Published in final edited form as:
Stat Med. 2012 November 10; 31(25): 2955–2972. doi:10.1002/sim.5404.

Bayesian Clinical Trials in Action

J. Jack Lee, Ph.D. and **Caleb T. Chu, M.D., M.P.H.**
Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030 U.S.A

Abstract

Although the frequentist paradigm has been the predominant approach to clinical trial design since the 1940s, it has several notable limitations. The alternative Bayesian paradigm has been greatly enhanced by advancements in computational algorithms and computer hardware. Compared to its frequentist counterpart, the Bayesian framework has several unique advantages, and its incorporation into clinical trial design is occurring more frequently. Using an extensive literature review to assess how Bayesian methods are used in clinical trials, we find them most commonly used for dose finding, efficacy monitoring, toxicity monitoring, diagnosis/decision making, and for studying pharmacokinetics/pharmacodynamics. The additional infrastructure required for implementing Bayesian methods in clinical trials may include specialized software programs to run the study design, simulation, and analysis, and Web-based applications, which are particularly useful for timely data entry and analysis. Trial success requires not only the development of proper tools but also timely and accurate execution of data entry, quality control, adaptive randomization, and Bayesian computation. The relative merit of the Bayesian and frequentist approaches continues to be the subject of debate in statistics. However, more evidence can be found showing the convergence of the two camps, at least at the practical level. Ultimately, better clinical trial methods lead to more efficient designs, lower sample sizes, more accurate conclusions, and better outcomes for patients enrolled in the trials. Bayesian methods offer attractive alternatives for better trials. More such trials should be designed and conducted to refine the approach and demonstrate its real benefit in action.

Keywords

adaptive trial design; Bayesian paradigm; clinical trial conduct; frequentist paradigm; trial efficiency; trial ethics

1. Introduction

A clinical trial is a prospective study that evaluates the effect of interventions in humans under prespecified conditions. Clinical trials provide the most definitive mechanism for assessing the outcome of interventions and form the foundation for evidence-based medicine through reliable data. Clinical trials also represent key components in research, with the potential to change the standard of care, improve quality of health, and control costs through careful comparison of alternative treatments. The results of the first modern clinical trial, which involved the use of streptomycin to treat pulmonary tuberculosis, were published in 1948 in the United Kingdom [1]. That trial involved randomizing patients into treatment and control groups and assessing the outcome without knowledge of the treatment assignment.

Correspondence to: J. Jack Lee, Ph.D., Department of Biostatistics, The University of Texas MD, Anderson Cancer Center, 1400 Pressler Street, Unit 1411, Houston, Texas 77030, U.S.A. jjlee@mdanderson.org; Tel: 1-713-794-4158; Fax: 1-713-563-4243.

Since then, clinical trials have been widely applied in medicine for the advancement of science and the search for better treatments to improve health.

In the United States, the National Institutes of Health, particularly the National Heart, Lung, and Blood Institute (NHLBI) and the National Cancer Institute (NCI), have led the effort to develop and conduct clinical trials [2, 3]. *Clinical Trials: Past, Present and Future*, an NHLBI-sponsored workshop held in 2010, explored the significance of clinical trials by examining their historical development, surveying their present use and impact on medicine, and discussing the future direction of clinical trials. This paper examines the emerging use of Bayesian methods in clinical trials, focusing on their expanding implementation and impact on medicine.

From a statistical framework point of view, the frequentist paradigm has dominated the field of clinical trials over the past 60 years. Considering the treatment effect, θ, which is the unknown parameter of interest, the frequentist framework assumes that θ is fixed, yet unknown. Through clinical trials, we can collect data to inform θ. Hence, the inference on the treatment effect can be made by evaluating the probability: Prob(data|θ), where the data are considered to be random and the parameter θ is fixed. Conversely, the Bayesian framework assumes that the data are fixed and the unknown parameter θ is random. Bayesian inference is made by computing Prob(θ|data). Thanks to the work of R.A. Fisher, J. Neyman, and K. Pearson, among others, the frequentist theory was well developed in the early 1900s [4]. Compared to the Bayesian methods, frequentist probability calculation is simpler and less computationally intensive. As a result, the frequentist framework became the mainstream of statistics and was quickly adopted into clinical trials as they evolved. Despite its usefulness and proven success in clinical trials, the frequentist framework suffers from some major deficiencies. Most notably, frequentist inference on the parameter of interest, θ, is made indirectly as it calculates Prob(data|θ) and not Prob(θ|data), as Bayesian inference does. In-depth comparisons between the frequentist and Bayesian approaches can be found in the literature [5, 6]. In this paper, our focus is on the use of Bayesian methods in clinical trials; in particular, on their implementation and impact on medicine. The rest of the paper is organized as follows. Section 2 describes the unique strength of the Bayesian paradigm. Section 3 discusses the barriers for Bayesian clinical trials and efforts to overcome them. Section 4 gives a brief overview of the various schools of Bayesian methods. Section 5 shows the results of our literature review, through which we sought to determine how Bayesian methods are used in clinical trials. Section 6 presents the MD Anderson Cancer Center experience in the design and conduct of Bayesian clinical trials. Section 7 concludes with further discussion and a glimpse into the future.

2. Unique strengths of the Bayesian paradigm

From the historical account, the concept of the Bayesian approach by Reverend Thomas Bayes was published posthumously in 1763 (with the help of his friend Richard Price) – long before the frequentist methods became popular [7, 8]. The now famous Bayes theorem states that the posterior probability of θ can be calculated proportional to the product of the prior probability of θ and the data likelihood, i.e., Prob(θ|data) $\propto$ P(θ) Prob(data|θ). This plain yet profound theorem was largely ignored in the early days (with the notable exception of Pierre-Simon Laplace), but was reinvigorated in the mid 1900s thanks to the work of Jeffreys, de Finetti, Good, Savage, de Groot, Lindley, Cornfield, and Zeller, among many others [9]. Of note, Jerry Cornfield worked at the Public Health Service/National Cancer Institute from 1947 to 1958 and at the National Heart Institute from 1960 to 1967 and played a key role in bringing Bayesian thinking to clinical trial development [10, 11]. Ashby comprehensively reviewed the development of Bayesian statistical methodology in clinical trials [12], while Grieve gave his personal account of the use of Bayesian methods in the

pharmaceutical industry [13]. The Bayesian framework has several unique advantages over its frequentist counterpart. We describe the key strengths of the Bayesian method in this section.

2.1. Bayesian methods conform to the likelihood principle

The likelihood principle states that all evidence of an unknown parameter θ, which is obtained from an experiment, is contained in the likelihood function of θ for the given data. In other words, all relevant information for making inference on θ is contained in the observed data and not in other unobserved quantities [14]. This is simple and logical. However, many of the frequentist inferences, such as the ones based on the *P* value or the coverage probability of a confidence interval, violate the likelihood principle because the inference depends on the unobserved data. In contrast, the Bayesian approach conditions on the data and conforms to the likelihood principle. As a result, frequentist inference is valid only when the prespecified clinical trial design is followed. When the study conduct deviates from the original design, frequentist inference suffers and adjustments are difficult. On the other hand, Bayesian inference is conditioned on the data and not on the design, so it can still maintain validity as long as the prior distribution and the probability model are correctly specified.

2.2. Bayesian methods model the unknown parameter with a distribution and properly address various levels of uncertainty. The Bayesian approach is ideal for hierarchical models

Under the Bayesian framework, all unknown parameters are random and follow certain probability distributions. The distribution parameters, themselves, are also unknown and can be modeled with hyper priors. Thus, the Bayesian method is intrinsically hierarchical. Under such a model, the uncertainty of the random variable can be captured by the variance of the distribution. Hierarchical models are commonly used in clinical trials. For example, patients undergo operations conducted by surgeons, and surgeons work within hospitals. Although the frequentist method can apply to mixed models by modeling the treatment effect as fixed and the surgeon and hospital effects as random, the construct is somewhat arbitrary in designating a certain effect as fixed or random. The frequentist approach requires a stronger assumption that observations are identical and independent. By assuming the parameter is fixed, frequentist methods often underestimate the variability of the parameter of interest. Conversely, the Bayesian method appropriates different levels of variability naturally under the hierarchical model assumption. The de Finetti theorem states that subjects enrolled in the clinical trial are exchangeable if and only if the probability of the observed data can be expressed as the data likelihood given the parameter that is integrated over the prior distribution of the parameter. Exchangeability implies conditional independence of the data given the parameter, which nicely fits the clinical trial setting [15].

2.3. Bayesian methods formally incorporate prior information gathered before, during and outside of the trial

Typically, the concept for initiating a clinical trial does not arise from an information vacuum, but is developed because of intriguing information found before the trial. To design a trial, frequentists use the prior information in an *ad hoc* way to make assumptions on the parameter of interest; whereas Bayesians elicit the prior distribution for θ and formally incorporate it to make an inference. Although the prior distribution assumption may be subject to criticism, it is spelled out explicitly and its impact can be evaluated by the sensitivity analysis. In addition, the Bayesian framework allows for the incorporation of information of two types: that which accumulated in the trial and that which was obtained outside of the trial. Incorporating both types of information into the analysis strengthens the

evidence for making an inference. This provides an ideal tool for monitoring a clinical trial, as well as for synthesizing information across multiple trials as a meta-analysis [16, 17].

2.4. Bayesian methods allow for more frequent monitoring and interim decision making during the trial

By definition, Bayesian methods provide a platform for sequential learning. The prior distribution is updated by the data to form the posterior distribution. The formed posterior distribution then becomes the prior distribution for a future evaluation. Many clinical trials are conducted over an extended period of time, and it is desirable to frequently monitor the interim results of such trials so decisions can be made early when sufficient evidence has accumulated. Although group sequential methods have been well developed under the frequentist paradigm [18]; frequentist properties are directly affected by the number and timing of interim analyses. In contrast, Bayesian methods do not impose a penalty on sequential learning. Another main difference between the two approaches is that the frequentist approach makes interim decisions based on the conditional power, which is calculated by fixing the parameter of interest at a certain value. The Bayesian approach calculates the predictive probability by integrating the conditional power over the distribution of θ because the parameter θ is random. The predictive probability factors in the uncertainty of θ, whereas the conditional power does not [19].

2.5. Bayesian methods give direct answers to the questions that most people want to ask

By addressing the question directly, Bayesian methods calculate the probability of θ given the data, and can answer a question such as, “For the new treatment, what is the probability that the success rate is more than 80%?” or “What is the probability that the true success rate lies between the interval of (0.76, 0.92)?” Frequentists calculate the probability of the observed data given a certain hypothesis, but they cannot answer the question, “What is the probability that the null or alternative hypothesis is true?” Similarly, the frequentist confidence interval is random because the data are random. The frequentist approach can be used to calculate the probability that such an interval covers the true parameter if the process is repeated many times, i.e., the long-range frequentist property; however, it cannot be used to determine the coverage rate containing the true parameter for a given confidence interval. Frequentists have to constantly explain to non-statisticians that the *P* value is *not* the probability of the null hypothesis being true and that the 95% confidence interval does *not* contain the true parameter 95% of the time. In contrast, Bayesian methods deal with the problem head-on and give direct answers to the questions that most people want to ask.

2.6. Bayesian methods provide a uniform way to solve complex problems

Both frequentist and Bayesian approaches need to formulate the problem under investigation by specifying the probability model and identifying the parameter of interest. To deal with any problem, the Bayesian method is straightforward under the skeleton of Bayesian 1-2-3: specifying the prior distribution of the parameter of interest – observing the data – updating the information by computing the posterior distribution. This provides a consistent and coherent statistical framework under which to formulate research questions and quantify the information at hand to provide answers to those questions. This method can be universally applied to simple and complex problems.

2.7. Bayesian methods can incorporate the utility function for informed decision making

In the Bayesian theoretic approach, clinical trial investigators can specify the “utility” or “loss” of various events. For example, “what is the utility (or importance) of curing cancer and what is the negative utility (loss) of developing a long-term toxicity due to the treatment?” The optimal decision of the best treatment for a given patient can be made by

maximizing the utility function or minimizing the loss function. Bayesian methods allow subjective opinions to be incorporated into the specification of the prior distribution and the utility function. Different people can have different levels of prior belief and different preferences as they rate the relative importance of events, such as being cured or suffering treatment-related toxicity. Bayesian methods formulate these components explicitly and quantitatively to aid investigators in making an informed decision.

2.8. Bayesian methods use a "learn as we go" approach. This real-time learning feature forms the basis of adaptive clinical trial designs

As previously stated, the Bayesian method is a sequential learning method and takes a "learn as we go" approach. It naturally adapts to the data and to all relevant information at hand. Traditional clinical trial designs and conduct are less adaptive, often lead to large trials over an extended period of time, and are extremely expensive. Adaptive designs have been proposed with the aim of creating more efficient, more flexible, and more ethical designs by making design changes based on the interim data. The Bayesian framework naturally and ideally fits into the development of adaptive designs [20]. Bayesian adaptive approaches are especially useful in the following three areas: adaptive randomization, interim monitoring for early stopping, and sample size adjustments. (1) Outcome adaptive randomization is used to assign more patients into more effective treatment arms as data accumulate in the trial. (2) Interim monitoring of the study endpoint uses predictive probability. If there is convincing evidence that the new treatment is superior to the standard treatment, the trial can be stopped early for efficacy. On the other hand, if the data strongly suggest that the new treatment is worse or not better than the standard treatment, the trial can be stopped early for futility. (3) Adaptive sample size estimation can be achieved by calculating the probability for a successful trial given the current result and the proposed sample size. If the current result is promising but the sample size is not adequate, the study can be expanded by increasing the sample size to have a high probability of reaching a definitive conclusion at the end of study. In 2010, the U.S. FDA issued a guidance document for adaptive clinical trials. [21] Although much of its content is in the frequentist framework, it also points out the usefulness of the Bayesian approach in adaptive designs.

3. Barriers for Bayesian clinical trials and efforts to overcome them

Despite the much earlier work of Bayes in the 1760s and Laplace in the following decades, Bayesian approaches were largely limited to a philosophical and theoretical context until they were reinvigorated in the mid 1900s. Two major barriers have prevented Bayesian methods from becoming popular: the inherent computational demands and the use of subjective information. We discuss these two barriers and the efforts to overcome them.

3.1. Computational demands

In the first 200 years of its existence, the Bayesian approach could solve only a few special cases when conjugate priors were available. Calculating the posterior distribution was extremely difficult for general cases without good computing algorithms or the use of powerful computers. This two-century stagnancy changed in the 1980s to 1990s with the advent and development of the Markov chain Monte Carlo method (MCMC) [22, 23]. By constructing a Markov chain with the desired distribution as its equilibrium state, MCMC can construct complex posterior probability distributions based on Monte Carlo samples. During the development of MCMC, personal computers and workstations were also becoming more available and more powerful, which made computing faster and cheaper. The coincidental invention of efficient computing algorithms and the availability of massive computing power not only removed the inhibitory computation bottleneck, but also allowed for a surge in the development and application of Bayesian methods.

Another significant step forward in Bayesian computing was the development of the BUGS (Bayesian inference using Gibbs sampling) software [24]. The BUGS project was started in 1989 by a group at the Medical Research Council (MRC) Biostatistics Unit and Imperial College School of Medicine in the United Kingdom. BUGS was the first general purpose software available for Bayesian computing. Users could specify the model and the probability distributions of data from a rich set of commonly used distributions. By supplying the prior distribution and the data, the posterior distribution could be computed. Subsequently, an open source version called OpenBUGS was developed, which could run on different operating systems. WinBUGS was then developed by adding useful GUIs (graphics user interfaces) to facilitate its use in the Microsoft Windows environment. In addition, R2WinBUGS and BRugs were developed for users to run WinBUGS within R such that WinBUGS code could be integrated within the R environment to ease the generation, analysis, and reporting.

Another similar development was JAGS (just another Gibbs sampler). This program analyzed Bayesian hierarchical models using MCMC simulation. The unique features of JAGS include (1) a cross-platform engine for the BUGS language; (2) the ability for users to write their own functions, distributions and samplers; and (3) a platform for experimenting with Bayesian modeling.

To address the increasing use of Bayesian methods, SAS (SAS Institute, Cary, NC) added the BAYES statement in the GENMOD, LIFEREG, and PHREG procedures. In addition, starting from version 9.2, SAS introduced a new MCMC procedure. PROC MCMC is a flexible simulation-based procedure suitable for fitting a wide range of Bayesian models. Upon specifying a likelihood function for the data and a prior distribution for the parameters, PROC MCMC obtains samples from the corresponding posterior distributions. It also produces summary and diagnostic statistics.

Although general computation tools such as BUGS or WinBUGS are available, specialized computer programs are often needed to run a Bayesian study design, simulation, and analysis. Web-based applications are particularly useful for timely data entry and analysis. Web services are useful tools for exchanging information between the database module and the computing module. The success of a Bayesian clinical trial requires not only the development of proper tools, but also timely and accurate execution of data entry, quality control, adaptive randomization, outcome assessment, and Bayesian computation. With advancements in both computational algorithms and computer hardware, Bayesian computation in the 21st century is no longer formidable.

3.2. Using subjective information

Following theoretical developments in Bayesian methods and supportive advancements in computing, more and more clinical trialists began to incorporate Bayesian thinking into the study design, conduct, and analysis of clinical trials. Despite its growing popularity, one major impediment to the widespread use of Bayesian methods still exists: a debate on whether or how to incorporate subjective information in inference and decision making. The use of subjective information in clinical trials is a double-edged sword. When used properly, adding subjective information can greatly improve the trial efficiency and facilitate reaching a decision earlier. On the other hand, the improper use of prior information can bias the inference and lead to incorrect conclusions. Furthermore, what is most bothersome to clinical trialists and regulatory agencies, such as the Food and Drug Administration (FDA), is that, given the same data, different conclusions may be drawn if different priors are used. Hence, priors must be pre-specified in the study design and sensitivity analysis is warranted. Bayesian communities have taken different approaches to this problem over the years. Some argue that the Bayesian approach is inherently subjective; hence it should be used

accordingly [25]. Others stress the importance of being objective and propose an objective Bayesian approach by specifying objective priors [26]. Although the debate continues; the goal is one shared by both communities — to efficiently and accurately infer conclusions based on the data [27].

4. Schools of Bayesian approaches

Several schools of Bayesian approaches with different modeling frameworks have been proposed in theory and practice. According to Spiegelhalter, et al. [28], Bayesian approaches can be largely classified into four major types: empirical, reference, proper, and decision-theoretic Bayes.

1. The empirical Bayes approach derives the prior distribution from the data; whereas the standard Bayesian approach sets the prior before any data are observed. The empirical approach can be viewed as a hierarchical Bayes model in which the parameters at the highest level of the hierarchy are set to their most likely values, instead of being integrated out.
2. The reference Bayes approach uses an "objective" or a "reference" prior such that the inference is more objective. Some criticize this approach as "an attempt to make the Bayesian omelets without breaking the Bayesian eggs."
3. The proper Bayes approach uses informative prior distributions based on the available evidence, but summarizes conclusions by posterior distributions without explicit incorporation of the utility function. Some have called this a "stylist Bayes" approach.
4. The decision-theoretic or "full" Bayes approach uses explicit utility (or loss) functions and makes decisions based on maximizing the expected utility (or minimizing loss). One can argue that the decision-theoretic approach provides the ultimate answer to the research question. For example, in drug development, not only does the toxicity and efficacy of the drug need to be assessed, but the relative risk and benefit of the drug also need to be specified explicitly in the utility function. Furthermore, the cost of making a false positive decision (accepting a bad drug) and the cost of making a false negative decision (rejecting a good drug) need to be specified, as well. In complex settings with conflicting goals, the decision-theoretic approach can provide the best (optimal) answer after considering all the loss and gain of each decision. However, it is not easy to come up with a generally acceptable utility function. Additional requirements include the use of dynamic programming and backward induction to obtain the solution in a sequential decision-making process. The computations can be very complex and demanding when applied to real clinical trial situations. As a result, the decision-theoretic approach is rarely used in clinical trials. Currently, the reference Bayes and the proper Bayes approaches are most commonly used in clinical trials.

5. Literature review of Bayesian clinical trials

To survey the use and impact of Bayesian methods in clinical trials, we performed a limited literature review. Our main interest is to ascertain how Bayesian methods have been applied in the design and analysis of real trials. The methodology and results are reported below.

5.1. Literature search procedures

A computerized literature search was performed using two major medical indices (Ovid-Medline and Ovid Embase) for all articles published until September 2011. Using MEDLINE (Table I.A), we searched for the terms, "Bayes or Bayesian," then limited the

search to “clinical trial, all or clinical trial, phase I or clinical trial, phase II or clinical trial, phase III or clinical trial, phase IV or clinical trial or controlled clinical trial or multicenter study randomized controlled trial.” The search was further limited to “review articles and meta analysis or review,” producing results that were subsequently removed from the search. We also performed a similar search in Ovid Embase (Table I.B) using the terms “Bayes or Bayesian.” We then searched “clinical trial*” under subject heading and combined it with the previous “Bayes or Bayesian” search line. The search was limited to “Cochrane library and meta analysis or systematic review” and the results were removed from the final search. The publications obtained in the two literature searches were combined (a total of 2012 articles) and the references were imported into Endnote X4.

5.2. Study Selection

We further processed the articles and placed them into exclusion and inclusion categories (see Figure 1). Articles for exclusion were placed into four main categories: duplicates (electronically and manually obtained), journals (statistical, epidemiological, computer/engineering, conference papers), subjects (meta-analysis, review/opinion, observational/database, statistical methods, pharmacokinetics/pharmacodynamics), and additional exclusions (non-medical, non-human, non-English, non-Bayesian). Note that 256 articles were excluded because they were published in statistical journals, and 478 additional articles were excluded because they had a methodology focus, which included statistical strategies, algorithms, trial designs, method comparisons or demonstrations, tutorials, model development or validation, and simulations. In the reviews/opinion category, 224 articles were excluded. We also excluded 141 meta-analysis/systematic review articles because our focus was on individual trials.

We excluded 185 articles in the last major exclusion category, pharmacokinetics/pharmacodynamics (PK/PD). One of the early applications of Bayesian methods in clinical trials was the use of a nonlinear mixed-effects model in a PK study. The NONMEM program was developed in the late 1970s and quickly became the gold standard for the population-based PK studies [29–31]. Several subsequent PK/PD models and programs were developed [32]. PK/PD examples illustrate the importance of software development. Without appropriate computer software, even the most elegant methods could not be used. Accompanied by user-friendly software, new and even complicated statistical methods can be applied to clinical trials. We decided to exclude the Bayesian PK/PD studies because the goals for these trials were narrow and essentially constituted a distinct subgroup. Following all of the exclusions, 119 articles remained. In addition, we added three more papers that had not been identified by the search algorithm [33–35]. The final number of articles reviewed was 122.

Articles for inclusion were placed into three categories: The first category included studies that used a Bayesian design and analysis (BDA; n = 21), where the clinical trial was prospectively based upon a Bayesian design and the data were analyzed using Bayesian methods. The second included prospective studies that used a frequentist design with a Bayesian analysis (FDBA; n = 83). Bayesian reanalysis (BR) studies (n = 18) comprised the final category, which involved the use of Bayesian methods to retrospectively analyze data from a previously run clinical trial.

5.3. Data extraction and results

Information extracted from each article included title, name of journal, year of publication, whether the trial was BDA, FDBA or BR, type of clinical trial, type of re-analysis (if applicable), medical area of study, method of randomization, number of treatment arms/groups, actual sample size, control group, months of accrual, type of primary endpoint,

endpoint category, Bayesian method category, type of prior distribution, number of interim analyses and whether the trial was terminated early. The results are summarized in Tables II through IV.

As seen in Table II, publications prior to 1990 included only three clinical trials that used Bayesian methods. That number quickly jumped to 20 in the 1990s, and to 99 in the period since 2000. Most trials (67%) applied Bayesian methods for testing treatment efficacy; 8% of the trials applied them for testing treatment safety; 16% of the trials applied them in the areas of diagnostics and decision making. In terms of the medical fields, oncology led the pack (25%), followed by cardiovascular research (16%), and CNS research (11%). These 122 papers were published dispersedly in 92 journals, with nine in the *Journal of Clinical Oncology*, four in *PLoS One*, and three each in *The New England Journal of Medicine*, *JAMA*, *Cancer*, and *Complementary Therapies in Medicine* (data not shown).

From Table III we see that the vast majority of the trials are 2-arm (59%) or 1-arm (30%) studies. About 50% of the trials had a control group. Almost 60% of the studies were randomized trials, 48% of which applied equal randomization, and 6% of which applied fixed, but unequal randomization. Only 5% of the trials applied adaptive randomization. In terms of sample size, 23% were very small ($n \leq 30$) and 26% had sample sizes between 31 and 99. Sample sizes were between 100 to 499 and 500 to 999, respectively, for 37% and 9% of the trials. Only 6% of the trials had sample sizes of 1,000 or more patients. About 51% of the trials enrolled patients during 2 or fewer years; whereas 35% of trials spent 2 to 5 years enrolling patients. The remaining 14% of the trials had an accrual period of 6 to 8 years.

Table IV shows that continuous, binary, ordinal, and time-to-event variables, respectively, were used as the primary endpoints in 43%, 33%, 12%, and 12% of the trials. A number of Bayesian methods were applied in clinical trials, including for hypothesis testing (23%), estimation (19%), model selection (12%), prediction (12%), regression (10%), hierarchical modeling (10%), a decision-theoretic approach (5%), sensitivity analysis (5%), classification (3%), and a Bayesian network (2%). Informative priors were used in 43% of the trials; noninformative priors were used in 24% of the trials. The remaining 34% of the trials did not provide sufficient information regarding the priors that were used. A vast majority of the trials (87%) did not specify an interim analysis. Only 7%, 3%, and 4% of the trials had 1, 2, or 3–7 interim analyses, respectively. Twelve percent of the trials were stopped early; 6 due to futility, 4 due to efficacy, 1 for equivalence, and 1 for toxicity.

5.4. Limitations

Although efforts were made to best identify the use of Bayesian methods in clinical trials, our search has several limitations. First, we included the words “Bayes or Bayesian” and “clinical trials” in our search criteria. Therefore, articles without these words were excluded. The most prominent inadvertent exclusions were trials that used the continual reassessment method (CRM) or the escalation with over-dose control (EWOC) method that did not mention the words “Bayes or Bayesian.” We performed a separate literature search and identified 81 CRM trials and 9 EWOC trials (see Appendix 1 for the search algorithm and results). Only three of such trials were included in the 122 articles we reported. Most of the studies were dose-finding cancer studies, with a goal of determining the maximum tolerated dose. These were typically single-arm, open-label studies without a control group and with a total sample size of less than 60. Most studies applied Bayesian design and analysis. However, we decided not to include these studies not identified in the main search in the tabulation because, similar to the Bayesian PK/PD studies, they form a distinct subgroup.

6. MD Anderson experience with the design and conduct of Bayesian clinical trials

A recent review of Bayesian adaptive clinical trials that was published in 2011 indicated that a large portion of the papers reporting the use of Bayesian methods were published from the University of Texas MD Anderson Cancer Center [36]. Many papers included in that review were methodology papers, which were not included in this study, as explained in the previous section. Among the 331 papers identified in the review, MD Anderson Cancer Center contributed to 17.2%; whereas the next two single highest sources were the NCI, contributing 4.3%, and Harvard University, contributing 3.9% of the total publications. Four of the nine researchers who published the highest number of articles describing a Bayesian clinical trial were affiliated with MD Anderson. A review published in 2009 described 964 clinical protocols registered at MD Anderson between 2000 and early 2005 [37]. Bayesian designs and/or analyses had been used in about 20% of the total protocols reviewed, and in about 30% of the MD Anderson trials, but in only 7% of the multicenter protocols. Bayesian methods had been applied in 34% of the phase I or phase II trials. The majority of the Bayesian design and analysis features were found in non-mutually exclusive categories, which included efficacy monitoring (62%), toxicity monitoring (27%), adaptive randomization (10%), dose finding (9%), hierarchical modeling (7%), and determinations of predictive probability (6%).

To facilitate the conduct of Bayesian clinical trials, a proper infrastructure must be set up for registering patients, assigning patients to treatments, recording the outcomes, and providing interim and final analyses. At MD Anderson, we have developed a Clinical Trial Conduct (CTC) platform. This secure, web-based application allows users to register a new trial and select the type of design so that proper treatment assignment or randomization and monitoring can be implemented. In this “role-based” system, each user has different privileges, depending on his/her role in the trial. For example, the research nurse can verify patients’ eligibility criteria, register patients, and enter patients’ toxicity and efficacy outcomes. The statistician can read the treatment assignment and access the details of the statistical computations, such as the randomization probability, but cannot alter the data. As of August 2011, there were 133 trials and over 4,300 patients enrolled in the system. The most commonly used designs were outcome adaptive randomization (n = 44), Pocock-Simon baseline adaptive randomization (n = 42), the continual reassessment method (n = 29), and trials with time-to-event interim monitoring (n = 11).

In addition to the CTC platform, we have also built custom-made applications for certain specialized trials. For example, Figure 2 shows the schematic diagram of the web-based application for running the BATTLE trial [34, 38]. The top panel shows the study flow chart. Eligible patients are registered to the trial and have a biopsy taken for molecular marker analysis. The research molecular pathology laboratory analyzes the sample for mutations and uses the results of fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) expression analyses to determine the biomarker group for the patient. This process, from registration to reporting the biomarker results, is completed within two weeks. Based on the patient’s biomarker group and the cumulative outcome results, the patient is then adaptively randomized into one of the four treatment groups. Additional clinical visits are scheduled for the patient and the disease control status (primary endpoint of the study) is evaluated eight weeks after randomization. Patients are continually followed for secondary endpoints, such as progression-free survival and overall survival, until they are off the study. The middle panel of Figure 2 shows different modules in the application. Through a web-interface, data are entered into different modules. For example, research nurses enter the medical history, physical examination, adverse events, efficacy assessment, etc. The laboratory technician enters the results of the marker analysis. To

perform adaptive randomization, the patient's marker information, eligibility status, and up-to-date outcome information are passed to an R code through web services. The R code performs Bayesian computation to determine the randomization probability and randomize patients to eligible treatments accordingly. All the data are stored in the institutional database CORe and/or the study-specific SQL Server 2005 database. The application also has a report generation module that can automatically generate several reports for monitoring and quality assurance purposes. For example, an accrual report is generated to check the accrual rate. An outcome timeliness report can check whether the 8-week disease control status has been timely entered. If the assessment of the primary endpoint of a patient remains past due for two weeks, an automatically generated email will be sent to the study coordinator/research nurse. The toxicity and drug compliance reports can also be generated to ensure the safety and compliance of patients in the study.

7. Discussion and perspectives

Despite its early conception, Bayesian methods have lagged behind frequentist methods in both statistical theoretical development and application in clinical trials. Thanks to the relentless efforts of many diehard enthusiasts, the Bayesian approach has staged a strong comeback in the past 20 years. As shown in our review, the first major application of Bayesian methods in clinical trials was in the area of PK/PD studies as the result of the development of the popular NONMEM software in the late 1970s. The second major application was spurred by the development of the CRM and EWOC methods in dose-finding studies in the 1990s, which involved the development of software tools to facilitate the implementation of these methods [39]. Despite a slow adaption, it is apparent that Bayesian methods are increasingly being used in clinical trials. This trend will likely continue [36, 40]. We have also begun to see the impact of Bayesian applications among policymakers in regulatory agencies regarding the approval of new medical devices or drugs evaluated in studies that use Bayesian designs or methods. The Center for Devices and Radiological Health (CDRH) at the FDA issued a "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials" in 2010 [41]. The CDRH has approved more than 20 original Pre-Market Approvals (PMAs) and PMA supplements with a Bayesian analysis as the primary method. Many Investigational Device Exemptions (IDEs) and applications for "substantial equivalence" (510(k)s) that used Bayesian methods have also been approved [42]. On the drug side, the Center for Drugs and Experimental Research of the FDA approved Pravigard Pac (Bristol-Myers Squibb) based on Bayesian analyses of efficacy in 2003 [17]. As many clinical trials using Bayesian methods are underway, it is expected that more drugs and devices will be approved by the FDA based on Bayesian methods.

The development of newer and better clinical trial designs under the Bayesian paradigm continues to be an active area of statistical methodology research. The availability of accompanying software for the implementation of Bayesian methods is crucial for the use of these methods in clinical trials. Altman indicated that there is a delay of 4 to 6 years between the date when a statistical method is published and when that publication is cited 25 times in medical journals [43]. The time gap between the publication of a new trial design and its adoption still exists, but is closing rapidly. This is evident in the Bayesian dose-finding studies and adaptive designs [36, 40]. (We list some useful information/tools for learning Bayesian clinical trial methods and designs in Appendix 2.)

Bayesian methods hold great promise for improving the efficiency and flexibility of conducting clinical trials and are ideal for learning and adaptation. Bayesian methods provide excellent tools when searching for effective treatments and predictive markers in the quest for biomarker-based personalized medicine — with a goal of treating more patients

with more effective therapies. Good examples for such trials include the BATTLE trial [34], the currently ongoing BATTLE-2 trial, and the I-SPY 2 trial [44]. Successful implementations of Bayesian methods have been demonstrated in a wide range of clinical trial applications.

Bradley Efron, in his 2004 address as the president of the American Statistical Association, stated that the field of statistics was dominated by the Bayesian view in the 19th century and by the frequentist view in the 20th century. He suggested that statistics in the 21st century, challenged by greater magnitudes of data and complexity, will require a combination of both Bayesian and frequentist methods [45]. The following year, Roderick J. Little, in his presidential address, proposed the "calibrated Bayes" approach [46]. The calibrated Bayes approach uses frequentist methods for model development and assessment, and Bayesian methods for inference under a model. This capitalizes on the strengths of both paradigms and provides a useful roadmap for many problems of statistical modeling and inference.

The relative merit of the Bayesian and frequentist approaches continues to be the subject of debate in statistics and other scientific fields. Regarding the two paradigms, the past was combative, the present is competitive, and the future will be cooperative. After all, Bayesian and frequentist approaches offer complementary views and can learn from each other. Recently, more evidence can be found showing the convergence of the two camps, at least on a practical level [47]. Ultimately, better clinical trial methods lead to more efficient designs, lower sample sizes, more accurate conclusions, and better outcomes for patients enrolled in the trials. Bayesian methods offer an attractive alternative for better trials. More such trials should be conducted to refine the approach and demonstrate the real benefit of the Bayesian approach in action.

Acknowledgments

The authors thank Donald A. Berry for helpful discussions, Richard Herrick for providing information on the MD Anderson Clinical Trial Conduct web-based software application, and Ms. LeeAnn Chastain for editorial assistance.

References

1. Streptomycin-in-Tuberculosis-Trials-Committee. Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. British Medical Journal. 1948; 2:769–782.10.1136/bmj.2.4582.769 [PubMed: 18890300]
2. Halperin M, DeMets DL, Ware JH. Early methodological developments for clinical trials at the National Heart, Lung and Blood Institute. Statistics in Medicine. 1990; 9(8):881–892.10.1002/sim.4780090804 [PubMed: 2218191]
3. Gehan EA, Schneiderman MA. Historical and methodological developments in clinical trials at the National Cancer Institute. Statistics in Medicine. 1990; 9:871–880.10.1002/sim.4780090803 [PubMed: 2218190]
4. Stigler SM. The history of statistics in 1933. Statistical Science. 1996; 11:244–252.10.1214/ss/1032280216
5. Gehan EA. Methodological issues in cancer clinical trials: the comparison of therapies. Biomedicine and Pharmacotherapy. 1988; 42 (3):161–165.
6. Bayarri MJ, Berger JO. The interplay of Bayesian and frequentist analysis. Statistical Science. 2004; 19:58–80.10.1214/088342304000000116
7. Bayes T. An essay towards solving a problem in the doctrine of chances. 1763. M D Computing: Computers in Medical Practice. 1991; 8:157–171. [PubMed: 1857193]
8. Stigler SM. Thomas Bayes' Bayesian inference. Journal of the Royal Statistical Society, Series A. 1982; 145:250–258.10.2307/2981538
9. Bernardo, JM.; Smith, AFM. Bayesian Theory. Wiley; West Sussex: 1994.

10. Meier P. Jerome Cornfield and the methodology of clinical trials. Controlled Clinical Trials. 1981; 1(4):339–345. [PubMed: 7021066]
11. Ederer F. Jerome Cornfield's contributions to the conduct of clinical trials. Biometrics. 1982; 38 (Suppl):25–32. [PubMed: 7046822]
12. Ashby D. Bayesian statistics in medicine: a 25 year review. Statistics in Medicine. 2006; 25(21): 3589–3631.10.1002/sim.2672 [PubMed: 16947924]
13. Grieve AP. 25 years of Bayesian methods in the pharmaceutical industry: a personal, statistical bummel. Pharmaceutical Statistics. 2007; 6:261–281.10.1002/pst.315 [PubMed: 17955514]
14. Berger, JO.; Wolpert, RL. The Likelihood Principle. Gupta, SS., editor. Institute of Mathematical Statistics; Hayward, CA: 1984. http://projecteuclid.org/euclid.lnms/1215466210
15. Bernardo JM. The concept of exchangeability and its applications. Far East Journal of Mathematical Sciences. 1996; (Part II):111–121. Special Volume.
16. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Statistical Methods in Medical Research. 2001; 10(4):277–303.10.1177/096228020101000404 [PubMed: 11491414]
17. Berry DA. Bayesian clinical trials. Nature Reviews Drug Discovery. 2006; 5(1):27–36.10.1038/nrd1927
18. Jennison, C.; Turnbull, BW. Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC; Boca Raton, FL: 2000.
19. Dmitrienko A, Wang MD. Bayesian predictive approach to interim monitoring in clinical trials. Statistics in Medicine. 2006; 25(13):2178–2195.10.1002/sim.2204 [PubMed: 16007570]
20. Berry DA. Adaptive trial design. Clinical Advances in Hematology and Oncology. 2007; 5(7):522–524. http://www.clinicaladvances.com/article_pdfs/ho-article-200707-drugdev.pdf. [PubMed: 17679925]
21. Center for Drug Evaluation and Research, Food and Drug Administration. Guidance for industry: adaptive design clinical trials for drugs and biologics. U.S. Department of Health and Human Services; Rockville, MD: 2010. available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf
22. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984; PAMI-6:721–741.10.1109/TPAMI.1984.4767596 [PubMed: 22499653]
23. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association. 1990; 85:398–409.10.2307/2289776
24. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: evolution, critique and future directions. Statistics in Medicine. 2009; 28:3049–3067.10.1002/sim.3680 [PubMed: 19630097]
25. Goldstein M. Subjective Bayesian analysis: principles and practice. Bayesian Analysis. 2006; 1:403–420.10.1214/06-BA116
26. Berger J. The case for objective Bayesian analysis. Bayesian Analysis. 2006; 1:385–402.10.1214/06-BA115
27. Fienberg SE. Does it make sense to be an "objective Bayesian"? (Comment on articles by Berger and by Goldstein). Bayesian Analysis. 2006; 1:429–432.10.1214/06-BA116C
28. Spiegelhalter, DJ.; Abrams, KR.; Myles, JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Wiley; West Sussex: 2004.
29. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics. 1980; 8:553–571.10.1007/BF01060053 [PubMed: 7229908]
30. Sheiner LB, Ludden TM. Population pharmacokinetics/dynamics. Annual Review of Pharmacology and Toxicology. 1992; 32:185–209.10.1146/annurev.pa.32.040192.001153
31. Whiting B, Kelman AW, Grevel J. Population pharmacokinetics. Theory and clinical application. Clinical Pharmacokinetics. 1986; 11:387–401.10.2165/00003088-198611050-00004 [PubMed: 3536257]

32. Bauer RJ, Guzy S, Ng C. A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. The AAPS Journal. 2007; 910.1208/aapsj0901007

33. Buzdar AU, Ibrahim NK, Francis D, Booser DJ, Thomas ES, Theriault RL, Pusztai L, Green MC, Arun BK, Giordano SH, Cristofanilli M, Frye DK, Smith TL, Hunt KK, Singletary SE, Sahin AA, Ewer MS, Buchholz TA, Berry D, Hortobagyi GN. Significantly higher pathologic complete remission rate after neoadjuvant therapy with trastuzumab, paclitaxel, and epirubicin chemotherapy: results of a randomized trial in human epidermal growth factor receptor 2-positive operable breast cancer. Journal of Clinical Oncology. 2005; 23:3676–3685.10.1200/JCO.2005.07.032 [PubMed: 15738535]

34. Kim ES, Herbst1 RS, Wistuba II, Lee JJ Jr, GRB, Tsao A, Stewart DJ, Hicks ME Jr, JE, Gupta S, Alden CM, Liu S, Tang X, Khuri FR, Tran HT, Johnson BE, Heymach JV, Mao L, Fossella F, Kies MS, Papadimitrakopoulou V, Davis SE, Lippman SM, Hong WK. The BATTLE trial: personalizing therapy for lung cancer. Cancer Discovery. 2011; 1:44–53.10.1158/2159-8274.CD-10-0010 [PubMed: 22586319]

35. Julian TB, Blumencranz P, Deck K, Whitworth P, Berry DA, Berry SM, Rosenberg A, Chagpar AB, Reintgen D, Beitsch P, Simmons R, Saha S, Mamounas EP, Giuliano A. Novel intraoperative molecular test for sentinel lymph node metastases in patients with early-stage breast cancer. Journal of Clinical Oncology. 2008; 26:3338–3345.10.1200/JCO.2007.14.0665 [PubMed: 18612150]

36. Chevret S. Bayesian adaptive clinical trials: a dream for statisticians only? Statistics in Medicine. 2011 Sep 9. Epub ahead of print. 10.1002/sim.4363

37. Biswas S, Liu DD, Lee JJ, Berry DA. Bayesian clinical trials at the University of Texas M. D. Anderson Cancer Center. Clinical Trials. 2009; 6:205–216.10.1177/1740774509104992 [PubMed: 19528130]

38. Zhou X, Liu S, Kim ES, Herbst RS, Lee JJ. Bayesian adaptive design for targeted therapy development in lung cancer — a step toward personalized medicine. Clinical Trials. 2008; 5:181–193.10.1177/1740774508091815 [PubMed: 18559407]

39. Tighiouart M, Rogatko A. Dose finding with escalation with overdose control (EWOC) in cancer clinical trials. Statistical Science. 2010; 25:217–226.10.1214/10-STS333

40. Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. Journal of Clinical Oncology. 2007; 25:4982–4986.10.1200/JCO.2007.12.1012 [PubMed: 17971597]

41. Center for Devices and Radiological Health, Food and Drug Administration. Guidance for the use of Bayesian statistics in medical device clinical trials. U.S. Department of Health and Human Services; Rockville, MD: 2010. available at http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf

42. Campbell, G. Bayesian statistics at the FDA: the trailblazing experience with medical devices. Presentation at Rutgers, the State University of New Jersey; April 3, 2009; available at http://www.stat.rutgers.edu/iob/bioconf09/slides/Campbell.pdf

43. Altman DG, Goodman SN. Transfer of technology from statistical journals to the biomedical literature—past trends and future predictions. JAMA. 1994; 272:129–132.10.1001/jama.272.2.129 [PubMed: 8015123]

44. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: An adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clinical Pharmacology and Therapeutics. 2009; 86:97–100.10.1038/clpt.2009.68 [PubMed: 19440188]

45. Efron B. Bayesians, frequentists, and scientists. Journal of the American Statistical Association. 2005; 100:1–5.10.1198/016214505000000033

46. Little RJ. Calibrated Bayes: A Bayes/frequentist roadmap. American Statistician. 2006; 60:213–223.10.1198/000313006X117837

47. Arjas E. On future directions in statistical methodologies — some speculations. Scandinavian Journal of Statistics. 2011; 38:185–194.10.1111/j.1467-9469.2011.00737.x

48. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. Biometrics. 1990; 46:33–48.10.2307/2531628 [PubMed: 2350571]

49. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. Statistics in Medicine. 1998; 17:1103–1120.10.1002/(SICI)1097-0258(19980530)17:10<1103::AID-SIM793>3.0.CO;2-9 [PubMed: 9618772]

Appendix 1. Search algorithm and results for identifying clinical trials using continual reassessment method (CRM) and escalation with overdose control (EWOC)

Appendix Table I

Schema for literature search for continual reassessment method

A. Ovid Medline ® In-process and other non-indexed citations and Ovid MEDLINE ® 1948 to present	Search result
continual reassessment method.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	148
limit 1 to (clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or controlled clinical trial or multicenter study randomized controlled trial)	46
B. Embase Classic + Embase 1947 to 2011 September 22	Search result
Continual reassessment method.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword	169
clinical trial*.sh.	834574
1 and 2	77

Appendix Table II

Schema for literature search for escalation with overdose control method

A. Ovid Medline ® In-process & Other Non-Indexed Citations and Ovid MEDLINE ® 1948 to present	Search result
escalation with overdose control.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	9
limit 1 to (clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or controlled clinical trial or multicenter study randomized controlled trial)	2
B. Embase Classic + Embase 1947 to 2011 September 22	Search result
escalation with overdose control.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword	16
clinical trial*.sh.	834574
1 and 2	7

In addition, we performed reverse citation searches on two original articles for the CRM and EWOC methods, respectively [48, 49]. The results from all our searches were placed into EndNote X4 for duplicate removal and categorization.

Results for Continual Reassessment Method and Escalation with Overdose Control Search

The Ovid Medline and Embase search netted a total of 123 CRM and 9 EWOC results. Following the removal of 43 duplicates, 17 statistical journals, 1 computer/engineering journal, 7 review/opinion papers, 16 method papers, and 5 conference papers, we found 41 CRM and 2 EWOC papers.

Our reverse citation search netted 356 articles from J. O'Quigley and 109 from J. Babb. There were 95 duplicates, 119 statistical journals, 4 computer/engineering journals, 75 review/opinion papers, 100 method papers, 3 meta-analysis papers, 3 observational/database papers, and 5 non-CRM papers. From the resulting journals, we found 59 CRM and 9 EWOC papers.

The Ovid Medline/Embase search was combined with the reverse citation search and 21 duplicates were removed. Our final cohort consisted of 81 CRM and 9 EWOC papers.

Appendix 2. Useful information/tools for learning Bayesian clinical trials

A. Articles

1. Berry DA. Bayesian clinical trials. *Nature Reviews Drug Discovery* 2006; **5(1)**:27–36.
2. Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. *Clinical Trials* 2005; **2**:282–290.
3. Louis TA. Introduction to Bayesian methods II: fundamental concepts. *Clinical Trials* 2005; **2**:291–294.
4. Berry DA. Introduction to Bayesian methods III: use and interpretation of Bayesian tools in design and analysis. *Clinical Trials* 2005; **2**:295–300.
5. Berry DA. Bayesian statistics and the efficiency and ethics of clinical trials. Statistical Science 2004; 19:175–187.
6. Ashby D, Tan SB. Where's the utility in Bayesian data-monitoring of clinical trials? *Clinical Trials* 2005; **2**:197–205.
7. Casella G, George EI. Explaining the Gibbs sampler. *American Statistician* 1992; **46**:167–174.

B. Books

1. Berry DA, Stangl D. *Bayesian Biostatistics*. CRC Press: Boca Raton, FL, 1996.
2. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley: West Sussex, 2004.
3. Carlin BP, Louis TA. *Bayesian Methods for Data Analysis.* 3rd edition. Chapman & Hall/CRC: Boca Raton, FL, 2008.
4. Hoff PD. A First Course in Bayesian Statistical Methods. Springer: New York, 2009.
5. Albert J. *Bayesian Computation with R (Use R)*. 2nd edition. Springer: New York, 2009.
6. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. 2nd edition. Chapman & Hall/CRC: Boca Raton, FL, 2009.
7. Berry SM, Carlin BP, Lee JJ, Mueller P. *Bayesian Adaptive Methods for Clinical Trials*. Chapman & Hall/CRC: Boca Raton, FL, 2010.

C. Video tutorials

1. FDA and the Johns Hopkins University Workshop: Can Bayesian approaches to studying new treatments improve regulatory decision-making? http://webcasts.prous.com/bayesian2004/

D. Computer programs

1. General Bayesian computation tools
    a. BUGS, OpenBUGS, and WinBUGS: http://www.mrc-bsu.cam.ac.uk/bugs/
    b. JAGS: http://mcmc-jags.sourceforge.net/
2. Running WinBUGS from R
    a. BRugs: http://www.biostat.umn.edu/~brad/software/BRugs/
    b. R2WinBUGS: http://cran.r-project.org/web/packages/R2WinBUGS/index.html
3. Running WinBUGS from Stata - The winbugsfromstata package: http://www2.le.ac.uk/departments/health-sciences/research/ships/gen-epi/Progs/winbugs-from-stata
4. A collections of useful tools for Bayesian clinical trials, including CRM, BMA-CRM, EFF-TOX, Multc99, adaptive randomization, predictive probability, etc., can be downloaded from https://biostatistics.mdanderson.org/softwaredownload
5. Other CRM and EWOC design programs
    a. TITE-CRM http://roadrunner.cancer.med.umich.edu/wiki/index.php/TITE-CRM
    b. Modified CRM v2.0 http://www.cancerbiostats.onc.jhmi.edu/software.cfm
6. EWOC https://apps.winship.emory.edu/biostatistics/software_ewoc.php SAS Proc MCMC: http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mcmc_sect019.htm
7. Tessella and Berry Consultants' Fixed and Adaptive Clinical Trials Simulator v2 (FACTS 2) http://www.smarterclinicaltrials.com/wp-content/uploads/FACTS_introduction.pdf
8. Cytel's Compass: software for adaptive dose-finding trials http://www.cytel.com/Software/Compass.aspx

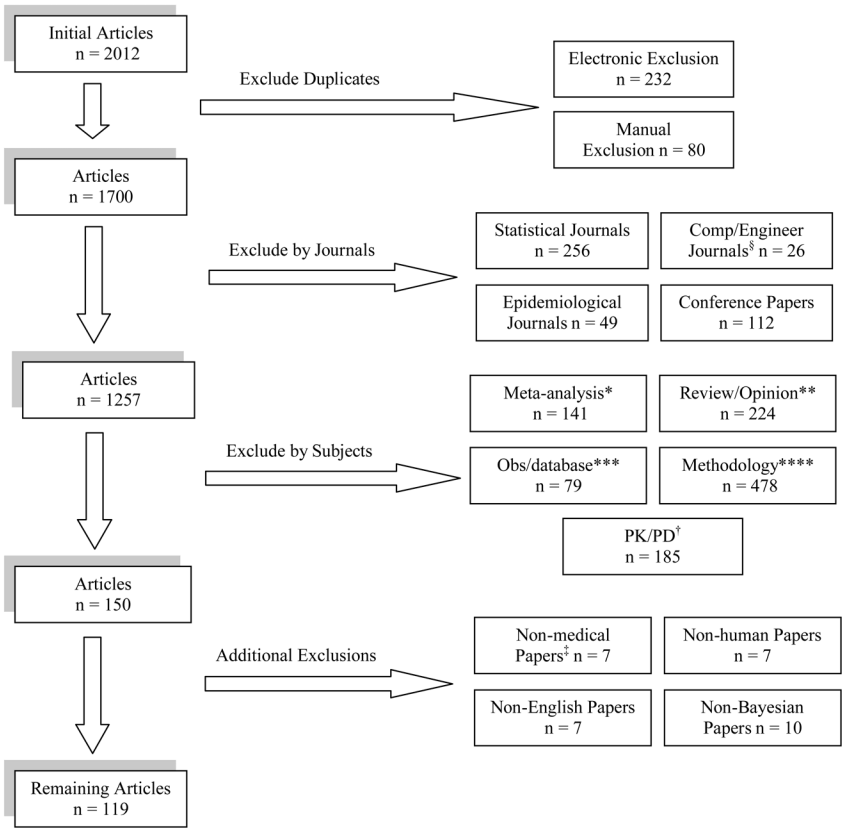

Figure 1.
Study selection algorithm
§ comp = computers
* also includes Cochrane journals and systematic reviews
** also includes commentaries, letters, replies, surveys, notes, guidelines and short articles
*** obs = observational; also includes registries and epidemiologic studies
**** methodology focused includes statistical strategies, algorithms, trial designs, method comparisons/demonstrations, tutorials, model development/validation, and simulations.
† pharmacokinetics/pharmacodynamics
‡ includes engineering, social science and policy making studies

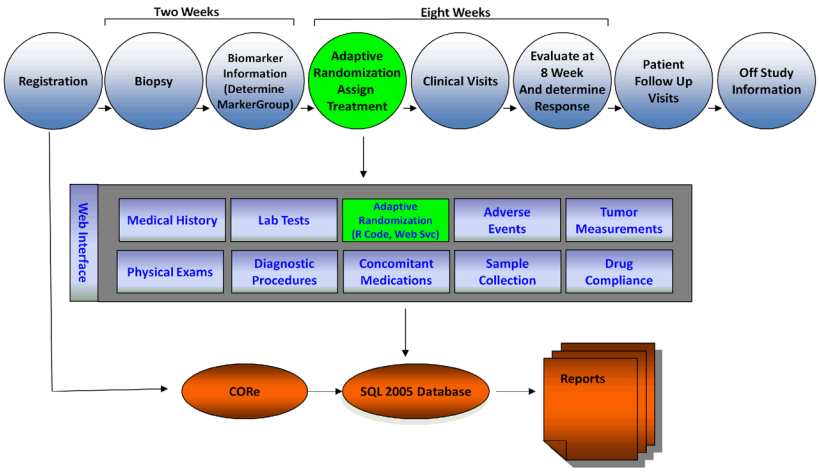

Figure 2.
Schematic diagram of the web-based database application for the conduct of the BATTLE trial

Table I

Schema for literature search

	A. Ovid Medline ® In-process and other non-indexed citations and Ovid MEDLINE ® 1948 to September 2011	Search result
1	(bayes or bayesian).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	21,310
2	limit 1 to (clinical trial, all or clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv or clinical trial or controlled clinical trial or multicenter study randomized controlled trial)	682
3	limit 2 to (“review articles” and (meta analysis or “review”))	3
4	2 not 3	679
	B. Embase Classic + Embase 1947 to 2011, September 22	**Search result**
1	(bayes or bayesian).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword	21,997
2	clinical trial*.sh.	829,875
3	1 and 2	1346
4	Limit 3 to (cochrane library and (meta analysis or “systematic review”))	13
5	3 not 4	1333

Table II

Characteristics of publications reviewed (n = 122): year of publication, type of clinical trial and medical area of study.

Variable	Frequency	Percentage (%)
Years		
1975–1989	3	2.5
1990–1994	6	4.9
1995–1999	14	11.5
2000–2004	27	22.1
2005–2011	72	59.0
Type of clinical trial		
Efficacy	82	67.2
Diagnostic/Decision making	19	15.6
Safety	10	8.2
Association studies	10	8.2
Cost-benefit	1	0.8
Medical areas of study		
Addiction	2	1.6
Auditory system	1	0.8
CNS	13	10.7
Cardiovascular system	20	16.4
Dentistry	1	0.8
Gastrointestinal system	3	2.5
Genetics	2	1.6
Genitourinary system	2	1.6
Geriatrics	2	1.6
Hematology	2	1.6
Infectious disease	10	8.2
Metabolic disorder	3	2.5
Obstetrics and gynecology	9	7.4
Oncology	30	24.6
Ophthalmology	1	0.8
Pain	2	1.6
Pediatrics	1	0.8
Pulmonary system	8	6.6
Radiology	2	1.6
Renal system	4	3.3
Transplant	4	3.3

Table III

Trial characteristics in publications reviewed (n = 122).

Variable	Frequency	Percentage (%)
Number of arms *		
1	35	30.4
2	68	59.1
3	7	6.1
≥ 4	5	4.3
The Use of Control Group *		
Yes **	57	49.6
None	58	50.4
Method of Randomization *		
Adaptive randomization	6	5.2
Equal randomization ***	55	47.8
Fixed unequal randomization ****	7	6.1
None	47	40.9
Actual Sample Size		
≤ 30	28	23.0
31–59	18	14.8
60–99	13	10.7
100–199	16	13.1
200–499	29	23.8
500–999	11	9.0
1000–9999	7	5.7
Accrual Period (in months) *****		
≤ 12	13	30.2
13–24	9	20.9
25–36	7	16.3
37–48	5	11.6
49–60	3	7.0
6–8 years	6	14.0

* Does not include seven papers with multiple studies

** 23 active; 19 placebo; 1 no treatment; 14 unspecified

*** includes: forty-seven 2-arm trials; five 3-arm trials; two 4-arm trials; one 5-arm trial

**** all 2-arm trials (one 1:1.5, one 2.8:1, four 2:1, one 3:1)

***** 79 studies did not specify

Table IV

Statistical attributes in publications reviewed (n = 122).

Variable	Frequency	Percentage (%)
Primary endpoint category		
Continuous	53	43.4
Binary	40	32.8
Ordinal	15	12.3
Time to event	14	11.5
Bayesian Method		
Hypothesis testing	28	23.0
Estimation	23	18.9
Forecast/prediction	15	12.3
Model selection/comparison	15	12.3
Regression model	12	9.8
Hierarchical model	12	9.8
Decision theory	6	4.9
Sensitivity analysis	6	4.9
Multivariate/classification	3	2.5
Bayesian network	2	1.6
Prior Distribution		
Informative	52	42.6
Non-informative	29	23.8
Unspecified	41	33.6
Number of Interim Analyses		
0	106	86.9
1	8	6.6
2	3	2.5
3	1	0.8
4	1	0.8
5	1	0.8
6	1	0.8
7	1	0.8
Trial Stopped Early		
No	110	90.2
Yes	12	9.8
Reason for Stopping early		
Futility	6	50.0
Efficacy	4	33.3
Equivalence	1	8.3
Toxicity	1	8.3