Generics, chemisimilars and biosimilars: is clinical testing fit for purpose?

John B. Warren

Medicines Assessment Ltd, London, UK

Correspondence

Dr John B. Warren MD FRCP, Medicines Assessment Ltd, 196 Rotherhithe Street, London SE16 7RB, UK.
Tel.: +44 7789 825 680
E-mail: info@medicinesassessment.com

Keywords

biomarkers, EMA, FDA, Hatch-Waxman, inhalers, pharmacokinetics

Received

11 March 2012

Accepted

3 May 2012

Accepted Article Published Online

11 May 2012

The effectiveness and safety of generic drugs are backed by sound physicochemical control and regulatory bioequivalence acceptance criteria. Statistical testing of bioequivalence, comparing the pharmacokinetic profiles of the test and reference products, was made possible by modern drug assays. When the pharmacokinetic profile correlates with the dose, such comparisons show assay sensitivity and readily detect differences in dose. For large biological molecules, different manufactured batches cannot be validated using pharmacokinetic data alone. For these biosimilars, there is a three-stage assessment of pharmaceutical quality, laboratory testing and clinical data. This approach has also been applied to certain chemical products, termed 'chemisimilars', which have variable or complex synthesis of the active substance, or complex formulation, or a complex delivery device. Although there may be no detectable difference between the test and reference on clinical testing, many of the outcome measures are insensitive to even large differences in dose. For testing to be fit for purpose it should distinguish important dose differences, but many clinical tests of chemisimilars and biosimilars do not. As pharmacokinetic and pharmacodynamic technology advances, the trend of replacing dose-insensitive clinical trial data with equivalence tests that show assay sensitivity can be expected to continue.

Introduction

The US Food and Drug Administration (FDA) pioneered the standardization of clinical testing requirements for copies of original drugs. Before 1984 the FDA required all applicants for generics to include a clinical dossier showing similar efficacy and safety data to that provided by the originator. This was expensive and also mostly failed to detect differences in dose. It gave little reassurance of equivalence, other than showing an efficacy difference from placebo and no major additional safety issues with the generic product.

This approach changed when the safety and efficacy of drugs were shown to depend upon their tissue concentration. Advances in the 1970s and 1980s, such as chromatography and radioimmunoassay, allowed more accurate quantification of medicines in blood samples. This was recognized by the US Drug Price Competition and Patent Term Restoration Act of 1984, Public Law 98-417, also known as the Hatch-Waxman Act, that allowed the efficacy and safety dossiers for generics to be replaced by evidence of equivalent pharmacokinetic exposure.

Showing equivalence between test and reference is not restricted to copies of innovator drugs that have come off

patent. It is also required when testing successive batches of an innovator during development or postmarketing, for comparing formulation changes or fixed dose combinations, or when new doses are developed [1].

The complexity of chemisimilar and biosimilar products has limited the replacement of unwieldy large-scale clinical trial data with simpler, single-dose, pharmacokinetic and pharmacodynamic comparisons between test and reference. This limitation should be recognized by defining the sensitivity of any test used to detect differences in dose.

Manufacturing quality

In pharmaceutical development, 'quality' is a term usually reserved for manufacturing and formulation, excluding the preclinical and clinical development. Quality control in mass production refers to the quality of all functions of manufacture. Quality assurance is the comparison of measurement to a standard, with preventive feedback, to ensure that a product meets customers' expectations.

Whether handcrafted objects are fit for purpose, from the Stone Age onwards, has depended on the skill of the

artisan. Now nearly all the objects in our daily lives are mass produced, and it is the skill of the manufacturing process, rather than the individual, that determines reliability. Quality assurance is the bedfellow of mass production. Whether making pins, nuclear missiles or pharmaceuticals, or designing a clinical trial, it requires set specification limits. Attempts to increase economic efficiency have spawned managerial and business processes that continue to improve manufactured quality.

The science of reliability engineering has been slow to transfer to the regulation of medicines. Historically, practitioners of the art of medicine have judged pharmaceutical quality to be adequate if any detectable pharmaceutical difference is insufficient to affect patient outcome. This approach contrasts with most manufacturing practice; for example, variability in aircraft component specification would not be accepted simply because deviation from a specification was not detectable by the pilot during a test flight. European and US pharmaceutical regulatory guidance has responded to the need for reliability specifications for simple chemical generics, but more complex products, such as chemisimilars and biosimilars, continue to pose a challenge.

Reliability engineering, a division of systems engineering, started in the 1930s with Walter Shewhart, a statistician who devised a quality control scheme to reduce the failure rate of Bell telephone equipment [2]. Statistics, probability theory and reliability theory were used to measure and predict reliability. William Edwards Deeming, a disciple of Shewhart, is considered a hero of the postwar success of the Japanese economy [3]. Interest in quality control was spurred by intercontinental ballistic nuclear missile design in 1962, with Bell Laboratories' Fault Tree Analysis [4], later adopted by Boeing in 1966 for civil aircraft. Motorola's 1986 Six Sigma quality improvement process for electronics became a global standard in 1990 [5]. It refers to the mean specification dimension that is at least six standard deviations from the specification limit, giving a failure rate of 3.4 per million, or less. There is less stringency in the pharmaceutical industry, where tolerance margins for a logarithmic dose response are not as tight as for electronics.

The message from reliability engineering for pharmaceuticals is that setting statistical acceptance/rejection criteria has a major impact on quality. Boeing 747 planes transported 3.5 billion passengers, a distance equivalent to 100 000 return trips to the moon, with a fatality rate of less than 1 per million passengers despite each aircraft being composed of 6 million parts. Yet most people find flying more unnerving than swallowing a pill. It is obvious that better planes have fewer crashes, but unfortunately a sub-standard pharmaceutical can be difficult to detect [6].

Although pharmaceutical specification limits are defined by the best available technology, the FDA Orange Book accepts less sensitive assays of clinical therapeutic equivalence: 'Drug products are considered to be thera-

peutic equivalents only if they are pharmaceutical equivalents and if they can be expected to have the same clinical effect and safety profile when administered to patients under the conditions specified in the labelling' [7]. Here, showing no difference can be a blunt instrument of quality assurance, particularly if the efficacy of the innovator product was only narrowly distinguishable from placebo in the approved original clinical database.

Defining a standard

Standardizing measures of length, weight, volume and time has amused and taunted great intellects, whose interest was often initiated by commerce. Think of Archimedes' 'eureka!' moment on the discovery of volume displacement to verify the content of a gold crown, or John Harrison's clocks guiding travel through longitude determination [8].

For pharmaceuticals, many measurements have to be accurate and reproducible. Active ingredients are manufactured to a standard and quantified with assays. Contaminants, excipients, formulations and shelf-life are specified. Nevertheless, biosimilars and chemisimilars remain a challenge for manufacturing and regulation. When inhalers deliver a range of fine particle size with large variations in drug dose that depend on the technique of the subject, or biosimilars contain large molecules that are hard to characterize fully, then care is required when defining acceptance/rejection criteria.

Simple generics

The testing of simple chemical generics of innovator drugs is generally fit for purpose, because pharmacokinetic assays distinguish important differences in dose. For most generics there are sensitive, precise and accurate assays to compare profiles in terms of the maximal plasma concentration and the area under the concentration-time curve set to limits of a 90% confidence interval of 80.00–125.00% [9–11]. Apart from the qualification of impurities, there is usually no need for preclinical testing and clinical trials. The '80–125' specification originates from a 1986 FDA 3 day public hearing, at which bioequivalence criteria were discussed, with 50 speakers and 800 participants [12], and most generic applications contain data well within this acceptance range [13].

When pharmacokinetics are linear there is reasonable confidence that a generic copy of a 40 mg dose strength is within a range of 32–50 mg. This is acceptable, for drugs other than those with a narrow therapeutic index, with no need for the Six Sigma tight control of the electronics industry. The setting of precise cut offs for acceptance criteria is reassuringly similar to other areas of manufacturing and arose from the adoption of statistical principles to

pharmaceutical testing [14, 15]. Although many clinicians find it hard to reject a negligible difference, say a lower limit of a confidence interval of 79.99% instead of 80.00%, not to do so permits drift, where increasing deviations from specifications have to be justified.

The setting of acceptance limits for bioequivalence is relatively new compared with other manufacturing, but the acceptance of pharmacokinetic bioequivalence testing was a breakthrough in terms of reducing costs and improving quality. The centralization of the FDA, compared with the diversity of EU regulation, favoured the standardization of bioequivalence testing in the USA. The EU eventually adopted a uniform approach to chemical generics with the 2009 release of EU bioequivalence guidance [10].

Chemisimilars: inhalers as exemplars

When proving batch-to-batch consistency, or equivalence of generics, some innovator synthetic chemical drugs are complex. This arises with controlled release formulations, variable or complex synthesis, or chemicals where the composition of subcomponents are not well defined. The parallels with biosimilars are obvious, and sometimes hybrid applications are required, where comparative data are combined with stand-alone efficacy and safety clinical data. Many such products are alternatives to the innovator products without being substitutable. Chemisimilar is a useful term sometimes used by EU regulators to describe these complex chemical generics [16]. These include semi-synthetic antibiotics, some liposomal formulations and some slow-release transdermal patches. To discuss the manufacturing and regulatory challenges of chemisimilars, inhaler products are used here as an example. Similarity of inhalers is a challenge that became particularly important with the opening in 1989 of the Montreal Protocol on substances that deplete the ozone layer (a protocol to the Vienna Convention for the protection of the ozone layer). This initiated the phasing out of halogenated aerosols in these products; the replacements were required to have a different formulation yet retain the risk-benefit profile of the originals.

In 1996 the only EU guidance on inhaler equivalence was a small section in the brief EU Notes for Guidance on the clinical requirements for locally applied locally acting products containing known constituents, CPMP/EWP/239/95 [17]. The short section on inhalers recommended that a comparison should include a three-arm trial with placebo, but there was no requirement for a dose response, nor any definition of acceptance/rejection criteria. By 2004 the EU Points to Consider on the requirements for clinical documentation for orally inhaled products (OIP), CPMP/EWP/4151/00 [18], stated that therapeutic equivalence must be established and, for the first time,

accepted that *in vitro* testing might be acceptable when products were identical.

In 2009 substantial EU recommendations were published on inhaler therapeutic equivalence, with the new guidance replacing EMEA/CHMP/EWP/48501/2008 Appendix 1 and CPMP/EWP/4151/00 of 2007 [19]. This recommended a three-stage approach of pharmaceutical quality, lung distribution and clinical testing. The guidance acknowledges that if two inhalers are identical in all aspects of formulation and device, then pharmaceutical quality bench testing alone might suffice. This circumstance is rare, because so many innovator inhalers have multiple device patents and are thus difficult to copy. For pulmonary deposition, acceptance criteria for key parameters are defined in terms of systemic kinetics, though scintigraphy is less amenable to numerical quantification, making objective and reproducible assessment a challenge.

If quality and pulmonary deposition evidence of equivalence are insufficient, then clinical testing is required. It is possible that an inhaler that is not equivalent on quality and deposition criteria might be approved on less strict clinical criteria. Previously, clinical testing could be limited to one dose strength in dose-insensitive studies. Now at least two non-zero dose strengths of the test compared with two doses of the reference are requested, and this should be within the steep part of the dose-response curve.

The clinical requirements of the latest EMA inhaler guidance are multiple, and few have defined acceptance criteria. Standards that are set using statistical probability can be complicated by multiplicity; requiring multiple tests to be within a 90 or 95% confidence interval increases the probability of a finding arising from chance. The placebo effects of inhalers can be similar to the effects of active medication in patients with asthma [20]. In the past, similar effects on lung function and asthma symptoms have led to approval [21] even when pharmacokinetic analyses readily show a difference between two inhalers [22]. There is considerable literature to support the sensitivity of pharmacokinetic assays and scintigraphy to detect inhaler differences for changes in dose, flows, breaths, delivered dose, particle size, excipients, patients and healthy volunteers. Conversely, no inhaler study has shown a difference between inhalers in measures of lung function, or symptoms, without showing a difference detectable by pharmacokinetic parameters.

Highly trained volunteers in well-controlled laboratory conditions are required, because much depends on inhaler technique. Such variance has to be minimized in order to detect any difference between inhalers. Intrasubject variation between breaths is in the region of 8–10% for highly trained volunteers, but 32–52% for those given initial training only [23–28]. With adequate training, differences in fine particle dose for the same product are detectable with only a few volunteers [29].

Even with training, inhaler bioequivalence can be difficult to prove. A recent approach used a spacer device and slow, deep inspirations to minimize variability when bioequivalence was not proved by conventional dosing [30].

The idea persists that differences observed with sensitive tests, such as pharmacokinetic analysis, may be considered clinically unimportant [31], although clinical relevance is not definable. Many patients are not compliant with long-term medication, and for them variations in manufacture are irrelevant. Even for established therapies it can be tough to show clinical differences from placebo. In the TORCH trial of some 6000 chronic obstructive pulmonary disease patients studied for 3 years, inhaled fluticasone was associated with increased mortality and pneumonia; when combined with salmeterol, it decreased exacerbations [32]. The trial had to be large to detect these differences from placebo; a much bigger trial would be needed to show differences between doses of the same drug.

The problems with inhaler equivalence are acknowledged by the FDA. Their guidance notes: 'Currently, bioequivalence for oral inhalation products is demonstrated through *in vitro* testing for device performance, pharmacodynamic studies of lung function for local delivery, and pharmacokinetic studies for systemic exposure. Due to the difficulty in demonstrating bioequivalence by passing all of these tests, as well as other factors, FDA receives few applications for these kinds of products, even though many of the older MDI products are on the market without patent or exclusivity protection. FDA has identified many of the scientific challenges that need to be addressed to develop generic versions of these products.' [33]. This lack of clear advice leaves applicants with uncertainty, and in many ways clinical relevance tests for generic inhalers are not fit for purpose. The lack of unequivocal pharmacodynamic tests of equivalence, for example for inhaled steroids, continues to provoke debate [34, 35].

Biosimilars

Biosimilar medicinal products are large, complex molecules that deliver activity to a binding site. A test protein may have the same primary, secondary and tertiary structure as a reference protein. Yet activity at, or delivery to, the binding site may be affected by other chemical differences, such as glycosylation, nitrosylation, phosphorylation, deamination, oxidation or PEGylation. Different binding sites may be responsible for efficacy, adverse events, metabolism and excretion pathways. This complexity is a challenge for biosimilar regulation. Small differences in manufacture or formulation can have serious clinical consequences, as documented with some biosimilar epoetins and insulins [36]. Although such experiences mitigate against acceptance of any alteration in the amino acid

sequence of a biosimilar, minor changes might not alter activity. To try to accommodate this, FDA recent draft biosimilar guidance requires the same identity of primary amino acid sequence, but allows minor differences in the extent of N-terminal or C-terminal processing [37]. This makes a small exception to the FDA ruling on generics that any new covalent bond represents a new chemical entity. The issue is a hot debate within EU regulation, because it touches on the fundamental definition of a new drug, with consequences for marketing authorization exclusivity, as well as overlapping with definitions used in patent protection.

The European Medicines Agency (EMA) published a general guideline on biosimilars in 2005 [38] listing relevant sources of information and advising that product-specific guidance would follow. The EMA recommended a case-by-case approach and invited requests for scientific advice. The first approval of a biosimilar was the European authorization of Omnitrope, a growth hormone, in 2006 [39]. The clinical biosimilar comparison of growth hormones used height in children as an end-point, an accurate measure with assay sensitivity. Untreated, or treated, growth charts of growth hormone deficiency are available. After 9 months, treatment with Omnitrope and Genotropin gave similar increases in height and speed of growth, equivalent to 10.7 cm per year with both medicines. The product replaces a hormone deficiency, and the adequacy of this replacement is supported by pharmacokinetic data.

The EMA subsequently developed specific guidelines for individual products, taking into account the differences in molecular characterization and indications of the range of licensed biological products. These guidelines adopt the logical approach of pharmaceutical quality, laboratory testing and clinical trials. Although a lack of significant difference between test and reference is reassuring, in some instances the assay sensitivity of such clinical testing is not established. Some innovator biological products have shown only a small difference from placebo in trials, sometimes without publicly available dose-response data. More recent EMA biosimilar guidelines have responded to this issue. The clinical assessment has moved away from establishing no significant difference in end-points that refer to the indication, towards favouring end-points that may better detect potential differences between test and reference.

The EMA 2011 Guideline on monoclonal antibodies [40] represents a major change in this respect. It requests that pharmacokinetic studies are performed in a homogeneous population to reduce variability, sample size and simplify interpretation; single doses in healthy volunteers are acceptable, if justified. These studies may be with a different population from the efficacy studies, the choice being determined by the population with the greatest sensitivity to pharmacokinetic assessment. Highly sensitive pharmacodynamic studies are requested to test for comparability in a clinically relevant manner. If this is not pos-

sible, scientifically appropriately sensitive human models should be used, not necessarily within the licensed indication [41].

The balance between sensitive testing and clinical tests that show no difference in the licensed indication continues to be problematic for biosimilar regulation. The recent guideline on biosimilar interferon β (IFN- β) for multiple sclerosis (MS) refers to dose-sensitive models [39]. Healthy volunteers are considered acceptable for pharmacokinetic studies, with the dose chosen from a sensitive part of the dose–concentration curve. The difficulty of detecting the product in serum is acknowledged. Possible pharmacodynamic markers are mentioned, although there is no validated biological marker related to the mechanism of action. The guideline requests a parallel design, adequately powered, clinical efficacy trial using relapse rate as the primary end-point, although offering magnetic resonance imaging (MRI) as a potential alternative. A homogeneous population of patients, whose disease is the most sensitive to detect differences, is recommended. Assay sensitivity is requested in reference to differentiation from placebo, but not in reference to dose, with clinical outcome data of not less than 12 months duration. No formal equivalence clinical outcome test is required, though clinical outcome should show the ‘same trend’ as the MRI data.

The challenge for a biosimilar IFN- β programme is considerable [42]. Ethically, a prolonged placebo arm in a trial of MS is not justified. The MRI changes have not been numerically quantified sufficiently to set acceptance and rejection limits. The public assessment report (EPAR) for Rebif [43] states that in a trial of 560 patients with relapsing MS the two doses tested were indistinguishable after 2 years of treatment. After 4 years the relapse rate was reduced by 22 and 29% by low-dose and high-dose Rebif, respectively, when compared with placebo. In progressive MS, Rebif had no significant effect on the progression to disability. Effects in the population age 12–18 years are uncertain. With such small changes in large-scale, prolonged trials, our current state of knowledge is insufficient to give a confidence limit on the clinical activity of a biosimilar IFN- β .

Attempts to write guidelines in the USA by conflicting interests caused political frustration. International manufacturing requirements are set in The International Conference on Harmonisation (ICH) guidance Q5E *Comparability of Biotechnological/Biological Products Subject to Changes in Their Manufacturing Process*. The FDA recently described a route for an abbreviated pathway for the approval of biosimilars to eliminate unnecessary testing in animals and man [44]. They recommend a stepwise approach that takes into account the totality of the evidence, making it difficult to preset acceptance/rejection criteria [45]. The guidance stresses the importance of functional assays related to the mechanism of action for both animal and clinical data. Clinical end-points are requested to be relevant to detect meaningful differences in safety or efficacy. It is acknowl-

edged that certain pharmacodynamic end-points, such as the International Normalized Ratio for anticoagulants, can be more sensitive to detect differences between test and reference than clinical outcome end-points. There is no mention in this guidance of sensitivity of tests to detect dose differences, although it is requested that doses for test/reference comparison are on the steep part of the dose–response curve.

Testing in children

Most clinical studies of equivalence are experiments designed to maximize the potential of detecting formulation differences between products. Studies in children are particularly unsuited to this task. There is the ethical dilemma of asking children to participate in trials for drugs which are already available, merely to provide an alternative product. Repeated venepuncture can be difficult, traumatic and hard to justify. Urine collections are usually incomplete. Some testing is justifiable, for example taste testing, whether a product can be swallowed, or the age limit at which a child can successfully handle a delivery advice. The EU EMA Bioequivalence Guideline discourages paediatric studies [10]. EMA Biosimilar Guidance [40] takes a similar stance and recommends that: ‘Clinical studies in special populations like the paediatric population or the elderly are normally not required since the overall objective of the development programme is to establish biosimilarity, and therefore the selection of the primary patient population is driven by the need for homogeneity and sensitivity.’ In contrast, the EMA Oral Inhaled Product Guideline requests limited paediatric data, on the grounds that inhaled drug delivery differs in children. Although true, most formulation differences are readily detectable in healthy volunteers, and the poor quality of paediatric trials limits the usefulness of these data. There is the additional problem of how to deal with a dossier where the adult data are equivalent, but the paediatric data are not.

Conclusion

Originally, all generic drug applicants were required to collect extensive efficacy and safety clinical data before marketing. The Hatch–Waxman Act of 1984 changed US legislation to allow the use of pharmacokinetic data alone to bridge generic drugs to the clinical dossiers of innovator drugs. This greatly reduced the costs of generic drug development, because most products can be tested in single-dose, crossover, pharmacokinetic trials in normal volunteers. This cost saving had the added advantage that pharmacokinetic studies with few missing data points, in contrast to large, long-term clinical studies, could readily detect differences between doses. Accurate pharmacokinetic assays for drugs with linear kinetics are widely

accepted as fit-for-purpose tests of bioequivalence for many generic drugs and are sensitive to differences in dose.

In contrast, the assay sensitivity of long-term, expensive, clinical safety and efficacy comparative studies is often blunted by missing data. Many innovator drugs show small differences from placebo, making it difficult to distinguish between doses, and some biological products have no published dose response.

The trend to compare formulations for equivalence with tests that show assay sensitivity has continued to gain popularity, aided by improvements in drug analysis, particularly the advances in mass spectrometry. For regulators responsible for drafting guidelines for chemisimilars and biosimilars, a key question is whether any proposed test has the capability to distinguish between even a twofold difference in dose. Often a bench or pharmacodynamic measure has greater assay sensitivity than long-term clinical trials in the target population. This change in philosophy is reflected in recent European regulatory guidance for inhaled drugs and for biosimilar monoclonal antibodies in the search for clinical tests that are fit for purpose.

Competing Interests

J.W. worked previously as a medical assessor for the MHRA and helped represent the UK at the European Medicines Agency. He currently works as a consultant and has advised multiple pharmaceutical companies on drug development and regulation. This article is based on a lecture that was funded in part by Sandoz GmbH. The views expressed are the personal views of the author and do not represent the view of an agency or regulatory authority.

I am most grateful for helpful comments from Dr Alfredo García-Arieta (Agencia Española de Medicamentos y Productos Sanitarios, Madrid), Mr George Wade (formerly of the European Medicines Agency, London) and Professor James Ritter (London).

REFERENCES

- Warren JB, Oselin K. Comment on: generics and therapeutic substitution: a viewpoint on achieving best practice in Europe by Johnston *et al.* BJCP 2011; 72: 733–34.
- Shewhart WA. Economic Control of Quality of Manufactured Product. New York: Van Nostrand, 1931.
- Biography. The man. W. Edwards Deming Institute. 2012. Available at <http://www.deming.org> (last accessed 25 May 2012).
- Ericson C. Fault tree analysis – a history. Proceedings of the 17th International Systems Safety Conference. 1999.
- Available at <http://www.fault-tree.net/papers/ericson-fta-history.pdf>. (last accessed 26 February 2012).
- Montgomery DC. In: Statistical Quality Control: A Modern Introduction, 6th edn. Hoboken, NJ: John Wiley & Sons, 2009.
- Okie S. Multinational medicines – ensuring drug quality in an era of global manufacturing. N Engl J Med 2009; 361: 737–40.
- U.S. Food and Drug Administration. Orange Book: Approved Drug Products with Therapeutic Equivalence Evaluations, 31st edn. Rockville, MD: U.S. Food and Drug Administration, 2011. Available at <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM071436.pdf> (last accessed 26 February 2012).
- Sobel D. Longitude: The True Story of A Lone Genius Who Solved the Greatest Scientific Problem of His Time. New York: Penguin Group, 1996.
- Frank RG. The ongoing regulation of generic drugs. N Engl J Med 2007; 357: 1993–6.
- European Medicines Agency. Guideline on the investigation of bioequivalence, CPMP/EWP/QWP/1401/98 rev. 1/ corr. London. 2010. Available at http://www.emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070039.pdf (last accessed 26 February 2012).
- Food and Drug Administration. Bioequivalence recommendations for specific products. 2012. Available at <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm075207.htm> (last accessed 24 May 2012).
- Williams RL. Therapeutic Equivalence of Generic Drugs – Response to National Association of Boards of Pharmacy. Silver Spring, MD: FDA, 1997. Available at <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/ucm073224.htm> (last accessed 26 February 2012).
- Davit BM, Nwakama PE, Buehler GJ, Conner D, Haidar SH, Patel DT, Yang Y, Yu LX, Woodcock J. Comparing generic and innovator drugs: a review of 12 years of bioequivalence data from the United States Food and Drug Administration. Ann Pharmacol 2009; 43: 1583–97.
- Westlake WJ. Use of confidence intervals in analysis of comparative bioavailability trials. J Pharm Sci 1972; 61: 1340–1.
- Schuurmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinetic Biopharm 1987; 15: 657–80.
- Wade G. Generics in the centralised procedure. Progress and Current Issues. 2011. Available at http://www.ema.europa.eu/docs/en_GB/library/Presentation/2011/06/WC5001078document_72.pdf (last accessed 26 February 2012).
- European Medicines Agency. Notes for Guidance on the clinical requirements for locally applied locally acting

- products containing known constituents. CPMP/EWP/239/95. London, 1996. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003687.pdf (last accessed 26 February 2012).
- 18 European Medicines Agency. Points to consider on the requirements for clinical documentation for orally inhaled products (OIP), CPMP/EWP/4151/00. London, 2004. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003558.pdf (last accessed 26 February 2012).
 - 19 European Medicines Agency. The guideline on the requirements for clinical documentation for orally inhaled products (OIP) including the requirements for demonstration of therapeutic equivalence between two inhaled products for use in the treatment of asthma and chronic obstructive pulmonary disease (COPD) in adults and for use in the treatment of asthma in children and adolescents, CPMP/EWP/4151/00 Rev. 1. London, 2009. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003504.pdf (last accessed 26 February 2012).
 - 20 Moerman DE. Meaningful placebos – controlling the uncontrollable. *N Engl J Med* 2011; 365: 171–2.
 - 21 Milanowski J, Qualtrough J, Perrin VL. Inhaled beclomethasone with non-CFC propellant (HFA 134a) is equivalent to BDP-CFC for the treatment of asthma. *Resp Med* 1999; 93: 245–51.
 - 22 Lipworth BJ, Jackson CM. Pharmacokinetics of CFC & HFA MDIs of beclomethasone dipropionate. *Br J Clin Pharmacol* 1999; 48: 866–8.
 - 23 Hindle M, Chrystyn H. Determination of the relative bioavailability of salbutamol to the lung following inhalation. *Br J Clin Pharmacol* 1992; 34: 311–5.
 - 24 Hindle M, Chrystyn H. Relative bioavailability of salbutamol to the lung following inhalation using metered dose inhalation methods and spacer devices. *Thorax* 1994; 49: 549–53.
 - 25 Hindle M, Newton DAG, Chrystyn H. Dry powder inhalers are bioequivalent to metered-dose inhalers: a study using a new urinary albuterol (salbutamol) assay technique. *Chest* 1995; 107: 629–33.
 - 26 Tomlinson HS, Corlett SA, Allen MB, Chrystyn H. Assessment of different methods of inhalation from salbutamol metered dose inhalers by urinary drug excretion and methacholine challenge. *Br J Clin Pharmacol* 2005; 60: 605–10.
 - 27 Tomlinson HS, Corlett SA, Chrystyn H. Dose-response relationship and reproducibility of urinary salbutamol excretion during the first 30 min after an inhalation. *Br J Clin Pharmacol* 2003; 56: 225–7.
 - 28 Aswania O, Ritson S, Iqbal SM, Bhatt J, Rigby AS, Everard ML. Intra-subject variability in lung dose in healthy volunteers using five conventional portable inhalers. *J Aerosol Med* 2004; 17: 231–8.
 - 29 Richardson CH, de Matas M, Hosker H, Mukherjee R, Wong I, Chrystyn H. Determination of the relative bioavailability of salbutamol to the lungs following inhalation from dry powder inhaler formulations containing drug substance manufactured by supercritical fluids and micronization. *Pharm Res* 2007; 24: 2008–17.
 - 30 MHRA. Public assessment report. Decentralised procedure. Neoven/Sereflo CFC-free inhaler 25 ug per actuation pressurised inhalation suspension (salmeterol xinafoate). 2011. Available at <http://www.mhra.gov.uk/home/groups/par/documents/websiteresources/con120337.pdf> (last accessed 26 February 2012).
 - 31 Singh D, Collarini S, Poli G, Acerbi D, Madasi A, Rusca A. Effect of AeroChamber Plus™ on the lung and systemic bioavailability of beclomethasone dipropionate/formoterol pMDI. *Br J Clin Pharmacol* 2011; 72: 932–9.
 - 32 Calverley PM, Anderson JA, Celli B *et al.* Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *NEJM* 2007; 356: 775–89.
 - 33 Food and Drug Administration. Critical path opportunities for generic drugs, Office of Generic Drugs, CDER, May 1, 2007, 4.3.1 Bioequivalence of Inhalation Products. 2007. Available at <http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077250.htm> (last accessed 26 February 2012).
 - 34 Garcia-Arieta A. Sensitive studies with a significant dose-response curve for inhaled corticosteroids to investigate equivalent relative potency are feasible. *Br J Clin Pharmacol* 2011; 72: 832–3.
 - 35 Lipworth B, Clearie K. Guidelines to establish relative potency are not in touch with reality. *Br J Clin Pharmacol* 2012; 72: 834–5.
 - 36 Kuhlmann M, Marre M. Lessons learned from biosimilar epoetins and insulins. *Br J Diabetes Vasc Dis* 2010; 10: 90–7.
 - 37 FDA. Draft guidance for industry on quality considerations in demonstrating biosimilarity to a reference protein product. 2012. Available at <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm> (last accessed 30 April 2012).
 - 38 European Medicines Agency. Guideline on Similar Biological Medicinal Products. London, 2005. Available at <http://www.ema.europa.eu/pdfs/human/biosimilar/043704en.pdf> (last accessed 26 February 2012).
 - 39 European Medicines Agency. EPAR for omnitrope. London, 2006. Available at <http://www.ema.europa.eu/humandocs/.../EPAR/omnitrope/omnitrope.htm> (last accessed 26 February 2012).
 - 40 European Medicines Agency. Guideline on similar biological medicinal products containing monoclonal antibodies. London, 2010. EMA/CHMP/BMWP/403543/2010. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/11/WC500099361.pdf 2011. EMA/CHMP/BMWP/403543/2010. (last accessed 26 February 2012).
 - 41 European Medicines Agency. Guideline on similar biological medicinal products 5 containing interferon beta. London, 2012. Available at http://www.ema.europa.eu/docs/en_GB/

document_library/Scientific_guideline/2012/01/WC500120652.pdf (last accessed 26 February 2012).

- 42** Smith-Parker J. Pharma majors' multiple sclerosis drugs safe from EU competition for several years. Financial Times 3rd February, 2012.
- 43** European Medicines Agency. EPAR Rebif, interferon beta-1a. 2009. Available at http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000136/human_med_001018.jsp&mid=WC0b01ac058001d124&jsenabled=true (last accessed 6 June 2012).
- 44** Kozlowski S, Woodcock J, Sherman RB, Midthun K. Developing the nation's biosimilars program. N Engl J Med 2011; 365: 385–8.
- 45** FDA. Guidance for industry. Scientific considerations in demonstrating biosimilarity to a reference product. 2012. Available at <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm> (last accessed 26 February 2012).