Pharm Res (2014) 31:837–846
DOI 10.1007/s11095-013-1259-1

EXPERT REVIEW

Current Challenges in Bioequivalence, Quality, and Novel Assessment Technologies for Topical Products

Avraham Yacobi • Vinod P. Shah • Edward D. Bashaw • Eva Benfeldt • Barbara Davit • Derek Ganes • Tapash Ghosh • Isadore Kanfer • Gerald B. Kasting • Lindsey Katz • Robert Lionberger • Guang Wei Lu • Howard I. Maibach • Lynn K. Pershing • Russell J. Rackley • Andre Raw • Chinmay G. Shukla • Kailas Thakker • Nathalie Wagner • Elizabeta Zovko • Majella E. Lane

Received: 12 August 2013 / Accepted: 5 December 2013 / Published online: 7 January 2014

ABSTRACT This paper summarises the proceedings of a recent workshop which brought together pharmaceutical scientists and dermatologists from academia, industry and regulatory agencies to discuss current regulatory issues and industry practices for establishing therapeutic bioequivalence (BE) of dermatologic topical products. The methods currently available for assessment of BE were reviewed as well as alternatives and the advantages and disadvantages of each method were considered. Guidance on quality and performance of topical products was reviewed and a framework to categorise existing and alternative methods for evaluation of BE was discussed. The outcome of the workshop emphasized both a need for greater attention to quality, possibly, *via* a Quality-By-Design (QBD) approach and a need to develop a "whole toolkit" approach towards the problem of determination of rate and extent in the assessment of topical bioavailability. The discussion on the BE and clinical equivalence of topical products revealed considerable concerns about the variability present in the current methodologies utilized by the industry and regulatory agencies. It was proposed that academicians, researchers, the pharmaceutical industry and regulators work together to evaluate and validate alternative methods that are based on both the underlying science and are adapted to the drug product itself instead of single "universal" method.

A. Yacobi
DOLE Pharma LLC, New York, New York, USA

V. P. Shah
Pharmaceutical Consultant
North Potomac, Maryland, USA

E. D. Bashaw • B. Davit • T. Ghosh • R. Lionberger • A. Raw • C. G. Shukla
U.S. Food and Drug Administration, Silver Spring, Maryland, USA

E. Benfeldt
Department of Dermatology, University of Copenhagen
Roskilde Hospital, Copenhagen, Denmark

D. Ganes
Ganes Pharma Inc., Toronto, Ontario, Canada

I. Kanfer
Rhodes University, Grahamstown, South Africa

G. B. Kasting
University of Cincinnati, Cincinnati, Ohio, USA

L. Katz
Biostudy Solutions LLC, Wilmington, North Carolina 28405, USA

G. W. Lu
Allergan, Irvine, California, USA

H. I. Maibach
University of California–San Francisco, San Francisco, California, USA

L. K. Pershing
University of Utah Health Sciences Center, Salt Lake City, Utah, USA

R. J. Rackley
Mylan Pharmaceuticals Inc, Morgantown, West Virginia, USA

K. Thakker
Tergus Pharma LLC, Durham, North Carolina, USA

N. Wagner
Galderma R&D, Sophia Antipolis, France

E. Zovko
Forest Research Institute, New York, USA

M. E. Lane (✉)
University College London School of Pharmacy, London, UK
e-mail: majella.lane@btinternet.com

KEY WORDS bioequivalence · dermatologic topical product · generic · *in vitro* · *in vivo* · quality-by-design

INTRODUCTION

This report provides a summary of the workshop entitled "Evaluation of Topical Drug Products—Current Challenges in Bioequivalence, Quality, and Novel Assessment Technologies" held on March 12–14, 2013, Rockville, Maryland, USA. The workshop was organised by the Product Quality Research Institute (PQRI) and was co-sponsored by the American Association of Pharmaceutical Scientists (AAPS), the European Federation for Pharmaceutical Scientists (EUFEPS), the International Pharmaceutical Federation (FIP) and the United States Pharmacopoeia (USP).

In assessing generic formulations, regulatory agencies require the demonstration of bioequivalence (BE) to a reference drug product. The US Food and Drug Administration (FDA) guidelines note that, taken together with the confirmation of pharmaceutical equivalence, establishing BE allows for a regulatory conclusion of therapeutic equivalence (1). BE may be documented by performance of different types of studies (2), in declining order of preference: (i) Pharmacokinetic measurements (ii) Pharmacodynamic measurements (PD) (iii) Comparative clinical trials and (iv) *In vitro* tests. Following a citizen's petition response (3) the FDA has recently issued a draft guidance and ANDA approval has been granted (4) utilizing *in vitro* characterization to establish pharmaceutical equivalence of acyclovir ointment and grant a waiver for a clinical end-point bioequivalence study.

As the target site of dermatological formulations is the skin, most topical products produce none to at best very low measurable amounts of drug in blood or plasma. Currently, comparative clinical trials are used to establish BE for most dermatological formulations except in the case of topical corticosteroids which are tested for BE by PD measurements (5) and a few specific cases where alternative approaches are recommended. This is in contrast to the development and implementation of regulatory guidance for oral products (6) based on pharmacokinetic measurements and the Biopharmaceutics Classification System (BCS).

Determining BE of topical dermatologic products clearly remains a challenging area for the regulatory authorities and for the pharmaceutical industry alike. This represents a formidable challenge to generic competition and a huge additional cost to consumers (7). There is a need for BE studies using alternate approaches which are faster, less expensive, more reproducible and sensitive to differences in topical dermatological products. For these reasons this workshop was organized to facilitate an open dialogue between all stakeholders on (i) Current regulatory issues and industry practices and (ii) Possibilities for exploring alternative new methodologies and guidance for BE assessment of topical dermatologic products.

The goals and objectives of the workshop were to:

1. Develop a science based regulatory approach for development and evaluation of topical dermatologic products – product quality, performance and BE determination.
2. Review the current approaches and novel technologies for determining bioequivalence of topical dermatologic products.
3. Identify and suggest the methodologies that may be used by the regulatory agencies to assess BE as part of registration dossiers.
4. Evaluate product uniformity and stability and impact on BE of commercial products.
5. Determine the value of *in vitro* drug release in semisolid dosage form development and in assessment of product quality and BE.
6. Develop this summary report of the workshop discussions and recommendations.

CHARACTERISTICS OF TOPICAL PRODUCTS

A topical dermatologic product is designed to deliver drug into the various layers of skin for treating dermal disorders that is, the skin is the target organ. The amount of a semisolid formulation applied is approximately 2 mg/cm^2 and the thickness of the product applied to the diseased/involved skin will vary depending on the individual excipients and type of formulation (*i.e.* cream, lotion, ointment, gel, *etc.*). In most cases the area of application is not occluded and there may be loss or evaporation of some of the components to the atmosphere as well as into the skin. It is also likely that some of the product may be removed by clothing. Unoccluded topical formulations may be "dynamic" and as soon as they are applied to the skin there may be changes in the formulation thermodynamic activity and the other physicochemical factors dictating drug delivery into the skin.

More recently, a number of patches have been developed which contain the active in a suitable adhesive matrix for topical *vs.* systemic delivery. In contrast to traditional topical products discussed above, these patches are occlusive and thus drug and excipients will be exposed only to the skin and not to the atmosphere. Thus, apart from some preliminary increase of skin hydration, the drug thermodynamic activity and physicochemical factors remain relatively constant for the duration of the patch application time.

For all topical products delivery of most of the active ingredient locally to the skin, rather than through the skin, is the ideal scenario. It is well understood that, the delivery of

drug into the skin is a complex phenomenon which includes the phsyicochemistry of the active ingredient itself, the effect of formulation, and the disease state. Also, in addition to loss of formulation components to the environment, there will also be transport of excipients into the skin. These excipients may modify the diffusion and/or partition properties of the active in the skin. Thus effective delivery to the target site will be a function of the unique composition of these preparations and how the ratio of these ingredients changes in relation to the active ingredient as time progresses following application.

Depending on the disease condition the target site for the topical product will vary. Fungal infections require treatment of the stratum corneum (SC); for management of eczema and psoriasis drug action is needed in the viable epidermis; amelioration of muscle strains and sprains requires penetration to the deeper tissues. The duration and magnitude of the therapeutic response will depend on (i) release of the active from the formulation; (ii) penetration or uptake of the drug into the skin; (iii) activation of the pharmacological effect; and (iv) the intradermal persistence of effective levels.

A number of methodologies have been used for assessment of drug delivery from topical products into the skin and these will be discussed further in the next section together with their respective advantages and limitations.

CLASSIFICATION OF TOPICAL GENERIC PRODUCTS WITH REFERENCE TO THE RLD

When comparing a generic product with the reference listed drug (RLD) the following classifications are used:

- Q1 means qualitative similarity between generic and reference listed products with respect to composition of the individual ingredients;
- Q2 represents quantitative similarity of each ingredient;
- Q3 products encompass Q1 and Q2 requirements but also have structural similarity to the RLD with the same arrangement of matter and state of aggregation of the product.

Examples of products that are Q1 and Q2 similar to each other, but differ in Q3 would be a solid dosage forms that differ only in crystal structure or suspensions or emulsions that differ only in particle size distribution or gels that differ only in the extent of cross-linking. For pharmaceutical dosage forms that are in thermodynamic equilibrium, such as solutions of small molecules, specification of Q2 and the conditions of temperature and pressure uniquely determine Q3. For the many dosage forms that are not in thermodynamic equilibrium, including most topical formulations, their Q3 or arrangement of matter depends on their history which includes the manufacturing processes and the conditions of storage. Differences in Q3 could manifest themselves as differences in physical properties such as rheology or *in vitro* release (dissolution) rate.

ASSESSMENT OF BIOEQUIVALENCE OF TOPICAL PRODUCTS

Clinical Endpoint Studies

For the majority of topical drug products comparative clinical endpoint studies are used to demonstrate BE to the RLD. The use of clinical endpoints to determine BE of topical products, although providing a direct assessment in patients that is reassuring to clinicians, is associated with a number of challenges as well. Clinical endpoints are associated with high variability and low sensitivity that make such studies less reliable and less efficient. In general a clinical response to a drug is known to be quite variable largely due to patho-physiological and environmental factors which influence the performance of a given agent. These factors influence the elicited effect by a drug, particularly with topical products which are applied on the skin and the quantity of the dose often is not standardized and determined by the patients. For topical skin products the variability in response to a particular drug product is compounded by the disease state which often spreads in a non-homogeneous way over a small or large area. Thus, clinical scientists have always sought to find alternative ways to assess response *via* standard methods, using a standard dose and utilizing a surrogate marker or quantitated response. Also, in these studies, formulation differences might not be detected efficiently and the number of patients enrolled can be quite large. Recommendations for specific products have been provided by the FDA as a guide to the pharmaceutical industry to conduct specific clinical BE studies for regulatory filing including for example, creams, gels and ointments (8). Data from one investigation site or even from few patients receiving the RLD, the test product or the placebo may, in some cases, determine the outcome of very large clinical trials with a large group of investigators.

Pharmacokinetic Trials

The use of pharmacokinetic studies to demonstrate BE for topical products is limited to some special cases where significant systemic absorption of the drug occurs. A recent example includes the FDA recommendation for lidocaine patches in a draft guidance (9,10) followed by an approved product (11). Pharmacokinetic studies are largely used to determine the safety of topical products as the majority of the safety issues with them, outside of local irritant effects, arise from systemic absorption and the inability, at this time, to determine drug concentrations at the site of action *i.e.*, the skin itself apart from the plasma.

Pharmacodynamics (Vasoconstriction Assay)

Topical corticosteroids (also known as glucocorticoids) produce a skin blanching response associated with drug induced constriction of cutaneous blood vessels (vasoconstriction) at the site of application. The original skin blanching methodology, commonly referred to as the Stoughton-McKenzie vasoconstriction assay employed visual evaluation of the degree of blanching by trained observers (12). In order to address the highly subjective nature of the assessment, instrumentation such as the chromameter has subsequently been recommended and used to evaluate the blanching response.

A BE guidance for corticosteroids was introduced by the FDA in 1995 (5). Specifications in the guidance require a dose related vasoconstriction response estimation by the use of a Minolta chromameter in a preliminary pilot study to determine the parameters for use in a pivotal BE study. The pilot study characterizes the dose-duration response relationship for the drug in terms of an E_{max} model and is conducted solely with the RLD. The E_{max} model describes some measure of effect (E) in terms of a baseline effect (E_0), a maximal effect (E_{max}) and a dose (D) at which the effect is half-maximal (ED_{50}):

$$E = E_0 + \frac{E_{max} \times D}{ED_{50} + D}$$

The dose duration method is based on three dose durations: ED_{50}, D_1, and D_2. The comparison of test and reference products in the pivotal study is conducted at a dose duration approximately equal to the population ED_{50} determined in the pilot study. Sensitivity in the pivotal study is established through dosing of the RLD calibrators at two dose durations, D_1 (the shorter dose duration calibrator) and D_2 (the longer dose duration calibrator). The guidance recommends that D_1 equal approximately 0.5 times ED_{50}, and D_2 equal approximately two times ED_{50} determined from the pilot study. Because each subject becomes a 'detector' in the study, only the data of those subjects whose D_2/D_1 ratio of area under effect curve (AUEC) measured up to 24 h after product removal meets a value ≥ 1.25 may be included in the data and statistical analyses supporting *in vivo* BE. Vasoconstriction is measured periodically over 24 h after product removal for both pilot and pivotal studies and the respective $AUEC_{0\text{-}24}$ time values are calculated. The 90% confidence interval is calculated for the ratio of the average AUEC response due to the test product to the average AUEC response associated with the reference product exposure at the ED_{50} dose. Demonstration of BE is based on statistical analysis showing that the 90% confidence interval on the test:reference ratio is within 0.80–1.25.

While this method works well with semi-solid formulations, a number of practical problems have been reported with the VCA using gels and sometimes collapsible foams. One of the most commonly reported issues is that of high inter-subject variability which may be >50% in some studies. Additionally, it can be difficult to determine the correct ED_{50} with gel and solution formulations in which the drug is fully soluble. Volunteers may have low qualification rates *i.e.* they cannot discriminate well between $2*ED_{50}$ and $0.5*ED_{50}$ and thus do not meet the ratio of vasoconstriction requirement of $AUEC_{0\text{-}24}$ $D_2/D_1 \geq 1.25$ to qualify for inclusion in the final dataset for assessing BE. Other criticisms of the pharmacodynamic method include the necessity to include all subjects in the pilot study, independent of their demonstrated dose-dependent response, but subsequent removal of non-detector (lack of dose-dependent response) subjects from the pivotal study. Consequently, with some products, large numbers of subjects may have to be enrolled (*e.g.*, $n = 100$) in the pivotal study to achieve the required number of evaluable subjects for BE assessment.

ALTERNATIVE/NOVEL METHODS TO ASSESS BIOEQUIVALENCE

Dermatopharmacokinetics

Since the target organ of topical products is the skin it would seem logical that determining drug concentration in skin should provide an indication of topical BE between products. This is the underlying principle of the Dermatopharmacokinetic (DPK) technique. Using this approach, only the SC not the deeper epidermal and dermal skin layer is evaluated for drug content. The Office of Generic Drugs (OGD) proposed a DPK test as a universal method for demonstrating BE of all topical drug products in June 1998 (13) but withdrew the draft guidance 4 years later in May 2002 (14). In the original draft guidance SC at the product exposed site is removed by the sequential application and removal of 12 pieces of adhesive tape, of which the first two are discarded and the remaining ten are combined and quantified for drug. The guidance further specified that topical product performance be quantified by the time integration of the drug amount in the harvested SC over time (AUC) to evaluate BE in terms of the maximum amount per unit area in the SC (Q max), the time at which Q max is first observed (T max). To establish the kinetic profiles, the guidance indicated that the amount of drug in the SC should be determined in no less than eight sites: at least four sites exposed to the product for different exposure periods to capture the rate and extent of drug uptake into skin, and four sites exposed to drug for the longest exposure period followed by four different clearance periods to capture the elimination phase of a typical pharmacokinetic profile.

The guidance was not pursued further and was withdrawn when evaluation of the DPK method with commercially available tretinoin gel products in two different laboratories

showed contradictory results (15,16). The studies were conducted with different protocols, rendering any comparison improper. A further concern was that the DPK test did not correlate with clinical effects at local sites of action. However, no direct comparison was conducted to deal with this concern. Efforts to refine the DPK approach continued post 2002 with FDA initiated studies to identify and evaluate sources of variability in the method and to optimize the protocol procedure(s). These studies evaluated an anti-fungal drug in a semi-solid vehicle formulation, which primarily targets the SC, (17,18). Several modifications to the original FDA guidance were evaluated. The number of sampling times was reduced, measurements were duplicated in each subject, transepidermal water loss (TEWL) measurements were made to assess the fraction of the SC removed by the tape-stripping procedure, a new cleaning procedure reduced variability by improving removal of residual drug before tape stripping, and drug present on all tape strips was included in the comparison of the amounts taken up into the SC from different products. In more recent publications which examined a topical corticosteroid (19) and an antifungal (20) the importance of an appropriate dose duration was highlighted and the analysis of each tape strip and normalization of the amount of SC removed was further emphasized.

Notwithstanding the refinements suggested in these reports, there are a number of other limitations associated with the DPK approach. The drug concentration at the site of action is not measured for drugs whose target is not the SC and thus bioavailability is not assessed. Even where the target site of the drug is the SC, the DPK approach will not distinguish between the amount of drug which is therapeutically available and that which may have crystallised out or been "stranded" in the skin (21,22). In addition DPK evaluation is conducted on healthy skin which is a poor model of diseased skin. The application area must be standardised as this contributed to the variability observed in the earlier DPK studies. Attention has also focussed on the tape itself employed to remove the SC. At present, the lack of assurance of "tape equivalence" from one batch, year of manufacture, manufacturer, to another, is a major weakness in terms of reproducibility. The tape selected may vary depending on the investigator or laboratory; there may be age-related changes in tape adhesive as well as changes in the adhesive system itself implemented by manufacturers. All of these factors and a process to control for them need to be considered as part of any validation strategy.

Microdialysis

Historically, microdialysis was developed as a tool to study tissue biochemistry in animals, specifically transmitter release in the rodent brain (23) and it has been available as a preclinical and clinical tool for human drug pharmacokinetic studies for more than two decades. The method consists of placing an ultrathin, semipermeable hollow fiber structure in the dermis and perfusing this fiber, called a probe, with a tissue-compatible sterile buffer (physiological Ringer's solution or phosphate buffered saline) at a very low rate (between 0.1 and 5 μl/min) by means of a microdialysis pump (a very precise syringe driver). The probe functions as an "artificial vessel" in the dermis, exchanging diffusible molecules backwards and forwards between the probe and the tissue. The technique permits direct and continuous measurement of unbound drug. With reference to skin, dermal microdialysis (DMD) has been successfully used to quantify endogenous (24) as well as the absorption of exogenous substances (25). More recently, DMD has been used to assess drug permeation in both healthy and damaged skin (26,27) and to determine the BE of certain topical formulations (28,29).

Probe designs are available with different membrane pore sizes which determine the upper molecular size of analytes which can be sampled by that particular probe. Probes are also available in a range of designs: linear, loop, concentric, side-by-side. In DMD it is the linear probe which has been predominantly used. In low flux systems laboratory-grade linear probes may sample molecules up to 10 kDa and in high flux systems the range is 40–70 kDa. Commercially available membranes may sample higher molecular weight compounds and have cut-off values up to 3,000 kDa. Once diffusion of molecules has taken place across the linear probe membrane the perfusate is termed dialysate.

As the probe is continuously perfused with fresh perfusate a steady state rate of exchange across the membrane is rapidly attained. The "absolute recovery" is the actual amount of analyte collected by the probe in a finite time period. The ratio of analyte between the dialysate and the extracellular fluid is termed the relative recovery (RR). The RR is influenced by a number of factors including the physicochemical properties of the analyte (molecular weight, configuration, lipophilicity, protein binding), tissue temperature, probe material, probe design, probe surface area, probe depth, perfusate and perfusate flow rate. In DMD, the perfusate is most often an isotonic saline solution or a Ringer solution. Depending on the lipophilicity of the drug, the perfusate medium may have to be modified to allow more lipophilic substances to enter the probe.

Calibration is necessary for quantitative DMD. *In vitro* calibration is an important step when studying a new molecule with DMD although it cannot replace *in vivo* calibration. The former will give an indication of the RR and subsequently the feasibility of *in vivo* DMD. It is important to establish the solubility properties of the analyte in the perfusate, any adherence of the analyte to the probe material and analyte loss *versus* recovery. A number of methods may be used for *in vivo* calibration: the no-net flux method, stop flow/flow rate

method and retrodialysis, further details of these approaches are outlined elsewhere (30).

DMD is comparatively more invasive than DPK. On the other hand, for most drugs, DPK data may not correlate with amount of drug at the site of action whereas DMD can provide detailed chronological pharmacokinetic data. More importantly, these observations can be obtained in subjects with the disease itself, without relying on extrapolations from normal skin, as is the case for DPK. In addition, for DMD, several sampling sites can be studied simultaneously in the same volunteer. The method can be challenging when used for sampling very lipophilic or highly protein-bound drugs because of low recovery of these molecules. One option for the measurement of lipophilic molecules is the use of lipid emulsion as the perfusate instead of the conventional aqueous buffer. Training of laboratory personnel is necessary to ensure low variability.

The FDA has actively sought information, presentations and co-operation with DMD researchers and submission of DMD studies has been encouraged by the FDA. Pre-and post-approval DMD data has also been requested to support *in vivo* bioavailability trials.

Open Flow Microperfusion

Open flow microperfusion (OFM) was introduced in 1997 as a tool for metabolic research (31). The first report of the application of OFM in dermal sampling appeared in 2011 (32). Typical flow rates are ≥0.5 μl/min. In contrast to the probes used in DMD, the probe design for OFM has an open exchange area and thus samples interstitial fluid directly. For this reason OFM does not have the same limitations as DMD specifically with reference to molecular size, drug protein-binding and drug lipophilicity. Thus, ideally, it could be used to establish a relationship between the quantity of drug substance in the dermis and the pharmacological response by measuring simultaneously the drug product and pharmacodynamic marker (*e.g.* skin levels of cytokines). However the OFM probes require active push-pull pumps in order to avoid the loss of perfusate to the tissue and the risk of oedema formation. Recently the technique has been used to determine the dermal delivery of a lipophilic molecule (BCT 194) in psoriatic patients (33). The molecule was applied as a cream (0.5%) to a lesional and a non-lesional application site once daily for 8 days in nine patients. Multiple sampling was performed for 25 h from each site on day 1 and day 8. OFM was well tolerated and demonstrated significant drug concentrations in lesional as well as non-lesional skin after 8 days, but did not show significant differences between tissues. While OFM may offer advantages in relation to the range of sampled substance, OFM samples are generally more complex than DMD samples and require pretreatment before analysis.

Confocal Raman Spectroscopy

Raman spectroscopy is an optical method based on inelastic light scattering. Confocal Raman Spectroscopy (CRS) combines the principles of confocal microscopy and Raman spectroscopy. The device comprises a high-performance dispersive spectrometer with laser excitation and a confocal measurement stage. This technique is capable of providing detailed information concerning the molecular composition of the SC and specifically the water gradient across the SC (34). CRS has been used in the personal care industry for a number of years to determine SC thickness, to profile endogenous skin components and to probe the influence of formulations on the SC (35–37). It is a completely non-invasive technique and a comprehensive safety assessment of its use in humans has been conducted (38). The application of CRS in profiling topical drug delivery in humans was reported recently by Mateus and co-workers (39). CRS was used to examine the disposition of ibuprofen after application from simple formulations which were previously investigated using the DPK approach. The results confirmed that ibuprofen distribution profiles in the SC were comparable to previously published data from tape stripping experiments. The authors proposed CRS as a non-invasive alternative to DPK for evaluation of topical pharmaceutical formulations.

In Vitro Skin Permeation Studies

In vitro skin permeation studies are routinely carried out in topical formulation development and to evaluate how a generic product compares with the RLD. The rate of permeation of a drug from the generic and RLD formulation and the time course over which permeation occurs should be comparable where both products have the same qualitative and quantitative composition and have the same physicochemical properties *i.e.* Q3. Skin samples are typically mounted in vertical diffusion cells and are securely clamped between a donor and receptor compartment. The donor compartment is open to the atmosphere. The amount of a semi-solid formulation applied in the donor compartment should simulate typical clinical application conditions where the patient applies ~2 mg/cm^2 to the skin. The receptor compartment contains a fluid designed to mimic the physiological scenario *e.g.* isotonic buffered saline, maintained at a suitable temperature to ensure that the skin temperature is 32°C. Of note, the Organization for Economic Co-Operation and Development (OECD) published guideline (#428) in 2004 regarding the absorption of a test substance applied to excised skin using diffusion cells (40). The OECD Guidance Document has been prepared to facilitate the selection of appropriate procedures for use in specific circumstances, to ensure the reliability of results obtained by this method.

However, the inherent variability associated with use of human skin samples has already been highlighted in an earlier report by Shah and co-workers (41) and the limitations associated with this approach have not yet been fully explored. In addition to excised human skin, porcine skin has also been used to investigate drug permeability from topical products. Fresh samples as well as samples which have been frozen have been used for such studies. The tissue may be heat separated, surgically separated or dermatomed before being mounted in the Franz cells; full thickness skin samples have also been used.

Franz and co-workers (42) evaluated the *in vitro* skin permeation of seven generic topical drug products and compared the data with the corresponding reference products during preclinical development. The relevant clinical data from BE trials became available post regulatory approval of the generic products and were also compared with the *in vitro* data. In five of the seven comparisons the same lots of both test and reference products used *in vivo* had been used *in vitro*. Generic versions of the following products were studied and compared with the RLD products: tretinoin (RetinA ® 0.01% and 0.025% gel), alclometasone dipropionate (Aclovate ® 0.05% cream and ointment), halobetasol propionate (Ultravate ® 0.05% cream and ointment), and mometasone furoate (Elocon ®0.1% ointment). All products were evaluated on skin sections from the same donors. The rate of absorption and total absorption were calculated for all formulations. In six of the seven cases the *in vitro* test:reference ratio for total absorption was close to one and indicated that the products were equivalent, in agreement with the clinical data. Results from the seventh case, in which the test:reference ratio was only 0.63, indicated that the *in vitro* model actually had greater sensitivity than the clinical method to detect small differences between products.

In Vitro Release Testing

An *in vitro* release test (IVRT) for semisolid preparations has been recognized in the FDA's SUPAC-SS guidance as a test for product sameness after certain manufacturing related changes (43). Also, a guidance relative to the conduct of the IVRT has been issued (6). Recently, this IVRT is also recognized as a reasonable and useful test to be considered as a product release and stability test (44). Can this principle of IVRT be extended to other semisolid drug products when Q_1, Q_2 and Q_3 are the same? Just as lower strengths of oral dosage forms are eligible for biowaiver when they are dose proportional to higher strengths and meet dissolution profile similarity (f_2) criteria, can the same principles be applied for lower strengths of topical dosage forms? These are important points which need to be explored.

Modelling/Pharmacometric Approaches

There are a number of applications of modelling and pharmacometrics in topical formulation development including (i) Modelling systemic concentrations for safety (ii) Modelling and simulation approaches for disposition in specific population such as paediatrics and geriatrics (iii) PK/PD models to link different methodologies, for example skin concentrations of clobetasol obtained using open flow microperfusion with biomarkers of inflammation (iv) Modelling for QBD in formulation development and optimization and prediction of vehicle effects. However, the role of modelling and pharmacometrics in establishing BE is relatively unexplored at the present time.

ASSESSMENT OF QUALITY AND PERFORMANCE OF TOPICAL DRUG PRODUCTS

Testing of the quality and performance of semi-solid products provides assurance of batch-to-batch quality, reproducibility, reliability and performance throughout the shelf life. Quality testing provides confidence that (i) The product maintains its physical integrity throughout its shelf life (ii) The product retains chemical and microbiological stability during the shelf life (iii) Container closure integrity/compatibility do not vary over the shelf life. To determine physical integrity the following assessments are recommended: determination of visual separation and/or chemical separation; colour change; pH; presence of crystals; general appearance (lumps, air, smell *etc.*); viscosity and spreadability for semi-solid products. For chemical stability the following should be conducted: assays of the active content or potency; tube/container uniformity, assay of the preservative content or potency; determination of the presence of degradation products generated during the manufacturing process or over the shelf life of the product; evaluation of any impurities of the active, inactive ingredients or product. Microbiological stability is examined by microbial limit testing preservative effectiveness testing. Container/closure incompatibilities should be tested by assessment of: peeling of labels or any laminated material; leakage of vehicles/solvents; migration of actives or inactive materials; penetration of moisture or oxygen; colour changes due to instability, degradation or oxidation; adherence of container closure or tube; package deformation. An official acceptance criterion for analytical tube uniformity testing has recently been included in chapter <3> of the current USP.

Product performance evaluation for semisolid dosage forms may be conducted using an IVRT. In 1997, the FDA released the Scale Up and Post Approval Changes (SUPAC-SS) Guidance (43). The Guidance recommends *in vitro* diffusion cell (VDC) testing for semisolid dosage forms

when minor formulation and/or process changes have been made to an approved topical dosage form. It is also used to bridge products manufactured at two different sites. If any of a variety of formulation, process or manufacturing site changes occurs, the Guidance requires an *in vitro* release test using a VDC to compare the pre- and post-change release rates of the product. An IVRT is thus a performance verification test for test and reference semisolid dosage forms. IVRTs may also be used to support formulation development, to compare generic with innovator formulations and to characterise release data from various formulations used in clinical trials.

Typically, *in vitro* drug release is measured using a VDC with a synthetic membrane. Cellulose, Fluoropore®, nylon, polycarbonate and Supor® are examples of the types of membranes which have been used. The synthetic membrane serves as an inert support membrane separating the product from a receptor phase. Selection of the appropriate membrane and receptor phase are important factors in the design of an appropriate IVRT. The membrane must be compatible with the formulation and not interfere with recovery of the active. The solubility of the API in the receptor phase must be sufficient to allow for sink conditions. The IVRT should be developed and properly validated keeping in mind that it should be sensitive enough to capture differences in composition of formulation, dosage form strength, particle size of API, viscosity and stress conditions. With reference to validation the following attributes must be addressed: precision; accuracy/sameness; dose proportionality; sensitivity to changes in excipient type, amount of excipient, size of batch and method of manufacture. Other aspects to be considered during validation include mass balance, back diffusion of alcohol from the receptor medium to the formulation, effect of temperature, inter- and intra-day instrument variation and inter- and intra-operator variation.

PROPOSED DECISION TREE STRAWMAN OF TOPICAL PRODUCTS FOR ASSESSMENT OF BIOEQUIVALENCE

In reviewing the information presented during this workshop the participants brainstormed a decision tree "strawman" to envision both the various merits of the methods available for BE evaluation and how they "might" be applied once validation of the methods was achieved. (Table I). The consensus opinion of participants was that consideration of any method would need a rational linkage to the clinical endpoint and/or site of action which is reflected in the table. The process of decision making should start with the Q1, Q2, Q3 evaluation of the generic and RLD products.

For the example where, under this rubric, the generic product is Q1Q2 equivalent with the RLD, *in vitro* characterization and performance testing sufficient to show Q3 equivalence may be conducted and *in vivo* testing may be waived depending on the *in vitro* results. If the generic is Q1 equivalent but not Q2 then *in vivo* tests are required where the Q2 difference is potentially significant. Where the product is not Q1Q2, then *in vitro* and *in vivo* tests will be required to demonstrate no formulation effect on absorption.

Table I Draft Decision Tree Strawman for Determination of Topical BE

Q1 Q2 Q3 + option	>	Evaluate biowaiver based on similarity and formulation complexity
v		
Amenable to VCA	>	Evaluate based on accepted VCA method
v		
Amenable to PK characterization	>	Evaluate PK-based BE, alone or in conjunction with other methods
v		
Site of action, SC/epidermis	>	- Clinical Endpoint Study - DPK - Skin permeability - Confocal Raman Spectroscopy
v		
Site of action Epidermis/dermis	>	- Clinical Endpoint Study - Microdialysis - Skin permeability - Confocal Raman Spectroscopy
v		
Site of action, Below dermis	>	- Clinical Endpoint Study - Microdialysis - PK-based BE

The VCA method is available where the product is a corticosteroid. Assuming the product results in significant plasma or tissue levels, pharmacokinetic measurements are an option and may be supplemented with other methods. Where the target site of action of the drug is the SC, clinical end point evaluation, and alternative methodologies based on DPK and Confocal Raman Spectroscopy may be appropriate. For products intended to act in the epidermis, *in vitro* skin permeation testing may also provide data to supplement clinical end point evaluation, or microdialysis, or Confocal Raman Spectroscopy methods. Clinical endpoint studies, or microdialysis, could be considered where the therapeutic effect is required in the dermis. Clinical endpoint studies, microdialysis or PK-based methods may be considered for target sites below the dermis. The objective of this summary table is to demonstrate that determination of topical bioequivalence will require a multi-faceted approach, tailored to the drug, disease, product interface. The "one-size fits all" model is an outdated model and for dermal delivery science to advance we must move beyond this narrow perspective.

CONCLUSIONS

The scientific presentations and discussions at this workshop clearly underlined the need to re-evaluate the present methods and approaches to determine BE of topical dermatological products. There is a need for new creative approaches to optimise the existing available methodologies and to explore the possibility of using alternative methodologies which may facilitate the development, registration and ultimately approval of semi-solid products. The tools discussed, and others that may be developed in the future, should not be viewed in isolation, but as part of a larger "complimentary toolkit of methods". In order to do so, there is also a need for early and timely communication between all stakeholders in both the development and deployment of these technologies. For this is not a regulators' issue alone, any more than it is an issue for the academician or manufacturers alone.

A key issue to consider for any new method is that of validation and the necessity to demonstrate sensitivity of same. Clinical trials, with rare exception, remain at this time the "standard" for establishing clinical equivalence for topical products. If we are to move forward to the methods represented here, then an assurance of reproducibility (ie., the validation) of the method becomes paramount to acceptance. As mentioned previously, we need to develop a whole toolkit of approaches based on science that is itself based on timely collaboration and communication. The time to have this communication is not at the time of submission but now.

ACKNOWLEDGMENTS AND DISCLOSURES

The views expressed here are the views of the scientists and do not represent the policy of their respective agencies or organization.

The Product Quality Research Institute (PQRI) is a non-profit consortium of organizations working together to generate and share timely, relevant, and impactful information that advances drug product quality and development.

By virtue of its diverse membership, PQRI provides a unique forum to focus critical thinking, conduct research, exchange information, and propose methodology or guidance to pharmaceutical companies, regulators, and standard setting organizations.

REFERENCES

1. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Guidance for Industry. Bioavailability and bioequivalence studies for orally administered drug products—general considerations.2003.
2. Code of Federal Regulations. Title 21 Part 320 0.24 (b) (1–4)
3. www.regulations.gov/#!documentDetail;D=FDA-2012-P-0779-0004
4. Abbreviated New Drug Application 202459.
5. Guidance: topical dermatologic corticosteroids: in vivo bioequivalence, Office of Generic Drugs, Center for Drug Evaluations and Research, FDA. 1995.
6. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Guidance for Industry: Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System, 2000.
7. Generic Drug Savings in the US. Fourth Annual Edition. Generic Pharmaceutical Association. 2012.
8. BE Recommendation for Specific Products webpage: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm075207.htm.
9. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/
10. Guidances/ucm086293.pdf).www.regulations.gov/#!documentDetail;D=FDA-2006-P-0346-0017
11. Abbreviated New Drug Application 200675.
12. McKenzie AW, Stoughton RB. Method of comparing percutaneous absorption of steroids. Arch Dermatol. 1962;86:608–10.
13. US FDA. Guidance for Industry: Topical dermatological drug product NDAs and ANDAs-in vivo bioavailability, bioequivalence, in vitro release, and associated studies. Draft Guidance, June 1998, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. 1998.
14. US FDA. Guidance for industry on special protocol assessment; availability. Fed Regist. 2002;67:35122.
15. Pershing LK. Bioequivalence assessment of three 0.025% tretinoin gel products: Dermatopharmacokinetic vs. Clinical Trial Methods, Transcribed presentation to the Advisory Committee for Pharmaceutical Sciences Meeting. Rockville: Center for Drug Evaluation and Research, FDA; 2001.
16. Franz TJ. Study #1, Avita Gel 0.025% vs Retin-A Gel 0.025%, Transcribed presentation, Advisory Committee for Pharmaceutical Sciences Meeting. Rockville: Center for Drug Evaluation and Research, FDA; 2001.
17. Bunge AL, N'Dri-Stempfer B, Navidi WC, Guy RH. Dermatopharmacokinetics: improvement of methodology for assessing bioequivalence of topical dermatological drug products, Revised Final Report, Award No. D3921303, Submitted to Department of Health and Human Services. Golden: FDA, Colorado School of Mines; 2006.
18. Bunge AL, N'Dri-Stempfer B, Navidi WC, Guy RH. Therapeutic Equivalence of Topical Products, Final Report, Award No. 223-04-3004, Submitted to Department of Health and Human Services, FDA, Colorado School of Mines, Golden, CO, January 30, 2007 (Revision submitted June 2008).
19. Au WL, Skinner M, Kanfer I. Comparison of tape stripping with the human skin blanching assay for the bioequivalence assessment of topical clobetasol propionate formulation. J Pharm Pharmaceut Sci. 2010;13:11–20.
20. Parfitt NR, Skinner M, Bon C, Kanfer I. Bioequivalence of topical clotrimazole formulations: an improved tape stripping method. J Pharm Pharmaceut Sci. 2011;14:347–57.
21. Stoughton RB. Dimethylsulfoxide (DMSO) induction of a steroid reservoir in human skin. Arch Dermatol. 1965;91:657–60.
22. Santos P, Watkinson AC, Hadgraft J, Lane ME. Oxybutynin permeation in skin: the influence of drug and solvent activity. Int J Pharm. 2010;384:67–72.
23. Ungerstedt U. Measurement of neurotransmitter release by intracranial dialysis. In: Marsden CA, editor. Measurement of neurotransmitter release in vivo. Chichester: Wiley; 1984. p. 81–105.
24. Petersen LJ, Kristensen JK, Bülow J. Microdialysis of the interstitial water space in human skin in vivo: quantitative measurement of cutaneous glucose concentrations. J Invest Dermatol. 1992;99:357–60.

25. Anderson C, Andersson T, Molander M. Ethanol absorption across human skin measured by in vivo microdialysis technique. Acta Derm Venereol. 1991;71:389–93.
26. Ortiz PG, Hansen SH, Shah VP, Menné T, Benfeldt E. The effect of irritant dermatitis on cutaneous bioavailability of a metronidazole formulation, investigated by microdialysis and dermatopharmacokinetic method. Contact Dermatitis. 2008;59: 23–30.
27. Ortiz PG, Hansen SH, Shah VP, Menné T, Benfeldt E. Impact of adult atopic dermatitis on topical drug penetration: assessment by cutaneous microdialysis and tape stripping. Acta Derm Venereol. 2009;89:33–8.
28. Benfeldt E, Hansen SH, Vølund A, Menné T, Shah VP. Bioequivalence of topical formulations in humans: evaluation by dermal microdialysis sampling and the dermatopharmacokinetic method. J Invest Dermatol. 2007;127:170–8.
29. Tettey-Amlalo RN, Kanfer I, Skinner MF, Benfeldt E, Verbeeck RK. Application of dermal microdialysis for the evaluation of bioequivalence of a ketoprofen topical gel. Eur J Pharm Sci. 2009;36: 219–25.
30. Holmgaard R, Nielsen JB, Benfeldt E. Microdialysis sampling for investigations of bioavailability and bioequivalence of topically administered drugs: current state and future perspectives. Skin Pharmacol Physiol. 2010;23:225–43.
31. Trajanoski Z, Brunner GA, Schaupp L, Ellmerer M, Wach P, Pieber TR, *et al*. Open-flow microperfusion of subcutaneous adipose tissue for on-line continuous ex vivo measurement of glucose concentration. Diabetes Care. 1997;20:1114–21.
32. Bodenlenz M, Hoefferer C, Priedl J, Dragatin C, Korsatko S, Liebenberger L, *et al*. A novel certified dermal sampling system for efficient clinical research. J Invest Dermatol. 2011;131 Suppl 2: S44.
33. Bodenlenz M, Höfferer C, Magnes C, Schaller-Ammann R, Schaupp L, Feichtner F, *et al*. Dermal PK/PD of a lipophilic topical drug in psoriatic patients by continuous intradermal membrane-free sampling. Eur J Pharm Biopharm. 2012;81: 635–41.
34. Caspers PJ, Lucassen GW, Bruining HA, Puppels GJ. Automated depth-scanning confocal Raman microspectrometer for rapid in vivo determination of water concentration profiles in human skin. J Raman Spectrosc. 2000;31:813–8.
35. Crowther JM, Sieg A, Blenkiron P, Marcott C, Matts PJ, Kaczvinsky JR, *et al*. Measuring the effects of topical moisturizers on changes in stratum corneum thickness, water gradients and hydration in vivo. Br J Dermatol. 2008;159:567–77.
36. Caspers PJ, Lucassen GW, Carter EA, Bruining HA, Puppels GJ. In vivo confocal Raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles. J Invest Dermatol. 2001;116:434–42.
37. Mohammed D, Crowther JM, Matts PJ, Hadgraft J, Lane ME. Influence of niacinamide containing formulations on the molecular and biophysical properties of the stratum corneum. Int J Pharm. 2013;441:192–201.
38. Puppels GJ, Sterenborg HJCM. Laser safety aspects of the use of the Model 3510 Skin Composition Analyzer (SCA) in *in vivo* studies of human subjects. River Diagnostics. 2007. 23 pp.
39. Mateus R, Abdalghafor H, Oliveira G, Hadgraft J, Lane ME. A new paradigm in dermatopharmacokinetics - Confocal Raman spectroscopy. Int J Pharm. 2013;444:106–8.
40. OECD (2004) Test No. 428: Skin Absorption: In vitro method. OECD 2004 8 pp.
41. Skelly JP, Sha VP, Maibach HI, Guy RH, Wester RC, Flynn G, *et al*. FDA and AAPS report of the workshop on principles and practices of in vitro percutaneous penetration studies: relevance to bioavailability and bioequivalence. Pharm Res. 1987;4:265–7.
42. Franz TJ, Lehman PA, Raney SG. Use of excised human skin to assess the bioequivalence of topical products. Skin Pharmacol Physiol. 2009;22:276–86.
43. FDA. SUPAC-SS nonsterile semisolid dosage forms, scale-up and postapproval changes: chemistry, manufacturing, and controls; in vitro release testing and in vivo bioequivalence documentation. U.S. Department of Health and Human Services, Washington,DC. 1997. http://www.fda.gov/cder/guidance/1447fnl.pdf[accessed 09 April 2007).
44. Chang R-K, Raw A, Lionberger Y. L. Generic development of topical dermatologic products: Formulation development, process development, and testing of topical dermatologic products. AAPS Journal. 2013;15:41–52.