Special Issue Paper

Statistics in Medicine

Received 14 March 2012, Accepted 15 March 2012 Published online 18 June 2012 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5404

Bayesian clinical trials in action

J. Jack Lee*† and Caleb T. Chu

Although the frequentist paradigm has been the predominant approach to clinical trial design since the 1940s, it has several notable limitations. Advancements in computational algorithms and computer hardware have greatly enhanced the alternative Bayesian paradigm. Compared with 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 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, all of 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 Bayesian trials should be designed and conducted to refine the approach and demonstrate their real benefit in action. Copyright © 2012 John Wiley & Sons, Ltd.

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 pre-specified 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 UK [1]. That trial involved randomizing patients into treatment and control groups and assessing the outcome without knowledge of the treatment assignment. 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 USA, the National Institutes of Health, particularly the National Heart, Lung, and Blood Institute and the National Cancer Institute, have led the effort to develop and conduct clinical trials [2, 3]. *Clinical Trials: Past, Present and Future*, a National Heart, Lung, and Blood Institute-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 parameter of interest,

Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, U.S.A.
**Correspondence to: J. Jack Lee, Department of Biostatistics, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1411, Houston, TX 77030, U.S.A.*
†E-mail: jjlee@mdanderson.org

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 is fixed and the unknown parameter θ is modeled as a random variable of a probability distribution. 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 with 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. We organize the rest of the paper 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, used to illustrate how Bayesian methods have been practically applied 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 roles that Bayesian methods may play in clinical trials.

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 proportionally to the product of the prior probability of θ and the data likelihood, that is, Prob(θ|data) ∝ 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 the arena of clinical trial development [10, 11]. Ashby comprehensively reviewed the development of Bayesian statistical methodology in clinical trials [12], whereas Grieve gave his personal account on 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 is based upon the observed data and thereby 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 to make. On the other hand, Bayesian inference is conditioned on the data and not on the design of the trial, so it can still maintain validity as long as the prior distribution and the probability model are correctly specified. Note that under the Bayesian reference prior analyses, it is possible for two models to yield similar proportional likelihoods but different reference priors, hence result in different analysis outcomes [15]. We give more discussions later on the choice of prior.

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*

By assuming the parameter is fixed, frequentist methods often underestimate the variability of the parameter of interest. On the other hand, 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. Different levels of variability are appropriated 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 [16].

2.3. *Bayesian methods give direct answers to the questions that most people want to ask and provide a uniform way to solve complex problems*

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 aforementioned questions. The frequentist confidence interval is random because the data observed is also 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, that is, 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. Problems of any sort can be approached in a straightforward three-step Bayesian formulation: first, specifying the prior distribution of the parameter of interest; second, observing the data and; third, 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.

2.4. *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. In the reverse manner, 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. The robustness of Bayesian analysis can also be improved by considering a class of plausible priors, allowing for heavy tailed distributions, or by constructing hierarchical mixture models [17–19]. In addition, the Bayesian framework allows for the incorporation of information accumulated in the trial, acquired outside of the trial, and can synthesize information across multiple trials as required in a meta-analysis [20, 21].

2.5. *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 data updates the prior distribution to form the posterior distribution. The formed posterior distribution then becomes the prior distribution for a future evaluation. Frequent monitoring of interim results are desirable in clinical trials because many of the trials are conducted over an extended period. This allows for decisions to be made early when sufficient evidence has accumulated. Although group sequential methods have been well developed under the frequentist paradigm [22], 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 on the basis of 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 assumes that θ is fixed [23].

2.6. 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?' Maximizing the utility function or minimizing the loss function can make the optimal decision of the best treatment for a given patient. 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 or 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.

2.7. 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 that are less adaptive often lead to large trials over an extended period. Adaptive designs have been proposed with the aim of creating more efficient, more flexible, and more ethical designs by making design changes on the basis of the interim data. The Bayesian framework naturally and ideally fits into the development of adaptive designs [24]. Bayesian adaptive approaches are especially useful in the following three areas: (i) outcome adaptive randomization to assign more patients into more effective treatment arms as data accumulates in the trial; (ii) interim monitoring for early stopping as a result of futility or efficacy; and (iii) adaptive sample size estimation by calculating the probability for a successful trial given the current result and the sample size required to reach a definitive conclusion at the end of study. In 2010, the U.S. Food and Drug Administration (FDA) issued a guidance document for adaptive clinical trials [25]. 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 early work of Bayes in the 1760s and Laplace in the following decades, Bayesian approaches have been 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 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-old stagnancy changed during the 1980s and the 1990s with the advent and development of the Markov Chain Monte Carlo (MCMC) [26, 27]. By constructing a Markov chain with the desired distribution as its equilibrium state, MCMC can construct complex posterior probability distributions on the basis of Monte Carlo samples. During the development of MCMC, personal computers and workstations were also becoming more available and powerful, allowing for both cheaper and faster computational processing. 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 [28]. A group at the Medical Research Council Biostatistics Unit and Imperial College School of Medicine in the UK started the BUGS project in 1989. 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, we could compute the posterior distribution. Subsequently, an open source version called OpenBUGS was developed, which could run on different operating systems. WinBUGS was then developed by adding useful GUIs 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 the WinBUGS code could be integrated within the R environment, allowing for easier code writing, analysis, and reporting.

Another similar development was JAGS (just another Gibbs sampler). This program analyzed Bayesian hierarchical models by using MCMC simulation. The unique features of JAGS include (i) a cross-platform engine for the BUGS language; (ii) the ability for users to write their own functions, distributions, and samplers; and (iii) 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 design and run a Bayesian study. Web-based applications are particularly useful for timely data entry and analysis. Web services can be called for exchanging information between the database module and the computing module. The success of a Bayesian clinical trial also requires timely and accurate execution of data entry, quality control, adaptive randomization, outcome assessment, and Bayesian computation.

3.2. Using subjective information

Following theoretical and the computational developments in Bayesian methods, 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 into 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 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 [29]. Others stress the importance of being objective and propose an objective Bayesian approach by specifying objective priors [30]. Although the debate continues, the goal is one shared by both communities—to efficiently and accurately infer conclusions on the basis of the data [31].

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*. [32], 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, where parameters at the top of the hierarchy are set to their most likely values, instead of being integrated out.
(2) The reference Bayes approach uses an ‘objective’ or ‘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 on the basis of 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 on the basis of 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) needs 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. 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. In practice, a combination of subjective and objective priors is often used. Because some software does not allow for improper priors, vague proper priors are sometimes used instead. The consequence in such analyses should be carefully assessed (see, e.g., Section 4.2 of Berger, (2006) [30]).

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 was to ascertain how Bayesian methods have been applied in the design and analysis of real clinical trials. The methodology for the literature search is reported in Appendix A. A total of 2012 articles were obtained.

5.1. Study selection

We processed the obtained articles by placing them into exclusion and inclusion categories (see Figure 1). We placed articles for exclusion into four main categories: duplicates (electronically and manually identified), journals (statistical, epidemiological, computer/engineering, and conference papers), subjects (meta-analysis, review/opinion, observational/database, statistical methods, and pharmacokinetics/pharmacodynamics), and additional exclusions (non-medical, non-human, non-English, and non-Bayesian). Note that we excluded 256 articles because they were published in statistical journals and 479 additional articles 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, we excluded 224 articles. We also excluded 141 meta-analysis/systematic review articles because our focus was on individual trials. We excluded 175 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 [33–35]. Several subsequent PK/PD models and programs were developed [36]. 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, 117 articles remained. In addition, we added four known papers that had not been identified by the search algorithm because words 'Bayesian' and 'Bayes' were not mentioned in the keywords nor were they found in the abstract [37–40]. The final number of articles reviewed was 121.

5.2. Data extraction and results

We extracted and summarized key elements of these trials in Tables I through III.

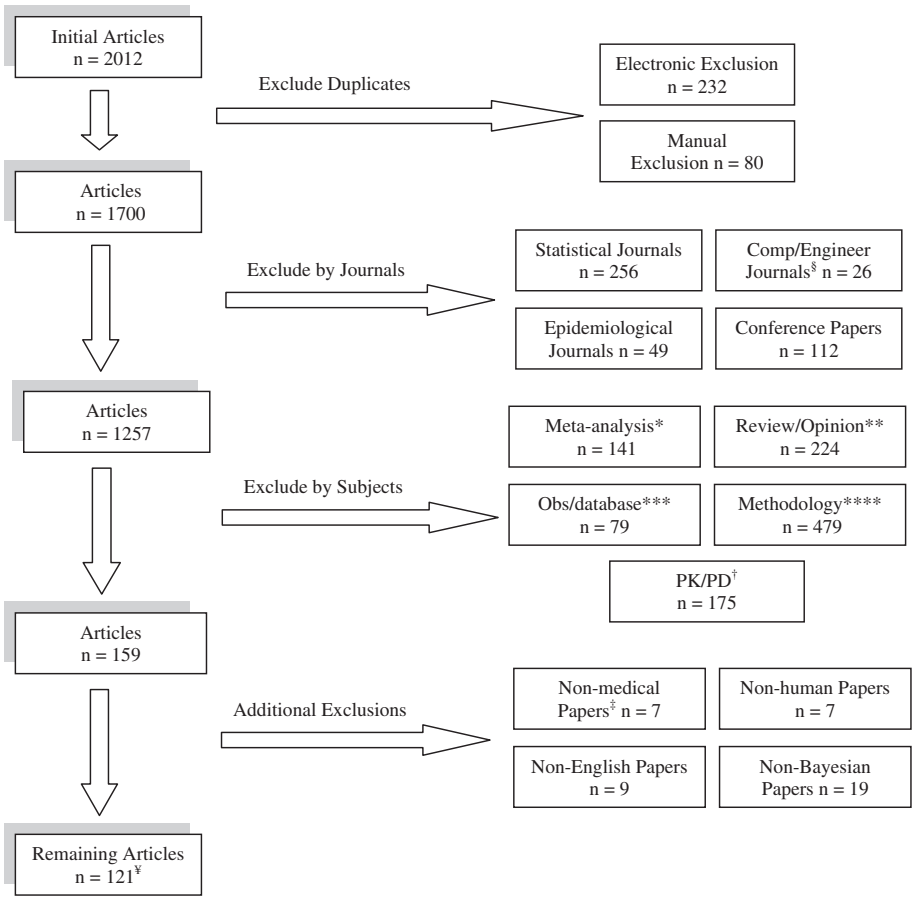

Figure 1. Study selection algorithm. §comp, computers; *also includes Cochrane journals and systematic reviews; **also includes commentaries, letters, replies, surveys, notes, guidelines, and short articles; ***obs, observational; also includes registries and epidemiologic studies; ****methodology focused studies include: statistical strategies, algorithms, trial designs, method comparisons/demonstrations, tutorials, model development/validation, and simulations; † pharmacokinetics/pharmacodynamics; ‡ includes engineering, social science, and policy making studies; ¥ includes 4 manually added articles.

As seen in Table I, publications prior to 1990 included only three clinical trials that used Bayesian methods. That number quickly jumped to 19 in the 1990s and to 99 in the period since 2000. Most trials (62%) applied Bayesian methods for testing treatment efficacy; 12% of the trials applied them for testing treatment safety; 12% of the trials applied them in the areas of medical decision making/cost-benefit analysis; and 10% were association studies. In terms of the medical fields, oncology led the pack (30%), followed by cardiovascular research (16%), and central nervous system research (10%). These 121 papers were published dispersedly in 91 journals, with 11 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).

In Table II, we placed articles into three categories: (i) clinical trials prospectively applied Bayesian design and analysis ($n = 31$); (ii) studies that used a frequentist design with a Bayesian analysis ($n = 72$); and (iii) Bayesian reanalysis studies ($n = 18$), which involved the use of Bayesian methods to retrospectively analyze data from a previously run clinical trial.

We see that the vast majority of the trials are two-arm (57%) or one-arm (32%) studies. About 50% of the trials had a control group. Almost 60% of the studies were randomized trials, 47% 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, 24% were very small ($n \leqq 30$) and 27% had sample sizes between 31 and 99. Sample sizes were between 100 to 499 and 500 to 999, respectively, for 35% and 8% of the trials. Only 6% of the trials had sample sizes of 1000 or more patients. About 46%

Table I. Characteristics of publications reviewed ($n = 121$): Year of publication, type of clinical trial, and medical area of study.

Variable	Frequency	Percentage
Years		
1975–1989	3	2.5
1990–1994	5	4.1
1995–1999	14	11.6
2000–2004	25	20.7
2005–2011	74	61.2
Type of clinical trial		
Association studies	12	9.9
Efficacy	75	62.0
Efficacy and safety	6	5.0
Medical decision making, cost–benefit	14	11.6
Safety*	14	11.6
Medical areas of study		
Addiction	2	1.7
Auditory system	1	0.8
Central nervous system	12	9.9
Cardiovascular system	19	15.7
Dentistry	1	0.8
Gastrointestinal system	3	2.5
Genetics	2	1.7
Genitourinary system	1	0.8
Geriatrics	1	0.8
Hematology	2	1.7
Infectious disease	10	8.3
Metabolic disorder	2	1.7
Obstetrics and gynecology	9	7.4
Oncology	36	29.8
Ophthalmology	1	0.8
Pain	2	1.7
Pediatrics	1	0.8
Pulmonary system	7	5.8
Radiology	2	1.7
Renal system	3	2.5
Transplant	4	3.3

* Includes five continuous reassessment model and two escalation with overdose control papers.

of the trials enrolled patients during 2 or fewer years; whereas 40% of trials spent 2 to 5 years enrolling patients. The remaining 15% of the trials had an accrual period of 6 to 8 years.

Table III shows that continuous, binary, ordinal, and time-to-event variables, respectively, were used as the primary endpoints in 45%, 31%, 12%, and 12% of the trials. A number of Bayesian methods were applied in clinical trials, including for estimation (28%), hypothesis testing (22%), prediction (12%), regression (9%), hierarchical modeling (8%), model selection (7%), sensitivity analysis (4%), a decision-theoretic approach (3%), and a Bayesian network (2%). Informative priors were used in 45% of the trials; noninformative priors were used in 24% of the trials. The remaining 31% 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. Ten percent of the trials were stopped early; six due to futility, four due to efficacy, one for equivalence, and one for toxicity.

5.3. Limitations and overall assessment

Although we made all efforts to best identify the use of Bayesian methods in clinical trials, our search had several limitations. First, we included the words ‘Bayes or Bayesian’ and ‘clinical trials’ in our search criteria. Therefore, we excluded articles without these words. The most prominent inadvertent

Table II. Trial characteristics in publications reviewed ($n = 121$).

Variable	Frequency	Percentage
Bayesian usage		
Bayesian design/analysis	31	25.6
Frequentist design/Bayesian analysis	72	59.5
Bayesian re-analysis	18	14.9
Number of arms*		
1	36	31.6
2	65	57.0
3	7	6.1
⩾4	6	5.3
Use of control group*		
Yes**	53	46.5
None	61	53.5
Method of randomization*		
Adaptive randomization	6	5.3
Equal randomization***	54	47.4
Fixed unequal randomization****	7	6.1
None	47	41.2
Actual sample size		
⩽ 30	29	24.0
31–59	21	17.4
60–99	12	9.9
100–199	16	13.2
200–499	26	21.5
500–999	10	8.3
1000–9999	7	5.8
Accrual period (in months)*****		
⩾12	13	27.1
13–24	9	18.8
25–36	8	16.7
37–48	7	14.6
49–60	4	8.3
6–8 years	7	14.6

*Does not include seven papers with multiple studies.
**Twenty three active, 18 placebo, 1 no treatment, and 11 unspecified.
***Includes 46 two-arm trials, 5 three-arm trials, 2 four-arm trials, and 1 five-arm trial.
****All two-arm trials (one 1:1.5, one 2.8:1, four 2:1, and one 3:1).
*****Seventy three studies did not specify.

exclusions were trials that used the continual reassessment method (CRM) [41] or the escalation with overdose control (EWOC) method [42] 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 B for the search algorithm and results). We included only seven of such trials in the 121 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. We decided to not include these studies, which were not identified in the main search, in the tabulation because similar to the Bayesian PK/PD studies, they form a distinct subgroup.

Overall speaking, there is a growing trend in the application of Bayesian methods in clinical trials albeit its percentage of use is still very small among all trials. A wide variety of Bayesian methods are increasingly being used for assessing efficacy, toxicity, diagnostic and medical decision-making, and so on. Bayesian methods tend to be used more in early-phase trials to assess PK/PD or in dose finding but have been rapidly expanding in its use to provide answers for a myriad of statistical and medical questions as well.

Table III. Statistical attributes in publications reviewed ($n = 121$).

Variable	Frequency	Percentage
Primary endpoint category		
Continuous	54	44.6
Binary	38	31.4
Ordinal	15	12.4
Time to event	14	11.6
Bayesian method		
Estimation	34	28.1
Hypothesis testing	26	21.5
Prediction/forecast	14	11.6
Regression model	11	9.1
Hierarchical model	10	8.3
Model selection/comparison	8	6.6
CRM/EWOC*	7	5.8
Sensitivity analysis	5	4.1
Decision theory	4	3.3
Bayesian network	2	1.7
Prior distribution		
Informative	54	44.6
Non-informative	29	24.0
Unspecified	38	31.4
Number of interim analyses		
0	105	86.8
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	109	90.1
Yes	12	9.9
Reason for stopping early		
Futility	6	50.0
Efficacy	4	33.3
Equivalence	1	8.3
Toxicity	1	8.3

*Five CRM, continuous reassessment model; and two EWOC, escalation with overdose control.

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

A recent review of Bayesian adaptive clinical trials 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 [43]. 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 National Cancer Institute, contributing 4.3%, and Harvard University, contributing 3.9% of the total publications. Four of the nine researchers who had 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 [44]. 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 or 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 2011 August, there were 133 trials and over 4300 patients enrolled in the system. The most commonly used designs were outcome adaptive randomization ($n = 44$), Pocock–Simon baseline adaptive randomization ($n = 42$), the CRM ($n = 29$), 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 [38, 45]. 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 and immunohistochemistry expression analyses to determine the biomarker group for the patient. This process, from registration to reporting the biomarker results, is completed within 2 weeks. On the basis of 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 8 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, and so on. The laboratory technician enters the results of the marker analysis. In order 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

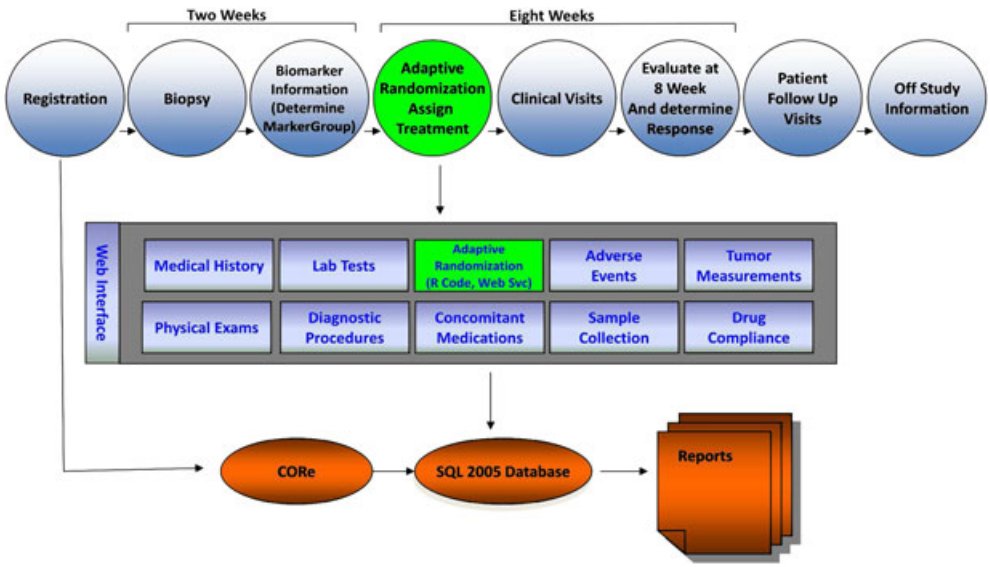

Figure 2. Schematic diagram of the web-based database application for the conduct of the BATTLE trial.

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 2 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 die-hard 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 and software in dose-finding studies in the 1990s [46]. Despite a slow adaption, it is apparent that Bayesian methods are increasingly being used in clinical trials. This trend will likely continue [43, 47]. We have also begun to see the impact of Bayesian applications among regulatory agencies regarding the evaluation and approval of new medical devices or drugs. For example, in safety monitoring, while there are thousands of potential adverse effects, the observed events can be sparse, which make the estimation unreliable. Bayesian hierarchical models can be applied to incorporate prior information, allow borrowing of information, and shrink the estimates toward the mean to yield more reliable inference. The FDA has indicated that 'Safety assessment is one area where frequentist strategies have been less applicable. Perhaps Bayesian approaches in this area have more promise [48].' Bayesian methods have already been successfully applied to provide risk-stratified and real-time safety monitoring in the interventional cardiovascular procedures and the medical device areas [49, 50].

The Center for Devices and Radiological Health at the FDA issued a 'Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials' in 2010 [51]. They indicated that

> Bayesian hierarchical modeling is a specific methodology you may use to combine results from multiple studies to obtain estimates of safety and effectiveness parameters. In a regulatory setting, hierarchical models can be very appealing: They reward having good prior information on device performance by lessening the burden in demonstrating safety and effectiveness. At the same time, the approach can protect against over-reliance on previous studies that turn out to be overly optimistic for the pivotal study parameter.

The Center for Devices and Radiological Health has already approved more than 20 original Pre-Market Approvals (PMAs) and PMA supplements with a Bayesian analysis as the primary method. Many investigational device exemptions and applications for 'substantial equivalence' (510(k)s) that used Bayesian methods have also been approved [52]. On the drug side, the Center for Drugs and Experimental Research of the FDA approved Pravigard Pac (Bristol-Myers Squibb) on the basis of Bayesian analyses of efficacy in 2003 [21]. As many clinical trials using Bayesian methods are underway, it is expected that the FDA will approve more drugs and devices 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 [53]. 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 [43, 47]. (We list some useful information/tools for learning Bayesian clinical trial methods and designs in Appendix C.)

Bayesian methods hold great promise for improving the efficiency and flexibility of conducting clinical trials and are ideal for learning and adaptation [54]. 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 [38], the currently ongoing BATTLE-2 trial, and the I-SPY 2 trial [54, 55]. Successful implementations of Bayesian methods have already been demonstrated in a wide range of clinical trial applications.

'Opposites are not contradictory but complementary.'—Niels Bohr

As pointed out by James Berger in his 2001 Fisher Lecture at the Joint Statistical Meeting, Ronald Fisher, Harold Jeffreys, and Jerz Neyman disagreed on what the correct foundation for statistics are but often agreed on which statistical procedure to actually use. Statistical methods such as Bayesian and frequentist approaches often lead to similar estimation procedures but very different results for hypothesis testing. Berger examined the conditional frequentist approach and attempted to unify the various approaches for testing [56]. 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 [57]. The following year, Roderick J. Little, in his presidential address, proposed the 'calibrated Bayes' approach [58]. 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 [59]. 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 Bayesian trials should be designed and conducted to refine the approach and demonstrate the real benefit of the Bayesian approach in action.

Appendix A. Literature search procedures

We performed a computerized literature search using two major medical indices (Ovid-Medline and Ovid Embase) for all articles published until September 2011. Using Medline (Table A.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.' We further limited the search to 'review articles and meta analysis or review,' producing results that we subsequently removed from the search. We also performed a similar search in Ovid Embase (Table A.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. We limited the search to 'Cochrane library and meta analysis or systematic review' and removed the results from the final search. We combined the publications obtained in the two literature searches (a total of 2012 articles) and imported the references into Endnote X4.

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

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.

In addition, we performed reverse citation searches on two original articles for the CRM and EWOC methods, respectively [41, 42]. The results from all our searches were placed into EndNote X4 for duplicate removal and categorization.

Our reverse citation search netted 362 articles from J. O'Quigley and 109 from J. Babb [41,42]. There were 95 duplicates, 119 statistical journals, 4 computer/engineering journals, 75 review/opinion papers, 100 method papers, 3 meta-analysis papers, 2 observational/database papers, and 5 non-CRM papers. From the resulting journals, we found 59 CRM and 9 EWOC papers.

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

Table A.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 22 September 2011		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 A.II. 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
1	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
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)	46
B. Embase classic + Embase 1947 to 22 September 2011		Search result
1	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
2	clinical trial*.sh.	834 574
3	1 and 2	77

Table A.III. Schema for literature search for escalation with overdose control method.

A. Ovid Medline® in-process and other non-indexed citations and Ovid Medline® 1948 to present		Search result
1	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
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)	2
B. Embase classic + Embase 1947 to 22 September 2011		Search result
1	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
2	clinical trial*.sh.	834 574
3	1 and 2	7

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

A. Articles

1. Berry DA. Bayesian clinical trials. *NatureReviews 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)*. 2^{nd} edition. Springer: New York, 2009.
6. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. 2^{nd} 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.phpSAS 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

Acknowledgements

The work was supported in part by the grants CA016672 and CA097007 from the National Cancer Institute. The authors thank Donald A. Berry for helpful discussions, Richard Herrick for providing information on the MD Anderson CTC web-based software application, and Ms. LeeAnn Chastain for editorial assistance. The authors also thank the reviewers and the associate editor for their constructive suggestions.

References

1. Streptomycin-in-Tuberculosis-Trials-Committee. Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *British Medical Journal* 1948; **2**:769–782. DOI: 10.1136/bmj.2.4582.769.
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. DOI: 10.1002/sim.4780090804.
3. Gehan EA, Schneiderman MA. Historical and methodological developments in clinical trials at the National Cancer Institute. *Statistics in Medicine* 1990; **9**:871–880. DOI: 10.1002/sim.4780090803.
4. Stigler SM. The history of statistics in 1933. *Statistical Science* 1996; **11**:244–252. DOI: 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. DOI: 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.
8. Stigler SM. Thomas Bayes' Bayesian inference. *Journal of the Royal Statistical Society, Series A* 1982; **145**:250–258. DOI: 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.
11. Ederer F. Jerome Cornfield's contributions to the conduct of clinical trials. *Biometrics* 1982; **38 Suppl**:25–32.
12. Ashby D. Bayesian statistics in medicine: a 25 year review. *Statistics in Medicine* 2006; **25**(21):3589–3631. DOI: 10.1002/sim.2672.
13. Grieve AP. 25 years of Bayesian methods in the pharmaceutical industry: a personal, statistical bummel. *Pharmaceutical Statistics* 2007; **6**:261–281. DOI: 10.1002/pst.315.
14. Berger JO, Wolpert RL. *The Likelihood Principle*. SS Gupta, ed. Institute of Mathematical Statistics: Hayward, CA, 1984. http://projecteuclid.org/euclid.lnms/1215466210.
15. Bernardo JM. Bayesian Reference Analysis, A Postgraduate Tutorial Course Universitat de València, Valencia, Spain. (http://www.uv.es/~bernardo/Monograph.pdf.
16. Bernardo JM. The concept of exchangeability and its applications. *Far East Journal of Mathematical Sciences* 1996; **4**:111–121.
17. Betrò B, Bodini A, Guglielmi A. Generalized moment theory and Bayesian robustness analysis for hierarchical mixture models. *Annals of the Institute of Statistical Mathematics* 2006; **58**(4):721–738.
18. Brutti P, De Santis F, Gubbiotti S. Robust Bayesian sample size determination in clinical trials. *Statistics in Medicine* 2008; **27**(13):2290–2306.
19. Andrade JAA, O'Hagan A. Bayesian robustness modelling of location and scale parameters. *Scandinavian Journal of Statistics* 2011; **38**(4):691–711.
20. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research* 2001; **10**(4):277–303. DOI: 10.1177/096228020101000404.
21. Berry DA. Bayesian clinical trials. *Nature Reviews Drug Discovery* 2006; **5**(1):27–36. DOI: 10.1038/nrd1927.
22. Jennison C, Turnbull BW. *Group Sequential Methods with Applications to Clinical Trials*. Chapman & Hall/CRC: Boca Raton, FL, 2000.
23. Dmitrienko A, Wang MD. Bayesian predictive approach to interim monitoring in clinical trials. *Statistics in Medicine* 2006; **25**(13):2178–2195. DOI: 10.1002/sim.2204.
24. 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.
25. 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. Accessed 10/14/2011; available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf.
26. 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. DOI: 10.1109/TPAMI.1984.4767596.
27. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 1990; **85**:398–409. DOI: 10.2307/2289776.
28. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: evolution, critique and future directions. *Statistics in Medicine* 2009; **28**:3049–3067. DOI: 10.1002/sim.3680.
29. Goldstein M. Subjective Bayesian analysis: principles and practice. *Bayesian Analysis* 2006; **1**:403–420. DOI: 10.1214/06-BA116.
30. Berger J. The case for objective Bayesian analysis. *Bayesian Analysis* 2006; **1**:385–402. DOI: 10.1214/06-BA115.

31. 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. DOI: 10.1214/06-BA116C.
32. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley: West Sussex, 2004.
33. 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. DOI: 10.1007/BF01060053.
34. Sheiner LB, Ludden TM. Population pharmacokinetics/dynamics. *Annual Review of Pharmacology and Toxicology* 1992; **32**:185–209. DOI: 10.1146/annurev.pa.32.040192.001153.
35. Whiting B, Kelman AW, Grevel J. Population pharmacokinetics. Theory and clinical application. *Clinical Pharmacokinetics* 1986; **11**:387–401. DOI: 10.2165/00003088-198611050-00004.
36. 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; **9**:E60–E83. DOI: 10.1208/aapsj0901007.
37. 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. DOI: 10.1200/JCO.2005.07.032.
38. Kim ES, Herbst 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. DOI: 10.1158/2159-8274.CD-10-0010.
39. 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. DOI: 10.1200/JCO.2007.14.0665.
40. de Lima M, Champlin RE, Thall PF, Wang X, Martin TG 3rd, Cook JD, McCormick G, Qazilbash M, Kebriaei P, Couriel D, Shpall EJ, Khouri I, Anderlini P, Hosing C, Chan KW, Andersson BS, Patah PA, Caldera Z, Jabbour E, Giralt S. Phase I/II study of gemtuzumab ozogamicin added to fludarabine, melphalan and allogeneic hematopoietic stem cell transplantation for high-risk CD33 positive myeloid leukemias and myelodysplastic syndrome. *Leukemia* 2008; **22**(2):258–264.
41. 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. DOI: 10.2307/2531628.
42. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in Medicine* 1998; **17**:1103–1120. DOI: 10.1002/(SICI)1097-0258(19980530)17:10<1103::AID-SIM793>3.0.CO;2-9.
43. Chevret S. Bayesian adaptive clinical trials: a dream for statisticians only? *Statistics in Medicine* 2011. DOI: 10.1002/sim.4363. [Sept 9 Epub ahead of print].
44. 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. DOI: 10.1177/1740774509104992.
45. 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. DOI: 10.1177/1740774508091815.
46. Tighiouart M, Rogatko A. Dose finding with escalation with overdose control (EWOC) in cancer clinical trials. *Statistical Science* 2010; **25**:217–226. DOI: 10.1214/10-STS333.
47. 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. DOI: 10.1200/JCO.2007.12.1012.
48. Chi G, Hung HMJ, O'Neill R. Some comments on "Adaptive Trials and Bayesian Statistics in Drug Development" by Don Berry. *In Pharmaceutical Report*, 2002. Vol 9, 1–11.
49. Resnic FS, Zou KH, Do DV, Apostolakis G, Ohno-Machado L. Exploration of a Bayesian updating methodology to monitor the safety of interventional cardiovascular procedures. *Medical Decision Making* 2004; **24**(4):399–407.
50. Matheny ME, Ohno-Machado L, Resnic FS. Exploration of a Bayesian updating tool to provide real-time safety monitoring for new medical devices. *AMIA. Annual Symposium proceedings/AMIA Symposium*, AMIA Symposium, 2005; 1045.
51. 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. Accessed 10/14/2011; available at http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf.
52. 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. Accessed 10/14/2011; available at http://www.stat.rutgers.edu/iob/bioconf09/slides/Campbell.pdf.
53. Altman DG, Goodman SN. Transfer of technology from statistical journals to the biomedical literature—past trends and future predictions. *JAMA* 1994; **272**:129–132. DOI: 10.1001/jama.272.2.129.
54. Berry DA. Adaptive clinical trials in oncology. *Nature Reviews Clinical Oncology* 2012; **9**:199–207. DOI: 10.1038/nrclinonc.2011.165.
55. 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. DOI: 10.1038/clpt.2009.68.

56. Berger JO, Christensen R. Could Fisher, Jeffreys and Neyman have agreed on testing? *Statistical Science* 2003; **18**(1):1–32.
57. Efron B. Bayesians, frequentists, and scientists. *Journal of the American Statistical Association* 2005; **100**:1–5. DOI: 10.1198/016214505000000033.
58. Little RJ. Calibrated Bayes: a Bayes/frequentist roadmap. *American Statistician* 2006; **60**:213–223. DOI: 10.1198/000313006X117837.
59. Arjas E. On future directions in statistical methodologies—some speculations. *Scandinavian Journal of Statistics* 2011; **38**:185–194. DOI: 10.1111/j.1467-9469.2011.00737.x.