Journal of Controlled Release 387 (2025) 114190

Contents lists available at ScienceDirect

Journal of Controlled Release

journal homepage: www.elsevier.com/locate/jconrel

Raman spectroscopy as a tool to accelerate development of complex medicinal products

Panagiota Zarmpi [a,b,1], Dimitrios Tsikritsis [c,1], Andrew C. Watson [a], Jean-Luc Vorng [c], Vasundhara Tyagi [c], Natalie A. Belsey [b,c], Elena Rantou [d], Priyanka Ghosh [e], Annette L. Bunge [f], Timothy J. Woodman [a], M. Begoña Delgado-Charro [a], Richard H. Guy [a,*]

[a] Department of Life Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK
[b] School of Chemistry & Chemical Engineering, University of Surrey, Guildford GU2 7XH, UK
[c] National Physical Laboratory, Teddington TW11 0LW, UK
[d] Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, White Oak Campus, Silver Spring, MD 20993, USA
[e] Office of Research and Standards, Office of Generic Drugs, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, White Oak Campus, Silver Spring, MD 20993, USA
[f] Department of Chemical & Biological Engineering, Colorado School of Mines, Golden, CO 80401, USA

ARTICLE INFO

Keywords:
Raman spectroscopy
Cutaneous pharmacokinetics
Topical drug bioavailability
Topical drug product bioequivalence
Regulatory science

ABSTRACT

Regulatory approval of complex generic drug products that are applied topically to treat skin disease can be challenging because a standard method to assess cutaneous pharmacokinetics in vivo is not available. As a result, expensive and prolonged clinical trials are often necessary, and this has proved to be a significant disincentive to generic product development and restrictive of patient access to more affordable medicines. Here Raman spectroscopy is used to quantify the relative spatiotemporal disposition of metronidazole (a drug used to treat rosacea) within the skin and proximal to its site of pharmacological action. Analysis of the data permits pharmacokinetic metrics to be extracted and compared between different gel and solution formulations with the potential to establish bioequivalence and enable more efficient complex generic drug product development.

1. Introduction

The regulatory authorities responsible for the approval of medicinal products (including the U.S. Food & Drug Administration (FDA) and the European Medicines Agency (EMA)) typically define the bioavailability (BA) of a drug as the rate and extent at which it is absorbed from a formulation and becomes available at or near the site of action [1,2]. The demonstration of reproducible BA of the active pharmaceutical ingredient (API) is a key component of a drug product's safety, efficacy and quality. Measurement of BA relies on a number of approaches with clinical studies assessing certain pharmacokinetic metrics, in parallel with pharmacodynamic outcome evaluations, being of particular importance.

When a generic drug product is developed, bioequivalence (BE) to an identified Reference-Listed Drug (RLD) must be demonstrated – the RLD is defined as *"an approved drug product to which new generic versions are compared to show that they are bioequivalent"* [3–5]. For many systemically-acting oral generic products, BE can be established in a relatively straightforward head-to-head pharmacokinetic study comparing the reference listed drug product with the generic, it being accepted that, in so doing, therapeutic equivalence will be assured, and no efficacy study is required [2,6].

However, for complex drug products, like topical formulations used to treat skin disease, an efficient pharmacokinetic-based approach is not available and, until relatively recently (with the exception of semi-solid corticosteroid formulations and certain liquids), BE could only be demonstrated through a comparative clinical endpoint study [7]. The latter, which are typically long, expensive and often considered less sensitive, provided a strong disincentive to the development of generic, topical drug products [8]. Lately, though, the FDA has initiated an important effort to widen public access to high-quality generic medicinal products by publishing several guidances to support the

* Corresponding author.
E-mail address: r.h.guy@bath.ac.uk (R.H. Guy).
[1] These authors contributed equally to this research.

https://doi.org/10.1016/j.jconrel.2025.114190
Received 16 July 2025; Received in revised form 25 August 2025; Accepted 1 September 2025
Available online 2 September 2025

development and approval of safe and effective generic drug products [9], and by increased funding of research into alternative methods to assess topical BE without recourse to comparative clinical endpoint studies [10].

Very recently, the EMA has published a full guideline on quality and equivalence of "locally applied, locally acting cutaneous products" [11]. As well as providing guidance on when the demonstration of therapeutic equivalence may or may not require studies with clinical endpoints, the EMA document includes quite detailed guidance for the design of potentially useful elements of product specific equivalent protocols for in vitro release tests (IVRT), in vitro human skin permeation tests (IVPT), stratum corneum (SC) sampling, and the vasoconstriction assay for corticosteroids.

Concomitant with these initiatives, a major research effort has been directed towards further methods with the potential to provide alternative *in vivo* (and, ideally, minimally or non-invasive) assessments of topical drug BA [7], including the use of vibrational – and, in particular, Raman – spectroscopy [12]. However, although there is instrumentation available to track drug disposition in vivo in the SC – the thin (~20 μm, on average) and most superficial skin layer – the detection of an API below the SC in the living layers of the viable epidermis and beyond, where most dermatological drugs elicit their pharmacological action, has proven difficult for a number of reasons.

First and foremost is that Raman spectroscopy, while generally a 'label-free' technique, is typically less sensitive in terms of detection limit than other, more conventional approaches (e.g., those using fluorescent tags) [12]. When employed on the skin, this limitation is exacerbated by the presence of a significant background provided by the tissue itself and, potentially, by other, spectrally-interfering constituents of the formulation of the topical drug product [13,14]. In addition, even though Raman signals can be acquired confocally and non-invasively from sub-surface layers of the skin (at depths and resolutions that depend on the type of Raman spectroscopy employed [12,14]), they are subject to attenuation, due to absorbance and scattering, of the signal reaching the detector [13,15,16].

Nevertheless, there are frequencies at which the Raman spectrum of skin is essentially transparent [17] and, in the fingerprint region, certain molecular bond vibrations originating from exogenously-applied chemicals can be isolated from the endogenous background. The former opportunity has been exploited using both a model compound, 4-cyanophenol (CP), and active ingredients (tazarotene, a Janus kinase inhibitor, and crisaborole for atopic dermatitis) by taking advantage of their strong vibrational signals – C≡N (for CP) and C≡C stretching (for the two drugs), respectively – to track their uptake into and clearance from the skin using both confocal Raman and stimulated Raman scattering (SRS) spectroscopies [13,18]. There has also been progress with respect to CP again and to another topical drug (diclofenac) for which it has proven possible to extract information on skin uptake and disposition from Raman measurements and then to correlate these positively with mass spectroscopy imaging, an approach of inherently much higher sensitivity [19,20]. There is clearly also considerable scope to apply spectral unmixing algorithms to identify the individual component contributions to complex Raman spectra where the drug signal overlaps with those from either endogenous skin species or other (non-drug) formulation components. The most common of these tools use regression or non-matrix factorisation techniques, to identify components from spectral features based on pre-existing Raman information [21]. For instance, the multivariate curve resolution-alternating least squares approach has successfully retrieved the distribution of cutaneously distributed CP in the presence of background interference [22].

In terms of Raman signal attenuation with depth into the skin, it has been shown that normalisation with the strong Amide *I* vibration originating predominantly from the endogenous keratin present in skin (the concentration of which has been deduced to be remarkably constant across the tissue from SC surface to below the viable epidermis) is a valid approach [13]. While this does not permit absolute quantitation of the compound of interest from the Raman data (unless a validated calibration method is available), the results are perfectly suitable for comparative purposes, including the assessment of equivalence, or not, between the rate and extent of drug delivery from two different formulations [13]. A further important step towards validation of the Raman approach has been to identify whether, in such experiments, a Raman signal measured at a particular frequency (and presumably reflecting the presence of the API) is or is not quantifiable [13,17]. This method establishes objectively a limit of detection by taking into account the standard deviation of the background 'noise' (originating from the skin and non-drug components of the applied product) found at frequencies above and below that believed to indicate the presence of the specific drug involved [17,23].

The positive progress reported to-date has been based primarily on experimental data acquired on excised porcine skin (which is accepted as an excellent model for human skin [24,25]); however, further translation of the Raman approach as a tool to provide contributory evidence for the regulatory assessment/approval of a generic drug product (bioequivalent to that of the innovator's reference formulation) requires proof that its ultimate in vivo application will be sufficiently discriminatory. Here, this issue is addressed head-on in an ex vivo study – essential before moving into human subjects – involving the measurement of the cutaneous disposition and pharmacokinetics of metronidazole (MTZ, a drug used to treat rosacea) following application of three marketed gel products and two laboratory-made formulations. One of the latter, at least, was expected to demonstrate clear inequivalence to the commercially available gels which, themselves, were considered likely to be equivalent. MTZ was selected as the drug of choice because, unlike the compounds studied in much of the literature in this area so far, its Raman spectrum lacks a 'stand-out' vibration in the silent frequency range of the skin and therefore provides a much more exacting and relevant test of the technology for most dermatological actives.

2. Materials and methods

2.1. Materials

Metronidazole (MTZ), propylene glycol (PG), other solvents and chromatography reagents were from Sigma Aldrich (Dorset, UK). Agar powder was from Merck (Darmstadt, Germany). Two approved and marketed MTZ topical gel 0.75 % *w*/w drug products (Prasco Laboratories, Mason OH, USA, NDC code 66993–962-45 and Tolmar Inc., Fort Collins CO, USA, NDC code 0115–1474-46) were purchased from WEP Clinical (Morrisville, NC, USA); the third gel (also 0.75 % *w*/w) (Galderma UK Ltd., London, UK, Marketing authorisation number PL 10590/0016) was acquired from AAH Pharmaceuticals (Coventry, UK). Fresh abdominal porcine skin from four different animals was obtained from a tissue supplier and dermatomed (Zimmer®, Hudson, OH, USA) to a nominal thickness of 750 μm within 48 h of slaughter. Visually obvious hairs were carefully cut with scissors, and the tissue was stored at −20 °C until it was thawed at room temperature shortly before use.

2.2. Metronidazole formulations

The composition of the formulations investigated are in Table 1. The gel by Prasco Laboratories is the reference-listed product (RLD) in the US and that by Tolmar is an FDA-approved generic that is bioequivalent to the RLD. The RLD product in the UK is the one manufactured by Galderma. The products are subsequently referred to as RLD-US, Generic-US and RLD-UK in the rest of the manuscript. In addition, two laboratory-made, saturated MTZ solutions were prepared using different ratios of water and PG (specifically 90/10 and 30/70 *v*/v). PG is known to influence the permeation of molecules across skin, either alone or with a co-solvent, in topical formulations [26]. The high PG levels in the solutions, together with the saturation of the drug, were designed to

Table 1
Composition of the MTZ formulations studied.

Formulation	MTZ (%*w*/w)	Excipients
Commercial gels		
RLD-US [27]	0.75	Carbomer 940, edetate disodium, methyl paraben 0.08 %, propyl paraben 0.02 %, PG, purified water, NaOH
Generic-US [28]	0.75	
RLD-UK [29]	0.75	Carbomer 940, edetate disodium, methyl paraben 0.08 %, propyl paraben 0.02 %, PG 3 %*, purified water, NaOH
Laboratory-made formulations		
90:10 v/v water:PG	~1.10 (saturated)	PG, Milli Q water
30:70 v/v water:PG	~2.75 (saturated)	

* PG is considered a functional excipient and therefore its amount is specified in the UK/EU product, in accordance with current regulatory guidelines, which require explicit disclosure of functional excipient quantities [29,30].

maximise the flux of the API and therefore to outperform the marketed products (i.e., to be inequivalent).

Equilibrium solubilities of MTZ in the 90/10 and 30/70 *v*/v water:PG mixtures were found, using the shake flask method [31], to be 11.0 ± 0.2 and 27.5 ± 0.4 mg mL^{-1}, respectively. Excess solid drug was separately introduced into glass vials containing 1 mL of formulation. The samples were maintained at 32 °C and shaken for 24 h. A sample of supernatant was then withdrawn, filtered (4-mm diameter, 0.45-μm pore size regenerated cellulose filters; SMI-LabHut Ltd., Maisemore, UK), diluted with 60/40 %*v*/v acetonitrile/phosphate buffer (pH 3) and analysed by HPLC-UV (Shimadzu LC-2010, Milton Keynes, UK) using a modified version of a previously published method [32].

2.3. In vitro assessment of MTZ uptake into and clearance from the skin

Application of either 20 mg cm^{-2} of the three gels, or 300 μL of the two solutions, was made to 2.01 cm^2 of abdominal pig skin mounted in a vertical Franz diffusion cell (PermeGear, Inc., Bethlehem, USA). MTZ uptake into the unoccluded skin was determined after 6 and 12 h in two separate experiments. The RLD-US experiments were completely replicated to provide an internal control (and are reported henceforth, when necessary, as RLD-US-R1 and RLD-US-R2). The receptor chamber of the diffusion cell was filled with 7.4 mL of pH 7.4 phosphate-buffered saline (PBS) and the cells were thermostatted in an oven at 32 °C throughout the uptake period. After 6 or 12 h, any residual formulation on the skin surface was removed with absorbent paper and the skin surface was then cleaned with alcoholic wipes (70 % isopropyl alcohol, FastAid pre-injection swabs, Robinson Healthcare, Worksop, UK). The diffusion cell was then disassembled and about 25 % of the treated skin area was subjected immediately to analysis by confocal Raman spectroscopy (see below).

The remaining portion of the skin that had been treated for 12 h was placed on a 2 % agar gel 'island' floating on 20 mL of PBS in a petri dish thermostatted at 32 °C in an oven (the gel was prepared by dissolving 1 g of agar in 50 mL of water and heating to 80 °C with continuous stirring; the solution was then cooled to room temperature forming a gel). After 2 h, a second portion, approximately 25 % of the initially treated skin was sectioned and immediately taken for analysis by confocal Raman; this procedure was repeated for the remaining skin piece after a further period of 2 h. These measurements – effectively made at 12, 14 and 16 h after the initial application of the products enabled the clearance of MTZ from the skin to be assessed (i.e., at 0, 2 and 4 h post-termination of the uptake phase, respectively). All experiments were performed in quadruplicate using skin acquired from three different pigs (resulting in $n = 12$ time points for each formulation studied).

A further set of 6-h uptake experiments were performed following skin treatment as before with the RLD-US and Generic-US gels, and with the 90/10 *v*/v water:PG solution of MTZ. In this case, post-cleaning of the skin (as above), the samples were immediately frozen on dry ice for short-term storage before their analysis by stimulated Raman scattering (SRS; see below). These experiments were performed in quadruplicate using skin from a single pig (different from those used for the confocal Raman experiments).

2.4. Confocal Raman spectroscopy

MTZ disposition in the skin following its application in different formulations was assessed using confocal Raman spectroscopy (Renishaw RM1000 Raman microscope running v4.4 WIRE software, Renishaw plc, Wotton-Under-Edge, UK) in a 'top-down' approach [13]. Skin sections post-treatment were mounted over a small (~0.5 mL) water-filled well in a custom-built sample holder to maintain tissue hydration. A diode laser (785 nm) was used with a 1200-line/mm grating providing spectral resolution of 1 cm^{-1}. The 520 cm^{-1} signal from a silicon wafer provided a calibration reference sample to correct Raman frequency shifts or laser beam de-alignment. Spectra were acquired in the frequency range 718–1803 cm^{-1}, using a long, 50× working distance objective lens (NA = 0.90), 40 accumulations (each of 10 s), and a laser power of 100 %. Reference spectra of untreated skin, solid MTZ, the commercial gels and the laboratory-made solutions are shown in Fig. 1. Signals with maximum intensities at 1192 cm^{-1} (C—N stretching) and 840 cm^{-1} (C-OH stretching) were used to track the disposition of MTZ and propylene glycol (PG), respectively, in the skin. The maximum intensities of the endogenous Amide *I* (C═O stretching at 1652 cm^{-1}; primarily from keratin) and phenylalanine (phenyl ring breathing at 1004 cm^{-1}) were also recorded. The C—N stretching band of MTZ was sensitive to the physical form of the drug providing a maximum signal at 1192 cm^{-1} from the dissolved form of the active ingredient as compared to one at 1185 cm^{-1} from the solid material (Fig. 1). 'Top-down' profiles of MTZ and PG were measured at nominal depths of 5, 10, 15, 20, 25 and 35 μm from the skin surface at one specific location, taking care to avoid skin 'crevices' and appendages (such as hair follicles); Amide *I* and phenylalanine signals were similarly recorded. Spectra from untreated skin were additionally acquired for all three pig sources in quadruplicate.

2.5. Normalisation and critical detection level of confocal Raman spectra

In the first instance, data analysis was performed manually. Baseline subtraction (intelligent polynomial fitting) and cosmic ray removal (width of features) were applied to all spectra, using WIRE software. The MTZ, PG, Amide *I* and phenylalanine bands were then fitted. Peak fitting was performed manually using the built-in function within WIRE software. Subsequently, to automate analysis of the results, data pre-processing was undertaken removing cosmic ray spikes by normalising a copy of the spectra by its maximum intensity (usually the spike, if present) and smoothing the spectra using the Savitzky-Golay filter to minimise the impact of noise. Prominent peaks with a normalised intensity exceeding 0.8 and FHWM ≤ 5 cm^{-1} were considered to be a cosmic ray spike and replaced by the linear interpolation of its immediate neighbours across a window size of 2 m + 1 where m = 10 cm^{-1}. The spectra was then baselined, and peak fitting was performed using Python software developed in-house, based upon the LMFIT library [33]. The two approaches provided, in general, excellent agreement. However, in a small subset of cases (less than 3 %), further refinement was necessary through adjustments to tolerance settings, specifically the parameter that guides whether closely adjacent peaks, initially identified via continuous wavelet transform (CWT), are merged or fitted separately, and helps to ensure accurate peak identification near the provided seed wavenumbers resulting in improved fitting outcomes. This parameter was adjusted to the range between 2 and 6 for both the MTZ and PG peaks, accommodating the intricacies of confocal Raman

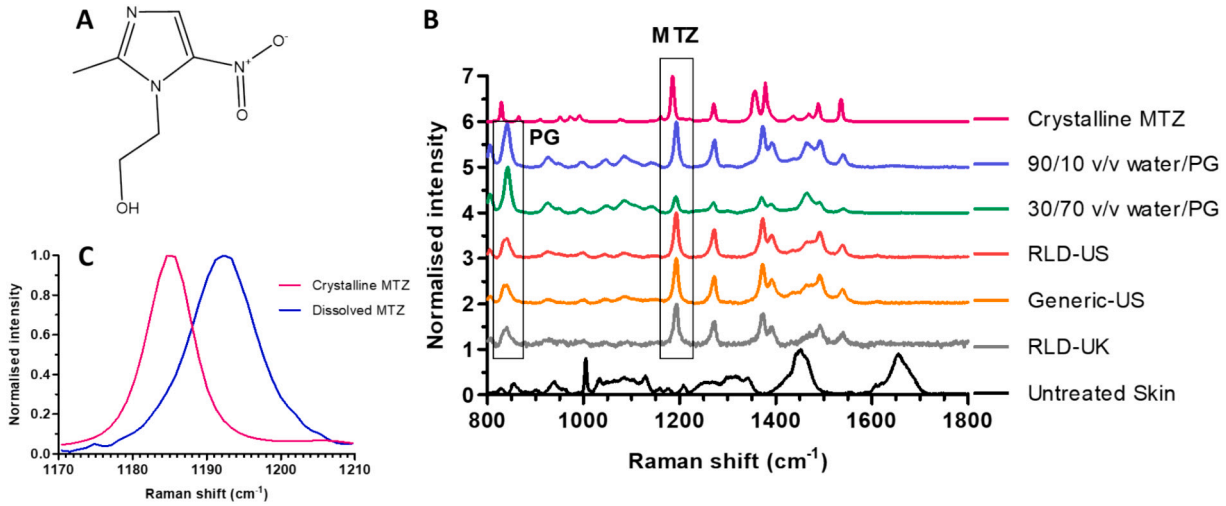

Fig. 1. [A] MTZ structure (ChemDraw Professional 18.1.0.535, PerkinElmer, Massachusetts, US). [B] Reference spectra (normalised by observed maximum intensity) of crystalline MTZ, two saturated MTZ solutions in water:PG, three commercial gels, and untreated skin. [C] Raman C—N stretching band from dissolved and solid MTZ (again normalised by maximum intensity).

spectral resolution. Other parameters, including the peak threshold (maintained at 0.1), were unchanged. The total counts in each spectrum were calculated by summing all intensities in the recorded spectral window. The standard deviation of the background noise (SD_{noise}) in a fixed spectral range (1720–1800 cm^{-1}), in which no drug or skin signals were expected, was calculated. In a few cases where cosmic ray spikes were present in this range and could not be removed, the SD_{noise} was calculated in a smaller frequency window. The critical (i.e., smallest) level of detection (A_c) was then set as (3 x SD_{noise}) [34]. The workflow involved is summarised in **Fig. S1**. A_c values were determined for every skin sample, i.e., for each replicate experiment with each MTZ formulation at each time and depth in each pig. Of course, it is impossible to say with certainty that any signals less than A_c correspond to the complete absence of the chemical and replacing non-detectable values with zero was therefore considered inappropriate. Instead, MTZ and PG signals below the A_c were replaced by $A_c/2$ for that skin sample, a conservative estimate of the chemical's intensity that assumes that the analyte may be present in the tissue [35,36]. The 0 h data were replaced with a zero value, as the skin was not treated with MTZ.

Normalised Raman intensities were calculated by dividing the maximum signals of interest (MTZ, PG and phenylalanine) by that of Amide *I* in each spectrum to correct for signal attenuation with depth; the Amide *I* was used as an internal standard because its concentration is reasonably constant with skin depth [13]. The ratio of Amide *I* to total counts was also examined in experiments of untreated skin. Data were plotted in 2D as function of measurement time using GraphPad Prism 5 (version 9.3.1, San Diego, CA) and in 3D as functions of both skin depth and time using OriginPro (2018, Massachusetts, USA).

2.6. Determination of cutaneous BA/BE metrics from confocal Raman data

Because the 6 h uptake data for each pig were from four diffusion cell experiments that were different to those from which the 12 h uptake and 2 and 4 h clearance data were acquired, it is inappropriate to calculate the area under the normalised intensity vs time of measurement curve (AUC_t) for both drug absorption and clearance in the same skin sample at each depth. Instead, the arithmetic averages of the normalised intensities of the four replicates at each time of measurement (6, 12, 14 and 16 h) and depth (5, 10, 15, 20, 25 and 35 μm) were calculated for each product in each pig. Then, the AUC_t at each depth for each product in each pig was calculated using the trapezoidal method. In addition, the volume under the normalised intensity vs time of measurement and depth into the skin surface (VUS) for each product in each pig was calculated with Eq.1:

$$VUS = \sum_{n=1}^{Nt}\sum_{i=1}^{Nd}\frac{NI_{d_i,t_n} + NI_{d_{i+1},t_n} + NI_{d_i,t_{n+1}} + NI_{d_{i+1},t_{n+1}}}{4} \times (d_{i+1} - d_i) \times (t_{n+1} - t_n) \quad (1)$$

where *NI* is normalised intensity, d_i is the i^{th} depth into the skin (μm), t_n is the n^{th} time of measurement (h), *Nt* is the total number of time points, and *Nd* is the total number of depths.

AUC_t and VUS values were natural log-transformed and the differences between these values for the test and reference (chosen to be RLD-US-R1) formulations in each pig were calculated, providing three results for each comparison (i.e., one per pig). The arithmetic average, standard deviation and the upper and lower 90 % confidence intervals (CIs) of these three results were calculated. Finally, the means and the upper and lower 90 % CIs were anti-logged to enable determination of the BE ratio and CI of the geometric means.

Normalised MTZ signals were measured at six depths within each skin sample at each of four experimental time points, permitting calculations of the area under the normalised intensity versus depth into the skin curve (AUC_z), again with the trapezoidal method, for each replicate in each pig. The within-pig standard deviation in the AUC_z measurements from the reference product (s_{wr}) was assessed for the natural log-transformed data as follows:

$$s_{wr} = \sqrt{\frac{1}{n(nr-1)} \times \sum_{j=1}^{n}\sum_{k=1}^{nr}\left[\ln(AUC_{z,jk}) - \ln(\overline{AUC}_{z,j})\right]^2} \quad (2)$$

where $AUC_{z,jk}$ is the k^{th} replicate in pig *j*, and $\overline{AUC}_{z,j}$ is the geometric mean of *nr* replicates (*n* = number of pigs, *r* = number of replicates). If s_{wr} is less than or equal to 0.294, the traditional average BE (ABE) can be used to assess BE. However, if the s_{wr} is greater than 0.294, indicating high data variability, the reference scaled average bioequivalence (SABE) assessment is recommended [6,37].

2.7. Stimulated Raman scattering microscopy

SRS imaging was also performed in a 'top-down' configuration on a Leica SP8 laser scanning microscope (Leica Microsystems, Wetzlar, Germany) coupled with a PicoEmerald-S laser system (APE, Berlin,

Germany). Samples were removed from dry ice and the tissue block was immediately cut horizontally with a scalpel to create a thin slice (100–200 μm containing the SC and part of the viable skin). An imaging spacer of double-sided adhesive tape (Grace Bio-Labs, SecureSeal™, Sigma Aldrich, Burlington, USA), into which a circular opening had been cut, was centrally positioned on a glass cover-slip, thickness #1.5 (0.16–0.19 mm thick, 26 × 76 mm, Menzel Glasser, ThermoScientific, Massachusetts, USA). The skin slice was placed in the opening, with the SC in direct contact with the glass, and was then overlaid with another cover-slip of the same size (minimising tissue dehydration and preventing sample movement during the experiment) [18].

During imaging, the samples were cooled to and maintained at 10 °C on a temperature-controlled stage (Okolabs s.r.l., Pozzuoli, NA, Italy), permitting a relatively high laser power (60 %) to be used to boost signal intensities, while minimising heating and sample dehydration. The sample was placed and secured on the temperature-controlled stage with two magnetic clamps to prevent movement and flexing of the glass cover slips. SRS images were acquired on a laser scanning microscope, as described previously [18] and were performed in the forward direction, with signals collected through a 25×/0.95 numerical aperture (NA) water immersion objective lens and an air condenser lens (Leica S1 50515) with a NA of 0.90. Second harmonic generation (SHG) was also acquired as reported previously [18].

The pump beam was tuned for SRS contrast at 1195 cm^{-1} and 1666 cm^{-1} for the on-resonance signals of MTZ (C═N stretching) and Amide *I* (C═O stretching), respectively. Spurious signals were identified and removed by acquiring and subtracting the corresponding off-resonance frequencies at 1181 and 1723 cm^{-1}, respectively. The off-resonance frequency for MTZ also served to eliminate any contribution of overlapping endogenous skin signals to that from MTZ (which results in negative grey scales when MTZ is not present); while this may have, also, eventually removed part of the drug signal, it ensured that any signal observed at 1195 cm^{-1} reflected MTZ alone. Complementary SHG imaging was performed at 437.75 nm (collection wavelength) correlatively with the imaging at 1723 cm^{-1} (875.50 nm excitation wavelength) with a Semrock 414/16 nm BrightLine single-band bandpass filter (Idex Health and Science, New York, USA). The SRS silicon detector gain settings were optimised at 6.5 *V* for MTZ and Amide *I*, while a gain of 726.6 V was applied to the photomultiplier tube for the SHG collection. This value was selected using a skin sample with the highest MTZ uptake (saturated solution in 90/10 *v*/v water:PG) and consistently applied to all samples to prevent signal saturation and to allow for comparison between treated and untreated samples. 3D image stacks were acquired using a depth increment of 1 μm on single tile areas (465 × 465 μm), again excluding skin crevices and appendages.

SRS image analysis was performed in ImageJ [38]. The "top-down" images for each signal were first stacked as a function of depth. Small translational drifts in the sample position were observed as different channels were acquired sequentially, but image registration was not performed for reasons explained in the literature [18]. The skin surface was identified using a previously published approach [18] and the grey scale values of off-resonance images were subtracted from those of on-resonance for MTZ and Amide *I* using the image calculator tool. Consistent display settings (such as colour window level) were applied to MTZ and Amide *I* signals to enable comparison between applications of the different formulations. The processed images were then colour merged. For further data analysis, on- and off-resonances were colour-stacked in one tile and SRS signal intensities were assessed from the average grey scale values of each component in 12 lines (20-pixel length) as function of depth for each map. Grey scales were graphically displayed as function of depth, and any drifts observed in the 12 lines examined were adjusted accordingly by monitoring the maximum rate of change of each component. Normalised MTZ grey scales were calculated by dividing the off-resonance corrected MTZ signal by the corresponding Amide *I* value. SRS data were plotted as function of depth in GraphPad Prism 5. The AUC_z was calculated using the trapezoidal method.

3. Results

3.1. Normalisation of confocal Raman signals as a function of skin depth

The utilisation of a novel methodology for evaluating the cutaneous BA of a topically applied drug – and, by extrapolation, the BE between generic and RLD topical drug products – demands appropriate validation. This is particularly true for Raman spectroscopy as described here because normalisation of the detected drug signal as a function of skin depth is essential to correct for attenuation by absorption and scattering [13]. A key question to answer, therefore, concerns the consistency and variability of the Amide *I* signal between different skin 'donors'.

The approach chosen to address this issue was to characterise the confocal Raman profiles with depth of not only the Amide *I* signal, from the three animals which provided skin for this study, but also that of phenylalanine (which has a strong Raman shift at 1004 cm^{-1}) and of the total counts in the chosen spectral range. The results of these experiments in untreated skin samples ($n = 4$ from each of 3 pigs) are presented in Fig. 2. The profiles of Amide *I*, phenylalanine and total counts appear qualitatively very similar and, for each profile, the agreement between the data from different pigs is also very good. Furthermore, when the phenylalanine signals are divided by the corresponding Amide *I* values, or the Amide *I* by the total counts, the derived ratios are remarkably constant across the skin depths interrogated.

3.2. Confocal Raman spectroscopy

Representative spectra of skin treated with each formulation are presented in **Fig. S2**. Normalised MTZ maximum intensities as function of time following skin treatment with three commercial gel products (including a replicate of the RLD) and two saturated solutions of the drug in different water:PG mixtures are shown in Fig. 3. The 0 h values are from untreated skin, 6 and 12 h results report on drug uptake, and 14 and 16 h on drug clearance.

The initial experiments performed did not include measurements at a skin depth of 20 μm; as a result, such data are missing (only for skin samples from pig 1) from untreated controls and from those treated with the two laboratory-made solutions. The maximum Amide *I* intensities from the entire series of MTZ experiments are displayed in **Fig. S3**. For MTZ, maximum intensities below A_c (which was ~200–400 a.u.) were replaced by $A_c/2$ before normalisation with Amide *I*. Other approaches were considered for treating these non-quantifiable values, including their replacement by either zero, or A_c, or $A_c/\sqrt{2}$. While different alternatives did have an impact on the absolute values of the cutaneous PK metrics (and their CIs) that were eventually deduced from the results (see below), there was no impact on the overall conclusions that could be drawn. Therefore, while using $A_c/2$ may not necessarily be the most rigorous approach to handle 'non-detects' from experiments such as those presented here, it has a practical *v*alue. An objective assessment of alternative solutions – while beyond the scope of the present study – merits further, careful consideration [35,36].

The MTZ uptake-clearance profiles in Fig. 3 show that drug was detectable at depths of up to 35 μm; that is, well beyond the stratum corneum (SC) layer of the skin. In general, and entirely consistent with expectations, the MTZ signals attenuated with increasing depth due to both the effects of absorption and scattering and the drug's concentration gradient across the skin. For the gels and the 30/70 *v*/v water:PG solution, MTZ uptake increased from 6 to 12 h; in contrast, for the 90/10 v/v water:PG solution, the change measured between 6 and 12 h was less clear, with the 6 h uptake being noticeably higher than that for all other formulations. Post-removal of the gels and solutions at 12 h, MTZ levels at all skin depths were lower at 14 h and had then decreased further at 16 h as the drug cleared from the skin. The overall PK profiles of MTZ delivered from the three gels and from the 30/70 *v*/v water:PG

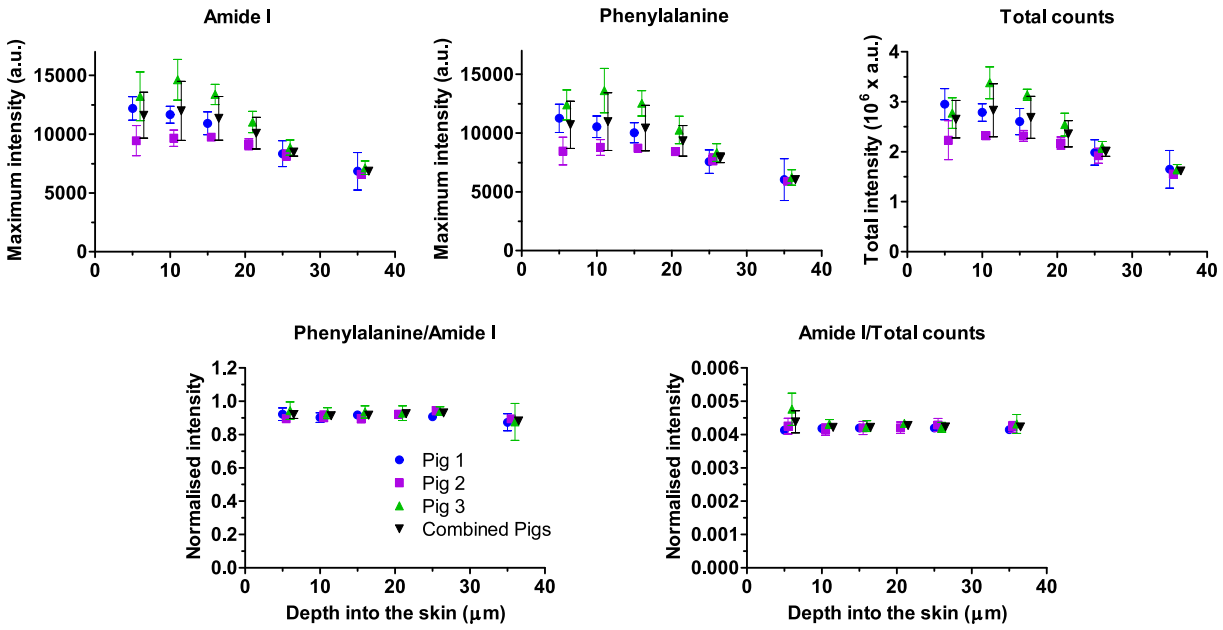

Fig. 2. Amide *I* and phenylalanine maximum intensities, total counts in each spectrum, phenylalanine/Amide *I* and Amide *I*/total counts ratio as function of depth for untreated skin from three different pigs ($n = 4$ replicates per pig) and their average. Data are means ± standard deviation (SD); some data points have been shifted on the x-axis to facilitate visualisation.

solution were therefore similar, while that from the 90/10 v/v mixture was different, primarily due to the enhanced uptake of the drug apparent at 6 h. This distinction was apparent up to a skin depth of 25 μm (Fig. 3).

The disposition of PG, an excipient present in all the studied formulations, was tracked simultaneously with that of MTZ. While the PG signals from the gels were very low and close to values (at 840 cm^{-1}) recorded from untreated skin (data not shown), those from the two laboratory-made solutions were measurable at depths to 35 μm as shown in Fig. 4. A clear differentiation in the uptake of PG was observed between the 90/10 and the 30/70 v/v water:PG solutions: for the former, PG uptake peaked at 6 h at all skin depths but had decreased following a longer uptake period of 12 h; for the latter, on the other hand, while PG levels detected in the skin matched those from the 90/10 solution at 6 h, uptake continued to increase when the application was extended to 12 h, presumably a reflection of the greater PG content available for absorption from the 30/70 solution. Once uptake was terminated at 12 h, PG signals, independent of the solution applied, fell over the next 4 h as the cosolvent cleared from the skin.

3.3. Bioequivalence (BE) analysis of confocal Raman data

The AUC_t ratios (i.e., 'test'/RLD) with the corresponding 90 % CIs at each skin depth considered are presented in Fig. 5; the specific AUC_t values are in **Table S1**. The RLD was arbitrarily chosen to be RLD-US-R1 and the 'test' formulations were the RLD-US-R2, Generic-US and RLD-UK gels, and the two laboratory-made water:PG solutions of MTZ. In general, 'test' and RLD treatments may be considered to be BE *"if the 90% confidence interval of the ratio of a log-transformed exposure measure (AUC_t in this case) falls completely within the range 0.8-1.25"* [39]; this range is included in each panel comprising Fig. 5. Apart from the 35 μm data (for which the MTZ signal was often non-quantifiable), Fig. 5 shows that the mean 'test'/RLD ratios (BE ratios) for the three 'test' gels fell within 0.8–1.25, whereas those for the solutions generally did not (the 30/70 at 20 μm did but with a very large CI). However, in no instance did the 90% CIs remain within 0.8–1.25.

The AUC_z values for each of the 4 replicates (of each porcine skin donor) were also determined at the two uptake (6, 12 h) and two clearance (14, 16 h) times (Table 2). The within-pig standard deviation of the RLD-US-R1 exceeded 0.29 at all skin depths, indicating high variability, for which the SABE assessment is recommended [40]; detailed explanations of SABE calculations have been previously published [37]. The results of this analysis for the data acquired here are summarised in Table 2.

The VUS is an alternative metric to AUC_t and AUC_z that simultaneously captures temporal and spatial data, yielding a single numerical value with which to compare the gels and solutions. As the inter-pig variability in the MTZ normalised intensities was relatively small, product comparisons were made across the n = 12 set of measurements for this analysis, rather than an assessment animal-by-animal. The calculated VUS values, and the BE ratios (and 90% CIs) are in Table 3.

3.4. Stimulated Raman scattering (SRS) microscopy

The average intensity grey scales of Amide *I* and MTZ signals determined by SRS, and the MTZ/Amide *I* normalised ratio, after a 6 h application of the RLD-US and Generic-US gels, and of the saturated MTZ solution in 90/10 *v*/v water:PG, are shown in Fig. 6A, together with the corresponding data from untreated skin. Representative, tiled SRS images from the same experiments are shown in Fig. 6B and videos of these depth scans are in **Movies S1**. As for the confocal Raman data discussed above, normalisation of the MTZ signal by that of Amide *I* at each skin depth permits correction for signal attenuation so that the relative concentration profile of the drug is correctly revealed. The lower Amide *I* intensities near the skin surface post-treatment with the laboratory-made solution may be attributed to the reported optical clearing effects of PG [41]. Qualitatively, the data in **Fig. 6A** and **B** confirm that MTZ uptake into the skin is higher from the solution compared to the two gels, between which no difference is observed.

The normalised MTZ grey scale intensities were positive for the two gels only to skin depths of 7–8 μm. In contrast, for the water:PG solution, these values remained positive up to 35 μm into the skin. The areas under the normalised grey scale intensities versus skin depth (AUC_z in units of μm, mean ± SD, $n = 4$) from 0 to 10 μm were: −0.34 ± 0.50 for untreated skin, 0.22 ± 0.38 for the RLD-US gel, 0.23 ± 0.18 for the Generic-US product, and 2.11 ± 0.51 for the 90/10 solution. AUC_z for the solution was significantly greater than those of both untreated skin and of the two gels (2-way ANOVA followed by Tukey's test, $p < 0.05$); for the two gels, the AUC_z values were not significantly different either

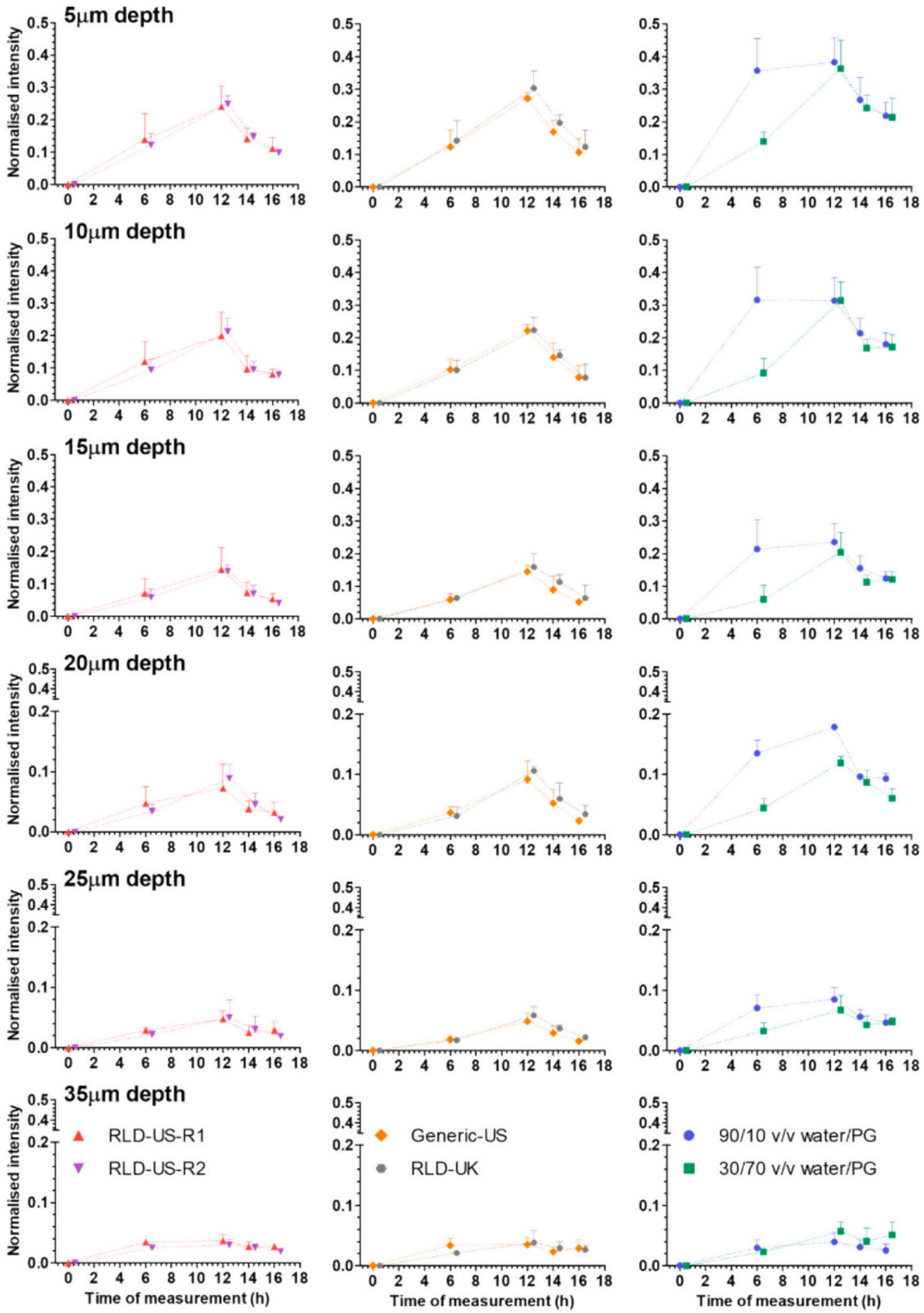

Fig. 3. Amide *I* normalised MTZ maximum Raman intensities as functions of time and skin depth. Data are the mean of the average normalised intensities ($n = 4$) in each pig + SD ($n = 3$ pigs, except for the two solutions for which no data were collected at 20 μm for pig 1, and for RLD-US R1 for which the result at 10 μm at 6 h in pig 1 is missing). To facilitate visualisation, some data points have been shifted on the x-axis, and the y-axis in the last three rows are segmented.

from the untreated control or from each other.

Finally, Fig. 6C shows that, as with confocal Raman, SRS also detected the presence of solid MTZ on the skin surface post-application of the 90/10 *v*/v water/PG solution; the SRS peak frequency of MTZ showed a clear distinction between the dissolved (1195 cm^{-1}) and solid forms (1188 cm^{-1}) of the drug.

4. Discussion

The application of different vibrational spectroscopic tools, including Raman in particular, to the assessment of drug disposition and PK in the skin has been well-documented ex vivo and in vivo over recent years [13,42]. However, the field has yet to overcome two important

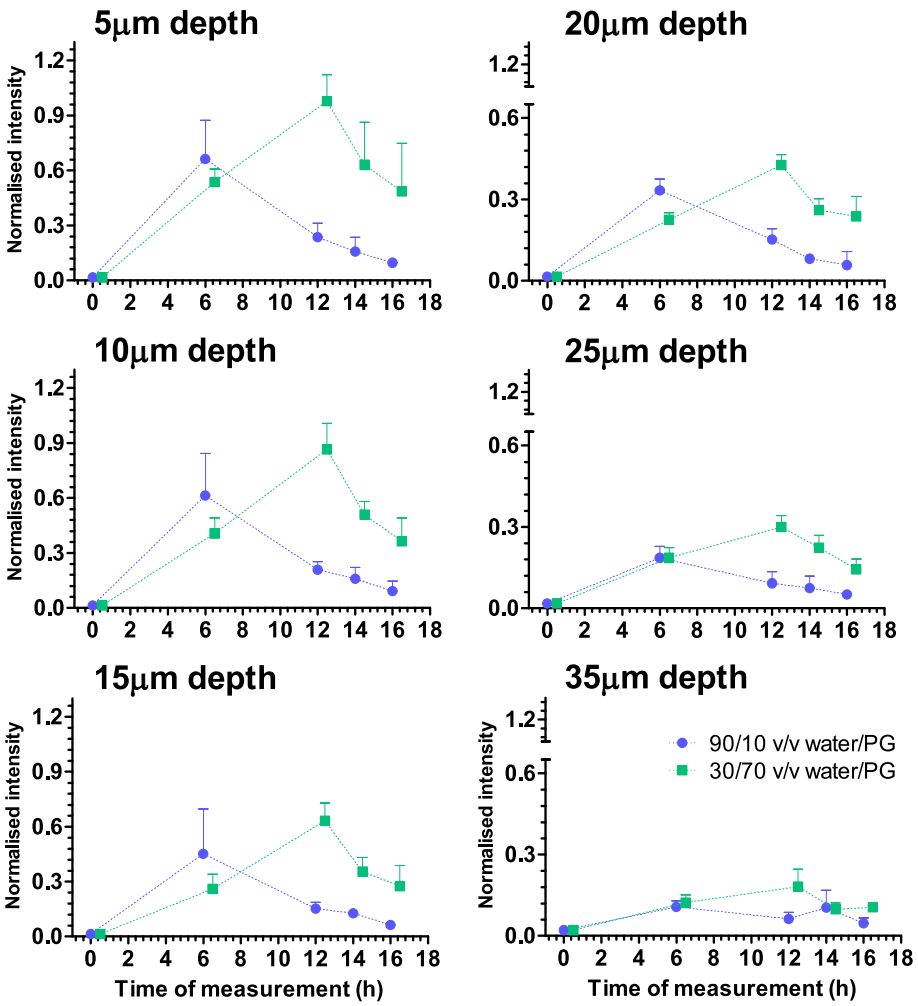

Fig. 4. Amide *I* normalised PG maximum Raman intensities as functions of time and skin depth following treatment with the two laboratory-made solutions of MTZ in water:PG mixtures. Data are the mean of the average normalised intensities ($n = 4$) in each pig + SD ($n = 3$ pigs, except that no data were collected at 20 μm for pig 1). To facilitate visualisation, some data points have been shifted on the x-axis, and the y-axis in the last three depths are segmented.

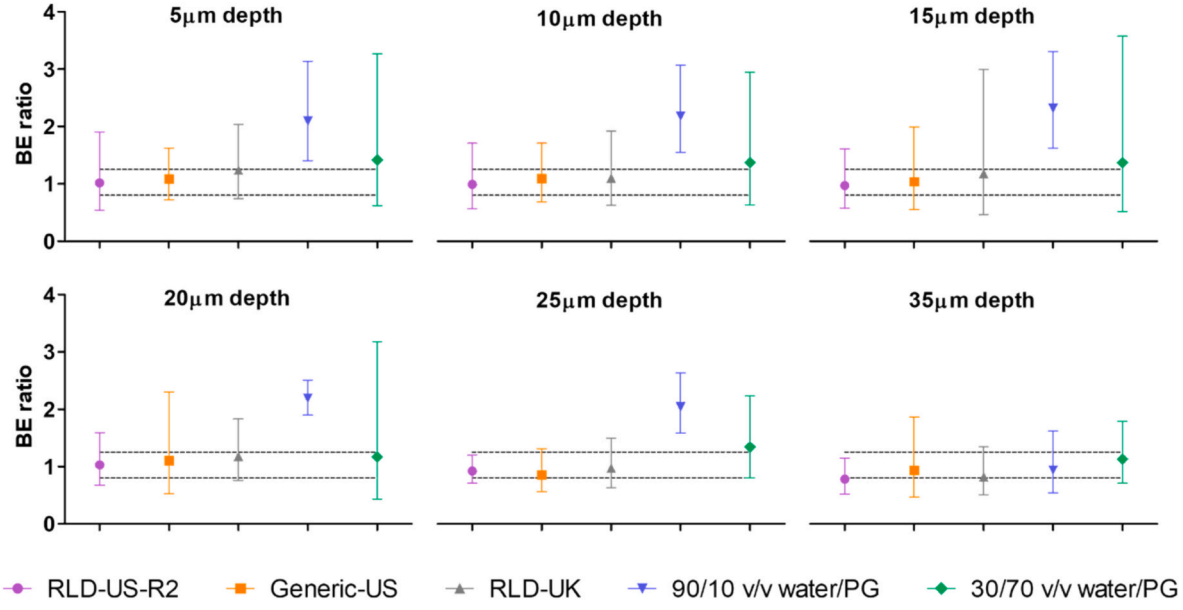

Fig. 5. Mean BE ratios ('test'/RLD-US-R1) based on AUC_t values ±90% CIs at each skin depth. Dotted lines indicate the 0.80–1.25 window. As no data were collected at 20 μm for the two solutions in pig 1, the BE comparison in this case only includes results from pigs 2 and 3.

Table 2
Scaled average bioequivalence (SABE) assessments between the test and reference (RLD-US-R1) formulations for AUC_z at two uptake and two clearance times. Formulations are considered not to be different if the BE ratio is between 0.80 and 1.25, and the upper bound on the scaled average 95 % confidence interval (SCI_{UB}) is less than zero. The 90 % lower and upper bound values of the traditional average bioequivalence (ABE) assessment are also presented. For comparisons between the RLD-US-R1 and the two MTZ solutions, the 20 μm depth data were excluded from all pigs. Data are the mean of the average AUC_z ($n = 4$) in each pig ($n = 3$) ± SD.

RLD-US-R1		Test formulation	AUC_z (μm)	BE ratio	SCI_{UB}	90% lower CI	90% upper CI
AUC_z (μm)	s_{wr}						
6 h uptake							
1.91 ± 0.93	0.63	RLD-US-R2	1.52 ± 0.43	0.91	0.21	0.49	1.67
		Generic-US	1.60 ± 0.45	0.98	−0.13	0.70	1.37
		RLD-UK	1.57 ± 0.33	0.93	−0.04	0.61	1.40
1.93 ± 0.93	0.61	90/10 v/v water:PG	4.92 ± 1.60	2.85	2.30	1.63	5.00
		30/70 v/v water:PG	1.69 ± 0.74	0.98	−0.14	0.76	1.27
12 h uptake							
3.22 ± 1.20	0.73	RLD-US-R2	3.35 ± 0.63	1.08	0.07	0.59	1.97
		Generic-US	3.51 ± 0.47	1.13	0.30	0.55	2.31
		RLD-UK	3.82 ± 0.69	1.32	0.26	0.78	2.24
3.34 ± 1.20	0.69	90/10 v/v water:PG	5.33 ± 1.20	1.89	0.52	1.40	2.55
		30/70 v/v water:PG	4.96 ± 1.14	1.62	1.31	0.71	3.67
2 h clearance							
1.71 ± 0.48	0.77	RLD-US-R2	1.77 ± 0.54	1.17	−0.10	0.79	1.74
		Generic-US	2.16 ± 0.72	1.39	0.03	1.00	1.93
		RLD-UK	2.50 ± 0.41	1.61	0.07	1.31	1.97
1.76 ± 0.50	0.79	90/10 v/v water:PG	3.61 ± 0.87	2.35	0.56	2.07	2.66
		30/70 v/v water:PG	2.92 ± 0.50	1.92	0.36	1.52	2.42
4 h clearance							
1.46 ± 0.41	0.38	RLD-US-R2	1.18 ± 0.08	0.81	0.38	0.50	1.31
		Generic-US	1.30 ± 0.35	0.88	0.43	0.48	1.64
		RLD-UK	1.48 ± 0.59	0.91	0.85	0.37	2.21
1.50 ± 0.40	0.39	90/10 v/v water:PG	2.96 ± 0.60	2.02	0.79	1.57	2.60
		30/70 v/v water:PG	3.03 ± 0.63	1.90	0.42	1.74	2.08

challenges, namely, (a) low sensitivity – that is, being able to detect the drug against the skin background and interference from co-applied excipients – and (b) the ability to measure drug presence at or near the site of action, typically below the SC in the viable epidermis. To-date, the value of the Raman approach has typically been demonstrated using model compounds or deuterated species, and a few drug molecules which possess C≡C and C≡N functionalities that have strong Raman vibrational signals in a very silent region of the skin spectrum [13]; without this advantage, the detection of a molecule delivered to depths in the skin below the SC has been quite limited. This investigation aimed to demonstrate that the identified challenges are surmountable with a careful experimental methodology and a rigorous data analysis approach.

Table 3
VUS values of the MTZ gels and solutions (mean values calculated from the average (± SD) normalised intensities ($n = 4$) in each pig ($n = 3$) at all measurement times and depths. BE ratios of the VUS (and 90 % CIs) provide comparisons between the test and reference (always RLD-US-R1) formulations.

Test formulation	VUS (μm. h)	BE ratio	ABE 90 % lower CI	ABE 90 % upper CI
RLD-US-R1	29.3 ± 9.8[a]	–	–	–
RLD-US-R2	27.2 ± 3.1	0.97	0.67	1.51
Generic-US	29.3 ± 3.8	1.04	0.64	1.68
RLD-UK	31.2 ± 2.9	1.11	0.64	1.93
90/10 v/v water: PG	61.1 ± 15.3[b]	2.08	1.72	2.51
30/70 v/v water: PG	38.9 ± 5.1[b]	1.35	0.65	2.79

[a] VUS = 30.0 ± 10.0 with the 20 μm depth data excluded.
[b] 20 μm depth data excluded.

MTZ was chosen as a drug with moderate, but not exceptional, skin permeability and with a Raman spectrum with all notable features lying within the finger-print region of that of the skin. Specific vibrational frequencies associated with MTZ were identified, particularly C—N stretching at 1192 cm^{-1} and C═C at 1535 cm^{-1}. The former was selected due to its higher intensity. Selection of MTZ also provided the study with access to both reference-listed gels in the UK and US, and to a US-approved generic product; the latter two formulations are expected to be bioequivalent (Table 3). These were supplemented by two laboratory-made solutions, one of which at least was designed to be non-BE.

The first element of the study design was to validate earlier conclusions (which applied to both confocal Raman and SRS) [13,18] that the measured Amide *I* maximum intensity in the experiments performed was an appropriate signal with which to normalise those from both the drug and the key functional excipient (PG) and to account for spectral attenuation with increasing depth into the skin (due to absorption and scattering of the radiation). This step was important to confirm that any variability in the Amide *I* signal is instrumentation-related (which would therefore have an equal impact on the drug signal as well) and is not caused by significant differences in Amide *I* "concentration" – that would be due to inconsistency in the keratin level – within different skin samples or sources. This question was addressed in replicate experiments on untreated skin in which, as a function of skin depth, signals from both Amide *I* and phenylalanine (which has a strong signal at 1004 cm^{-1}), as well as the total counts over the entire spectral range examined, were recorded simultaneously. The results showed that the signal intensity ratios of [phenylalanine/Amide *I*] and [Amide *I*/total counts] were remarkably constant with increasing skin depth (Fig. 2). Phenylalanine is a constituent amino acid of keratin, and there is also a 'free' component present in the so-called natural moisturising factor of the stratum corneum [43]. The fact that the ratio of their signals - in addition to the ratio of Amide *I* to total counts - both remained constant with very limited variabilities from the SC surface to the upper viable epidermis, lends strong support to the contention that any Amide *I* intensity variability is due to instrumental factors. It is noted that, although it might appear that total counts could be an alternative option for normalisation, this would not be preferred for experiments in which the skin is treated with a product. In this scenario, components of the formulation (drug and excipients) will be taken up into the skin and potentially contribute significantly to the total counts, undermining its potential value, therefore, for normalisation.

The core objective of the research reported here was to demonstrate the potential of Raman spectroscopy in general, and the confocal Raman approach in particular, to assess the equivalence or lack thereof between complex drug products intended for topical application to the skin.

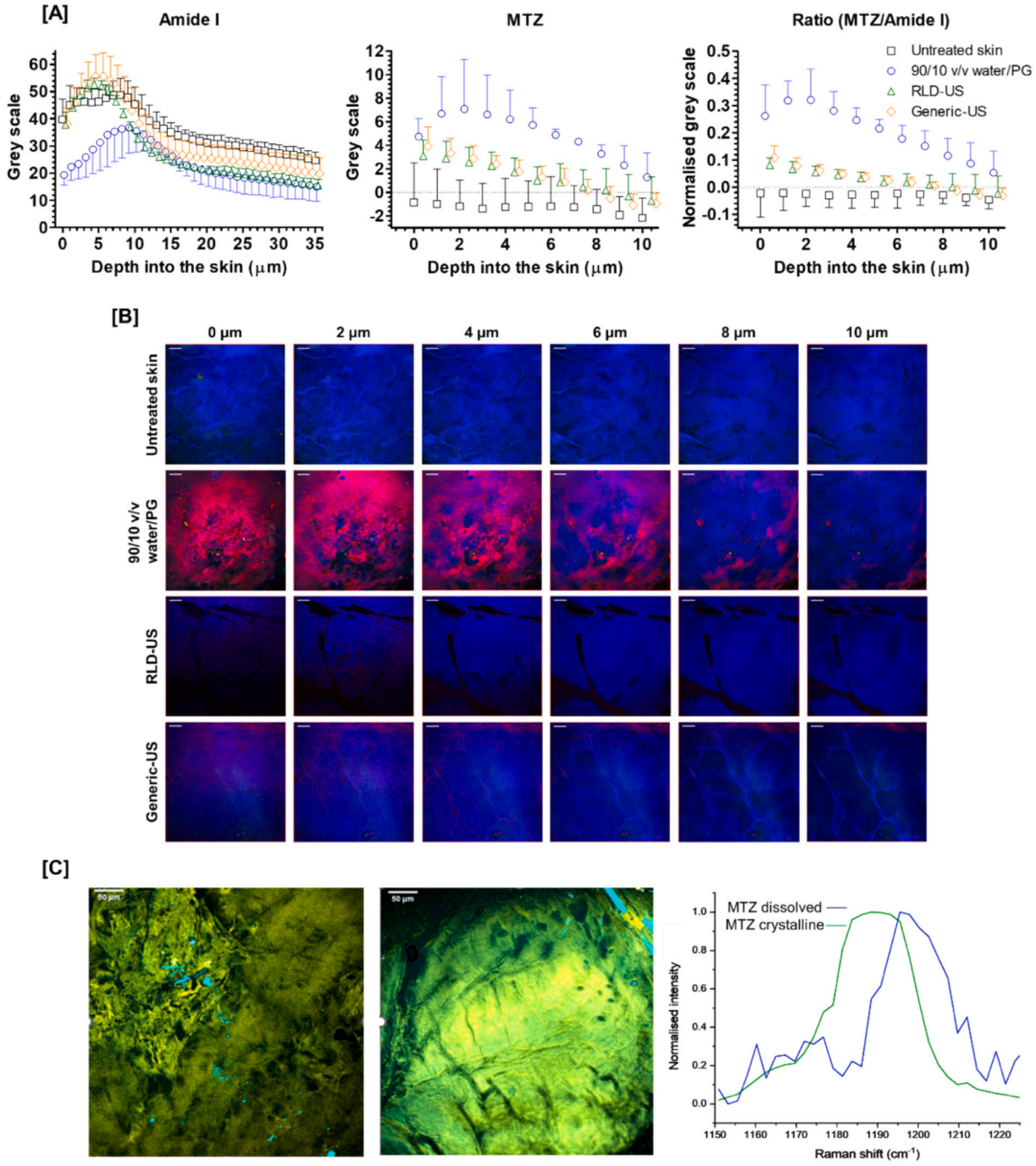

Fig. 6. **[A]** Amide *I* and MTZ signals, and their normalised ratios, measured "top-down" by SRS following 6 h application of a fully saturated solution of MTZ in 90/10 v/v water:PG (blue circles), RLD-US (green triangles) and Generic-US (orange diamonds) gels, and untreated skin (white squares). Data points are the average (+/-SD) from four different skin samples from the same pig source (different from those used for confocal Raman experiments) per formulation, with each sample providing the average of 12 selected lines (20 pixels). **[B]** Tiled SRS images acquired from untreated skin and after the 6 h application of the three MTZ formulations. Superimposed SRS signals from MTZ (red) and Amide *I* (blue) are shown as a function of depth into the skin; scale bar = 50 μm. **[C]** Left and middle panels: "Top-down" SRS images acquired after application of the 90/10 MTZ solution for 6 h. Superimposed SRS signals (normalised by their maximum intensity) from MTZ crystals (cyan) and Amide *I* (yellow) are shown; scale bar = 50 μm. Right panel: SRS-assessed frequencies of the characteristic C—N stretching band for the dissolved and solid MTZ detected on the skin surface. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Three commercial MTZ gels and two laboratory-made solutions were included in the study. The gels included the RLD-US, experiments with which were duplicated as an internal control, a FDA-approved generic product, and a third that is the RLD in the UK. Following validation of the procedure to normalise the Raman signals, cutaneous uptake and clearance PK profiles of MTZ post-application of the three gels were acquired as functions of both time of measurement and skin depth (Fig. 3). Analyses of these results, either in terms of areas under the normalised Raman signal versus time curves (AUC_t) at different depths (Fig. 5), or the areas under the signal versus skin depth curves (AUC_z) at different times (Table 2), revealed – given the variability in the data – that no differences between the gels could be discerned, suggesting (but

not proving) their equivalence. Indeed, the mean BE ratios based on AUC_t of the 'test' products (RLD-US-R2, Generic-US, RLD-UK) to the reference (RLD-US-R1) all fell in the acceptable window of 0.8 to 1.25. However, the 90 % confidence intervals on these ratios mostly extended beyond this range and hence the results could not be concluded to establish BE in a formal sense; the same overall deduction could be reached based on the SABE analysis of the AUC_z results (Table 2). This is perhaps not too surprising given that, (a) the practical resources (time, personnel) were limited, (b) the variability typically associated with percutaneous absorption measurements of any kind is significant, and (c) this proof-of-principle study was not sufficiently powered with only three skin 'donors'. It does not seem unreasonable, though, to anticipate that, with an appropriate number of replicates, the three gels would be found BE and (as further discussed below) at least the 90/10 *v*/v water: PG solution would not. The estimated number of pigs required to achieve a statistical power of at least 80 % for the SABE assessment of AUC_t values between the 'test' products and RLD-US-R1, assuming they are BE, is at least 12 for RLD-US-R2, 13 for Generic-US and 16 for RLD-UK (details provided in the Supplementary Information).

The laboratory-made MTZ solutions were included in the study with the expectation that they would behave differently (with respect to cutaneous PK) than the commercial gel products. The results acquired with both confocal Raman (Figs. 3 and 5) and (to a more limited extent) SRS (Fig. 6) show that this was certainly the case for the 90/10 v/v water:PG solution, while that for the 30/70 was less striking. Either way, the mean BE ratios in these cases mostly fell outside the 0.8–1.25 window, as did the associated 90 % CI values (Fig. 5). Of course, these solutions likely contain more PG than the gels – the RLD-UK product is known to contain 3 % by weight [29]. That the 30/70 solution sustains MTZ delivery for a longer period is not surprising as there is a larger source of PG available to diffuse into and across the skin and to enhance and prolong drug permeation.

For the 90/10 solution, on the other hand, the rapid depletion of PG via uptake into the skin (Fig. 4), coupled with evaporation of water from the skin surface, means that the composition of this formulation is changing quite dramatically post-application - a process often referred to as product 'metamorphosis' [44]. The probable consequence of these events is that MTZ uptake into the skin will proceed rapidly at first but, with PG depletion and water loss by evaporation (and also its movement into the skin), there will inevitably come a time when there are insufficient amounts of the cosolvents to keep the remaining drug fully dissolved and precipitation will occur (even, perhaps with a transient period of supersaturation as has been postulated before [44]). The results, both in terms of cutaneous PK profiles and microspectroscopic observations (confocal Raman and SRS imaging) fully support this contention.

The confocal Raman results from MTZ application in the two solutions were further analysed to track the cutaneous disposition of PG. The data (Fig. 4) reinforces the above conclusions showing that PG was taken up quickly from the 90/10 solution after 6 h but that its presence in the skin had then fallen away after 12 h. In contrast, from the 30/70 solution, PG uptake matched that from the 90/10 at 6 h but then increased further when the application continued to 12 h. The local PK of MTZ mirror rather closely this behaviour of the functional excipient.

Returning to the broader issue addressed in this work about the value of Raman spectroscopy as a potential tool to assess BE between complex drug products intended for skin application, some positive information has emerged. To begin with, it has been possible using confocal Raman to produce comparative, cutaneous PK profiles of different formulations of a drug. From these data, attention has first focused on two area-under-the-curve metrics: AUC_t and AUC_z, the subscripts indicating with respect to time and skin depth, respectively. The former is immediately recognizable from conventional (and typically oral) assessments of BA and BE, but here offers a metric for comparison at several nominal skin depths that were interrogated at each time point. The particular advantage afforded by this extra 'dimension', of course, is that the disposition of the drug can be examined in different 'compartments' of the skin – such as within the SC, or below this layer in the upper, viable epidermis – that may be in more or less proximity to the site of pharmacological action. The AUC_z offers a different perspective, relevant to the point just made, and assesses the cutaneous drug availability at specific depths into the skin at different times of uptake and clearance. In this way, attention might be focused on whether a 'test' product can mirror the RLD in terms of the achieved residence time for the drug in a specific location within the skin.

A formal SABE analysis of the AUC_t results was considered inappropriate because the 6-h uptake data were derived from different diffusion cells (i.e., from different skin sections, even though they were from the same pigs) than the 12-h uptake and 14- and 16-h clearance measurements. Nonetheless, for illustrative purposes, it is possible to undertake such an analysis by assuming that the 6-h uptake data can be paired appropriately with the others based on the sequence in which the experiments were performed (**Tables S2 and S3**, respectively, for RLD-US-R1 and RLD-US-R2 as the RLD); i.e., as if a corrected experimental design had been followed. The derived BE analysis from this exercise (for which, like the AUC_z, an SABE assessment is recommended) is entirely consistent with those from the approach reported in Fig. 5. A further advantage of this illustration is that the number of pigs required to adequately power an assessment of BE to the chosen RLD can be calculated based on the observed variability within subjects, and between subjects for the compared formulations. Such an analysis shows that SABE assessments of test and reference formulations that did not satisfy the requirements for BE – even though their average BE ratio is within the 0.8 and 1.25 range and BE is therefore expected - require more than three skin donors (and for some comparisons more than 20 are needed) to sufficiently power the assessment (**Table S4**).

Parenthetically, while specific attention on the drug clearance measurements has been relatively limited, it is worth noting that – when collection of more data points (above the A_C) are possible – a further PK metric, specifically an 'elimination' rate constant at different depths, may be extracted from the results and used to compare, for example, whether subtle differences in the excipient composition of different products might impact on the retention time of the active drug within the skin. In the case of MTZ and the formulations studied in this work, however, it must be said that visual inspection of the clearance profiles in Fig. 3 does not support any obvious difference in the clearance kinetics of MTZ delivered from the gels and solutions. A further comment on this point is that the Raman information acquired may also be used – as has been reported with SC sampling experiments using adhesive tape-stripping [13,37] – to generate relative flux measurements of the drug from one nominal skin position into the next from the difference between selected 'uptake' and 'clearance' values of the normalised Raman signals.

A third metric has also been explored that attempts to capture the spatiotemporal nature of the experimental data provided by confocal Raman, that is, the volume under the 3-dimensional surface (VUS) created by simultaneously visualizing the normalised spectroscopic signals as functions of both time of measurement and depth into the skin (Fig. 7). When the VUS results are analysed for BE between the different formulations studied, the conclusions are entirely consistent with those drawn using AUC_t and AUC_z. The benefit of the VUS approach is that it is a single metric encompassing product performance and requires drug disposition in position and time in the skin by a 'test' formulation to match that of the RLD.

As previously acknowledged, the scope of this proof-of-concept study was, in fact, insufficiently powered to draw any definitive conclusions about BE. However, that was not a specific goal of this work; rather, the objective was to determine whether a Raman-based approach could realistically characterise a drug's cutaneous PK profile and enable, potentially, conclusions about the BA of the molecule *at or near its site of action* to be drawn. In that respect, this capability has indeed been shown, and the results provide the basis for future refinements to study

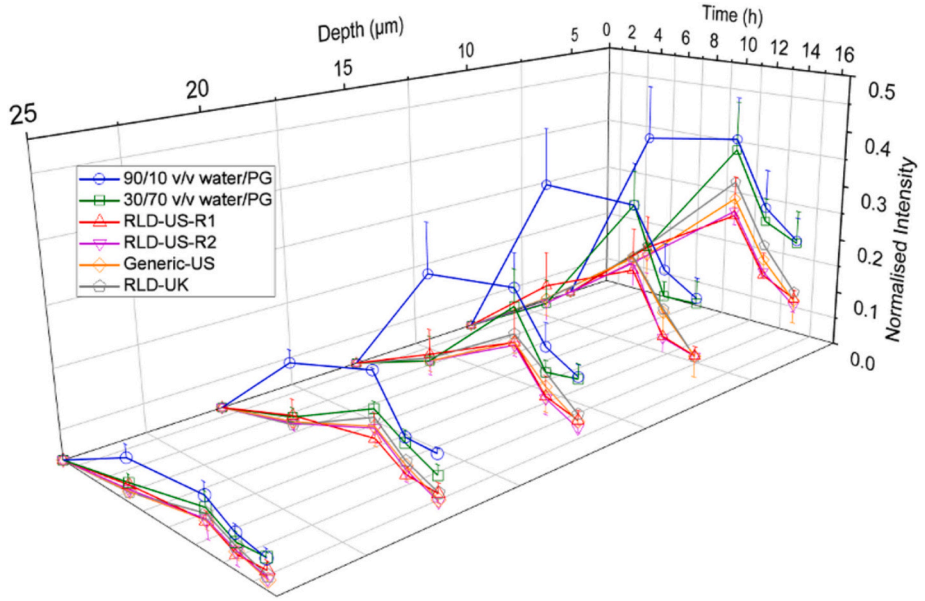

Fig. 7. MTZ maximum Raman intensities, normalised by Amide *I*, as functions of time of measurement and depth into the skin for the gels and solutions studied. Data from three pigs are combined and are presented as mean *v*alues + or – SD. [missing data points are explained elsewhere in the paper].

design, appropriately powering the replicate measurements needed, and data analysis (conventional BE metrics, use of SABE, etc.) to be made.

Finally, the SRS experiments offered at least partial validation of the confocal Raman results and, with its rapid scan and enhanced axial resolution capabilities, provided detailed visualisation of the skin samples under examination. A less extensive set of representative studies was performed, limited to 6 h MTZ uptake measurements post-application of the RLD-US and Generic-US gels and the 90/10 *v*/v water:PG solution. In terms of signal analysis, in addition to the normalisation with Amide *I* [18] as performed with confocal Raman, the SRS data were also corrected for any spurious signals by subtracting the intensity of an off-resonance frequency from that of the on-resonance for MTZ. The off-resonance frequency was selected to be one where (i) a background presence from the skin was present, and (ii) there was a slight overlap with the MTZ signal. This automatically ensures that spurious signals are removed but also means that a part of the MTZ signal is lost. On the other hand, the conservative approach guarantees that the signal measured at the MTZ frequency can only originate from the drug (even though sensitivity is somewhat compromised).

As with confocal Raman, the AUC_z values (over the range 0–10 μm) for the two gels and the one solution (Fig. 6) were clearly distinguishable: that for the solution was significantly greater (and the signal persisted to greater skin depths) than those for both gels, and the latter were different from one another or from the untreated skin control. These comparisons were based primarily on observations from the SC alone but with depth resolutions of ~1 μm. Finally, and again supporting previous findings [44], SRS clearly delineated the presence of solid MTZ on the skin surface post-application and metamorphosis of the 90/10 water:PG solution (Fig. 6). It is also worth mentioning that relative to previous studies using SRS [13,18], MTZ uptake into the skin was lower and its Raman signals were smaller. For measurable signals to be detected, therefore, a higher laser power was required for SRS imaging, and this was only achievable, without compromising the skin (dehydration and burning) through the use of a temperature-controlled stage that kept the tissue at 10 °C. This contingency, of course, means that translation of the SRS approach to in vivo use is not immediately foreseeable, particularly for drugs that are only taken up into the skin to the same extent as (or even less than) MTZ.

5. Conclusion

It has been demonstrated that confocal Raman spectroscopy is a useful tool with which to assess the cutaneous BA of a topically applied drug and, by extrapolation, to determine the BE between RLD and prospective generic drug products. The investigation showed that the approach may have broader applicability to topical drugs than previously demonstrated and that drug disposition at depths closer to typical sites of action can be evaluated. While a definitive assessment of BE between different drug products was not possible, it seems clear that, with a suitably powered experimental design, this will be achievable. The next step must be translation to an in vivo confirmation of the results presented and proof-of-principle that this advance in regulatory science can accelerate and widen patient access to higher quality generic medicines. In a broader context, the Raman approaches presented here offer potentially useful non-invasive and non-destructive tools with which to facilitate the development and optimisation of novel formulations, complementing more conventional methodologies in current use, and accelerating the route-to-market of innovative drug products.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jconrel.2025.114190.

CRediT authorship contribution statement

Panagiota Zarmpi: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Dimitrios Tsikritsis:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Andrew C. Watson:** Writing – review & editing, Visualization, Validation, Investigation, Formal analysis. **Jean-Luc Vorng:** Writing – review & editing, Validation, Methodology, Investigation. **Vasundhara Tyagi:** Writing – review & editing, Visualization, Validation, Methodology, Investigation. **Natalie A. Belsey:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Elena Rantou:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis. **Priyanka Ghosh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation,

Conceptualization. **Annette L. Bunge:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Timothy J. Woodman:** Writing – review & editing, Validation, Methodology, Investigation. **M. Begoña Delgado-Charro:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Richard H. Guy:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

This project is supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award 1-U01-FD 006533 (RHG) totalling $1.25 M with 100 % funded by FDA/HHS. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.

Declaration of competing interest

None.

Acknowledgements

Valuable insight from and stimulating discussions with Drs. Sam Raney and Markham Luke from the FDA's Office of Generic Drugs, and from Professor Jane White at the University of Bath, are gratefully acknowledged. NAB thanks the Community for Analytical Measurement Science for a 2020 CAMS Fellowship Award funded by the Analytical Chemistry Trust Fund.

Data availability

Data will be made available on request.

References

[1] FDA, Guidance for Industry: bioavailability and bioequivalence studies submitted in NDAs or INDs — General considerations. https://www.fda.gov/files/drugs/published/Bioavailability-and-Bioequivalence-Studies-Submitted-in-NDAs-or-INDs-%E2%80%94-General-Considerations.pdf, 2014 (accessed 17 January 2023).
[2] EMA, Guideline on the investigation of bioequivalence. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-bioequivalence-rev1_en.pdf, 2010 (accessed 04 November 2024).
[3] M.-L. Chen, V. Shah, R. Patnaik, W. Adams, A. Hussain, D. Conner, et al., Bioavailability and bioequivalence: an FDA regulatory overview, Pharm. Res. 18 (2001) 1645–1650, https://doi.org/10.1023/A:1013319408893.
[4] FDA, Drugs@FDA Glossary of Terms. https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-glossary-terms#RLD, 2017 (accessed 12 December 2024).
[5] FDA, Physicochemical and structural (Q3) characterization of topical drug products: Submitted in ANDAs guidance for industry. https://www.fda.gov/media/162471/download, 2022 (accessed 26 October 2023).
[6] FDA, Bioequivalence studies with pharmacokinetic endpoints for drug submitted under an ANDA guidance for industry. https://www.fda.gov/media/87219/download, 2021 (accessed 30 October 2024).
[7] V.P. Shah, G.L. Flynn, A. Yacobi, H.I. Maibach, C. Bon, N.M. Fleischer, et al., Bioequivalence of topical dermatological dosage forms–methods of evaluation of bioequivalence, Pharm. Res. 15 (1998) 167–171, https://doi.org/10.1023/a:1011941929495.
[8] M. Miranda, Z. Volmer, A. Cornick, A. Goody, C. Cardoso, A.A.C.C. Pais, et al., *In vitro* studies into establishing therapeutic bioequivalence of complex topical products: weight of evidence, Int. J. Pharm. 656 (2024) 124012, https://doi.org/10.1016/j.ijpharm.2024.124012.
[9] FDA, Product-specific guidances for generic drug development. https://www.fda.gov/drugs/guidances-drugs/product-specific-guidances-generic-drug-development, 2022 (accessed 05 November 2024).
[10] FDA, FY 2022 GDUFA science and research report. https://www.fda.gov/media/164843/download, 2022 (accessed 28 October 2023).
[11] EMA, Guideline on quality and equivalence of locally applied, locally acting cutaneous products. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-quality-equivalence-locally-applied-locally-acting-cutaneous-products_en.pdf, 2024 (accessed 05 November 2024).
[12] M.E. Darvin, Optical methods for non-invasive determination of skin penetration: current trends, advances, possibilities, prospects, and translation into *in vivo* human studies, Pharmaceutics 15 (2023) 1–35, https://doi.org/10.3390/pharmaceutics15092272.
[13] P. Zarmpi, M.A.M. Tabosa, P. Vitry, A.L. Bunge, N.A. Belsey, D. Tsikritsis, et al., Confocal raman spectroscopic characterization of dermatopharmacokinetics *ex vivo*, Mol. Pharm. 20 (2023) 5910–5920, https://doi.org/10.1021/acs.molpharmaceut.3c00755.
[14] D. Tsikritsis, E.J. Legge, N.A. Belsey, Practical considerations for quantitative and reproducible measurements with stimulated raman scattering microscopy, Anal 147 (2022) 4642–4656, https://doi.org/10.1039/D2AN00817C.
[15] L. Franzen, M. Windbergs, Applications of raman spectroscopy in skin research — from skin physiology and diagnosis up to risk assessment and dermal drug delivery, Adv. Drug Deliv. Rev. 89 (2015) 91–104, https://doi.org/10.1016/j.addr.2015.04.002.
[16] N. Everall, Optimising image quality in 2D and 3D confocal raman mapping, J. Raman Spectrosc. 45 (2014) 133–138, https://doi.org/10.1002/jrs.4430.
[17] M.A. Maciel Tabosa, P. Vitry, P. Zarmpi, A.L. Bunge, N.A. Belsey, D. Tsikritsis, et al., Quantification of chemical uptake into the skin by vibrational spectroscopies and stratum corneum sampling, Mol. Pharm. 20 (2023) 2527–2535, https://doi.org/10.1021/acs.molpharmaceut.2c01109.
[18] P. Zarmpi, D. Tsikritsis, J.L. Vorng, N.A. Belsey, A.L. Bunge, T.J. Woodman, et al., Evaluation of chemical disposition in skin by stimulated raman scattering microscopy, J. Control. Release 368 (2024) 797–807, https://doi.org/10.1016/j.jconrel.2024.02.011.
[19] N.A. Belsey, A. Dexter, J.L. Vorng, D. Tsikritsis, C.J. Nikula, T. Murta, et al., Visualisation of drug distribution in skin using correlative optical spectroscopy and mass spectrometry imaging, J. Control. Release 364 (2023) 79–89, https://doi.org/10.1016/j.jconrel.2023.10.026.
[20] V. Tyagi, A. Dexter, J.-L. Vorng, D. Tsikritsis, I.S. Gilmore, R.H. Guy, M. Boncheva Bettex, N.A. Belsey, Correlative optical spectroscopy and mass spectrometry imaging methodology to visualise drug distribution in a soft tissue section, J. Vis. Exp. 220 (2025) e67383, https://doi.org/10.3791/67383 jove.com/video/67383.
[21] M. Bergholt, A. Serio, M. Albro, Raman spectroscopy: guiding light for the extracellular matrix, Front. Bioeng. Biotechnol. 7 (2019) 1–16, https://doi.org/10.3389/fbioe.2019.00303.
[22] A. Goel, D. Tsikritsis, N.A. Belsey, R. Pendlington, S. Glavin, T. Chen, Measurement of chemical penetration in skin using stimulated Raman scattering microscopy and multivariate curve resolution - alternating least squares, Spectrochim. Acta A Mol. Biomol. Spectrosc. 296 (2023) 122639, https://doi.org/10.1016/j.saa.2023.122639.
[23] ISO, Surface chemical analysis — X-ray photoelectron spectroscopy — Estimating and reporting detection limits for elements in homogeneous materials. https://www.iso.org/obp/ui/#iso:std:iso:19668:ed-1:v1:en, 2017 (accessed 29 February 2023).
[24] I. Jakasa, S. Kezic, Evaluation of in-vivo animal and in-vitro models for prediction of dermal absorption in man, Hum. Exp. Toxicol. 27 (2008) 281–288, https://doi.org/10.1177/0960327107085826.
[25] G.A. Simon, H.I. Maibach, The pig as an experimental animal model of percutaneous permeation in man: qualitative and quantitative observations–an overview, Skin Pharmacol. Appl. Ski. Physiol. 13 (2000) 229–234, https://doi.org/10.1159/000029928.
[26] C. Herkenne, A. Naik, Y.N. Kalia, J. Hadgraft, R.H. Guy, Effect of propylene glycol on ibuprofen absorption into human skin *in vivo*, J. Pharm. Sci. 97 (2008) 185–197, https://doi.org/10.1002/jps.20829.
[27] L. Prasco, 0.75% Metronidazole topical gel. https://nctr-crs.fda.gov/fdalabel/services/spl/set-ids/49b16ff1-299d-4dd0-82a7-fdb9c6e7c7a0/spl-doc?hl=metronidazole%20topical%20gel, 2025.
[28] I. Tolmar, 0.75% Metronidazole topical gel. https://fda.report/DailyMed/2ede12d7-421c-4580-93ad-289a5c616297, 2025.
[29] L. Galderma, 0.75% Rosex gel summary of product characteristics. https://mhraproducts4853.blob.core.windows.net/docs/d004fa0c82e41860018a63df4b3cef8d1bc0b5e8, 2022.
[30] EMA, Annex to the European Commission guideline on 'Excipients in the labelling and package leaflet of medicinal products for human use'. https://www.ema.europa.eu/en/annex-european-commission-guideline-excipients-labelling-package-leaflet-medicinal-products-human-use, 2024 (accessed 30 October 2024).
[31] OECD, Test No. 105: water solubility, OECD guidelines for testing of chemicals, section 1. https://www.oecd.org/en/publications/test-no-105-water-solubility_9789264069589-en.html, 1995 (accessed 05 March 2025).
[32] T.P. de Araujo, I.M. Fittipaldi, D.C.G. Bedor, M.L. Duarte, S.F. Cordery, R.H. Guy, et al., Topical bio(in)equivalence of metronidazole formulations *in vivo*, Int. J. Pharm. 541 (2018) 167–172, https://doi.org/10.1016/j.ijpharm.2018.02.032.
[33] M. Newville, R. Otten, A. Nelson, T. Stensitzki, A. Ingargiola, D. Allan, et al., lmfit/lmfit-py: 1.3.2, Zenodo, 2024, https://doi.org/10.5281/zenodo.12785036.
[34] R.L. McCreery, Signal-to-Noise in Raman Spectroscopy. Raman Spectroscopy for Chemical Analysis, Wiley-Interscience, New York, 2000, pp. 49–71.
[35] R.J. Keizer, R.S. Jansen, H. Rosing, B. Thijssen, J.H. Beijnen, J.H. Schellens, et al., Incorporation of concentration data below the limit of quantification in population pharmacokinetic analyses, Pharmacol. Res. Perspect. 3 (2015) e00131, https://doi.org/10.1002/prp2.131.
[36] EPA, Guidance for data quality assessment; practical methods for data analysis EPA QA/G-9. https://www.epa.gov/sites/default/files/2015-06/documents/g9-final.pdf, 2000 (accessed 23 September 2024).

[37] A. Pensado, W.S. Chiu, S.F. Cordery, E. Rantou, A.L. Bunge, M.B. Delgado-Charro, et al., Stratum corneum sampling to assess bioequivalence between topical acyclovir products, Pharm. Res. 36 (2019) 1–16, https://doi.org/10.1007/s11095-019-2707-3.

[38] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH image to ImageJ: 25 years of image analysis, Nat. Methods 9 (2012) 671–675, https://doi.org/10.1038/nmeth.2089.

[39] D.J. Schuirmann, A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability, J. Pharmacokinet. Biopharm. 15 (1987) 657–680, https://doi.org/10.1007/bf01068419.

[40] FDA, Statistical approaches to establishing bioequivalence Guidance for industry. https://www.fda.gov/media/163638/download, 2022 (accessed 25 October 2024).

[41] S. Osseiran, J.D. Cruz, S. Jeong, H. Wang, C. Fthenakis, C.L. Evans, Characterizing stratum corneum structure, barrier function, and chemical content of human skin with coherent raman scattering imaging, biomed, Opt. Express 9 (2018) 6425–6443, https://doi.org/10.1364/BOE.9.006425.

[42] P.J. Caspers, C. Nico, T.C. Bakker Schut, J. de Sterke, P.D.A. Pudney, P.R. Curto, et al., Method to quantify the *in vivo* skin penetration of topically applied materials based on confocal raman spectroscopy, Transl. Biophotonics. 1 (2019) 1–10, https://doi.org/10.1002/tbio.201900004.

[43] J.P. Sylvestre, C.C. Bouissou, R.H. Guy, M.B. Delgado-Charro, Extraction and quantification of amino acids in human stratum corneum *in vivo*, Br. J. Dermatol. 163 (2010) 458–465, https://doi.org/10.1111/j.1365-2133.2010.09805.x.

[44] P. Zarmpi, A. Pensado, S.N. Gordeev, K.A. Jane White, A.L. Bunge, R.H. Guy, et al., Physicochemical characterization of the metamorphosis of film-forming formulations of betamethasone-17-valerate, Int. J. Pharm. 664 (2024) 124595, https://doi.org/10.1016/j.ijpharm.2024.124595.