

# Comparing the Efficacy of Drug Regimens for Pulmonary Tuberculosis: Meta-analysis of Endpoints in Early-Phase Clinical Trials

Laura J. Bonnett,<sup>1</sup> Gie Ken-Dror,<sup>1</sup> Gavin C. K. W. Koh,<sup>2</sup> and Geraint R. Davies<sup>3</sup>

<sup>1</sup>Department of Biostatistics, University of Liverpool, <sup>2</sup>Diseases of the Developing World, GlaxoSmithKline, Uxbridge, and <sup>3</sup>Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, United Kingdom

(See the Editorial Commentary by Phillips on pages 55–6.)

**Background.** A systematic review of early clinical outcomes in tuberculosis was undertaken to determine ranking of efficacy of drugs and combinations, define variability of these measures on different endpoints, and to establish the relationships between them.

**Methods.** Studies were identified by searching PubMed, Medline, Embase, LILACS (Latin American and Caribbean Health Sciences Literature), and reference lists of included studies. Outcomes were early bactericidal activity results over 2, 7, and 14 days, and the proportion of patients with negative culture at 8 weeks.

**Results.** One hundred thirty-three trials reporting phase 2A (early bactericidal activity) and phase 2B (culture conversion at 2 months) outcomes were identified. Only 9 drug combinations were assessed on >1 phase 2A endpoint and only 3 were assessed in both phase 2A and 2B trials.

**Conclusions.** The existing evidence base supporting phase 2 methodology in tuberculosis is highly incomplete. In future, a broader range of drugs and combinations should be more consistently studied across a greater range of phase 2 endpoints.

**Keywords.** tuberculosis; efficacy; meta-analysis.

First-line tuberculosis therapy has remained unchanged for 40 years. Whereas “short-course” treatment is effective in clinical trials, in practice the 6 months required for successful cure is burdensome for patients and tuberculosis programs. Identifying new and ultra-short regimens will require identification of suitable surrogate outcomes to facilitate progression of novel treatment regimens through phase 2 to phase 3 trials and de-risk drug development [1].

The current “gold standard” phase 3 endpoint is a composite of treatment failure and relapse up to 24 months following treatment completion. Use of this binary outcome, which is rare in the comparator arm (<5% with standard short-course regimens), mandates large sample sizes to adequately power clinical trials. The prolonged follow-up that is required further adds to trial costs, making the definitive outcome unsuitable for extensive evaluation of drug combinations or dose-finding.

Numerous surrogate outcomes have been used for these purposes in phase 2 studies. Phase 2A studies of early bactericidal activity (EBA) based on quantitative sputum bacteriology enroll small patient numbers for up to 2 weeks [2]. While the original rationale for such studies was dose-finding for single agents, more recent studies have evaluated drug combinations [3, 4]. This concept has been extended into larger phase 2B studies with combination therapy lasting up to 2 months [5]. The most studied phase 2B outcome has been sputum culture conversion at fixed time-points, usually 2 months [6–8]. This endpoint is supported by regulators for conditional approval of novel drugs [9], but there remains a lack of consensus among trialists as to the utility of EBA studies and of other approaches to intermediate bacteriological data such as time-to-event and regression modeling [10].

A complete understanding of the performance of tuberculosis treatment regimens in early-phase clinical trials is critical to understanding their usefulness in predicting phase 3 trial results and in calibrating preclinical models of treatment. Although the goals of historical phase 2A and 2B regimens are distinct, with the former focusing on proof-of-concept for individual drugs and the latter on identifying the best combinations of drugs, it seems important to understand whether this information can be transmitted rationally through these phases. We undertook a systematic review of early clinical outcomes in tuberculosis (within the first 2 months of treatment), focusing on the key drugs comprising

Received 9 October 2016; editorial decision 27 January 2017; accepted 30 January 2017; published online April 10, 2017.

Correspondence: G. R. Davies, Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, 8 West Derby St, Liverpool L69 7BE, UK (gdavies@liverpool.ac.uk).

Clinical Infectious Diseases® 2017;65(1):46–54

© The Author 2017. Published by Oxford University Press for the Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com). DOI: 10.1093/cid/cix247

modern and historical first-line treatment regimens, to determine the overall ranking of efficacy of drugs and combinations, to define the variability of these measures of effect on the different endpoints used, and to establish the relationships between them.

## METHODS

The review included randomized clinical trials (RCTs) including patients with smear- and culture-positive pulmonary tuberculosis, being treated for the first time or with known isoniazid-monoresistant organisms, and including regimens containing any combination of historic or novel drugs used or proposed for use in first-line treatment regimens. Predefined outcomes of interest were EBA over 2 (EBA<sub>0-2</sub>), 7 (EBA<sub>0-7</sub>), and 14 days (EBA<sub>0-14</sub>), and the proportion of patients with negative culture results at 8 weeks. A systematic search of databases was conducted on 12 December 2016 (see Supplementary Appendix 1). Risk of bias was considered [11]. Pooled estimates of each outcome for each drug or combination were obtained. Meta-regression was used to examine the impact of clinical covariates on the effect size of culture results at 8 weeks. Analyses were performed using R version 2.14.1 [12]. Full methodology is detailed in the Supplementary Appendix 1.

## RESULTS

### Included Studies

Figure 1 shows the number of studies included at each stage of the review. The main reasons for exclusion were failure to meet the inclusion criteria, specifically previously treated or drug-resistant patients, and study design other than RCT.

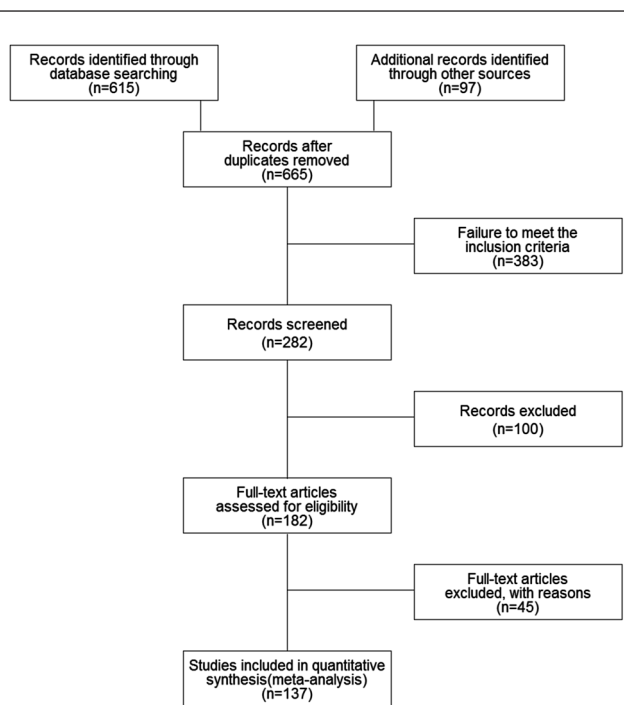


Figure 1. Literature review process.

One hundred thirty-three relevant studies were identified that reported outcomes of interest. Of these, 37 were phase 2 studies and 96 were phase 3 studies reporting intermediate bacteriological outcomes. All 96 phase 3 studies contributed to the phase 2B outcome only—no phase 3 studies contributed data to the phase 2A outcomes. Together these studies provide data relating to 37 173 patients and 67 drug combinations.

Figure 2 summarizes numbers of studies and patients pertaining to each drug combination across all the outcomes of interest. Drug abbreviations are as follows: E, ethambutol; G, gatifloxacin; H, isoniazid; J, bedaquiline; L, levofloxacin; M, moxifloxacin; O, ofloxacin; P, para-aminosalicylic acid; Pa, PA-824/pretomanid; R, rifampicin; Rb, rifabutin; Rp, rifapentine; S, streptomycin; T, thiacetazone; Z, pyrazinamide. Because only combinations including drugs of interest to the review are summarized, in some cases only data concerning the control arms of trials are presented. Additionally, studies with  $\geq 2$  trial arms were analyzed separately. The composition of each drug combination refers only to the period preceding the endpoint of interest. Therefore, for the 8-week culture outcome, usually only drugs used in the initiation phase of treatment are reported without the associated continuation-phase drugs. Where the initiation phase was  $< 2$  months, however, continuation drugs are also listed. The regimens for which most data were available were HRZE, SHRZ, and HRZ combination therapy.

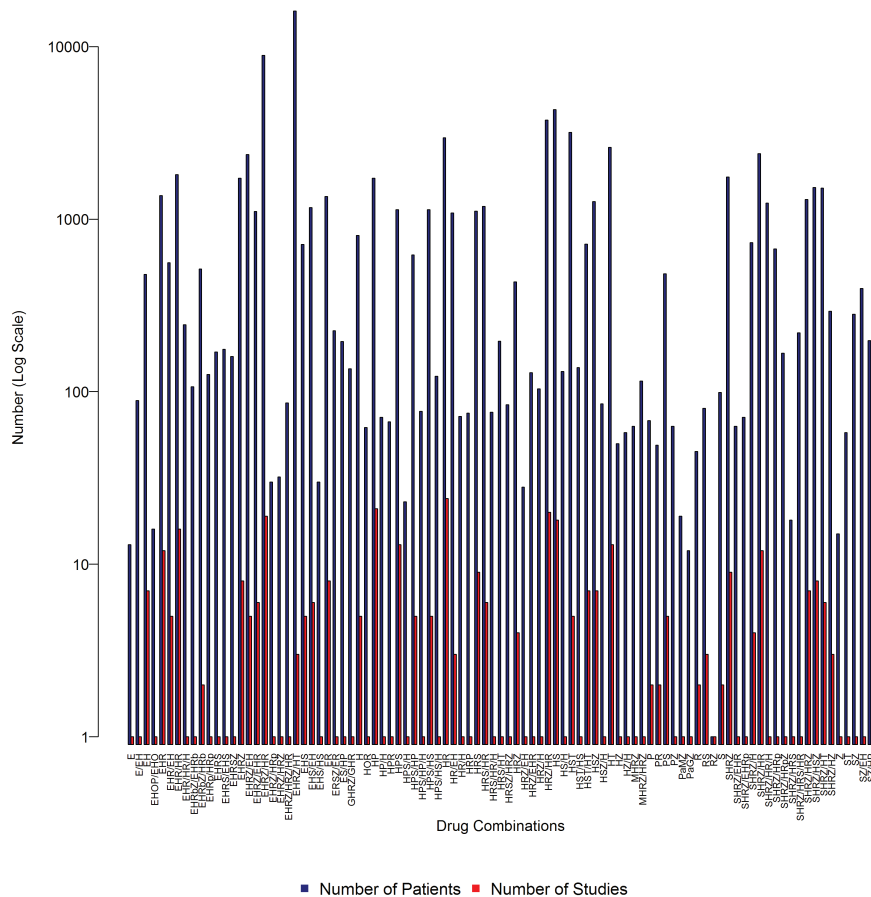
Twenty-four studies reported phase 2A outcomes. EBA<sub>0-2</sub> was reported in all 24 studies (141 trial arms, 35 drug combinations, 1424 patients). In some cases, studies considered the same drug combination but different treatment strategies and dosing intervals. Others considered single-formulation treatments vs combined formulations, and some considered multiple dosages of a drug. EBA<sub>0-7</sub> and EBA<sub>0-14</sub> were reported in only 6 (23 trial arms, 14 drug combinations, 296 patients) and 8 studies (46 trial arms, 27 drug combinations, 449 patients), respectively.

The proportion of patients who were culture negative by 8 weeks was reported in 104 studies considering phase 2B outcomes. These studies investigated 45 different drug combinations in 34 418 patients. One study reported both phase 2A and phase 2B outcomes [13].

Forest plots for each drug combination and outcome and associated numerical results are presented in the Supplementary Appendices 2 and 3. Results shown in Figures 3–8 are graphical summaries based on standard doses recommended in treatment guidelines in the case of historic drugs, or doses going into phase 3 trials in the case of novel drugs.

### Risk of Bias

Eighty-seven (65%) studies provided information on sequence generation. In most cases (94%) patients were “randomly allocated,” so studies were classified as unclear risk of bias. Some publications mentioned stratifying by factors such as severity, or used permuted block designs, random tables, or similar and were classified as low risk of bias. Five studies referred to



**Figure 2.** Drug combinations of included studies. Abbreviations: E, ethambutol; G, gatifloxacin; H, isoniazid; J, bedaquiline; L, levofloxacin; M, moxifloxacin; O, ofloxacin; P, para-aminosalicylic acid; Pa, PA-824/pretomanid; R, rifampicin; Rb, rifabutin; Rp, rifapentine; S, streptomycin; T, thiacetazone; Z, pyrazinamide.

quasi-randomization and were therefore classified as high risk of bias.

Only 30 (23%) studies mentioned allocation concealment. Of these, 28 (93%) studies used sealed envelopes and were classified as low risk. Ninety-seven (73%) studies either reported that the study was unblinded, or did not specify blinding procedures and were classified as high risk of bias. Fifteen included studies were of a double-blinded nature including the use of telephone randomization, or prearranged lists, although 1 study stated that it was double-blind during the maintenance phase of treatment only [14]. Most other blinded studies mentioned that radiographers or laboratory staff were blinded to treatment; these were considered as single-blinded designs and classified as low risk of bias.

Ninety-seven (73%) studies were published before the Consolidated Standards of Reporting Trials (CONSORT) when selective reporting had not been raised as a possible source of bias. In all studies published post-CONSORT, the risk of bias is unclear as there is insufficient information to determine whether the published reports include all expected outcomes, including those prespecified. Eighty-five (64%) studies reported reasons for exclusions, or numbers lost to follow-up.

Due to the limited number of high-quality studies, sensitivity analysis assessing the impact of risk of bias was not performed.

### Phase 2A Studies

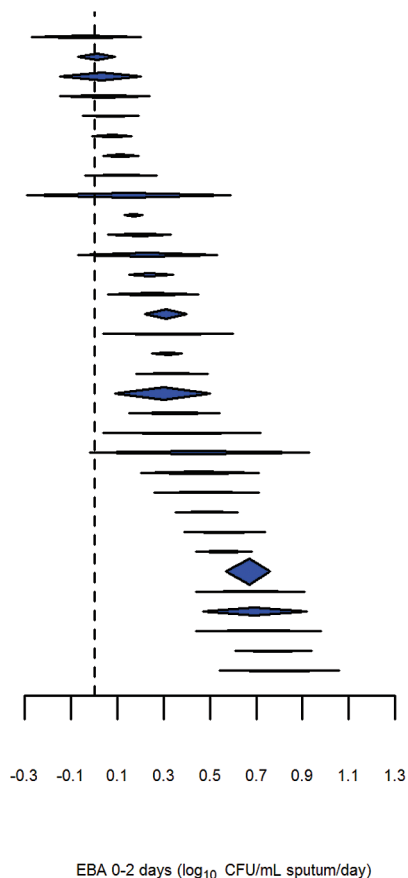
#### EBA0-2

Pooled results for the  $EBA_{0-2}$  outcome can be seen in Figure 3. Of the 32 drugs and combinations, only 5 were studied in >30 patients. Hence, the confidence intervals on pooled estimates of effect are wide and frequently overlap. Some drugs (Rb, Z, J, and S) do not demonstrate any significant efficacy on this endpoint and can also clearly be distinguished from H, the most commonly studied and precisely estimated drug. However, even quite commonly studied combinations containing H such as HRZ, HRZE, and SHRZ do not appear significantly different from H monotherapy using the  $EBA_{0-2}$  endpoint. Similarly, it does not appear to be possible to separate the effect of HRZE from any of its component drugs, with the exception of Z.

#### EBA0-7 and EBA0-14

Pooled results for the  $EBA_{0-7}$  and  $EBA_{0-14}$  outcomes can be seen in Figures 4 and 5, respectively. Although the number of studies reporting these endpoints was fewer, variability of these

Combination	Culture	No. Patients	No. Regimens
Rb	Solid	12	1
Z	Solid	34	3
J	Solid	41	3
S	Solid	13	2
ES	Solid	4	1
JZ	Solid	15	1
JPa	Solid	14	1
SZ	Solid	4	1
Pa	Solid	29	2
PaZ	Solid	15	1
PaGZ	Solid	12	1
R	Solid	20	2
Rp	Solid	16	1
E	Solid	13	2
O	Solid	53	5
HRS	Solid	4	1
PaMZ	Solid	15	1
RS	Solid	4	1
HRZE	Solid	66	7
G	Solid	10	1
SH	Solid	4	1
M	Solid	18	2
L	Solid	10	1
EHR	Solid	4	1
HZ	Solid	4	1
ER	Solid	4	1
HRZ	Solid	9	1
H	Solid	145	15
ESHRZ	Solid	4	1
SHRZ	Solid	43	3
HR	Solid	4	1
EH	Solid	4	1
HSZ	Solid	4	1



**Figure 3.** Results of early bactericidal activity over 2 days (fixed effects, generalized inverse variance method). Abbreviations: CFU, colony-forming units; E, ethambutol; EBA, early bactericidal activity; G, gatifloxacin; H, isoniazid; J, bedaquiline; L, levofloxacin; M, moxifloxacin; O, ofloxacin; P, para-aminosalicylic acid; Pa, PA-824/pretomanid; R, rifampicin; Rb, rifabutin; Rp, rifapentine; S, streptomycin; T, thiacetazone; Z, pyrazinamide.

endpoints appeared lower than for  $EBA_{0-2}$ . Even so, it did not appear possible to distinguish statistically between the drugs and regimens studied, including combinations such as HRZE and drugs as diverse as H, Pa, and J.

#### Phase 2B Studies

Figures 6–8 summarize the pooled estimates of the proportion of patients who were culture negative at 8 weeks, grouped by culture method. The overall point estimates of culture conversion for the most effective rifampicin-containing regimens on this endpoint (HRZ, SHRZ, and HRZE) exceeded 85%, whereas for most nonrifampicin regimens this estimate was no better than 50%. Though relatively precise estimates were obtained for frequently studied modern short-course regimens such as HRZ, SHRZ, and HRZE, the performance of these regimens was not statistically distinguishable from historical regimens comprising similar numbers of trials and patients, such as HS, HPS, or SHR. This appeared to reflect high intertrial variability within regimens as measured by  $I^2$  and  $\tau^2$  estimated using 2 approaches (see Supplementary Appendices 2 and 3).

Fewer data were available relating to culture conversion as measured by liquid culture. Although variability appeared lower

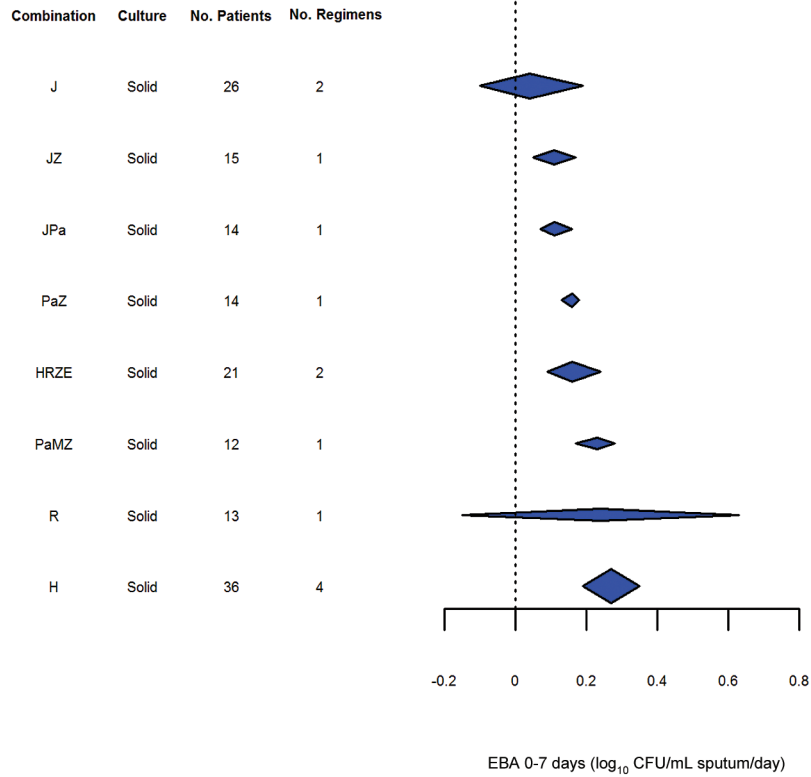
than with solid culture, this could be due to the smaller number of studies included in this analysis. Variability increased when pooled results based on solid and liquid culture methods were reported and confidence intervals for all the regimens tested using liquid culture results overlapped.

#### Meta-regression Analyses

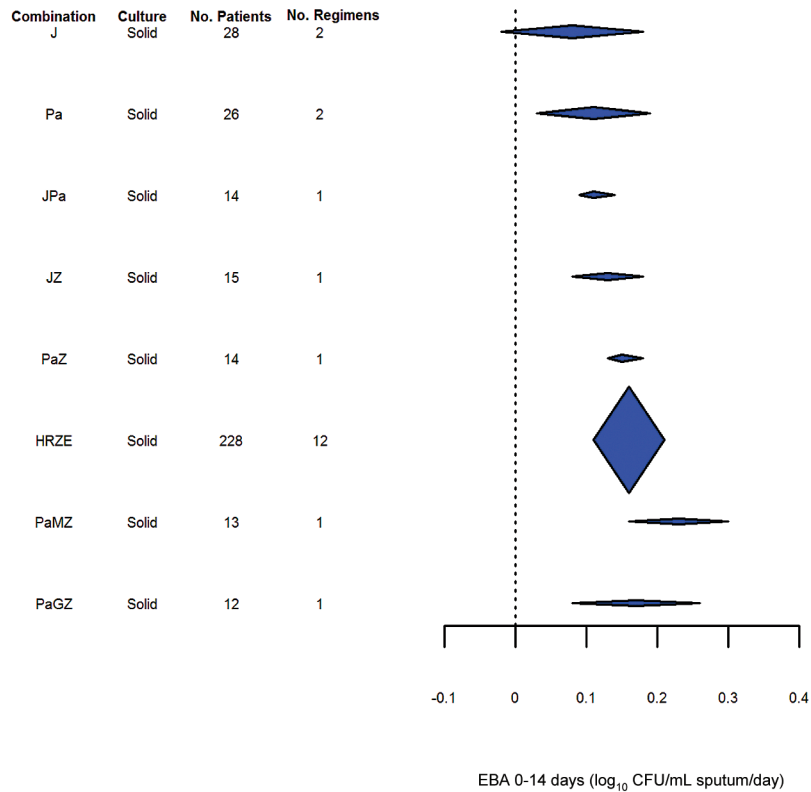
The results of meta-regression analyses can be seen in (Table 1). For the selected drug combinations HRZE, HRZ, and SHRZ, neither year of publication nor geographical location were statistically significant. Human immunodeficiency virus (HIV) coinfection could only be examined for the drug combination HRZE as there was insufficient data for HRZ and SHRZ. It was also not significantly associated with the proportion of patients who were culture negative at either weeks. The inclusion of R and Z in a regimen independently explained significant heterogeneity among drug combinations (see Supplementary Appendix 4).

#### Ranking

Because it was difficult to discriminate between regimens in terms of formal statistical inference, we evaluated whether the rank order of regimens was consistent between different

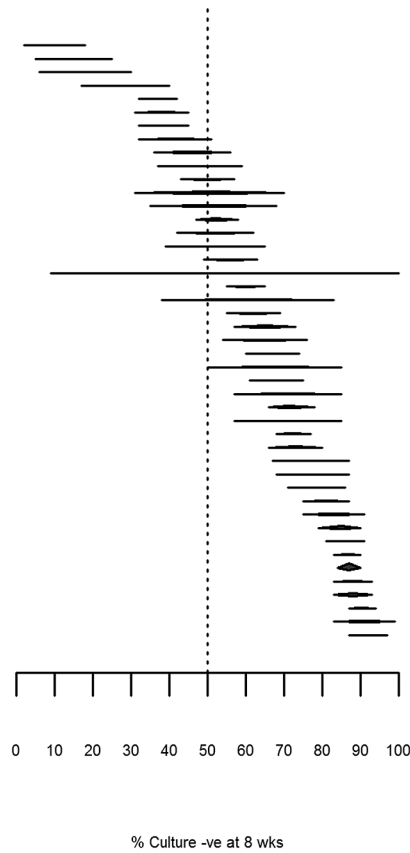


**Figure 4.** Results of early bactericidal activity over 7 days (fixed effects, generalized inverse variance method). Abbreviations: CFU, colony-forming units; E, ethambutol; EBA, early bactericidal activity; H, isoniazid; J, bedaquiline; M, moxifloxacin; Pa, PA-824/pretomanid; R, rifampicin; Z, pyrazinamide.



**Figure 5.** Results of early bactericidal activity over 14 days (fixed effects, generalized inverse variance method). Abbreviations: CFU, colony-forming units; E, ethambutol; EBA, early bactericidal activity; G, gatifloxacin; H, isoniazid; J, bedaquiline; M, moxifloxacin; Pa, PA-824/pretomanid; R, rifampicin; Z, pyrazinamide.

Combination	Culture	No. Patients	No. Regimens
P	Solid	58	1
S	Solid	54	1
PS	Solid	37	1
HRpSZ	Solid	83	2
SZ/EH	Solid	214	2
ES/HP	Solid	195	1
SZ/HZ	Solid	198	1
HP	Solid	424	5
HT	Solid	522	5
SHP/HP	Solid	83	1
SZ	Solid	206	3
SHP	Solid	969	7
H	Solid	533	6
SH	Solid	849	8
HST	Solid	442	5
OHRZ	Solid	53	1
EHS/EH	Solid	214	2
EH	Solid	38	2
HS/HS	Solid	327	2
HSZ	Solid	319	3
HPS/HPS	Solid	171	1
HR	Solid	1415	10
R	Solid	77	2
SHRZ/HT	Solid	162	1
ER	Solid	135	2
SHRZ/HSZ	Solid	191	1
EHS	Solid	157	3
HRS	Solid	1500	13
PR	Solid	42	1
HRS/HR	Solid	413	4
EHR	Solid	919	13
HOR	Solid	62	1
ERZ	Solid	71	2
HPR	Solid	120	2
EHRZ	Solid	171	1
GHRZ	Solid	793	2
HRZ	Solid	1492	15
EHRpZ	Solid	198	1
MHRZ	Solid	688	4
SHRZ	Solid	3479	25
EHRZ	Solid	153	1
HRZE	Solid	1818	8
SHRZ/HR	Solid	278	4
EMRZ	Solid	701	2
EHRbZ	Solid	107	1



**Figure 6.** Culture negativity at 8 weeks: solid culture (random effects, DerSimonian and Laird method). Abbreviations: -ve, negative; E, ethambutol; G, gatifloxacin; H, isoniazid; J, bedaquiline; L, levofloxacin; M, moxifloxacin; O, ofloxacin; P, para-aminosalicylic acid; Pa, PA-824/pretomanid; R, rifampicin; Rb, rifabutin; Rp, rifapentine; S, streptomycin; T, thiacetazone; Z, pyrazinamide.

endpoints. The ranking, however, was highly constrained by the limited number of drugs and regimens studied in both phases, principally because of the ethical unacceptability of prolonged monotherapy in 2-month studies and the lack of historical combination EBA studies.

Nine distinct regimens (HRZE, H, J, JPa, JZ, Pa, PaMZ, PaZ, and R) were considered on at least 2 phase 2A endpoints. All 9 were considered when examining EBA<sub>0-2</sub>, and all except Pa were considered when examining EBA<sub>0-7</sub>. Seven were considered (H, HPa, HZ, Pa, PaMZ, and PaZ) when examining EBA<sub>0-14</sub>. Only 3 of these regimens (HRZE, H, and R) were considered for 2-month culture conversion on solid media.

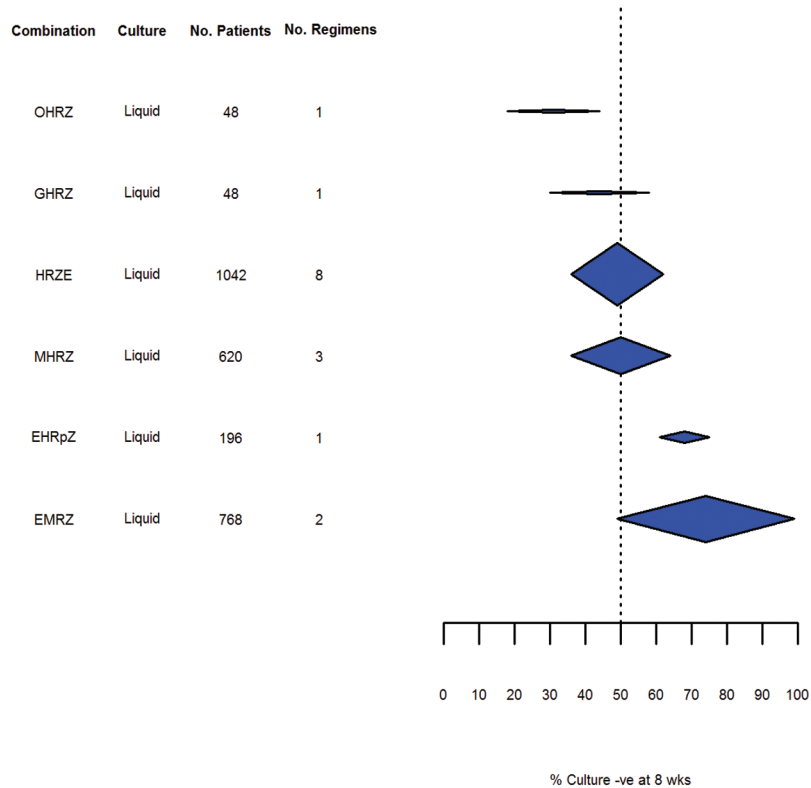
Using the relative order among the drugs in the common sets and basing results only on solid culture, the rankings are shown in (Table 2). Though qualitative rankings for the available regimens were reasonably consistent, the dataset was too small to be able to draw conclusions about their usefulness for decision making.

## DISCUSSION

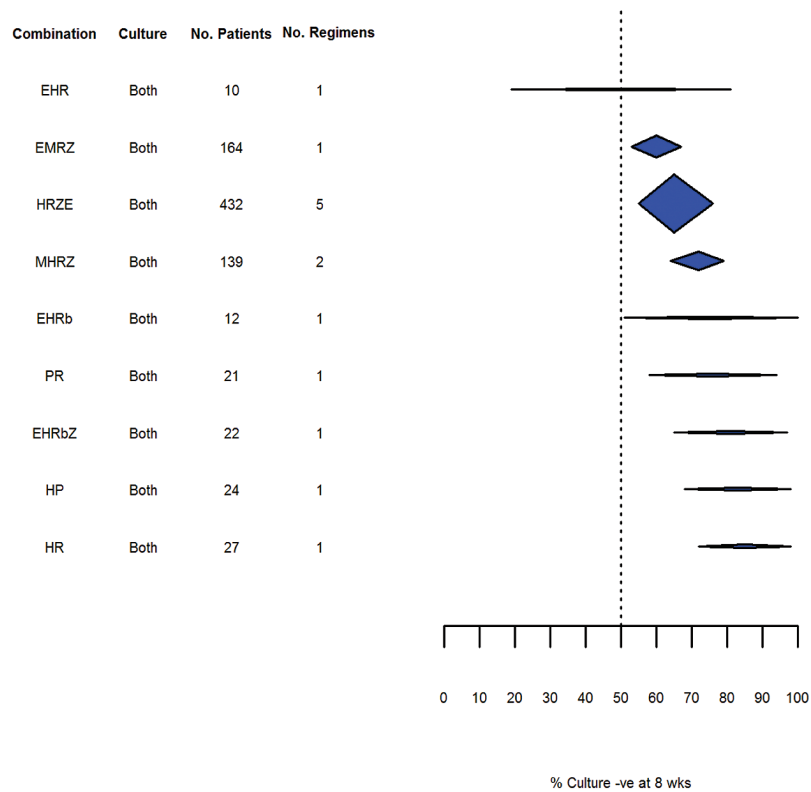
This review is the first to systematically appraise the performance of single drugs and combination regimens across early clinical endpoints in trials of treatment of pulmonary tuberculosis.

Though we focussed only on the set of drugs most relevant to historical and modern first-line therapy, we identified 133 trials reporting phase 2A and 2B outcomes comprising >37 000 patients and 67 drug combinations. However, the diversity of treatment regimens represented in 14-day phase 2A studies was much lower, with only 9 drug combinations assessed on >1 phase 2A endpoint and only 3 of these combinations assessed in both phase 2A and 2B trials. While these findings partly reflect the history, development, and differing goals of such trials, the narrowness of this evidence base is concerning and suggests a potentially serious gap in rational translation between these 2 critical phases of development. While rankings of the efficacy of treatment appeared reasonably consistent on different phase 2A endpoints, the existing dataset does not provide convincing support to current practices, and intertrial variability was high in many cases.

We selected 4 outcome measures for our review based on those most commonly reported in the included studies [15]. However, there was large variation in reporting, particularly of EBA measures, with many unique to a single study. Overall the quality of reporting, particularly of phase 2A studies, made data extraction and synthesis challenging and imposes limitations in



**Figure 7.** Culture negativity at 8 weeks: liquid culture (random effects, DerSimonian and Laird method). Abbreviations: -ve, negative; E, ethambutol; G, gatifloxacin; H, isoniazid; M, moxifloxacin; O, ofloxacin; R, rifampicin; Rp, rifapentine; Z, pyrazinamide.



**Figure 8.** Culture negativity at 8 weeks: solid and liquid culture (random effects, DerSimonian and Laird method). Abbreviations: -ve, negative; E, ethambutol; H, isoniazid; M, moxifloxacin; P, para-aminosalicylic acid; R, rifampicin; Rb, rifabutin; Z, pyrazinamide.



**Table 1. Meta-regression Results for the Impact of Selected Variables on the Proportion of Patients Who Were Culture Negative at 8 Weeks**

Drug Combinations <sup>a</sup>	Year of Publication	Covariate: $\beta$ Regression Coefficient (P Values)	
		Proportion of Patients With HIV Coinfection	Geographical Location (Africa or Not)
HRZE	-0.00 (.27)	0.00 (.79)	0.01 (.87)
HRZ	0.01 (.24)	Insufficient data <sup>b</sup>	-0.12 (.45)
SHRZ	0.00 (.57)	Insufficient data <sup>c</sup>	-0.08 (.17)

$\beta$  coefficients describe how the outcome variable changes with a unit increase in the explanatory variable.

Abbreviation: HIV, human immunodeficiency virus.

<sup>a</sup>Drug abbreviations: E, ethambutol; H, isoniazid; R, rifampicin; S, streptomycin; Z, pyrazinamide.

<sup>b</sup>Only 1 study presented proportion of patients with HIV coinfection.

<sup>c</sup>No studies presented proportion of patients with HIV coinfection.

interpretation of the data. The striking feature of the available dataset is the variability of pooled estimates of effect for all the endpoints examined. For EBA<sub>0-2</sub> and 2-month culture conversion, this variability was particularly marked, with overlapping confidence intervals for the majority of regimens. Though there were appreciable differences between the best-performing regimens on these endpoints (H, HRZ, HRZE, and SHRZ) and the worst (Z, S, SH, and SHP), this suggests that such trials may lack the power to formally discriminate between regimens where differences in treatment effect are more modest but still clinically relevant. The reasons for this variability were difficult to explore using the data available, given the quality and consistency of reporting.

We used the data as reported—some studies adopted an intention-to-treat approach to analysis and included patients with missing or contaminated culture results, while others used a per-protocol approach and excluded these patients. Poor-quality reporting meant it was mostly impossible to distinguish these situations. This may account for some of the observed heterogeneity.

**Table 2. Ranking of Drugs Across Outcomes Based on a Subset of Regimens for Which at Least 2 of the Early Bactericidal Activity Results Were Available**

Drug Combinations <sup>a</sup>	Ranking and No. of Patients (No. of Regimens)			
	EBA <sub>0-2</sub>	EBA <sub>0-7</sub>	EBA <sub>0-14</sub>	2 Months
H	1 149 (16)	1 36 (4)	...	3 533 (6)
HRZE	2 51 (6)	4 21 (2)	2 50 (6)	1 1618 (8)
PaMZ	3 15 (1)	3 12 (1)	1 13 (1)	...
R	4 28 (3)	2 13 (1)	...	2 77 (2)
PaZ	5 15 (1)	5 14 (1)	3 14 (1)	...
Pa	6 29 (2)	...	6 26 (2)	...
JPa	7 14 (1)	6 14 (1)	5 14 (1)	...
JZ	8 15 (1)	7 15 (1)	4 15 (1)	...
J	9 41 (3)	8 26 (2)	7 28 (2)	...

Abbreviations: EBA<sub>0-2</sub>, early bactericidal activity over 2 days; EBA<sub>0-7</sub>, early bactericidal activity over 7 days; EBA<sub>0-14</sub>, early bactericidal activity over 14 days.

<sup>a</sup>Drug abbreviations: E, ethambutol; H, isoniazid; J, bedaquiline; M, moxifloxacin; Pa, PA-824/pretoamanid; R, rifampicin; Z, pyrazinamide.

The variance of the pooled estimates for EBA endpoints may be inflated by the regression coefficients being based on different numbers of observations. There were no such methodological problems for the 2-month culture conversion results, suggesting that the observed heterogeneity is likely to be a real clinical effect. Among the most likely sources of this within-regimen heterogeneity are pharmacological (ethnic differences in absorption, elimination, and clearance), bacteriological (differences in initial bacterial burden, and virulence), and patient factors (disease stage such as presence of cavities; comorbidities such as malnutrition, diabetes, and HIV). Our meta-regression analysis was able to explore a very limited subset of such variables for a few of the most common regimens on a single endpoint.

Because the review incorporated all reported trials over the last 60 years, evolution in the efficiency and standardization even of solid culture methods may have contributed to variation in 2-month culture results, with older studies tending to produce numerically higher rates (not statistically significant) of culture conversion due to lower assay sensitivity compared to more modern methods. We tested this via meta-regression for drug combinations HRZE, HRZ, and SHRZ and in all 3 cases, year of publication was not significant. Isoniazid resistance, whether known or undetected, may have tended to increase heterogeneity in outcomes, but this is difficult to assess due to small numbers of patients. While it is known that, at least for modern regimens such as HRZE, the risk of poor outcome for patients with H-resistant organisms is only modestly increased [16, 17], this could be more important for older regimens, although there were insufficient data to test this in a meta-regression. Finally, patient factors—in particular, chronicity of disease, geographical location, and HIV coinfection—could also have increased inter-trial variability within regimens, although there were insufficient data to test this for chronicity of disease. It has been observed that culture conversion at 2 months may vary widely even between study sites in individual trials [18] and may be influenced by the lower sputum bacillary load observed in HIV-infected patients.

We propose 3 approaches that could help to overcome some of the limitations this review identifies in the existing evidence base for phase 2 trials in tuberculosis. First, assembly of a database of individual patient data relating to the trials identified would facilitate reanalysis of the trials and also enable computation of endpoints not reported in the original study publication, which may help to address the lack of diversity of regimens on each endpoint. Such an effort is currently in progress by the model-based preclinical development of anti-tuberculosis drug combinations (PreDiCT-TB) consortium, and we anticipate an update of this review based on individual patient data when that process has been completed.

Second, development of a core outcome set for tuberculosis trials that could be applied to new studies in the field would likely assist both investigators and systematic reviewers in choosing and reporting endpoints in such a way that the



contribution of each trial to the overall evidence base is maximized [19]. This would provide a minimum but not exhaustive set of clearly defined outcomes to be reported in each study.

Finally, we also suggest wider use of novel, more efficient, adaptive screening trial designs that would enable a broader range of regimens to be studied in phase 2 than was previously possible. However, such trials may also impact any meta-analysis to which they contribute in terms of the endpoints that they prioritize and any bias they might introduce due to early termination.

Our review shows that the existing evidence base supporting phase 2 methodology in tuberculosis is highly incomplete. To truly understand and improve drug development in tuberculosis, it is desirable that a broader range of drugs and combinations be more consistently studied across a greater range of phase 2 endpoints than is currently available and that these regimens be rigorously compared in a cumulative meta-analytic framework. Although this review forms an initial contribution, achieving this goal will require a coordinated and multidisciplinary effort by the tuberculosis trials community.

### Supplementary Data

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

### Notes

**Acknowledgments.** The authors thank Sarah Donegan for her review of this manuscript.

**Financial support.** The research leading to these results has received funding from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115337, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007–2013) and European Federation of Pharmaceutical Industries and Associations companies' in kind contribution.

**Potential conflicts of interest.** G. C. K. W. K. is an employee of and holds shares in GlaxoSmithKline. No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

### References

1. Lienhardt C, Raviglione M, Spigelman M, et al. New drugs for the treatment of tuberculosis: needs, challenges, promise, and prospects for the future. *J Infect Dis* **2012**; 205(suppl 2):S241–9.
2. Donald PR, Diacon AH. The early bactericidal activity of anti-tuberculosis drugs: a literature review. *Tuberculosis (Edinb)* **2008**; 88(suppl 1):S75–83.
3. Jindani A, Aber VR, Edwards EA, Mitchison DA. The early bactericidal activity of drugs in patients with pulmonary tuberculosis. *Am Rev Respir Dis* **1980**; 121:939–49.
4. Diacon AH, Dawson R, von Groote-Bidlingmaier F, et al. 14-day bactericidal activity of PA-824, bedaquiline, pyrazinamide, and moxifloxacin combinations: a randomised trial. *Lancet* **2012**; 380:986–93.
5. Rustomjee R, Lienhardt C, Kanyok T, et al; Gatifloxacin for TB (OFLOTUB) study team. A phase II study of the sterilising activities of ofloxacin, gatifloxacin and moxifloxacin in pulmonary tuberculosis. *Int J Tuberc Lung Dis* **2008**; 12:128–38.
6. Phillips PP, Fielding K, Nunn AJ. An evaluation of culture results during treatment for tuberculosis as surrogate endpoints for treatment failure and relapse. *PLoS One* **2013**; 8:e63840.
7. Wallis RS, Wang C, Meyer D, Thomas N. Month 2 culture status and treatment duration as predictors of tuberculosis relapse risk in a meta-regression model. *PLoS One* **2013**; 8:e71116.
8. Wallis RS, Peppard T, Hermann D. Month 2 culture status and treatment duration as predictors of recurrence in pulmonary tuberculosis: model validation and update. *PLoS One* **2015**; 10:e0125403.
9. Center for Drug Evaluation and Research. Pulmonary tuberculosis: developing drugs for treatment. Silver Spring, MD: US Food and Drug Administration, **2013**.
10. Davies GR, Phillips PP, Jaki T. Adaptive clinical trials in tuberculosis: applications, challenges and solutions. *Int J Tuberc Lung Dis* **2015**; 19:626–34.
11. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. Available at: [www.cochrane-handbook.org](http://www.cochrane-handbook.org); Cochrane Collaboration, **2011**.
12. R Cran Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, **2015**.
13. Kennedy N, Fox R, Kisyombe GM, et al. Early bactericidal and sterilizing activities of ciprofloxacin in pulmonary tuberculosis. *Am Rev Respir Dis* **1993**; 148(6 pt 1):1547–51.
14. Snider DE, Graczyk J, Bek E, Rogowski J. Supervised six-months treatment of newly diagnosed pulmonary tuberculosis using isoniazid, rifampin, and pyrazinamide with and without streptomycin. *Am Rev Respir Dis* **1984**; 130:1091–4.
15. Bonnett LJ, Davies GR. Quality of outcome reporting in phase II studies in pulmonary tuberculosis. *Trials* **2015**; 16:518.
16. Aber VR, Nunn AJ. Factors affecting relapse following short-course chemotherapy. *Bull Int Union Tuberc Lung Dis* **1978**; 53:260–4.
17. Espinal MA, Kim SJ, Suarez PG, et al. Standard short-course chemotherapy for drug-resistant tuberculosis: treatment outcomes in 6 countries. *JAMA* **2000**; 283:2537–45.
18. Burman WJ, Goldberg S, Johnson JL, et al. Moxifloxacin versus ethambutol in the first 2 months of treatment for pulmonary tuberculosis. *Am J Respir Crit Care Med* **2006**; 174:331–8.
19. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* **2007**; 8:39.