Biopharmaceutical Section

American Statistical Association

Biopharmaceutical Report

Volume 12, No. 1 Spring 2004

Chair: *Keith Soper* **Editors:** *Demissie Alemayehu, Kevin W. Anderson, and Neal Thomas*

Vaccine Clinical Trials – A Statistical Primer

Devan V. Mehrotra

Clinical Biostatistics, Merck Research Laboratories, Blue Bell, PA 19422

1. Introduction

The remarkable success of many vaccines and their impressive safety record, along with the eradication of smallpox are regarded amongst the greatest public health achievements of the 20th century. Statisticians have contributed (and continue to contribute) significantly toward the research and development of vaccines worldwide. In this article, we discuss some of the statistical issues that arise in all phases of vaccine development, and, where necessary, contrast drug and vaccine clinical trials. A more detailed treatment of this subject is provided by Chan, Wang and Heyse [1].

In the past three decades, there has been an incredible transformation in our understanding of the human immune system and its functions. While statisticians working on vaccine clinical trials are not expected to keep abreast with the latest advances in cellular and molecular immunology, understanding of the basics is essential for proper development of design and analysis strategies. Accordingly, before moving on to statistical issues, we provide a brief review of basic immunology. For more advanced reading, see Abbas, Lichtman and Pober [2].

There are two forms of immunity - innate and adaptive. Innate immunity, the principal components of which include blood proteins and phagocytic cells, provides the first line of defense against microbes (bacteria, viruses, parasites, fungi, etc.). The pathogenicity of microbes is related to their ability to defeat the soldiers of innate immunity. The other form of immunity is called adaptive (or specific) immunity, and evolves as a response to infection. There are two types of adaptive immunity: humoral and cellular. Until the 1970s, only humoral immunity had been well understood. It is mediated by antibodies that primarily defend against *extracellular* microbes. Specifically, recognition of microbes triggers white blood cells called B cells to multiply and secrete antibodies that destroy microbes before they infect host cells. In 1996, two microbiologists (Peter

Contents

Doherty and Rolf Zinkernagel) won a Nobel Prize for deciphering how the cellular immune system works. It is mediated by white blood cells called T cells that defend against intracellular microbes. Specifically, T cells (mostly CD8+ T cells) seek out and destroy cells that have already become infected with the specific microbe. T cells can detect the presence of *intracellular* microbes because infected cells display on their surfaces peptide fragments derived from the pathogens' proteins. The foreign proteins are delivered to the cell surface by specialized host cell glycoproteins called MHC or HLA molecules. The adaptive immune system "remembers" each encounter with a specific microbe, through the establishment of memory B and/or T cells. Subsequent microbe-specific encounters stimulate increasingly effective defense mechanisms, and this *immunological memory* serves as the basis of protective vaccination against microbes.

Virtually all vaccines in use today have been licensed using antibody-based endpoints. More recently, research has intensified on developing vaccines that stimulate cellular immunity (or both). But regardless of whether the vaccine is intended to induce humoral or cellular immunity, the operational goal of vaccination is the same: to *simulate* a microbe-specific exposure so that the host's immune system will generate a pool of memory B and/or T cells to protect against potential real exposures later on. The simulation is accomplished via inoculation of the host by a vaccine that contains either a weakened version of the microbe, or a DNA plasmid or viral vector encoding certain gene(s) of the microbe, and so on.

Understanding the "mechanism of action" of the vaccine is critical for identifying appropriate study endpoints and statistical analyses in clinical trials. For example, several T cell mediated immunity-based vaccines targeted against HIV-1 are currently being developed worldwide. Such vaccines *may* not prevent acquisition of HIV-1 infection, but will hopefully prevent or significantly delay the progression to AIDS among subjects who become infected despite vaccination. From a statistical perspective, this poses a plethora of challenges for the design and analysis of current HIV-1 vaccine trials, including the selection of study endpoints.

In the following sections, we provide an overview of the key statistical issues in each phase of vaccine development.

2. Preclinical phase

Before a vaccine can be tested in humans, it undergoes extensive testing in animals. This is similar to what is done in the preclinical phase for drugs. However, for vaccines there is additional emphasis on the development and validation of *bioassays* to measure the *immunogenicity* of the vaccine, i.e., the ability of the vaccine to induce specific immune responses. The statistical characteristics of an ideal assay include accuracy, unbiasedness, reliability, reproducibility, precision, and ruggedness. In addition, a good assay should have high levels of specificity and sensitivity for the hypothesized *biomarker* of interest (antibody level, T cell response, etc.). Standard statistical tools used in assay development and validation include classic design of experiments (e.g., D-optimal factorial designs), linear and non-linear regression, the four parameter logistic model, concordance correlation, and variance component models. Schofield [3,4] provides an excellent review of this topic.

An important by-product of the assay validation process is identification of what constitutes a *positive* (or perhaps more accurately, a *non-negative*) response to vaccination for each biomarker of interest. The positivity criterion is often (but not always) one dimensional, such as the 99.9th percentile of the estimated distribution of biomarker responses in the absence of vaccination. In such a case, the vaccine is considered *minimally immunogenic* for a given subject if his or her biomarker response is greater than the positivity cut-off. Note that this does not necessarily imply that the vaccine will subsequently provide protection from infection and/or disease for that subject. Often times, the response has to be notably higher than the positivity immunogenicity cut-off for the vaccine to induce a protective effect; we revisit this issue in section 4.

Vaccines are advanced to phase I clinical testing if they are deemed to be generally safe in animals, and for which an adequate proportion of animals exhibit a minimally immunogenic post-vaccination response.

3. Phase I (clinical safety and immunogenicity)

Phase I vaccine clinical trials are small, typically enrolling 30 to 100 human volunteers across multiple investigational centers. They are usually double-blind, placebo controlled trials that study different doses and/or vaccination schedules of the experimental vaccine. The primary focus is on safety and tolerability, but the trials are designed to also provide preliminary assessments of immunogenicity. Note that drug trials typically enroll healthy subjects in phase I, but move to the target population (patients requiring treatment) in phase II and beyond. In contrast, vaccine trials, not surprisingly, involve healthy volunteers in all phases of development. Exceptions include so-called "therapeutic vaccination" studies, which are not discussed here. The statistical challenges there are even greater, since it is difficult to quantify and conclusively demonstrate the benefits of vaccination in subjects that are already infected with the microbe of interest.

Safety in phase I is commonly summarized using the incidence of serious vaccine-related adverse events (if any), along with data on injection-site reactions, body temperatures, systemic adverse events, and laboratory measures. The sparseness of safety data from an individual phase I trial make them more suited for descriptive rather than formal inferential statistical analyses. The decision to proceed to a subsequent trial is therefore based primarily on sound clinical judgment, with input from a safety evaluation committee (if necessary), and regulatory agencies such as the Center for Biologics Evaluation and Research (CBER) for US-based trials.

While statisticians may have a smaller role for safety analyses in phase I, they play a pivotal role in the analysis of immunogenicity. Two types of immunogenicity summaries are reported for the biomarker(s) of interest: the proportion of subjects with a post-vaccination response above the predefined positivity cut-off ("responders"), and the (geometric) mean post-vaccination biomarker response. The small sample sizes in phase I trials pose a multitude of statistical challenges for analyzing immunogenicity. Some of these are readily tackled using a prudent selection of methods from the statistician's existing tool kit. Others present opportunities for innovative analytical solutions and further methodological research. Some of the key statistical issues encountered in phase I analyses of immunogenicity are discussed below under subheadings.

Cross-validation of "positivity" criterion

As mentioned earlier, what represents a "positive" response

to vaccination is determined before phase I clinical trials are begun, in conjunction with the assay validation for the biomarker of interest. It is important to use the accumulating immunogenicity data to either confirm the validity of the positivity criterion, or modify it if necessary. For example, Mogg et al [5] used baseline (pre-vaccination) responses from **559** subjects to cross-validate the two dimensional positivity criterion for the HIV-1 gag specific ELISPOT assay that had been established before phase I clinical trials began. Specifically, they used binomial score intervals, robust parametric methods, and nonparametric density estimates with bootstrap-based confidence intervals to estimate the proportion of "non-responders" that are incorrectly classified as "responders". All three methods converged to a common conclusion, namely that the false positive rate associated with the ELISPOT positivity criterion used by Merck Research Laboratories is estimated to be less than 1% with high confidence.

Stratification

Stratification is often used in vaccine clinical trials; either pre-stratification at the enrollment stage, or post-stratification at the time of analysis. Interestingly, investigational center is rarely used as a stratification factor in phase I because of the small (sometimes zero) sample size per treatment group at each center. Instead, stratification is limited to one or two key prognostic factors that are likely to influence the response to vaccination in a *systematic* way. For example, it is well-known that the ability of a vaccine to induce an antibody-based immune response diminishes with increasing age. Failure to incorporate this important information at either the design or analysis stage can result in a biased and/or inefficient statistical analysis, particularly for small trials! The summary table below reinforces this point. In this hypothetical phase I trial, vaccine A is observed to be more immunogenic than vaccine B for both younger (18-45 years) and older (> 45 years) subjects. However, naïve "pooling" of the results, i.e., failing to adjust for an age effect, yields a result which paradoxically suggests that vaccine B is better!

Hypothetical Data (% Responders)			
Age	Vaccine A	Vaccine B	A - B
≤ 45 years	79.2% (19/24)	64.7% (22/34)	14.5%
> 45 years	12.5% (2/16)	0.0% (0/6)	12.5%
"Pooled"	52.5% (21/40)	55.0% (22/40)	-2.5%

An overview of stratification issues in clinical trials, including references to some recently developed analytic strategies, is provided elsewhere (Mehrotra [6], [7]). The key point here is that stratification-based adjustment for prognostic factors is important for phase I vaccine trials, particularly since the sample sizes are quite small.

Minimum effective dose

As mentioned earlier, phase I vaccine trials often involve multiple dose levels of a vaccine. Interest lies in quantifying the dose-response association, and in identifying the smallest dose that provides adequate immunogenicity. It is usually (but not always) expected that, within the range of doses studied, a higher dose of the vaccine will be at least as immunogenic as a lower dose. Given the small sample sizes in phase I, it is important to use statistical methods that capitalize on this additional biological information to help identify the minimum effective dose. For example, consider the analysis of response proportions. A simple way to proceed is to compare each dose group with placebo using an exact score test for two independent binomials (Suissa and Shuster [8]), and assess the resulting p-values for statistical significance after a multiplicity adjustment (Dunnett [9], Hochberg [10], etc.). However, a more efficient way is to use a step-up trend testing strategy, such as an exact Cochran-Armitage trend test [11-12] embedded within the NOSTATSOT closed-testing procedure (Tukey, Ciminera and Heyse [13]). The gains in statistical efficiency using a trend testing approach over the pairwise approach can be considerable, especially when there are three or more dose levels in a small study (Shirley [14]). This is of particular relevance for phase I vaccine trials. The reason is that larger doses of a vaccine can be substantially more costly to manufacture compared with lower doses. As a result, use of a suboptimal statistical approach can have negative economic ramifications if it contributes to a selection of doses for further study that are considerably larger than the truly minimum effective dose.

"Missing" immunogenicity data

Most vaccine regimens include a sequence of one or more "priming" inoculations followed by a "booster" shot later. In phase I trials, blood samples are collected at one or more time points after each inoculation and assayed for immune activity. The primary analysis focuses on statistical estimation and inference involving the mean post-boost response of the biomarker of interest (μ), and the true proportion of post-boost responders (p). However, the post-boost response is occasionally "missing" for some subjects at the time of analysis. This happens because subjects either drop out of the study prior to the booster or, more commonly, the analysis is an interim look at the data when the subjects in question have received priming inoculations but not yet been boosted.

This situation is similar to the incomplete longitudinal data problem for drug trials. However, there are two key differences. First, while the missing data resulting from dropouts in vaccine trials are typically missing completely at random (MCAR), they are more likely to be either missing at random (MAR) or non-ignorably missing (NM) for drug trials. The reason is that patients often drop out from drug trials because they are not responding favorably to their assigned treatment (e.g., high blood pressure not declining); this concept is generally not applicable for vaccine trials! The second key difference is that the ability to predict or impute the missing data at, say, the last scheduled visit may be better for vaccine trials compared with drug trials. This happens because subjects in vaccine trials are inherently less heterogeneous that patients in drug trials. Moreover, basic immunology tells us that successful priming bodes well for successful boosting, i.e., if the post-prime immune responses are positive, they will almost always be positive post-boost.

So, how should we estimate μ and p? A simple (and common) way is to use a "complete case analysis", i.e., exclude subjects with missing post-boost data. This approach is unbiased under MCAR, but it is also inefficient because it fails to utilize the rich post-prime information of the excluded subjects. A better alternative is to use principled methods for longitudinal data analysis like restricted maximum likelihood (REML), generalized estimating equations (GEE), or multiple imputation, all of which are readily available in standard software. The gains in

efficiency of the latter approaches over the complete case analysis can be significant when the amount of missing data is large (say >20%), as illustrated by Li, Mehrotra and Barnard [15].

4. Phase II/III (clinical immunogenicity, efficacy and safety)

After phase I, there is continued assessment of the immunogenicity and safety of the one or two doses of the vaccine selected for further study. However, the primary focus shifts towards evaluation of vaccine efficacy, and to determine if the biomarker(s) used to advance the vaccine beyond phase I are correlated with efficacy. In this section, we discuss the key statistical issues encountered in phase II/III. Interestingly, for drug clinical trials there is usually a clear demarkation between phase II and phase III, but this is less common for vaccine trials.

Assessing Vaccine Efficacy

After a candidate vaccine has been demonstrated to be immunogenic and generally safe and well tolerated in phase I, controlled clinical trials are conducted to evaluate vaccine efficacy (VE). The "efficacy" of a vaccine refers to its ability to either prevent infection (e.g., for an antibody-based prophylactic vaccine) or reduce the incidence and/or severity of the associated disease in the target population (e.g., for a T cell immunity-based vaccine). Two types of strategies are used in practice. In the first, a "small" phase II proof-of-concept efficacy trial is conducted to get preliminary evidence of vaccine efficacy before moving to a "large" phase III confirmatory trial. In the second, researchers proceed directly to a large pivotal trial (phase II/III combined). The sample sizes required to demonstrate vaccine efficacy trials depend on a multitude of factors, and can range from several hundred subjects to tens of thousands of subjects. O'Neill [16], and Chan and Bohidar [17] describe methodology for sample size estimation to establish vaccine efficacy.

A commonly used measure of efficacy for a vaccine designed to prevent infection is given by $VE = 1 - \left(\lambda_V \div \lambda_C\right)$, where λ_V and λ_C denote the true incidence or hazard rates for the vaccine and control arms, respectively. A vaccine is 100% efficacious if VE = 1. Since a licensed vaccine could ultimately be administered to millions of healthy subjects, it is usually insufficient to demonstrate that the vaccine efficacy is merely greater than zero. Instead, there is a requirement of "super efficacy", i.e., a need to demonstrate with high confidence that the true vaccine efficacy is greater than some pre-specified non-zero lower bound, say VE*. The choice of VE* is influenced by several factors, both statistical and non-statistical; this is analogous to the choice of the non-inferiority or equivalence bound for drug trials. The statistical tools used to quantify vaccine efficacy are context dependent, and include time-to-event analyses based on the Cox model, and conditional and unconditional tests for incidence ratios (Chan [18], Ewell [19]). Interestingly, as is the case for drug trials, there is often a debate on whether the primary analysis should be an "intent-to-treat" analysis or a "per protocol" analysis (Horne et al [20]). Fortunately, the two sets of analyses in vaccine efficacy trials have historically yielded very similar results.

Defining and demonstrating efficacy for a vaccine that is designed to attenuate disease but not necessarily prevent infection is a difficult issue that is beyond the scope of this article. Some progress has been made in this area, for example, by Gilbert et al [21] and Hudgens et al [22] for evaluation of T cell immunity-based HIV-1 vaccines, but more work remains to be done. Other statistical tools that appear promising for the evaluation of such vaccines include the "burden-of-illness" statistic (Chang et al [23]) and the "two part model" (Lachenbruch [24]), both of which provide for a composite evaluation of incidence and severity of disease.

It should be noted that the above discussion of vaccine efficacy has implicitly focused on the *direct* effects of the vaccine. In addition to direct effects, vaccines often confer *indirect* effects through "herd immunity". Related statistical issues are discussed by several authors (e.g., Haber et al [25]), and omitted here for brevity.

Surrogate Markers or "Correlates of Protection"

Vaccine efficacy trials provide valuable data for determining whether the immune biomarker (e.g., antibody or T-cell response) used to assess immunogenicity can also serve as a surrogate marker for vaccine efficacy. For example, suppose that the vaccine is observed to have no efficacy in subjects with low biomarker responses, but has near perfect efficacy in those with high responses (e.g., at a level that is much higher than the positivity cut-off discussed earlier). In this case, use of the Prentice criterion [26] and related approaches [27-28] will easily help formally establish the validity of the biomarker as a surrogate for vaccine efficacy. In contrast, it is very difficult to establish the biomarker as a valid surrogate for vaccine efficacy if the vaccine is highly immunogenic in all subjects, or if the vaccine efficacy is close to 1, for obvious reasons. This was indeed the case for Wyeth-Lederle's PREVNAR®, a seven-valent vaccine licensed in February 2000 to protect infants and children from pneumococcal disease. In a large efficacy trial, all protocol-defined cases of disease occurred in the placebo arm. So, the vaccine was 100% efficacious, but a correlation between immune response and protection from disease could not be determined. Vaccine researchers use the term "correlates of protection" to describe surrogate markers of vaccine efficacy. The availability of such surrogates allows for significantly more efficient evaluation of newer (e.g., 2nd generation) vaccines, since vaccine efficacy can be indirectly demonstrated through the surrogate. A detailed discussion of correlates of protection, including other real examples and a useful bibliography is provided by Chan et al [1].

Assessing Vaccine Safety

As mentioned earlier, vaccines are developed for potential administration to millions of healthy subjects worldwide. Accordingly, the assessment of safety is of paramount importance, and requires a comprehensive evaluation to ensure that the benefits of vaccination outweigh the potential risks. The methods and measurements chosen to establish the safety of a vaccine depend on many factors, including the type of vaccine and its mechanism of action.

Common reactions to vaccines are readily identified in phase I, and continue to be tracked in phases II and III. These include swelling, tenderness, and redness at the injection site (e.g., arm), and are almost always attributable to the vaccine. Systemic reactions, such as fevers or muscle aches, are also fairly common for some types of vaccines (and placebo!) It is important to stress that the large volume of safety data, either for a single phase III trial or an integrated summary of safety

across several trials, calls for careful statistical analysis and interpretation. For example, systemic adverse events (AEs) are typically evaluated using between-group p-values for every AE encountered within each of several body systems. If the p-values are interpreted without multiplicity considerations, there is a potential for an excess of false positive findings. This can needlessly complicate the safety profile of the vaccine under study. Mehrotra and Heyse [29] have recently proposed a novel method for taming the multiplicity artifact in such situations. Their method involves a two-step application of adjusted p-values based on the Benjamini and Hochberg [30] false discovery rate methodology. They use real data from three moderate to large vaccine trials to illustrate their proposed "Double FDR" approach, and to reinforce the potential impact of failing to account for multiplicity.

Phase II/III vaccine trials are usually well powered for comparative analyses of common but less serious adverse events. However, determining the sample size required to rule out less common but more serious adverse events is a challenging issue that requires context-dependent solutions. For example, Sadoff et al [31] described the study design considerations necessary to detect an increased risk of intussusception in a randomized, placebo-controlled trial of a rotavirus vaccine. They proposed extensive monitoring for intussusception cases through multiple stopping boundaries, and used Monte Carlo simulation methods to justify a study size of at least 60,000 infants. A more detailed discussion of statistical design and analysis issues involving vaccine safety are provided by Ellenberg [32].

Other Pre-Licensure Issues

The licensing application for a new vaccine is called the Biological License Application (BLA); it is analogous to the New Drug Application (NDA) for a drug. In order for the license to be approved by a regulatory agency (like CBER in the US), the BLA must provide convincing data to support the safety and efficacy of the vaccine. In addition, it must demonstrate that the product meets regulatory standards of purity and potency, and consistency of manufacturing (Lachenbruch et al [33]). Evidence for the latter is obtained through a "lot consistency" study. Such studies typically use three lots of vaccine made from the same manufacturing process. The goal is to demonstrate that the three lots evoke "similar" immune responses. Similarity is concluded if a pre-specified clinically significant difference between any two pairs of lots can be ruled out with high confidence, with respect to both the proportion of responders and the (geometric) mean response for the primary biomarker. Statistical methods for lot consistency studies are discussed by Wiens and Iglewicz [34].

5. Post-Licensure Issues

Phase IV studies are conducted after licensure to collect additional information on the safety, immunogenicity, and/or efficacy of the vaccine to meet regulatory commitments or post-marketing objectives. These include so-called bridging studies, persistence studies, and post-licensure safety studies. Some of the attendant statistical issues are briefly discussed here. Chan et al [1] and Halloran [35] provide more detail.

Bridging Study

After the vaccine has been licensed, the manufacturing process, storage conditions, or dosing schedule may be altered to enhance production yield, vaccine stability, or convenience of vaccination schedule, respectively. Regulatory requirements mandate that sponsors conduct studies to demonstrate that such changes have no material impact on vaccine effectiveness. This is accomplished via immunogenicity bridging trials designed to demonstrate similarity of the modified vaccine/process to the current vaccine/process in a manner analogous to that for lot consistency trials. Of note, it is presumed that the biomarker used to establish similarity in a bridging study is sufficiently correlated with efficacy. Lack of such a correlation makes it harder to justify the use of the biomarker, since the ultimate goal is to (indirectly) ensure that the vaccine efficacy is unaffected.

Immunological Persistence Study

It is important to have an understanding of how long vaccine-induced immunity lasts. For example, if the protective efficacy of a vaccine is known to last for ten years, then giving a booster shot every ten years might be reasonable. However, such information is rarely available before the vaccine is licensed. The reason is that the expected duration of vaccine-induced immunity is usually much longer than the duration of the clinical trials that are included in the BLA. Accordingly, immunological long-term persistence studies are often conducted post-licensure. These are typically open label studies in which vaccinated subjects provide blood samples over time (usually annually) for determination of immune responses. The resulting data can be analyzed using standard time-to-event methodology. Modeling strategies have also been proposed to predict the duration of vaccine-induced immunity based on extrapolation of observed antibody or cellular immune responses from clinical trials [36-37].

Post-licensure safety surveillance

The Food and Drug Administration (FDA) and the Centers for Disease Control (CDC) have created the Vaccine Adverse Event Reporting System (VAERS) for post-licensure safety surveillance [38]. This system accepts reports of adverse events that may be associated with U.S. licensed vaccines from health care providers, manufacturers, and the public. The reports are continually monitored for any unexpected patterns or changes in rates of adverse events. Post-marketing safety evaluations are often complicated and contentious, particularly when they are based on retrospective analyses or involve data collected via potentially biased reporting systems. Bayesian data mining methods have been proposed by DuMochel [39] and implemented in practice by Niu et al [40]. See also Brewer and Colditz [41] for an informative discussion on post-marketing safety issues.

6. Concluding remarks

In this article, we have provided an overview of the key statistical issues that arise in all phases of vaccine development. We have stressed the importance of understanding the science behind the numbers, including how the vaccine is intended to work, as well as the bioassays that measure whether or not the vaccine is immunogenic. Since licensed vaccines are administered to millions of healthy people, we have highlighted the importance of establishing vaccine safety in a large number of subjects, and explained the concept of super efficacy studies. Finally, we have noted the importance of establishing that the

vaccine manufacturing process produces vaccine lots that evoke statistically similar post-vaccination immune responses. Recent advances in genetic engineering and pharmacogenetics are spawning a new generation of vaccine modalities to protect against HIV/AIDS, cancer, malaria, anthrax, plague, and so on. Development of such vaccines will pose additional statistical challenges that will require innovative thought and creative solutions.

References

1. Chan, I.S.F., Wang, W.W.B., Heyse, J.F. (2003). Vaccine Clinical Trials. *Encyclopedia of Biopharmaceutical Statistics.* Second Edition, Edited by Chow, S.C., Marcel Dekker, Inc.: New York, 1005-1022.
2. Abbas, A.K., Lichtman, A.H., Pober, J.S. (1997). *Cellular and molecular immunology.* Third edition. W.B. Saunders Company, Philadelphia, PA.
3. Schofield, T.L. (2003). Assay Development. *Encyclopedia of Biopharmaceutical Statistics.* Second Edition, Edited by Chow, S.C., Marcel Dekker, Inc.: New York, 55-62.
4. Schofield, T.L. (2003). Assay Validation. *Encyclopedia of Biopharmaceutical Statistics.* Second Edition, Edited by Chow, S.C., Marcel Dekker, Inc.: New York, 63-71.
5. Mogg, R., Fan, F., Li, X., Dubey, S., Fu, T-M., Shiver, J., Mehrotra, D. (2003). *Statistical Cross-Validation of Merck's IFN-gamma ELISPOT Assay Positivity Criterion.* AIDS Vaccine Conference, New York, NY.
6. Mehrotra, D.V. (2001). Stratification issues with binary endpoints. *Drug Information Journal*, 35, 1343-1350.
7. Mehrotra, D.V. (2002). Stratified Comparative clinical trials – analysis and interpretation issues. *Proceedings of the 21st International Biometric Conference*, Freiburg, Germany, 201-215.
8. Suissa, S., Shuster, J.J. (1985). Exact unconditional sample sizes for the 2X2 binomial trial. *Journal of Royal Statistical Society*, Series A, 148, 317-327.
9. Dunnett, C.W. (1955). A multiple comparison procedure for comparing several treatments with a control. *JASA*, 50, 1096-1121.
10. Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75, 800-802.
11. Cochran, W.G. (1954). Some methods for strengthening the common c2 tests. *Biometrics*, 10, 417-451.
12. Armitage, P. (1955). Tests in linear trends in proportions and frequencies. *Biometrics*, 11, 375-386.
13. Tukey, J.W., Ciminera, J.L., Heyse, J.F. (1985). Testing the statistical certainty of a response to increasing doses of a drug. *Biometrics*, 41, 295-301.
14. Shirley, E.A. (1985). The value of specialized tests in studies where ordered group means are expected. *Statistics in Medicine*, 4, 489-496.
15. Li, X., Mehrotra, D.V., Barnard. (2003). Analysis of incomplete longitudinal binary data. Presented at the Joint Statistical Meetings, San Francisco, CA.
16. O'Neill, R.T. (1988). On sample sizes to estimate the protective efficacy of a vaccine. *Statistics in Medicine*, 7, 1279-1288.
17. Chan, I.S.F., Bohidar, N.R. (1998). Exact power and sample size for vaccine efficacy studies. *Communications in Statistics - Theory and Methods*, 27, 1305-1322.
18. Chan, I.S.F. (1998). Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies. *Statistics in Medicine*, 17, 1403-1413.
19. Ewell, M. (1996). Comparing methods for calculating confidence intervals for vaccine efficacy. *Statistics in Medicine*, 15, 2379-2392.
20. Horne, A.D., Lachenbruch, P.A., Goldenthal, K.L. (2001). Intent-to-treat analysis and preventive vaccine efficacy. *Vaccine*, 19, 319-326.
21. Gilbert, P.B., DeGruttola, V.G., Hudgens, M., Self, S.G., Hammer, S.M., Corey, L. (2003). What constitutes efficacy for an HIV vaccine that ameliorates viremia: issues involving surrogate endpoints in phase III trials. *Journal of Infectious Diseases*, 188, 179-93.
22. Hudgens, M.G., Hoering, A., Self, S.G. (2003). On the analysis of viral load endpoints in HIV vaccine trials. *Statistics in Medicine*, 22, 2281-2298.
23. Chang, M.N., Guess, H.A., Heyse, J.F. (1994). Reduction in burden of illness: a new efficacy measure for prevention trials. *Statistics in Medicine*, 13, 1807-1814.
24. Lachenbruch, P.A. (2001). Power and sample size requirements for two-part models. *Statistics in Medicine*, 20, 1235-1238.
25. Haber, M., Longini, I.M., Halloran, M.E. (1991). Measures of the effects of vaccination in a randomly mixing population. *International Journal of Epidemiology*, 20, 300-310.
26. Prentice, R.L. (1989). Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in Medicine*, 8, 431-440.
27. Lin, D.Y., Fleming, T.R., DeGruttola, V. (1997). Estimating the proportion of treatment effect explained by a surrogate marker. *Statistics in Medicine*, 16, 1515-1527.
28. Buyse, M., Molenberghs, G. (1998). The validation of surrogate endpoints in randomized experiments. *Biometrics*, 54, 1014-1029.
29. Mehrotra, D.V., Heyse, J.F. (2004). Use of the false discovery rate for evaluating clinical safety data. *Statistical Methods in Medical Research*, in press.
30. Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, Series B, 57, 289-300.
31. Sadoff, J., Heyse, J.F., Heaton, P., Dallas, M. (2000). Designing a study to evaluate whether the Merck rotavirus vaccine is associated with a rare adverse reaction (intussusception). *U.S. Food and Drug Administration Workshop: Evaluation of new vaccines: how much safety data?* Bethesda, MD.
32. Ellenberg, S.S. (2001). Safety considerations for new vaccine development. *Pharmacoepidemiology and Drug Safety*, 10, 1-5.
33. Lachenbruch, P.A., Horne, A.D., Lynch, C.J., Tiwari, J., Ellenberg, S.S. (2000). Biologics. *Encyclopedia of Biopharmaceutical Statistics*. Second Edition, Edited by Chow, S.C., Marcel Dekker, Inc.: New York, 47-54.
34. Wiens, B.L., Iglewicz, B. (2000). Design and analysis of three treatment equivalence trials. *Controlled Clinical Trials*, 21, 127-137.
35. Halloran, M.E. (2001). Vaccine studies. *Biostatistics in*

Clinical Trials, Edited by Armitage P., Colton T., John Wiley & Sons, New York, 479-486.
36. Wiens, B.L., Bohidar, N., Pigeon, J., Egan, J., Hurni, W., Brown, L., Kuter, B., Nalin, D. (1996). Duration of protection from clinical hepatitis A disease after vaccination with VAQTA. *Journal of Medical Virology*, 49, 235-241.
37. Pigeon, J.G., Bohidar, N.R., Zhang, Z.X., Wiens, B.L. (1999). Statistical models for predicting the duration of vaccine-induced protection. *Drug Information Journal*, 33(3), 811-819.
38. Chen, R.T., Rastogi, S.C., Mullen, J.R., Hayes, S.W., Cochi, S.L., Donlon, J.A., Wassilak S.G. (1994). The vaccine adverse reporting system (VAERS). *Vaccine*, 12, 42-50.
39. DuMouchel, W. (1999). Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *The American Statistician*, 53, 177-190.
40. Niu, M.T., Erwin, D.E., Braun, M.M. (2001). Data mining in the US vaccine adverse event reporting system (VAERS): early detection of intussusception and other events after rotavirus vaccination. *Vaccine*, 19, 4627-4634.
41. Brewer, T., Colditz, G.A. (1999). Postmarketing surveillance and adverse drug reactions: current perspectives and future needs. *JAMA*, 281(9), 824-829. ■

Letter from the Chair

Keith Soper

The Biopharmaceutical Section of ASA was formed in 1968 to provide a forum for "statisticians dealing with the quantitative aspects of drugs in the areas of experimental therapeutics, clinical pharmacology, and other bio-pharmaceutical sciences". That covered a lot of ground in 1968 and today biopharmaceutical statistical science has broadened further to include high throughput screening, genomics, formulation, vaccines, bioequivalence, manufacturing, post-marketing surveillance, and health economics, among other areas. Our Section has grown to become one of the largest and most active Sections in ASA. We provide a combination of services to members that you cannot get anywhere else.

Much of our effort, naturally, is directed to making the Joint Statistical Meetings as useful and relevant as possible to our members each year. In 2004 our Section will sponsor six invited sessions, over a dozen luncheon roundtables, over two dozen regular and special topic contributed sessions, as well as continuing education classes designed to keep your technical skills sharp. We also help organize sessions on biopharmaceutical topics at the ENAR meeting, and work behind the scenes to help coordinate topics at JSM and ENAR with other conferences such as ASQ (Deming) and the MidWest Biopharmaceutical (Muncie) Workshop.

We seek to recognize and promote excellence in our profession in many ways, including awards for the best student papers each year, and for the best presentation of a biopharmaceutical contributed paper at JSM. If you know of a leader in our profession who has not yet been elected a fellow of ASA, we can help you prepare a strong nomination for him or her (and if they've been active supporters of the Section, write a letter on their behalf).

Each year we partner with the FDA to organize an FDA/Industry workshop where hundreds of statisticians from the industry and FDA can meet to hear statistical research directly relevant to our work, discuss regulatory issues, or just get to know our colleagues in a relaxed setting. If you do not have the opportunity to travel to the Washington, D.C. area, you can get our "Biopharmaceutical Report" published periodically via e-mail.

If you are already a member of our Section, you know the benefits. If not, the modest annual dues will give a great return on your investment. We hope to see many of you at our annual meeting in Toronto at JSM this year. ■

2004 Biopharmaceutical Section Corporate Sponsors

Allergan

Amgen

Astra Zeneca

Aventis

Biogen IDEC

Bristol-Myers Squibb

Centocor

Cytel Software

Eli Lilly

Genentech

Genzyme

GlaxoSmithKline

Hoffman LaRoche

ICON

Insightful Corporation

Johnson & Johnson

Merck

Novartis

Pfizer

PharmaNet

Purdue Pharma

Rho, Inc.

Sanofi-Synthelabo

SAS

Schering-Plough

Takeda

3M Pharmaceuticals

The 27th Annual Midwest Biopharmaceutical Statistics Workshop

May 24 – 26, 2004
Ball State University, Muncie, Indiana

Final Program

Monday, May 24

8:30 am – 5:00 pm
Workshop Registration
Fee: $140 until May 1 ($45 for students), $160 after May 1

9:00 am – 1:00 pm
Short Course
(Separate Registration Fee: $55)
Presenter: *Georges Grinstein, University of Massachusetts*
Topic: High Dimensional Data Visualization for Solving Complex Data Exploration and Discovery Problems

2:15 pm – 2:30 pm
Introduction and Welcome
Tim Schofield, Merck
Michael A. Maggiotto, Dean, College of Sciences and Humanities, Ball State University

2:30 pm – 4:30 pm
Plenary Session
Speaker: *Harald Martens, Matforsk/Norwegian Food Research Institute*
Topic: **Analysis of Chemometric Data**

Tuesday Morning, May 25

Concurrent Sessions
8:30 am – 11:30 am

A. Statistical Analysis of New Studies Required in Preclinical Safety Testing
Organizer/Chair: *Steve Bailey, Wyeth*

1. "New Methodologies for QT Interval Prolongation Adjustments in Preclinical Safety Pharmacology Studies," Lori Mixson, Merck
2. "Design and Analysis of Juvenile Animal Toxicology Studies in Support of Pediatric Drug Products," Edmund Kadyszewski, Pfizer
3. "Statistical Aspects of the Auditory Startle Experiments in Behavioral Toxicology Studies," Wherly Hoffman, Eli Lilly
4. "Design and Analysis of CNS/FOB Studies in Preclinical Safety Pharmacology Testing," Steven Bailey, Wyeth

B. Novel Designs and Their Analyses in Early Clinical Trials
Organizer/Chair: *Steve Gulyas, Pfizer*

1. "Design and Analysis Considerations for a Clinical Endpoint Based on Count Data," Mani Lakshminarayanan and Aparna Raychaudhuri, Centocor
2. "How we Broke the Rules and Got Away with It: Use of a Latin Square to Evaluate Symptomatic Treatment of AD," Wayne Ewy, Pfizer
3. "Biomarkers of Stroke: A Bivariate Outcome Design and **Mixed Model Analysis**," *Jessica Mancuso, Pfizer*
4. Discussant: Tim Montague, GlaxoSmithKline

C. Tools and Methods for Large Data
Organizer/Chair: *Kjell Johnson, Pfizer*

1. "Approaches to Analyzing Large, Drug Discovery Data Sets," Kimberly Crimin and Thomas Vidmar, Pfizer
2. "Autonomous Fast Classifiers For Pharmaceutical Data Sets," Paul Komarek, Carnegie Mellon University
3. "Top Mistakes with Large Data," John Elder, Elder Research

Tuesday Afternoon, May 25

Poster Session
12:00 pm – 1:30 pm
Chair: *Jackie Reisner, Pfizer*
Posters will be accepted on any biopharmaceutical statistical topic. Abstracts must be received by May 1. Students may qualify for the Charlie Sampson poster award if abstract, poster panels and a paper briefly describing the poster are received by May 1.
For more information contact Jackie at (734) 622-4996 or *jacqueline.k.reisner@pfizer.com*

Concurrent Sessions
1:30 pm – 4:30 pm

A. Equivalence Testing for CMC: Approaches and Challenges
Organizer/Chair: *Douglas Lee, Pfizer*

1. "Effectively Incorporating Experimental Design and

The 27th Annual Midwest Biopharmaceutical Statistics Workshop

May 24 – 26, 2004
Ball State University, Muncie, Indiana

Final Program *(continued)*

Analysis into Analytical Method Transfer Exercises," Greg Steeno, Pfizer
2. "Assessment of Population Bioequivalence Criteria Applied to CMC in Vitro Data," Beth Morgan, Glaxo-SmithKline
3. "Assessing In Vitro Bioequivalence for Profile Data: A New Modeling Approach," Bin Cheng, University of Wisconsin
4. "Multivariate Equivalence Testing with an Elliptical Hypothesis Boundary," Mark Berry, Pfizer

B. Robust Statistical Methods for Clinical Trials
Organizer/Chair: *Robin Mogg, Merck*

1. "Some Recent Advances Related to Robust ANOVA and Regression," Rand Wilcox, University of Southern California
2. "RAVE Analysis of Longitudinal Clinical Trials," Devan V. Mehrotra, Merck
3. "Robust Estimation in Linear Mixed-Effects Models Using the Multivariate t-Distribution," Jose Pinheiro, Novartis Pharmaceuticals
4. Discussant: George Milliken, Kansas State University

C. Multiple Data Block Analysis
Organizer/Chair: *Stan Young, NISS*

1. "Integrative Analysis of High Dimensional Gene Expression, Metabolite and Protein Data," Raymond Lam, Lei Zhu, Kwan Lee, Amit Bhattacharyya, Alan Menius, Biomedical Data Sciences, GlaxoSmithKline
2. "Three-way Analysis: Micro Array, Biological Potency, and Molecular Descriptors," Jack Liu and S. Stanley Young, NISS
3. "Three-way Analysis," Harald Martens, Matforsk/ Norwegian Food Research Institute

Tuesday Evening Banquet
Announcement of Student Winner of Charlie Sampson Poster Award
Speaker: *Tony Lachenbruch, CBER*
Topic: **Lessons from a Life in Statistics**

Wednesday Morning, May 26

Concurrent Sessions
8:30 am – 11:30 am

A. Methodologies and Challenges in Assay Validation
Organizer/Chair: *Kristi Griffiths, Eli Lilly*

1. "Challenges in a Priori Acceptance Criteria Establishment for Assay Validations," Randy Rafferty, Eli Lilly
2. "Unique Challenges in the validation and Routine Monitoring of Bioassays," Charles Tan, Merck
3. "Challenges and Issues with LOD, LOQ and Reporting Threshold Establishment and Use," Dave LeBlond, Abbott

B. Multiple Comparisons
Organizer/Chair: *Alex Dmitrienko, Eli Lilly*

1. "Decision-theoretic Views on Switching between Superiority and Non-inferiority Testing," Peter Westfall, Texas Tech University
2. "Applying the Partitioning Principle to Dose-response and Multiple Endpoint Problem," Jason Hsu and Haiyan Xu., Ohio State University
3. "Multiple Co-primary Endpoints where All Must Achieve Statistical Significance," Walt Offen, Eli Lilly, Paul Stryszak, Schering-Plough, and Alex Dmitrienko, Eli Lilly

C. Subgroup Analysis
Organizer/Chair: *Alan Menius, GlaxoSmithKline*

1. "Data Mining Large Clinical Databases: Searching for Important Subgroups," Daniel Park and Kwan Lee, Glaxo-SmithKline
2. "Practical Considerations in Analyzing Pharmacogenetic Data Sets," Michael Man, Pfizer
3. "Competing Tree Techniques for Subgroup Analysis," Joe Boyer, North Carolina State University

For more information on the workshop, please contact MIR ALI, Ball State University, (765) 285-8670, Email: *mali@bsu.edu* or Ying Zhang, Quintiles, (816) 767-4679, Email: *yings.zhang@quintiles.com*. The preliminary program will be updated periodically at the web site *www.mbswonline.com/*. ■

Let's Hear from You!

If you have any comments or contributions, contact **Editor:** Neal Thomas, Pfizer, Clinical Biostatistics, Eastern Point Road/MS 8260-2227, Groton, Connecticut, 06340; Phone 860-715-0268; email: *snthomas99@yahoo.com,* **Editor:** Kevin W. Anderson, Biostatistical Consultant, 3400 Iroquois Way, Ambler, Pennsylvania 19002; Phone (215) 646-4453; email: *kwanderson@rcn.com*; or **Editor:** Demissie Alemayehu, Biostatistics & Reporting, 205-9-13, Pfizer Inc., New York, New York 10017; Phone 212-573-2084; email: *alemad@pfizer.com.*

The ***Biopharmaceutical Report*** is a publication of the Biopharmaceutical Section of the American Statistical Association.

Printed in the United States of America

Biopharmaceutical Section
American Statistical Association

Biopharmaceutical Report
c/o American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3415
USA