

Why Single Pharmacogenomic Trials Should be Very Significant?

Clive E Bowman*

School of Mathematical Sciences, The University of Reading, United Kingdom

*Corresponding author: Clive E Bowman, School of Mathematical Sciences, The University of Reading, Whiteknights, Reading, RG6 6AH, United Kingdom, Tel: 44 7810 506362; E-mail: c.e.bowman@reading.ac.uk

Received date: May 28, 2014, Accepted date: May 29, 2014, Published date: June 6, 2014

Copyright: © 2014 Bowman CE. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Recasting traditional frequentist thinking into Bayesian beliefs illustrates that requirements for 'strength of evidence' should drive the sizing of pharmacogenomic trials not hopeful hyperbolae. There is no 'free lunch' in pharmacogenomic investigations.

Keywords Pivotal; Phase 3

Editorial

Many authors over the years have held up the promise that pharmacogenomic clinical trials can be smaller than their non-pharmacogenomic equivalents [1-3]. This is certainly true if the pharmacogenomically targeted population has a smaller variance in any efficacy or safety outcome compared to a heterogeneous 'all-comers' sample [4]. Clinical trial sizes (n) scale with the variance of measurements i.e. with the square of typical measurement error [5]. A homogenous population segment of patients selected on a specific biomarker should have a smaller error variance and thus need a smaller trial to investigate. However, to non-aficionados of statistics this has led to the belief that *per se* pharmacogenomic trials are something special beyond the realms of ordinary regulatory concerns about trial sizing. Often pharmaceutical executives mistakenly hear, or may be unscrupulously led to believe, that a segmented medicine will always be in some sense cheaper or quicker to develop. Considering the weight of evidence requirements [6] is an illuminating way of understanding that the sizing of pharmacogenomic trials are not exceptional when it comes to regulatory sufficient proof.

A positive pharmacogenomic result in a clinical trial arises when a hypothesis of 'no-effect' of the genomic biomarker on the outcome is rejected by the data. Statisticians frequently refer to this as 'rejecting the null' in favour of 'the alternative' (where the alternative hypothesis is that there is some pharmacogenomic effect). One could look at these hypotheses as encapsulating the experimenter's subjective belief about the world.

Given these two mutually exclusive and exhaustive pharmacogenomic hypotheses H_0 (belief in the null) and H_1 (belief in the alternative) then Bayes rule gives the posterior probability of the null hypothesis (i.e. after the trial data has been collected) as:

$$p(H_0|D) = \frac{p(D|H_0) \cdot p(H_0)}{p(D|H_0) \cdot p(H_0) + p(D|H_1) \cdot p(H_1)}$$

Where, | means 'given', D is the trial's data and $p(H)$ is the prior probability of the hypothesis. Before an experiment, the belief in either hypothesis may be equally likely (i.e. $p(H) = 0.5 \forall H$) or as in the crucial equipose philosophy of standard frequentist sample size calculations). This Bayesian approach belies the fact that H_0 is a point hypothesis

while the alternative spreads the probability out according to the extra parameter(s) and selection specification(s) included in H_1 [7], but is a useful framework here for illustration.

How can this argument be used to approximate the support for belief in hypotheses when designing a pharmacogenomic clinical trial programme?

Consider, that a pivotal phase 3 pharmacogenomic trial is planned such that the *propter hoc* type 1 error is 0.05 one-sided and the *propter hoc* type 2 error is 20%. The trial runs and all that one knows is that it rolls out a 'significant' p value for the observed pharmacogenomic result - the actual values of the estimated parameters being unknown.

Then from the above, as the probability of obtaining a result at least as extreme as this obtained under the null belief is $\leq 0:05$, and under the alternative was planned as $\leq 0:8$, then the approximately

$$posterior(H_0|data) \leq \frac{0.05}{0.8+0.05} = \frac{1}{\frac{0.8}{0.05} + 1} = 0.059$$

Since the equal priors of the two hypotheses cancel, and approximately

$$posterior(H_1|data) \geq \frac{0.8}{0.8+0.05} = \frac{1}{1 + \frac{0.05}{0.8}} = 0.941$$

Note the key importance of the relative weight of evidence ($\frac{0.8}{0.05}$, the surrogate for the actual observed likelihood ratio).

In terms of a probability ratio or Bayes Factor [8] in favour of the alternative, this is $=0.941/0.059 = 16$ (or $\frac{\ln(16)}{\ln(2)}$ [7]). As a guide Evett et al. [9] this probability ratio would be considered as 'Moderate evidence to support' by Aitken et al. [10] and in bits as 'Strong' by Jeffreys [7]. It is equivalent to a weight of evidence ($=10 \times \log_{10}$ [probability ratio]) of 12.04 decibans [7], being 'Moderate to strong' according to Good [11]. It maps to the general statistical theory 'deviance' ($=$ twice natural logarithm of likelihood ratio) used for hypothesis testing in linear modeling of 5.55.

$$\chi^2_{1,\alpha=0.05} = 3.84, \chi^2_{1,\alpha=0.025} = 5.02, \chi^2_{1,\alpha=0.01} = 6.63$$

In accordance with general regulatory practice in pharmaceutical development, a second (independent) replicate pivotal trial is planned with the same assumptions as the first and let us assume that it itself rolls out with an unspecified 'significant' pharmacogenomic result. Then, putting aside for now issues of whether the actual parameter values found are the same as before or not, given this extra data, it follows again from the Bayes argument above that the new posterior for belief in the alternative is

$$p(H_1 | data) \geq \frac{\left(\frac{0.8 \times 0.8}{0.85}\right)}{\left(\frac{0.05 \times 0.05}{0.85}\right) + \left(\frac{0.8 \times 0.8}{0.85}\right)} = 0.996$$

And the *new* belief in the null hypothesis posterior is ≤ 1 minus this.

By now, evidence for the alternative pharmacogenomic hypothesis (i.e. that there is a pharmacogenomic effect) is: $0.996/0.004 =$ a Bayes Factor of 249 ('Moderately strong evidence to support [10]), or 7.96 bits ('Decisive' [7]), or 23.96 decibans ('Strong' [11]). This now maps to an impressive deviance of 11.04.

$$\chi^2_{1,\alpha=0.01} = 6.63, \chi^2_{1,\alpha=0.005} = 7.79, \chi^2_{1,\alpha=0.001} = 10.83$$

Now imagine instead that a pharmaceutical company executive believes that pharmacogenomic trials are somehow special and can be done differently? Perhaps just with one 'smaller, cheaper' single pivotal trial? What obtained 'significance' level matching this planned trial type 1 error would generate at least the equivalent alternative hypothesis posterior to executing the two trials (given a planned 80% power)?

For a single planned pivotal trial with a pre-specified one sided type 1 error and 80% power, the answer is given through simple algebraic equivalence as $\frac{0.8}{\alpha+0.8} = 0.996$ to which the solution is a planned (and obtained) $\alpha = 0.003$ - close to heuristically assuming the resultant support is 0.05 (for the first trial) times 0.05 (for the second trial) and very near the standard 'extreme p value' of 0.001 desired in regulatory guidelines for single pivotal trials [12]. One assumes that regulators require a little lower type one error for confidence in resultant assertions as they have now lost the reliability of 2 'independent' sets of pharmacogenomic investigators investigating the drug. This extra burden of proof also compensates for the fact that despite the consilience in rejecting the null belief, the actual parameter values whilst considered the same as before will vary in practice in each trial, as well as allowing for any perceived issue of multiple testing. As clinical trial sizes (\sqrt{n}) scale non-linearly with desired significance [5] - required numbers in the study explode for such very small type 1 errors! A smaller cheaper trial is not indicated.

A recent systematic review has highlighted that many pharmacogenomic trials being carried out are in fact small [13] - some with additional design failings too [14]. The above algebraic outline does not contain any special term for the use of a pharmacogenomic biomarker - whether the trial is pharmacogenomic or not, the force of logic still applies. A single trial to generate proof needs to be much larger than expected compared to the size of one of an equivalent pair of clinical trials - there is 'no free lunch'.

Note (*most usefully*): If the planned significance level changes (or the planned power is altered) for the second pharmacogenomic trial - that can be fed into the above approximate argument. Similarly the above algebra can also be used for the actual significance levels obtained and the *post hoc* power obtained versus the planned alternative for each and every pharmacogenomic trial to calculate more accurate support. Similarly, should alternatives change for the second replicate trial then recalculation of the *post hoc* power of that trial back to the original first trial's alternative can be made to ensure appropriate synthesis (or alternatively the original trial's evidence be recast to the new alternative). One can use this approach including 'failed' trials, more than 2 replicate trials etc. Any pharmacogenomic trial type can be used and folded into the same argument about beliefs. Of course, for final formal support over any particular hypothesis, the *actual data likelihood* ($p(D|H)$) values over the whole possible hypothesis space (i.e. not just restricted to these two didactic exclusive and exhaustive hypotheses) should be used for each and every trial.

These views are personal and self-financed. They should not be construed as representing in any way those of the University of Reading, Daiichi-Sankyo Development Ltd, nor the Royal Society, London.

References

1. Shah J (2004) Concept and Putative Application of Pharmacogenetics and Pharmacogenomics. Health Law Review 12: 3-10.
2. Shah RR (2005) Pharmacogenetics in drug regulation: promise, potential and pitfalls. Philos Trans R Soc Lond B Biol Sci 360: 1617-1638.
3. Gladding PA (2013) Clinical applications of pharmacogenetics: Present and near future. Cleveland Clinic Journal of Medicine 80: 477-482.
4. Ross S, Anand SS, Joseph P, Pare G (2012) Promises and challenges of pharmacogenetics: an overview of study design, methodological and statistical issues. JRSM Cardiovascular Disease 1: 1-13.
5. Machin D, Campbell MJ, Tan SB, Tan SH (2009) Sample Size Tables for Clinical Studies. (3rd edn), Wiley-Blackwell, United Kingdom.
6. Bowman CE, Ligensa T (2013) Appropriate, adequate and reliable - A view on statistical and regulatory evidence in clinical trials. Regulatory Rapporteur 10: 5-8.
7. Jeffreys H (1961) The Theory of Probability. Oxford University Press, New York.
8. Kass RE, Raftery AE (1995) Bayes Factors. Journal of the American Statistical Association 90: 773-795.
9. Evett IW, Jackson G, Lambert JA, McCrossan S (2000) The impact of the principles of evidence interpretation on the structure and content of statements. Sci Justice 40: 233-239.
10. Aitken C G G, Stoney D A (1991) The use of statistics in forensic science. (1st edn), Ellis Horwood Limited, England.
11. Good I J (1988) The interface between statistics and philosophy of science. Statist Sci 3: 386-397.
12. Committee for Proprietary Medicinal Products (2001) Points to Consider on Application with 1. Meta-analyses; 2. One Pivotal Study. The European Agency for the Evaluation of Medicinal Products.
13. Holmes MV, Shah T, Vickery C, Smeeth L, Hingorani AD, Casas JP (2009) Fulfilling the Promise of Personalized Medicine? Systematic Review and Field Synopsis of Pharmacogenetic Studies. PLoS ONE 4: e7960.
14. Stein MA, McGough JJ (2008) The Pharmacogenomic Era: Promise for Personalizing ADHD Therapy. Child Adolesc Psychiatr Clin N Am 17: 475-490.