ORIGINAL ARTICLE

Anti-inflammatory potency testing of topical corticosteroids and calcineurin inhibitors in human volunteers sensitized to diphenylcyclopropenone

Kristian F. Mose[1,2,*] (iD), Flemming Andersen[1,2], Mads A. Røpke[3], Lone Skov[4], Peter S. Friedmann[5] and Klaus E. Andersen[1,2,6] (iD)

[1]Department of Dermatology and Allergy Centre, Odense University Hospital, University of Southern Denmark, 5000 Odense C, Denmark, [2]Dermatological Investigations Scandinavia, J.B. Winsløwsvej 9, 5000 Odense C, Denmark, [3]Department of Clinical Pharmacology, LEO Pharma A/S, 2750 Ballerup, Denmark, [4]Department of Dermato-Allergology, Herlev and Gentofte Hospital, University of Copenhagen, 2900 Hellerup, Denmark, [5]Division of Infection, Inflammation & Immunity, Sir Henry Wellcome Laboratories, Southampton University Hospitals NHS Trust, Southampton SO16 6YD, UK, and [6]Centre for Innovative Medical Technology, Institute of Clinical Research, University of Southern Denmark, 5000 Odense C, Denmark

*Principal investigator: Kristian F. Mose

AIMS
To quantify the anti-inflammatory potency of topical corticosteroids and topical calcineurin inhibitors by measuring the contact allergic response to a diphenylcyclopropenone (DPCP) challenge in *de novo* sensitized human volunteers.

METHODS
Two randomized, double-blind, vehicle-controlled studies were performed encompassing 76 volunteers: 29 in the first and 47 in the second study. Topical drugs were applied pre- and/or post-treatment in block designs. The compounds were tested simultaneously under occluded patch tests covering DPCP-induced dermatitis. Inhibitory responses were assessed by visual scoring and measurements of the oedema thickness with ultrasound.

RESULTS
When applied both before and after the DPCP challenge, significant anti-inflammatory effects were seen in descending order for tacrolimus 0.1% ointment, clobetasol propionate ointment, betamethasone valerate ointment and hydrocortisone butyrate ointment, while pimecrolimus cream, hydrocortisone ointment and vehicles had no significant effect. Only tacrolimus ointment ($P < 0.01$) demonstrated a consistent significant pre-treatment inhibitory effect compared with an untreated DPCP control.

CONCLUSIONS
This human testing method in which the inflammation of experimentally induced allergic patch test reactions is quantified by objective measurement allows an analysis of the anti-inflammatory potency of not only topical corticosteroids, but also of drugs that have no effect on vasoconstriction. The method allowed comparison of the potencies of four topical corticosteroids and two calcineurin inhibitors.

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- The potency of topical corticosteroids is determined by the human vasoconstrictor assay, which is a surrogate method for anti-inflammatory effect.
- Human methods for testing the anti-inflammatory potency of topical corticosteroids are reported; however, there is no standardized quantitative human test method for simultaneous testing of both corticosteroids and nonsteroidal drugs.

WHAT THIS STUDY ADDS

- These studies provide a standardized method for testing and comparing the anti-inflammatory potencies of topical corticosteroids and calcineurin inhibitors in human volunteers.

Introduction

Corticosteroids are the major topical anti-inflammatory drugs used in the treatment of a wide range of inflammatory skin diseases. There is a need for new therapies and for reliable human testing methods that allow the anti-inflammatory potencies of topical corticosteroids to be compared, as well to compare topical non-steroidal anti-inflammatory drugs with topical corticosteroids. Traditionally, the potency of topical corticosteroids has been determined using the human vasoconstrictor assay in which cutaneous pallor is evaluated after occluded application of agents on healthy skin [1]. This classical assay has inherent limitations as follows: (i) the vasoconstrictor assay methodology relies primarily on the subjective nature of clinical inspection to estimate the blanching effect; (ii) it uses the degree of skin blanching as an indicator of drug potency and, as such, only functions as a surrogate assessment of the anti-inflammatory effect; and (iii) it is not possible to compare the effect of topical corticosteroids with nonsteroidal therapies such as topical calcineurin inhibitors, which have no vasoconstrictor activity [2, 3]. At present, tacrolimus and pimecrolimus are only approved for the treatment of atopic dermatitis; however, there is evidence to support their efficacy in the treatment of other types of eczema including seborrheic dermatitis and allergic contact dermatitis [4–8]. Several studies have compared the therapeutic potency of topical corticosteroids by assessing their effects on allergic reactions elicited by environmental allergens such as nickel or poison ivy in spontaneously sensitized patients [2, 9–13], while use of a standardized immunological challenge with an experimental sensitizer, dinitrochlorobenzene (DNCB), in healthy volunteers has only been reported once as a tool to quantify the potency of anti-inflammatory agents [3].

To develop a human testing method to quantify the relative potencies of topical anti-inflammatory drugs, we chose diphenylcyclopropenone (DPCP) as the experimental allergen. We recently showed that following sensitization of healthy volunteers with DPCP, reactivity to repeated challenges with DPCP initially became progressively stronger, but following the second epicutaneous challenge, the responses become stable and reproducible as assessed by both visual scoring and as measured by skin fold thickness [14]. This clinical finding was supported by histopathology and immunohistochemistry staining together with microarray gene expression profiling using skin biopsies taken from DPCP-challenged sites [15].

We aimed to develop a human testing method for a T-cell mediated inflammatory process in the form of a quantified and standardized allergic contact hypersensitivity response, to evaluate the relative potencies of a variety of topical anti-inflammatory drugs. We designed the study so that the inflammatory challenge used for testing anti-inflammatory effects was not too aggressive and therefore had the best chance of revealing weaker drug effects. This was achieved by pretesting volunteers to determine the ability of the DPCP challenge concentrations to elicit measurable inflammatory responses. We also wanted to compare the sensitivity of the test system to detect potency differences when the agents were applied before or after the DPCP challenge, or were applied both before and after the challenge. We performed two randomized, double blind, vehicle-controlled studies in which tacrolimus, pimecrolimus and four topical corticosteroids of different strengths/*potencies* were compared in randomized block designs.

Methods

Volunteers

Healthy volunteers were recruited by advertising on internal notice boards. Twenty-nine volunteers participated (13 women and 16 men, aged 20–43 years, median 25 years) in study I, and 47 volunteers (16 women and 31 men, aged 20–44 years, median 26 years) in study II. Participants comprised two different groups; pre-sensitized volunteers ($n = 5$) from a previous DPCP study [14, 15] and first-time volunteers (Study I: $n = 24$; Study II: $n = 23$). Informed consent was obtained from all participants. Exclusion criteria included the following: (i) active skin disease; (ii) a history of atopic dermatitis; (iii) endocrine or immune system disorders; (iv) pregnancy or breastfeeding; and (v) active or prior (i.e. 30 days before study inclusion) treatment with UV radiation, systemic/topical corticosteroids or other immunosuppressive agents that might influence the treatment response. The studies were approved by the Regional Scientific Ethical Committees for Southern Denmark (Study-ID: S-20140149 and S-20130074) and conducted in accordance with the Helsinki declaration.

Induction of DPCP contact allergy

Induction of allergic sensitivity was performed using a single DPCP dose of 30 µg cm^{-2} (22.8 µl of 0.125% DPCP in acetone) applied to a filter-paper-lined, 12-mm Finn chamber

(SmartPractice, Phoenix, AZ, USA). The patch test chamber was attached to the skin on the upper buttock then left in place for 48 h [16]. Presensitized volunteers (*n* = 5) from a previous DPCP study [14, 15] and all participants in studies I and II were sensitized using the same sensitization protocol. All DPCP (CAS no. 886–38-4; Alfa Aesar, Karlsruhe, Germany) solutions used in the study were provided by Central Pharmacy, Odense University Hospital.

The initial elicitation challenge

The initial elicitation challenge was carried out four weeks after the induction phase. Volunteers received a DPCP dose series consisting of seven doses (dose per unit areas) that were increased in 60% increments: 0.484, 0.774, 1.24, 1.98, 3.17, 5.1 and 8.12 $\mu g\ cm^{-2}$. The DPCP doses were applied to filter-paper-lined 8-mm Finn chambers as 10 μl of the appropriate concentration in acetone to give the required dose-range. A (negative) control of acetone only was also applied in a separate patch test chamber. The patch test chambers were loaded immediately before their attachment to the upper back. The elicitation patches were removed after 6 h and the patch test sites were marked with a skin marker. At 24 and 48 h, the elicitation responses were assessed visually and by skin ultrasound. In study I, two consecutive DPCP doses were selected for each volunteer at the final (48-h) reading. This was done to: (i) ensure that a clinically detectable and quantifiable degree of DPCP reactivity was elicited in the subsequent topical treatment phase; and (ii) examine the inhibitory effects of topical treatments on patch test reactions of different inflammatory levels elicited by lower versus higher DPCP doses. Based on the findings from study I, we selected a single DPCP dose in study II for each volunteer that elicited a patch test reaction, corresponding to a doubtful/weak positive reaction or higher, at the final (48-h) visual reading.

Study designs

Both studies were conducted as randomized, double-blind, vehicle-controlled trials comprising four phases as follows: (i) induction of DPCP contact allergy in first-time volunteers; (ii) initial elicitation challenge with a DPCP dose series to determine the desired dose (i.e. a mild to moderate positive reaction) to be used in the drug treatment phase; (iii) topical drug treatment of DPCP-induced allergic patch test reactions; and (iv) visual scoring and skin ultrasound measurement of test sites at 48 and 72 h. The placement of test drugs was rotated (randomized) between volunteers.

In study I, 3 topical corticosteroid ointments were evaluated in comparison with a negative control [i.e. an occluded (empty) patch test chamber] and vehicle ointment on both sides of the upper back in each individual. A randomized complete block design was applied for this study.

In study II, the anti-inflammatory effect of six topical agents was evaluated using a similar patch test method, and in comparison with an untreated (positive) DPCP control, negative control, and vehicle ointment and cream. The number of DPCP patches on each volunteer was limited to six on each side of the upper back; therefore, due to the increased number of tested agents, a balanced incomplete block design was applied for this study.

Visual scoring

An extended version of the International Contact Dermatitis Research Group scoring scale was used (Table 1) [17]. The responses were graded as negative (–), doubtful (?+), weak positive (1+), definite positive (2+), strong positive (3+) or extremely strong positive (4+).

Skin ultrasound

High-resolution (20 MHz) skin ultrasound measurements of patch test reactions were done using the DermaLab (SkinLab) Combo Instrument (Cortex Technology, Hadsund, Denmark). The ultrasound image contains a specific curve (a *super A-scan*), which describes the total intensity of the scanned skin area as a function of the depth into the skin. The full ultrasound image is composed of the accumulated average of 188 A-scans. The average thickness of the dermis was calculated based on the super A-scan. Dermal inflammation was determined by recording two dermal thickness scans at each test site at both 48 h and 72 h after treatment (Figure 1). The means of the two dermal thickness measurements were used in the data analysis. For all recordings, the same operating conditions were used for all volunteers.

Treatment of DPCP-induced allergic contact dermatitis with test drugs

Study I assessed the effects of the topical anti-inflammatory drugs applied after removal of the DPCP challenge (post-treatment) because this mimics normal use in clinical practice. Volunteers were rechallenged 5 days after the initial elicitation challenge on either side of the upper back with the two consecutive DPCP doses that had elicited mild to moderate reactions in the first challenge. These two doses were applied in two separate panels, each containing four patches and a vehicle (acetone) control patch. The DPCP doses were applied to filter-paper-lined, 8-mm Finn chambers and were attached to the skin for 6 h. After removal, elicitation sites were marked with a skin marker. Large (12-mm) Finn chambers were then loaded with approximately 45 mg (corresponding to a dose per area of $\approx$ 40 mg cm^{-2}) ointment and placed as a treatment to three of the highlighted

Table 1

Modified version of the International Contact Dermatitis Research Group clinical scoring system

Symbol	Numerical value	Morphology
-	0	Negative reaction
?+	1	Doubtful reaction; faint erythema only
1+	2	Weak positive reaction; erythema with no infiltration, possibly papules
2+	3	Definite positive reaction; erythema, infiltration, follicular papules, no vesicles
3+	4	Strong positive reaction; intense erythema, infiltration and vesicles
4+	5	Extremely strong positive reaction; coalescing vesicles and/or bullae
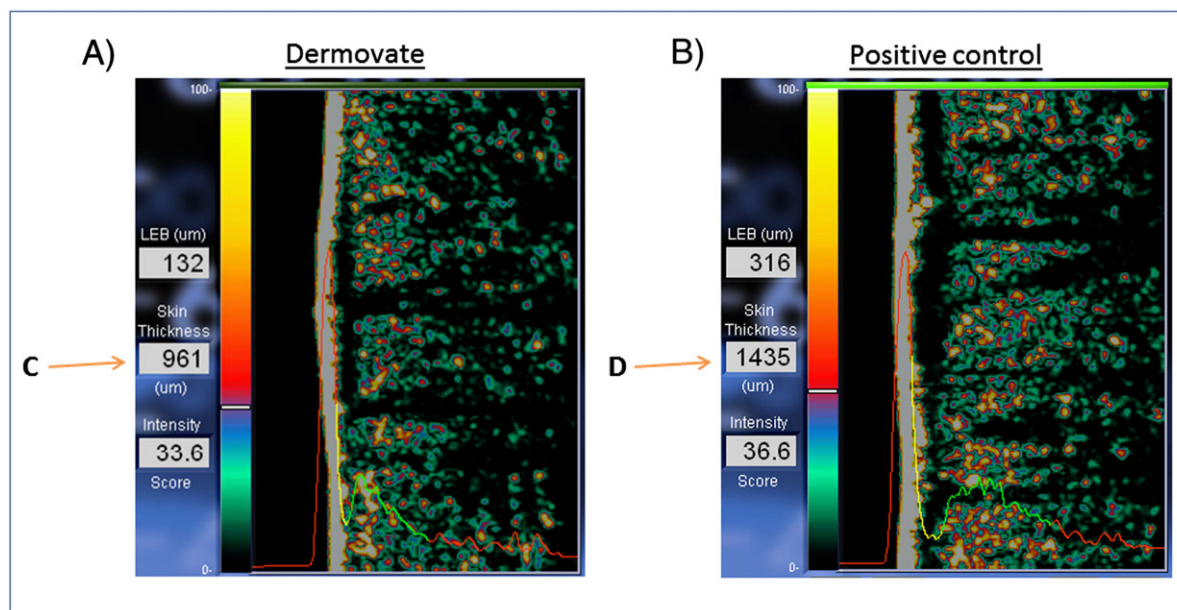
Figure 1

Automated skin ultrasound measurements of dermal thickness in allergic contact dermatitis (ACD) and control reactions. The *red left* part of the ultrasound curve shows the intensity of the ultrasound passing through the water chamber in the probe. The high left peak originates from the combined film/epidermal reflection. The *red right* part of the curve shows the intensity of the ultrasound passing through the subcutaneous layers. The *yellow* part of the curve indicates the less reflective part of the dermis, while the *green* part indicates the more reflective part of the dermis. The average thickness of the dermis is calculated based on the super A-scan. The dermal thickness measurements (in μm) are shown next to the ultrasound images (i.e. C and D) and represent the *yellow* and *green* part of the curves. (A) Ultrasound scan of ACD after Dermovate (clobetasol propionate) treatment at 72 h and (B) untreated (positive) diphenylcyclopropenone control scan at 72 h. Data from a participant in Study I are shown

elicitation/challenge sites, while the fourth DPCP elicitation area was treated with a vehicle (white soft paraffin) patch. The vehicle (acetone) control area was occluded with an empty patch test chamber (i.e. negative control). All treatment patches were removed after 48 h. The response at each test site was assessed visually and measured with skin ultrasound immediately and after a further 24 h.

In study II, the anti-inflammatory effects of six topical drugs (four corticosteroid ointments, tacrolimus 0.1% ointment, pimecrolimus 1% cream, and vehicle ointment and cream) were compared following pre-treatment alone and following combined pre- and post-treatment. The combined pre- and post-treatment was used in order to maximize the possible effect. The same patch test technique was used as described for study I. Two panels were applied on each side of the upper back (i.e. a left and right panel), each with six test sites. According to the incomplete block design, the test sites on each volunteer were pretreated for 24 h with four of the topical products applied in duplicate in 12-mm Finn chambers. In addition, four empty Finn chambers were applied. All chambers were removed after 24 h and the test sites were marked. Subsequently, the selected DPCP dose was applied in filter-paper-lined, 8-mm Finn chambers and placed for 6 h on the eight pretreated sites and on two untreated sites. The two remaining sites were occluded with empty Finn chambers. After removal of the DPCP patches, the right panel test sites were post-treated for 48 h with the same agents as had been used for the pre-exposure treatment. The response at each test site was assessed visually and measured with skin

ultrasound immediately and after a further 24 h. For volunteers who had participated in study I, the challenges were performed on previously nonchallenged sites to avoid increased reactivity at previous DPCP-challenged sites [18]. A flow chart and schematic diagrams illustrating the study designs are shown in Figure 2.

Sample size

The sample size estimation was performed using the statistical software, SAS (software package: Proc Power, version 9.2, SAS Institute Inc., Cary, NC, USA). Sample size estimation was performed using a two-tailed paired t test and a level of significance of 0.05. Skin ultrasound was used as the primary measurement variable. In study I, a randomized complete block design was chosen, and a sample size of 24 volunteers was required to achieve 90% power and enabling detection of a 0.3-mm mean treatment difference based on a standard deviation of 0.3 mm. In study II, because of the larger number of test drugs included and to limit the number of DPCP patches on each volunteer, a balanced incomplete block design was chosen [19]. The parameters of the balanced incomplete block design were: t (total number of treatments) = 8, b (number of blocks) = 14, k (block size) = 4, and r (number of repetitions) = 3. The design contained $4 \times 14 = 56$ subexperiments and had a relative efficacy (variance of pairwise comparison) of 0.86 corresponding to a complete block design with 56/0.86 = 65 subexperiments. To achieve the desired power equal to $24 \times 8 = 192$ subexperiments, the
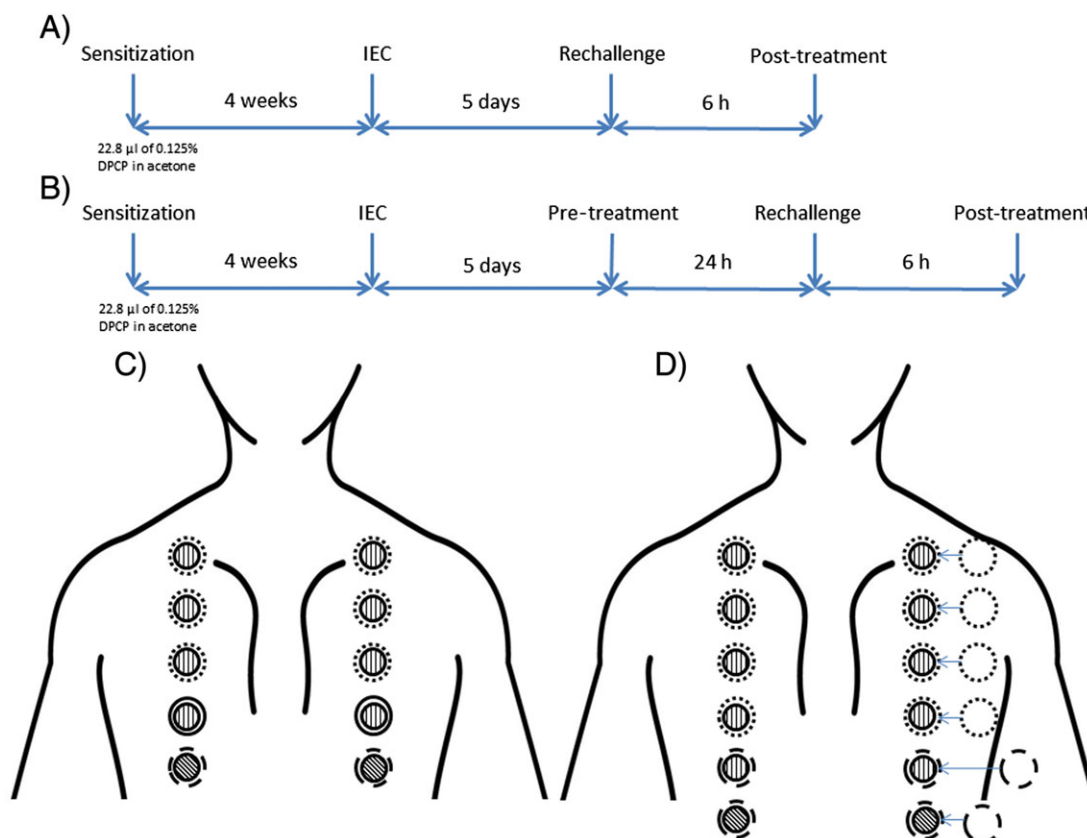
Figure 2

Protocols for sensitization, challenges and topical drug treatment. (A) *Flow chart (study I):* Volunteers were sensitized on the upper buttock for 48 h. Four weeks later, an initial elicitation challenge (IEC) was performed with a diphenylcyclopropenone (DPCP) dose-series. After IEC, volunteers were rechallenged with two DPCP doses applied in two separate panels on the upper back followed by post-treatment with topical drugs. (B) *Flow chart (study II):* The same setup was used as in study I except that after IEC, test sites were pretreated with four topical products on either side of the upper back (left and right panel). This was followed by rechallenge with a select DPCP dose and finally post-treatment with topical drugs. (C) *Schematic setup of rechallenge and post-treatment with topical drugs (study I):* The two panels each consisted of four small (8-mm) DPCP patches (shown as vertical hatchings) and a vehicle (acetone) control patch (cross-hatching). Upon removal of these chambers, three large (12-mm) chambers loaded with topical drugs (dotted rings), a large vehicle ointment chamber (solid ring), and a large empty (negative control) chamber (dashed ring) were placed on the elicitation areas. (D) *Schematic setup of pre-treatment, rechallenge, and combined pre- and post-treatment with topical drugs (study II):* The two pre-treatment panels each consisted of six test sites comprising four large Finn chambers filled with topical drugs (dotted rings) and two empty chambers (dashed rings). Small (8-mm) DPCP chambers (vertical hatchings) were placed on the eight pretreated sites and on two untreated sites. The two remaining sites were occluded with empty Finn chambers (cross-hatchings). The right panel was post-treated with the same drugs as used for pre-treatment, again using large Finn chambers (dotted rings)

balanced incomplete block design had to be repeated (192/65) three times. Hence, the total number of participants required for this study was 3 × 14 = 42.

Data and statistical analysis

Statistical analyses were performed with the statistical software package GraphPad Instat (version 3.00, GraphPad Software, San Diego, CA, USA). Experiments and data collection were done by operators blinded to the treatment identity. The blinding of treatments was broken after completion of all tests. Visual scores were transformed to numerical values, and non-parametric tests were applied for statistical analysis. In study I, the visual scores were analysed globally using a Friedman test, whereas a Kruskal–Wallis test was applied in

study II. In the case of global significance, *posthoc* multiple comparison tests (Dunn's post-test in study I and Dunnett's post-test in study II) were further applied to examine values between reference points [vehicle ointment in study I and untreated (positive) DPCP control in study II] and each topical drug.

In study I, the skin thickness values of test areas treated with drugs measured at 48 h and 72 h, respectively, were expressed as percentages of the vehicle ointment value (reference point). Vehicle values were arbitrarily set at 100%. In study II, untreated (positive) DPCP control was used as the reference point instead of vehicle ointment due to the incomplete block design. Skin thickness data were expressed as a percentage of the absolute values at each time-point in order to allow for accurate statistical analysis. One-way

analysis of variance with *posthoc* test (Dunnett's post-test), adjusted for multiple comparisons, was applied for the ultrasound data. The level of significance was set to $P < 0.05$. The data and statistical analysis comply with the recommendations on experimental design and analysis in pharmacology [20].

Materials (topical anti-inflammatory agents)

The topical agents were: clobetasol propionate 0.5 mg g^{-1} ointment (Dermovate; GlaxoSmithKline Pharma, Brøndby, Denmark), betamethasone-17-valerate 1 mg g^{-1} ointment (Betnovate; GlaxoSmithKline Pharma, Brøndby, Denmark), hydrocortisone-17-butyrate 1 mg g^{-1} ointment (Locoid; Astellas Pharma, Kastrup, Denmark), hydrocortisone 10 mg g^{-1} ointment (Hydrocortisone *DnE*; Den Norske Eterfabrikk, Oslo, Norway), tacrolimus 1 mg g^{-1} ointment (Protopic; Astellas Pharma, Kastrup, Denmark) and pimecrolimus 10 mg g^{-1} cream (Elidel; Meda, Allerød, Denmark). Vehicle ointment [Apotekets Vaseline (Ph. Eur.); Apotekernes Amba, Skovlunde, Denmark] and cream (Helo creme; Faaborg Pharma, Faaborg, Denmark) were included. The volunteers and investigators were blinded to all topical agents and vehicles. The Central Pharmacy, Odense University Hospital, conducted the blinding and coding.

Results

Response to initial DPCP elicitation challenge

In study I, positive contact hypersensitivity to DPCP was induced in 20 of 24 first-time volunteers (sensitization rate: ≈ 83%) based on the positive patch test responses at the initial elicitation challenge. Of the five presensitized volunteers, four responded to the dose series at the initial elicitation challenge; thus, 24 volunteers were included in the study. In study II, sensitization to DPCP was obtained in 21 of 23 first-time participants (sensitization rate: ≈ 91%). All presensitized volunteers from study I responded to the dose-series in the initial elicitation challenge; thus, 45 volunteers participated in study II.

The anti-inflammatory effect assessed by visual scoring

In the post-treatment study (study I), assessment by visual scoring revealed a significant anti-inflammatory effect for clobetasol propionate and betamethasone valerate ($P < 0.001$), while hydrocortisone was similar to vehicle. The same degree of inhibitory effects of clobetasol propionate and betamethasone valerate was found for both DPCP doses (Table 2).

In study II, visual scoring of DPCP-induced reactions pretreated with anti-inflammatory agents, only demonstrated a significant effect for tacrolimus 0.1% ointment at 48 h ($P < 0.001$) and at 72 h ($P < 0.01$) compared with untreated (positive) DPCP control. For DPCP-induced reactions, which received both pre- and post-treatment, the three strongest groups of steroids and tacrolimus showed the greatest overall inhibition of response at 48 h ($P < 0.001$; Table 3).

The anti-inflammatory effect assessed by skin ultrasound

In study I, significant anti-inflammatory effects ($P < 0.01$) were shown at both 48 h and 72 h for betamethasone valerate and clobetasol propionate on challenge sites elicited by higher DPCP doses. Betamethasone valerate and clobetasol propionate reduced the inflammatory thickness up to 23.3% and 22.5%, respectively. For lower DPCP doses, the anti-inflammatory effects were less pronounced (Figure 3A, B).

In Study II (pre-treatment of DPCP challenge sites), the anti-inflammatory effects of tacrolimus were significant ($P < 0.01$) at both 48 h and 72 h, with a 26% and 18% reduction in the inflammatory thickness, respectively. Betamethasone valerate ($P < 0.05$) displayed significant inhibition (11.5%) only at 48 h compared with untreated (positive) DPCP control. For combined pre- and post-treatment, there was a significant anti-inflammatory effect of tacrolimus (28% reduction in inflammatory thickness at 48 h and 26.5% at 72 h; $P < 0.01$), clobetasol propionate (28% at 48 h and 19% at 72 h; $P < 0.01$), betamethasone valerate (22% at 48 h and 16% at 72 h; $P < 0.01$), and hydrocortisone butyrate (11% at 72 h; $P < 0.01$), while the effects of pimecrolimus and hydrocortisone were not significant in this model (Figure 3C-F).

Table 2

Anti-inflammatory effect of topical drugs evaluated by visual scoring. Results 48 h and 72 h post-treatment [using lower and higher diphenylcyclopropenone (DPCP) doses]

Post-treatment												
Lower DPCP dose							**Higher DPCP dose**					
Vehicle/ Topical drug	**48 h**			**72 h**			**48 h**			**72 h**		
	Median	**Range**	*P* value	**Median**	**Range**	*P* value	**Median**	**Range**	*P* value	**Median**	**Range**	*P* value
Vehicle ointment	3	0–5	-	3	1–5	-	3	2–5	-	3	2–5	-
Hydrocortisone	3	1–5	NS	3	0–5	NS	3	1–4	NS	3	1–5	NS
Betamethasone valerate	2	0–4	<**0.001**	2	0–4	<**0.001**	2	0–5	<**0.001**	2	1–4	<**0.001**
Clobetasol propionate	2	0–4	<**0.001**	2	0–4	<**0.001**	2	1–5	<**0.01**	2	0–4	<**0.001**

DPCP, diphenylcyclopropenone. *P*-values <0.05 are considered statistically significant (NS = not significant)

Table 3

Anti-inflammatory effect of topical drugs evaluated by visual scoring. Pre-treatment results and combined pre- and 48 h and 72 h post-treatment results

| Pre-treatment | | | | | | | Combined pre- and post-treatment | | | | | |
| Positive DPCP/ Vehicle/ Topical drug | 48 h | | | 72 h | | | 48 h | | | 72 h | | |
	Median	Range	P value	Median	Range	P value	Median	Range	P value	Median	Range	P value
Positive DPCP	3	1–5	-	3	1–5	-	3	2–5	-	3	2–5	-
Vehicle ointment	4	2–5	NS	3	2–4	NS	4	2–5	NS	3	2–4	NS
Vehicle cream	3	1–5	NS	3	1–5	NS	4	2–5	NS	3	2–5	NS
Hydrocortisone	3	2–5	NS	3	2–5	NS	4	1–5	NS	3	2–5	NS
Pimecrolimus	3	2–5	NS	3	2–5	NS	3	0–5	<**0.05**	3	0–5	NS
Hydrocortisone butyrate	3	1–5	NS	3	1–5	NS	2	0–5	<**0.001**	2	1–5	NS
Betamethasone valerate	3	0–5	NS	3	2–5	NS	2	1–5	<**0.001**	3	1–5	NS
Clobetasol propionate	3	2–5	NS	3	0–5	NS	2	0–4	<**0.001**	2	0–4	<**0.01**
Tacrolimus	1	0–3	<**0.001**	2	0–3	<**0.01**	0	0–3	<**0.001**	0	0–3	<**0.001**

P-values <0.05 are considered statistically significant (NS = not significant)
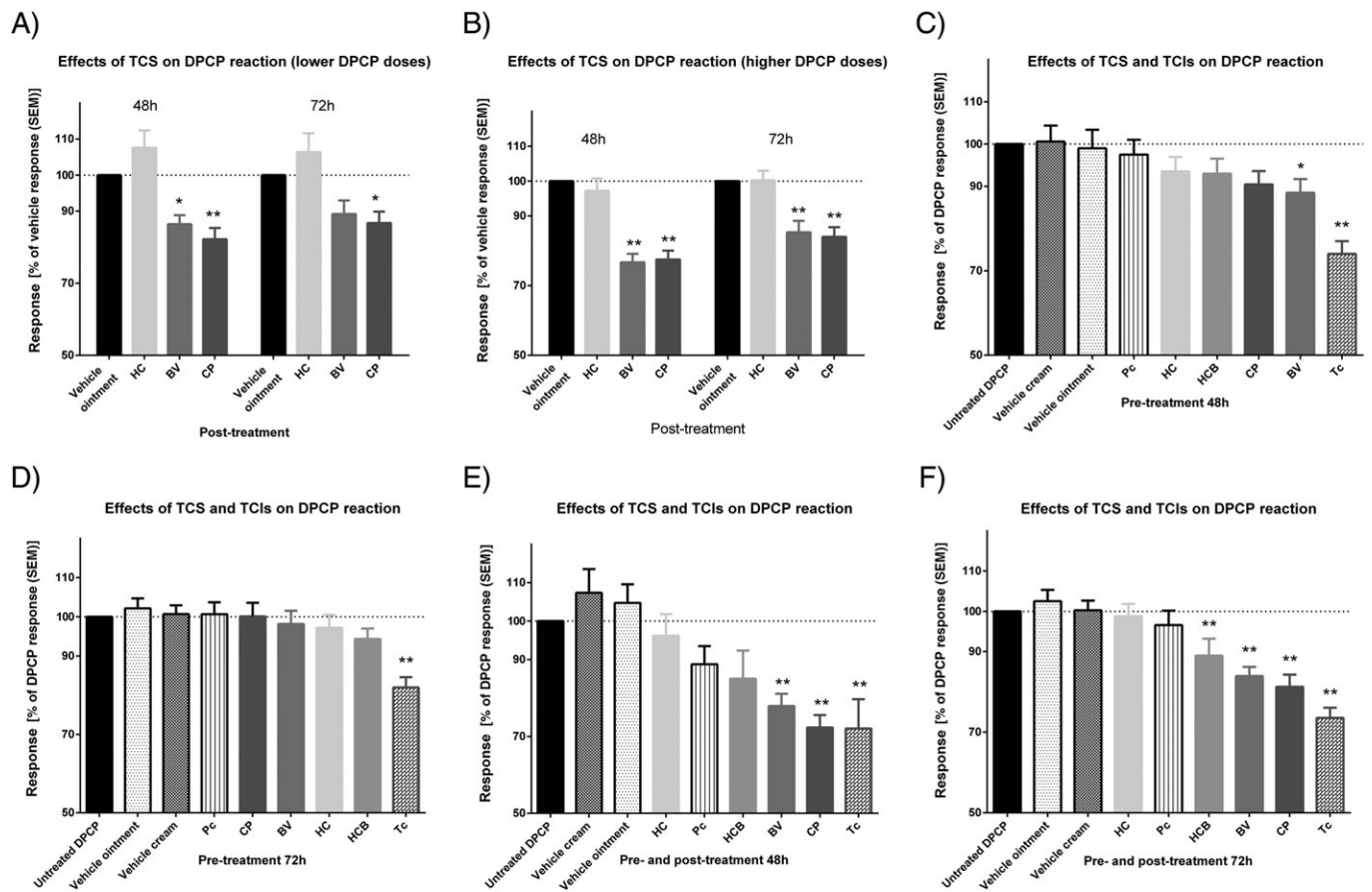
Figure 3

Anti-inflammatory effects of topical drugs evaluated by skin ultrasound. (A, B) Post-treatment at 48 h and 72 h using lower and higher diphenylcyclopropenone (DPCP) doses; (C, D) pre-treatment at 48 h and 72 h; (E, F) combined pre- and post-treatment for 48 h and 72 h. SEM: standard error of the mean; TCS: topical corticosteroids; TCIs: topical calcineurin inhibitors; Pc: pimecrolimus; Tc: tacrolimus; HC: hydrocortisone; HCB: hydrocortisone butyrate; BV: betamethasone valerate; CP: clobetasol propionate. Asterisks indicate * $P < 0.05$ and **$P < 0.01$

Discussion

These studies showed that an *in vivo* human model using an inflammatory response driven by contact hypersensitivity to DPCP succeeded in ranking the anti-inflammatory effect of four topical corticosteroid ointment preparations and two topical calcineurin inhibitors. Both visual assessment and objective measurement of DPCP induced skin inflammation (skin ultrasound thickness) were used. Three modes of application of the test drugs were explored; post-treatment, pre-treatment, and combined pre- and post-treatment of DPCP-induced allergic patch test reactions. Post-treatment was used in the first study; however, this study only revealed a significant anti-inflammatory effect for clobetasol propionate and betamethasone valerate. In the second study, two calcineurin inhibitors were included in addition to the four corticosteroids. The only topical drug showing a significant effect upon pre-treatment was tacrolimus ointment, while the combined pre-and post-treatment design was successful in ranking all six products in the following order: tacrolimus 0.01% ointment ≥ clobetasol propionate ointment ≥ betamethasone valerate ointment ≥ hydrocortisone-17-butyrate ointment > pimecrolimus cream = hydrocortisone ointment = vehicle ointment and cream and untreated (positive) DPCP control.

DPCP elicited controlled, reliably reproducible, and acceptable degrees of allergic patch test reactions in sensitized volunteers. There were no dropouts, and several volunteers participated in 2–3 studies without noticeable adverse side effects. We have previously demonstrated that repeated DPCP challenges in newly sensitized volunteers results in a clinically reproducible and constant level of contact allergy measured by several *in vivo* and *in vitro* methods [14, 15]. This constitutes the ethical and scientific basis for the possibility of asking volunteers to participate in repeated tests for the purpose of developing a human model for testing topical anti-inflammatory drugs. Previous studies attempted to develop appropriate *in vivo* models to assess the anti-inflammatory effects of drugs using various types of inflammatory challenges. Experiments in human volunteers using skin irritants (e.g. nonanoic acid, sodium lauryl sulphate and sodium dodecyl sulphate) as model inflammatory agents showed that irritant skin reactions cannot be used as a tool in models testing the anti-inflammatory effects of topical drugs [21–23]. Possible explanations for these failures include difficulty in standardization with regards to the choice of irritant, determining an adequate irritant concentration to elicit suitable irritant reactions, a pronounced interindividual variation in skin irritancy studies, and the lack of effect of systemically administered prednisone on irritant patch test reactions [21–25].

Several *in vivo* models have been developed that aimed to quantify the relative anti-inflammatory potencies of topical corticosteroids using allergic hypersensitivity reactions elicited by environmental allergens in spontaneously sensitized patients. Most models used post-treatment application of topical drugs because this represents the normal situation for clinical usage.

Some studies have showed differences in the anti-inflammatory effects of topical drugs on patch test reactions. Alomar *et al.* [5] used visual scoring of nickel-induced contact allergic reactions in a randomized, ointment-controlled,

double-blind study. It was shown that topical tacrolimus 0.1% and mometasone furoate 0.1% applied under occlusion for 48 h as post-treatment had significant inhibitory effects [5]. Queille-Roussel *et al.* also used nickel-induced allergic reactions and reported that twice daily open application of pimecrolimus 0.2% and 0.6% was superior to vehicle and equal to betamethasone valerate 0.1% [4].

Only a few studies have used DNCB- and nickel-induced allergic reactions to test the relative efficacy and safety of tacrolimus *in vivo* in humans by means of visual scoring [7, 26, 27]. These studies showed that tacrolimus was more effective than the corresponding vehicle in ameliorating allergic reactions to nickel and DNCB, respectively [7, 26, 27]. Not only is tacrolimus effective in suppressing allergic skin reactions in various animals and in man during the elicitation phase [28, 29], there is also additional evidence to suggest that it may be efficacious when skin is treated before elicitation of allergic reactions [26, 29]. This prechallenge treatment effect may be explained on the basis that calcineurin inhibitors will exert significant suppressive effects on the cytokine response of resident memory T cells in the skin, thereby reducing their capacity to initiate the inflammatory process following exposure to an allergen [30, 31].

While the use of an experimental challenge with contact allergens would seem to be a highly appropriate and relevant form of inflammatory reaction for analysis of the anti-inflammatory effects of drugs, use of spontaneously arising contact allergies in humans presents problems. First, suitable experimental subjects must be selected from the patient population who have spontaneous contact allergy to the desired allergen. Secondly, the process of eliciting reproducible responses with an appropriate strength of response is not straightforward. Many dose–response studies in nickel-sensitive subjects have shown large intraindividual variation in the threshold dose that elicits a positive patch test reaction. The highest dose difference observed was 250-fold [32]. Therefore, to obtain responses of similar intensity across test subjects, it is necessary to perform similar prechallenges to those we undertook. Thus, a prechallenge with a dose-series could indicate the appropriate challenge doses to elicit mild to moderate inflammatory responses. The advantage of an experimental sensitizer such as DPCP is that it can be used on all healthy human volunteers. DPCP is a strong sensitizer that is able to sensitize virtually 100% of healthy humans. Stable and reproducible responses to DPCP skin challenge develop after the second exposure. By dose-titration, it is possible to elicit low level inflammation, which creates a suitable inflammatory reaction to be able to discriminate between the different topical anti-inflammatory drugs. Strong contact hypersensitivity reactions present too big a therapeutic challenge for existing topical anti-inflammatory drugs, which reduces the possibility of revealing differences in their potency.

Another important aspect of the DPCP model is the short application time (6 h) of the allergen challenge. This means that the treatment with the drugs can be started before the reaction is fully developed and increases the possibility of detecting effects that inhibit the processes involved in generating the inflammation. The main problem with initiating anti-inflammatory treatment 48 h after the activation of the contact hypersensitivity response is that the reaction is virtually fully developed by that time, so the *task* for the

anti-inflammatory drugs is harder or possibly even different. Our data show that the combination of pre- and post-treatment augments the sensitivity of detection.

Because of the incomplete block design in study II, we could not select vehicle ointment as a reference point as we did in study I. It appears from Figure 3E and 3F that vehicle (ointment and cream) enhanced the DPCP reactions, probably due to increased absorption.

This is the first human *in vivo* model that has succeeded in comparing and ranking the anti-inflammatory potency of various types of topical drugs, including both corticosteroids and calcineurin inhibitors. The DPCP model in healthy volunteers is a promising alternative to the vasoconstrictor assay, which is only relevant for corticosteroids. Furthermore, the traditional classification of topical corticosteroids has been questioned due to the fact that a strong vasoconstrictor effect, as determined by the vasoconstrictor assay, is not necessarily equivalent to a strong anti-inflammatory effect [33].

To increase the chances of the model being able to detect more subtle differences between different active agents design modifications such as use of repeated applications of the drugs under test may be introduced to allow for a cumulative effect over time. Also, effects may be distinguished over longer repetitive challenges; hence, combining repeated DPCP exposures and drug applications could increase the sensitivity. This has proven useful in a few studies with nickel and tacrolimus [7, 27]. However, when several topical drugs are compared in each individual, we consider that occluded patches are needed to effectively control exposure.

In conclusion, the DPCP model developed here could detect significant anti-inflammatory effects for the three strongest groups of corticosteroids and tacrolimus, and offers a highly relevant alternative to the vasoconstrictor assay for topical drug potency testing.

Competing Interests

Contributors

K.F.M. and K.E.A. initiated the present studies. K.F.M. drafted the protocols and prepared the experimental setup under the guidance and input from all authors. K.F.M. performed the clinical experiments with technical assistance. K.F.M. drafted the manuscript in cooperation with K.E.A. and P.S.F. All authors critically revised the manuscript. The final version of this manuscript has been read and approved by all authors.

References

1 McKenzie AW, Stoughton RB. Method for comparing percutaneous absorption of steroids. Arch Dermatol 1962; 86: 608–10.

2 Kaidbey KH, Kligman AM. Assay of topical corticosteroids: efficacy of suppression of experimental rhus dermatitis in humans. Arch Dermatol 1976; 112: 808–10.

3 Rees JL, Matthews JNS, Friedmann PS. Quantifying anti-inflammatory agents' potency by measurement of response to dinitrochlorobenzene challenge. J Dermatol Sci 1992; 4: 1–5.

4 Queille-Roussel C, Graeber M, Thurston M, Lachapelle JM, Decroix J, De Cuyper C, *et al*. SDZ ASM 981 is the first non-steroid that suppresses established nickel contact dermatitis elicited by allergen challenge. Contact Dermatitis 2000; 42: 349–50.

5 Alomar A, Puig L, Gallardo CM, Valenzuela N. Topical tacrolimus 0.1% ointment (Protopic) reverses nickel contact dermatitis elicited by allergen challenge to a similar degree to mometasone furoate 0.1% with greater suppression of late erythema. Contact Dermatitis 2003; 49: 185–8.

6 Meshkinpour A, Sun J, Weinstein G. An open pilot study using tacrolimus ointment in the treatment of seborrheic dermatitis. J Am Acad Dermatol 2003; 49: 145–7.

7 Saripalli YV, Gadzia JE, Belsito DV. Tacrolimus ointment 0.1% in the treatment of nickel-induced allergic contact dermatitis. J Am Acad Dermatol 2003; 49: 477–82.

8 Warshaw EM, Wohlhuter RJ, Liu A, Zeller SA, Wenner RA, Bowers S, *et al*. Results of a randomized, double blind, vehicle controlled efficacy trial of pimecrolimus 1% for the treatment of moderate to severe facial seborrheic dermatitis. J Am Acad Dermatol 2007; 57: 257–64.

9 De Lacharriere O, Lachaise K, Kalis B. Evolution of skin inflammation and barrier function parameters during contact allergy to nickel after application of anti-inflammatory agents. J Invest Dermatol 1989; 92: 497.

10 Queille-Roussel C, Duteil L, Padilla JM, Poncet M, Czernielewski J. Objective assessment of topical anti-inflammatory drug activity on experimentally induced nickel contact dermatitis: comparison between visual scoring, colorimetry, laser Doppler velocimetry and transepidermal water loss. Skin Pharmacol 1990; 3: 248–55.

11 Seidenari S, Di Nardo A, Giannetti A. Assessment of topical corticosteroid activity on experimentally induced contact dermatitis: echographic evaluation with binary transformation and image analysis. Skin Pharmacol 1993; 6: 85–91.

12 Di Nardo A, Giusti G, Mantovani L, Bianchi B, Seidenari S. Inhibition of elicitation of contact dermatitis in humans by mometasone furoate: evaluation by means of 20-MHz B scanning associated with image analysis. Dermatology 1997; 195: 137–41.

13 Seidenari S, Di Nardo A, Mantovani L, Giannetti A. Parallel intraindividual evaluation of the vasoconstrictory action and the anti-allergic activity of topical corticosteroids. Exp Dermatol 1997; 6: 75–80.

14 Mose KF, Andersen F, Skov L, Røpke MA, Litman T, Friedmann PS, *et al*. Repeated monthly epicutaneous challenges with diphenylcyclopropenone result in a clinically reproducible level of contact allergy in de novo sensitized individuals. Br J Dermatol 2017; 176: 1095–7.

15 Mose KF, Burton M, Thomassen M, Andersen F, Kruse TA, Tan Q, *et al*. The gene expression and immunohistochemical time-course of diphenylcyclopropenone induced contact allergy in healthy humans following repeated epicutaneous challenges. Exp Dermatol 2017; 26: 926–33.

16 Harper Smith AD, Coakley SL, Ward MD, Macfarlane AW, Friedmann PS, Walsh NP. Exercise-induced stress inhibits both the induction and elicitation phases of *in vivo* T-cell-mediated immune responses in humans. Brain Behav Immun 2011; 25: 1136–42.

17 Johansen JD, Aalto-Korte K, Agner T, Andersen KE, Bircher A, Bruze M, *et al*. European society of Contact Dermatitis guideline for diagnostic patch testing – recommendations on best practice. Contact Dermatitis 2015; 73: 195–221.

18 Hindsén M, Bruze M, Christensen OB. The significance of previous allergic contact dermatitis for elicitation of delayed hypersensitivity to nickel. Contact Dermatitis 1997; 37: 101–6.

19 Cochran WG, Cox GM. Experimental Designs. New York: John Wiley and Sons, 1957.

20 Curtis MJ, Bond RA, Spina D, Ahluwalia A, Alexander SPA, Giembycz MA, *et al*. Experimental design and analysis and their reporting: new guidance for publication in BJP. Br J Pharmacol 2015; 172: 3461–71.

21 Van der Valk PG, Maibach HI. Do topical corticosteroids modulate skin irritation in human beings? Assessment by transepidermal water loss and visual scoring. J Am Acad Dermatol 1989; 21: 519–22.

22 Le TK, De Mon P, Schalkwijk J, van der Valk PG. Effect of a topical corticosteroid, a retinoid and a vitamin D3 derivative on sodium dodecyl sulphate induced skin irritation. Contact Dermatitis 1997; 37: 19–26.

23 Clemmensen A, Andersen F, Petersen TK, Hagberg O, Andersen KE. Applicability of an exaggerated forearm wash test for efficacy testing of two corticosteroids, tacrolimus and glycerol, in topical formulations against skin irritation induced by two different irritants. Skin Res Technol 2011; 17: 56–62.

24 Judge MR, Griffiths HA, Basketter DA, White IR, Rycroft RJ, McFadden JP. Variation in response of human skin to irritant challenge. Contact Dermatitis 1996; 34: 115–7.

25 Anveden I, Lindberg M, Andersen KE, Bruze M, Isaksson M, Lidén C, *et al*. Oral prednisone suppresses allergic but not irritant patch test reactions in individuals hypersensitive to nickel. Contact Dermatitis 2004; 50: 298–303.

26 Lauerma AI, Maibach HI, Granlund H, Erkko P, Kartamaa M, Stubb S. Inhibition of contact allergy reactions by topical FK506. Lancet 1992; 340: 556.

27 Belsito D, Wilson DC, Warshaw E, Fowler J, Ehrlich A, Anderson B, *et al*. A prospective randomized clinical trial of 0.1% tacrolimus ointment in a model of chronic allergic contact dermatitis. J Am Acad Dermatol 2006; 55: 40–6.

28 Meingassner JG, Stütz A. Immunosuppressive macrolides of the type FK506: a novel class of topical agents for treatment of skin diseases? J Invest Dermatol 1992; 98: 851–5.

29 Lauerma AI, Stein BD, Homey B, Lee CH, Bloom E, Maibach HI. Topical FK506: suppression of allergic and irritant contact dermatitis in the guinea pig. Arch Dermatol Res 1994; 286: 337–40.

30 Schreiber SL, Crabtree GR. The mechanism of action of cyclosporine A and FK506. Immunol Today 1992; 13: 136–42.

31 Tocci MJ, Matkovich DA, Collier KA, Kwok P, Dumont F, Lin S, *et al*. The immunosuppressant FK506 selectively inhibits expression of early T cell activation genes. J Immunol 1989; 143: 718–26.

32 Hindsén M, Bruze M, Christensen OB. Individual variation in nickel patch test reactivity. Am J Contact Dermat 1999; 10: 62–7.

33 Humbert P, Guichard A. The topical corticosteroid approach called into question: towards a new approach. Exp Dermatol 2015; 24: 393–5.