

BIOEQUIVALENCE IMPLIES THERAPEUTIC EQUIVALENCE.

I. BIOSTATISTICAL APPROACH

FLORENTINA GHERGHICEANU¹, ROXANA SANDULOVICI², IRINA PRASACU^{3*},
VALENTINA ANUȚA³, CONSTANTIN MIRCIOIU³

¹PhD student, Faculty of Pharmacy, "Carol Davila" University of Medicine and Pharmacy, Bucharest, Romania

²Faculty of Pharmacy, "Titu Maiorescu" University, Bucharest, Romania

³Faculty of Pharmacy, "Carol Davila" University of Medicine and Pharmacy, Bucharest, Romania

*corresponding author: irinaprasacu@yahoo.com

Manuscript received: December 2015

Abstract

The axiom that "if two drugs achieve same plasma levels for the active substance, they have the same therapeutic effect" is accepted in all specific legislations and it is not object of dispute. So, that bioequivalence implies therapeutically equivalence. Following economic interest or following real concern regarding the effect, but unawareness about the biopharmaceutical and pharmacokinetic component of drug actions, about methodology of bioequivalence studies, a large number of papers appeared along time, alarming about possible risks associated with generic substitution for particular drugs or for particular diseases. The aim of the paper was to present rigorous mathematical arguments for establishing that bioequivalence implies therapeutic equivalence. The confidence intervals for the difference between a tested and a reference drug are simulated starting from the intra-subjects variability and the actual difference between pharmacokinetic parameters. Performing a concrete calculus starting from the confidence interval statistical method for testing bioequivalence, it is shown that the difference between two bioequivalent drugs is most likely all the time lower than 10%. It concerns the fear of clinicians, concerning the highly variable drugs, but the problem is false since, for clinicians, variability is, in fact, inter-subjects variability and error term in the bioequivalence testing formula depend on intra-subjects variability, which is much lower than inter-subjects variability, in all cases. Finally, is shown that estimation of the number of subjects required for a statistical significant proof that two bioequivalent drugs are not therapeutic equivalent, leads to numbers of hundreds of thousands. Theoretically and practically, bioequivalence (BE) implies therapeutic equivalence and there is no scientific argument or practical example that cancels this conclusion.

Rezumat

Axioma "două medicamente care realizează aceleași concentrații plasmatice au același efect terapeutic" este acceptată în toate legislațiile specifice și nu a fost niciodată un subiect de dispută. Ca urmare, dacă două medicamente sunt bioechivalente, ele vor fi și terapeutic echivalente. Urmare însă a unor interese economice sau a unor îngrijorări reale în ceea ce privește efectul medicamentelor generice dar și a unei necunoașteri a componentelor biofarmaceutice și farmacocinetice ale acțiunii medicamentelor și a metodologiei studiilor de bioechivalență, au apărut în timp o mulțime de lucrări trăgând un semnal de alarmă privind posibilele riscuri asociate cu substituția generică în cazul unor medicamente și boli particulare. Scopul lucrării a fost de a prezenta argumente matematice riguroase că bioechivalența implică echivalență terapeutică și nici un argument științific sau exemplu nu vine să contrazică această concluzie. Intervalele de încredere pentru diferențele între parametrii farmacocinetici ai medicamentului de referință și ai celui testat au fost simulate pornind de la date de variabilitate intra-subiecți și diferența între parametrii farmacocinetici ai formulărilor. Efectuarea unor calcule concrete pornind de la metoda intervalului de încredere pentru testarea statistică a bioechivalenței, se arată că diferența între două medicamente bioechivalente este, cu o foarte mare probabilitate, mai mică de 10%. În ceea ce privește teama clinicienilor în cazul medicamentelor cu variabilitate mare, problema este falsă, deoarece clinicienii se referă la variabilitatea intersubiecți, iar formula pentru testarea bioechivalenței se bazează pe variabilitatea intra-subiecți, care este practic întotdeauna mult mai mică decât intervariabilitatea. În final se arată că estimarea numărului de subiecți necesari a demonstra statistic semnificativ că două medicamente bioechivalente nu sunt terapeutic echivalente este foarte mare, de ordinul sutelor de mii. Teoretic și practic, bioechivalența implică echivalență terapeutică. Nu există nici un argument științific sau contra exemplu care să contrazică această concluzie.

Keywords: bioequivalence, therapeutic equivalence, confidence interval, biostatistics approach

Introduction

Axiom: If two drugs achieve same plasma levels for the active substance, they have the same therapeutic effect.

This axiom, like all axioms, is not required to and cannot be demonstrated. So, that it is accepted in all specific legislations and it is not object of dispute. If two drugs are realizing same plasma level profiles of active substances, are called

bioequivalent. So, that bioequivalent drugs are accepted as therapeutically equivalent.

Bioequivalence (BE) clinical trials (CT) as surrogate of safety and efficacy clinical trials, for generic, drugs represent one of the most significant scientific, economic, political and ethical contribution to clinical science and practice in the last half of the twentieth century, in direction of assurance a fundamental human right: drugs for all, right to access at drugs for poor people, for poor countries.

The idea came from America, where the Drug Price Competition and Patent Term Restoration Act of 1984 adopted by the Congress (Hatch-Waxman Amendments) tried to balance the interest of society to benefit from new, more active drugs developed by innovator companies on the one hand and to assure lower price for the generic version of these drugs after a reasonable time protection of patents for the recovery of the invested money.

Some twenty years later in a presentation made in the Senate, an FDA representative [9] concluded that "availability and access to generic drugs will save \$35 billion over 10 years".

The appearance of the generic drugs implies a diminishing of the profit of innovator companies, so, it is normal that these ones are trying, as much as possible, to induce the idea that generic drug is a more or less imperfect "copy" and consequently could have less efficacy and less safety.

On the other hand, a real complication appears as a consequence of biological variability. "Same plasma profiles" have to be understood as statistical "sameness" and understanding the mathematical foundations is not too accessible for the majority of medical specialists.

$$P\left(0.8 < \frac{\mu_T^{AUC_{0-\infty}}}{\mu_R^{AUC_{0-\infty}}} < 1.25\right) \geq 0.9 \quad \text{and} \quad P\left(0.8 < \frac{\mu_T^{C_{max}}}{\mu_R^{C_{max}}} < 1.25\right) \geq 0.9$$

Since C_{max} is more variable, an extension of the interval of acceptance can be considered. Also, difficulties appear in case of drugs with enterohepatic circuit, since C_{max} in such cases is not well defined, being replaced with a zone of maximum, with a "saw teeth" aspect [12].

The origin of the definition was the firm belief of the FDA clinicians that a difference in plasma levels less than 20 % is not clinically significant. Usually, data are logarithmated and since

$$CI_{0.90} = \left((\bar{T} - \bar{R}) - t(1 - \alpha/2, n_1 + n_2 - 2) \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right), \left((\bar{T} - \bar{R}) + t(1 - \alpha/2, n_1 + n_2 - 2) \hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

where $\bar{T}$ and $\bar{R}$ are the least square means of measured pharmacokinetic parameters (AUC or C_{max}),

$$\left((\bar{T} / \bar{R} - 1) - t(1 - \alpha/2, n_1 + n_2 - 2) \frac{\hat{\sigma}_d}{\bar{R}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right), \left((\bar{T} / \bar{R} - 1) + t(1 - \alpha/2, n_1 + n_2 - 2) \frac{\hat{\sigma}_d}{\bar{R}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

Also, apparently normal is to think that a drug product which was not tested concerning efficacy and safety, but only it in terms of pharmacokinetics, is in fact not completely characterized.

Following economic interests or real concern regarding the effect but unawareness of the biopharmaceutical and pharmacokinetic component of the drug actions, about the methodology of bioequivalence studies, a lot of papers were published during time, alarming about the possible risks associated with generic substitution for particular drugs or for particular diseases. For example only for epilepsy, a recent paper [1] cited 81 papers discussing this subject.

So that, in spite of the law clear assertions, in clinical practice there is a significant resistance to its application and it is necessary to underline with more strong arguments, that all, numerous doubts have not a real support. In this context, the aim of this paper was to present rigorous mathematical arguments that the bioequivalence implies therapeutic equivalence.

Statistical evaluation of bioequivalence

In spite of the fact that a huge number of people are confronted with bioequivalence problems, and many of them speak about bioequivalence or, much more, perform bioequivalence studies, the understanding of the quantitative statistical testing of BE and its implications is rather incomplete. A tested drug T is bioequivalent with a reference drug R if the 90 % confidence interval (CI) for ratios of means μ of main pharmacokinetic parameters - area under curve (AUC) and maximum concentrations C_{max} are included in the 0.8 - 1.25 interval.

$$\ln \mu_T^{AUC_{0-\infty}} - \ln \mu_R^{AUC_{0-\infty}} = \ln \frac{\mu_T^{AUC_{0-\infty}}}{\mu_R^{AUC_{0-\infty}}}$$

the rule concerning a difference less than 20 % was transformed in the rule concerning ratios. For the simplicity of calculus, in the following will be considered that data are normally distributed and will consider the difference of direct data, without logarithmic transformation.

Confidence interval calculated based on the hypothesis that data are normally distributed is:

If we divide both limits with $\bar{R}$, the relationship becomes:

$\hat{\sigma}_d$ is the variance of the mean difference between T and R, estimated from an Analysis of Variance (ANOVA) calculus

$$\hat{\sigma}_d^2 = \frac{MSE}{2} = \frac{\hat{\sigma}_e^2}{2}$$

(MSE = Mean Square Error in ANOVA), and $t(0.95, n_1 + n_2 - 2)$ is the $1 - \alpha/2$ quantile of Student distribution with $n_1 + n_2 - 2$ degrees of freedom, α the risk of patient, the risk that non-bioequivalent products could be found as bioequivalent and is taken in all cases 10 %, n_1 and n_2 are the numbers of subjects in the two sequences (experiment is considered cross-over with two periods and two sequences).

Almost all concerns appeared mainly from misunderstanding how the constants 20% and risk are influencing final decision concerning bioequivalence and this problem it is treated in the next chapter.

The inclusion of CI between acceptance limits implies that differences in mean values of the main PK parameters are much less than 20 %.

Inclusion of the CI in the acceptance limits as quantitative condition for bioequivalence is presented in Figure 1.

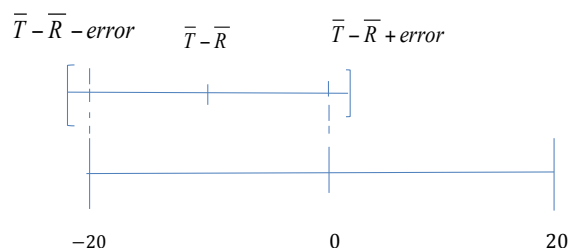


Figure 1.

Inclusion of the CI in the acceptance limits as quantitative condition for bioequivalence

Pointwise estimation of the difference $\bar{T} - \bar{R}$, the centre of confidence interval has to be very closed to zero. Otherwise, the interval will overflow on limits. This can be easier understood, for example, after a calculus of the confidence interval for a “model drug” at the frontier, between moderate and high variability: 30 % coefficient of variance.

The number of subjects included in the bioequivalence clinical trials is determined starting from statistics, economic and ethical reasons, being usual in the 12 – 50 interval. So, that let us consider a 24 subjects study.

Supposing the coefficient of variation (CV) is 30 %, and calculating the error

$$CV = 100 * \frac{\sigma_e}{\mu_R} = 100 * \frac{\sqrt{2\hat{\sigma}_d^2}}{\mu_R} = 100 * \frac{\sqrt{2}\hat{\sigma}_d}{\mu_R} = 30$$

and, if we take,

$\mu_R = 100$, it results in $\hat{\sigma}_d = 0.3 / \sqrt{2}$ and

$$error = t \hat{\sigma}_d \sqrt{\frac{2}{n}} = 1.68 * 0.3 / \sqrt{2} * \sqrt{\frac{2}{24}} = 0.103$$

If for example $\bar{T} - \bar{R} = -10\% * \bar{R}$, i.e products differ within 10 %, the confidence interval will be CI = [-20.3% + 0.3 %] and products are not bioequivalent. If the difference is in favour of reference drug $\bar{T} - \bar{R} = +10\% * \bar{R}$, the confidence interval will be [-0.7 %, + 20.3 %] and products are again non-bioequivalent.

So that, if drugs are bioequivalent, a difference of more than 10 % is theoretically highly unlikely. As will be shown further, practice proved that usually differences are lower than 7%.

FDA retrospective analysis shown that difference is less than 10 %

A retrospective analysis compared the generic and innovatory bioequivalence measures, from 2070 single-dose clinical bioequivalence studies of orally administered generic drug products approved by the Food and Drug Administration (FDA) from 1996 to 2007 (12 years) [3].

The mean $\pm$ SD of the geometric mean ratios from the 2070 studies was 1.00 ± 0.06 for C(max) and 1.00 ± 0.04 for AUC. The average difference in C(max) and AUC between generic and innovatory products was 4.35% and 3.56%, respectively. In addition, in nearly 98% of the bioequivalence studies conducted during this period, the generic product AUC differed from that of the innovatory product by less than 10%.

Highly variable drugs. The false problem.

Although it was never presented in a clinical study, proving that two bioequivalent drugs are not therapeutic equivalent, the fear of clinicians exists and, in practically all cases, is connected with highly variable drugs. The problem is false since, for clinicians, variability is in fact inter-subjects variability. In testing bioequivalence using cross-

over design, in the above presented formula, $\hat{\sigma}_d$ is an estimation of the intra-subjects variability, since every subject is its own control. Difference “d” is the difference between plasma levels of the active substance in the same subject after administration of reference or tested drug. Highly variable drugs for clinicians are not necessary highly variable drugs for the bioequivalence studies. High inter-subjects variability doesn’t imply automatically high intra-subjects variability.

Data of 113 human bioequivalence (BE) studies performed on immediate release (IR) formulations of 74 active pharmaceutical ingredients (APIs)

showed that no significant correlation was observed between intra-variability and inter-variability of all drugs [10].

Highly within subject variable drugs and Narrow Therapeutic Index (NTI) drugs. Real problem.

The “Anti-generic papers” worry about NTI and variability, but in fact ignore the within subjects variability. Fortunately, NTI drugs have low within subjects variability. For example, warfarin, levothyroxine, carbamazepine and phenytoin generic drugs approved by FDA in the period 1996 – 2008 presented a mean CV% between 5.7% and 9% [4].

Intra-subject variability $\hat{\sigma}_d$ is, in fact, the critical factor in determining the error term in the confidence interval formula. Intra-variability was found, for example, to increase in association with a decrease in drug permeability, for Biopharmaceutical Classification System (BCS) class 2 drugs, which have a limited solubility. Intra-subject variability was high for endogenous drugs [5,6], for drugs with significant first pass metabolism [7] for drugs metabolized by CYP3A4, while inter-subject variability was high for drugs metabolized by CYP2D6 [10].

Unfortunately, the real problem of the high variability in dissolution of drugs and precipitation in intestine is neglected in design of clinical trials, this information usually missing completely from Investigator's Brochures.

A calculus similar with what was presented above for a drug with CV = 60%, gave an error equal to 0.206 and even in case when products are identical and it obtains $\bar{T} - \bar{R} = 0$, test for bioequivalence fails. A solution for proving bioequivalence is to increase the number of subjects, to a maximum acceptable $n = 50$. The error will become 0.143. The maximum delta would become less than 5.6%.

In fact, when intra-subjects variability is high, BE studies fail. For example a statistics involving 1200 BE studies in FDA showed that, when CV was less than 10%, the percent of failed studies was less than 6 % and for drugs with CV > 30%, majority of studies (62%) failed [11, 12]. Consequently, it is expected that NTI drugs are not intra-subjects highly variable drugs (HVD).

For rare cases of drugs which are really NTI and highly variable, FDA conceived special Survey on Quality and Standards, which include product design and manufacturing, drug assay, content uniformity, dissolution, stability, more narrow acceptance intervals, strict control of variability of both reference and tested drug, etc. [4].

Why bioequivalence CT instead of clinical efficacy CT? Non-ethical, too large, foolish Clinical Trial: testing the therapeutically equivalence of two bioequivalent drugs.

Since many clinicians consider that testing the bioequivalence have to be completed with clinical trials based on the effect, let us estimate the number of patients required for obtaining a statistical significant results in comparing two bioequivalent drugs. We will consider as example two bio-equivalent formulations of aspirin for cardiologic use. As end point, we will consider the overall survival analysis at one year post myocardial infarct. Clinical trials [8] put in evidence that the proportion of death, is in the first year, 10% and the treatment with aspirin lead to a 20% reduction, i.e. to some 8%.

Let us accept that the difference between bioavailability of the two formulations A1 and A2 is: 20%.

The effect of A1 will be 20% and the effect of A2 will be 80% of 20%, i.e. 16% .

$$10\% \xrightarrow{-16\%} 8,4\%$$

The difference between the effects of the formulations will be

$$\Delta = p_1 - p_2 = 0,020 - 0,016 = 0,004$$

$$p_2 = 0,084 \Rightarrow q_2 = 1 - p_2 = 0,916$$

The pooled estimation of the death probability in the total population of patients, in case of equal sizes treatment groups $n_1 = n_2 = n$ will be

$$\hat{p} = \frac{p_1 + p_2}{2} = 0,082 \Rightarrow \hat{q} = 1 - \hat{p} = 0,918$$

The usual formula for calculating the number of subjects n take into consideration the risk

$$n = \frac{\left[z_{\frac{\alpha}{2}} \sqrt{2\hat{p}(1-\hat{p})} + z_{\beta} \sqrt{p_1(1-p_1) + p_2(1-p_2)} \right]^2}{(p_1 - p_2)^2}$$

where, $z_{\frac{\alpha}{2}}$ and z_{β} are the quantiles associated to

normal standard distribution $P(Z < z_{\frac{\alpha}{2}}) = z_{\frac{\alpha}{2}}$ and $P(Z < z_{\beta}) = z_{\beta}$

Usual notations are:

- the probability of type I error:

$$\alpha = \Pr(\text{reject } H_0 / H_0 \text{ true})$$

- the probability of type II error

$$\beta = \Pr(\text{accept } H_0 / H_0 \text{ false})$$

- $\pi = 1 - \beta$, called “power of the test” is the probability of rejecting a false hypothesis: and β consequently $\pi = 1 - \beta$ are at the choice of producer being its risk.

$$\pi = \Pr(\text{reject } H_0 / H_0 \text{ false})$$

The decrease of the number of subjects can be made only by increasing the risk of producer. Risk of patient is usually limited by guidances. In bioequivalence studies is considered $\alpha = 0.05$.

It is to note that the critical parameter is the "clinical significance difference" which is tested $\Delta = p_1 - p_2$.

Replacing our data in the formula, it leads to:

$$n \geq \left(\frac{1.96\sqrt{2 \cdot 0.082 \cdot 0.918} + 1.28\sqrt{0.08 \cdot 0.92 + 0.084 \cdot 0.916}}{0.004} \right)^2 = 98776$$

So, such a study was never and will be never performed.

As was presented by Leslie Benet, "no prospective study has ever found that an FDA approved generic product does not show the same clinical efficacy and safety as the innovator product, even when special populations (e.g., elderly, women, severely sick patients) are studied" [2].

Conclusions

The fear of clinicians concerning the possibility that two BE drugs could be not TE is based essentially on the 20% rule. But this value is only the border between significant and not significant differences in plasma levels of two drugs. Both theoretical and practical differences between BE drugs are under 10%.

Since proving the non-TE of two BE drugs would require hundreds of thousands of patients, no study has ever been performed to find that a generic product does not show the same clinical efficacy and safety as the innovatory product.

The fear concerning HVD is not justified. Real problem regards NTI and intra-HVD which is connected mainly with usually ignored biopharmaceutical problems. The solution in these cases takes into consideration all factors of variability starting from manufacturing to biopharmaceutical and pharmacokinetics critical factors.

Acknowledgement

This work was supported by a grant of the Romanian National Authority for Scientific Research, CNDI-UEFISCDI, project PN-II-PT-PCCA-2013-4-2071, contract no. 139/2014.

References

1. Atif M., Azeem M., Rehan Sarwar M., Potential problems and recommendations regarding substitution of generic antiepileptic drugs: a systematic review of literature, Springer Plus 2016; 5:182.
2. Concerns of Healthcare Practitioners and Patients, in AAPS/FIP/RA/SSFR International Regulatory Workshop on Bioequivalence and Dissolution, Bucharest, December 4-5, 2003.
3. Davit B.M., Nwakama P.E., Buehler G.J., Conner D.P., Haidar S.H., Patel D.T., Yang Y., Yu L.X., Woodcock J., Comparing generic and innovator drugs: a review of 12 years of bioequivalence data from the United States Food and Drug Administration. *Ann Pharmacother.* 2009 Oct; 43(10): 1583-97.
4. Lawrence X.Y., Quality and Bioequivalence Standards for Narrow Therapeutic Index Drugs, GPhA 2011 Fall Technical Workshop, www.fda.gov/.../Drugs/.../HowDrugsareDeveloped.
5. Mircioiu C., Perju A., Neagu A., Pharmacokinetics of progesterone in postmenopausal women: 1. Pharmacokinetics following intravaginal administration *European journal of drug metabolism and pharmacokinetics*, 1998; 23(3), 391-396.
6. Mircioiu C., Perju A., Griu E., Pharmacokinetics of progesterone in postmenopausal women - 2. Pharmacokinetics following percutaneous administration *European journal of drug metabolism and pharmacokinetics*, 1998; 23(3), 397-402.
7. Mircioiu C., Ionica G., Danilceac A., Pharmacokinetic and mathematical outliers for drugs with active metabolites. Note I. Model independent analyses for pentoxifylline, *Farmacia* 2010; 58(3), 264-278
8. Salim Y., Overview of Result of Randomized Clinical Trials in Heart Disease. II. Unstable Angina, Heart Failure, Primary Prevention with Aspirin and Risk Factor Modification, *JAMA*, 1988; 260(15), 2259-2263.
9. Statement of Daniel E. Troy Chief Counsel U.S. Food and Drug Administration before the Senate Committee on the Judiciary august 1, 2003.
10. Sugihara M., Takeuchi S., Sugita M., Higaki K., Kataoka M., Yamashita S., Analysis of Intra- and Intersubject Variability in Oral Drug Absorption in Human Bioequivalence Studies of 113 Generic Products. *Mol Pharm.* Dec 7, 2015; 12(12): 4405-13.
11. Tanguay M., When Will a Drug Formulation Pass or Fail the Bioequivalence Criteria? Experience from 1200 Studies AAPS Abstract, November 2002.
12. Tvrdonova M., Dedik L., Mircioiu C., Physiologically Motivated Time-Delay Model to Account for Mechanisms Underlying Enterohepatic Circulation of Piroxicam in Human Beings *Basic & Clinical Pharmacology & Toxicology*, 104(1), 35-42.