Tuberculosis 94 (2014) 317–322

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

DIAGNOSTICS

How we determined the most reliable solid medium for studying treatment of tuberculosis☆

Charles M. Heilig [a,*], Pei-Jean I. Feng [a], Moses L. Joloba [b,c], John L. Johnson [c,d], Karen Morgan [c], Phineas Gitta [c], W. Henry Boom [c,d], Harriet Mayanja-Kizza [c], Kathleen D. Eisenach [d,e], Lorna Bozeman [a], Stefan V. Goldberg [a]

[a] Division of Tuberculosis Elimination, Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Atlanta, GA, USA
[b] Department of Medical Microbiology, College of Health Sciences, Makerere University, Kampala, Uganda
[c] Uganda-Case Western Reserve University Research Collaboration, Kampala, Uganda
[d] Tuberculosis Research Unit, Department of Medicine, Case Western Reserve University School of Medicine, Cleveland, OH, USA
[e] Department of Pathology, University of Arkansas for Medical Sciences, Little Rock, AR, USA

ARTICLE INFO

Article history:
Received 19 April 2013
Received in revised form
20 February 2014
Accepted 21 February 2014

Keywords:
Middlebrook agar culture media
Lowenstein–Jensen culture medium
Composite reference standard
Latent-class model
Mycobacterium tuberculosis

SUMMARY

Phase 2 clinical trials for tuberculosis (TB) treatment require reliable culture methods to determine presence or absence of *Mycobacterium tuberculosis* (Mtb) over the course of therapy, as these trials are based primarily on bacteriological endpoints. We evaluate which of 5 solid media is most reliable: Lowenstein–Jensen (LJ) egg-base medium and 4 Middlebrook agar media (nonselective 7H10 and 7H11 and selective 7H10 and 7H11). We analyze 393 specimens from 50 HIV-negative Ugandan adults with newly-diagnosed, pulmonary TB and high acid-fast bacillus smear grade. Specimens were collected every 2–4 weeks during the first 12 weeks of therapy. We compare the results for each culture to 2 composite reference standards—one that was deemed positive if any solid culture was positive for Mtb and another based on latent-class analysis. Both reference standards established that the 2 selective Middlebrook media most reliably determine the presence or absence of Mtb ($P < 0.003$), largely because of their lower contamination rates. We also showed that results on Middlebrook media were similar to each other, while LJ was most frequently discordant. Contaminated results appeared more likely to be truly negative than to harbor undetected Mtb.

Published by Elsevier Ltd.

1. Introduction

Phase 2 clinical trials for tuberculosis (TB) treatment require reliable culture methods to determine presence or absence of *Mycobacterium tuberculosis* (Mtb) over the course of therapy, as phase 2 trial endpoints are based primarily on bacteriological status at baseline and multiple time points during treatment and follow-up. Historically, most trials have involved culture on solid media, particularly egg-base, locally made Lowenstein–Jensen media. In contemporary trials, sputum often is cultured on both solid and liquid media. Even with the advent of more sensitive liquid culture methods and other methods for detecting Mtb, solid media will continue to provide an important bridge to historical data linking solid media culture results during therapy to clinical outcomes such as treatment failure and relapse. Many solid media are available for mycobacterial culture, and little systematic research has been done to compare the utility of different solid media to determine which is best for use in TB treatment trials.

We conducted a prospective cohort study comparing 5 solid media for recovery of Mtb during treatment of adults with pulmonary TB with standard chemotherapy: Lowenstein–Jensen (LJ) egg-base medium and 4 Middlebrook agar media (nonselective 7H10 and 7H11 and selective 7H10 and 7H11 [7H10S and 7H11S]). Each of these has been used for both diagnostic and research purposes. Which one is the most reliable? By reliable we mean that the culture method (1) correctly indicates Mtb growth when it is present in a sputum specimen, (2) correctly indicates a lack of growth

☆ The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

* Corresponding author. Division of Tuberculosis Elimination, Centers for Disease Control and Prevention, 1600 Clifton Road NE, MS E-10, Atlanta, GA 30329, USA.
E-mail address: cqh9@cdc.gov (C.M. Heilig).

http://dx.doi.org/10.1016/j.tube.2014.02.006
1472-9792/Published by Elsevier Ltd.

when Mtb is absent from a sputum specimen, within the medium's limits of detection, and (3) produces evaluable results as often as possible by minimizing contamination and other sources of interference. Several culture methods reliably detect Mtb for diagnostic purposes, that is, prior to beginning TB therapy. There is, however, no consensus regarding which method is the most reliable for repeated assessments within a clinical trial as the bacillary burden diminishes in response to therapy. We need a benchmark or standard against which to compare candidate culture methods, with the intention to select the method that best conforms to the benchmark reference standard.

In this paper, we explore data-derived benchmarks, also known as composite reference standards [2, 21 (p. 19–21)], constructed from the combined results of the 5 separate culture methods. In synthesizing these discrete results, we must have a principled way to reconcile apparently contradictory patterns, wherein we can observe any mixture of positive, contaminated, and negative cultures from a single specimen. Consequently, we pose 2 methodological questions in connection with constructing reference standards to characterize the reliability of solid culture media: First, how do we resolve the situation in which results on different media contradict each other? Second, what information is conveyed by contaminated culture results?

To address our research questions, we considered 2 composite reference standards: one method that infers the presence of Mtb when any of the 5 solid media yields a positive culture result, and a second method using latent-class analysis (LCA) to optimize information about the presence of Mtb from all 5 solid media culture results. LCA has been used in other applications to TB [12,16,23,24], and several overviews are available [8,13,21].

We examined how each of these constructs informs our understanding of the reliability of each culture method, provides clues about how to resolve apparently contradictory culture results, and yields statistical information about contaminated results, with further implications about the impact of contamination on statistical analysis of culture-based study outcomes. In a separate manuscript [14], we applied LCA and concluded that the 2 selective media were the most reliable, with 7H11S showing superior ability to detect Mtb when it is present. In this manuscript, we expand on our methods, provide additional motivation for our latent-class approach, and give a more complete justification of our conclusions.

2. Methods

2.1. Setting

From August 2009 to August 2010, 50 participants were enrolled in an observational study conducted at the National Tuberculosis Treatment Centre, Mulago Hospital, Kampala, Uganda. All participants were HIV-negative adults over 18 years old with newly-diagnosed, cavitary, pulmonary TB with sputum that is positive for acid-fast bacilli (AFB). We selected participants with AFB smear grade 3+ or 4+ (≥1 AFB per high-power field [20]) because they have a high number of tubercle bacilli in the sputum at diagnosis, are the most frequent type of person enrolled in TB treatment trials, and are more likely to experience poor treatment outcomes. Participants received standard combination anti-TB treatment in accordance with CDC-ATS-IDSA guidelines [6]; all available specimens were included in this analysis, even if a participant's regimen was changed because of drug resistance or other factors. The study protocol was approved by research ethics committees at the Joint Clinical Research Centre, Case Western Reserve University, and the US Centers for Disease Control and Prevention, and by the Ugandan National Council for Science and Technology. All participants freely consented in writing.

2.2. Sputum collection

Clinic staff instructed patients in standard sputum collection procedures at baseline and follow-up visits. Two on-the-spot, deep-cough sputum specimens were collected prior to treatment. During TB treatment, 1 on-the-spot sputum specimen was obtained at weeks 2, 4, and 6, 2 specimens at week 8, and 1 specimen every 4 weeks thereafter through to the end of TB treatment (week 24 or 36). Each sputum specimen was cultured on 5 solid media (LJ, 7H10, 7H11, 7H10S, and 7H11S) and in liquid MGIT medium (BACTEC MGIT, Becton, Dickinson, and Company, Franklin Lakes NJ USA) (See Ref. [14] for the details of sputum preparation and culturing). For this analysis, we used data from specimens collected from baseline through week 12. Culture results on each medium were categorized as Mtb-positive, contaminated, or Mtb-negative. If a culture showed both Mtb growth and contamination, both were reported but it was categorized as Mtb-positive; cultures with contamination and no detectable Mtb growth were regarded as contaminated.

2.3. Data analysis

We characterized the pairwise discordance between solid media by the frequency with which the 2 media directly contradicted each other—that is, where 1 medium was Mtb-positive, and the other was Mtb-negative.

We constructed 2 reference standards by using information from all 5 solid media and compared the results of individual solid media to these reference standards. Our analytic tasks for each reference were to determine which solid medium agreed with this constructed reference most often, how to interpret discrepancies for specimens that yielded contradictory results, how to interpret contaminated cultures relative to the constructed standard, and whether the construct was ultimately credible.

We first constructed an intuitively appealing reference, called here the **any-positive construct**, which was positive if any of the 5 solid media showed Mtb growth and negative if none showed Mtb growth. For this construct, we assumed that, with proper quality control, observed Mtb growth was real. Furthermore, when 1 medium showed growth and another was Mtb-negative or contaminated, the negative result could reflect inhomogeneous sputum preparation or the insensitivity of 1 medium relative to another. We looked closer at this construct by considering patterns in which exactly 1 culture medium was positive (and at least 1 other was negative) and patterns in which exactly 1 culture medium was negative (and at least 1 other was positive), with particular attention to how often each medium tends to agree or disagree with the other media. We also computed the pairwise concordance between solid media.

We next constructed a reference, called here the **latent-class construct**, which used the statistical information from all 5 solid media. The latent-class construct was positive if, according to a statistical model based on latent-class analysis (LCA), the pattern of all 5 results on solid media was more likely to indicate the presence of Mtb than its absence. The observed combinations of solid media results were separated into 2 sets, based on latent classes, in a way that optimized the statistical likelihood function. Compared to the any-positive construct, the latent-class procedure can give greater insight into the variability in the data.

We used the bootstrap resampling method [7] to characterize the joint variability of the parameter estimates in both our models. Technical details of model assumptions, model selection (including