contexte, nous avons adapté la méthodologie à des essais cliniques avec un critère de jugement binaire. Ce travail permet d'obtenir les tailles d'échantillon et erreurs de type I optimales pour un essai randomisé à 2 bras selon plusieurs paramètres à pré-spécifier dont la sévérité de maladie, la taille de la population cible et l'effet attendu du traitement.

Nous avons ensuite illustré cette méthodologie dans le contexte des FHV et notamment de la maladie à virus Ebola. Pour cela, nous avons d'abord conduit une revue systématique afin de récolter, pour chaque FHV, les nombres de cas, mortalité, traitements et séquelles reportés dans la littérature ainsi que sur les sites de l'OMS et du CDC. Nous avons ensuite appliqué les chiffres et hypothèses de la maladie à virus Ebola dans notre méthodologie BDA en prenant le taux de mortalité comme critère de jugement. Nous avons trouvé que pour un essai 2 bras à taille d'échantillon fixe, les erreurs de type I et les tailles d'échantillon optimales peuvent varier grandement selon la taille de la population cible et les taux de létalité de la maladie.

Ce travail permet d'apporter un nouvel outil aux futurs chercheurs souhaitant évaluer un nouveau traitement, notamment pour des maladies très mortelles avec peu de traitements disponibles, tels que les FHV.

Mots-clés: *Bayesian decision analysis, essai clinique, maladie à virus Ebola, fièvre hémorragiques virales, erreur de type I, taille d'échantillon*

Abstract

Bayesian decision analysis to design clinical trials in the context of viral hemorrhagic fevers

Viral hemorrhagic fevers (VHFs) are often associated with high mortality rates and unpredictable outbreaks, leading to difficulties to conduct methodologically strong clinical trials (such as a randomized controlled trial). As very few therapeutic options are available against VHFs, new experimental drugs need to be assessed. Due to a conservative statistical significance threshold (5% threshold), investigators face reduced numbers of included patients with an impact on the statistical power of the trial. Therefore, in the context of VHF, with high mortality rates and few effective therapies available, a 5% type I error may not be appropriate and may raise ethical questions.

The Bayesian Decision Analysis (BDA) is a method that incorporate the burden of disease and patient preferences when designing a clinical trial. It allows taking into account the context of the disease to calculate optimal sample size and type I error for a clinical trial in a given disease context. As the mortality rate is the most commonly used outcome in our context, we extended the BDA framework to clinical trials with a binary outcome, specifically those targeting VHF therapeutics. Finally, this work allows obtaining optimal values of sample size and type I error for a two-arm randomized clinical trial given several parameters to pre-specify including the severity of the disease, the size of the target population and the expected treatment effect.

We illustrated this methodology in the context of VHFs, more particularly Ebola virus disease. We conducted a systematic review to retrieve, for each VHF, the number of cases, mortality, treatments and sequelae reported in the literature and in the WHO and CDC websites. Then we applied Ebola's numbers and hypotheses to our BDA framework with the mortality rate as primary outcome. For a two-arm fixed-sample clinical trial, we found that optimal type I errors and sample sizes vary greatly given the size of the target population and the severity of the disease.

Through this BDA adaptation, we provide a new tool to researchers for designing clinical trials with a binary primary outcome to assess therapeutic options, which can be particularly useful for diseases with high mortality rates and few therapies available.

Keywords: *Bayesian decision analysis, clinical trial, Ebola virus disease, viral hemorrhagic fevers, type I error, sample size*

Sommaire

Remerciements	3
Sommaire	11
Productions scientifiques liées à la thèse	12
1. Introduction	13
1.1. Fièvres hémorragiques virales.....	13
1.2. Quelques éléments de contexte statistique.....	17
1.3. Cas de la maladie à virus Ebola.....	20
1.4. Challenges associés à la recherche clinique dans un contexte d'épidémie de FHV .	21
1.5. Remise en question du seuil de significativité à 5%.....	24
1.6. Bayesian Decision Analysis	25
1.7. Objectifs de la thèse	29
2. Nombres de cas, mortalité, traitements et séquelles des fièvres hémorragiques virales	32
2.1. Résumé	32
2.2. Article 1 publié dans PLOS Neglected Tropical Diseases.....	34
2.3. Séquelles des fièvres hémorragiques virales	60
3. Adaptation de la Bayesian Decision Analysis pour la mise en place d'essai clinique avec un critère de jugement binaire	65
3.1. Résumé	65
3.2. Article 2 en révision dans Journal of Biopharmaceutical Statistics.....	69
4. Discussion et perspectives	90
4.1. Première revue systématique publiée sur le nombre de cas, la mortalité et les traitements des FHV.....	90
4.2. Comment prendre en compte le contexte des FHV lors du design d'un essai clinique ?	91
5. Bibliographie	96
Annexe 1 : Manchon <i>et al.</i> BMC Medical Research Methodology	102
Annexe 2 : protocole de la revue systématique publié dans PROSPERO	112
Liste des tableaux.....	117
Liste des figures.....	117
Table des matières	118

Productions scientifiques liées à la thèse

Publications principales

- Belhadi D, Cho J, Malvy D, Mentré F, Lo AW , Laouéan C. Adaptation of Bayesian decision analysis for clinical trial design to binary outcomes: Illustration for Ebola Virus Disease Treatment. BMC Medical Research Methodology (*en révision*)
- Belhadi D, El Baied M, Mulier G, Malvy D, Mentré F, Laouéan C. The number of cases, mortality and treatments of viral hemorrhagic fevers: A systematic review. PLOS Neglected Tropical Diseases. 2022 Oct 31;16(10):e0010889.
- Manchon P,* Belhadi D,* Mentré F, Laouéan C. Evaluation by simulation of clinical trial designs for evaluation of treatment during a viral haemorrhagic fever outbreak. BMC Medical Research Methodology. 2021 Dec;21(1):1-0.
* *co-premiers auteurs*

Communications orales

- 15^{ème} conférence francophone d'Épidémiologie CLINique (EPICLIN 2021) et 28^{èmes} Journées des Statisticiens des Centres de Lutte Contre le Cancer – Juin 2021, Marseille (France) – “Evaluation by simulation of several clinical trial designs for evaluation of treatment during a viral haemorrhagic fever outbreak”
- 43rd Annual Conference of the International Society for Clinical Biostatistics (ISCB) – Août 2022, Newcastle (Royaume-Uni) – “The number of cases, mortality and treatments of viral hemorrhagic fevers: a systematic review”

Communication affichée

- 44th Annual Conference of the International Society for Clinical Biostatistics (ISCB) – Août 2023, Milan (Italie) – “Adaptation of Bayesian decision analysis for clinical trial design to binary outcomes: Illustration for Ebola Virus Disease Treatment”

1. Introduction

1.1. Fièvres hémorragiques virales

Les fièvres hémorragiques virales (FHV) sont des maladies fébriles causées par des virus à ARN (acide ribonucléique). Les virus à ARN, ou ribovirus, sont des virus utilisant l'ARN comme matériel génétique ou dont la réplication du génome passe par un ARN intermédiaire. Les FHV sont causées par 4 familles de virus à ARN : *arenaviridae*, *filoviridae*, *bunyaviridae* et *flaviviridae* [1]. Elles désignent un ensemble de maladies qui peuvent être bénignes ou très sévères, présentant pour certaines des taux de mortalité extrêmement élevés. Elles peuvent être caractérisées, selon la gravité, par de la fièvre, des douleurs musculaires et articulaires, des malaises, des vomissements, des saignements des muqueuses et gastro-intestinaux, un état de choc dû à une perte de sang, des œdèmes et de l'hypotension [2].

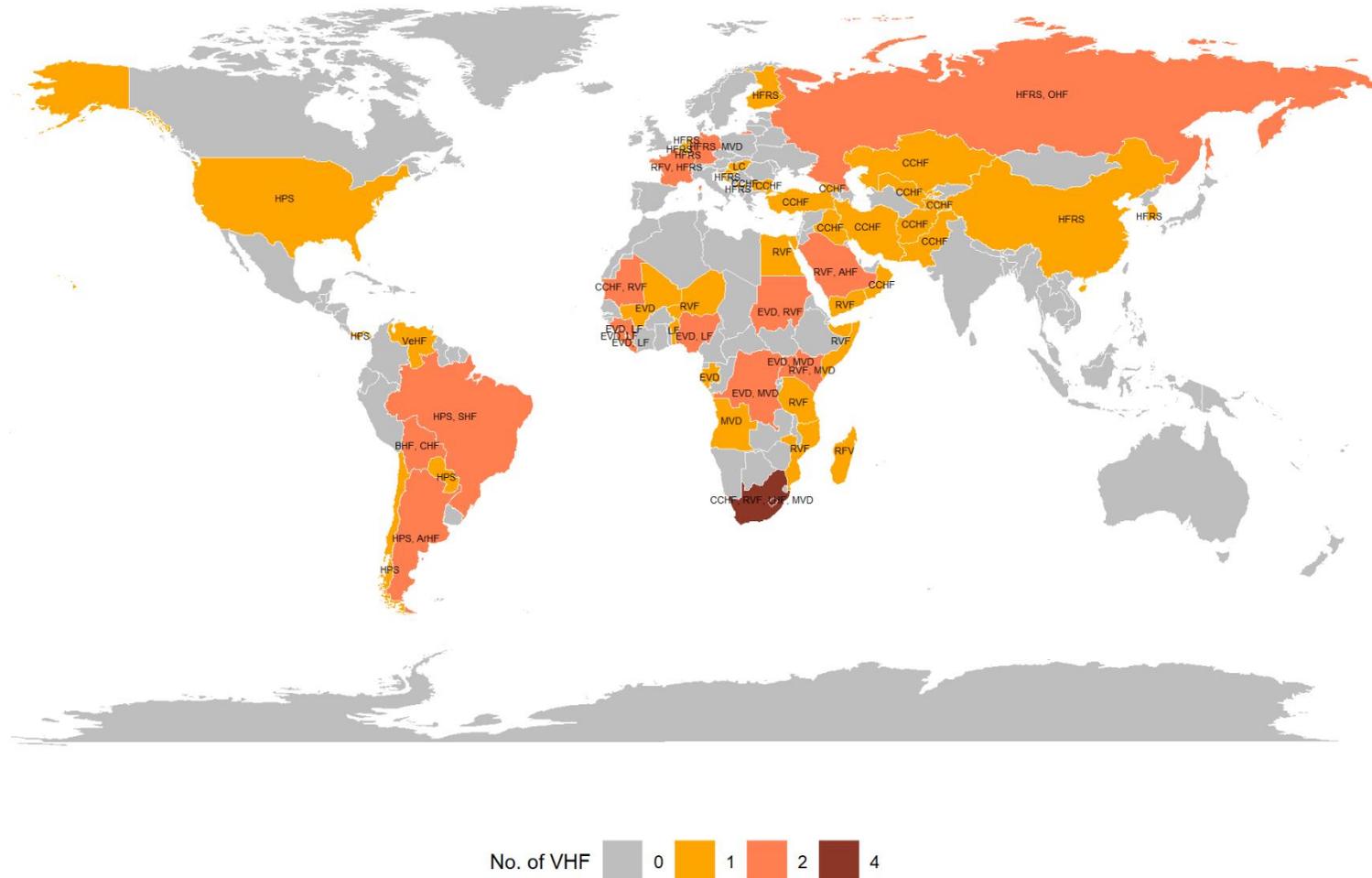
Ces virus infectieux sont essentiellement zoonotiques, c'est-à-dire qu'ils existent naturellement dans les populations animales ou arthropodes. Les modes de transmission sont variés. La transmission peut être interhumaine par contact avec une personne symptomatique via les fluides corporels, ou via des interactions avec des animaux infectés : consommation de viande crue ou de lait non pasteurisé provenant d'animaux infectés ou encore piqûres de moustiques ou de tiques [3]. La période d'incubation varie selon la FHV entre 2 et 35 jours [2,4]. Cette période d'incubation commence avec une phase de premiers symptômes non-spécifiques tels que de la fièvre, des malaises ou des symptômes gastro-intestinaux. Cette phase est suivie par une augmentation de la réplication virale menant à une libération excessive de cytokines pouvant causer une activation endothéliale, une augmentation de la perméabilité vasculaire, une vasodilatation suivie d'hypotension, des défaillances multi-viscérales et le décès [4].

A ce jour, aucun traitement spécifique ou vaccin n'est disponible pour la plupart des FHV [5]. La difficulté de la recherche dans ce domaine réside dans le fait que les FHV sont souvent associées à des épidémies difficiles à prédire et très mortelles. Les épidémies de FHV ont généralement lieu dans les zones du monde présentant les réservoirs du virus associé. Ainsi, bien que les FHV sont présentes sur tous les continents (Figure 1), ce sont très souvent des maladies endémiques c'est-à-dire que celles-ci sont présentes dans des régions données [2]. Les zones touchées par chaque FHV sont souvent délimitées par la présence d'un réservoir

de virus à proximité. Par exemple la maladie à virus Ebola touche essentiellement l'Afrique alors que la fièvre hémorragique de Crimée-Congo touche plus généralement l'Asie ou l'Europe de l'Est [2,5].

Bien que cela soit peu documenté pour la plupart des FHV, les survivants peuvent présenter des séquelles. Si l'on prend l'exemple de la maladie à virus Ebola, des études ont été menées sur les complications au long cours chez des patients ayant guéri de l'infection [6,7]. Plus de la moitié des survivants suivis présentaient au moins un symptôme un an après l'épisode aigu. Ces symptômes consistaient essentiellement en des douleurs articulaires, des maux de tête et de la fatigue.

Figure 1. Distribution des FHV à travers le monde. La couleur indique le nombre de FHV avec une épidémie rapportée jusqu'en janvier 2020 dans chaque pays et l'acronyme représente la FHV – Source : Belhadi et al. PLoS Negl Trop Dis 2022 [5]



Notations : AHF, Alkhurma hemorrhagic fever; ArHF, Argentine hemorrhagic fever; BHF, Bolivian hemorrhagic fever; CHF, Chapare hemorrhagic fever; CCHF, Crimean-Congo hemorrhagic fever; EVD, Ebola Virus Disease; FHV, fièvre hémorragique virale; HPS, Hantavirus Pulmonary Syndrome; HFRS, Hemorrhagic fever with renal syndrome; LF, Lassa fever; LHF, Lujo hemorrhagic fever; MVD, Marburg virus disease; No. of VHF, nombre de FHV reportant au moins une épidémie dans le pays en question; OHF, Omsk hemorrhagic fever; RVF, Rift Valley fever; SHF, Sabia hemorrhagic fever; VeHF, Venezuelan hemorrhagic fever

Les épidémies de FHV les plus récentes rapportées jusqu'en juin 2023 incluent des épidémies de maladie à virus Ebola, de maladie à virus Marburg et de fièvre de Lassa. La dernière épidémie de maladie à virus Ebola a eu lieu entre septembre 2022 et janvier 2023 en Ouganda avec 142 cas confirmés et 55 décès [8]. Contrairement à la grande épidémie de 2013-2016 en Afrique de l'Ouest, celle-ci a été causée par la souche du *Sudan ebolavirus*. Ceci s'explique par le fait que le *Sudan ebolavirus* est présent dans les réservoirs animaux de la région. Bien que de nouveaux traitements (deux cocktails d'anticorps monoclonaux REGN-EB3 et mAb114) et vaccins (un vaccin vivant atténué recombinant constitué du virus de la stomatite vésiculaire et deux vaccins à adénovirus) [9] soient disponibles de façon limitée contre la souche du *Zaire ebolavirus*, ce n'est pas le cas pour le *Sudan ebolavirus*.

Concernant la maladie à virus Marburg, les dernières épidémies ont eu lieu en 2023 en Guinée Equatoriale et en Tanzanie avec respectivement 17 et 8 cas confirmés, incluant 12 et 6 décès [10,11]. Ce sont les premières épidémies rapportées de maladie à virus Marburg dans ces deux régions. Aucun vaccin ou traitement antiviral n'a été approuvé à ce jour pour lutter contre cette maladie. Ainsi, des soins de support standards (réhydratation orale ou via des solutés intraveineux), des traitements symptomatiques ou des traitements de co-infections ont été essentiellement recommandés pour améliorer la survie des patients. Le traitement antiviral remdesivir a également été utilisé à titre compassionnel en Guinée Equatoriale [10]. Un certain nombre de traitements ont été évalués durant ces épidémies, dont des produits sanguins, des immunothérapies ou des traitements médicamenteux [10,12].

Concernant la fièvre de Lassa, la dernière épidémie rapporte un total de 4702 cas suspects avec 877 cas confirmés dont 152 décès au Nigéria en 2023 [13]. Cette maladie est endémique au Nigéria, avec des épidémies saisonnières notamment entre décembre et avril. Des soins de support standards précoces avec de la réhydratation et des traitements symptomatiques sont recommandés pour augmenter la survie des patients. Le traitement par ribavirine (un antiviral à large spectre contre les virus à ARN) est aussi utilisé pour traiter la fièvre de Lassa. Cependant son efficacité et sa tolérance n'ont à ce jour jamais été démontrés de façon robuste. En effet, plusieurs études rapportent des taux de mortalité plus faibles chez les patients traités par ribavirine, notamment si le traitement est administré de façon précoce, mais ces études présentent de sérieux risques de biais méthodologiques [14]. Ainsi, l'OMS souligne la nécessité de produire des résultats fiables sur l'efficacité de la ribavirine pour le traitement de la fièvre de Lassa [15].

Par conséquent, il apparaît nécessaire d'accroître les connaissances autour des FHVs et d'évaluer de nouveaux traitements contre ces maladies. Pour cela nous allons tout d'abord revenir aux principes de base des tests statistiques afin de mieux comprendre les éléments méthodologiques à prendre en compte dans ce contexte.

1.2. Quelques éléments de contexte statistique

a. Généralités sur les tests statistiques

En statistiques, un test, ou test d'hypothèse, est une procédure de décision entre deux hypothèses : l'hypothèse nulle H_0 que l'on considère comme vraie *a priori* et sa complémentaire l'hypothèse alternative H_1 . Un test statistique peut être de 2 formes :

- Unilatéral : on considère une hypothèse alternative H_1 où le signe de la différence entre les deux valeurs que l'on teste est connu ;
- Bilatéral : on considère une hypothèse alternative H_1 où le signe de la différence entre les deux valeurs que l'on teste est inconnu.

Si l'on prend l'exemple d'un test statistique pour démontrer l'efficacité d'un traitement sur un critère de jugement continu, avec μ_1 la moyenne pour le traitement expérimental et μ_2 la moyenne pour le comparateur, appelé « contrôle », alors le test peut être défini comme suit :

$$H_0 : \mu_1 = \mu_2 \text{ (traitement expérimental inefficace)}$$

$$\text{Test bilatéral : } H_1 : \mu_1 \neq \mu_2 \text{ (traitement expérimental efficace ou délétère)}$$

$$\text{Test unilatéral : } H_1 : \mu_1 < \mu_2 \text{ (ou } \mu_1 > \mu_2) \text{ (traitement expérimental efficace)}$$

Pour définir la règle de décision du test statistique, nous utilisons la notion de statistique de test dont la formule varie selon le test utilisé (test de Student, test du χ^2 , ANOVA...). La valeur de cette statistique de test observée dans notre échantillon est ensuite comparée à la valeur critique, c'est-à-dire sa valeur sous l'hypothèse nulle H_0 .

Lors de l'interprétation du résultat d'un test statistique, on accepte un certain risque d'erreur. En effet, il existe deux types d'erreur possibles :

- l'erreur de type I, correspondant au fait de rejeter l'hypothèse nulle H_0 alors que celle-ci est vraie. La probabilité α de commettre cette erreur correspond au seuil de significativité permettant d'interpréter le résultat du test statistique. Le seuil le plus

courant utilisé est 0.05 pour un test bilatéral (0.025 pour un test unilatéral), ce qui signifie que nous acceptons d'avoir 5% de chances de rejeter H_0 à tort.

- l'erreur de type II, correspondant au fait de ne pas rejeter l'hypothèse nulle H_0 alors que celle-ci est fautive. La probabilité β de commettre cette erreur est directement liée à la puissance statistique du test définie par : $Puissance = 1 - \beta$. Ainsi, la puissance statistique du test correspond à la probabilité de rejeter H_0 lorsque celle-ci est fautive.

b. Description de différents designs d'essai clinique utilisés lors d'une épidémie de FHV

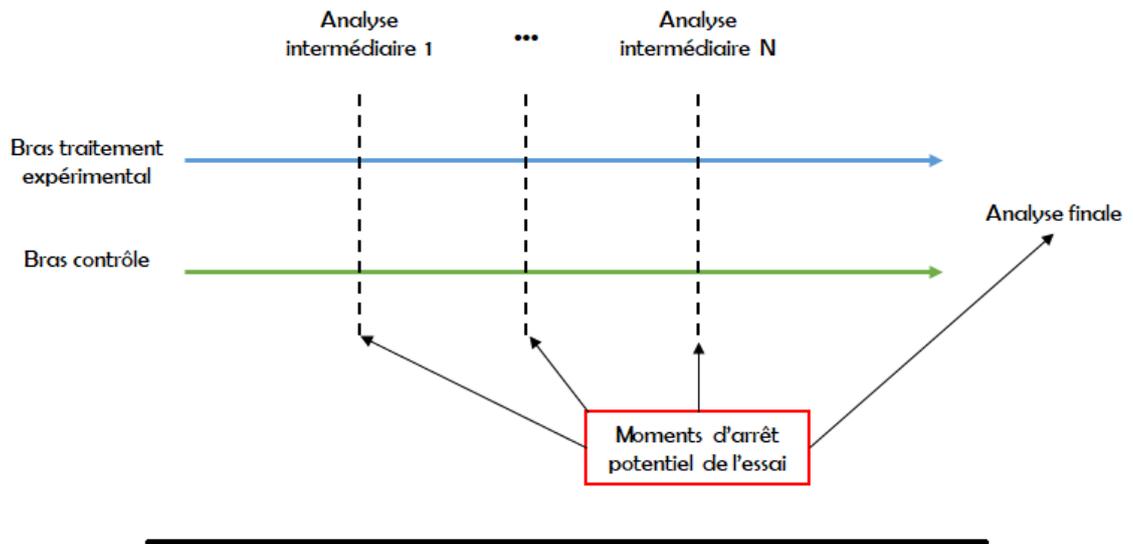
Pour tester l'efficacité et la tolérance d'un traitement expérimental, conduire un essai clinique est nécessaire. Généralement, le critère de jugement principal utilisé dans les essais cliniques de FHV est binaire et correspond au décès. La conception de l'essai clinique ou « design » varie d'un essai à l'autre en fonction notamment :

- Du nombre de groupes ou « bras » de traitement :
 - Essai simple-bras : essai avec un seul bras de traitement correspondant au groupe de patients recevant le traitement expérimental. La proportion observée est généralement comparée à une proportion historique chez des patients avant le début de l'essai n'ayant pas reçu le traitement expérimental.
 - Essai double-bras : essai avec deux bras de traitement, un bras où les patients reçoivent le traitement expérimental et un bras contrôle où les patients reçoivent soit un placebo soit les soins de support standards. Dans un essai clinique randomisé les patients sont aléatoirement assignés à un des deux bras. Les proportions observées dans chacun des deux bras sont statistiquement comparées pour évaluer l'effet du traitement expérimental.
- Du plan expérimental
 - Essai fixe : un nombre de sujets nécessaires est calculé avant la mise en place de l'essai, ce nombre correspond au nombre de patients qu'il faut inclure pour avoir la puissance statistique suffisante pour montrer une différence attendue entre les deux bras sur le critère de jugement principal. Les inclusions sont arrêtées lorsque ce nombre est atteint. Si l'essai est arrêté avant la fin des inclusions, par exemple dans la situation où le nombre de cas de la maladie diminue trop fortement conduisant à une incapacité à inclure de nouveaux patients dans l'essai, alors le résultat sera potentiellement non-significatif sur le critère de jugement principal dû à un manque de puissance statistique.

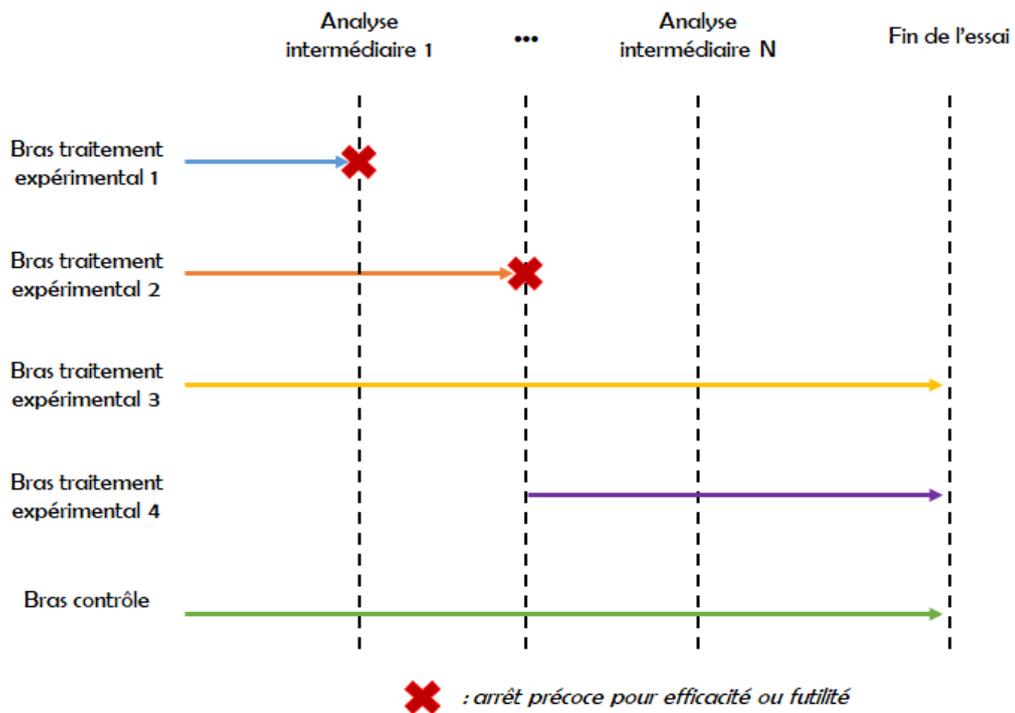
- Essai séquentiel par groupes (cf. Figure 2.A.) : des analyses intermédiaires durant les inclusions sont programmées en amont. Des règles d'arrêt précoce pour futilité ou efficacité sont également définies à l'avance. Ainsi, l'essai peut être arrêté suite aux résultats d'une de ces analyses intermédiaires. Etant donné que la méthodologie de ces analyses intermédiaires et que les règles d'arrêt sont définies en amont du début des inclusions, ces éléments sont pris en compte lors des différentes analyses afin de maintenir la puissance globale, de prendre en compte la multiplicité des tests statistiques et d'éviter une inflation de l'erreur de type I globale.
- Essai multi-bras multi-étapes (cf. Figure 2.B.) : plusieurs traitements expérimentaux sont comparés à un bras contrôle. Des analyses intermédiaires et des règles d'arrêt précoce sont également définies en amont. Chaque traitement expérimental peut être arrêté au cours de l'étude (pour efficacité ou futilité démontrée lors d'une analyse intermédiaire) et de nouveaux traitements expérimentaux peuvent également être ajoutés en cours d'essai.

Figure 2 Schéma pour : A. un essai séquentiel par groupes; B. un essai multi-bras multi-étapes

A. Essai séquentiel par groupes



B. Essai multi-bras multi-étapes



1.3. Cas de la maladie à virus Ebola

Un certain nombre d'études a été conduit afin d'identifier des traitements efficaces contre la maladie à virus Ebola. Plusieurs essais cliniques ont été menés durant l'importante épidémie de 2013-2016 en Afrique de l'Ouest, où trois antiviraux (favipiravir, brincidofovir, TKM-13083), un cocktail d'anticorps monoclonaux (ZMapp), et des traitements par plasma de convalescents ont été évalués [16–20]. La majorité de ces essais cliniques consistaient en des essais simple-bras ou des essais non-randomisés. Ce fut le cas par exemple de l'essai clinique

JIKI évaluant le traitement antiviral favipiravir en Guinée, dont la méthodologie a été assurée par France Mentré et Cédric Laouénan au sein de notre laboratoire de recherche. Un design séquentiel simple-bras avec comparateur historique a été choisi dû à la réticence des professionnels de santé ainsi que des populations locales à accepter un essai randomisé où le patient ne recevrait potentiellement pas le traitement expérimental [17]. Sur les 111 patients analysés, les résultats n'ont pas permis de conclure à l'efficacité du favipiravir aux doses administrées en termes de mortalité à 14 jours après l'initiation du traitement. Cependant, il a été démontré que les concentrations plasmatiques de favipiravir auxquelles ont été exposés les patients étaient inférieures à celles initialement définies comme susceptibles d'exercer une activité antivirale significative [21]. Il a donc été conclu que des doses plus élevées de favipiravir devraient être testées.

Tous les essais conduits durant cette épidémie ont débuté les inclusions après le pic épidémique. Ainsi le nombre de sujets nécessaires calculé avec une erreur de type I de 5% n'a pas été atteint pour la plupart de ces essais, entraînant une incapacité à mettre en évidence un potentiel effet du traitement. Notamment l'essai contrôlé randomisé PREVAIL II a identifié un effet bénéfique du cocktail d'anticorps monoclonaux ZMapp mais non-significatif, certainement dû au manque de puissance statistique lié au fait que le nombre de sujets nécessaires n'ait pas été atteint [16]. Plus récemment un essai multi-bras multi-étapes a été mené en République Démocratique du Congo pour évaluer trois traitements expérimentaux (remdesivir un traitement antiviral, Mab114 un anticorps monoclonal et REGN-EB3 un cocktail d'anticorps monoclonaux) en comparaison avec le ZMapp [22]. Ils ont ainsi pu conclure à l'efficacité de deux traitements (Mab114 et REGN-EB3) en comparaison avec ZMapp en termes de mortalité à 28 jours.

1.4. Challenges associés à la recherche clinique dans un contexte d'épidémie de FHV

Conduire un essai clinique durant une épidémie est associé à un certain nombre de challenges, notamment lorsqu'il s'agit de maladies avec de forts taux de mortalité et peu d'options thérapeutiques disponibles. Une autre complexité liée à la conduite d'un essai clinique lors d'une épidémie de FHV concerne la durée de l'épidémie qui est difficile à prédire et qui peut même être « trop courte » pour pouvoir mettre en place un essai et recruter le nombre de patients nécessaires. En effet, un nombre de patients réduit aura un impact sur la puissance statistique de l'essai et empêchera de démontrer un potentiel effet bénéfique du traitement.

Dans ce contexte, quel type d'essai clinique est-il le plus judicieux de conduire lors d'une épidémie de FHV ?

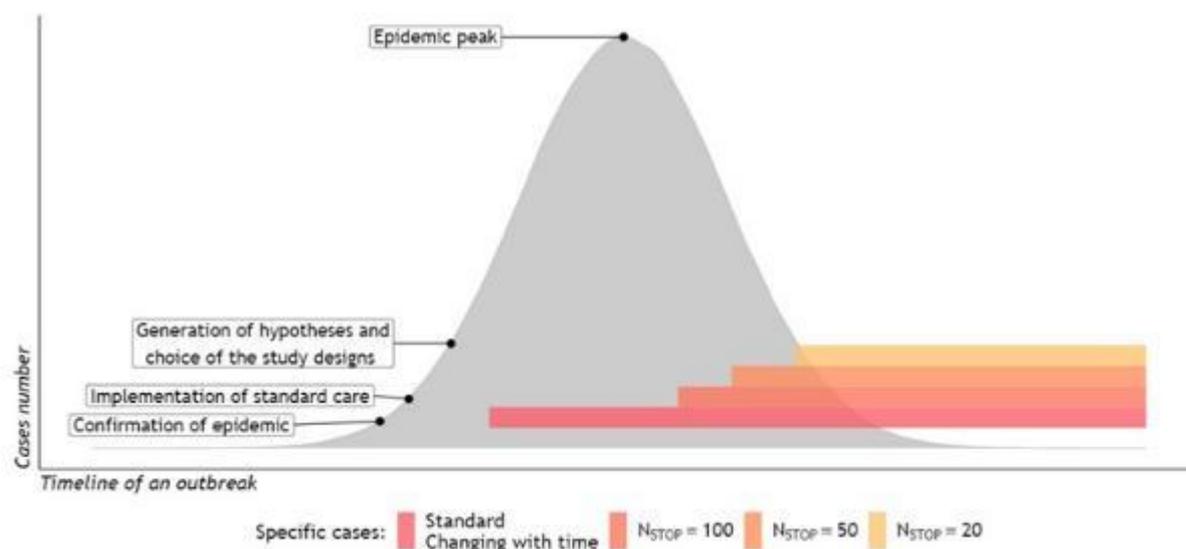
Nous avons voulu répondre à cette question en évaluant plusieurs scénarios afin d'aider les futurs chercheurs à choisir le design le plus adapté à leur situation. Nous avons ainsi conduit une étude de simulation d'essais cliniques comparant plusieurs designs en considérant plusieurs scénarios d'épidémie de FHV. Ce travail a été publié dans *BMC Medical Research Methodology* (Annexe 1) [23].

Lors de la conception d'un essai clinique, le choix du design est important. Les chercheurs doivent par exemple faire le choix parmi les designs suivants : un essai simple-bras versus un essai double-bras ; un essai fixe avec un nombre de sujets nécessaires calculé au moment de la conception de l'essai versus un essai séquentiel par groupes avec un nombre de sujets final qui peut être réduit durant l'essai en cas d'arrêt précoce suite à une analyse intermédiaire. Ainsi, nous avons décidé de comparer quatre designs d'essai clinique : un essai fixe simple-bras, un essai séquentiel simple-bras, un essai fixe double-bras et enfin un essai séquentiel par groupes double-bras. Nous avons considéré le décès à 14 jours comme critère de jugement principal.

Concernant les scénarios, nous avons décomposé la chronologie générale d'une épidémie en deux phases principales : avant et après le pic épidémique. Au début de l'épidémie, des soins de support standards sont mis en œuvre, ce qui peut permettre de diminuer progressivement la mortalité liée à la maladie. Après le pic épidémique, il est de plus en plus difficile d'inclure des nouveaux patients dans l'essai dû au nombre de cas qui diminue. Ainsi, nous avons considéré trois scénarios (cf. Figure 3) :

- Le cas « standard » correspondant à une mise en place de l'essai au début de l'épidémie, où il est possible d'inclure le nombre de sujets nécessaires calculé initialement et avec une mortalité fixe tout au long de l'épidémie ;
- Le cas « évoluant au cours du temps » (« changing with time » sur la Figure 3) similaire au cas « standard » mais où la mortalité diminue au cours de l'épidémie ;
- Le cas « arrêt du recrutement » (« N_{STOP} » sur la Figure 3) correspondant à une mise en place de l'essai après le pic de l'épidémie et où il n'est pas possible d'inclure le nombre de sujets nécessaires calculé initialement. Nous avons considéré trois arrêts possibles (correspondants aux N_{STOP}) : au bout de 20, 50 et 100 patients inclus.

Figure 3 Scénarios considérés au cours d'une épidémie pour l'étude de simulation – *Source* : Manchon et al. BMC Med Res Methodol 2021 [23]



Notation : N_{STOP} , nombre de patients maximal qu'il est possible d'inclure dû au nombre de cas qui diminue

Les résultats de cette étude montrent que la phase de l'épidémie durant laquelle est mis en place l'essai clinique a un impact important sur le design qu'il est préférable de choisir. Au début de l'épidémie, un essai séquentiel par groupes double-bras devrait être préféré. En effet, avant le pic de l'épidémie, le nombre de cas augmente permettant ainsi d'avoir un nombre important de patients éligibles à l'inclusion et donc de pouvoir conduire un essai contrôlé randomisé méthodologiquement plus robuste. De plus, le côté adaptatif d'un design séquentiel par groupes rend possible un arrêt précoce de l'essai si le traitement évalué est jugé futile, délétère ou efficace lors d'une analyse intermédiaire. Par contre, après le pic de l'épidémie, un essai fixe simple-bras devrait être considéré car durant cette phase le nombre de cas diminue. En effet, il est préférable d'atteindre le nombre de sujets nécessaires d'un essai fixe simple-bras, que de conduire un essai méthodologiquement plus robuste mais présentant un manque de puissance statistique dû à une incapacité à inclure le nombre de sujets nécessaires. Cependant, cela nécessite malgré tout un niveau de confiance important concernant la proportion historique (chez des patients non inclus dans l'essai et n'ayant pas reçu le traitement expérimental) qui sera utilisée comme comparateur à la proportion observée dans l'essai pour évaluer le traitement expérimental. Ce choix est notamment discutable si l'on se trouve dans le cas « évoluant au cours du temps » où la mortalité diminue au cours de l'épidémie impliquant forcément pour un essai simple-bras que l'hypothèse faite sur la proportion historique est incorrecte.

1.5. Remise en question du seuil de significativité à 5%

Notre étude de simulation décrite en section 1.4 nous a permis de déterminer quel design d'essai clinique il est préférable de mettre en place selon la situation épidémique dans laquelle nous nous trouvons. Cependant, lorsque l'on doit faire face à ces situations où le nombre de cas est limité, où les taux de mortalité sont très élevés et pour lesquelles peu de traitements sont disponibles, un autre point important à considérer est l'utilisation du seuil habituel de significativité à 5%. En effet, la question se pose si dans ces cas-là utiliser une erreur de type I de 5% dans nos essais cliniques est approprié, voire même éthique.

En 2019, plus de 800 scientifiques ont signé une tribune dans le journal *Nature* appelant à ne plus considérer la notion de « significativité statistique » [24]. Pour cause, il est courant que les études publiées concluent à une absence d'association dès lors que les résultats sont non-significatifs au seuil de 5%. Or, dans certains cas si l'on observe de plus près la taille d'effet du traitement et la p-valeur associée, les résultats ne sont pas aussi catégoriques. Les auteurs nous conseillent ainsi d'accepter le fait qu'il y a de l'incertain dans nos résultats. Ils nous invitent à ne pas prendre de décision catégorique à partir d'un seuil de significativité qui peut ne pas être adapté au cas dans lequel nous nous trouvons, mais plutôt à apprendre à mieux comprendre et interpréter nos estimations, voire même de parler d'intervalle de « compatibilité » plutôt que d'intervalle de confiance.

On pourrait répondre que nous avons besoin de cette notion de seuil de significativité statistique pour obtenir une réponse claire à une question posée. Dans la pratique, si l'on prend l'exemple d'un dossier de soumission à une autorité de santé pour un nouveau traitement, de nombreux autres éléments, que seulement un résultat significatif, sont pris en compte pour prendre une décision finale. L'American Statistical Association (ASA) a également publié un bilan clarifiant les principes sous-jacents à l'utilisation et l'interprétation correcte de la p-valeur [25]. Une conclusion scientifique ne doit pas être fondée seulement sur la significativité d'une p-valeur par rapport à un seuil spécifique. En effet, les chercheurs se doivent d'inclure des facteurs contextuels, tels que le design de l'étude, la qualité de la mesure, les éléments extérieurs et enfin la validité des hypothèses dans leur analyse et conclusion.

Dans des essais cliniques évaluant le bénéfice-risque d'un traitement expérimental, cette convention consistant à prendre une décision à partir d'un seuil de 5% ne prend pas en compte les préférences des patients, c'est-à-dire leur propre arbitrage bénéfice-risque

concernant le traitement, ni de l'impact de cette décision sur la santé publique. Ainsi, pour pallier cette difficulté, plusieurs méthodes ont été proposées. Nous avons décidé de nous intéresser à l'une de ces méthodes, la Bayesian Decision Analysis (BDA), qui permet de prendre en compte ces éléments contextuels dans la conception même d'un essai clinique [26–28].

1.6. Bayesian Decision Analysis

Plusieurs études récentes ont exploré l'utilisation de la méthode de la BDA pour calculer des valeurs optimales de tailles d'échantillon et de valeurs critiques du test à échantillon fixe pour une maladie donnée [26–28]. Ces valeurs optimales sont calculées en prenant en compte le fardeau de la maladie et les préférences des patients en termes de bénéfice-risque de recevoir un traitement expérimental. Il est ainsi possible d'adapter la méthodologie au contexte de la maladie lors de la conception d'un essai clinique.

La BDA est une méthode qui emprunte l'idée du principe de la statistique bayésienne d'incorporer des informations *a priori* lors du design de l'essai clinique, tout en conservant une analyse fréquentiste pour l'essai clinique en lui-même. Cette méthode est fondée sur la définition de « coûts » associés à un essai clinique sous l'hypothèse nulle H_0 et sous l'hypothèse alternative H_1 . Le terme de coût utilisé ici fait référence aux conséquences de prendre une décision correcte ou incorrecte concernant le traitement évalué sur la santé des patients inclus dans l'essai (représentés par n , la taille d'échantillon par bras dans l'essai) et ceux qui pourraient bénéficier ou non du traitement dans le futur (représentés par N , la taille de la population cible, c'est-à-dire le nombre de personnes atteintes par la maladie sur une période de temps définie). Ces coûts ne font donc pas référence à des coûts financiers. Des *a priori* sont ainsi attribués à l'hypothèse nulle H_0 et à l'hypothèse alternative H_1 , pour pouvoir définir le coût engagé par l'essai clinique.

La fonction de coût associée à un essai clinique peut être décomposée en deux groupes : les coûts pendant l'essai et les coûts post-essai. Les coûts pendant l'essai sont indépendants de la conclusion finale de l'essai et dépendent par contre du nombre de patients inclus dans l'essai. Les coûts post-essai dépendent essentiellement du résultat de l'essai et sont supposés indépendants du nombre de patients inclus. Tous ces coûts prennent en compte les effets indésirables du traitement et le fardeau de la maladie via notamment sa prévalence, sa mortalité et les séquelles associées chez les survivants.

Ces différents coûts peuvent être représentés via une matrice de coût bayésienne. Le Tableau 1 rapporte la matrice pour un critère de jugement continu comme définie dans l'article d'Isakov *et al.* [26].

Tableau 1 Matrice de coût bayésienne pour un critère de jugement continu (Bayesian Decision Analysis) - *Source* : Isakov *et al.* J Econom 2019 [26]

	Coûts post-essai		Coûts pendant l'essai
	$\hat{H} = H_0$ (ne pas conclure à l'efficacité du traitement)	$\hat{H} = H_1$ (conclure à l'efficacité du traitement)	
$H = H_0$ (traitement inefficace)	0	Nc_1	nc_1
$H = H_1$ (traitement efficace)	Nc_2	0	$n\gamma Nc_2$

Notations : H , vraie hypothèse sous-jacente ; \hat{H} , conclusion à la fin de l'essai (rejet ou non de l'hypothèse nulle) H_0 , hypothèse nulle pour un test à échantillon fixe correspondant à un traitement inefficace ; H_1 , hypothèse alternative pour un test à échantillon fixe correspondant à un traitement efficace ; N , taille de la population cible ; n , taille d'échantillon par bras dans l'essai ; c_1 , coût lié aux effets indésirables du traitement ; c_2 , coût lié au fardeau de la maladie ; γ , coût marginal lié à l'ajout d'un patient en plus dans chaque bras de l'essai.

A partir de cette matrice de coût, le coût engagé par l'essai, pour un test à échantillon fixe et un critère de jugement continu, peut être défini par :

$$C(H, Z_n, fxd(n, \lambda_n)) = \begin{cases} Nc_1 \mathbb{1}_{\{Z_n \geq \lambda_n\}} + nc_1, & H = H_0 \\ Nc_2 \mathbb{1}_{\{Z_n < \lambda_n\}} + n\gamma Nc_2, & H = H_1 \end{cases}$$

- n est la taille d'échantillon par bras
- Z_n est la statistique de test observée dans l'essai
- λ_n est la valeur critique de la statistique de test
- $fxd(n, \lambda_n)$ correspond à la comparaison de Z_n à λ_n pour un n donné
- $\mathbb{1}_{\{\dots\}}$ est la fonction indicatrice qui prend la valeur 1 si la condition entre accolades est réalisée, 0 sinon.

Les valeurs optimales de taille d'échantillon et de valeur critique de la statistique du test à échantillon fixe sont ensuite obtenues en minimisant la fonction de coût associée à l'essai clinique :

$$(n^*, \lambda_n^*) = \underset{n \in \mathbb{N}, \lambda_n \in \mathbb{R}}{\operatorname{argmin}} C(fxd(n, \lambda_n)) \text{ soumis à } 1 - \beta \leq \operatorname{Power}_{max}$$

- n^* est la taille d'échantillon optimale par bras
- λ_n^* est la valeur critique optimale de la statistique de test

- $C(fxd(n, \lambda_n))$ est défini par l'espérance du coût engagé par l'essai : $\mathbb{E}[C(H, Z_n, fxd(n, \lambda_n))]$
- β est l'erreur de type II
- $Power_{max}$ est la borne supérieure de la puissance.

Cette méthodologie a été développée pour des critères de jugement continus ou censurés. En 2019, un article a été publié décrivant la méthodologie pour un critère de jugement continu qui a été ensuite appliquée aux 30 principales causes de mortalité prématurée aux Etats-Unis en 2010 [26]. Le Tableau 2 présente les résultats de cette étude. D'un côté, sur des maladies très sévères avec peu de traitements disponibles, tel que le cancer du pancréas, on peut voir que l'erreur de type I calculée par BDA est de 23.9% (comparée au seuil standard de 2.5% pour un test unilatéral). De l'autre côté, pour des maladies moins mortelles avec une prévalence élevée, tel que pour le diabète, l'erreur de type I calculée par BDA est de 1.3%. Ceci montre l'impact que peut avoir la prise en compte des éléments contextuels des maladies dans le design de l'essai clinique. Cela a également un impact sur la taille d'échantillon nécessaire pour montrer un potentiel effet bénéfique du traitement. En effet, lorsque l'erreur de type I calculée par BDA est plus élevée que le seuil standard, alors la taille d'échantillon nécessaire sera plus faible que celle calculée de façon standard et inversement si l'erreur de type I calculée par BDA est plus faible que le seuil standard.

Tableau 2 Nombre de sujets nécessaires, valeur critique, erreur de type I et puissance optimaux calculés par BDA pour un effet du traitement attendu $\delta_0 = \sigma/\delta$ (où σ représente l'erreur standard dans chaque bras) - Source : Isakov et al. J Econom 2019 [26]

YLL Rank	Disease Name	Prevalence (Thousands)	Severity	Sample Size	Critical Value	α (%)	Power (%)
1	Ischemic heart disease	8,895.61	0.12	806	1.853	3.2	74.4
2	Lung cancer	289.87	0.45	521	1.094	13.7	82.2
3a	Ischemic stroke	3,932.33	0.15	767	1.755	4.0	75.6
3b	Hemorrhagic/other non-ischemic stroke	949.33	0.16	751	1.723	4.2	75.8
4	Chronic obstructive pulmonary disease	32,372.11	0.06	940	2.182	1.5	70.1
7	Diabetes	23,694.90	0.05	958	2.226	1.3	69.5
8	Cirrhosis of the liver	78.37	0.49	491	1.024	15.3	82.5
9	Alzheimer's disease	5,145.03	0.18	727	1.652	4.9	76.8
10	Colorectal cancer	798.90	0.15	752	1.727	4.2	75.7
11a	Pneumococcal pneumonia	84.14	0.30	596	1.351	8.8	79.0
11b	Influenza	119.03	0.20	679	1.584	5.7	76.4
11c	H influenzae type B pneumonia	21.15	0.26	545	1.378	8.4	75.4
11d	Respiratory syncytial virus pneumonia	14.90	0.07	—	—	—	—
13	Breast cancer	3,885.25	0.05	951	2.218	1.3	69.4
16	Chronic kidney disease	9,919.02	0.04	981	2.288	1.1	68.5
18	Pancreatic cancer	22.67	0.71	384	0.711	23.9	84.6
20	Cardiomyopathy	416.31	0.17	729	1.677	4.7	76.1
21	Hypertensive heart disease	185.26	0.27	633	1.429	7.6	78.7
22	Leukemia	139.75	0.21	671	1.551	6.0	77.0
23	HIV/AIDS	1,159.58	0.10	830	1.926	2.7	73.3
24	Kidney cancers	328.94	0.12	794	1.864	3.1	73.4
25	Non-Hodgkin lymphoma	282.94	0.13	766	1.792	3.7	74.4
27	Prostate cancer	3,709.70	0.05	967	2.259	1.2	68.8
28	Brain and nervous system cancers	59.76	0.30	585	1.339	9.0	78.8
30	Liver cancer	31.27	0.44	492	1.080	14.0	81.1

Notation : YLL, Years of Life Lost due to premature death; α , erreur de type I.

Dans un contexte différent, dans un article publié en 2022, les chercheurs de la même équipe se sont intéressés au traitement expérimental AMX0035 contre la sclérose latérale amyotrophique (SLA) [27]. La SLA est une maladie neurodégénérative mortelle sans traitement curatif disponible. Le traitement AMX0035 est composé de deux molécules, le phénylbutyrate de sodium, qui est un inhibiteur de l'histone-désacétylase, et l'acide tauroursodésoxycholique. La combinaison de ces deux molécules permettrait de réduire le stress cellulaire et la mort des cellules nerveuses [29]. Or, une controverse est née autour du traitement AMX0035 suite à la publication de l'essai de phase II (137 patients inclus : 89 patients dans le bras traitement et 48 patients dans le bras contrôle) présentant des résultats positifs en termes de réduction du risque de progression de la maladie (avec une p-valeur obtenue à 3%) [29]. Bien que les résultats de l'essai de phase III ne soient pas encore disponibles, des associations de patients ont soumis, sans succès, une pétition avec plus de 50 000 signatures à la US Food and Drug Administration (FDA) demandant de rendre ce traitement disponible.

Les auteurs ont ainsi décidé d'utiliser la BDA afin de déterminer l'erreur de type I optimale permettant de faire un compromis entre le besoin de la FDA de limiter l'impact des effets indésirables du traitement et le besoin des patients d'avoir accès à un traitement potentiellement efficace. Ils ont appliqué la BDA à un critère de jugement censuré, des paramètres liés à la chronologie de la recherche et au temps nécessaire à la FDA pour relire le dossier et prendre une décision ont également été pris en compte en plus des paramètres liés au fardeau de la maladie. Le Tableau 3 rapporte les principaux résultats de leur analyse. Plusieurs scénarios ont été considérés : une probabilité que l'hypothèse nulle ne soit pas rejetée p_0 variant entre 10% et 95%, un rapport « signal sur bruit » (*i.e.* rapport entre un effet réel du traitement et les fluctuations aléatoires de l'effet du traitement) ρ variant entre 0.25 et 0.50, et un fardeau de la maladie mesuré soit en utilisant les YLD (Years lived with Disability) soit les DALY (Disability-Adjusted Life Years). Ainsi pour $p_0 \leq 70\%$ et $\rho < 0.5$, l'erreur de type I optimale calculée par BDA est supérieure à 10%, soit plus élevée que la p-valeur de 3% rapportée dans l'essai de phase II. Ces résultats donnent un point de vue intéressant sur la possibilité d'incorporer des éléments contextuels (ici une maladie mortelle sans traitement curatif disponible) dans l'évaluation de l'effet thérapeutique de l'AMX0035.

Tableau 3 Erreurs de type I optimales calculées par BDA en utilisant les données de l'essai de phase II évaluant AMX0035 et en prenant 89 patients dans le bras traitement et 48 patients dans le bras contrôle – *Source* : Xu et al. Amyotroph Lateral Scler Frontotemporal Degener 2022 [27]

p_0 (%)	ρ	α^* (%)	Power (%)
Burden of Disease Measure: YLD			
10	0.25	29.0	80.0
10	0.33	15.4	80.0
10	0.50	2.6	80.0
30	0.25	29.0	80.0
30	0.33	15.4	80.0
30	0.50	2.6	80.0
50	0.25	26.3	77.7
50	0.33	15.4	80.0
50	0.50	2.6	80.0
70	0.25	10.7	56.1
70	0.33	11.8	75.1
70	0.50	2.6	80.0
95	0.25	0.3	8.9
95	0.33	1.0	32.7
95	0.50	1.5	73.3
Burden of Disease Measure: DALY			
10	0.25	29.0	80.0
10	0.33	15.4	80.0
10	0.50	2.6	80.0
30	0.25	29.0	80.0
30	0.33	15.4	80.0
30	0.50	2.6	80.0
50	0.25	29.0	80.0
50	0.33	15.4	80.0
50	0.50	2.6	80.0
70	0.25	29.0	80.0
70	0.33	15.4	80.0
70	0.50	2.6	80.0
95	0.25	3.8	35.1
95	0.33	5.6	60.8
95	0.50	2.6	80.0

Notations : p_0 , probabilité que l'hypothèse nulle ne soit pas rejetée ; ρ rapport « signal sur bruit » ; α^* , erreur de type I optimale ; YLD, Years lived with Disability ; DALY, Disability-Adjusted Life Years.

A travers ces différents résultats, nous nous sommes demandés si nous ne pouvions pas adapter cette méthode au cas des FHVs. Il serait ainsi possible de prendre en compte le contexte des FHVs lors du design des essais cliniques afin de pouvoir pallier la difficulté de trouver un traitement efficace pour des maladies mortelles dont la majorité n'a pas d'option thérapeutique disponible.

1.7. Objectifs de la thèse

L'objectif des travaux de cette thèse porte sur l'étude d'une nouvelle approche méthodologique pour l'aide à la mise en place d'un essai clinique lors d'une épidémie de FHV. En effet, du fait d'un seuil de significativité statistique à 5% conservateur, les nombres de

sujets à inclure peuvent être relativement élevés en comparaison du nombre de cas de la maladie.

Dans une première partie nous avons réalisé une revue systématique de la littérature afin de collecter les informations disponibles pour chaque FHV. Nous nous sommes concentrés sur 18 FHVs et avons recueilli les données sur les épidémies, le nombre de cas, les taux de mortalité, les traitements et les séquelles de chacune de ces FHVs.

Dans une seconde partie nous avons développé la méthodologie de la BDA afin de pouvoir l'utiliser dans le contexte des FHVs. Pour cela, étant donné que la mortalité à un temps donné est généralement le critère de jugement le plus utilisé pour les FHVs, nous avons adapté la BDA à des essais cliniques avec un critère de jugement binaire. Nous avons ensuite illustré cette méthode en utilisant le cas de la maladie à virus Ebola. Ce travail a été effectué en collaboration avec l'équipe du professeur Andrew Lo, directeur du Laboratory of Financial Engineering au MIT Sloan School of Management à Boston (Etats-Unis), qui est à l'origine des différents travaux sur le développement et l'utilisation de la BDA dans la conception d'essais cliniques [26–28].

2. Nombres de cas, mortalité, traitements et séquelles des fièvres hémorragiques virales

2.1. Résumé

Dans cette première partie, l'objectif est de proposer un état des lieux des informations disponibles sur les FHVs en termes de nombre de cas, de mortalité, de traitements disponibles et des séquelles. Le protocole de cette revue systématique a été publié dans le registre prospectif international des revues systématiques PROSPERO (CRD42020167306 ; cf. Annexe 2) selon les recommandations PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis – Protocols) [30].

Les critères PICOS (Participants, Intervention, Comparison, Outcomes and Study types) suivants ont été utilisés:

- **Participants** : humains infectés par un agent pathogène d'une FHV de la liste ci-dessous définie à partir des FHVs disponibles sur le site du CDC (Centers for Disease Control and Prevention) en décembre 2019 [31] :
 - Fièvre hémorragique d'Alkhurma
 - Fièvre hémorragique argentine
 - Fièvre hémorragique bolivienne
 - Fièvre hémorragique du Chapare
 - Fièvre hémorragique de Crimée-Congo
 - Maladie à virus Ebola
 - Syndrome pulmonaire à hantavirus
 - Fièvre hémorragique avec syndrome rénal
 - Fièvre de Kyasanur
 - Fièvre de Lassa
 - Fièvre hémorragique virale due au virus Lujo
 - Chorioméningite lymphocytaire
 - Maladie à virus Marburg
 - Fièvre hémorragique d'Omsk
 - Fièvre de la vallée du Rift
 - Fièvre hémorragique brésilienne (due au virus de Sabiá)

- Méningoencéphalite à tiques (d'un variant hémorragique)
- Fièvre hémorragique vénézuélienne.
- **Interventions/Comparateurs** : la revue ne se restreint à aucune intervention en particulier.
- **Critères de jugement** : toute référence rapportant des nombres de cas, des taux de létalité, des taux de mortalité associés ou non à des traitements d'une FHV et des informations sur les séquelles.
- **Types d'étude** : informations officielles venant de l'OMS, du CDC ou de sites de santé nationaux ; études transversales, de cohorte ou cas-contrôle ; rapports descriptifs ; essais cliniques.

La revue a été divisée en deux parties. La première recherche a été conduite en décembre 2019 sur les sites de l'OMS et du CDC et la seconde recherche était une revue systématique de la littérature conduite en janvier 2020 sur les bases de données MEDLINE, Embase and CENTRAL (Cochrane Central Register of Controlled Trials). Pour cette seconde recherche, chaque publication a été revue par moi-même et un autre chercheur de façon indépendante, tout d'abord sur titres et abstracts puis sur articles complets en utilisant les critères PICOS définis précédemment. Les discordances ont été résolues par un troisième chercheur.

Pour la première recherche, 57 rapports de l'OMS et 96 rapports du CDC ont été identifiés, pour un total de 141 rapports retenus dans notre revue systématique. Pour la seconde recherche, 4461 publications ont été identifiées, pour un total de 126 publications répondant aux critères PICOS et retenues dans notre revue systématique. Le risque de biais de chaque publications incluse a été analysé à travers l'outil RoB (Risk of Bias) décrit dans les recommandations Cochrane (Cochrane Handbook for Systematic Reviews of Interventions) pour les essais randomisés ; et le Newcastle-Ottawa Quality Assessment Form pour les autres types d'étude [32,33].

La plupart des références (rapports et publications) ont été publiées après 2010 et concernaient essentiellement le continent africain. Parmi les 38 références rapportant des résultats sur des traitements, 32% (n = 12) étaient des études rétrospectives, 26% (n = 10) étaient des essais contrôlés randomisés, 16% (n = 6) étaient des essais non-randomisés et les 26% restants (n=10) correspondaient à d'autres types d'études (études cas-contrôle, rapports descriptifs...). La plupart des références portaient sur un nombre limité de FHVs, dont la

maladie à virus Ebola, la fièvre hémorragique de Crimée-Congo ou encore le syndrome pulmonaire à hantavirus. Nous avons identifié des références rapportant au moins une épidémie de FHV dans 55 pays. La plupart de ces pays, essentiellement en Afrique, rapportaient des épidémies de deux FHVs différentes. Les nombres de cas et taux de mortalité varient grandement pour une même FHV selon l'épidémie et le pays.

Nous avons donc identifié seulement 38 publications évaluant l'efficacité de traitements sur la mortalité liée aux FHVs et cela ne concernait que 6 FHVs : la fièvre hémorragique argentine, la fièvre hémorragique de Crimée-Congo, la maladie à virus Ebola, le syndrome pulmonaire à hantavirus, la fièvre hémorragique avec syndrome rénal et la fièvre de Lassa. De plus, une grande partie de ces publications ont été classées comme ayant un risque important de biais. Peu d'essais contrôlés randomisés ont été identifiés impliquant un manque de preuve robuste d'efficacité des traitements évalués. Cela souligne le besoin de conduire des essais cliniques méthodologiquement robustes dans le contexte des FHVs.

Les résultats de cette revue systématique sur les nombres de cas, la mortalité et les traitements de chaque FHV ont fait l'objet d'une publication dans *PLOS Neglected Tropical Diseases* en octobre 2022. Les données sur les séquelles associées aux FHVs (non publiées) sont présentées dans la section 2.3 de ce mémoire de thèse.

2.2. Article 1 publié dans PLOS Neglected Tropical Diseases

RESEARCH ARTICLE

The number of cases, mortality and treatments of viral hemorrhagic fevers: A systematic review

Drifa Belhadi^{1,2*}, Majda El Baied², Guillaume Mulier², Denis Malvy^{3,4}, France Mentré^{1,2}, Cédric Laouénan^{1,2}

1 Université Paris Cité, Inserm, IAME, Paris, France, **2** AP-HP, Hôpital Bichat, Département d'Epidémiologie Biostatistiques et Recherche Clinique, Paris, France, **3** UMR 1219 Inserm/EMR 271 IRD, University of Bordeaux, Bordeaux, France, **4** Department for Infectious and Tropical Diseases, University Hospital Center Pellegrin, Bordeaux, France

* drifa.belhadi@aphp.fr



Abstract

Background

Viral hemorrhagic fevers (VHFs) are a group of diseases, which can be endemo-epidemic in some areas of the world. Most of them are characterized by outbreaks, which occur irregularly and are hard to predict. Innovative medical countermeasures are to be evaluated but due to the field specificities of emerging VHF, challenges arise when implementing clinical studies. To assess the state of the art around VHFs, we conducted a systematic review for all reports and clinical studies that included specific results on number of cases, mortality and treatment of VHFs.

Methods

The search was conducted in January 2020 based on PRISMA guidelines (PROSPERO CRD42020167306). We searched reports on the WHO and CDC websites, and publications in three international databases (MEDLINE, Embase and CENTRAL). Following the study selection process, qualitative and quantitative data were extracted from each included study. A narrative synthesis approach by each VHF was used. Descriptive statistics were conducted including world maps of cases number and case fatality rates (CFR); summary tables by VHF, country, time period and treatment studies.

Results

We identified 141 WHO/CDC reports and 126 articles meeting the inclusion criteria. Most of the studies were published after 2010 ($n = 97$ for WHO/CDC reports and $n = 93$ for publications) and reported number of cases and/or CFRs ($n = 141$ WHO/CDC reports and $n = 88$ publications). Results varied greatly depending on the outbreak or cluster and across countries within each VHF. A total of 90 studies focused on Ebola virus disease (EVD). EVD outbreaks were reported in Africa, where Sierra Leone (14,124 cases; CFR = 28%) and Liberia (10,678 cases; CFR = 45%) reported the highest cases numbers, mainly due to the 2014–

OPEN ACCESS

Citation: Belhadi D, El Baied M, Mulier G, Malvy D, Mentré F, Laouénan C (2022) The number of cases, mortality and treatments of viral hemorrhagic fevers: A systematic review. *PLoS Negl Trop Dis* 16(10): e0010889. <https://doi.org/10.1371/journal.pntd.0010889>

Editor: Mabel Carabali, McGill University Faculty of Medicine and Health Sciences, CANADA

Received: June 3, 2022

Accepted: October 14, 2022

Published: October 31, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pntd.0010889>

Copyright: © 2022 Belhadi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

2016 western Africa outbreak. Crimean-Congo hemorrhagic fever (CCHF) outbreaks were reported from 31 studies in Africa, Asia and Europe, where Turkey reported the highest cases number (6,538 cases; CFR = 5%) and Afghanistan the last outbreak in 2016/18 (293 cases; CFR = 43%).

Regarding the 38 studies reporting results on treatments, most of them were non-randomized studies (mainly retrospective or non-randomized comparative studies), and only 10 studies were randomized controlled trials. For several VHFs, no specific investigational therapeutic option with strong proof of effectiveness on mortality was identified.

Conclusion

We observed that number of cases and CFR varied greatly across VHFs as well as across countries within each VHF. The number of studies on VHF treatments was very limited with very few randomized trials and no strong proof of effectiveness of treatment against most of the VHFs. Therefore, there is a high need of methodologically strong clinical trials conducted in the context of VHF.

Author summary

Viral hemorrhagic fevers are a group of febrile illnesses that constitutes a challenge to public health. Viral hemorrhagic fevers are endemic in some areas of the world and are often associated with high morbidity and mortality. Due to high mortality rates and outbreaks that are irregular and hard to predict, difficulties arise when undertaking clinical studies to assess new treatments against viral hemorrhagic fevers. To provide an extensive overview of viral hemorrhagic fevers, we conducted a systematic review to retrieve available information on number of cases, mortality and treatments of viral hemorrhagic fevers. We found that cases number and mortality varied greatly across outbreaks and across countries within each viral hemorrhagic fever. The number of studies on viral hemorrhagic fever treatments was very limited with very few methodologically strong studies. Moreover, for several viral hemorrhagic fevers we did not identify specific investigational therapeutic option with strong proof of effectiveness on mortality. Therefore, there is a high need to conduct methodologically strong studies to investigate treatments against viral hemorrhagic fevers.

Introduction

Viral hemorrhagic fevers (VHFs) are a group of febrile illnesses caused by four families of RNA viruses: arenaviridae, filoviridae, bunyaviridae and flaviviridae. [1] These highly infectious viruses are mainly zoonotic; meaning they naturally exist in animal or insect populations. [2] When a person encounter an infected animal or insect, the virus can spread through spillover into the human population, and subsequently is transmitted from person-to-person through contact with blood or other body fluids. Whatever their capacity to drive paramount hemorrhagic manifestations, many VHFs can cause severe, life-threatening disease. The agents that are causative of VHF are often classified as Biosafety Level 4 (BSL-4) pathogens that require special laboratory facilities with the highest level of safety measures. [2,3] VHFs are distributed worldwide and are often associated with high morbidity and mortality. Most of them

are characterized by clusters or even outbreaks occurring irregularly and almost resulting from spillover or more recently from human reservoirs constituted by immunologically preserved sanctuaries where the virus may persist after recovery of convalescent survivors.

Patient outcomes are highly associated with the timing of curative treatment with improved outcomes when the specific or supportive therapy is started early. [4,5] Hence, VHF care remains essentially supportive and some VHFs are treated only with basic medical care that is not always reaching the optimized level of standards aimed to prevent or control the multi-systemic disorders that account for bad outcome. [2,5,6] Available VHF drugs are limited and clinical data on the efficacy of VHF drugs is restricted. [4,5] New investigational treatments need to be evaluated but due to the field specificities of emerging VHF, difficulties arise when conducting clinical studies. Indeed, hard to predict outbreak duration leads to limited number of recruited patients. Moreover high case fatality rate (CFR) leads to reluctance to use methodologically strong trial design such as randomized controlled trials as part of the patients will not receive the potentially beneficial treatment. [7]

VHFs have recently caused various outbreaks around the world. To assess the state of the art around VHFs, we systematically reviewed the World Health Organization (WHO) and Centers for Disease Control (CDC) websites and published literature for all reports and clinical studies that included specific results on number of cases, mortality and treatments of VHFs. We decided to focus mostly on VHFs caused by a selection of arenaviridae, filoviridae, bunyaviridae and flaviviridae and did not look at some other important VHF conditions such as the severe dengue and yellow fever.

Methods

The systematic review was registered on the International prospective register of systematic reviews (PROSPERO 2020 CRD 42020167306). The objective of the systematic review was to review the case fatality rates, number of cases and treatment options of VHFs. The initial protocol also included the review of sequelae, which will not be presented here.

Eligibility criteria

The PICOS (Participants, Intervention, Comparison, Outcomes and Study types) framework was used to identify relevant data.

Participants. Humans infected with a pathogen causative of a VHF from the following list: [8]

- Alkhurma hemorrhagic fever (AHF)
- Argentine hemorrhagic fever (ArHF)
- Bolivian hemorrhagic fever (BHF)
- Chapare hemorrhagic fever (CHF)
- Crimean-Congo hemorrhagic fever (CCHF)
- Ebola virus disease (EVD)
- Hantavirus pulmonary syndrome (HPS)
- Hemorrhagic fever with renal syndrome (HFRS)
- Kyasanur Forest disease (KFD)
- Lassa fever (LF)

- Lujo hemorrhagic fever (LHF)
- Lymphocytic choriomeningitis (LCM)
- Marburg virus disease (MVD)
- Omsk hemorrhagic fever (OHF)
- Rift Valley fever (RVF)
- Sabia-associated hemorrhagic fever (SHF)
- Tick-borne encephalitis (from a hemorrhagic variant)
- Venezuelan hemorrhagic fever (VeHF)

Interventions/Comparisons. This review did not focus on any specific intervention.

Outcomes. Any studies reporting number of cases and/or CFR and/or mortality rates associated with treatments of a VHF were included.

Study types. Official information from WHO/CDC and national health websites, cross-sectional, cohort/case-control studies, descriptive reports and clinical trials were included.

Study language. Included studies were limited to studies published in English or French. The following types of study were deemed ineligible; case reports, case series, systematic reviews and meta-analyses.

Search strategy

The systematic review consisted in two parts. The first search was conducted in December 2019 on the WHO and CDC websites to identify the number of cases and deaths associated with each disease by year and by country. The second search consisted in a systematic literature search of bibliographic databases based on the PRISMA guideline. Relevant studies were identified by searching MEDLINE, Embase and the Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library. We searched the electronic databases until January 21st, 2020. Search terms were developed using a combination of MeSH/EMTREE terms and free-text terms to capture the relevant populations, outcomes and study types (cf. [S1 Fig](#) in Supplementary Material). Reference lists of included studies were not reviewed. Additional hand searches were performed on national health websites.

Study selection

For the systematic literature search, after removing duplicates, two researchers (DB and GM) independently evaluated all identified citations based on titles and/or abstracts. In case of disagreement, a third researcher (CL) served as tiebreaker. Full-text publications of studies included based on title and abstract were retrieved and reviewed by two researchers (DB and ME) to assess eligibility based on the inclusion/exclusion criteria. A third researcher (CL) served as a tiebreaker for any discordant decisions. Justification for study exclusion was documented.

Data extraction

Information identified from the WHO and CDC websites were extracted for each disease by period and by country. For the systematic literature search, following the study selection process, qualitative and quantitative data were extracted from each of the included articles. Numbers were extracted as reported in each study. If the CFR was not reported in a study, the number of cases and the number of deaths were used, if available, to estimate the missing CFR.

Risk of bias assessment

Quality assessment was performed using the Newcastle-Ottawa Quality Assessment Form [9] to evaluate cohort studies, and the Risk of Bias (RoB) tool described in the Cochrane Handbook for Systematic Reviews of Interventions [10] to evaluate randomized clinical trials.

Strategy for data synthesis

A narrative synthesis approach by disease was used for each outcome. Descriptive statistics were conducted to describe the published articles and WHO/CDC reports. World map of the total number of cases and CFR were produced and summary tables of the number of cases and CFR by VHF, country and time period were also produced. Summary tables reporting the characteristics and results of each study on VHF treatments were produced.

World maps were produced using R software version 3.6.0.

Results

A total of 57 reports were identified from the WHO website, and 96 reports from the CDC website. After removing duplicates, 141 reports were extracted. After removing duplicates, 4 461 publications were identified from Medline, Embase, CENTRAL and through hand searches. We excluded 4 153 publications based on the screening of titles and abstracts. A total of 308 publications were included for full text review. Finally, 126 publications met the inclusion criteria and were extracted. The selection process and the numbers at each stage are shown in Fig 1. The exhaustive lists of included publications and WHO and CDC reports are reported in S1 and S2 Tables in Supplementary Material.

Characteristics of included studies

The characteristics of the included studies are summarized in Table 1. Most of the studies were published after 2010 (69%, n = 97 for WHO and CDC reports and 74%, n = 93 for publications) and were conducted mainly in Africa (57%, n = 81 for WHO and CDC reports and

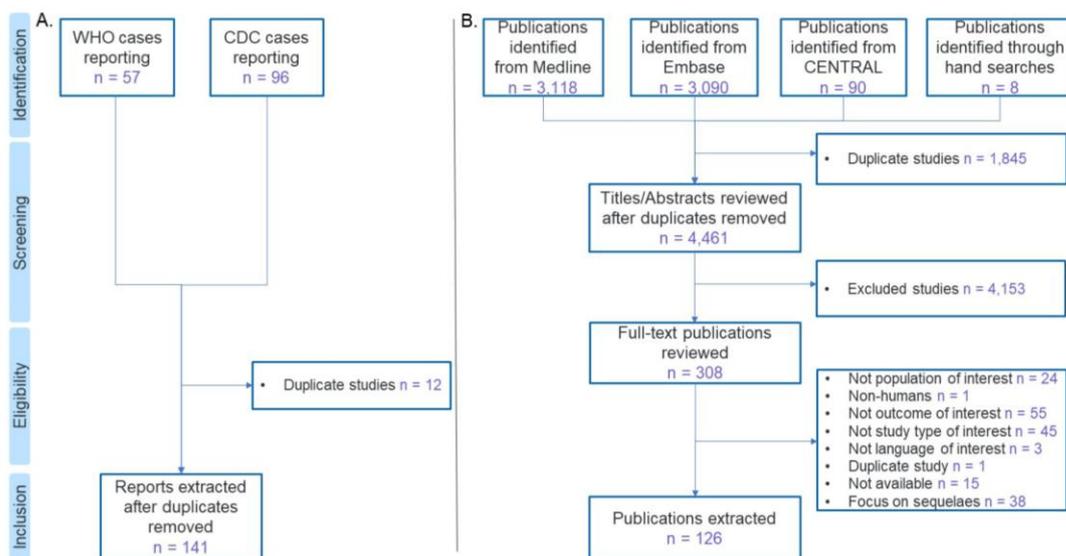


Fig 1. Study selection flow diagram: A. For the WHO and CDC websites search; B. For the systematic literature search.

<https://doi.org/10.1371/journal.pntd.0010889.g001>

48%, $n = 60$ for publications). The majority of studies reporting number of cases and/or CFR were descriptive reports (100%, $n = 141$ for WHO and CDC reports and 90%, $n = 79$ for publications) followed by retrospective studies (10%, $n = 9$ for publications). Regarding studies reporting results on treatments, 32% ($n = 12$) of them were retrospective studies, followed by 26% ($n = 10$) of randomized controlled trial and 16% ($n = 6$) of non-randomized comparative studies. Most of WHO and CDC reports focused on EVD (33%, $n = 47$), followed by HPS (18%, $n = 26$), LF (17%, $n = 24$) and RVF (16%, $n = 23$). In terms of publications, most of them focused also on EVD (33%, $n = 43$), followed by CCHF (21%, $n = 27$), HFRS (14%, $n = 18$) and HPS (10%, $n = 13$). No relevant data were identified for KFD, LCM and tick-borne encephalitis (from a hemorrhagic variant).

Risk of bias of included studies

On the 38 studies reporting results on treatments, 10 were randomized trials and 28 were non-randomized. Regarding the randomized comparative studies, only two were classified as having a “low risk of bias”, both published after 2010, and the other eight studies were classified as having “some concerns” (cf. [S3 Table](#)). Regarding the non-randomized studies, 13 were classified as “poor quality”, 5 as “fair quality” and 10 as “good quality” (cf. [S4 Table](#)). On those 28 non-randomized studies, 6 were published before 2010 and 67% of them ($n = 4/6$) were classified as “poor quality”, compared with 41% ($n = 9/22$) for those published after 2010.

The 88 studies reporting number of cases and CFR were evaluated using the Newcastle-Ottawa Quality Assessment Form: 59 were classified as “poor quality” and 29 as “fair quality” (cf. [S5 Table](#)). On these 88 studies, 31 were published before 2010 and a similar proportion than those published after 2010 were classified as “poor quality” (65%, $n = 20/31$, for studies before 2010, compared with 68%, $n = 45/66$, for studies published after 2010).

Number of cases and CFR

The worldwide distribution of VHF by country is reported in [Fig 2](#). We identified studies reporting outbreaks of at least one VHF across 55 countries. We found studies on 4 different VHFs in South Africa: CCHF, RVF, MVD and LHF. Based on identified studies, a total of 16 countries were associated with 2 VHFs, mainly in Africa with 9 countries, followed by America with 3 countries and Asia and Europe with 2 countries each. More detailed results are reported below for the VHFs with at least 5 studies. They are presented by alphabetical order.

Crimean-Congo hemorrhagic fever (CCHF)

The distribution of CCHF cases and CRFs are reported in [Fig 3](#). Cases were reported in Africa, Asia and Europe in the following countries: Afghanistan, Bulgaria, Georgia, Iran, Iraq, Kazakhstan, Kosovo, Mauritania, Oman, Pakistan, South Africa, Tajikistan, Turkey and Uzbekistan. The first documented cases were reported in 1948–1969 in Kazakhstan (89 cases, CFR = 25%). At the date of the review, the country with the highest reported number of cases was Turkey with a total of 6,538 cases (CFR = 5%) and the last documented cases reported in 2016–2018 in Afghanistan (293 cases, CFR = 43%). More details by country and period are reported in [S6 Table](#).

Ebola virus disease (EVD)

The distribution of EVD cases and CRFs are reported in [Fig 4](#). Outbreaks were reported in sub-saharan Africa in the following countries: Democratic Republic of the Congo, Gabon, Republic of Guinea, Liberia, Mali, Nigeria, Sierra Leone, Sudan and Uganda. The first reported

Table 1. Characteristics of included studies.

	WHO/CDC reports (N = 141)	Publications (N = 126)
Year of publication		
Before 2000	23 (16%)	13 (10%)
Between 2000 and 2010	21 (15%)	20 (16%)
After 2010	97 (69%)	93 (74%)
Region of interest		
Africa	81 (57%)	60 (48%)
America	33 (23%)	18 (14%)
Asia	11 (8%)	40 (32%)
Asia and Europe	-	1 (1%)
Europe	16 (11%)	6 (5%)
Worldwide	-	1 (1%)
Study type		
<i>Studies reporting number of cases and/or CFR</i>	<i>N = 141</i>	<i>N = 88</i>
Descriptive report	141 (100%)	79 (90%)
Retrospective study	-	9 (10%)
<i>Studies reporting results on treatments</i>		<i>N = 38</i>
Case-Control Study	-	2 (5%)
Descriptive report	-	3 (8%)
Non-randomized comparative study	-	6 (16%)
Prospective and retrospective study	-	1 (3%)
Randomized Controlled trial	-	10 (26%)
Retrospective study	-	12 (32%)
Single-arm trial	-	4 (10%)
Disease of interest (alphabetical order)		
Alkhurma hemorrhagic fever (AHF)	1 (1%)	2 (2%)
Argentine hemorrhagic fever (ArHF)	0 (0%)	3 (2%)
Bolivian hemorrhagic fever (BHF)	0 (0%)	1 (1%)
Chapare hemorrhagic fever (CHF)	2 (1%)	0 (0%)
Crimean-Congo hemorrhagic fever (CCHF)	4 (3%)	27 (21%)
Ebola Virus Disease (EVD)	47 (33%)	43 (33%)
Hantavirus Pulmonary Syndrome (HPS)	26 (18%)	13 (10%)
Hemorrhagic fever with renal syndrome (HFRS)	0 (0%)	18 (14%)
Kyasanur Forest Disease (KFD)	0 (0%)	0 (0%)
Lassa fever (LF)	24 (17%)	7 (5%)
Lujo hemorrhagic fever (LHF)	1 (1%)	1 (1%)
Lymphocytic choriomeningitis (LCM)	0 (0%)	0 (0%)
Marburg virus disease (MVD)	13 (9%)	1 (1%)
Omsk hemorrhagic fever (OHF)	0 (0%)	1 (1%)
Rift Valley fever (RVF)	23 (16%)	9 (7%)
Sabia-associated hemorrhagic fever/Sabia hemorrhagic fever (SHF)	0 (0%)	1 (1%)
Tick-borne Encephalitis (from a hemorrhagic variant) (TBE)	0 (0%)	0 (0%)
Venezuelan hemorrhagic fever (VeHF)	0 (0%)	2 (2%)

<https://doi.org/10.1371/journal.pntd.0010889.t001>

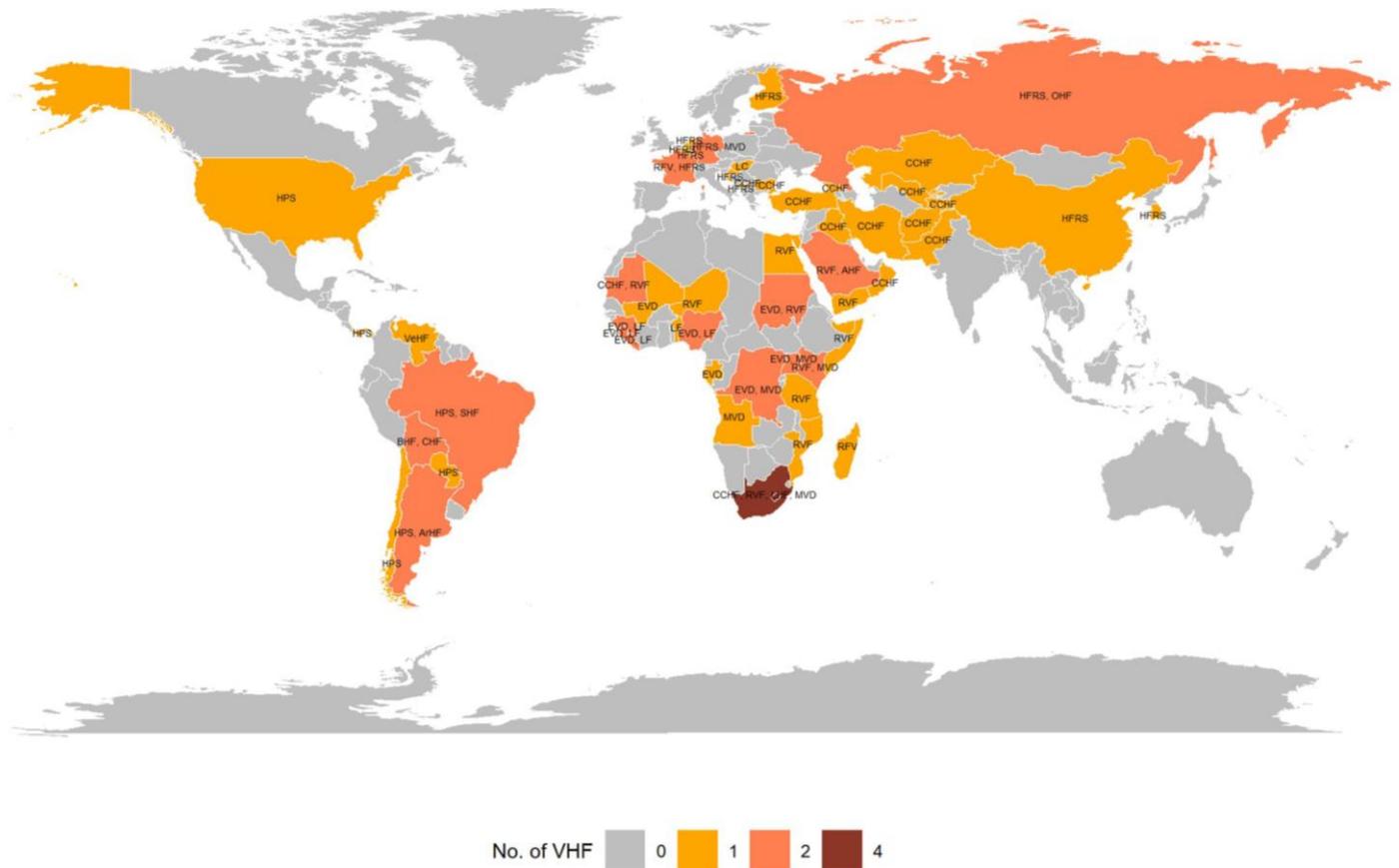


Fig 2. Worldwide distribution of VHF. The color denotes the number of VHF with at least one outbreak reported by country and the acronyms denote the names of the VHFs. Note: AHF, Alkhurma hemorrhagic fever; ArHF, Argentine hemorrhagic fever; BHF, Bolivian hemorrhagic fever; CHF, Chapare hemorrhagic fever; CCHF, Crimean-Congo hemorrhagic fever; EVD, Ebola Virus Disease; HPS, Hantavirus Pulmonary Syndrome; HFRS, Hemorrhagic fever with renal syndrome; LF, Lassa fever; LHF, Lujo hemorrhagic fever; MVD, Marburg virus disease; OHF, Omsk hemorrhagic fever; RVD, Rift Valley fever; SHF, Sabia hemorrhagic fever; VeHF, Venezuelan hemorrhagic fever. Maps were generated with the 'maps' R package using data from the Natural Earth Project (the 1:50m resolution version).

<https://doi.org/10.1371/journal.pntd.0010889.g002>

outbreak took place in 1976 in the Democratic Republic of the Congo (318 cases, CFR = 88%) and Sudan (284 cases, CFR = 53%). At the date of the review, the countries with the highest reported number of cases were Sierra Leone with a total of 14,124 cases (CFR = 28%) and Liberia with 10,678 cases (CFR = 45%), due to the 2014–2016 western Africa outbreak.

The reported CFRs since 2010 varied greatly across countries (cf. [S7 Table](#)). In total, 28% in Sierra Leone in 2014–2016 (14,124 cases), 40% in Nigeria in 2014 (20 cases), 41% in Uganda in 2012 (17 cases), 45% in Liberia (10,678 cases), 66% in the Democratic Republic of the Congo in 2018 (3,470 cases), 67% in Guinea in 2014–2016 (3,811 cases) and 75% in Mali in 2014 (8 cases).

Hantavirus pulmonary syndrome (HPS)

The distribution of HPS cases and CRFs are reported in [Fig 5](#). Cases were reported in America in the following countries: Argentina, Brazil, Chile, Panama, Paraguay and USA. The first documented cases were reported in 1993 in the USA (48 cases, CFR = 56%) and Brazil (884 cases from 1993 to 2006, CFR = 39%). At the date of the review, the countries with the highest documented number of cases were Brazil with 2,370 cases (CFR = 39%) and the last documented cases were reported in 2018 in USA (3 cases, CFR = 67%). More details by country and period are reported in [S8 Table](#).



Fig 3. Distribution of number of cases and CFR of Crimean-Congo hemorrhagic fever. The size of the bubble is proportional with the number of cases reported and the color denotes the level of CFR by country. Note: Maps were generated with the ‘maps’ R package using data from the Natural Earth Project (the 1:50m resolution version).

<https://doi.org/10.1371/journal.pntd.0010889.g003>

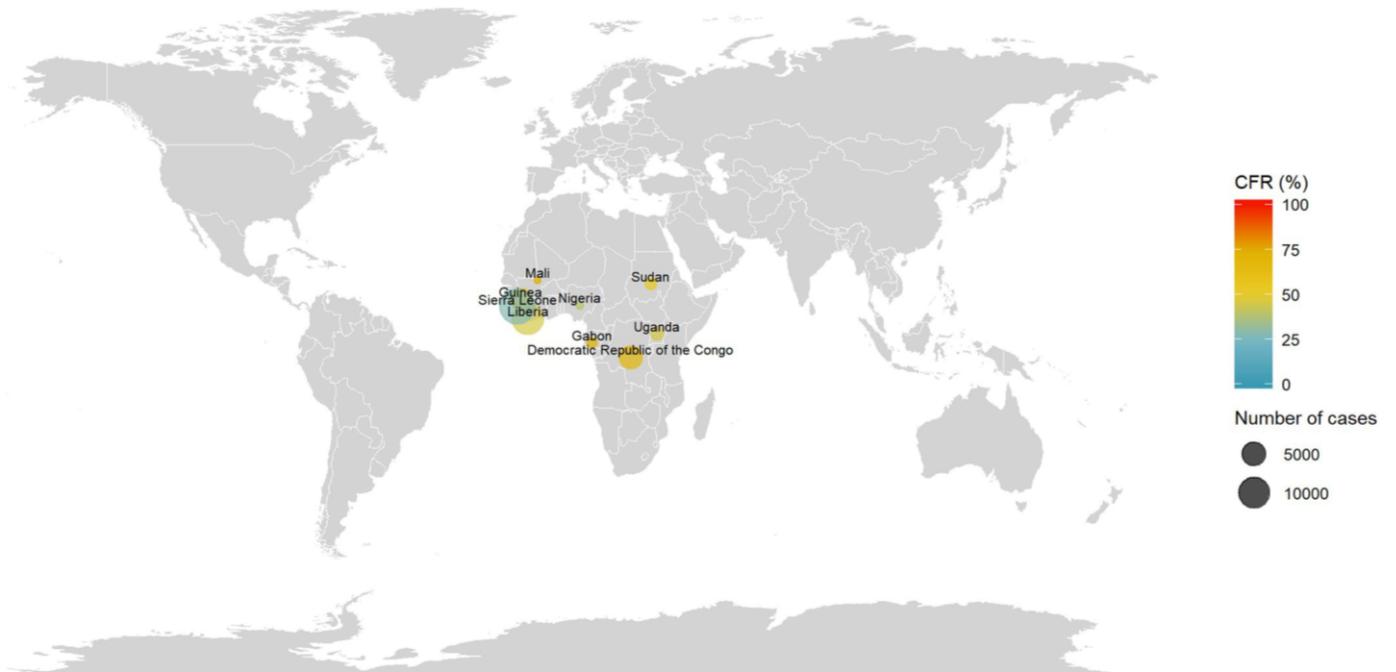


Fig 4. Distribution of number of cases and CFR of Ebola Virus Disease. The size of the bubble is proportional with the number of cases reported and the color denotes the level of CFR by country. Note: Maps were generated with the ‘maps’ R package using data from the Natural Earth Project (the 1:50m resolution version).

<https://doi.org/10.1371/journal.pntd.0010889.g004>



Fig 5. Distribution of number of cases and CFR of Hantavirus Pulmonary Syndrome. The size of the bubble is proportional with the number of cases reported and the color denotes the level of CFR by country. Note: Maps were generated with the 'maps' R package using data from the Natural Earth Project (the 1:50m resolution version).

<https://doi.org/10.1371/journal.pntd.0010889.g005>

Hemorrhagic fever with renal syndrome (HFRS)

HFRS cases were reported in Asia and Europe in the following countries: China, Croatia, Finland, Belgium, France, Germany, Netherlands, Luxembourg, South Korea, Montenegro and Russia. The first documented cases were reported in 1931–1941 in China (10,000 cases, CFR = 30%). At the date of the review, the country with the highest reported number of cases was China with 1,306,812 cases (CFR = 3%) and the last documented cases reported in 2000–2017 in Russia (131,590 cases, CFR = 0.4%). More details by country and period are reported in [S9 Table](#).

Lassa fever (LF)

The distribution of LF cases and CRFs are reported in [Fig 6](#). Cases were reported in western Africa in the following countries: Benin, Liberia, Nigeria, Sierra Leone and Guinea. We identified a study which reported the first outbreak in 1996–1999 in Guinea (22 cases, CFR = 18%). At the date of the review, the country with the highest reported number of cases was Nigeria with 2,287 cases (CFR = 23%) and the last documented cases reported in 2019 in South-West Nigeria (554 cases, CFR = 22%). More details by country and period are reported in [S10 Table](#).

Marburg virus disease (MVD)

MVD cases were reported in Africa and Europe in the following countries: Angola, Democratic Republic of the Congo, Germany, ex-Yugoslavia, South Africa, Kenya and Uganda. The first documented cases were reported in 1967 in Germany (29 cases, CFR = 24%) and ex-Yugoslavia (2 cases, CFR = 0%) and relied to lab.-accidental transmission in settings that were used to import monkeys from Central Africa. At the date of the review, the country with the highest reported number of cases was Angola with 374 cases (CFR = 88%) and the last



Fig 6. Distribution of number of cases and CFR of Lassa fever. The size of the bubble is proportional with the number of cases reported and the color denotes the level of CFR by country. Note: Maps were generated with the 'maps' R package using data from the Natural Earth Project (the 1:50m resolution version).

<https://doi.org/10.1371/journal.pntd.0010889.g006>

documented cases reported in 2017 in Uganda (2 cases, CFR = 100%). More details by country and period are reported in [S11 Table](#).

Rift Valley fever (RVF)

The distribution of RVF cases and CRFs are reported in [Fig 7](#). Cases were reported in Africa and Asia in the following countries: Egypt, Kenya, Madagascar, Mauritania, Mayotte (France), Mozambique, Niger, Saudi Arabia, Somalia, South Africa, Sudan, Tanzania and Yemen. The first documented cases were reported in 1977–1978 in Egypt (18,000 cases, CFR = 3%). At the date of the review, the country with the highest documented number of cases was Egypt with 18,148 cases (CFR = 3%) and the last documented cases reported in 2018–2019 in Mayotte (129 cases, CFR not reported). More details by country and period are reported in [S12 Table](#).

Other VHFs

The remaining VHFs reported cases in one country each: Saudi Arabia for AHF (335 cases, CFR = 2%), Argentina for ArHF (981 cases, CFR not reported), Bolivia for BHF (690 cases, CFR = 23%), South Africa for LHF (5 cases, CFR = 80%), Russia for OHF (1144 cases, CFR = 14%), Brazil for SHF (4 cases, CFR = 2%) and Venezuela for VeHF (728 cases, CFR = 23%). More details by VHF and period are reported in [S13 Table](#).

VHFs treatments

Publications evaluating the efficacy of specific treatments on mortality were identified for Argentine hemorrhagic fever, Crimean-Congo hemorrhagic fever, Ebola Virus Disease, Hantavirus Pulmonary Syndrome, Hemorrhagic fever with renal syndrome and Lassa fever (cf.



Fig 7. Distribution of number of cases and CFR of Rift Valley Fever. The size of the bubble is proportional with the number of cases reported and the color denotes the level of CFR by country. Note: Maps were generated with the 'maps' R package using data from the Natural Earth Project (the 1:50m resolution version).

<https://doi.org/10.1371/journal.pntd.0010889.g007>

Tables 2 and 3). More details on each investigational treatment are reported in S14 Table. Detailed results are reported below for the VHF with at least one study. They are presented by alphabetical order.

Argentine hemorrhagic fever (ArHF)

One study on treatment was identified for ArHF: a randomized controlled trial published in 1979 assessing intravenous immune plasma obtained from convalescent donors.[11] This trial showed a significantly decrease in mortality when immune plasma is given before the ninth day of the disease (CFR = 1.1%) compared with normal plasma obtained from donors without a history of ArHF (CFR = 16.5%).

Crimean-Congo hemorrhagic fever (CCHF)

Nine studies on treatments were identified for CCHF: two randomized controlled trials, two non-randomized comparative studies, two case-control studies, two retrospective and/or prospective studies and one descriptive report. Most of the trials assessed oral or intravenous ribavirin, one trial assessed ribavirin +/- corticosteroids and one trial assessed immune globulins + ribavirin. Ribavirin was associated with inconsistent results across studies.

Two studies reported significant results of ribavirin and ribavirin +/- corticosteroids on mortality. However, those two studies had a relatively weak design (a non-randomized comparative study [19] and a prospective and retrospective study [13]). In the randomized trial comparing ribavirin versus standard therapy alone [18], ribavirin was associated with no positive effect on mortality (CFR = 6.3% versus 5.6%). The other randomized trial was a small study (40 patients, [20]) assessing polyvalent immune globulins + ribavirin which showed with no positive effect of the treatment on mortality compared with ribavirin alone (CFR = 25% versus 11%). Therefore, no strong proof of effectiveness of specific treatment against CCHF was identified.

Table 2. Characteristics of included studies reporting results on treatments by VHF.

Study	Study design	Randomisation	Blinding	Country	Treatment	Comparator
Argentine hemorrhagic fever						
Maiztegui 1979 [11]	Randomized Controlled trial	Yes	Double-blind	Argentina	Immune plasma	Normal plasma obtained from donors without a history of AHF who were residents of Buenos Aires, a city located outside the endemic area of the disease.
Crimean-Congo hemorrhagic fever						
Elaldi 2009 [12]	Non-randomized comparative study	No	NA	Turkey	Ribavirin + Supportive therapy	Supportive therapy (patients who were diagnosed in 2003)
Dokuzoguz 2013 [13]	Prospective and retrospective study	No	NA	Turkey	Ribavirin +/- Corticosteroid	No ribavirin +/- corticosteroid
Cevik 2008 [14]	Descriptive report	No	NA	Turkey	Ribavirin + Supportive therapy	Supportive therapy
Yilmaz 2016 [15]	Retrospective study	No	NA	Turkey and Iran	Ribavirin + Supportive therapy	Supportive therapy
Izadi 2009 [16]	Case-Control Study	No	NA	Iran	Ribavirin + Supportive therapy	NR
Tulek 2012 [17]	Case-Control Study	No	NA	Turkey	Ribavirin + Supportive therapy	Supportive therapy (patients with CCHF of the other infectious diseases department in the same hospital)
Koksal 2010 [18]	Randomized Controlled trial	Yes	NR	Turkey	Ribavirin + Supportive therapy	Supportive therapy
Mardani 2003 [19]	Non-randomized comparative study	No	NA	Iran	Ribavirin	No ribavirin (historical controls)
Salehi 2013 [20]	Randomized Controlled trial	Yes	Single-blind	Iran	Immune globulins (IGIV) + ribavirin + Supportive therapy	Ribavirin + Supportive therapy
Ebola Virus Disease						
Aluisio 2019 [21]	Retrospective study	No	NA	Liberia and Sierra Leone	Intravenous fluid (IVF) + Supportive therapy	NR
Aluisio 2019 [22]; Aluisio 2019 [23]	Retrospective study	No	NA	Liberia and Sierra Leone	Vitamin A supplementation + Supportive therapy	Supportive therapy
Aluisio 2020 [24]	Retrospective study	No	NA	Liberia and Sierra Leone	Cephalosporin + Supportive therapy	Supportive therapy
Bai 2016 [25]	Retrospective study	No	NA	Sierra Leone	Favipiravir + Supportive therapy	Supportive therapy
Sissoko 2016 [26]	Single-arm trial	No	NA	Guinea	Favipiravir + Supportive therapy	Historical data (540 patients hospitalized between 15 September and 15 December 2014, the 3 months preceding the start of the JIKI study)
Yam 2020 [27]	Retrospective study	No	NA	Sierra Leone and Liberia	Multivitamin supplementation + Supportive therapy	Supportive therapy
Dunning 2016 [28]	Single-arm trial	No	NA	Liberia	Brincidofovir + Supportive therapy	NR
Gignoux 2015 [29]	Retrospective study	No	NA	Liberia	• Artemether-Lumefantrine + Supportive therapy • Artesunate-Amodiaquine + Supportive therapy	Supportive therapy
Garbern 2019 [30]	Retrospective study	No	NA	Liberia and Sierra Leone	Artesunate-amodiaquine (ASAQ) + Supportive therapy	Supportive therapy

(Continued)

Table 2. (Continued)

Study	Study design	Randomisation	Blinding	Country	Treatment	Comparator
Dunning 2016 [31]	Single-arm trial	No	NA	Sierra Leone	TKM-130803 + Supportive therapy	NR
Sahr 2017 [32]	Non-randomized comparative study	No	NA	Sierra Leone	Convalescent whole blood (CWB) + Supportive therapy	Supportive therapy
Konde 2017 [33]	Single-arm trial	NA	NA	Guinea	IFN-beta 1a + Supportive therapy	Supportive therapy (Historical cohort). The historical control patients were admitted to the Coyah ETU during the same time period as the IFN beta-1a patients with RT-PCR confirmed blood EBOV. Also included 17 patients who matched the IFN treated patients for eligibility criteria based on <6 days from symptom onset, age, under care in a Guinean treatment centre, who were better matched for baseline CT values.
Sadek 1999 [34]	Descriptive report	NA	NA	Democratic Republic of the Congo	Whole blood transfusion from convalescent patients + Supportive therapy	NA
Van Griensven 2016 [35]	Non-randomized comparative study	No	NA	Guinea	Transfusion of convalescent plasma + Supportive therapy	Supportive therapy: Patients who had been admitted to the ETU during the preparatory period of the study while the system for apheresis and pathogen reduction was being set up and those for whom ABO compatible convalescent plasma was not available during the study. At the start of recruitment, there was a sufficient amount of convalescent plasma available to treat all the patients, so a protocol amendment was approved for the control group to consist of patients who were treated at the same ETU before the initiation of the trial.
Davey 2016 [36]	Randomized Controlled trial	Yes	No	Liberia, Sierra Leone, Guinea and USA	ZMapp + Supportive therapy	Supportive therapy
Mulangu 2019 [37]	Randomized Controlled trial	Yes	NA	Democratic Republic of the Congo	<ul style="list-style-type: none"> •Remdesivir + Supportive therapy •Mab114 + Supportive therapy •REGN-EB3 + Supportive therapy 	ZMapp (intravenous 50mg per kilogram of body weight every third day beginning on day 1 for a total of 3 doses) + Supportive therapy
Kerber 2019 [38]	Retrospective study	NA	NA	Guinea	Favipiravir + Supportive therapy	Supportive therapy
Hantavirus Pulmonary Syndrome						
Chapman 1999 [39]	Non-randomized comparative study	No	Open-label	United States	Ribavirin	No ribavirin (contemporaneous patients)
Mertz 2004 [40]	Randomized Controlled trial	Yes	Double-blind	USA and Canada	Ribavirin	Placebo
Vial 2013 [41]	Randomized Controlled trial	Yes	Double-blind	Chile	Methylprednisolone	Placebo
Wernly 2011 [42]	Retrospective study	NA	NA	USA	Extracorporeal membrane oxygenation (ECMO) support	Patients intubated when they became hypoxic and placed on ECMO when they became hemodynamically unstable
Vial 2015 [43]	Non-randomized comparative study	No	NA	Chile	Immune plasma infusion	No immune plasma infusion
Hemorrhagic fever with renal syndrome						

(Continued)

Table 2. (Continued)

Study	Study design	Randomisation	Blinding	Country	Treatment	Comparator
Gui 1987 [44]	Randomized Controlled trial	Yes	Single-blind	China	Recombinant interferon α -2 + Supportive therapy	Placebo + Supportive therapy
Du 2013 [45]	Retrospective study	No	NA	China	Renal Replacement Therapy (RRT)	No RRT
Huggins 1991 [46]	Randomized Controlled trial	Yes	Double-blind	China	Ribavirin	Placebo
Lassa fever						
Ilori 2019 [47]	Descriptive report	No	NA	Nigeria	Ribavirin + Supportive therapy	Supportive therapy
McCormick 1986 [48]	Randomized Controlled trial	Yes	No	Sierra Leone	PHASE 1: No therapy VS Oral Ribavirin VS Lassa convalescent plasma PHASE 2: intravenous ribavirin +/- lassa plasma	No therapy

NA, Not applicable; NR, Not reported

<https://doi.org/10.1371/journal.pntd.0010889.t002>

Ebola virus disease (EVD)

Seventeen studies on treatments were identified for EVD: two randomized controlled trials, two non-randomized comparative studies, four single-arm trials, eight retrospective studies and one descriptive report. Among these studies, four reported positive results on mortality. The two first studies showed that early vitamin A supplementation or IFN-beta 1a therapy may be associated with reduced mortality compared with no vitamin A supplementation or historical control (Relative risk reduction of mortality with vitamin A supplementation within 48h = 0.77 [0.59 to 0.99]; 21-day survival based on Kaplan Meier for IFN-beta 1a therapy versus historical control: 67% versus 19%). However, those results are to be taken with caution based on the weak design of the studies (retrospective study [22,23] and single-arm trial [33]).

The third study was a retrospective study assessing favipiravir [25] and reported that the treatment was associated with prolonged survival compared with standard therapy alone (CFR = 44% versus 65%). However, this result was challenged by another retrospective study (adjusted odds ratio = 0.48 [0.20 to 1.01]) [38] and a single-arm trial (CFR = 54% versus 55%), [26] which reported no significant results of favipiravir on mortality.

The last positive study was a randomized controlled trial assessing three treatments, the antiviral remdesivir, and the antibody-based therapies Mab114 and REGN-EB3 against ZMapp. [37] The study showed that both MAb114 and REGN-EB3 were superior to ZMapp (difference between MAb114 and ZMapp = -14.6% [-27.2 to -1.7]; difference between REGN-EB3 and ZMapp = -17.8% [-28.9 to -2.9]; in reducing mortality from EVD (with stringent findings among patients presenting with high levels of viral load).

Hantavirus pulmonary syndrome (HPS)

Five studies on treatments were identified for HPS: two randomized controlled trials, two non-randomized comparative studies, and one retrospective study. Only two studies reported weakly significant results. The first study was a retrospective study evaluating the impact on survival of extracorporeal membrane oxygenation (ECMO) support in patients with HPS refractory to medical treatment and the associated complications. [42] This study reported a lower mortality in patients who had elective insertion of vascular sheaths and were almost concurrently intubated and placed on ECMO when they decompensated (CFR = 20%) compared with patients intubated when they became hypoxic and placed on ECMO when they became

Table 3. Mortality results associated with VHF treatments reported in the included studies.

Study	Comparison	Treatment group—No.	Comparator group—No.	Time	Treatment group—Mortality rate (%)	Comparator group—Mortality rate (%)	p-value	Outcome measure	Outcome result	P-value
Argentine hemorrhagic fever										
Maiztegui 1979 [11]	Immune plasma VS normal plasma	91	97	NR	1.1	16.5	<0,01	NR	NR	NR
Crimean-Congo hemorrhagic fever										
Elaldi 2009 [12]	Ribavirin VS no ribavirin	126	92	First 8 days	7.1	11.9	0.283	Hazard ratio	10.33 (1.65–64.71)	NR
				After 8 days (to 30 days)	-	-	-	Hazard ratio	0.1 (0.01–8.80)	NR
Dokuzoguz 2013 [13]	Ribavirin +/- corticosteroids VS no treatment	Ribavirin +/- corticosteroids: 235	46	NR	7.7	10.9	NR	Adjusted odds ratio	0.04 (0.004–0.48)	0.01
	Corticosteroids + Ribavirin VS no treatment	Corticosteroids + Ribavirin: 44	217	NR	20.5	4.1	NR	Adjusted odds ratio	0.22 (0.039–1.27)	0.092
Cevik 2008 [14]	Ribavirin VS no ribavirin	9	16	NR	55.5	43.7	0.571	NR	NR	NR
Yilmaz 2016 [15]	Ribavirin VS no ribavirin	198	345	NR	7.6	8.4	0.733	NR	NR	NR
Izadi 2009 [16]	Patients receiving ribavirin after onset of bleeding VS before onset of bleeding or had no bleeding at all	39	24	NR	35.9	8.3	NR	Odds ratio	6.2 (1.3–30.3)	0.018
	Patients receiving ribavirin after the 4th day of disease onset VS within the first 4 days of disease onset	25	38	NR	40.0	15.8	NR	Odds ratio	3.6 (1.1–11.6)	0.031
Tulek 2012 [17]	Ribavirin VS no ribavirin	91	152	NR	1.1	5.3	0.096	NR	NR	NR
Koksal 2010 [18]	Ribavirin VS supportive therapy	64	72	NR	6.3	5.6	0.86	NR	NR	NR
Mardani 2003 [19]	Ribavirin VS historical control	69	12	NR	11.6	58.3	NR	Relative risk in confirmed patients	0.20 (0.09–0.45)	<0.001
Salehi 2013 [20]	Immune globulins + ribavirin VS ribavirin	12	28	Until one week after hospitalisation	25	11	0.24	NR	NR	NR
Ebola Virus Disease										
Aluisio 2019 [21]	Intravenous fluid (IVF)	70	354	28 days	58.7	55.1	0.583	NR	NR	NR
Aluisio 2019 [22]; Aluisio 2019 [23]	Vitamin A supplementation VS no vitamin A supplementation	330	94	During Ebola Treatment Unit care	55	71.9	NR	Relative risk reduction of mortality with vitamin A treatment within 48 h	0.77 (0.59–0.99)	0.041
Aluisio 2020 [24]	Cephalosporin VS no treatment	360	64	48h	58.7	55.1	0.583	NR	NR	NR
Bai 2016 [25]	Favipiravir VS supportive care	39	85	60 days	43.6	64.7	0.027	NR	NR	NR

(Continued)

Table 3. (Continued)

Study	Comparison	Treatment group—No.	Comparator group—No.	Time	Treatment group—Mortality rate (%)	Comparator group—Mortality rate (%)	p-value	Outcome measure	Outcome result	P-value
Sissoko 2016 [26]	Favipiravir VS historical control	111	NA	On-trial	54.05	55	NR	NR	NR	NR
Yam 2020 [27]	Multivitamin supplementation VS no multivitamin supplementation	261	163	48h	53.6	63.8	NR	NR	NR	NR
Dunning 2016 [28]	Brincidofovir	4	NA	14 days	100	NA	NR	NR	NR	NR
Gignoux 2015 [29]	No Antimalarial Drug Prescription VS Artemether-Lumefantrine	194	63	NR	64.4	65.1	NR	Unadjusted risk ratio	1.01 (0.82–1.25)	0.92
	Artesunate-Amodiaquine VS Artemether-Lumefantrine	71	194	NR	50.7	64.4	NR	Unadjusted risk ratio	0.79 (0.61–1.01)	0.06
Garbern 2019 [30]	Artesunate-amodiaquine VS no treatment	22	402	NR	45.5	72.7	NR	Risk of death in the matched cohort	0.63 (0.37–1.07)	0.086
Dunning 2016 [31]	TKM-130803	14	1820 (individual-level data from patients with PCR-confirmed Ebola infection from the 2014–2015 outbreak)	14 days	79	55	NR	NR	NR	NR
Sahr 2017 [32]	Convalescent whole blood VS routine care	44	25	NR	27.9	44	NR	Odds ratio for survival	2.3 (0.8–6.5)	
Konde 2017 [33]	IFN-beta 1a VS historical control	9	38	21 days	33.3	81	NR	21-day survival based on Kaplan Meier	67% in treated patients VS 19% for the control	0.026
Sadek 1999 [34]	Whole blood transfusion from convalescent patients	8	NA	NR	12.5	NR	NR	NR	NR	NR
Van Griensven 2016 [35]	Convalescent plasma VS no treatment	84	418	16 days	31	38	NR	Risk difference	-7% (-18 to 4)	NR
Davey 2016 [36]	ZMapp VS standard of care	35	36	28 days	22	37	NR	Absolute difference	-15% (-37 to 7)	NR
Mulangu 2019 [37]	•Remdesivir VS ZMapp	175	169	28 days	53.1	49.7	NR	Difference between groups	3.4% (-7.2 to 14)	NR
	•Mab114 VS ZMapp	174	169	28 days	35.1	49.7	NR	Difference between groups	-14.6% (-27.2 to -1.7)	0.007

(Continued)

Table 3. (Continued)

Study	Comparison	Treatment group—No.	Comparator group—No.	Time	Treatment group—Mortality rate (%)	Comparator group—Mortality rate (%)	p-value	Outcome measure	Outcome result	P-value
	•REGN-EB3 VS ZMapp	155	169	28 days	33.5	49.7	NR	Difference between groups	-17.8% (-28.9 to -2.9)	0.002
Kerber 2019 [38]	Favipiravir VS no treatment	72	90	NR	42.5	57.8	0.053	Adjusted odds ratio	0.48 (0.20–1.01)	0.11
Hantavirus Pulmonary Syndrome										
Chapman 1999 [39]	Ribavirin VS no treatment	30	34	1 year	46.7	50	NR	NR	NR	NR
Mertz 2004 [40]	Ribavirin VS placebo	10	13	28 days	20	15	1	NR	NR	NR
Vial 2013 [41]	Methylprednisolone VS placebo	30	30	28 days	27	40	0.41	Relative risk	0.67 (0.32–1.39)	NR
Wernly 2011 [42]	Patients who had elective insertion of vascular sheaths and were almost concurrently intubated and placed on ECMO when they decompensated VS patients intubated when they became hypoxic and placed on ECMO when they became hemodynamically unstable.	25	26	NR	20	46	0.048	NR	NR	NR
Vial 2015 [43]	Immune plasma infusion VS no treatment	29	199	30 days	14	32	0.049	Odds ratio	0.35 (0.12–0.99)	NR
Hemorrhagic fever with renal syndrome										
Gui 1987 [44]	Recombinant interferon α -2 VS placebo	25	25	On-trial	16	16	NS	NR	NR	NR
Du 2013 [45]	RRT VS no RRT	Total (both groups): 77	NA	NR	34.3	70	0.031	NR	NR	NR
Huggins 1991 [46]	Ribavirin VS placebo	126	117	On-trial	2.4	8.55	0.01	NR	NR	NR
Lassa fever										
Ilori 2019 [47]	Ribavirin VS no treatment	334	21	During the outbreak	20.7	71.4	<0.001	NR	NR	NR
McCormick 1986 [48]	Group of AST level>150: IV ribavirin VS no therapy	63	60	NR	19	55	0.00003	NR	NR	NR
	Group of AST level>150: Oral ribavirin VS no therapy	14	60	NR	14	55	0.006	NR	NR	NR
	Group of AST level>150: Plasma VS no therapy	28	60	NR	50	55	0.3	NR	NR	NR

(Continued)

Table 3. (Continued)

Study	Comparison	Treatment group—No.	Comparator group—No.	Time	Treatment group—Mortality rate (%)	Comparator group—Mortality rate (%)	p-value	Outcome measure	Outcome result	P-value
	Group of patients with virus level superior $10^{3.6}$ TCID ₅₀ /mL: IV ribavirin VS no therapy	31	46	NR	32	76	0.00015	NR	NR	NR
	Group of patients with virus level superior $10^{3.6}$ TCID ₅₀ /mL: Oral ribavirin VS no therapy	10	46	NR	30	76	0.008	NR	NR	NR
	Group of patients with virus level superior $10^{3.6}$ TCID ₅₀ /mL: Plasma VS no therapy	21	46	NR	57	76	0.12	NR	NR	NR
	Group of patients with virus level inferior $10^{3.6}$ TCID ₅₀ /mL: IV ribavirin VS no therapy	32	111	NR	9	28	0.02	NR	NR	NR
	Group of patients with virus level inferior $10^{3.6}$ TCID ₅₀ /mL: Oral ribavirin VS no therapy	29	111	NR	7	28	0.01	NR	NR	NR
	Group of patients with virus level inferior $10^{3.6}$ TCID ₅₀ /mL: Plasma VS no therapy	32	111	NR	16	28	0.12	NR	NR	NR

NA, Not applicable; NR, Not reported

<https://doi.org/10.1371/journal.pntd.0010889.t003>

hemodynamically unstable (CFR = 46%). The second study was a non-randomized comparative study, which compared immune plasma infusion versus no treatment. [43] The study reported a weakly significant decrease in mortality associated with immune plasma infusion (Odds ratio = 0.35 [0.12 to 0.99]).

Therefore, no strong proof of effectiveness of treatment against HPS was identified.

Hemorrhagic fever with renal syndrome (HFRS)

Three studies on treatments were identified for HFRS: two randomized controlled trials and one retrospective study. Only two studies reported positive results. The first study was a retrospective study comparing Renal Replacement Therapy (RRT) versus no RRT [45] and showed that RRT is associated with a decrease in mortality (CFR = 34% versus 70%). However, those results are to be taken with caution based on the weak design of the study. The second study was a randomized controlled trial assessing ribavirin [46] and showed a significant reduction in mortality among patients treated with ribavirin (CFR = 2% versus 9%).

Lassa fever (LF)

Two studies on treatments were identified for LF: a randomized controlled trial and a descriptive report. The first study published in 1986 was a randomized controlled trial and evaluated

ribavirin and convalescent plasma compared with no therapy in several subgroups [48]. The study showed that ribavirin was associated with a significantly lower mortality than no therapy (CFR = 21% versus 71%). Moreover, the second study which was a descriptive report published in 2019, [47] also reported that ribavirin was associated with a decrease in mortality.

Discussion

To our knowledge, this is the first comprehensive systematic review to summarize all published information available on worldwide cases numbers, mortality and treatments of a range of VHFs, excluding severe dengue and yellow fever.

Only a few number VHF systematic reviews were previously published and focused on single VHF, mainly EVD, CCHF and LF. Some meta-analyses were also conducted but often associated with heterogeneity issues. A previous meta-analysis on EVD found a pooled CFR of 60% in Africa [49]. However, this result was associated with a very high level of heterogeneity. This is consistent with our findings, which showed that CFR varied greatly across countries in outbreaks since 2010 ranging from 28% (2014–2016 outbreak in Sierra Leone) to 75% (2014 in Mali). In terms of specific treatments, a recent systematic review [50] focused mainly on the randomized controlled trial assessing remdesivir, Mab114 and REGN-EB3 against ZMapp. [37] The authors concluded as well that both MAb114 and REGN-EB3 were superior to ZMapp in reducing mortality from EVD with differences depending on the viral load at baseline.

Regarding CCHF, we identified in our review no strong proof of effectiveness of treatment. A previous meta-analysis on the efficacy of ribavirin in CCHF patients showed that ribavirin decreased the mortality rate compared with patients not treated with ribavirin.[51] However, this meta-analysis included an important number of low-quality studies such as case series. Therefore, the results should be considered with caution. Regarding Lassa fever, a previous meta-analysis on the efficacy of ribavirin showed that ribavirin was associated with lower risk of death than patients not treated with ribavirin.[52] However, heterogeneity was identified across studies and the results are mainly based on retrospective studies.

Our systematic reviews also has some limitations. One limitation of our review on the number of cases and CFR identified is that our findings are based on numbers registered on the WHO and CDC websites or published numbers, which can underestimate the reality. We also decided, when available, to prioritize laboratory confirmed cases numbers. Moreover, we restricted our review of the grey literature to national health websites and references reported on the WHO and CDC websites; data on clinical trials registries (e.g. ClinicalTrials.gov) were not included here. It is also important to mention that we did not stratify our results according to species or the strain of the virus. For example, in the case of Ebola disease investigational therapeutic options, the trials of importance were conducted during the 2014–2016 western Africa outbreak and the 2018–2020 North-Kivu (Democratic Republic of the Congo) outbreak that were related to Ebola virus (species Zaire Ebolavirus) and Makona and Kikwit specific strain respectively.[36,37] With respect to treatment of other ebolavirus diseases (e.g. Sudan and Bundibugyo virus diseases), options are even more limited. Notably, the three monoclonal antibody treatments tested in the PALM trial [37], ZMapp, REGN-EB3, and MAb114, have a narrow spectrum and are ineffective against other filovirus infections. In the case of other VHFs such as HPS or LF, the sparse comparative trials assessed mainly a nucleoside inhibitor (i.e., ribavirin). No strong proof of effectiveness of ribavirin was identified for HPS. Regarding LF, the efficacy of ribavirin was not considered as specific to Lassa virus lineage. Besides, considerable uncertainty was recently even more raised about its activity as an anti-infectious agent in the management of the condition.[52–54] Another limitation of our review is that we focused our results on the mortality of the VHF. However, the mortality does not account for

transmissibility of the virus, contagiousness and immune escape. Mortality rates varied also greatly across countries, especially in Africa, which can be explained by the fact that treatment facilities in some places in Africa may be limited. The study context also has an impact on the results of our review. Some authors report the challenge they can face of studying a relatively rare disease that affect widely dispersed rural areas.[41] Regarding VHFs such as EVD, authors reported difficulties during the course of their study, with for example interruption of participating centers due to violence from local community or paramilitary groups who can be suspicious of the activities in those facilities.[37] Moreover, a large number of included studies on treatment evaluation in our review were associated with a high risk of bias. This highlights the need to conduct clinical trials with a methodologically strong design.

This is of most importance to adapt the methodology of clinical trials in the specific context of VHFs. A recent study explored the application of Bayesian Decision Analysis (BDA) in order to incorporate the burden of disease and disease context into clinical trials, especially for the deadliest diseases in the US, such as cancers or liver cirrhosis. [55,56] This framework allows taking into account the disease context when determining the sample size and critical value of a fixed-sample test. Therefore, an interesting next step of our review would be to use these results to conduct a BDA to evaluate the optimal sample sizes and type I errors for future VHF clinical trials.

Conclusion

We observed that number of cases and mortality varied greatly across VHFs as well as across countries within each VHF. The number of studies on VHFs treatments was very limited with very few randomized trials and no strong proof of effectiveness of treatment against most of the VHFs. Therefore, there is a high need of methodologically strong clinical trials conducted in the context of VHF.

Supporting information

S1 Fig. MEDLINE Search strategy.

(DOCX)

S1 Table. References details of each included publication.

(DOCX)

S2 Table. References details of each included WHO or CDC report.

(DOCX)

S3 Table. Quality Assessment of studies reporting results on treatments–Randomized Trials.

(DOCX)

S4 Table. Quality Assessment of studies reporting results on treatments–Non-Randomized Trials

(DOCX)

S5 Table. Quality Assessment of studies reporting number of cases and/or case fatality rates

(DOCX)

S6 Table. Number of cases and CFRs of Crimean-Congo hemorrhagic fever by country (alphabetical order) and period

(DOCX)

S7 Table. Number of cases and CFRs of Ebola Virus Disease by country and period
(DOCX)

S8 Table. Number of cases and CFRs of Hantavirus Pulmonary Syndrome by country and period
(DOCX)

S9 Table. Number of cases and CFRs of Hemorrhagic fever with renal syndrome by country and period
(DOCX)

S10 Table. Number of cases and CFRs of Lassa fever by country and period
(DOCX)

S11 Table. Number of cases and CFRs of Marburg virus disease by country and period
(DOCX)

S12 Table. Number of cases and CFRs of Rift Valley fever by country and period
(DOCX)

S13 Table. Number of cases and CFRs of Alkhurma hemorrhagic fever, Argentine hemorrhagic fever, Bolivian hemorrhagic fever, Chapare hemorrhagic fever, Lujo hemorrhagic fever, Omsk hemorrhagic fever, Sabia hemorrhagic fever and Venezuelan hemorrhagic fever by country and period
(DOCX)

S14 Table. Description of the treatment and comparator groups of each included studies reporting results on VHF treatments
(DOCX)

Author Contributions

Conceptualization: Drifa Belhadi, France Mentré, Cédric Laouénan.

Data curation: Drifa Belhadi.

Formal analysis: Drifa Belhadi, Majda El Baied, Guillaume Mulier.

Methodology: Drifa Belhadi.

Project administration: Drifa Belhadi, France Mentré, Cédric Laouénan.

Supervision: France Mentré, Cédric Laouénan.

Validation: Cédric Laouénan.

Visualization: Drifa Belhadi.

Writing – original draft: Drifa Belhadi.

Writing – review & editing: Drifa Belhadi, Majda El Baied, Guillaume Mulier, Denis Malvy, France Mentré, Cédric Laouénan.

References

1. McFee RB. Viral hemorrhagic fever viruses. *Dis Mon.* 2013; 59(12):410–25. <https://doi.org/10.1016/j.disamonth.2013.10.003> PMID: 24314803
2. CDC. Diseases Caused by High-Consequence Viruses [Internet]. [cited 2022 Jun 1]. Available from: <https://www.cdc.gov/ncezid/dhcpp/vspb/diseases.html>

3. Cleri DJ, Ricketti AJ, Porwancher RB, Ramos-Bonner LS, Vernaleo JR. Viral hemorrhagic fevers: current status of endemic disease and strategies for control. *Infect Dis Clin*. 2006; 20(2):359–93.
4. Ippolito G, Feldmann H, Lanini S, Vairo F, Di Caro A, Capobianchi MR, et al. Viral hemorrhagic fevers: advancing the level of treatment. *BMC Med*. 2012; 10(1):1–8. <https://doi.org/10.1186/1741-7015-10-31> PMID: 22458265
5. Malvy D, McElroy AK, de Clerck H, Günther S, van Griensven J. Ebola virus disease. *The Lancet*. 2019; 393(10174):936–48.
6. Iannetta M, Di Caro A, Nicastrì E, Vairo F, Masanja H, Kobinger G, et al. Viral hemorrhagic fevers other than Ebola and Lassa. *Infect Dis Clin*. 2019; 33(4):977–1002. <https://doi.org/10.1016/j.idc.2019.08.003> PMID: 31668201
7. Lee JS, Adhikari NK, Kwon HY, Teo K, Siemieniuk R, Lamontagne F, et al. Anti-Ebola therapy for patients with Ebola virus disease: a systematic review. *BMC Infect Dis*. 2019; 19(1):1–11.
8. CDC. Diseases Caused by High-Consequence Viruses [Internet]. [cited 2019 Dec 1]. Available from: <https://www.cdc.gov/ncezid/dhcpp/vspb/diseases.html>
9. Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ott Ott Hosp Res Inst*. 2011; 2(1):1–12.
10. Higgins JP, Savović J, Page MJ, Elbers RG, Sterne JA. Assessing risk of bias in a randomized trial. *Cochrane Handb Syst Rev Interv*. 2019;205–28.
11. Maiztegui JI, Fernandez NJ, de Damilano AJ. Efficacy of immune plasma in treatment of Argentine haemorrhagic fever and association between treatment and a late neurological syndrome. *Lancet Lond Engl*. 1979 Dec 8; 2(8154):1216–7. [https://doi.org/10.1016/s0140-6736\(79\)92335-3](https://doi.org/10.1016/s0140-6736(79)92335-3) PMID: 92624
12. Elaldi N, Bodur H, Ascioğlu S, Celikbas A, Ozkurt Z, Vahaboglu H, et al. Efficacy of oral ribavirin treatment in Crimean-Congo haemorrhagic fever: a quasi-experimental study from Turkey. *J Infect*. 2009 Mar; 58(3):238–44. <https://doi.org/10.1016/j.jinf.2009.01.014> PMID: 19246100
13. Dokuzoguz B, Celikbas AK, Gök ŞE, Baykam N, Eroglu MN, Ergönül Ö. Severity scoring index for Crimean-Congo hemorrhagic fever and the impact of ribavirin and corticosteroids on fatality. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2013 Nov; 57(9):1270–4. <https://doi.org/10.1093/cid/cit527> PMID: 23946218
14. Cevik MA, Elaldi N, Akinci E, Ongürü P, Erbay A, Buzgan T, et al. A preliminary study to evaluate the effect of intravenous ribavirin treatment on survival rates in Crimean-Congo hemorrhagic fever. *J Infect*. 2008 Oct; 57(4):350–1. <https://doi.org/10.1016/j.jinf.2008.07.007> PMID: 18722669
15. Yilmaz G., Sunbul M., Yapar D., Baykam N., Hasanoglu I., Guner R., et al. Ribavirin in treatment of crimean-congo hemorrhagic fever (CCHF): An international multicenter retrospective analysis. *Open Forum Infect Dis*. 2016; 3.
16. Izadi S, Salehi M, S. I. M. S. Evaluation of the efficacy of ribavirin therapy on survival of Crimean-Congo hemorrhagic fever patients: A case-control study. *Jpn J Infect Dis*. 2009 Jan; 62(1):11–5.
17. Tulek N., Ozturk B., Bulut C., Tuncer Ertem G., Erdinc F.S., Altun S., et al. The evaluation of ribavirin use in patients with Crimean-Congo haemorrhagic fever. *Clin Microbiol Infect*. 2012;18((Tulek N.; Ozturk B.; Bulut C.; Tuncer Ertem G.; Erdinc F.S.; Altun S.; Kinikli S.; Demiroz A.P.) Ankara, Turkey):579–80.
18. Koksali I, Yilmaz G, Aksoy F, Aydin H, Yavuz I, Iskender S, et al. The efficacy of ribavirin in the treatment of Crimean-Congo hemorrhagic fever in Eastern Black Sea region in Turkey. *J Clin Virol Off Publ Pan Am Soc Clin Virol*. 2010 Jan; 47(1):65–8. <https://doi.org/10.1016/j.jcv.2009.11.007> PMID: 19962342
19. Mardani M, Jahromi MK, Naieni KH, Zeinali M. The efficacy of oral ribavirin in the treatment of crimean-congo hemorrhagic fever in Iran. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2003 Jun 15; 36(12):1613–8. <https://doi.org/10.1086/375058> PMID: 12802764
20. Salehi H, Salehi M, Adibi N, Salehi M. Comparative study between Ribavirin and Ribavirin plus Intravenous Immunoglobulin against Crimean Congo hemorrhagic fever. *J Res Med Sci Off J Isfahan Univ Med Sci*. 2013 Jun; 18(6):497–500. PMID: 24250699
21. Aluisio A.R., Yam D., Peters J.L., Cho D.K., Perera S.M., Kennedy S.B., et al. Impact of Intravenous Fluid Therapy on Survival Among Patients with Ebola Virus Disease: An International Multisite Retrospective Cohort Study. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2019;ciz344.
22. Aluisio AR, Perera SM, Yam D, Garbern S, Peters JL, Abel L, et al. Vitamin A Supplementation Was Associated with Reduced Mortality in Patients with Ebola Virus Disease during the West African Outbreak. *J Nutr*. 2019 Oct 1; 149(10):1757–65. <https://doi.org/10.1093/jn/nxz142> PMID: 31268140
23. Aluisio A.R., Yam D., Peters J., Cho D., Perera S., Kennedy S., et al. Association between vitamin a supplementation and mortality in Ebola virus disease: A multisite cohort study. *Acad Emerg Med*. 2019; 26((Aluisio A.R.) Department of Emergency Medicine, Alpert Medical School, Brown University, Providence, United States):S67.

24. Aluisio A.R., Perera S.M., Yam D., Garbern S., Peters J.L., Abel L., et al. Association Between Treatment with Oral Third-Generation Cephalosporin Antibiotics and Mortality Outcomes in Ebola Virus Disease: A Multinational Retrospective Cohort Study. *Trop Med Int Health TM IH*. 2020; <https://doi.org/10.1111/tmi.13369> PMID: 31912627
25. Bai C.-Q., Mu J.-S., Kargbo D., Song Y.-B., Niu W.-K., Nie W.-M., et al. Clinical and Virological Characteristics of Ebola Virus Disease Patients Treated with Favipiravir (T-705)—Sierra Leone, 2014. *Clin Infect Dis*. 2016; 63(10):1288–94. <https://doi.org/10.1093/cid/ciw571> PMID: 27553371
26. Sissoko D, Laouenan C, Folkesson E, M'Lebing AB, Beavogui AH, Baize S, et al. Experimental Treatment with Favipiravir for Ebola Virus Disease (the JIKI Trial): A Historically Controlled, Single-Arm Proof-of-Concept Trial in Guinea. *PLoS Med*. 2016 Mar 1; 13(3):e1001967–e1001967. <https://doi.org/10.1371/journal.pmed.1001967> PMID: 26930627
27. Yam D., Aluisio A.R., Perera S.M., Peters J.L., Cho D.K., Kennedy S.B., et al. Association between multivitamin supplementation and mortality among patients with Ebola virus disease: An international multi-site cohort study. *Afr J Emerg Med*. 2020;23–9. <https://doi.org/10.1016/j.afjem.2019.11.001> PMID: 32161708
28. Dunning J, Kennedy SB, Antierens A, Whitehead J, Ciglenecki I, Carson G, et al. Experimental Treatment of Ebola Virus Disease with Brincidofovir. *PLoS One*. 2016 Sep 9; 11(9):e0162199–e0162199. <https://doi.org/10.1371/journal.pone.0162199> PMID: 27611077
29. Gignoux E., Azman A.S., Ciglenecki I. Artesunate-amodiaquine is associated with reduced ebola mortality. *Am J Trop Med Hyg*. 2015; 93(4):446.
30. Garbern SC, Yam D, Aluisio AR, Cho DK, Kennedy SB, Massaquoi M, et al. Effect of Mass Artesunate-Amodiaquine Distribution on Mortality of Patients With Ebola Virus Disease During West African Outbreak. *Open Forum Infect Dis*. 2019 May 24; 6(7):ofz250–ofz250. <https://doi.org/10.1093/ofid/ofz250> PMID: 31281856
31. Dunning J, Sahr F, Rojek A, Gannon F, Carson G, Idriss B, et al. Experimental Treatment of Ebola Virus Disease with TKM-130803: A Single-Arm Phase 2 Clinical Trial. *PLoS Med*. 2016 Apr 19; 13(4):e1001997–e1001997. <https://doi.org/10.1371/journal.pmed.1001997> PMID: 27093560
32. Sahr F, Ansumana R, Massaquoi TA, Idriss BR, Sesay FR, Lamin JM, et al. Evaluation of convalescent whole blood for treating Ebola Virus Disease in Freetown, Sierra Leone. *J Infect*. 2017 Mar; 74(3):302–9. <https://doi.org/10.1016/j.jinf.2016.11.009> PMID: 27867062
33. Konde MK, Baker DP, Traore FA, Sow MS, Camara A, Barry AAA, et al. Interferon β -1a for the treatment of Ebola virus disease: A historically controlled, single-arm proof-of-concept trial. *PLoS ONE*. 2017 Feb; 12(2):e0169255–e0169255. <https://doi.org/10.1371/journal.pone.0169255> PMID: 28225767
34. Sadek RF, Khan AS, Stevens G, Peters CJ, Ksiazek TG. Ebola hemorrhagic fever, Democratic Republic of the Congo, 1995: determinants of survival. *J Infect Dis*. 1999 Feb; 179 Suppl 1:S24–7. <https://doi.org/10.1086/514311> PMID: 9988161
35. van Griensven J, Edwards T, de Lamballerie X, Semple MG, Gallian P, Baize S, et al. Evaluation of Convalescent Plasma for Ebola Virus Disease in Guinea. *N Engl J Med*. 2016 Jan 7; 374(1):33–42. <https://doi.org/10.1056/NEJMoa1511812> PMID: 26735992
36. Davey R, Dodd L, Proschan M, Neaton J, Neuhaus Nordwall J, Koopmeiners J, et al. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J Med*. 2016; 375(15):1448–1456. <https://doi.org/10.1056/NEJMoa1604330> PMID: 27732819
37. Mulangu S, Dodd LE, Davey RT Jr, Tshiani Mbaya O, Proschan M, Mukadi D, et al. A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *N Engl J Med*. 2019 Dec 12; 381(24):2293–303. <https://doi.org/10.1056/NEJMoa1910993> PMID: 31774950
38. Kerber R, Lorenz E, Duraffour S, Sissoko D, Rudolf M, Jaeger A, et al. Laboratory Findings, Compassionate Use of Favipiravir, and Outcome in Patients With Ebola Virus Disease, Guinea, 2015—A Retrospective Observational Study. *J Infect Dis*. 2019 Jun 19; 220(2):195–202. <https://doi.org/10.1093/infdis/jiz078> PMID: 30788508
39. Chapman LE, Mertz GJ, Peters CJ, Jolson HM, Khan AS, Ksiazek TG, et al. Intravenous ribavirin for hantavirus pulmonary syndrome: safety and tolerance during 1 year of open-label experience. *Ribavirin Study Group. Antivir Ther*. 1999; 4(4):211–9.
40. Mertz G, Miedzinski L, Goade D, Pavia A, Hjelle B, Hansbarger C, et al. Placebo-controlled, double-blind trial of intravenous ribavirin for the treatment of hantavirus cardiopulmonary syndrome in North America. *Clin Infect Dis*. 2004; 39(9):1307–1313. <https://doi.org/10.1086/425007> PMID: 15494907
41. Vial P, Valdivieso F, Ferres M, Riquelme R, Rioseco M, Calvo M, et al. High-dose intravenous methylprednisolone for hantavirus cardiopulmonary syndrome in Chile: a double-blind, randomized controlled clinical trial. *Clin Infect Dis*. 2013; 57(7):943–951. <https://doi.org/10.1093/cid/cit394> PMID: 23784924
42. Wernly JA, Dietl CA, Tabe CE, Pett SB, Crandall C, Milligan K, et al. Extracorporeal membrane oxygenation support improves survival of patients with Hantavirus cardiopulmonary syndrome refractory to

- medical treatment. *Eur J Cardio-Thorac Surg Off J Eur Assoc Cardio-Thorac Surg*. 2011 Dec; 40(6):1334–40. <https://doi.org/10.1016/j.ejcts.2011.01.089> PMID: 21900022
43. Vial P.A., Valdivieso F., Calvo M., Riaseco M.L., Riquelme R., Aranedo A., et al. A non-randomized multicentre trial of human immune plasma for treatment of hantavirus cardiopulmonary syndrome caused by Andes virus. *Antivir Ther*. 2015; 20(4):377–86. <https://doi.org/10.3851/IMP2875> PMID: 25316807
 44. Gui X, Ho M, Cohen M, Wang Q, Huang H, Xie Q. Hemorrhagic fever with renal syndrome: treatment with recombinant alpha interferon. *J Infect Dis*. 1987; 155(5):1047–1051. <https://doi.org/10.1093/infdis/155.5.1047> PMID: 2881968
 45. Du H., Li J., Yu H.-T., Jia Z.-S., Yu D.-H., Wang J.-P., et al. The optimal timing of RRT for critical patients with hemorrhagic fever with renal syndrome. *Ther Apher Dial*. 2013; 17(5):A2.
 46. Huggins JW, Hsiang CM, Cosgriff TM, Guang MY, Smith JI, Wu ZO, et al. Prospective, double-blind, concurrent, placebo-controlled clinical trial of intravenous ribavirin therapy of hemorrhagic fever with renal syndrome. *J Infect Dis*. 1991 Dec; 164(6):1119–27. <https://doi.org/10.1093/infdis/164.6.1119> PMID: 1683355
 47. Ilori E.A., Furuse Y., Ipadeola O.B., Dan-Nwafor C.C., Abubakar A., Womi-Eteng O.E., et al. Epidemiologic and clinical features of lassa fever outbreak in Nigeria, January 1–May 6, 2018. *Emerg Infect Dis*. 2019; 25(6):1066–74. <https://doi.org/10.3201/eid2506.181035> PMID: 31107222
 48. McCormick JB, King IJ, Webb PA, Scribner CL, Craven RB, Johnson KM, et al. Lassa fever. Effective therapy with ribavirin. *N Engl J Med*. 1986 Jan 2; 314(1):20–6. <https://doi.org/10.1056/NEJM198601023140104> PMID: 3940312
 49. Kawuki J, Musa TH, Yu X. Impact of recurrent outbreaks of Ebola virus disease in Africa: a meta-analysis of case fatality rates. *Public Health*. 2021; 195:89–97. <https://doi.org/10.1016/j.puhe.2021.03.027> PMID: 34077889
 50. Sivanandy P, Jun PH, Man LW, Wei NS, Mun NFK, Yii CAJ, et al. A Systematic Review of Ebola Virus Disease Outbreaks and an Analysis of the Efficacy and Safety of Newer Drugs Approved for the Treatment of Ebola Virus Disease by the US Food and Drug Administration from 2016 to 2020. *J Infect Public Health*. 2022; <https://doi.org/10.1016/j.jiph.2022.01.005> PMID: 35085865
 51. Arab-Bafrani Z, Jabbari A, Mostakhdem Hashemi M, Arabzadeh AM, Gilanipour A, Mousavi E. Identification of the crucial parameters regarding the efficacy of ribavirin therapy in Crimean–Congo haemorrhagic fever (CCHF) patients: a systematic review and meta-analysis. *J Antimicrob Chemother*. 2019; 74(12):3432–9. <https://doi.org/10.1093/jac/dkz328> PMID: 31369086
 52. Eberhardt KA, Mischlinger J, Jordan S, Groger M, Günther S, Ramharter M. Ribavirin for the treatment of Lassa fever: A systematic review and meta-analysis. *Int J Infect Dis*. 2019; 87:15–20. <https://doi.org/10.1016/j.ijid.2019.07.015> PMID: 31357056
 53. Salam AP, Cheng V, Edwards T, Olliaro P, Sterne J, Horby P. Time to reconsider the role of ribavirin in Lassa fever. *PLoS Negl Trop Dis*. 2021 Jul 8; 15(7):e0009522. <https://doi.org/10.1371/journal.pntd.0009522> PMID: 34237063
 54. Duvignaud A, Jaspard M, Etafo IC, Gabillard D, Serra B, Abejegah C, et al. Lassa fever outcomes and prognostic factors in Nigeria (LASCOPE): a prospective cohort study. *Lancet Glob Health*. 2021 Apr 1; 9(4):e469–78. [https://doi.org/10.1016/S2214-109X\(20\)30518-0](https://doi.org/10.1016/S2214-109X(20)30518-0) PMID: 33740408
 55. Isakov L, Lo AW, Montazerhodjat V. Is the FDA too conservative or too aggressive?: A Bayesian decision analysis of clinical trial design. *J Econom*. 2019; 211(1):117–36.
 56. Montazerhodjat V, Chaudhuri SE, Sargent DJ, Lo AW. Use of Bayesian decision analysis to minimize harm in patient-centered randomized clinical trials in oncology. *JAMA Oncol*. 2017; 3(9):e170123–e170123. <https://doi.org/10.1001/jamaoncol.2017.0123> PMID: 28418507

2.3. Séquelles des fièvres hémorragiques virales

Une partie des données recueillies dans notre revue systématique n'ont pas été intégrées dans notre publication. Il s'agit des données concernant les séquelles des différentes FHVs. Un total de 39 références a été identifié à ce sujet, dont une référence traitant de plusieurs FHVs. La grande majorité étudiait les séquelles de la maladie à virus Ebola (n=29), suivi de la fièvre hémorragique avec syndrome rénal (n=5), du syndrome pulmonaire à hantavirus (n=2), de la fièvre de Lassa (n=2), de la fièvre hémorragique de Crimée-Congo (n=1) et de la chorioméningite lymphocytaire (n=1). Le Tableau 4 détaille les séquelles rapportées dans les références identifiées pour chacune de ces FHVs. Comme nous pouvons le voir, il y a un manque important de données disponibles pour la majorité des FHVs.

Tableau 4 Résultats sur les séquelles rapportées par FHV

FHV	Nombre de références	Références	Résumé des résultats
Maladie à virus Ebola	29	De St. Maurice 2018 [34] ; Tiffany 2016 [35] ; Toure 2017 [36] ; Mwanza 2016 [37] ; Billioux 2017 [38] ; Clark 2015 [39] ; Pearson 2016 [40] ; Etard 2017 [6] ; Godwin 2019 [41] ; Sneller 2019 [42] ; Wilson 2018 [43] ; Kelly 2019 [44] ; Etard 2016 [45] ; Mattia 2016 [46] ; Keita 2019 [47] ; Hereth-Hebert 2017 [48] ; Shantha 2018 [49] ; Fallah 2016 [50] ; Pers 2017 [51] ; Tozay 2019 [7] ; Vandy 2018 [52] ; Howlett 2017 [53] ; Pers 2018 [54] ; Jagadesh 2018 [55] ; Steptoe 2017 [56] ; Pers 2016 [51] ; Scott 2016 [57] ; Guetiya	<ul style="list-style-type: none"> • 29 références, incluant des publications mais également des abstracts de conférences, ont évalué la présence de séquelles chez des survivants de la maladie à virus Ebola. Parmi ces références plusieurs concernent 2 cohortes principales : Postebogui et PREVAIL III. • L'étude observationnelle Postebogui évaluait les séquelles à long-terme de 802 survivants ayant été infectés durant l'épidémie de 2013-2016 en Guinée. La médiane entre la sortie du centre de traitement de la maladie à virus Ebola et l'inclusion dans la cohorte est de 350 jours. 76% (606 patients) des patients présentaient au moins un symptôme : <ul style="list-style-type: none"> ➔ 40% (324 patients) avaient des symptômes généraux (fatigue, fièvre, anorexie) ➔ 38% (303 patients) avaient des symptômes musculo-squelettiques (arthralgie, myalgie) ➔ 37% (298 patients) avaient des troubles neurosensoriels (céphalées) ➔ 18% (142 patients) avaient un trouble oculaire (conjonctivite, déficience visuelle, douleurs oculaires) ➔ 2% (19 patients) souffraient de surdit�. <p>Cependant une limite de cette �tude observationnelle est l'absence de groupe contr�le.</p> • L'�tude observationnelle PREVAIL III �tudiait les s�quelles sur 12 mois de 966 survivants compar� � 2350 contacts proches n'ayant pas �t� infect�s (groupe contr�le). <ul style="list-style-type: none"> ➔ 6 sympt�mes ont �t� rapport�s comme significativement plus pr�sents chez les survivants : <ul style="list-style-type: none"> ○ Pollakiurie (15% vs. 3%) ○ C�phal�es (48% vs. 36%) ○ Myalgie (23% vs. 10%)

FHV	Nombre de références	Références	Résumé des résultats
		Wadoum 2017 [58] ; Mohammed 2017 [59]	<ul style="list-style-type: none"> ○ Pertes de mémoire (29% vs. 5%) ○ Arthralgie (48% vs. 18%) <p>→ L'incidence de cas d'uvéïte était significativement plus élevée chez les survivants que chez le groupe contrôle (p<0.001)</p>
Fièvre hémorragique avec syndrome rénal	5	Stojanovic 2008 [60] ; Mäkelä 2010 [61] ; Miettinen 2006 [62] ; Miettinen 2009 [63] ; Mathes 2005 [64]	<ul style="list-style-type: none"> • 6 publications évaluant les séquelles chez les survivants de fièvre hémorragique avec syndrome rénal. • Les séquelles suivantes ont notamment été identifiées dans ces études : <ul style="list-style-type: none"> → Prévalence relativement élevée de troubles hormonaux (hypopituitarisme), par exemple 18% (11 survivants sur 60) ayant un trouble endocrinien [60]. Des troubles hormonaux ont également été retrouvés chez les patients ayant eu une néphropathie épidémique (une forme particulière de fièvre hémorragique avec syndrome rénal) [61]. → 6 ans après la phase aiguë, prévalence plus élevées chez 37 survivants de néphropathie épidémique comparés à 38 patients contrôles, de protéinurie (p=0.006), débit de filtration glomérulaire plus élevé (p=0.016) et tension systolique moyenne plus élevée (p=0.012) [62,63].
Syndrome pulmonaire à hantavirus	2	Gracia 2010 [65]; Pergam 2009 [66]	<ul style="list-style-type: none"> • 2 publications évaluant les symptômes pulmonaires persistants, les anomalies dans la fonction pulmonaire et les complications rénales chez les survivants du syndrome pulmonaire à hantavirus. • 14 survivants du syndrome pulmonaire à hantavirus causé par le virus Choclo au Panama et 9 survivants du syndrome pulmonaire à hantavirus causé par le virus Sin Nombre au Nouveau Mexique (Etats-Unis) ont été suivis jusqu'à 8 ans après l'infection [65]. Il a été trouvé que la dyspnée à l'effort persistait 1-2 ans après la phase aiguë (43% chez les patients au Panama et 77% chez les patients au Nouveau Mexique). Le débit expiratoire moyen était également réduit chez les survivants. Cependant tous ces symptômes se corrigent avec le temps pour la plupart des survivants.

FHV	Nombre de références	Références	Résumé des résultats
			<ul style="list-style-type: none"> • 30 survivants ont été suivis jusqu'à 5 ans après l'infection pour évaluer les potentielles séquelles rénales [66]. La moitié des patients ont eu une protéinurie durant le suivi. La prévalence de protéinurie semble augmenter au fur et à mesure des visites. 16 patients sur les 30 ont des résultats répondant à la définition d'une insuffisance rénale chronique.
Fièvre de Lassa	2	Wilson 2014 [67] ; Rapport de l'OMS de 2019 [68]	<ul style="list-style-type: none"> • 1 publication décrivant les caractéristiques de plusieurs FHVs, dont la fièvre de Lassa et 1 rapport de l'OMS sur la fièvre de Lassa. • Les 2 références rapportent la surdité (à différents degrés) comme séquelle potentielle chez les survivants. L'OMS rapporte notamment que cela touche environ 25% des survivants avec un retour à la normale de l'audition après 1-3 mois chez 50% de ces survivants [68]. • L'ataxie cérébrale est également rapportée comme séquelle potentielle sans plus de détails [67].
Fièvre hémorragique de Crimée-Congo	1	Gul 2012 [69]	<ul style="list-style-type: none"> • 1 publication turque évaluant la qualité de vie liée à la santé et la présence de trouble de stress post-traumatique chez des survivants 12 mois après la récupération. • Sur les 54 patients inclus, 48% avaient des symptômes de trouble de stress post-traumatique et 19% avaient un diagnostic de trouble de stress post-traumatique. Ces résultats étaient plus marqués chez les patients ayant eu un séjour en réanimation, ayant eu une hémorragie et ceux ayant eu une transfusion sanguine.
Chorioméningite lymphocytaire	1	Wilson 2014 [67]	<ul style="list-style-type: none"> • 1 publication décrivant les caractéristiques de plusieurs FHVs, dont la chorioméningite lymphocytaire. • Des séquelles neurologiques ont été rapportées chez des survivants de chorioméningite lymphocytaire. Chez des femmes enceintes infectées, l'infection peut se propager chez le fœtus résultant en un avortement précoce ou à des malformations selon le moment où l'infection a eu lieu pendant la

FHV	Nombre de références	Références	Résumé des résultats
			grossesse. Les enfants nés par la suite peuvent souffrir de déficience mentale, d'épilepsie, de cécité ou de troubles moteurs.

3. Adaptation de la Bayesian Decision Analysis pour la mise en place d'essai clinique avec un critère de jugement binaire

3.1. Résumé

a. Adaptation à un critère de jugement binaire

Dans cette seconde partie de thèse, nous avons tout d'abord adapté la méthodologie de la BDA pour la mise en place d'un essai clinique fréquentiste avec un critère de jugement binaire. En effet, le critère de jugement le plus souvent utilisé dans un contexte de FHV est le décès à un temps donné. Nous avons considéré un essai fixe randomisé à deux bras. L'effet du traitement, δ , est défini par :

$$\delta = p_{drug} - p_{control}$$

avec p_{drug} la proportion de décès observée dans le bras expérimental et $p_{control}$ celle observée dans le bras contrôle. Puis, nous avons défini la magnitude de l'effet attendu du traitement par :

$$|\delta_0| = p_{control} \times Ratio$$

avec *Ratio* la valeur utilisée pour définir la réduction attendue de la mortalité telle que :

$$p_{drug} = p_{control} \times (1 - Ratio)$$

L'objectif ici est de développer une méthode permettant de calculer les erreurs de type I et tailles d'échantillon optimales pour un cas donné qui minimisent le coût lié au fait de conduire cet essai. Ainsi, nous avons défini notre matrice de coût en prenant en compte deux types de coût lié à cet essai :

- Un coût post-essai correspondant à deux situations : ne pas approuver un traitement efficace ou approuver un traitement inefficace
- Un coût pendant l'essai correspondant à deux situations : celle où les patients reçoivent un traitement inefficace pendant l'essai et celle où des patients sont dans le bras contrôle alors que le traitement expérimental est efficace

A partir de cela nous avons défini la fonction de coût suivante :

$$C(fxd(n, \lambda_n)) = p_0 \times Loss(H = H_0) + p_1 \times Loss(H = H_1)$$

- n est la taille d'échantillon par bras ;
- λ_n est la valeur critique de la statistique de test ;
- $fxd(n, \lambda_n)$ correspond à la comparaison de la statistique de test observée dans l'essai Z_n à la valeur critique λ_n pour un n donné
- p_0 est la probabilité *a priori* que le traitement ne soit pas efficace ;
- $Loss(H = H_0)$ est le coût associé au fait d'évaluer et d'approuver un traitement inefficace ;
- p_1 est la probabilité *a priori* que le traitement soit efficace ;
- $Loss(H = H_1)$ est le coût associé au fait de ne pas approuver un traitement efficace.

Les valeurs optimales de la taille d'échantillon, de la valeur critique et de l'erreur de type I sont ainsi calculées en résolvant le problème d'optimisation non-linéaire multivarié sous contrainte suivant (la puissance maximale étant fixée à 90%) :

$$(n^*, \lambda_n^*) = \underset{n \in \mathbb{N}, \lambda_n \in \mathbb{R}}{\operatorname{argmin}} C(fxd(n, \lambda_n)) \text{ soumis à } 1 - \beta \leq \text{Puissance}_{max}$$

- n^* est la taille d'échantillon optimale
- λ_n^* est la valeur critique optimale du test à échantillon fixe $fxd(n, \lambda_n)$
- β est l'erreur de type II
- Puissance_{max} est la borne supérieure de la puissance (ici fixée à 90%).

Ce travail a ensuite été codé sous R (R software v.4.2.2) et une fonction R a été développée avec :

- en entrée : la taille de la population cible, la mortalité de la maladie, l'effet attendu du traitement évalué, la probabilité que le traitement ne soit pas efficace, les effets indésirables du traitement, le pourcentage de séquelles et leur fardeau.
- en sortie : les valeurs optimales de la taille d'échantillon, de la valeur critique et de l'erreur de type I.

b. Illustration pour un essai clinique évaluant un traitement contre la maladie à virus Ebola

Nous avons illustré notre méthodologie dans le cadre de la conception d'un potentiel essai clinique évaluant un traitement contre la maladie à virus Ebola.

Les différents paramètres à pré-spécifier dans notre fonction de coût ont été essentiellement obtenus via deux sources :

- Des données générales venant des études du Global Burden of Disease [70] ;
- Des données recueillies lors de la revue systématique présentée en section 2 [71].

Nous avons également fait varier les paramètres correspondants à la taille de la population cible, à la mortalité et au ratio utilisé pour définir l'effet attendu du traitement à l'étude. Cela nous permet d'envisager les différentes valeurs optimales possibles selon les différents cas d'épidémies qui peuvent survenir.

Pour des tailles de population cible faibles, l'erreur de type I calculée par BDA est plus élevée que l'erreur usuelle de 2.5% (pour un test unilatéral). En effet, dans notre modèle, la fonction de coût prend en compte en parallèle les effets indésirables du traitement et le fardeau de la maladie. Ici nous avons considéré la maladie à virus Ebola qui est caractérisée par des taux de mortalité élevés. Ainsi, le coût lié aux effets indésirables du traitement peut être très faible en comparaison du fardeau de la maladie, ce qui est encore plus prononcé pour des tailles de population cible faibles étant donné que les coûts sont dépendants de la taille de cette population.

Dans les cas où l'on suppose un effet fort du traitement évalué et une taille de population cible élevée, l'erreur de type I calculée par BDA est proche voire parfois plus petite que l'erreur usuelle de 2.5% (test unilatéral), en particulier pour des taux de mortalité élevés. Dans ces cas-là, la taille d'échantillon calculée par BDA est proche voire plus élevée que celle calculée avec un test standard. Ce résultat reflète le fait que lorsque les nombres de cas et la mortalité sont élevés, si le traitement est supposé être très efficace alors on veut réduire l'erreur de type I.

Ce travail a pour objectif de proposer un nouvel outil aux chercheurs voulant mettre en place un essai clinique évaluant un traitement sur un critère de jugement binaire, notamment pour des maladies avec des taux de mortalité élevés et peu de traitements disponibles.

Cette méthodologie ainsi que son illustration dans le contexte de la maladie à virus Ebola a été soumise dans *Journal of Biopharmaceutical Statistics*.

3.2. Article 2 en révision dans Journal of Biopharmaceutical Statistics

1 **Adaptation of Bayesian decision analysis for clinical trial design to binary outcomes: Illustration for**
2 **Ebola Virus Disease Treatment**

3 Drifa Belhadi^{1,2}, Joonhyuk Cho^{3,4,8}, Denis Malvy^{5,6}, France Mentré^{1,2}, Andrew W. Lo^{3,7,8,9,10¶}, Cédric
4 Laouénan^{1,2*¶}

5 ¹ Université Paris Cité, Inserm, IAME, F-75018 Paris, France

6 ² AP-HP, Hôpital Bichat, Département d'Epidémiologie Biostatistiques et Recherche Clinique, F-75018
7 Paris, France

8 ³ MIT Laboratory for Financial Engineering, Cambridge, MA, USA

9 ⁴ MIT Department of Electrical Engineering and Computer Science, Cambridge, MA, USA

10 ⁵ UMR 1219 Inserm/EMR 271 IRD, University of Bordeaux, Bordeaux, France

11 ⁶ Department for Infectious and Tropical Diseases, University Hospital Center Pellegrin, Bordeaux,
12 France

13 ⁷ MIT Operations Research Center, Cambridge, MA, USA

14 ⁸ MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA, USA

15 ⁹ MIT Sloan School of Management, Cambridge, MA, USA

16 ¹⁰ Sante Fe Institute, Santa Fe, NM, USA

17 * Corresponding author: Cédric Laouénan; email address: cedric.laouenan@inserm.fr

18 ¶ Co-last authors

19

20 **Short title:** BDA to design an Ebola trial with a binary outcome

21 Abstract

22 When designing trials for diseases with high mortality and limited therapeutic options available—such
23 as Ebola virus disease (EVD)— using the conventional 5% type I error threshold (α) to determine the
24 trial size may not be ideal from the patients' perspective. A higher α may be more desirable for
25 seriously ill patients that have no other therapeutic alternatives. Bayesian decision analysis (BDA) is a
26 well-known technique that captures this phenomenon by incorporating the burden of disease and
27 patient preferences in determining the optimal α and sample size of the trial. In this study, we extend
28 the BDA framework to trials with binary outcomes, specifically those targeting EVD therapeutics. We
29 consider frequentist two-arm randomized trials with a binary outcome. Target population size (N),
30 mortality rate ($p_{control}$), expected treatment effect (*Ratio*) and parameters for treatment side effects
31 and disease severity need to be specified in the loss model. We set parameters to fit an EVD outbreak
32 context using mortality as outcome. BDA-optimal type I errors (α^*) and sample sizes per arm (n^*) are
33 compared to a standard one-sided test with $\alpha_{ref} = 0.025$ and sample size n_{ref} . In most cases, α^* are
34 higher than α_{ref} and n^* are lower than n_{ref} . In other word, EVD patients benefit the most when the
35 trial is designed with lower n^* and higher α^* . This study provides a new tool for designing trials with a
36 binary outcome. A standard α may not be always appropriate, especially for diseases with high
37 mortality and few therapies available.

38

39 Keywords

40 Bayesian decision analysis; clinical trial; binary outcome; mortality; Ebola virus disease; viral
41 hemorrhagic fevers; burden of disease; type I error; sample size.

42

43 Introduction

44 Designing a clinical trial during an outbreak of an emerging infectious disease is a major challenge. Viral
45 hemorrhagic fevers (VHFs), including Ebola virus disease, are a group of febrile diseases that recently
46 caused several outbreaks [1,2]. As very few therapeutic options are available against VHFs, new
47 experimental drugs need to be assessed. The unpredictability of the outbreaks and high mortality rates
48 lead regularly to restricted clinical capacity for care and research and reluctance to use
49 methodologically strong clinical trial design, such as a randomized controlled trial [3,4]. When
50 conducting clinical trials to assess VHF therapeutic options, researchers face a reduced number of
51 included patients and non-significant results due to the very conservative threshold of statistical
52 significance [5].

53 Previous Ebola virus disease clinical trials used different methods to adjust their design in order to take
54 into account high mortality rates, unpredictable outbreaks, and the importance of rapidly finding
55 effective treatments. In a trial comparing four therapeutic options (an antiviral agent remdesivir, a
56 single monoclonal antibody MAb114, a triple monoclonal antibody REGN-EB3 and a triple monoclonal
57 antibody ZMapp) for Ebola virus disease [6], they decided to use two Boschloo's exact one-tailed tests
58 at a type I error of 0.05 for each pairwise comparison without adjusting for multiple comparisons. In
59 another trial comparing the triple monoclonal antibody ZMapp to the standard of care, they faced
60 statistical considerations when designing the trial [7,8]. They decided to use a "less strict than usual"
61 type I error and a "barely Bayesian" adaptive trial design [9]. However, this method has previously
62 been questioned due to its properties leading to type I error levels higher than the conventional ones
63 and low power [10].

64 The use of an adaptive design in this context, such as a group-sequential design allowing for early trial's
65 termination for futility or efficacy, has been extensively discussed [11]. Simulation studies were
66 conducted to evaluate the use of an adaptive approach to assess experimental treatments for Ebola

67 virus disease [12]. However, running out of patients to include, especially after the outbreak peak,
68 leading to non-significant results, remains an issue even with an adaptive trial design [11].

69 We previously performed a simulation study comparing several clinical trial designs implemented at
70 different times during a VHF outbreak [13]. Four types of designs were compared: fixed/group-
71 sequential single-/two-arm trials. The conclusions of this study varied depending on the outbreak
72 timing. At the beginning of an outbreak, group-sequential two-arm trials should be preferred. Indeed,
73 at the beginning of an outbreak, the infected number of cases increases allowing to have more
74 potential patients to conduct a randomized controlled trial and the group-sequential design allows to
75 decide an early termination of the trial in case of futile or harmful experimental treatment. After the
76 epidemic peak, fixed single-arm design should be preferred, as the number of cases decreases.
77 However, this assumes a high level of confidence in the pre-trial historical outcome rate.

78 We can also argue that with high mortality rates and few effective therapies available, a 5% type I error
79 may not be appropriate and may raise ethical questions. Indeed, in that case, rejecting a new
80 treatment due to a lack of statistical power to show a significant result with a 5% threshold may be
81 complicated to accept for patients and healthcare workers. Moreover, when patients do not have any
82 other therapeutic options for a disease with a high mortality rate, they might also be willing to accept
83 a treatment with a certain efficacy even though this efficacy was not shown with a 5% significant
84 threshold. Recent studies have explored the application of Bayesian decision analysis (BDA) in order to
85 incorporate the burden of disease and patient preferences into clinical trial designs, especially for the
86 deadliest diseases in the US, such as cancers [14–17]. The BDA is a methodology that borrows the idea
87 from the Bayesian principle of incorporating prior information when designing the trial, while keeping
88 a frequentist design for the clinical trial itself. Therefore, we are considering here frequentist
89 randomized controlled trials. This method aims to minimize the overall expected harm to patients
90 participating in clinical trials as well as future patients, where “harm” consists of not receiving an
91 effective therapy and receiving an ineffective one. Burden of disease, mortality, sequelae, and

92 treatment side effects are incorporated in a loss matrix that is used to determine the optimal sample
93 size and critical value for a fixed-sample test. In previous works, continuous and survival outcomes,
94 such as survival duration, were considered as primary outcomes of the trials [14–17]. In the context of
95 VHF, especially Ebola virus disease, the most common primary outcome considered in clinical trials is
96 the mortality rate. Therefore, our objectives are to adapt the BDA model to a binary outcome, then to
97 illustrate the method to evaluate the optimal sample sizes and type I errors obtained in an Ebola virus
98 disease context. This work will be based on the Ebola virus disease data from the systematic review
99 we conducted to retrieve the number of cases, mortality, and treatments of VHF [1].

100

101 **Material and Methods**

102 The BDA model defined by Isakov et al. [14] focused on clinical trials with a continuous outcome. We
103 decided to extend this model to clinical trials with a binary outcome. Therefore, we reviewed the loss
104 model and the loss formulas from Isakov et al. to fit a binary outcome.

105 **Loss model and BDA-optimal fixed-sample test**

106 We consider a fixed two-arm trial with 1:1 randomization. The first step of designing a clinical trial is
107 to define the test that will be used to assess the effectiveness of the experimental treatment. Assuming
108 a certain magnitude of treatment effect of the experimental drug, we will then be able to calculate the
109 required sample size needed to show the expected effect of the experimental drug. Here we consider
110 a binary outcome assessing the effectiveness of a drug with a fixed-sample test such as a χ^2 test. The
111 outcome considered here is the mortality rate at a given time. The treatment effect, δ , is defined
112 as $\delta = p_{drug} - p_{control}$, with p_{drug} the outcome rate in the experimental arm and $p_{control}$ the
113 outcome rate in the control arm. The null hypothesis, H_0 , of the fixed-sample test corresponds to $\delta =$

114 0 (the drug is ineffective) and the alternative hypothesis, H_1 , corresponds to $\delta < 0$ (the drug is
115 effective). We define the magnitude of the treatment effect for an effective drug as:

$$116 \quad |\delta_0| = p_{control} \times Ratio, \quad (1)$$

117 where *Ratio* is the reduction ratio of the mortality rate associated with an effective drug.

118 We consider two types of loss associated with a clinical trial: post-trial loss and in-trial loss. All losses
119 or costs considered here are in terms of health consequences and not financial costs. While we focus
120 on the binary outcome for the clinical trial, our definition of “loss” in the loss function encompasses a
121 broad spectrum of health outcomes, including both mortality and patients’ quality of life. Two post-
122 trial losses are considered: the loss associated with not approving an effective treatment (type II error
123 cost) and the loss associated with approving an ineffective treatment (type I error cost). The in-trial
124 loss is defined as the loss of evaluating an ineffective treatment in a clinical trial or the loss for patients
125 in the control arm of not receiving the effective treatment. The loss matrix is reported in Table 1.

126 [Table 1 near here]

127 The expected value of the total loss associated with the given fixed-sample test $fxd(n, \lambda_n)$ is defined
128 with:

$$129 \quad C(fxd(n, \lambda_n)) = p_0 \times Loss(H = H_0) + p_1 \times Loss(H = H_1), \quad (2)$$

130 where n is the sample size per arm; λ_n is the critical value of the fixed-sample test; p_0 is the probability
131 that the treatment is not effective, $p_0 = p(H_0)$; and p_1 is the probability that the treatment is
132 effective, $p_1 = p(H_1) = 1 - p_0$.

133 Compared to the BDA model from Isakov et al. [14] for continuous outcomes, we had to review the
134 formulas of the losses associated with $H = H_0$ and $H = H_1$ to take into account the binary nature of
135 our outcome. Based on our loss matrix, the loss associated with evaluating and approving an
136 ineffective treatment is defined as:

137
$$Loss(H = H_0) = p_{0_H0}Nc_1 + nc_1, \quad (3)$$

138 where p_{0_H0} is the probability that the drug is approved under H_0 . As under H_0 the treatment effect is
 139 supposed to be null ($\delta = 0$), using the cumulative distribution function of a standard normal random
 140 variable ϕ , we defined p_{0_H0} as:

141
$$p_{0_H0} = \phi(-\lambda_n), \quad (4)$$

142 We defined the loss associated with not approving an effective treatment as:

143
$$Loss(H = H_1) = p_{1_H1}Nc_2 + nc_2, \quad (5)$$

144 where c_2 is the burden of disease per patient (overall severity of the disease) and p_{1_H1} is the
 145 probability that the drug is not approved under H_1 . We needed to define p_{1_H1} considering a binary
 146 outcome. Under H_1 the treatment effect is not null, therefore we considered the expected treatment
 147 effect δ_0 used to design the clinical trial. We defined the variance of δ_0 as
 148 $p_{drug}(1 - p_{drug}) + p_{control}(1 - p_{control})/n$ with n the sample size in each arm. Therefore, using
 149 the cumulative distribution function of a standard normal random variable ϕ , we defined p_{1_H1} as:

150
$$p_{1_H1} = \phi\left(\lambda_n - |\delta_0|/\sqrt{\frac{p_{drug}(1 - p_{drug}) + p_{control}(1 - p_{control})}{n}}\right) \quad (6)$$

151 We determine the optimal sample size per arm, n^* , and critical value, λ_n^* , by solving the following
 152 constrained nonlinear multivariate optimization problem:

153
$$(n^*, \lambda_n^*) = \arg \min_{n \in \mathbb{N}, \lambda_n \in \mathbb{R}} C(fxd(n, \lambda_n)) \text{ subject to } 1 - \beta \leq Power_{max}, \quad (7)$$

154 where $Power_{max}$ is the upper bound for the power level, which is set to 90%.

155 **General parameters specification**

156 The value of the parameter c_1 can be calculated based on global costs of adverse effects of medical
 157 treatment per patient from the Global Burden of Disease 2017 study [18].

158 The c_2 parameter is defined as:

$$159 \quad c_2 = \frac{D + YLD}{D + N}, \quad (8)$$

160 where D is the number of deaths caused by the disease per year defined as:

$$161 \quad D = N \times p_{control}, \quad (9)$$

162 and YLD is the years lived with disability defined as:

$$163 \quad YLD = w \times p_{seq} \times N, \quad (10)$$

164 where w is the disability weight associated with the disease sequelae, which can be calculated from
165 measured values of the Global Burden of Disease 2010 study [19] and p_{seq} is the percentage of
166 sequelae associated with the disease.

167 **Development of an R function**

168 This binary-BDA model is programmed using R software version 4.2.2 (see Additional file 1). To obtain
169 the optimal sample size, n^* , critical value, λ_n^* , and the associated type I error, α^* , the following set of
170 parameters need to be specified in the function: $(p_0, c_1, w, p_{seq}, N, p_{control}, Ratio)$.

171 To solve the optimization problem, a constrained optimization by linear approximations (COBYLA)
172 method is conducted using the 'nloptr' package. The COBYLA algorithm implemented by Powell [20,21]
173 is a derivative-free optimization with nonlinear inequality and equality constraints. At each iteration
174 of the algorithm, linear approximations of the objective and constraints function are constructed using
175 a simplex, and the approximations are optimized in a trust region at each step.

176 **Illustration for trials on Ebola virus disease treatment**

177 In this illustration, the parameters of the binary-BDA model are specified for the context of therapeutic
178 options evaluation for Ebola virus disease. Table 2 reports the full parameter specifications used in this
179 illustration.

180 [Table 2 near here]

181 The probability that the drug candidate is not effective, p_0 , is set to 0.5 to assume a non-informative
182 prior on the efficacy of the drug [14,17].

183 To calculate the overall severity of the disease, c_2 , the disability weight, w , associated with post-acute
184 consequences of infectious disease measured from the Global Burden of Disease 2010 study is used
185 [19]. The percentage of sequelae is $p_{seq} = 0.76$ as reported for Ebola virus disease [22,23]. This
186 percentage of sequelae is based on all possible sequelae reported by Ebola survivors. To assess the
187 impact of this percentage on the results, a sensitivity analysis with $p_{seq} = 0.20$ is conducted.

188 Regarding the parameters $(N, p_{control}, RR)$, several scenarios are performed in order to obtain the
189 optimal sample size per arm, n^* , and type I error, α^* , over a range of specified parameter values. Based
190 on a previous systematic review [1], over the last five years, the number of cases per year (when Ebola
191 virus disease cases were reported) varied from around 50 to 3500 and case fatality rates varied from
192 0.39 to 0.65. Therefore, for the various scenarios, N varies in {50; 100; 250; 500; 1000; 2500; 5000};
193 $p_{control}$ varies in {0.30; 0.35; 0.40; 0.45; 0.50; 0.55; 0.60; 0.65; 0.70}; and for the reduction ratio *Ratio*,
194 we considered three values of expected treatment efficacy used to design the trial: {0.20; 0.35; 0.50}.

195 The optimal type I error, α^* , and sample size per arm, n^* , obtained in each scenario are compared to
196 the associated reference scenario of a standard sample-size calculation for one-sided test for
197 proportions with a type I error $\alpha_{ref} = 0.025$, a power of 0.90 and the associated sample size, n_{ref} .

198 Results

199 Development of the binary-BDA model

200 The BDA model is adapted to be used for randomized clinical trials with a binary primary outcome. The
201 associated R function uses the following parameters as input: $(p_0, c_1, w, p_{seq}, N, p_{control}, Ratio)$ and
202 gives the BDA-optimal sample size, critical value and associated type I error and power. The input
203 parameters can be fixed values or included in a set of given values.

204 Illustration for Ebola virus disease treatment trial design

205 **Figure 1, Figure 2** and **Figure 3** illustrate the optimal sample sizes and type I error rates for the three
206 values of *Ratio* {0.20; 0.35; 0.50} across the different values of population sizes (*N*) and mortality rates
207 ($p_{control}$). Full tables with the results in all scenarios are reported in Additional file 1 (Tables S1, S2 and
208 S3).

209 [Figure 1 near here]

210 When the population size is low, the optimal type I error, α^* , is higher than the reference type I error,
211 $\alpha_{ref} = 0.025$, for all three scenarios of treatment efficacy. Therefore, the optimal number of patients,
212 n^* , to be included per arm in the clinical trial is lower than the sample size, n_{ref} , calculated in the
213 reference scenario. For example, with $N = 500$ reported cases per year and a mortality
214 rate $p_{control} = 0.45$, the optimal sample sizes obtained for the three scenarios of *Ratio* =
215 {0.20; 0.35; 0.50} are $n^* = \{43; 40; 40\}$, respectively, compared to the sample size calculated in the
216 reference scenario: $n_{ref} = \{620; 193; 88\}$.

217 [Figure 2 near here]

218 For large population sizes, when *Ratio* = 0.35 (see **Figure 2**), the optimal type I error, α^* , is close to
219 the reference type I error, α_{ref} , and the optimal number of patients, n^* , is lower or close to the
220 reference sample size, n_{ref} . For example with $N = 5000$ reported cases per year and a mortality
221 rate $p_{control} = 0.60$, the optimal sample size obtained for *Ratio* = 0.35 is $n^* = 116$ compared
222 to $n_{ref} = 114$.

223 [Figure 3 near here]

224 When the reduction ratio is expected to be higher, i.e., *Ratio* = 0.50 (see **Figure 3**), and the
225 population size is large (more than 2500 reported cases per year), the optimal type I error, α^* , is close
226 or lower than the reference type I error, α_{ref} , especially for very high mortality rates. Therefore, the
227 optimal number of patients per arm in the trial, n^* , is close or higher than the sample size,

228 n_{ref} , calculated in the reference scenario. This result translates the fact that when the number of cases
229 and mortality rates are very high, with a treatment which is expected to be very effective, we want to
230 reduce the type I error.

231 Figures for the power obtained in each scenario are provided in the Additional file 2 (see Figures S1,
232 S2 and S3). In most cases, the power is capped at 90%, except for lower Ratios and low population
233 sizes.

234 The results of the sensitivity analysis on the percentage of sequelae (from 0.76 before to 0.2 here) are
235 summarized in the Additional file 2 (see Figures S4 to S9). Results are very close to the results of the
236 base case scenarios with similar conclusions.

237

238 Discussion

239 This study reports a new adaptation of a BDA model for designing clinical trial when the outcomes are
240 binary. Through this adaptation, we developed a transparent, objective, and reproducible framework
241 for determining the optimal type I error threshold, which is particularly important when designing
242 clinical trials in the context of diseases with high mortality rates and few therapies available. This
243 methodology incorporates the post-trial loss associated with not approving an effective treatment or
244 approving an ineffective treatment; and the in-trial loss associated with the administration of a
245 potentially ineffective treatment to patients during the trial. Given a fixed two-arm trial, we showed
246 that the optimal type I error and sample sizes to be used to design a clinical trial vary across population
247 sizes and case fatality rates.

248 In the context of therapeutic options evaluations for diseases such as VHF, our study aims at helping
249 future researchers in their process of designing a clinical trial in order to find an effective treatment as
250 quickly as possible. Our results are illustrated using the Ebola virus disease as a case study. We specified
251 the model parameters and explored different scenarios to reflect the Ebola virus disease context based

252 on worldwide disease numbers over the last few years [1]. When the target population size is low,
253 which was the case in the last few years for Ebola virus disease, the optimal type I error is found to be
254 higher than the usual one-sided type I error of 0.025. Indeed, in the model, losses are a trade-off
255 between treatment side effects and burden of disease. We are considering Ebola virus disease which
256 is characterized by high mortality rates. Thus, the health consequences of treatment side effects is low
257 in comparison to the burden of disease. This is even more pronounced for low population size, as the
258 losses are dependent on the population sizes. Therefore, the number of patients to be included in the
259 clinical trial in order to identify a new therapeutic option will be lower than the sample size calculated
260 with a standard test. Reducing the number of subjects to be included in a clinical trial is a key point in
261 the context of deadly diseases with hard-to-predict outbreaks where the number of infected cases
262 decreases after the outbreak peak.

263 In contrast for large target population sizes and very high case fatality rates, the optimal type I error
264 was found to be lower than the usual type I error of 0.025, especially when we assess a drug candidate
265 that is expected to have a strong positive effect. Therefore, with more infected cases and higher
266 mortality rates, more patients than the sample size calculated with a standard test should be included
267 in the clinical trial in order to increase the evidence around the potentially very effective treatment.
268 Our methodology allows the clinical trial design to adapt to the fluctuations in the number of cases as
269 the epidemic progresses, with high numbers before the epidemic peak and low numbers after.
270 Previously, it has been recommended that a fixed single-arm design should be preferred after the
271 epidemic peak as the number of cases decreases [13]. By calculating optimal type I errors associated
272 with a low number of cases, our binary-BDA methodology could allow use of a fixed two-arm design
273 even after the peak, instead of a fixed single-arm design. This could improve the quality of therapeutic
274 options evaluation in this context and contribute to improved efficiency.

275 A limitation of our study is that the findings are based on previously reported population sizes and case
276 fatality rates. However, in a context of diseases characterized by disparate outbreaks, these numbers

277 vary greatly across years and the number of infected cases in the future years is difficult to predict.
278 That is why we focused our illustration on a wide range of Ebola virus disease numbers from the last
279 few years. The different scenarios give an overview of optimal type I errors and sample sizes across
280 several situations and show the influence of each key parameter on the optimal type I error and sample
281 size.

282 One discussion point of our methodology is that we decided to set the probability that the treatment
283 is not effective p_0 at 0.50. Indeed, based on the publication by Isakov et al., by setting a 50% chance
284 of having an ineffective treatment we assume a non-informative prior and therefore we follow the
285 “equipoise” principle of two-arm trials. [14] Therefore, this would reflect the uncertainty of the
286 medical community on the preferred treatment. [24] However, it would be interesting to assign prior
287 distributions on these parameters instead of fixed values. In that case we would therefore use a
288 Bayesian design as we would run the model for a certain number of iterations with at each iteration a
289 new value of the probability drawn from the prior distribution assigned to the parameter.

290 Regarding the use of our BDA framework, a potential idea would be to use it for future pandemic of
291 emergent diseases. In the last few years, the world experienced an important pandemic of an
292 infectious disease, the Covid-19. However, in that particular case, the use of the BDA would have been
293 limited especially at the beginning of the pandemic. Indeed, it was the emergence of a new disease for
294 which we had limited information regarding prevalence, mortality or sequelae. Therefore, it would
295 have been complicated to use our BDA framework to design a clinical trial. In that case, a real Bayesian
296 design might have been more appropriate.

297 This work can be extended to the evaluation of therapeutic options for other diseases when designing
298 clinical trials with a binary primary outcome by adapting the different parameters to fit the context of
299 each disease, especially for diseases with high mortality rates and few therapies available. Following
300 this study, we wish to develop an R package that would be available from the Comprehensive R Archive
301 Network (CRAN) for users to calculate optimal type I errors and sample sizes based on our method.

302 **Conclusions**

303 With this new adaptation of the BDA model for designing clinical trials with binary outcomes, we
304 explored alternatives to the use of the common type I error threshold. We illustrated this methodology
305 in the context of evaluating an experimental treatment for Ebola virus disease. We explored various
306 values of target population size, mortality rate, and expected treatment effect sizes to assess their
307 impact on the BDA-optimal type I error and sample size. The result shows that optimal type I error
308 threshold to minimize patient loss is higher than baseline value, 5%, when total number of outbreaks
309 is below 2500. Through this BDA adaptation, we provide a new tool to researchers for designing clinical
310 trials with a binary primary outcome to assess therapeutic options, which can be particularly useful for
311 diseases with high mortality rates and few therapies available.

312

313 **Acknowledgements**

314 We thank Jessie Y. Liu for helpful comments and discussion.

315 **Declarations of interest statement**

316 **Competing interests**

317 DB, JC, DM, FM, and CL declare that they have no competing interests. AWL reports personal
318 investments in private biotechnology companies, biotechnology venture capital funds, and mutual
319 funds. AWL is a co-founder and principal of QLS Advisors LLC, a healthcare investments advisor, and
320 QLS Technologies LLC, a healthcare analytics and consulting company; a director of AbCellera, Annual
321 Reviews, Atomwise, and BridgeBio Pharma; and an advisor to Apricity Health, Aracari Bio, BrightEdge
322 Impact Fund, Enable Medicine, FINRA, Health at Scale, Lazard, MIT Proto Ventures, Quantile Health,
323 Roivant Social Ventures, Swiss Finance Institute, Thalès, Uncommon Cures, and xCures. During the
324 most recent six-year period, AWL has received speaking/consulting fees, honoraria, or other forms of

325 compensation from AbCellera, AlphaSimplex Group, Annual Reviews, Apricity Health, Aracari Bio,
326 Atomwise, Bernstein Fabozzi Jacobs Levy Award, BridgeBio, Cambridge Associates, Chicago Mercantile
327 Exchange, Enable Medicine, Financial Times Prize, Harvard Kennedy School, IMF, Journal of Investment
328 Management, Lazard, National Bank of Belgium, New Frontier Advisors/Markowitz Award,
329 Oppenheimer, Princeton University Press, Q Group, QLS Advisors, Quantile Health, Research Affiliates,
330 Roivant Sciences, SalioGen Therapeutics, Swiss Finance Institute, and WW Norton.

331 **Funding**

332 The authors received no specific funding for this work.

333 **Authors' contributions**

334 DB, FM and CL designed the study. DB performed the adaptation of the method and drafted the
335 manuscript. DB, JC, FM and CL participated in the statistical analysis. All authors (DB, JC, DM, FM, AWL
336 and CL) read and approved the final manuscript.

337 **Availability of data and materials**

338 All data generated or analyzed during this study are included in this published article and its
339 supplementary information files.

340

341 **References**

- 342 1. Belhadi D, El Baied M, Mulier G, Malvy D, Mentré F, Laouénan C. The number of cases,
343 mortality and treatments of viral hemorrhagic fevers: A systematic review. *PLoS Negl Trop Dis.*
344 2022 Oct;16(10):e0010889.
- 345 2. Centers for Disease Control and Prevention. Diseases Caused by High-Consequence Viruses.
346 <https://www.cdc.gov/ncezid/dhcpp/vspb/diseases.html>. Accessed 5 February 2023.
- 347 3. World Health Organization. Guidance for managing ethical issues in infectious disease
348 outbreaks. [https://apps.who.int/iris/bitstream/handle/10665/250580/9789241549837-](https://apps.who.int/iris/bitstream/handle/10665/250580/9789241549837-eng.pdf)
349 [eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/250580/9789241549837-eng.pdf). Accessed 5 February 2023.

- 350 4. Ellenberg SS, Keusch GT, Babiker AG, Edwards KM, Lewis RJ, Lundgren JD, et al. Rigorous Clinical
351 Trial Design in Public Health Emergencies Is Essential. *Clin Infect Dis*. 2018 Apr 17;66(9):1467–9.
- 352 5. Lee JS, Adhikari NK, Kwon HY, Teo K, Siemieniuk R, Lamontagne F, et al. Anti-Ebola therapy for
353 patients with Ebola virus disease: a systematic review. *BMC Infect Dis*. 2019;19(1):1–11.
- 354 6. Mulangu S, Dodd LE, Davey RT Jr, Tshiani Mbaya O, Proschan M, Mukadi D, et al. A
355 Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *N Engl J Med*. 2019 Dec
356 12;381(24):2293–303.
- 357 7. Dodd LE, Proschan MA, Neuhaus J, Koopmeiners JS, Neaton J, Beigel JD, et al. Design of a
358 Randomized Controlled Trial for Ebola Virus Disease Medical Countermeasures: PREVAIL II, the
359 Ebola MCM Study. *J Infect Dis*. 2016 Jun 15;213(12):1906–13.
- 360 8. PREVAIL II Writing Group, Multi-National PREVAIL II Study Team, Davey RT Jr, Dodd L, Proschan
361 MA, Neaton J, et al. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J*
362 *Med*. 2016 Oct 13;375(15):1448–56.
- 363 9. Proschan MA, Dodd LE, Price D. Statistical considerations for a trial of Ebola virus disease
364 therapeutics. *Clin Trials*. 2016 Feb;13(1):39–48.
- 365 10. Whitehead J. A comparison of the barely Bayesian design with the triangular test for clinical
366 trials in infectious diseases. *Clin Trials*. 2016 Aug 1;13(4):451–3.
- 367 11. Whitehead J, Olliaro P, Lang T, Horby P. Trial design for evaluating novel treatments during an
368 outbreak of an infectious disease. *Clin Trials*. 2016 Feb;13(1):31–8.
- 369 12. Cooper BS, Boni MF, Pan-ngum W, Day NPJ, Horby PW, Olliaro P, et al. Evaluating clinical trial
370 designs for investigational treatments of Ebola virus disease. *PLoS Med*. 2015 Apr
371 14;12(4):e1001815–e1001815.
- 372 13. Manchon P, Belhadi D, Mentré F, Laouénan C. Evaluation by simulation of clinical trial designs
373 for evaluation of treatment during a viral haemorrhagic fever outbreak. *BMC Med Res*
374 *Methodol*. 2021 May 6;21(1):98.
- 375 14. Isakov L, Lo AW, Montazerhodjat V. Is the FDA too conservative or too aggressive?: A Bayesian
376 decision analysis of clinical trial design. *J Econom*. 2019;211(1):117–36.
- 377 15. Xu Q, Cho J, Ben Chaouch Z, Lo AW. Incorporating patient preferences and burden-of-disease in
378 evaluating ALS drug candidate AMX0035: a Bayesian decision analysis perspective. *Amyotroph*
379 *Lateral Scler Frontotemporal Degener*. 2023;24(3–4):281–8.
- 380 16. Montazerhodjat V, Chaudhuri SE, Sargent DJ, Lo AW. Use of Bayesian decision analysis to
381 minimize harm in patient-centered randomized clinical trials in oncology. *JAMA Oncol*.
382 2017;3(9):e170123–e170123.
- 383 17. Chaudhuri SE, Ben Chaouch Z, Hauber B, Mange B, Zhou M, Christopher S, et al. Use of Bayesian
384 decision analysis to maximize value in patient-centered randomized clinical trials in Parkinson’s
385 disease. *J Biopharm Stat*. 2023;1–20.
- 386 18. Institute for Health Metrics and Evaluation. Global Burden of Disease 2017 [Internet]. [cited
387 2023 Feb 5]. Available from: <https://gbd2017.healthdata.org/gbd-results/>

- 388 19. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in
389 assessing health outcomes from disease and injury: disability weights measurement study for
390 the Global Burden of Disease Study 2010. *Lancet*. 2012 Dec 15;380(9859):2129–43.
- 391 20. Powell MJD. A Direct Search Optimization Method That Models the Objective and Constraint
392 Functions by Linear Interpolation. In: Gomez S, Hennart JP, editors. *Advances in Optimization
393 and Numerical Analysis*. Springer, Dordrecht; 1994. p. 51–67. (Mathematics and Its
394 Applications; vol. 275).
- 395 21. Powell MJD. Direct search algorithms for optimization calculations. *Acta Numerica*. 1998
396 Jan;7:287–336.
- 397 22. Etard JF, Sow MS, Leroy S, Touré A, Taverne B, Keita AK, et al. Multidisciplinary assessment of
398 post-Ebola sequelae in Guinea (Postebogui): an observational cohort study. *Lancet Infect Dis*.
399 2017 May;17(5):545–52.
- 400 23. Tozay S, Fischer WA, Wohl DA, Kilpatrick K, Zou F, Reeves E, et al. Long-term Complications of
401 Ebola Virus Disease: Prevalence and Predictors of Major Symptoms and the Role of
402 Inflammation. *Clin Infect Dis*. 2020 Oct 23;71(7):1749–55.
- 403 24. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987 Jul
404 16;317(3):141–5.

405

406 Appendices

407 **Additional file 1: AdditionalFile1_R_function.pdf**

408 **Additional file 2: AdditionalFile2_SensitivityAnalysis.pdf**

409

410

411 **Tables**

412 **Table 1. Loss matrix**

	Post-trial loss		In-trial loss
	$\hat{H} = H_0$ (do not approve the treatment)	$\hat{H} = H_1$ (approve the treatment)	
$H = H_0$ (ineffective treatment)	0	Nc_1	nc_1
$H = H_1$ (effective treatment)	Nc_2	0	nc_2

413 *Notations: H_0 , null hypothesis of the fixed-sample test corresponding to an ineffective drug; H_1 ,*
 414 *alternative hypothesis of the fixed-sample test corresponding to an effective drug; N , size of the target*
 415 *population; n , sample size per arm; c_1 , health consequences of treatment side effects per patient; c_2 ,*
 416 *burden of disease per patient.*

417

418 **Table 2. Parameter specifications used in the illustration**

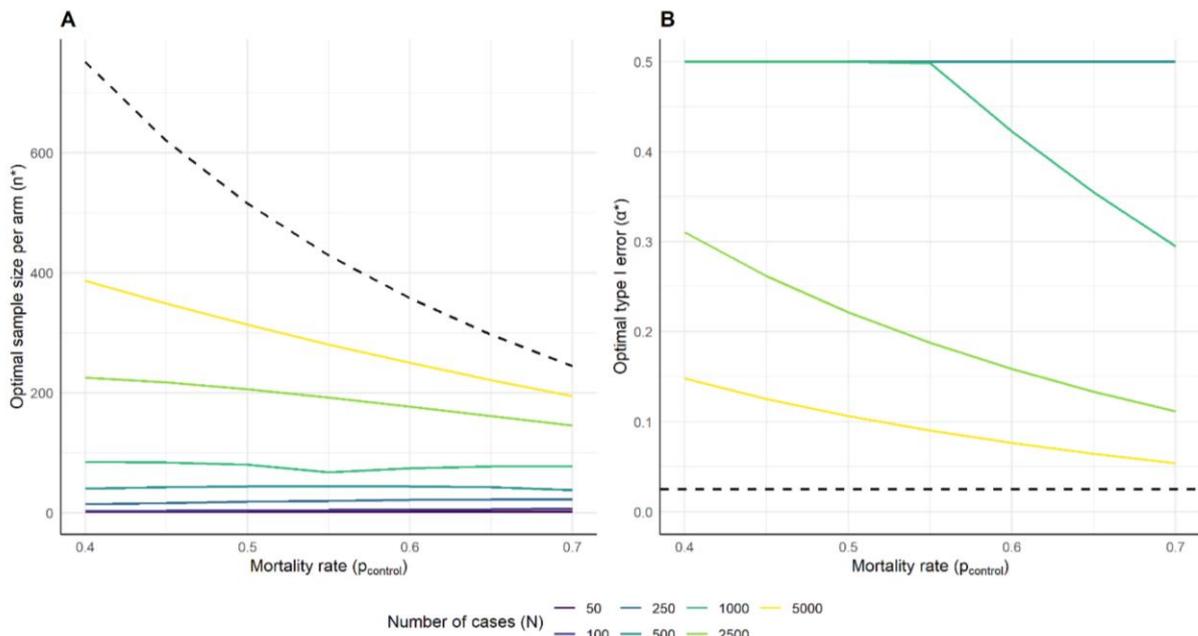
Parameter	Value	Definition	Source
p_0	0.5	Probability that the treatment is ineffective	Assumption (non-informative prior) [14]
c_1	0.17	Health consequences of side effects per patient	Global value based on the Global Burden of Disease 2017 study [18]
w	0.254	Disability weight associated with post-acute consequences of infectious disease	Global Burden of Disease 2010 study [19]
p_{seq}	0.76	Percentage of sequelae associated with the disease	Reported percentage of sequelae for Ebola virus disease [22,23]
N	{50; 100; 250; 500; 1000; 2500; 5000}	Size of the target population	Assumption
$p_{control}$	{0.30; 0.35; 0.40; 0.45; 0.50; 0.55; 0.60; 0.65; 0.70}	Control outcome rate (mortality rate)	Assumption
Ratio	{0.20; 0.35; 0.50}	Reduction ratio of the mortality rate associated with an effective drug	Assumption

419

420

421 **Figures**

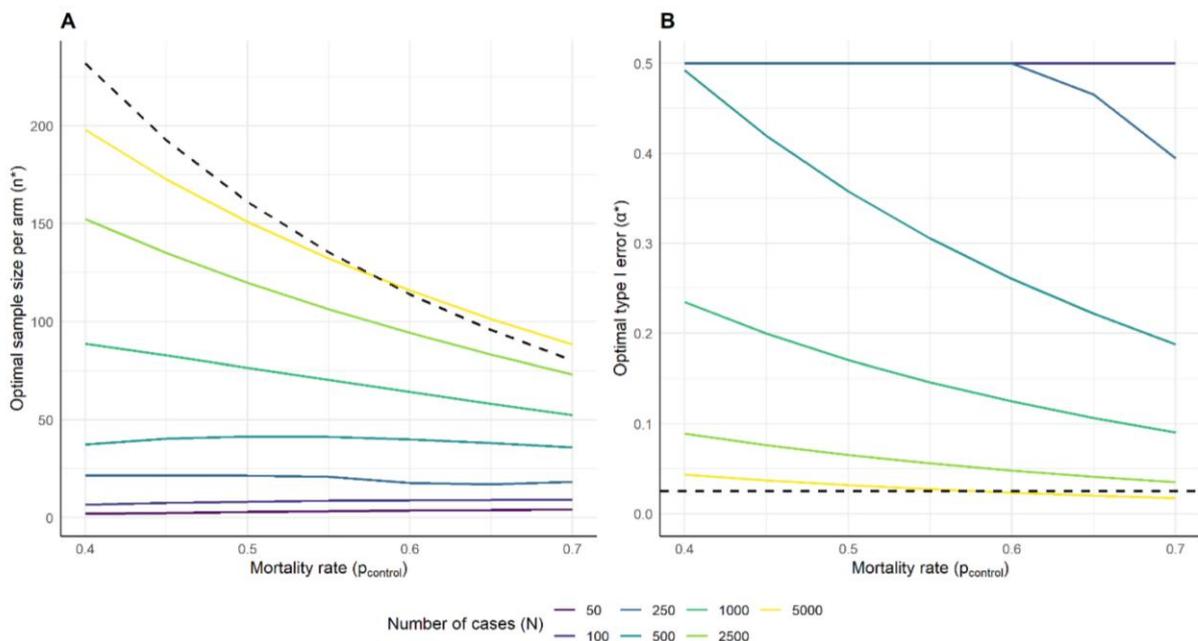
422 **Figure 1. Reduction ratio of 0.20 – Optimal sample sizes (panel A) and type I error rates (panel B)**



423

424 *Note: the dashed line corresponds to the reference for a one-sided test with a type I error=0.025 and a*
 425 *power=0.9*

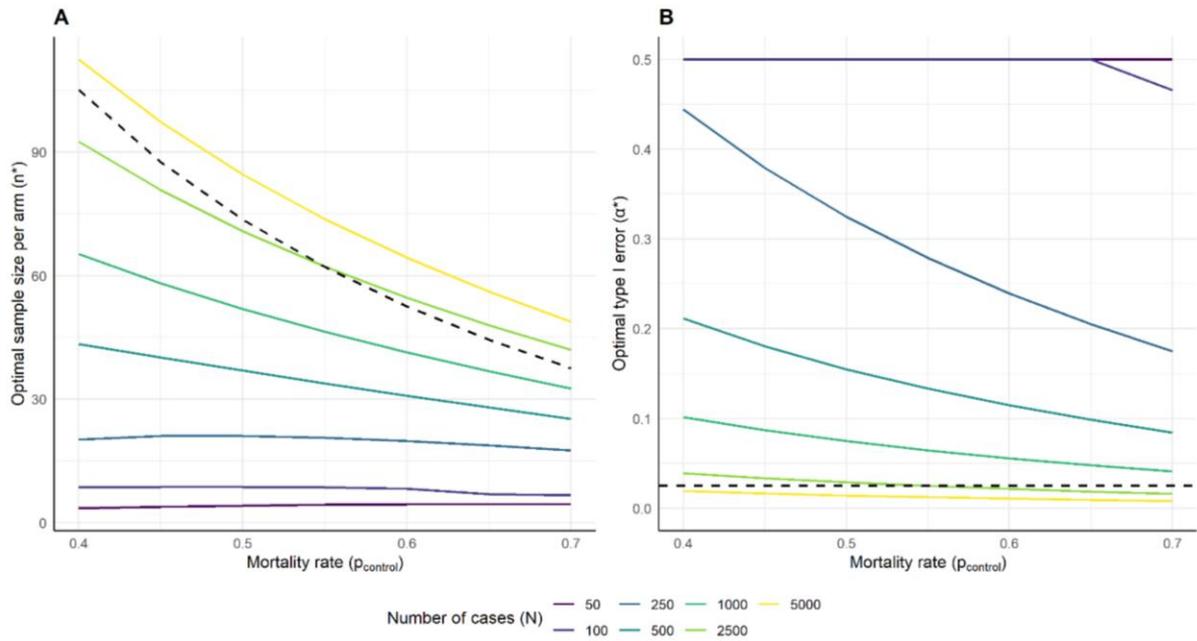
426 **Figure 2. Reduction ratio of 0.35 – Optimal sample sizes (panel A) and type I error rates (panel B)**



427

428 *Note: the dashed line corresponds to the reference for a one-sided test with a type I error=0.025 and a*
 429 *power=0.9*

430 **Figure 3. Reduction ratio of 0.50 – Optimal sample sizes (panel A) and type I error rates (panel B)**



431

432 *Note: the dashed line corresponds to the reference for a one-sided test with a type I error=0.025 and a*
433 *power=0.9*

434 Figure captions

435 **Figure 1. Reduction ratio of 0.20 – Optimal sample sizes (panel A) and type I error rates (panel B)**

436 *Note: the dashed line corresponds to the reference for a one-sided test with a type I error=0.025 and a*
437 *power=0.9*

438

439 **Figure 2. Reduction ratio of 0.35 – Optimal sample sizes (panel A) and type I error rates (panel B)**

440 *Note: the dashed line corresponds to the reference for a one-sided test with a type I error=0.025 and a*
441 *power=0.9*

442

443 **Figure 3. Reduction ratio of 0.50 – Optimal sample sizes (panel A) and type I error rates (panel B)**

444 *Note: the dashed line corresponds to the reference for a one-sided test with a type I error=0.025 and a*
445 *power=0.9*

4. Discussion et perspectives

L'objectif de cette thèse était de fournir un nouvel outil d'aide au design d'un essai clinique dans un contexte particulier tel que celui des FHV. La première étape a été d'identifier toutes les données scientifiques disponibles sur ces maladies via une revue systématique de la littérature. La deuxième étape a été d'adapter la méthodologie de la BDA afin de pouvoir l'utiliser avec un critère de jugement binaire puis de l'illustrer en utilisant les données issues de la revue systématique.

4.1. Première revue systématique publiée sur le nombre de cas, la mortalité et les traitements des FHV

A notre connaissance, nous avons conduit la première revue systématique publiée synthétisant les connaissances sur le nombre de cas, la mortalité et les traitements d'un nombre aussi important de FHV. Notre revue systématique a montré que les nombres de cas et la mortalité varient grandement pour une même FHV entre les épidémies et entre les pays. De plus, le nombre de références évaluant des traitements est très limité avec peu d'essais contrôlés randomisés. Pour plusieurs FHV, nous n'avons pas identifié de traitement spécifique associé à une preuve robuste d'efficacité sur la mortalité.

Ce travail présente certaines limites, notamment une sous-estimation potentielle des données réelles. En effet, nous avons recueilli les informations disponibles sur les sites de l'OMS et du CDC, sur les sites nationaux de santé et celles publiées dans la littérature scientifique. Les informations non rapportées par l'un de ces moyens n'ont donc pas été captées. De plus, il s'agit ici de maladies relativement rares affectant des zones rurales très dispersées, ce qui implique potentiellement une sous-notification des cas. Au-delà des données sous-notifiées, nous avons également décidé de nous concentrer sur la mortalité des FHV sans prendre en compte la transmissibilité du virus, la contagiosité et l'échappement immunitaire. Bien qu'importants, ces éléments n'ont donc pas été considérés dans notre synthèse sur les FHV.

Une autre limite de ce travail est que nous présentons les résultats par maladie, et n'avons pas stratifié nos résultats par espèce ou souche de virus. Cela est particulièrement notable pour le cas de la maladie à virus Ebola où les principaux essais cliniques conduits l'ont été pour le *Zaire ebolavirus* et plus spécifiquement les souches *Makona* et *Kikwit* [16,22]. En ce qui concerne les autres souches d'*ebolavirus* (par exemple *Soudan* ou *Bundibugyo*), les

connaissances et traitements disponibles sont encore plus limités. Par exemple, dans l'essai mené en République Démocratique du Congo pendant une épidémie de *Zaire ebolavirus* évaluant un traitement antiviral (remdesivir), un anticorps monoclonal (Mab114) et un cocktail d'anticorps monoclonaux (REGN-EB3) en comparaison avec le ZMapp, ces trois derniers (Mab114, REGN-EB3 et ZMapp) ont un spectre d'action restreint et sont potentiellement inefficaces contre d'autres souches d'*ebolavirus* [22].

La principale conclusion de cette première partie de thèse est qu'une grande partie des publications identifiées via notre revue systématique ont un risque important de biais et que peu d'essais contrôlés randomisés ont été publiés entraînant un manque de preuve robuste sur les traitements évalués.

4.2. Comment prendre en compte le contexte des FHV lors du design d'un essai clinique ?

La BDA est une méthode qui permet de prendre en compte le fardeau de la maladie et les préférences patients au niveau du bénéfice-risque de recevoir un nouveau traitement expérimental, en tant qu'information *a priori* lors du design d'un essai clinique fréquentiste [26–28]. Fardeau de la maladie, mortalité, séquelles et effets indésirables du traitement sont incorporés dans une matrice de coûts qui est ensuite utilisée via une fonction de coût pour calculer une erreur de type I et un nombre de sujets à inclure optimaux.

Nous avons voulu appliquer le modèle BDA dans un contexte de FHV pour le design d'un essai clinique évaluant un traitement sur la mortalité à un temps donné. Etant donné que la BDA a précédemment été développée pour des critères de jugement continus et censurés [26–28], nous avons adapté ce modèle aux critères de jugement binaires. Pour cela, il nous fallait proposer des formules de coût adéquates pour prendre en compte la nature binaire de notre critère de jugement. Nous avons également étudié, avec l'aide de l'équipe du professeur Andrew Lo, les types de coûts (post-essai et pendant l'essai) qu'il nous fallait considérer dans ce contexte. Nous avons ensuite réfléchi à la spécification des différents paramètres du modèle (taille de la population cible, mortalité de la maladie, taille d'effet attendu du traitement évalué, probabilité que le traitement ne soit pas efficace, effets indésirables du traitement, pourcentage de séquelles et leur fardeau). Pour finir, nous avons développé une fonction R prenant en entrée ces différents paramètres et donnant en sortie les valeurs optimales de la taille d'échantillon, de la valeur critique et de l'erreur de type I. Nous avons ensuite appliqué notre nouveau modèle au cas de la maladie à virus Ebola.

Une partie des paramètres (le coût des effets indésirables du traitement et le poids du handicap lié aux séquelles de la maladie) a pu être spécifiée grâce aux différentes études du Global Burden of Disease [70,71]. Via ces études, nous avons pu utiliser des valeurs assez générales de ces paramètres qui peuvent être considérées pour d'autres maladies au-delà de notre illustration sur la maladie à virus Ebola. En effet, nous avons utilisé un coût global pour les effets indésirables du traitement (toujours en termes d'impact sur la santé des personnes), qui n'est pas spécifique à un traitement ou à un pays en particulier [70]. En ce qui concerne le poids du handicap lié aux séquelles de la maladie, nous avons utilisé celui calculé par le Global Burden of Disease pour les conséquences post-phase aiguë des maladies infectieuses [71]. Pour les autres paramètres du modèle, ils sont dépendants du contexte de la maladie dans lequel nous nous trouvons. Pour notre illustration, nous avons utilisé les données recueillies sur la maladie à virus Ebola de notre revue systématique.

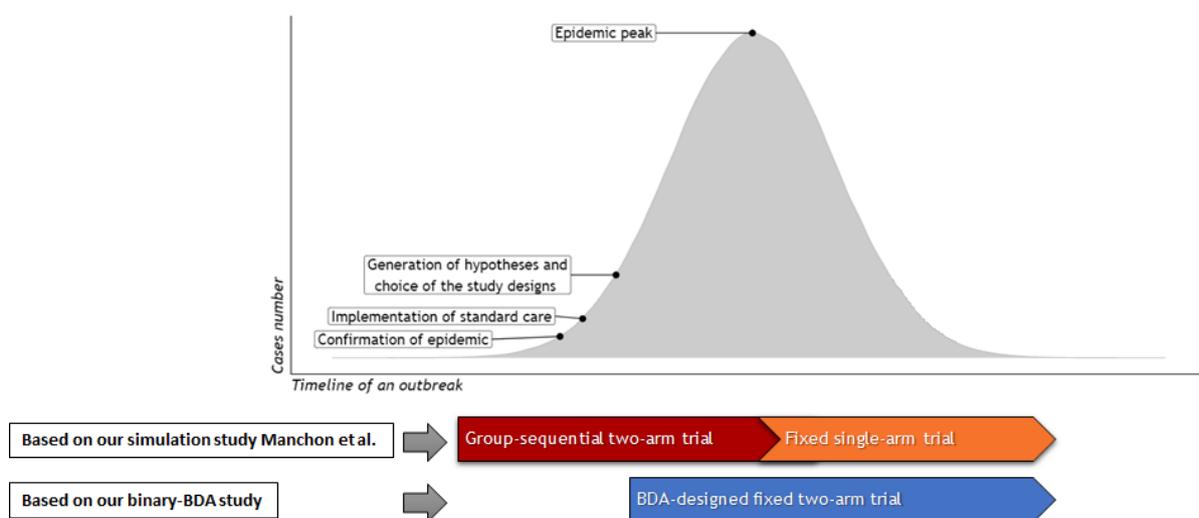
Nous proposons ainsi un outil permettant de calculer une erreur de type I et un nombre de sujets à inclure en prenant en compte les informations connues de la pathologie. Cela peut permettre dans certaines conditions de réduire le nombre de sujets à inclure pour évaluer un traitement potentiellement bénéfique. Cette méthode a l'avantage de calculer de façon objective et reproductible une erreur de type I qui sera utilisée pour le design d'un essai clinique adapté au contexte de la maladie. A travers notre illustration, nous avons montré l'influence de la taille de la population cible, de la mortalité et de l'effet attendu du traitement sur l'erreur de type I et le nombre de sujets à inclure dans le cadre d'une FHV.

Cependant, une limite de cette méthode est qu'elle se base sur des données précédemment publiées de taille de population cible et de mortalité. Or, dans un contexte de maladies infectieuses, ces nombres peuvent varier grandement au fil des années et le nombre de cas dans les futures épidémies est difficilement prédictible. Nous avons essayé de prendre en compte cette limite dans notre illustration en considérant un large éventail de scénarios possibles pour le cas de la maladie à virus Ebola d'après les dernières épidémies rapportées de la maladie. Ainsi, les divers scénarios nous donnent une vue d'ensemble des cas auxquels nous pourrions faire face lors de futures épidémies.

Lors du design d'un essai clinique, ici nous nous intéressons au cas des FHVs, notre méthode permettrait de prendre en compte les dernières informations rapportées selon le moment où l'on se trouve dans l'épidémie, avec un nombre de cas potentiel plus élevé avant le pic épidémique et plus faible après ce pic. Dans notre précédente étude de simulation, nous

avons montré qu'avant le pic épidémique un essai séquentiel double-bras devrait être privilégié, alors qu'après le pic épidémique un essai fixe simple-bras serait potentiellement préférable [23]. Avec cette nouvelle adaptation de la BDA que nous avons développée, nous pouvons maintenant considérer le fait de conduire un essai fixe double-bras optimisé par BDA quel que soit le moment de l'épidémie (Figure 4). Notamment, si l'on se trouve après le pic épidémique, en calculant une erreur de type I optimale, l'approche par BDA peut nous permettre d'éviter de conduire un essai fixe simple-bras mais plutôt un essai fixe double-bras optimisé par BDA. Ainsi, la qualité de l'évaluation du traitement expérimental en serait améliorée et la faisabilité de l'essai clinique serait augmentée.

Figure 4 Chronologie d'une épidémie de FHV avec les recommandations de Manchon et al. [23] et de notre étude BDA



Une perspective à ce travail serait de considérer les différents cas étudiés dans notre étude de simulation décrits en section 1.4 pour comparer un essai fixe double-bras optimisé par BDA aux autres types d'essais considérés (essai simple-bras versus essai double-bras ; essai fixe versus essai séquentiel par groupes). Il serait notamment intéressant de voir l'impact du cas « évoluant au cours du temps », où la mortalité diminue au cours de l'épidémie, sur un essai optimisé par BDA. Nous avons déjà discuté le problème que cela pose dans le cas d'un essai fixe simple-bras où l'hypothèse initiale faite sur la proportion historique utilisée comme comparateur est par conséquent incorrecte. Etant donné que nous utilisons les dernières informations disponibles avant le début de l'essai pour la mortalité dans le calcul de la BDA, on peut se demander quel sera l'impact d'une mortalité qui diminue au cours de l'épidémie sur un essai fixe double-bras optimisé par BDA.

Un autre point de discussion est que dans notre méthode BDA nous avons décidé de fixer la probabilité que le traitement soit efficace à 0.50. Nous nous sommes appuyés sur l'étude d'Isakov *et al.* qui considérait qu'en fixant à 50% le pourcentage de chance d'avoir un traitement efficace, on suppose un *a priori* non-informatif sur l'efficacité du traitement de telle sorte à suivre le principe d'« equivoque » pour des essais double-bras [26]. Cela reflète l'incertitude de la communauté médicale sur le traitement à préférer et donc confirme le besoin même de conduire un essai clinique pour évaluer ce traitement [72]. Dans notre cas, nous nous sommes concentrés sur un essai fixe double-bras fréquentiste. Une perspective intéressante serait de considérer un design bayésien afin d'attribuer une distribution *a priori* sur cette probabilité que le traitement soit efficace.

Une autre perspective à ce travail serait de considérer des designs adaptatifs, où l'on peut arrêter précocement l'essai si le traitement évalué est jugé futile, délétère ou efficace lors d'une analyse intermédiaire (par exemple un essai séquentiel par groupes). De plus, la littérature récente montre un intérêt grandissant pour les essais plateforme multi-bras multi-étapes où l'on peut arrêter un traitement ou ajouter de nouveaux traitements expérimentaux en cours d'essai. Il serait donc intéressant de combiner la méthodologie de la BDA avec ce type de designs.

A la suite de cette thèse, nous souhaitons développer un package R qui serait disponible sur le CRAN (Comprehensive R Archive Network) pour permettre aux futurs chercheurs de calculer des erreurs de type I et tailles d'échantillons optimales selon la BDA. Nous envisageons de faire cela en collaboration avec l'équipe du professeur Andrew Lo du MIT et ainsi proposer un package qui contiendrait la BDA pour des critères de jugement binaires mais également continus.

En conclusion, à travers notre revue systématique et notre approche de la BDA pour des critères de jugement binaires, nous avons développé une nouvelle méthodologie qui pourra permettre à des chercheurs de prendre en compte le contexte des FHV lors de la conception d'essais cliniques pour évaluer des traitements. Cela permettra ainsi de produire des résultats robustes pour des traitements contre des FHV potentiellement très mortelles avec peu de traitements disponibles.

5. Bibliographie

1. McFee RB. Viral hemorrhagic fever viruses. *Dis Mon.* 2013;59(12):410-25.
2. Fhogartaigh CN, Aarons E. Viral haemorrhagic fever. *Clin Med (Lond).* 2015;15(1):61.
3. Organisation mondiale de la Santé - Bureau régional de Méditerranée orientale. Fièvres hémorragiques virales [Internet]. [consulté le 18 juin 2023]. Disponible sur: <https://www.emro.who.int/fr/health-topics/haemorrhagic-fevers-viral/introduction.html>
4. Flórez-Álvarez L, de Souza EE, Botosso VF, de Oliveira DBL, Ho PL, Taborda CP, et al. Hemorrhagic fever viruses: Pathogenesis, therapeutics, and emerging and re-emerging potential. *Front Microbiol.* 2022;13.
5. Belhadi D, El Baied M, Mulier G, Malvy D, Mentré F, Laouénan C. The number of cases, mortality and treatments of viral hemorrhagic fevers: A systematic review. *PLoS Negl Trop Dis.* oct 2022;16(10):e0010889.
6. Etard JF, Sow MS, Leroy S, Touré A, Taverne B, Keita AK, et al. Multidisciplinary assessment of post-Ebola sequelae in Guinea (Postebogui): an observational cohort study. *Lancet Infect Dis.* mai 2017;17(5):545-52.
7. Tozay S, Fischer WA, Wohl DA, Kilpatrick K, Zou F, Reeves E, et al. Long-Term Complications of Ebola Virus Disease: Prevalence and Predictors of Major Symptoms and the Role of Inflammation. *Clin Infect Dis.* 6 nov 2019;ciz1062.
8. World Health Organization (14 January 2023). Disease Outbreak News; Ebola disease caused by Sudan ebolavirus – Uganda [Internet]. [consulté le 18 juin 2023]. Disponible sur: <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON433>
9. Liu CH, Hu YT, Wong SH, Lin LT. Therapeutic Strategies against Ebola Virus Infection. *Viruses.* 11 mars 2022;14(3):579.
10. World Health Organization (9 June 2023). Disease Outbreak News; Marburg virus disease - Equatorial Guinea [Internet]. [consulté le 18 juin 2023]. Disponible sur: <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON472>
11. World Health Organization (2 June 2023). Disease Outbreak News; Marburg virus disease - the United Republic of Tanzania. [Internet]. [consulté le 18 juin 2023]. Disponible sur: <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON471>
12. Sah R, Mohanty A, Reda A, Siddiq A, Mohapatra RK, Dhama K. Marburg virus re-emerged in 2022: recently detected in Ghana, another zoonotic pathogen coming up amid rising cases of Monkeypox and ongoing COVID-19 pandemic- global health concerns and counteracting measures. *Vet Q.* 31 déc 2022;42(1):167-71.
13. World Health Organization (1 May 2023). Disease Outbreak News; Lassa Fever – Nigeria [Internet]. [consulté le 18 juin 2023]. Disponible sur: <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON463>
14. Cheng HY, French CE, Salam AP, Dawson S, McAleenan A, McGuinness LA, et al. Lack of Evidence for Ribavirin Treatment of Lassa Fever in Systematic Review of Published and Unpublished Studies. *Emerg Infect Dis.* août 2022;28(8):1559-68.

15. World Health Organization. WHO Workshop - « Efficacy trials of Lassa Therapeutics: endpoints, trial design, site selection » [Internet]. 2018 avr [consulté le 13 juill 2023]. (WHO R&D blueprint). Disponible sur: <https://www.who.int/docs/default-source/blueprint/lassa-therapeutics-workshop-april-25-2018.pdf>
16. PREVAIL II Writing Group, Multi-National PREVAIL II Study Team, Davey RT Jr, Dodd L, Proschan MA, Neaton J, et al. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J Med*. 13 oct 2016;375(15):1448-56.
17. Sissoko D., Laouenan C., Folkesson E., M'Lebing A.-B., Beavogui A.-H., Baize S., et al. Experimental Treatment with Favipiravir for Ebola Virus Disease (the JIKI Trial): A Historically Controlled, Single-Arm Proof-of-Concept Trial in Guinea. *PLoS Med*. 2016;13(3).
18. van Griensven J, Edwards T, de Lamballerie X, Semple MG, Gallian P, Baize S, et al. Evaluation of convalescent plasma for Ebola virus disease in Guinea. *N Engl J Med*. janv 2016;374(1):33-42.
19. Dunning J, Kennedy SB, Antierens A, Whitehead J, Ciglenecki I, Carson G, et al. Experimental Treatment of Ebola Virus Disease with Brincidofovir. *PLoS One*. 9 sept 2016;11(9):e0162199-e0162199.
20. Dunning J, Sahr F, Rojek A, Gannon F, Carson G, Idriss B, et al. Experimental treatment of Ebola virus disease with TKM-130803: a single-arm phase 2 clinical trial. *PLoS Med*. 2016;13(4):e1001997.
21. Nguyen THT, Guedj J, Anglaret X, Laouénan C, Madelain V, Taburet AM, et al. Favipiravir pharmacokinetics in Ebola-Infected patients of the JIKI trial reveals concentrations lower than targeted. *PLoS Negl Trop Dis*. 23 févr 2017;11(2):e0005389-e0005389.
22. Mulangu S, Dodd LE, Davey RT Jr, Tshiani Mbaya O, Proschan M, Mukadi D, et al. A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *N Engl J Med*. 12 déc 2019;381(24):2293-303.
23. Manchon P, Belhadi D, Mentré F, Laouénan C. Evaluation by simulation of clinical trial designs for evaluation of treatment during a viral haemorrhagic fever outbreak. *BMC Med Res Methodol*. 6 mai 2021;21(1):98.
24. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567(7748):305-7.
25. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. Vol. 70, *The American Statistician*. Taylor & Francis; 2016.
26. Isakov L, Lo AW, Montazerhodjat V. Is the FDA too conservative or too aggressive?: A Bayesian decision analysis of clinical trial design. *J Econom*. 2019;211(1):117-36.
27. Xu Q, Cho J, Ben Chaouch Z, Lo AW. Incorporating patient preferences and burden-of-disease in evaluating ALS drug candidate AMX0035: a Bayesian decision analysis perspective. *Amyotroph Lateral Scler Frontotemporal Degener*. 2023;24(3-4):281-8.

28. Montazerhodjat V, Chaudhuri SE, Sargent DJ, Lo AW. Use of Bayesian decision analysis to minimize harm in patient-centered randomized clinical trials in oncology. *JAMA Oncol.* 2017;3(9):e170123-e170123.
29. Paganoni S, Macklin EA, Hendrix S, Berry JD, Elliott MA, Maiser S, et al. Trial of Sodium Phenylbutyrate–Taurursodiol for Amyotrophic Lateral Sclerosis. *N Engl J Med.* 3 sept 2020;383(10):919-30.
30. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev.* 2015;4(1):1-9.
31. CDC. Diseases Caused by High-Consequence Viruses [Internet]. [consulté le 1 déc 2019]. Disponible sur: <https://www.cdc.gov/ncezid/dhcpp/vspb/diseases.html>
32. Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Ottawa Hospital Research Institute. 2011;2(1):1-12.
33. Higgins JP, Savović J, Page MJ, Elbers RG, Sterne JA. Assessing risk of bias in a randomized trial. *Cochrane handbook for systematic reviews of interventions.* 2019;205-28.
34. De St Maurice A., Ervin E., Orone R., Choi M., Dokubo E.K., Rollin P.E., et al. Care of Ebola survivors and factors associated with clinical sequelae-Monrovia, Liberia. *Open Forum Infect Dis.* 2018;5(10).
35. Tiffany A., Vetter P., Mattia J., Dayer J.-A., Bartsch M., Kasztura M., et al. Ebola virus disease complications as experienced by survivors in Sierra Leone. *Clin Infect Dis.* 2016;62(11):1360-6.
36. Toure A., March L., Sow M.S., Etard J.F., Kpamou C., Savane I., et al. Long term evolution of the clinical complications of Ebola virus disease: Results of the Postebogui cohort. *Trop Med Int Health.* 2017;22:190.
37. Mwanza A.K., Mukadi D., Mukadi P., Hoff N.A., Doshi R.H., Wasiswa J., et al. Clinical characteristics of ebola survivors 40 years post infection in the democratic republic of congo. *Am J Trop Med Hyg.* 2016;95(5):338.
38. Billieux B.J., Mith B., Bowen L., Schindler M., Azodi S., Ohayon J., et al. Longitudinal cohort study of neurological sequelae in ebola virus disease survivors in liberia. *J Neurol Sci.* 2017;381:91-2.
39. Clark D.V., Kibuuka H., Millard M., Wakabi S., Lukwago L., Taylor A., et al. Long-term sequelae after Ebola virus disease in Bundibugyo, Uganda: A retrospective cohort study. *Lancet Infect Dis.* 2015;15(8):905-12.
40. Pearson D., Davis L., Frankhauser J., Caplan L. Musculoskeletal manifestations in ebola survivors. *Arthritis Rheum.* 2016;68:1664-5.
41. Godwin C.L., Wohl D.A., Fischer 2nd W.A., Singh K., Hawks D.A., Devore E.E., et al. Reproductive health sequelae among women who survived Ebola virus disease in Liberia. *Int J Gynecol Obstet.* 2019;146(2):212-7.

42. PREVAIL III Study Group, Sneller MC, Reilly C, Badio M, Bishop RJ, Eghrari AO, et al. A Longitudinal Study of Ebola Sequelae in Liberia. *N Engl J Med.* 7 mars 2019;380(10):924-34.
43. Wilson HW, Amo-Addae M, Kenu E, Ilesanmi OS, Ameme DK, Sackey SO. Post-Ebola Syndrome among Ebola Virus Disease Survivors in Montserrado County, Liberia 2016. *Biomed Res Int.* 28 juin 2018;2018:1909410-1909410.
44. Kelly JD, Hoff NA, Spencer D, Musene K, Bramble MS, McIlwain D, et al. Neurological, Cognitive, and Psychological Findings Among Survivors of Ebola Virus Disease From the 1995 Ebola Outbreak in Kikwit, Democratic Republic of Congo: A Cross-sectional Study. *Clin Infect Dis.* 8 avr 2019;68(8):1388-93.
45. Etard J.-F., Sow M.S., Leroy S., Msellati P., Toure A.I., Taverne B., et al. Sequelae of ebola virus disease in surviving patients in guinea: Postebogui cohort. *Top Antiviral Med.* 2016;24(E-1):30.
46. Mattia JG, Vandy MJ, Chang JC, Platt DE, Dierberg K, Bausch DG, et al. Early clinical sequelae of Ebola virus disease in Sierra Leone: a cross-sectional study. *Lancet Infect Dis.* mars 2016;16(3):331-8.
47. Keita M, Diallo B, Mesfin S, Marega A, Nebie KY, Magassouba N, et al. Subsequent mortality in survivors of Ebola virus disease in Guinea: a nationwide retrospective cohort study. *Lancet Infect Dis.* nov 2019;19(11):1202-8.
48. Hereth-Hebert E, Bah MO, Etard JF, Sow MS, Resnikoff S, Fardeau C, et al. Ocular Complications in Survivors of the Ebola Outbreak in Guinea. *Am J Ophthalmol.* mars 2017;175:114-21.
49. Shantha J.G., Weil N., Miller C., Bastien G., Yeh S. Ocular complications in IN pediatric ebola survivors. *J Women's Health.* 2018;27(11):1428-9.
50. Fallah M. A cohort study of survivors of ebola virus infection in liberia (prevail III). *Top Antiviral Med.* 2016;24(E-1):30-1.
51. Pers YM, Sow MS, Taverne B, March L, Izard S, Étard JF, et al. Characteristics of the musculoskeletal symptoms observed among survivors of Ebola virus disease in the Postebogui cohort in Guinea. *Rheumatology (Oxford).* 1 déc 2017;56(12):2068-72.
52. Vandy M.J., Chang J., Mattia J.G., Molleh B., Charles D.B., Dierberg K., et al. Burden, timing, and outcomes of uveitis from a retrospective cohort of Ebola survivors in Sierra Leone. *Invest Ophthalmol Vis Sci.* 2018;59(9).
53. Howlett P., Walder A., Lisk D., N'Jai A., Lado M., Brown C., et al. Neurological and psychiatric manifestations of post Ebola syndrome in Sierra Leone. *Lancet.* 2017;389:S48.
54. Pers Y.-M., Dubois A., Barry A., Sall M., Saliou Sow M., Taverne B., et al. Long-term rheumatic and musculoskeletal disorders associated to ebola virus infection. *Arthritis Rheum.* 2018;70:1059-60.
55. Jagadesh S, Sevalie S, Fatoma R, Sesay F, Sahr F, Faragher B, et al. Disability Among Ebola Survivors and Their Close Contacts in Sierra Leone: A Retrospective Case-Controlled Cohort Study. *Clin Infect Dis.* 6 janv 2018;66(1):131-3.

56. Steptoe PJ, Scott JT, Baxter JM, Parkes CK, Dwivedi R, Czanner G, et al. Novel Retinal Lesion in Ebola Survivors, Sierra Leone, 2016. *Emerg Infect Dis.* juill 2017;23(7):1102-9.
57. Scott JT, Sesay FR, Massaquoi TA, Idriss BR, Sahr F, Semple MG. Post-Ebola Syndrome, Sierra Leone. *Emerg Infect Dis.* avr 2016;22(4):641-6.
58. Guetiya Wadoum RE, Samin A, Mafopa NG, Giovanetti M, Russo G, Turay P, et al. Mobile health clinic for the medical management of clinical sequelae experienced by survivors of the 2013-2016 Ebola virus disease outbreak in Sierra Leone, West Africa. *Eur J Clin Microbiol Infect Dis.* nov 2017;36(11):2193-200.
59. Mohammed H, Vandy AO, Stretch R, Otieno D, Prajapati M, Calderon M, et al. Sequelae and Other Conditions in Ebola Virus Disease Survivors, Sierra Leone, 2015. *Emerg Infect Dis.* janv 2017;23(1):66-73.
60. Stojanovic M, Pekic S, Cvijovic G, Miljic D, Doknic M, Nikolic-Djurovic M, et al. High risk of hypopituitarism in patients who recovered from hemorrhagic fever with renal syndrome. *J Clin Endocrinol Metab.* juill 2008;93(7):2722-8.
61. Mäkelä S, Jaatinen P, Miettinen M, Salmi J, Ala-Houhala I, Huhtala H, et al. Hormonal deficiencies during and after Puumala hantavirus infection. *Eur J Clin Microbiol Infect Dis.* juin 2010;29(6):705-13.
62. Miettinen MH, Mäkelä SM, Ala-Houhala IO, Huhtala HSA, Kööbi T, Vaheri AI, et al. Ten-year prognosis of Puumala hantavirus-induced acute interstitial nephritis. *Kidney Int.* juin 2006;69(11):2043-8.
63. Miettinen MH, Makela SM, Ala-Houhala IO, Huhtala HS, Koobi T, Vaheri AI, et al. Tubular proteinuria and glomerular filtration 6 years after puumala hantavirus-induced acute interstitial nephritis. *Nephron Clin Pract.* 2009;112(2):c115-20.
64. Mathes RW, Page WF, Crawford HM, McBean AM, Miller RN. Long-term sequelae of hemorrhagic fever with renal syndrome attributable to hantaan virus in Korean War veterans. *Mil Med.* avr 2005;170(4):315-9.
65. Gracia F, Armien B, Simpson SQ, Munoz C, Broce C, Pascale JM, et al. Convalescent pulmonary dysfunction following hantavirus pulmonary syndrome in Panama and the United States. *Lung.* oct 2010;188(5):387-91.
66. Pergam SA, Schmidt DW, Nofchissey RA, Hunt WC, Harford AH, Goade DE. Potential renal sequelae in survivors of hantavirus cardiopulmonary syndrome. *Am J Trop Med Hyg.* févr 2009;80(2):279-85.
67. Wilson M.R., Peters C.J. Diseases of the central nervous system caused by lymphocytic choriomeningitis virus and other arenaviruses. *Handb Clin Neurol.* 2014;123:671-81.
68. World Health Organization. Introduction to Lassa fever - Managing infectious hazards [Internet]. 2019 déc [consulté le 24 juill 2023]. Disponible sur: <https://www.who.int/publications/i/item/introduction-to-lassa-fever>
69. Gul S, Gul EU, Yesilyurt M, Ozturk B, Kuscu F, Ergonul O. Health-related quality of life and the prevalence of post-traumatic stress disorder among Crimean-Congo hemorrhagic fever survivors. *Jpn J Infect Dis.* 2012;65(5):392-5.

70. Institute for Health Metrics and Evaluation. Global Burden of Disease 2017 [Internet]. [consulté le 5 févr 2023]. Disponible sur: <https://gbd2017.healthdata.org/gbd-results/>
71. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2129-43.
72. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 16 juill 1987;317(3):141-5.

RESEARCH ARTICLE

Open Access

Evaluation by simulation of clinical trial designs for evaluation of treatment during a viral haemorrhagic fever outbreak

Pauline Manchon^{1,2*†} , Drifa Belhadi^{1,2,3†}, France Mentré^{1,2,3} and Cédric Laouénan^{1,2,3}

Abstract

Background: Viral haemorrhagic fevers are characterized by irregular outbreaks with high mortality rate. Difficulties arise when implementing therapeutic trials in this context. The outbreak duration is hard to predict and can be short compared to delays of trial launch and number of subject needed (NSN) recruitment. Our objectives were to compare, using clinical trial simulation, different trial designs for experimental treatment evaluation in various outbreak scenarios.

Methods: Four type of designs were compared: fixed or group-sequential, each being single- or two-arm. The primary outcome was 14-day survival rate. For single-arm designs, results were compared to a pre-trial historical survival rate p_H . Treatments efficacy was evaluated by one-sided tests of proportion (fixed designs) and Whitehead triangular tests (group-sequential designs) with type-I-error = 0.025. Both survival rates in the control arm p_C and survival rate differences Δ (including 0) varied. Three specific cases were considered: “standard” (fixed p_C , reaching NSN for fixed designs and maximum sample size N_{Max} for group-sequential designs); “changing with time” (increased p_C over time); “stopping of recruitment” (epidemic ends). We calculated the proportion of simulated trials showing treatment efficacy, with $K = 93,639$ simulated trials to get a type-I-error $PI_{95\%}$ of [0.024;0.026].

Results: Under H_0 ($\Delta = 0$), for the “standard” case, the type-I-error was maintained regardless of trial designs. For “changing with time” case, when $p_C > p_H$, type-I-error was inflated, and when $p_C < p_H$ it decreased. Wrong conclusions were more often observed for single-arm designs due to an increase of Δ over time. Under H_1 ($\Delta = +0.2$), for the “standard” case, the power was similar between single- and two-arm designs when $p_C = p_H$. For “stopping of recruitment” case, single-arm performed better than two-arm designs, and fixed designs reported higher power than group-sequential designs. A web R-Shiny application was developed.

(Continued on next page)

* Correspondence: pauline.manchon@aphp.fr

†Pauline Manchon and Drifa Belhadi contributed equally to this work.

¹INSERM, Centre d'Investigation clinique-Epidémiologie Clinique 1425, Hôpital Bichat, F-75018 Paris, France

²Département Epidémiologie Biostatistiques et Recherche Clinique, AP-HP, Hôpital Bichat, F-75018 Paris, France

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Conclusions: At an outbreak beginning, group-sequential two-arm trials should be preferred, as the infected cases number increases allowing to conduct a strong randomized control trial. Group-sequential designs allow early termination of trials in cases of harmful experimental treatment. After the epidemic peak, fixed single-arm design should be preferred, as the cases number decreases but this assumes a high level of confidence on the pre-trial historical survival rate.

Keywords: Viral haemorrhagic fever outbreak, Clinical trial design, Simulation study

Background

Emerging infectious diseases such as influenza, cholera, coronavirus, varicella, meningitis and viral haemorrhagic fevers recently caused numerous outbreaks [1]. Since 2003, the coronavirus caused various outbreaks around the world. SARS-CoV-1 was first observed in Southeast Asia. In 2012, MERS-CoV caused a case-fatality rate of 37% in the Middle East [2]. Today SARS-CoV-2 is pandemic. Viral haemorrhagic fevers are among the most severe, including Crimean–Congo haemorrhagic fever, Ebola virus disease, Lassa fever and Marburg haemorrhagic fever, which are endemic in some areas of Africa, South America and Asia [3]. They are characterized by outbreaks with high mortality rate which occur irregularly and are hard to anticipate. The second largest outbreak of Ebola virus disease since the 2014–2016 one in West Africa, began in 2018 and ended in June 2020 in the Democratic Republic of the Congo [4].

Recently, a scoping review evaluated political, economic, administrative, regulatory, logistical, ethical and social (PEARLES) challenges associated with clinical research in the context of emergency epidemics [5]. The authors highlighted the challenges associated with the planning, conduct and dissemination of clinical research responses during an epidemic. They stressed the need for developing solutions to improve rapid clinical research deployment, delivery, and dissemination for future epidemics.

Difficulties often occur when implementing therapeutic trials in the context of viral haemorrhagic fever outbreak. Indeed, the duration of the outbreak is hard to predict and can be very short compared to the necessary delay to launch a trial and recruit the required number of patients. A reduced number of included patients impact the statistical power of the trial and can lead to study which cannot demonstrate a significant effect of their experimental treatment [6–8]. Evolution of the number of cases and case fatality rate during an outbreak period is also a dimension to be considered when designing a clinical trial in this context.

Moreover, the design of the study must be acceptable on the field by patients, healthcare workers and the global population. In the situation of an epidemic peak, local authorities and non-governmental organizations

managing treatment centres may argue that randomization is hard to implement due to the high case-fatality rate and the reluctance to give a less beneficial treatment to part of the patients [7, 9, 10]. A single-arm design may seem to be best suited for those ethical aspects, its feasibility on the field and its acceptability by the population affected directly or indirectly by the disease. Nevertheless, non-comparative trials can lead to substantially biased results if the pre-trial survival rate is incorrect, indeed its evolution over time is not taken into account.

Another aspect to consider is the choice of the design: fixed or group-sequential. Indeed, group-sequential designs sometimes lead to reducing the duration of the trial especially in case of early stop for futility or efficacy of the experimental treatment [11]. At each interim analysis, the efficacy of the treatment is tested to decide if the recruitment has to continue or if the experimental treatment is already shown to be effective or not. Group-sequential designs are interesting in an epidemic peak of infectious disease, under emergency conditions and with different treatment candidates.

During the Ebola outbreak in West Africa, between December 2013 and May 2016, three antivirals (favipiravir, brincidofovir, TKM-13083), one cocktail of antibodies, and convalescent plasma were evaluated in five therapeutic trials in Guinea, Liberia and Sierra Leone [6–8, 12, 13]. These trials were implemented during the second phase of the outbreak, when the number of cases decreased, and none of them achieved the number of patients required. For four trials, the primary outcome was the mortality at 14 days after the inclusion. This outcome was chosen because it requires no invasive procedures or special equipment, and most of the time deaths from Ebola occur within 14 days. Only one used the mortality at 28 days and it was the only randomized trial with an adaptive scheme [6]. The four other trials were non-randomized and the comparison was established with the pre-trial mortality based on historical data, and three trials used a group-sequential design [7, 12, 13]. More recently, a multi-arm multi-stage clinical trial was conducted with four investigational therapies for Ebola virus disease in the Democratic Republic of Congo, where, two of the four treatments compared were stopped based on interim analysis [14].

A simulation study by Cooper et al. [15] was conducted to investigate the choice of the study design in the context of the Ebola Virus disease. In particular, they worked on a multi-stage approach (MSA) comprising a single-arm phase II study followed by one or two phase III studies. They concluded that the MSA and group-sequential double-arm randomised trial led to substantially fewer deaths than a conventional two-arm randomised trial if the tested interventions were either highly effective or harmful. MSA was applied to design two clinical trials during the 2014–2015 West African Ebola outbreak [16].

The objectives of the present clinical trial simulation study was to compare different designs, during various outbreak scenarios to develop recommendations when designing a clinical trial for viral haemorrhagic fever. We considered several specific cases associated with the timeline of an outbreak. Indeed, at the beginning of an outbreak, standard cares are developed, which lead to increased survival rates. After the epidemic peak, it is more difficult to include patients in a trial due to a decreased number of cases. This crisis situation requires an accelerated research process and the study feasibility depends on the choice of the trial design. For this simulation study, we choose to compare four clinical trial designs: fixed single-arm, group-sequential single-arm, fixed two-arm and group-sequential two-arm designs. An online tool based on the results from this simulation study was also developed.

Methods

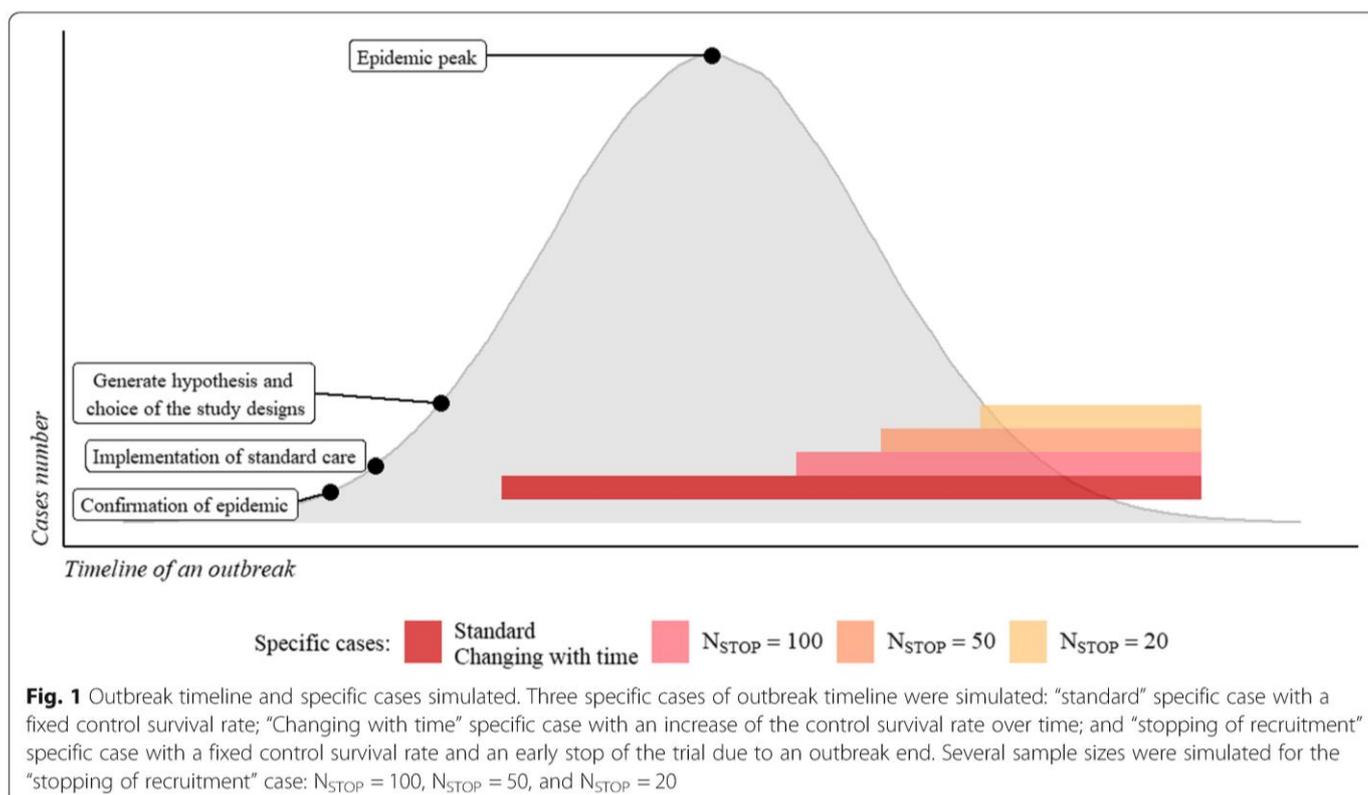
We performed clinical trial simulations to compare four designs (see Additional file 1): fixed single-arm trial (F1); group-sequential single-arm trial (S1); fixed two-arm trial (F2); and group-sequential two-arm trial (S2). A randomization 1:1 was considered for two-arm designs. The primary outcome of each simulated trial was the number of patients who survived at day 14 after inclusion. For single-arm designs, results were compared to a pre-trial historical survival rate.

Several outbreak scenarios were explored based on three scenario for the timeline of the outbreak and various values of control survival rate (p_C) and efficacy of the treatment (Δ). A scenario was defined as the combination of one specific case of outbreak timeline, one control survival rate and one survival rate difference.

Three specific cases of outbreak timeline (Fig. 1)

The “standard” case corresponded to a trial launch at the beginning of the outbreak, i.e. a fixed p_C and a number of subjects equal to the number of subjects needed (NSN) for fixed designs and the possibility to reach the maximum sample size (N_{Max}), if needed for group-sequential trials.

The “changing with time” case was defined to mimic increased proportion of patients who survived in the control group over time during the outbreak (p_C). An increase of +0.03 every 20 patients included was simulated, with a maximal total increase of 0.10. This scenario



was extrapolated from the real weekly evolution of the survival rate during the 2014's Ebola epidemic in West Africa [17].

The "stopping of recruitment" case was performed to mimic the effect of starting recruitment at different times during the time course of the outbreak after the epidemic peak. Several sample sizes were simulated: $N_{STOP} = 20$, $N_{STOP} = 50$, and $N_{STOP} = 100$, and we assumed a fixed p_C not changing with time. If required, the p -values in group-sequential designs were adjusted for underrunning (group-sequential trial which ended before that the stopping rule has been fulfilled) [18].

Statistical tests

For fixed designs (F1 and F2), the treatment efficacy was evaluated by a one-sided comparison test of proportions. A Yate's continuity correction was applied if needed. The experimental treatment was considered significantly superior if the p -value was inferior to 0.025 with the power $1-\beta = 0.90$. For group-sequential designs (S1 and S2), Whitehead triangular one-sided tests were conducted [19]. Stopping boundaries to conclude to the treatment efficacy were calculated using a type-I-error $\alpha = 0.025$ and a power $1-\beta = 0.90$. Group-sequential analyses were conducted every 20 patients included for both single and two-arm trials; and at each interim analysis the statistics Z (cumulative benefit of experimental treatment) and V (global information) from the Whitehead triangular test were calculated. The stopping rules for an interim analysis j were then defined as follows:

- If $Z_j \geq a + cV_j - 0.583\sqrt{V_j - V_{j-1}}$, the trial was stopped as the experimental treatment was shown to be significantly better than the control;
- If $Z_j \leq -a + 3cV_j + 0.583\sqrt{V_j - V_{j-1}}$, the trial was stopped due to futility/inefficacy of the experimental treatment;
- Otherwise, the trial continued.

The stopping rules for the final analysis J were defined by:

- If $Z_J \geq a + cV_J - 0.583\sqrt{V_J - V_{J-1}}$, the experimental treatment was considered as significantly better than the control;
- Otherwise, efficacy was not shown.

In case of premature termination of the group-sequential trial:

- If premature stopping of the trial occurred due to external reasons, i.e. independently of the cumulative information and the observed difference in efficacy between treatments, the final p -values were adjusted

by the method proposed by Whitehead for underrunning [18],

- Otherwise, efficacy was not shown.

Sample sizes

The basic scenario was used to calculate the sample sizes. It assumed a "standard" case for the outbreak, a proportion of patient who survived at day 14 $p_C = 0.5$ (control survival rate) for two-arm trials, and $p_H = 0.5$ (pre-trial historical survival rate) for single-arm trials, and an expected survival rate difference $\Delta = +0.2$. The number of subjects needed (NSN) for each fixed design and the maximum sample size for each group-sequential design (N_{MAX}), with a power $1-\beta = 0.9$ and a one-sided type-I-error $\alpha = 0.025$, were calculated and were: $NSN_{F1} = 60$, $NSN_{F2} = 248$, $N_{MAX, S1} = 91$, $N_{MAX, S2} = 378$.

For group-sequential trials, when the final analysis was performed with a cumulated number of subjects included equal to N_{MAX} , the final analysis J was conducted with the remaining patients, necessarily inferior to 20 patients (number of patients included at each intermediate analysis).

Those sample sizes were used for all simulated scenario under "standard" or "changing with time" cases.

Simulated values of parameters

All specific cases of outbreak timeline and all designs were simulated with different values of the survival rate in the control arm $p_C = (0.35; 0.50; 0.75)$ and survival rate difference $\Delta = (-0.10; 0; 0.10; 0.20)$. Of note $\Delta = 0$ corresponds to simulation under the null hypothesis. The experimental survival rate in the treatment arm was defined as $p_E = p_C + \Delta$. When $p_C + \Delta \leq 0$ then p_E was set at 0.01 in order to be able to calculate the log-odds ratio for triangular tests. Likewise, p_E was set at 1 when $p_C + \Delta > 1$.

For single-arm trials the fixed pre-trial historical survival rate p_H was assumed to be 0.50 regardless of the scenario and control survival rate.

Design evaluation

For each design and each scenario, defined by the specific case of outbreak timeline and parameters p_C and Δ , $K = 93,639$ simulated trials were simulated in order to ensure a 95% prediction interval width of 0.001 around $\alpha = 0.025$. The seed was chosen randomly and was the same for all the scenarios. We set the seed once at the beginning of the entire set of simulated studies.

To evaluate the performance of each design, for each scenario, we calculated the proportion of simulated trials showing the efficacy of the experimental treatment i.e. the proportion of significant tests (denoted p), which corresponds to the type-I-error when $\Delta = 0$, and the median [with 5th and 95th percentiles] number of subjects

included in group-sequential trials (to be compared with the NSN of the fixed trials).

Simulations were performed with R software version 3.2.1.

Results

Scenarios under H_0 ($\Delta = 0$)

For the basic scenario ($p_C = 0.50$ and $\Delta = 0$), the type-I-error was maintained regardless of the trial design (Table 1). For the “changing with time” case, the type-I-error was calculated for one-arm designs exclusively: for two-arm designs, control and experimental survival rates increased together. Moreover, the type-I-error for single-arm designs F1 and S1 increased respectively to $p = 0.069$ and $p = 0.062$. This was due to the incorrect assumption on the pre-trial historical survival rate set at $p_H = 0.50$ whereas the survival rate increased during the trial conduct, leading to a p_H inferior to the control survival rate p_C . The assumption on the pre-trial historical survival rate is very important. Indeed, a wrong assumption on the value of the pre-trial historical survival rate (p_H) has considerable impact on the results. For the “standard” and “changing with time” cases, when $p_C = 0.75$, whereas we assumed $p_H = 0.50$, i.e. $p_H < p_C$, and $\Delta = 0$, single-arm trials (either fixed or group-sequential) still had higher proportions of simulated trials showing a significant efficacy than two-arm trials. On the contrary, when $p_H > p_C$, with $p_C = 0.35$ and $\Delta = 0$, proportions of simulated trials showing a significant efficacy was very close to zero. Of note, for the “changing with time” case, when $p_C = 0.75$ and $\Delta = 0$, the type-I-error was lower for group-sequential two-arm design ($p = 0.015$) than fixed two-arm design ($p = 0.025$). This was due to the presence of 11% of inconclusive trials for the group-sequential design (data not shown).

Regarding the “stopping of recruitment” case, the only scenario where the type-I-error was maintained was

observed under H_0 for the group-sequential single-arm design with $p_C = 0.50$, $\Delta = 0$ and $N_{STOP} = 100$, a sample size very close to the $N_{max,S1}$ (Table 2). When inclusions were poor and $p_C = 0.50$, $\Delta = 0$, the type-I-error was always lower. For this specific case with $p_C = 0.75$, $\Delta = 0$ and $N_{STOP} = 50$, wrong assumption on p_H had a stronger impact on the type-I-error estimated at 0.95 to 0.97 for single-arm designs.

Scenarios under H_1 ($\Delta = 0.2$)

For the “standard” case, proportions of simulated trials with a significant efficacy were similar ($p = 0.90$) between single- and two-arm designs under H_1 when the control survival rate was equal to the pre-trial historical survival rate ($p_C = p_H = 0.50$) and $\Delta = 0.2$ (Table 1). However, when the pre-trial historical survival rate was incorrect, i.e. the control survival rate was lower ($p_C = 0.35$) than the pre-trial historical survival rate, power of single-arm designs was poor. Indeed, only 11–12% of simulated single-arm trials were significant when $p_C = 0.35$ and $p_E = 0.55$ ($\Delta = 0.20$), versus 88–89% of the simulated two-arm trials. For “changing with time” cases, proportions were overall similar between fixed and group-sequential designs (Table 1). Regarding “stopping of recruitment” cases, single-arm trials performed better than two-arm trials (Table 2). When $p_C = 0.50$, the proportions of simulated trials showing a significant efficacy were higher for single-arm trials, ranging from 41% ($N_{STOP} = 20$) to 99% ($N_{STOP} = 100$), compared to two-arm trials (from 5% for $N_{STOP} = 20$ to 54% for $N_{STOP} = 100$). Moreover, fixed designs reported higher significant efficacy than group-sequential designs when $\Delta = 0.20$.

Scenarios with $\Delta = -0.1$

When the treatment was harmful ($\Delta = -0.1$), the proportion of simulated trials showing the efficacy of the

Table 1 Proportion of trials showing significant improvement with the experimental treatment (“standard” and “changing with time” cases)

	Design	$p_C = 0.50$				$p_C = 0.35$		$p_C = 0.75$	
		$\Delta = 0.20$	$\Delta = 0.10$	$\Delta = 0$	$\Delta = -0.10$	$\Delta = 0.20$	$\Delta = 0$	$\Delta = 0$	$\Delta = -0.10$
Standard	F1	0.896	0.351	0.025	0.0002	0.120	0.000	0.985	0.662
	S1	0.893	0.339	0.024	0.0003	0.115	0.0001	0.982	0.657
	F2	0.897	0.347	0.024	0.0001	0.887	0.025	0.025	0.0001
	S2	0.904	0.339	0.025	0.0004	0.883	0.025	0.024	0.0002
Changing with time	F1	0.964	0.536	0.069	0.0011	0.241	0.0001	0.997	0.823
	S1 (*)	0.944 (0)	0.520 (0)	0.062 (0)	0.0008 (0)	0.229 (0)	0 (0)	0.991 (0)	0.796 (0)
	F2	0.926	0.370	0.025	0.0002	0.875	0.023	0.025	0.0001
	S2 (*)	0.958 (0.0007)	0.390 (0.0001)	0.024 (0)	0.0004 (0)	0.880 (0)	0.025 (0)	0.015 (0.11)	0 (0)

In bold, type-I-errors maintained in [0.024;0.026]

Abbreviations: p_C indicates control survival rate; Δ : simulated survival rate difference; F1: fixed single-arm design; S1: group-sequential single-arm design; F2: fixed two-arm design; S2: group-sequential two-arm design; Standard: specific case with a fixed control survival rate; Changing with time: specific case with an increase of the control survival rate over time

(*): Proportion of significant tests with proportion of inconclusive group-sequential trials for “changing with time” case

Table 2 Proportion of trials showing significant improvement for “stopping of recruitment” cases

Stopping of recruitment	Design	$p_C = 0.50$			$p_C = 0.75$		
		$\Delta = 0.20$	$\Delta = 0$	$\Delta = -0.10$	$\Delta = 0.20$	$\Delta = 0$	$\Delta = -0.10$
$N_{STOP} = 20$	F1	0.416	0.021	0.002	1	0.618	0.244
	S1	^(a) 0.414 (0.75)	0.021 (0.587)	0.002 (0.245)	1 (0.003)	0.615 (0.585)	0.242 (0.830)
		^(b) 0.234 + 0.180	0.006 + 0.015	0.0004 + 0.0012	0.997 + 0.002	0.411 + 0.204	0.116 + 0.126
	F2	0.054	0.006	0.002	0.048	0.005	0.001
	S2	^(a) 0.118 (1)	0.042 (1)	0.060 (1)	0.054 (1)	0.019 (1)	0.044 (1)
		^(b) 0 + 0.118	0 + 0.042	0 + 0.060	0 + 0.054	0 + 0.019	0 + 0.044
$N_{STOP} = 50$	F1	0.858	0.032	0.0005	1	0.972	0.619
	S1	^(a) 0.793 (0.136)	0.023 (0.020)	0.0004 (0.0002)	1 (0)	0.948 (0.054)	0.526 (0.181)
		^(b) 0.720 + 0.073	0.016 + 0.007	0.0003 + 0.0001	1 + 0	0.916 + 0.032	0.438 + 0.088
	F2	0.333	0.032	0.005	0.320	0.021	0.003
	S2	^(a) 0.234 (0.987)	0.016 (0.950)	0.003 (0.837)	0.420 (0.998)	0.020 (0.989)	0.017 (0.903)
		^(b) 0.013 + 0.221	0.0001 + 0.016	0 + 0.003	0.002 + 0.418	0 + 0.020	0 + 0.017
$N_{STOP} = 100$	F1	0.988	0.028	0	1	1	0.875
	S1	0.894 (0)	0.025 (0)	0.0005 (0)	1 (0)	0.981 (0)	0.661 (0)
		0.542	0.029	0.002	0.834	0.024	0.0008
	S2	^(a) 0.425 (0.732)	0.012 (0.529)	0.0004 (0.190)	0.747 (0.765)	0.012 (0.766)	0.0004 (0.287)
		^(b) 0.255 + 0.170	0.005 + 0.007	0.0001 + 0.0003	0.236 + 0.511	0.001 + 0.011	0 + 0.0004

In bold, type-I-errors maintained in [0.024;0.026]

Abbreviations: p_C indicates control survival rate; Δ : simulated survival rate difference; F1: fixed single-arm design; S1: group-sequential single-arm design; F2: fixed two-arm design; S2: group-sequential two-arm design; Stopping of recruitment: specific case with a fixed control survival rate and an early stop of the trial due to an outbreak end; For group-sequential trials, the total of trials demonstrating an efficacy of the experimental treatment was defined by the sum of trials with a significant test and trials with an adjusted p -value inferior to 0.025. The p -value was adjusted for underrunning using the method proposed by Whitehead [18]

^(a): Proportion of significant tests with proportion of inconclusive group-sequential trials for “changing with time” cases

^(b): Proportion of significant tests + proportion of trials with adjusted p -values inferior to 0.025 for group-sequential trials and “stopping of recruitment” cases

experimental treatment was close to 0 regardless of the specific case, except when $p_C = 0.75$ for single-arm designs, which was due to the wrong assumption on p_H (Tables 1 and 2).

Number of subjects included

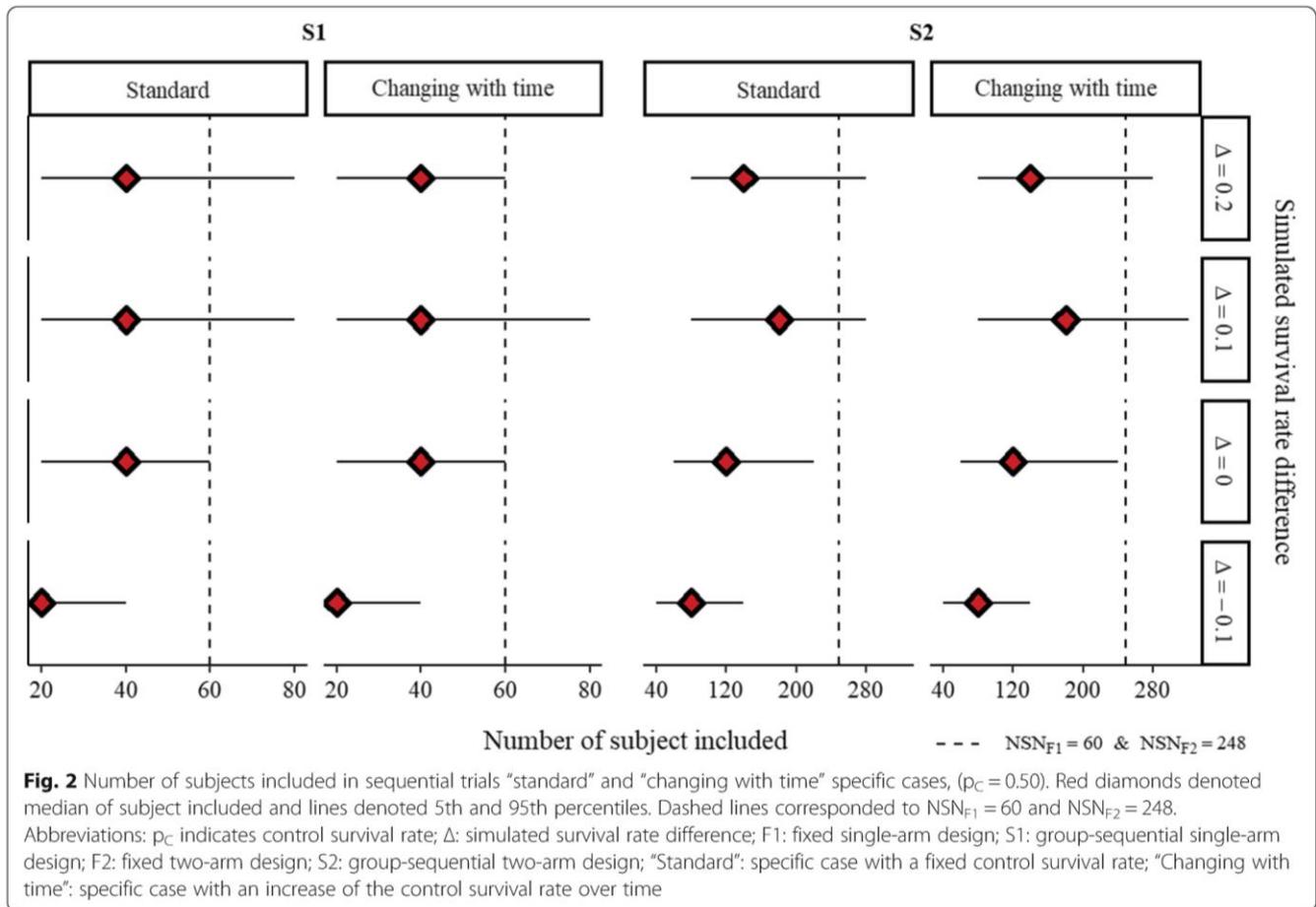
The median number [$P5^{th}$ - $P95^{th}$] of subject included in group-sequential trials for both “standard” and “changing with time” cases are reported in Fig. 2 with $p_C = 0.50$. The median number of included subjects was always lower in the group-sequential trials compared with the fixed trials, especially when the experimental treatment effect was harmful ($\Delta = -0.10$). Indeed, scenarios with $\Delta = -0.10$ or $\Delta = 0$ always presented a median number of included subjects lower than scenarios with $\Delta > 0$. Moreover, less than 10% of the group-sequential single- or two-arm trials reported a higher number of included subjects than single- or two-arm fixed trials respectively (Table 3). On the other hand, the NSN and N_{Max} determined for two-arm designs were greater to 200 patients whereas the NSN and N_{Max} for single-arm designs were lower to 100 patients. The rapidity to perform a single-arm trial compared with two-arm designs, which require more than the double of inclusions, is considerable.

Online tool to visualize results of simulated scenarios

Simulation results were reported online on the following dashboard: <https://pmn-bch.shinyapps.io/simu-vhf/>. This tool reports all scenarios for all designs (F1, F2, S1, S2), according to the choice of parameters: the control survival rate p_C and the survival rate difference Δ . Available values are, for p_C : 0.35, 0.40, 0.45, 0.50, 0.65, 0.75, and for Δ : -0.40, -0.20, -0.10, 0, 0.10, 0.20, 0.30, 0.40. Several specific cases of an outbreak can be selected: the “standard” one (with fixed p_C , the NSN for fixed designs and N_{Max} if required for group-sequential trials), the “changing with time” one which simulated the evolution of the survival rate and the “stopping of recruitment” which simulated an early stop of the trial (with $N_{STOP} = 20$; $N_{STOP} = 50$; $N_{STOP} = 100$). This tool will allow to evaluate all the elements before designing a clinical trial in the context of an outbreak of a viral haemorrhagic fever.

Discussion

We presented results from our simulation to help future investigators choosing the best clinical trial design (single- versus two-arm, fixed versus group-sequential trial)



based on the outbreak timeline of an emerging infectious disease.

At the beginning of an outbreak group-sequential two-arm trials should be preferred. A group-sequential design will often allow an early termination of the trial when the treatment does not perform better than the control or when treatment efficacy is large. Moreover, the number of infected cases increases until the epidemic peak. Thus, the required number of patients to perform a two-arm trial can be reached and the

potential evolution of the survival rate will be considered with the presence of the control arm.

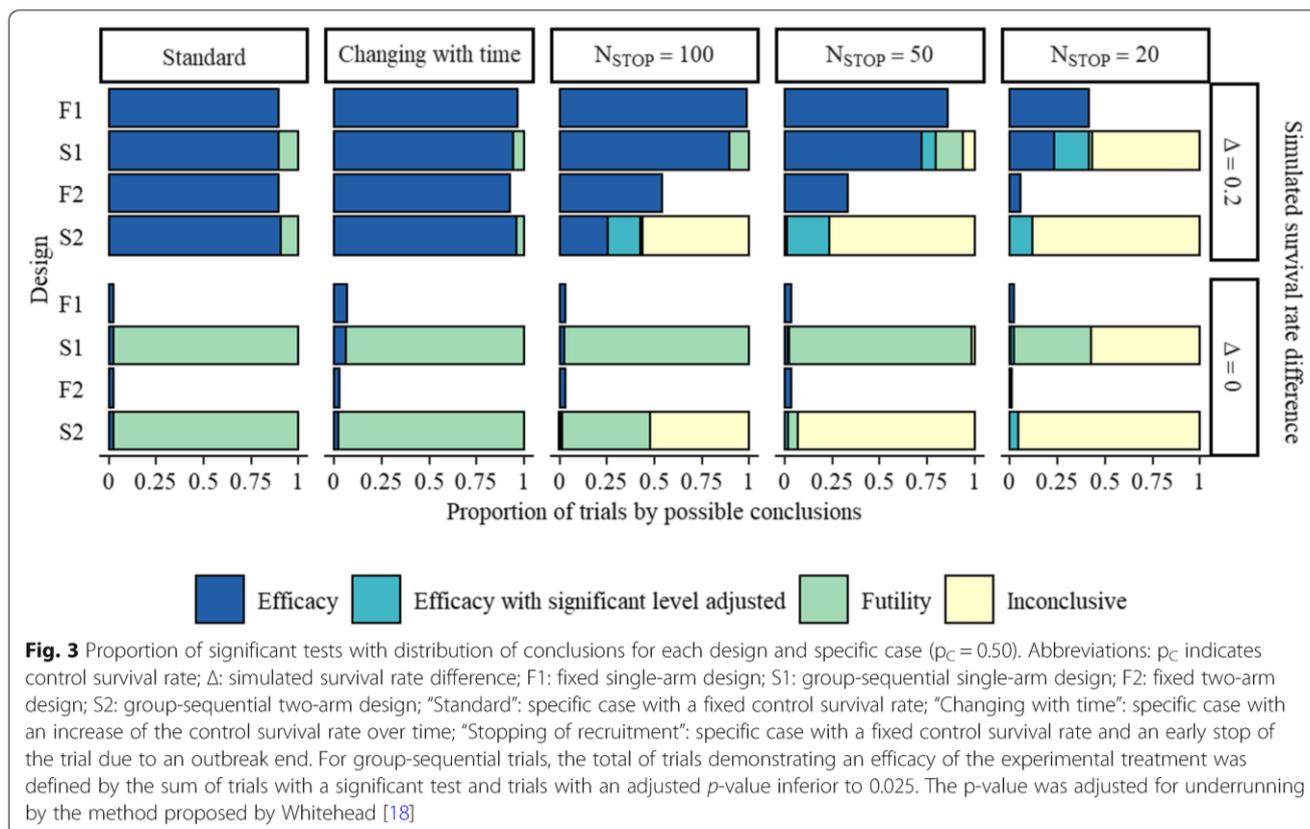
For trials beginning after the epidemic peak fixed single-arm design should be performed. As the number of cases decreases after the peak of the epidemic, single-arm designs would be preferred as they required a lower number of patients than two-arm designs. However, this assumes that the pre-trial historical survival rate was correctly estimated. We showed that with an incorrect assumption on this survival rate, the type-I-error is not maintained. Single-arm trials should only be conducted if a high level of confidence can be put on the pre-trial historical survival rate used.

From a methodological standpoint, concern must be taken when conducting a group-sequential trial. Indeed, for “stopping of recruitment” case presenting a premature termination due to an epidemic end, the adjustment for underrunning had an important impact on the results. The number of simulated trials showing a significant efficacy increased after adjusting for underrunning: the adjusted p -value led to conclude more often to efficacy of treatment (Fig. 3). Therefore, in the case of an epidemic end, where the cumulative information for group-sequential trials is too low due to the impossibility

Table 3 Proportion of group-sequential trials with subjects numbers superior to fixed designs NSN ($p_C = 0.50$)

		$\Delta = 0.2$	$\Delta = 0.1$	$\Delta = 0$	$\Delta = -0.1$
Standard	S1	0.07	0.13	0.02	0.0004
	S2	0.09	0.15	0.02	0.0003
Changing with time	S1	0.02	0.13	0.04	0.001
	S2	0.1	0.23	0.03	0.0005

Abbreviations: p_C indicates control survival rate; Δ : simulated survival rate difference; F1: fixed single-arm design; S1: group-sequential single-arm design; F2: fixed two-arm design; S2: group-sequential two-arm design; “Standard”: specific case with a fixed control survival rate; “Changing with time”: specific case with an increase of the control survival rate over time



of including patients, it is necessary to consider the adjusted p -value to limit the loss of information and to increase the probability to conclude (for futility or efficacy).

One limitation of our study is that our recommendation on using single-arm trial if it starts after the epidemic peak is based on the assumption made on the pre-trial historical survival rate used to design the trial. Indeed, we noted that results are substantially biased when this value is incorrect and the treatment is harmful. Double-blind randomised controlled trial is the best method to use to avoid those erroneous conclusions. However, conducting a randomised controlled trial is not always feasible. In those cases, single-arm trials should be used when there is relatively strong evidence behind the pre-trial historical survival rate used. Another limitation is that our analyses are based on the widely used 2.5% threshold for the type-I-error, for a one-sided test. However, in the context of viral haemorrhagic fever, where mortality rates are high and few effective therapies are available, a 2.5% threshold may not be appropriate and may even raise ethical questions. This threshold has previously been challenged, and recent studies, especially in oncology, explored new methods (Bayesian Decision Analysis) to choose suitable thresholds that minimize the overall expected harm to patients within clinical trials and future patients especially for the

deadliest diseases [20]. In this clinical trial simulation study, we decided to set the type-I-error at 2.5% for a one-sided test. By increasing the type-I-error, it is expected that the number of subjects included will decrease. However, the choice of this threshold is questionable. In particular, it could depend on the disease being studied, the potential outbreak context, the mortality rate, as well as the prevalence (e.g. rare disease).

An extension of this work would be to consider the case of multi-arm clinical trials. In the context of an outbreak, several treatments could be candidates. This was observed for Ebola virus, where various treatments were tested during the same outbreak period [6–8, 12, 13]. Each treatment was evaluated by one clinical trial and associated with a specific research institute and country. Multi-arm trial has the advantage of reducing the number of subjects to be included: indeed, a single control arm is needed against different experimental arms without loss of power. Launching several two-arm trials increases the number of control arms and the global sample size to find the right treatment [21, 22]. It also raises difficulties about acceptability for the participants and lead to a large loss of time due to competing clinical trials. The gain in the number of subjects is an important point in the context of an outbreak where the number of cases decreases following its peak and the choice of treatment to be tested is uncertain. Moreover, group-

sequential trials allow to minimise the expected number of subjects included regardless of the number of arms in the trial design. The same considerations explored in this simulation study for two-arm trials would apply to multi-arm trials; except for the case of group-sequential multi-arm design, where after premature termination of treatment arms for futility, more patients would have the possibility to receive the remaining treatment. Comparing the performance of a multi-arm group-sequential trial to a fixed trial at the end of an outbreak would be interesting, but would require to explore a multitude of scenarios, such as the efficacy or futility of each treatment.

In this simulation study, we decided to evaluate a randomization 1:1 for two-arm trial designs. An unequal allocation ratio in favour of the experimental treatment arm would have increased the probability of treatment exposure of the included patient. At first sight, this strategy is beneficial, especially at the end of an outbreak when the number of cases and the number of inclusions is decreasing substantially. However, the loss of power due to this mode of randomization is not negligible: despite an increase in the proportion of patients receiving treatment, this design would therefore require larger sample sizes to achieve the same level of statistical power [23].

Furthermore, this point questions the choice of designs for group-sequential rather than adaptive trials. This type of trial is based on the same principles as group-sequential trials: to adapt the trial based on data accumulated during the study. For example, in a group-sequential trial, a treatment arm will be stopped for futility. With an adaptive trial, it will be possible to re-evaluate during the study the number of subjects needed and change the probability of assigning a treatment (as in the case of unequal allocation ratios).

In practice, adaptive trials are mainly used for the evaluation of treatment doses. They allow, like multi-stage multi-arm trials (MAMS), the combination of phase II and phase III trials. Our simulation study focuses on the performance of phase III clinical trials. Further work could be conducted on adaptive methodologies and MAMS. New questions arise concerning the efficiency of these methods compared with phase II and III clinical trials, with the aim of accelerating the therapeutic evaluation process, using both Bayesian and frequentist approaches [15, 24]. In the case of a VHF outbreak, such as Ebola virus, one MAMS phase II – phase III trial was conducted with a multicentre, multi-outbreak, randomized controlled trial design [25]. In addition, by focusing on adaptive phase III clinical trials, the reasoning would no longer be short-term (a single outbreak wave) but medium and even long-term with withdrawal and/or addition of treatment arms according to the evolution of knowledge about the

disease studied, eligible treatments, and marketing authorisations specific to each country. Indeed, an adaptive trial, a MAMS trial, or a classic multi-arm trial seem to be possible over several outbreak waves (multi-outbreak).

Following this work, we wish to create a more complete Shiny App where users could enter other values than that presently available in the drop-down menus.

Conclusion

In conclusion, the choice of the clinical trial design to be conducted depends on the timeline of the outbreak of a viral haemorrhagic fever. At the beginning of the outbreak, group-sequential two-arm trials should be preferred, as the number of infected cases will increase until the epidemic peak allowing to conduct a strong randomised controlled trial. Moreover, a group-sequential design will allow an early termination of the trial in cases of harmful experimental treatment. The stopping for futility would be faster than the stopping for efficacy, which is an important aspect during first period of an outbreak, usually corresponding to a treatment screening phase.

Regarding trials beginning after the epidemic peak, fixed single-arm design should be performed, as the number of cases decreases after the peak of the outbreak, reducing the number of patients that could be included. However, this assumes that a high level of confidence can be put on the pre-trial historical survival rate used in the single-arm trial.

Abbreviations

F1: Fixed single-arm design; F2: Fixed two-arm design; MSA: Multi-stage approach; NSN: Numbers of subject needed; p_C : Control survival rate, the proportion of patient who survived in the control arm; p_E : Experimental survival rate, the proportion of patient who survived in the experimental arm; p_H : Pre-trial historical survival rate; S1: Group-sequential single-arm design; S2: Group-sequential two-arm design

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12874-021-01287-w>.

Additional file 1. Description of the assessed clinical trial designs.

Acknowledgements

Not applicable.

Authors' contributions

PM, FM and CL designed the simulation study. PM performed the simulation and drafted the manuscript. PM, FM and CL participated in the statistical analysis and DB helped to draft the manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. The R codes for simulations are available in GitHub: <https://github.com/PMN-BCH/simu-VHF>.

All results are available were in the online Shiny App: <https://pmn-bch.shinyapps.io/simu-vhf/>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹INSERM, Centre d'Investigation clinique-Epidémiologie Clinique 1425, Hôpital Bichat, F-75018 Paris, France. ²Département Epidémiologie Biostatistiques et Recherche Clinique, AP-HP, Hôpital Bichat, F-75018 Paris, France. ³Université de Paris, INSERM, IAME UMR 1137, F-75018 Paris, France.

Received: 2 November 2020 Accepted: 21 April 2021

Published online: 06 May 2021

References

- World Health Organization. Disease outbreaks. <http://www.who.int/emergencies/diseases/en/>. Accessed 24 Jan 2020.
- World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV). WHO. <http://www.who.int/emergencies/mers-cov/en/>. Accessed 22 Oct 2020.
- Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), Division of High-Consequence Pathogens and Pathology (DHCPP), Viral Special Pathogens Branch (VSPB). <https://www.cdc.gov/vhf/index.html>. Accessed 1 Jul 2019.
- Ebola health update - North Kivu/Ituri, DRC, 2018-2020. <https://www.who.int/emergencies/diseases/ebola/drc-2019>. Accessed 22 Oct 2020.
- Sigfrid L, Maskell K, Bannister PG, Ismail SA, Collinson S, Regmi S, et al. Addressing challenges for clinical research responses to emerging epidemics and pandemics: a scoping review. *BMC Med*. 2020;18(1):190. <https://doi.org/10.1186/s12916-020-01624-8>.
- PREVAIL II Writing Group, Multi-National PREVAIL II Study Team, Davey RT, Dodd L, Proschan MA, Neaton J, et al. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J Med*. 2016;375:1448–56.
- Sissoko D, Laouenan C, Folkesson E, M'Lebing A-B, Beavogui A-H, Baize S, et al. Experimental treatment with Favipiravir for Ebola virus disease (the JIKI trial): a historically controlled, Single-Arm Proof-of-Concept Trial in Guinea. *PLoS Med*. 2016;13:e1001967.
- van Griensven J, Edwards T, de Lamballerie X, Semple MG, Gallian P, Baize S, et al. Evaluation of convalescent plasma for Ebola virus disease in Guinea. *N Engl J Med*. 2016;374(1):33–42. <https://doi.org/10.1056/NEJMoa1511812>.
- World Health Organization. Guidance for managing ethical issues in infectious disease outbreaks. World Health Organization. 2016. <https://apps.who.int/iris/handle/10665/250580>. Accessed 1 Jul 2019.
- Adebamowo C, Bah-Sow O, Binka F, Bruzzone R, Caplan A, Delfraissy J-F, et al. Randomised controlled trials for Ebola: practical and ethical issues. *Lancet*. 2014;384(9952):1423–4. [https://doi.org/10.1016/S0140-6736\(14\)61734-7](https://doi.org/10.1016/S0140-6736(14)61734-7).
- Chow S-C, Chang M. Adaptive design methods in clinical trials - a review. *Orphanet J Rare Dis*. 2008;3(1):11. <https://doi.org/10.1186/1750-1172-3-11>.
- Dunning J, Kennedy SB, Antierens A, Whitehead J, Ciglenecki I, Carson G, et al. Experimental treatment of Ebola virus disease with Brincidofovir. *PLoS One*. 2016;11(9):e0162199. <https://doi.org/10.1371/journal.pone.0162199>.
- Dunning J, Sahr F, Rojek A, Gannon F, Carson G, Idris B, et al. Experimental treatment of Ebola virus disease with TKM-130803: a single-arm phase 2 clinical trial. *PLoS Med*. 2016;13(4):e1001997. <https://doi.org/10.1371/journal.pmed.1001997>.
- Mulangu S, Dodd LE, Davey RT, Tshiani Mbaya O, Proschan M, Mukadi D, et al. A randomized, controlled trial of Ebola virus disease therapeutics. *N Engl J Med*. 2019;381(24):2293–303. <https://doi.org/10.1056/NEJMoa1910993>.
- Cooper BS, Boni MF, Pan-ngum W, Day NPJ, Horby PW, Olliaro P, et al. Evaluating clinical trial designs for investigational treatments of Ebola virus disease. *PLoS Med*. 2015;12(4):e1001815. <https://doi.org/10.1371/journal.pmed.1001815>.
- Whitehead J, Olliaro P, Lang T, Horby P. Trial design for evaluating novel treatments during an outbreak of an infectious disease. *Clin Trials*. 2016;13(1):31–8. <https://doi.org/10.1177/1740774515617740>.
- Centers for Disease Control and Prevention. 2019. <https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/case-counts.html>. Accessed 26 Dec 2019.
- Whitehead J. Overrunning and underrunning in sequential clinical trials. *Control Clin Trials*. 1992;13(2):106–21. [https://doi.org/10.1016/0197-2456\(92\)90017-T](https://doi.org/10.1016/0197-2456(92)90017-T).
- Whitehead J. The design and analysis of sequential clinical trials. Chichester: Wiley; 1997. ISBN: 978-0-471-97550-2.
- Montazerhodjat V, Chaudhuri SE, Sargent DJ, Lo AW. Use of Bayesian decision analysis to minimize harm in patient-centered randomized clinical trials in oncology. *JAMA Oncol*. 2017;3(9):e170123. <https://doi.org/10.1001/jamaoncol.2017.0123>.
- Freidlin B, Korn EL, Gray R, Martin A. Multi-arm clinical trials of new agents: some design considerations. *Clin Cancer Res*. 2008;14(14):4368–71. <https://doi.org/10.1158/1078-0432.CCR-08-0325>.
- Parmar MKB, Carpenter J, Sydes MR. More multiarm randomised trials of superiority are needed. *Lancet*. 2014;384(9940):283–4. [https://doi.org/10.1016/S0140-6736\(14\)61122-3](https://doi.org/10.1016/S0140-6736(14)61122-3).
- Hey SP, Kimmelman J. The questionable use of unequal allocation in confirmatory trials. *Neurology*. 2014;82(1):77–9. <https://doi.org/10.1212/01.wnl.0000438226.10353.1c>.
- Chevret S. Bayesian adaptive clinical trials: a dream for statisticians only? *Stat Med*. 2012;31(11-12):1002–13. <https://doi.org/10.1002/sim.4363>.
- National Institute of Allergy and Infectious Diseases (NIAID). A Multicenter, Multi-Outbreak, Randomized, Controlled Safety and Efficacy Study of Investigational Therapeutics for the Treatment of Patients With Ebola Virus Disease. Clinical trial registration. clinicaltrials.gov; 2020. <https://clinicaltrials.gov/ct2/show/study/NCT03719586>. Accessed 24 Jan 2021.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Annexe 2 : protocole de la revue systématique publié dans PROSPERO

The mortality, prevalence, sequelaes and treatments of viral haemorrhagic fevers: a systematic literature review.

To enable PROSPERO to focus on COVID-19 submissions, this registration record has undergone basic automated checks for eligibility and is published exactly as submitted. This protocol has been amended since registration with changes to the PICOS criteria, data extraction, quality assessment, or data synthesis methods. Previous versions of the registration may be viewed for comparison. PROSPERO has never provided peer review, and usual checking by the PROSPERO team does not endorse content. Therefore, automatically published records should be treated as any other PROSPERO registration. Further detail is provided [here](#).

Citation

Drifa Belhadi, Cédric Laouénan, France Mentré. The mortality, prevalence, sequelaes and treatments of viral haemorrhagic fevers: a systematic literature review.. PROSPERO 2020 CRD42020167306 Available from: https://www.crd.york.ac.uk/prospERO/display_record.php?ID=CRD42020167306

Review question

Which data are available in terms of case fatality rates, prevalence, sequelaes and treatments of viral haemorrhagic fevers (VHF)?

Searches

We will conduct a search on the WHO and CDC websites to identify the number of cases and deaths associated with each disease by year and by country. Additional hand searches will be performed on websites from Ministries of Health.

We will also conduct a systematic literature search of bibliographic databases based on the PRISMA guideline. Relevant studies will be identified by searching MEDLINE, Embase and the Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library. Additional hand searches will be performed on databases not indexed in Embase, on ClinicalTrials.gov and other potentially relevant internet sources. Search terms will be developed using a combination of MeSH/EMTREE terms and free-text terms to capture the relevant populations, outcomes and study types.

Types of study to be included [1 change]

Any information source and study type providing prevalence, incidence, case-fatality rates and sequelaes for specified VHF and survival rates associated with treatments are considered. These include official information from WHO/CDC and Ministries, cross-sectional and cohort/case-control studies (including incidence/prevalence surveys) and clinical trials.

Studies are ineligible if any of the following apply:

- the study design is case-control, case reports, case series and incidence/prevalence estimates from the defined population cannot be calculated,
- the methodology is not described,
- reviews and meta-analyses are excluded.

Condition or domain being studied

Any VHF from the following list:

- Alkhurma hemorrhagic fever
- Argentine hemorrhagic fever
- Bolivian hemorrhagic fever
- Chapare hemorrhagic fever
- Crimean-Congo hemorrhagic fever
- Ebola (Ebola Virus Disease)
- Hantavirus Pulmonary Syndrome (HPS)
- Hemorrhagic fever with renal syndrome (HFRS)
- Kyasanur Forest Disease (KFD)
- Lassa fever
- Lujo hemorrhagic fever
- Lymphocytic choriomeningitis (LCM)
- Marburg hemorrhagic fever
- Omsk hemorrhagic fever
- Rift Valley fever
- Sabia-associated hemorrhagic fever
- Tick-borne Encephalitis (TBE)
- Venezuelan hemorrhagic fever

Participants/population

Humans infected with a VHF reported above.

Intervention(s), exposure(s)

This review will not focus on any specific intervention. Any studies reporting prevalence and/or case fatality rates and/or sequelae and/or treatments of a VHF will be included.

Comparator(s)/control

No comparator.

Main outcome(s)

Case fatality rates associated with each disease.

Measures of effect

Until the time of the search and at each outbreak.

Additional outcome(s)

Prevalence, sequelae, sequelae prevalence associated with each disease.

Survival rates associated with available treatments.

Measures of effect

Until the time of the search and at each outbreak.

Data extraction (selection and coding)

Information identified from the WHO and CDC websites will be extracted for each disease by outbreak and by country.

For the literature search, after removing duplicates, two researchers will independently evaluate all identified citations on the basis of titles and/or abstracts. In case of disagreement, consensus meeting will take place to discuss the study selection. Full-text publications of studies included based on title and abstract will be retrieved and reviewed by two researchers to assess eligibility based on the inclusion/exclusion criteria. Justification for study exclusion will be documented. Data extraction will then be conducted on the included studies.

If different data sources overlapped for a country over one period, only one data source will be used, and the following order will be used in the selection of the data source: WHO data, CDC data, national data and other data source.

Risk of bias (quality) assessment

Quality assessment will be independently performed by two researchers using the Newcastle-Ottawa Quality Assessment Form to evaluate cohort studies, and the Risk of Bias (RoB) tool described in the Cochrane Handbook for Systematic Reviews of Interventions to evaluate clinical trials.

Strategy for data synthesis

A narrative synthesis approach by disease will be used for the analyses. Descriptive tables and graphs will be used to report the findings.

Analysis of subgroups or subsets

None planned.

Contact details for further information

Drifa Belhadi
drifa.belhadi@aphp.fr

Organisational affiliation of the review

INSERM, University Paris Diderot

Review team members and their organisational affiliations

Ms Drifa Belhadi. APHP, Department of Epidemiology, Biostatistic and Clinical Research, Bichat Hospital, Paris France. INSERM, University Paris Diderot, IAME, UMR1137, France

Dr Cédric Laouénan. APHP, Department of Epidemiology, Biostatistic and Clinical Research, Bichat Hospital, Paris, France. INSERM, University Paris Diderot, IAME UMR1137, France

Professor France Mentré. APHP, Department of Epidemiology, Biostatistic and Clinical Research, Bichat Hospital, Paris, France. INSERM, University Paris Diderot, IAME UMR1137, France

Type and method of review

Epidemiologic, Narrative synthesis, Systematic review

Anticipated or actual start date

01 December 2019

Anticipated completion date [1 change]

31 October 2022

Funding sources/sponsors

None

Conflicts of interest

Language

English

Country

France

Stage of review [1 change]

Review Completed published

Details of final report/publication(s) or preprints if available [1 change]

Belhadi D, El Baied M, Mulier G, Malvy D, Mentré F, Laouénan C. The number of cases, mortality and treatments of viral hemorrhagic fevers: A systematic review. PLOS Neglected Tropical Diseases. 2022 Oct 31;16(10):e0010889.

<https://doi.org/10.1371/journal.pntd.0010889>

Subject index terms status

Subject indexing assigned by CRD

Subject index terms

Disease Outbreaks; Hemorrhagic Fevers, Viral; Humans; Prevalence

Date of registration in PROSPERO

28 April 2020

Date of first submission

30 January 2020

Stage of review at time of this submission [2 changes]

Stage	Started	Completed
Preliminary searches	Yes	Yes
Piloting of the study selection process	Yes	Yes
Formal screening of search results against eligibility criteria	Yes	Yes
Data extraction	Yes	Yes
Risk of bias (quality) assessment	Yes	Yes
Data analysis	Yes	Yes

Revision note

I corrected the stages as "started" and "completed".

The record owner confirms that the information they have supplied for this submission is accurate and complete and they understand that deliberate provision of inaccurate information or omission of data may be construed as scientific misconduct.

The record owner confirms that they will update the status of the review when it is completed and will add publication details in due course.

Versions

28 April 2020

04 August 2022

13 October 2023

13 October 2023

Liste des tableaux

TABLEAU 1 MATRICE DE COUT BAYESIENNE POUR UN CRITERE DE JUGEMENT CONTINU (BAYESIAN DECISION ANALYSIS) - SOURCE : ISAKOV ET AL. J ECONOM 2019 [26]	26
TABLEAU 2 NOMBRE DE SUJETS NECESSAIRES, VALEUR CRITIQUE, ERREUR DE TYPE I ET PUISSANCE OPTIMALES CALCULEES PAR BDA POUR UN EFFET DU TRAITEMENT ATTENDU $\delta_0 = \sigma_0$ (OU σ REPRESENTE L'ERREUR STANDARD DANS CHAQUE BRAS) - SOURCE : ISAKOV ET AL. J ECONOM 2019 [26]	27
TABLEAU 3 ERREURS DE TYPE I OPTIMALES CALCULEES PAR BDA EN UTILISANT LES DONNEES DE L'ESSAI DE PHASE II EVALUANT AMX0035 ET EN PRENANT 89 PATIENTS DANS LE BRAS TRAITEMENT ET 48 PATIENTS DANS LE BRAS CONTROLE – SOURCE : XU ET AL. AMYOTROPH LATERAL SCLER FRONTOTEMPORAL DEGENER 2022 [27]	29
TABLEAU 4 RESULTATS SUR LES SEQUELLES RAPPORTEES PAR FHV	61

Liste des figures

FIGURE 1. DISTRIBUTION DES FHV A TRAVERS LE MONDE. LA COULEUR INDIQUE LE NOMBRE DE FHV AVEC UNE EPIDEMIE RAPPORTEE JUSQU'EN JANVIER 2020 DANS CHAQUE PAYS ET L'ACRONYME REPRESENTE LA FHV – SOURCE : BELHADI ET AL. PLOS NEGL TROP DIS 2022 [5]	15
FIGURE 2 SCHEMA POUR : A. UN ESSAI SEQUENTIEL PAR GROUPE; B. UN ESSAI MULTI-BRAS MULTI-ETAPES	20
FIGURE 3 SCENARIOS CONSIDERES AU COURS D'UNE EPIDEMIE POUR L'ETUDE DE SIMULATION – SOURCE : MANCHON ET AL. BMC MED RES METHODOL 2021 [23]	23
FIGURE 4 CHRONOLOGIE D'UNE EPIDEMIE DE FHV AVEC LES RECOMMANDATIONS DE MANCHON ET AL. [23] ET DE NOTRE ETUDE BDA	93

Table des matières

Remerciements	3
Sommaire	11
Productions scientifiques liées à la thèse	12
1. Introduction	13
1.1. Fièvres hémorragiques virales.....	13
1.2. Quelques éléments de contexte statistique.....	17
a. Généralités sur les tests statistiques.....	17
b. Description de différents designs d'essai clinique utilisés lors d'une épidémie de FHV	18
1.3. Cas de la maladie à virus Ebola.....	20
1.4. Challenges associés à la recherche clinique dans un contexte d'épidémie de FHV .	21
1.5. Remise en question du seuil de significativité à 5%.....	24
1.6. Bayesian Decision Analysis	25
1.7. Objectifs de la thèse	29
2. Nombres de cas, mortalité, traitements et séquelles des fièvres hémorragiques virales	32
2.1. Résumé	32
2.2. Article 1 publié dans PLOS Neglected Tropical Diseases.....	34
2.3. Séquelles des fièvres hémorragiques virales	60
3. Adaptation de la Bayesian Decision Analysis pour la mise en place d'essai clinique avec un critère de jugement binaire	65
3.1. Résumé	65
a. Adaptation à un critère de jugement binaire.....	65
b. Illustration pour un essai clinique évaluant un traitement contre la maladie à virus Ebola	66
3.2. Article 2 en révision dans Journal of Biopharmaceutical Statistics.....	69
4. Discussion et perspectives	90
4.1. Première revue systématique publiée sur le nombre de cas, la mortalité et les traitements des FHV.....	90
4.2. Comment prendre en compte le contexte des FHV lors du design d'un essai clinique ?	91
5. Bibliographie	96
Annexe 1 : Manchon <i>et al.</i> BMC Medical Research Methodology	102
Annexe 2 : protocole de la revue systématique publié dans PROSPERO	112
Liste des tableaux.....	117
Liste des figures.....	117
Table des matières	118

L'Analyse Bayésienne Décisionnelle pour le design d'essais cliniques dans le contexte de fièvres hémorragiques virales

Les fièvres hémorragiques virales (FHV) sont souvent associées à une forte mortalité et des épidémies imprévisibles, entraînant des difficultés à conduire des essais cliniques. Du fait d'un seuil de significativité statistique conservateur de 5%, les investigateurs font face à des nombres réduits de patients inclus impactant la puissance statistique de l'essai. Ainsi, dans ce contexte, une erreur de type I de 5% peut ne pas être appropriée. La "Bayesian Decision Analysis" (BDA) permet de prendre en compte le contexte de la maladie pour calculer la taille d'échantillon et erreur de type I optimales lors du design d'un essai. Nous avons adapté la BDA pour pouvoir l'utiliser pour des FHV. Etant donné que la mortalité est le critère de jugement le plus utilisé dans ce contexte, nous avons adapté la méthodologie à des essais cliniques avec un critère binaire. Ce travail permet d'obtenir les tailles d'échantillon et erreurs de type I optimales pour un essai randomisé à 2 bras selon plusieurs paramètres à pré-spécifier dont la sévérité de maladie, la taille de la population cible et l'effet attendu du traitement. Nous avons illustré cette méthodologie dans le contexte des FHV, notamment la maladie à virus Ebola. Nous avons conduit une revue systématique pour récolter les nombres de cas, mortalité, traitements et séquelles pour chaque FHV. Nous avons ensuite appliqué notre BDA au cas de la maladie à virus Ebola. Ce travail permet d'apporter un nouvel outil aux chercheurs souhaitant évaluer un traitement, notamment pour des maladies très mortelles avec peu de traitements disponibles.

Mots-clés: *Bayesian decision analysis, essai clinique, maladie à virus Ebola, fièvre hémorragiques virales, erreur de type I, taille d'échantillon*

Bayesian decision analysis to design clinical trials in the context of viral hemorrhagic fevers

Viral hemorrhagic fevers (VHFs) are often associated with high mortality rates and unpredictable outbreaks, leading to difficulties to conduct clinical trials. Due to a conservative statistical significance threshold (5% threshold), investigators face reduced numbers of included patients with an impact to the statistical power of the trial. Therefore, in this context, a 5% type I error may be inappropriate. The Bayesian Decision Analysis (BDA) allows taking into account the context of the disease to calculate optimal sample size and type I error when designing a clinical trial. We adapted the BDA to be able to use it in the context of VHFs. As the mortality rate is the most commonly used outcome in our context, we adapted the methodology to clinical trials with a binary outcome. Finally, this work allows obtaining optimal values of sample size and type I error for a two-arm randomized trial given several parameters to pre-specify including the severity of the disease, the size of the target population and the expected treatment effect. We illustrated this methodology in the context of VHFs, more particularly Ebola virus disease. We conducted a systematic review to retrieve the number of cases, mortality, treatments and sequelae for each VHF. Then we applied our BDA framework in the context of Ebola virus disease. This work provides a new tool to researchers wishing to assess a new experimental treatment, especially for diseases with high mortality rates and few therapies available.

Keywords: *Bayesian decision analysis, clinical trial, Ebola virus disease, viral hemorrhagic fevers, type I error, sample size*