Contemporary Clinical Trials 33 (2012) 1132–1142

Contents lists available at SciVerse ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Antibiotic retreatment of Lyme disease in patients with persistent symptoms: A biostatistical review of randomized, placebo-controlled, clinical trials

Allison K. DeLong [a,*], Barbara Blossom [b], Elizabeth L. Maloney [c], Steven E. Phillips [d]

[a] Center for Statistical Sciences, Department of Biostatistics, Brown University, Providence, RI, USA
[b] Department of Statistics, Colorado State University, Fort Collins, CO, USA
[c] Partnership for Healing and Health, Ltd., Wyoming, MN, USA
[d] Greenwich Hospital, Greenwich, CT, USA

ARTICLE INFO

Article history:
Received 4 May 2012
Revised 16 July 2012
Accepted 14 August 2012
Available online 19 August 2012

Keywords:
Lyme disease
Neuroborreliosis
Randomized controlled trial
Statistical power
Minimum clinically important difference (MCID)
Treatment guidelines

ABSTRACT

Introduction: Lyme disease (Lyme borreliosis) is caused by the tick-borne spirochete *Borrelia burgdorferi*. Long-term persistent illness following antibiotic treatment is not uncommon, particularly when treatment is delayed. Current treatment guidelines for persistent disease primarily rely on findings from four randomized, controlled trials (RCTs), strongly advising against retreatment.

Methods: We performed a biostatistical review of all published RCTs evaluating antibiotic retreatment, focusing on trial design, analysis and conclusions.

Results: Four RCTs met the inclusion criteria; all examined the efficacy of intravenous ceftriaxone versus placebo at approximately 3 or 6 months. Design assumptions for the primary outcomes in the two Klempner trials and two outcomes in the Krupp trial were unrealistic and the trials were likely underpowered to detect clinically meaningful treatment effects. The Klempner trials were analyzed using inefficient statistical methods. The Krupp RCT was well-designed and analyzed for fatigue, finding statistically significant and clinically meaningful improvement. Fallon corroborated this finding. Fallon also found improvement in cognitive functioning, a primary outcome, at 12 weeks which was not sustained at 24 weeks; improvements in physical functioning and pain were demonstrated at week 24 as an interaction effect between treatment and baseline symptom severity with the drug effect increasing with higher baseline impairment.

Discussion: This biostatistical review reveals that retreatment can be beneficial. Primary outcomes originally reported as statistically insignificant were likely underpowered. The positive treatment effects of ceftriaxone are encouraging and consistent with continued infection, a hypothesis deserving additional study. Additional studies of persistent infection and antibiotic treatment are warranted.

1. Introduction

Reporting bias in clinical trials, particularly with respect to publishing bias toward significant findings [1,2] and interpretive "spin" to overemphasize a possible benefit while de-emphasizing non-significant findings [3] is receiving increased attention within the statistical and medical communities. A variation on interpretive bias deserves concern as well, namely the interpretation of statistically insignificant findings from small, underpowered, or poorly executed clinical trials as evidence of treatment inefficacy. Such trials may lead to the premature and erroneous conclusion that the treatment is ineffective, constituting a type II error. Concerns about such

* Corresponding author at: Center for Statistical Sciences, Department of Community Health, Brown University, Providence, RI 02912, USA. Tel.: +1 401 863 9697; fax: +1 401 863 9182.
E-mail address: adelong@stat.brown.edu (A.K. DeLong).

1551-7144/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cct.2012.08.009

errors may arise when disagreement and uncertainty exists in the medical community, as is the case with Lyme disease (Lyme borreliosis).

Lyme disease, caused by the tick-borne spirochete *Borrelia burgdorferi* sensu lato, is classified as an emerging infectious disease by the U.S. National Institute of Allergy and Infectious Diseases (NIAID) due to the relatively recent discovery of its causal agent (1982) [4] and its rapidly increasing incidence over the last two decades in the U.S. [5] and much of Europe [6]. The infection is multi-systemic, resulting in diverse physical and neuro-psychiatric symptoms and manifestations and causing mild to severe disease [7–13]. Although many patients respond to antibiotic treatment regimens of 2 to 4 week duration [9], it is well recognized that long-term persistent illness can occur following a 30-day course of treatment, particularly when treatment is delayed [7,9,14,15]. Multiple randomized trials found significant morbidity in their study populations, similar to that of multiple sclerosis or congestive heart failure. Although the trials employed different entrance criteria, none required this degree of physical disability as a condition of enrollment [16,17].

The management of patients with ongoing debilitating symptoms following antibiotic treatment for Lyme disease has generated debate within the medical community. The primary questions concern whether or not infection persists after standard antibiotic treatment and whether additional antibiotic treatment is of benefit [18,19]. Until a sensitive laboratory test for active infection is clinically available, clinical trials evaluating retreatment in persistently symptomatic Lyme disease patients provide the cornerstone of treatment guideline recommendations. Most guidelines for the diagnosis and management of Lyme disease [20–23] direct clinicians to limit the duration of antibiotic treatment, even in cases where ongoing symptoms compatible with a *B. burgdorferi* infection are present. These publications base their recommendations on a similar interpretation of the four randomized, blinded, placebo-controlled antibiotic retreatment trials funded by the U.S. National Institutes of Health (NIH) for patients with ongoing symptoms following standard Lyme disease treatment [16,17,24].

For this reason, a rigorous, independent evaluation of the findings from these trials is needed. The present study is a biostatistical review of the four NIH-funded clinical trials. By focusing on the trial design and analyses of primary and secondary outcomes in each trial, the review demonstrates weaknesses which limit the ability to draw strong conclusions regarding retreatment. This review will likely be of broad interest to medical practitioners, researchers, medical ethicists, and treatment guideline developers in Europe and North America.

2. Methods

The four NIH-funded Lyme disease retreatment trials were initially selected for evaluation in January 2009 through a review of current Lyme disease treatment guidelines, which identify these trials as the only published RCTs relevant to the question of retreatment [21,22]. To ensure that other relevant RCTs to date were not missed, a Cochrane Library search of the published literature was conducted on September 10, 2010, setting the limits of study type to “clinical trial” and requiring the use of “Lyme” or “Borrelia” in the title, abstract or in the manuscript's keywords. Additional studies were sought by searching ClinicalTrials.gov, a registry of both federally and privately funded clinical trials. The title and abstract of each selected publication were read by two authors (AKD and BB) and coded as a clinical trial and if it was a clinical trial evaluating retreatment of Lyme disease patients with persistent symptoms despite receipt of a standard course of antibiotics. The full text of all articles evaluating retreatment was read by all authors and eligibility was determined by consensus. All primary and secondary outcomes were tabulated for each clinical trial, including, where possible, the treatment effect and 95% confidence interval (CI) overall and by trial arm.

A review was conducted of each trial's design, execution, statistical analysis and conclusions. For trial design, attention was paid to the enrolled patient population, the definitions and measurements of primary and secondary outcomes, and the definition of clinically meaningful changes in those outcomes which determine power of the sample sizes to detect clinically meaningful treatment effects. For trial execution, patient dropout, masking of study medication, and interim analyses were considered. We evaluated the appropriateness of the statistical method chosen to estimate the treatment effect and the handling of patient dropouts. Since our objective is to place the findings from these trials within the current framework of Lyme disease as of 2012, the present review is also informed by research conducted after the retreatment trials were designed, executed, and/or published. Three important statistical concepts are used throughout the review: statistical power, interim analysis and stopping rules, and non-inferiority trials.

2.1. Statistical power

When designing a clinical trial, the sample size can only be calculated after researchers determine an appropriate and plausible design treatment effect δ, which is a hypothetical value of the effect of the treatment under investigation. In addition to selecting δ, trial design also requires an acceptable probability of declaring treatment effectiveness if δ is true (i.e. power, typically 80–90%). For a fixed power, a smaller δ would necessitate a study design with a larger sample size, and vice versa. Ideally δ should correspond to the minimum clinically important difference (MCID) for the disease and outcome measure studied. If the *true underlying treatment effect* is greater than the MCID, yet less than the design treatment effect δ, then the study is underpowered with an insufficient sample size, and thus inadequately designed to meet its stated goals, and the power may be far less than the nominal value set in the trial design. Such studies are likely to conclude an insignificant result although a true, clinically relevant treatment effect exists. Although MCID values are context-specific and difficult to ascertain, reasonable estimates are identified based on published knowledge of the disease studied or, when disease-specific data are not available, of studies of other similar diseases [25].

2.2. Interim analyses and stopping rules

Interim analyses are commonly used to gauge the success of a clinical trial, by analyzing outcome data at pre-defined points during the study instead of waiting until all patients