

Reproducibility of patch test results: a concurrent right-versus-left study using TRUE TestTM

S. IRIS ALE¹ AND HOWARD I. MAIBACH²

¹Department of Dermatology, School of Medicine, University of Uruguay, Arazati 1194, PC 11300, Montevideo, Uruguay, and ²Department of Dermatology, University of California, 90 Medical Center Way, Surge 110, PO Box 0989, San Francisco, CA 94143-0989, USA

Wide variations in the reproducibility rate of positive patch test responses have been reported. We hypothesized that a major source of non-reproducibility resides in the methodological pitfalls of routine patch testing. Simultaneous duplicate patch testing on opposite sides of the upper back was performed on 500 consecutive patients, using the TRUE TestTM system consisting of 2 panels, each one containing 12 standard allergens. A rigorous methodological design was applied and relevance was assessed for all discordant patch test reactions. A total of 435 positive patch test reactions were observed either on one or both sides in 289 patients (58.8%). Of these, 22 (5%) were discordant, i.e. interpreted as positive allergic on one side whilst negative or doubtful on the opposite side. The allergens responsible for discordant reactions were nickel sulfate (4 patients), cobalt chloride (3), lanolin alcohol (3), fragrance mix (2), carba mix (2), thiuram mix (2), colophonium (1), potassium dichromate (1), *p*-phenylenediamine (1), formaldehyde (1), balsam of Peru (Myroxylon pereirae resin) (1) and thimerosal (1). Of the 19 (4%) patients with discordant patch test reactions, the allergen was deemed to be of definite present or past relevance in 9 patients (1.8% of the total and 3.1% of all patients with positive patch test reactions) and of possible relevance in a further 2 patients. These data suggest that patch testing is a reasonably reproducible procedure as long as methodological error is minimized.

Key words: allergen; allergic contact dermatitis; clinical relevance; patch testing; reproducibility.
© Blackwell Munksgaard, 2004.

Accepted for publication 13 April 2004

Patch testing, an essential diagnostic and investigative tool for the study of cutaneous delayed hypersensitivity, has been in continuous development since its inception, yet still faces methodological and interpretative problems (1, 2). A perfect diagnostic test should be accurate and totally reproducible; *ergo*, the test results should vary only as a function of the phenomenon being investigated. However, inconsistencies in methodology or in interpretation contribute to the variability of most diagnostic tests.

Reproducibility, the capacity of a test to produce consistent results when being performed under the same conditions, should be studied and incorporated in validation of a test's diagnostic utility. Reproducibility of patch testing remains a matter of controversy; great variations in the reproducibility rate of positive responses have been reported (3–13). Golhausen et al. (3) double-tested concomitantly on the left and right sides of the upper back 35 patients with allergens

from the standard series and an ointment series and found that 43.8% of the positive allergic reactions were non-reproducible (Table 1). Subsequently (4), they reported a higher incidence of non-reproducibility of duplicate patch testing using allergens in petrolatum (pet.) and Finn ChambersTM (37.9%) when compared with TRUE TestTM (17.9%), and suggested that the low reproducibility of patch testing is inherent in the method. In a multicentre study from the German Contact Dermatitis Research Group (5), 1285 patients were double-tested concomitantly with 10 allergens from the standard series and manually loaded patch test systems. Non-reproducible allergic reactions were seen in 194 patients (15.1%). The authors concluded that the non-reproducible results were highly dependent on the allergen tested, whilst other factors, such as technical or methodological variations, were of minor importance. In contrast with these studies, Lindelöf (7) obtained a non-reproducibility rate

Table 1. Reproducibility of concomitant patch test reactions. Comparison between right and left sides of the upper back

Reference	Allergens and patch test system	Exposure time (h)	Reading time (h)	Number of patients	Results	% ^a	Commentaries
(3)	39 allergens from the standard series plus an ointment battery with Finn Chambers TM	48	48 and 72 h after application	35	32 (total number of positive reactions, either in one or both sides) 18 reproducible positive reactions (both sides) 14 non-reproducible positive reactions (one side only)	43.8	It is not stated whether standardized commercial grade allergens were used. If the responses to cetylpyridinium and gentamycin are not considered the rate of non-reproducibility would be 25%
(4)	12 allergens with TRUE Test TM patch test system	48	48 and 72	63	28 (total number of positive reactions) 23 reproducible positive reactions 5 non-reproducible positive reactions	17.9	
(4)	12 allergens with Finn Chambers TM	48	48 and 72	76	58 (total number of positive reactions) 36 reproducible positive reactions 22 non-reproducible positive reactions	37.9	
(5)	10 allergens from the standard series (Trolab) using either Finn Chambers TM , HAL or Leukotest	24 in 647 48 in 638	72	1285	1857 (total number of positive reactions)	15.1	The degree of non-reproducibility was dependent on the tested allergen Non-reproducibility was higher in subjects with more than 4 positive reactions
(6)	26 standard allergens with Finn Chambers TM	48	72 or 96	121	166 (total number of positive reactions) 131 reproducible positive reactions 35 non-reproducible positive reactions	21.1	
(7)	10 standard allergens in pet. (Chemotechnique) with Finn Chambers TM	48	72	220	84 (total number of positive reactions) 76 reproducible positive reactions 8 non-reproducible positive reactions	9.5	
(8)	ICDRG standard series plus wood tar	48	48 and 72	44	28 (total number of positive reactions) 26 reproducible positive reactions 2 non-reproducible positive reactions	7	All reproducible positive reactions had identical score grade
(10)	19 allergens (Trolab) with the Epiquick test system	48	72	100	70 (total number of positive reactions) 67 reproducible positive reactions 3 non-reproducible positive reactions	4.2	64 of the 67 (95.5%) reproducible positive reactions had identical score grade
(12)	10 allergens (standard series) with Finn Chambers TM	48	48 and 96	383	32 completely discordant reactions (definite positive on one side, definite negative on the other side) 70 partially discordant reactions	8	Discordant reactions were of clinical relevance in only 7 patients

^aRatio of non-reproducible positive reactions.

of 9.5%, Bousema et al. (8) reported 7% of non-reproducibility and Bourke et al. (12) reported a rate of 8% of completely discordant results. Studies using ready-to-apply patch test systems have reported even higher reproducibility rates. Lachapelle (10) using EpiquickTM observed a non-reproducibility rate of only 4.2% (Table 1).

Several factors influence the elicitation of a positive allergic result in a sensitized subject and many sources of variability may account for non-reproducible results, namely, patch test materials, technique and methodology and inherent biological variability of patch test responses (Table 2). We hypothesize that the main source of non-reproducibility resides in the methodological pitfalls of routine patch testing. This study reassesses the reproducibility of patch testing, applying a rigorous methodological design (strict criteria of subjects' inclusion, standardized patch test materials, optimized application of tests, defined application and reading times, standardized interpretation of the responses and correlation with clinical relevance) (Table 3) that enables discrimination between the sources of unreliability in diagnostic patch testing.

Subjects and Methods

The study was approved by the Ethics Committee of the University Hospital, Montevideo, Uruguay.

Subjects

500 consecutive patients attending the contact dermatitis clinic at the University Hospital in Uruguay over 11 months were included.

Criteria for inclusion in the study were:

- (1) Contact dermatitis suspected from history and clinical examination
- (2) Over 15 years old
- (3) No active dermatitis (except for mild localized dermatitis)
- (4) No active dermatitis of any degree on or around the test area
- (5) No systemic corticosteroid or immunosuppressive treatment during the previous 30 days
- (6) No topical treatment with corticosteroids during the previous 15 days on or around the test area
- (7) Not pregnant (by history)

Clinical data were recorded in a standardized manner, including details of the present dermatitis, history of previous dermatitis, personal and family atopy and/or other skin diseases, occupational and non-occupational exposures, including hobbies, topical and systemic therapies, personal skin care, etc. Patients were instructed not to shower, bathe or engage in sports or heavy physical efforts while the patches were on place. A complete and clear instruction form was provided to assist patient compliance.

Table 2. Sources of variability in diagnostic patch testing

Materials

- type of patch test system (30)
- different sources of patch test allergens
- different vehicles and concentrations for some allergens (34–36, 38–46)
- uneven distribution of allergens in the vehicle (44)

Methodology and Technical drawbacks

- criteria of patient's selection (27)
- application and reading times (47–54)
- amount of allergen applied (9, 31–33)
- errors in the sequence of consecutive allergens (9)
- dissimilar pressure supported by the system according to the area of application (56–58)
- partial or complete detachment of patches (9)
- interpretation of the responses (and)
 - intraobserver and interobserver variability (55)
 - scoring scale used

Biological

- regional variations in skin absorption and responsiveness (56–58, 64)
- cyclic variations throughout the menstrual cycle (62, 63)
- weak and doubtful responses
- summation of individual responses
- overlooked intercurrent factors:
 - drugs, sun exposure, etc.
 - active dermatitis
- skin hyporeactivity and 'silent back syndrome' (60)
- skin hyperreactivity and 'excited skin syndrome' (28, 59)

Table 3. Experimental design for studying reproducibility of patch testing

Sufficiently large population to obtain statistically significant data
Standardized criteria of patient's selection, specially regarding opportunity of patch testing
Complete and clear instructive to the patient to assure patient's compliance
Defined and standardized patch testing procedure:
– patch test system
– application time
– reading schedule
– score grading of the responses
– technical details
Evaluation of all discordant results:
– retesting in duplicate all non-reproducible reactions
– performing serial dilutions
– assessing clinical relevance (re-interrogation, ROAT, PUT)
Evaluation of weak and doubtful reactions (same as above)

ROAT = repeated open application test; PUT = provocative use test.

Materials

The TRUE TestTM ready-to-use patch test system (Pharmacia and Upjohn, Uppsala, Sweden), consisting of 2 panels (I and II) each one containing 12 allergens in a prearranged sequence, was used (14). The allergens are dissolved, emulsified or micronized into a gel, printed on a plastic sheet, dried and cut into 0.81 cm² (9 × 9 mm) patches. The patches with allergens are attached on a tape forming a test kit. This test has been pharmaceutically optimized concerning stability, solubility and bioavailability. The dose has been determined by dose–response studies and the amount per unit area standardized (14–22).

Methodology

Duplicate patch test series were applied simultaneously on the left and the right side of the upper back. Both panels were arranged symmetrically, although the allergen sequence was maintained, thus a complete mirror image was not obtained (Fig. 1). Special care was taken to obtain adherence. A 4-cm wide area of skin on either side of the midline was not used, as patches applied to that area frequently become detached. An extra strip of 2-inch-wide tape (MicroporeTM Surgical Tape, 3 M, St. Paul, MN, USA) was placed over the edges of the patch test units for reinforcement. No other allergens were applied at the same time.

The variables investigated were: responsiveness, i.e. the character and strength of the reactions; quality, i.e. the reaction area compared to the test area; tape adhesiveness; tape irritation; and other side effects. Patches were removed at 48 h, and readings were made at 49, 72, 96 h and 1 week after patch test application according to ICDRG criteria (23). All readings were performed by the same investigator (SIA). Readings were classified as negative, allergic, doubtful and irritant. Presumed allergic reactions were graded

as + (erythema, infiltration and possible papules), ++ (erythema, infiltration, papules and vesicles) or +++ (erythema, infiltration and coalescent vesicles). A reaction was interpreted as doubtful when only long-lasting erythema without infiltration was present or when an erythematous infiltrated reaction covered 50% or less of the test area. A patch test reaction was considered to be discordant when, being interpreted as true allergic on one side, it was negative, doubtful or deemed to be irritant on the other side. When discordant results were obtained, patients were retested 6 weeks later, to re-evaluate the results obtained at the first challenge. Clinical relevance was assessed for all positive reactions on the basis of the clinical history, examination and provocative testing according to published guidelines (24–27). Relevance was considered either present, when related to the current clinical symptoms, or past, when associated with previous clinical events, and was scored as likely or definite, possible, doubtful or

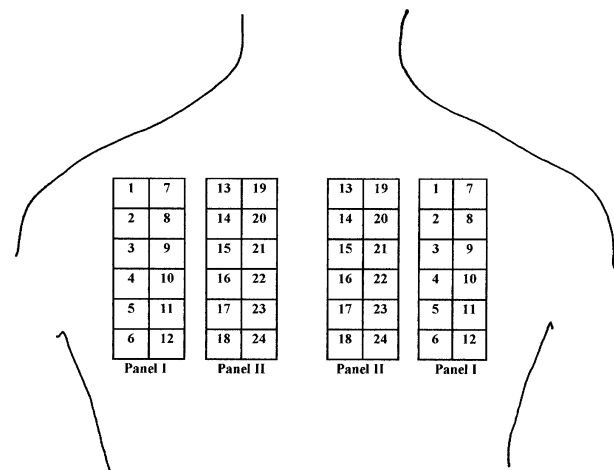


Fig. 1. Arrangement of the allergens (TRUE TestTM panels I and II) on the upper back.

unknown, according to a previously proposed scoring system (25).

The χ^2 -test and the two-sided Student's *t*-test were used for statistical comparisons between different subgroups of patients. A *P*-value of 0.05 or less was considered to indicate statistical significance.

Results

9 subjects were excluded from analysis for one of the following reasons: not available for a second reading (5 cases), multiple reactions compatible with excited skin syndrome (2 cases) (28), widespread skin eruption presumably related to medication (1 case), and patch test strip almost totally detached on one side (1 case).

Of the 491 subjects studied, 285 were women (58%) and 206 men (42%). The mean age was 42.4 years (range 15–77, SD 15.6). A total of 114 patients (23.2%) were considered to have an atopic background (29). Of these, 44 (8.9%) had a history of cutaneous atopy. The test was removed after 48 h, except in 2 cases in which it was removed at 72 h. The decisive reading (i.e. the reading utilized to assess the right–left reproducibility) was performed at 96 h in 459 patients at 72 h in another 31 and at 120 h in only 1 patient.

A total of 435 positive patch test reactions (+ to +++) were observed either on one or both

sides in 289 patients (58.8%). Of these, 413 reactions were concordant (interpreted as allergic on both sides) (Table 4); 22 reactions were discordant. The ratio of reproducible positive results to overall positive results calculated as a percentage and taken as a measure of reproducibility was 95%. 91% of the positive concordant reactions were scored identically in terms of intensity. Positive allergic reactions were observed for all allergens (Table 4). Allergic reactions to nickel were more frequent in women ($P < 0.01$), whilst potassium dichromate, carba mix and thiuram mix elicited more allergic reactions in men ($P < 0.05$). Doubtful and irritant reactions were seen in less than 1% of all test patches (Table 4). Discordant test reactions were observed in 19 (4%) patients. Of these, the allergen was deemed to be of definite present or past relevance in 9 patients (1.8% of the total and 3.1% of all patients with positive patch test reactions) and of possible relevance in 2 patients (Table 5). Patients with discordant patch tests were not different from the total study group in terms of epidemiological data. Neither side of the back was more likely to produce a discordant reaction. The discordant allergens were: nickel sulfate (4 patients), lanolin alcohol (3), cobalt chloride (3), carba mix (2), thiuram mix (2), fragrance mix (2), balsam of Peru (Myroxylon pereirae resin) (1), *p*-phenylenediamine (1), thimerosal (1), potassium

Table 4. Number of positive concordant and discordant, doubtful and irritant patch test responses

Allergen	Positive concordant	Positive discordant	Irritant or doubtful either on one or both sides
Nickel sulfate	83	4	14
Lanolin alcohol	7	3	4
Neomycin	13		1
Potassium dichromate	37	1	8
Caine mix	13		4
Fragrance mix	20	2	8
Colophonium	5	1	1
Epoxy resin	5		5
Quinoline mix	4		1
Balsam of Peru (Myroxylon pereirae resin)	19	1	7
Ethylenediamine	9		2
Cobalt	27	3	16
<i>p</i> - <i>tert</i> -butylphenol-formaldehyde resin	18		5
Paraben mix	4		2
Carba mix	24	2	4
Black rubber mix	2		3
MCI/MI	3		1
Quaternium-15	8		1
MBT	14		3
PPD	25	1	6
Formaldehyde	13	1	7
Mercapto mix	7		5
Thimerosal	8	1	2
Thiuram mix	45	2	8
Total	413	22	118

MCI/MI = methylchloroisothiazolinone/methylisothiazolinone; MBT = mercaptobenzoethiazole; PPD = *p*-phenylenediamine.

Table 5. Clinical data of the 19 patients with discordant patch test reactions

Sex	Age (years)	Occupation	Atopy	Dermatitis	Discordant allergen	Score (L/R)	Relevance	Retesting	Other reactions	Final diagnosis
M	61	Retired (clerk)	No	Hand (wrists) and foot dermatitis	PPD	?+/+	Definite (past)	Positive (+)	MBT (+++/++) Mercapto mix (+++/++) Nickel (+++/+)	ACD
M	71	Retired (clerk)	No	Leg eczema	Formaldehyde	+/-	Unknown	Negative	Carba mix (?+/-) Quinoline mix ++/++ Lanolin alcohol (+++/++) Fragrance mix (+++/+) Neomycin (+/+)	Stasis dermatitis + ACD (topical products)
F	41	Salesperson	No	Foot dermatitis	Potassium dichromate	+/?+	Possible	Positive (+)	Cobalt chloride (+/+) PTBPR (+++/++)	ACD shoes
F	27	Housewife	Yes	Hand dermatitis	Lanolin alcohol	+/-	Unknown	Negative	Nickel (IR/IR) Fragrance mix (+++/++) Nickel (+/+)	Atopic + allergic hand dermatitis
M	52	Construction worker	No	Hand and foot dermatitis	Thiuram mix	++/?+	Definite (present)	Positive (++)	Balsam of Peru (?+/?+) Dichromate (++/+)	ACD
F	29	Secretary	Yes	Face and hands dermatitis	Nickel sulfate	-/+	Definite (past)	Positive (+)	Nickel (IR/IR) Mercapto-mix (IR/-) Fragrance mix (++/++)	ACD (cosmetics) + hand dermatitis (irrit/atopic)
F	32	Housewife	Yes	Hand dermatitis	Lanolin alcohol	+/-	Unknown	Negative	Thiuram mix (+++/++) Nickel (+++/++)	ACD
F	19	Student	No	Metal contact areas	Nickel sulfate	-/+	Definite (present)	Positive (++)	Cobalt chloride (+/+)	ACD (metals)
M	33	Clerk	No	Arm and legs	Fragrance mix	+/-	Unknown	Negative		Nummular eczema (?)
F	44		No	Foot dermatitis	Cobalt chloride	-/+	Doubtful (past)	Negative	Dichromate (+++/++) Nickel sulfate (++/+)	ACD (shoes)
M	52	Cleaner		Hand dermatitis	Thiuram mix Carba mix	?+/++ -/+	Definite (present)	Positive (++) Positive (++)	Nickel sulfate (+/+)	ACD
F	31	Clerk	No	Face, hands and neck dermatitis	Fragrance mix	-/+	Definite (present)	Positive (+)	Balsam of Peru (+++/++)	ACD (fragrances)
F	51	Assistant nurse	No	Hand dermatitis	Nickel sulfate	++/-	Definite (past)	Not retested	Thiuram mix (+++/++) Carba mix (++/+)	Allergic hand dermatitis
M	66	Retired Clerk	No	Leg dermatitis	Cobalt chloride	+/-	Unknown	Negative	Balsam of Peru (++/++) Fragrance mix (+/+)	Stasis and ACD (topical products)
M	35	Electrical Engineer	Yes	Hand dermatitis	Lanolin alcohol	-/+	Unknown	Negative		Irritant dermatitis + atopic
F	47	Teacher	No	Hand and foot dermatitis	Nickel sulfate	-/+	Definite (past)	Positive (++)		Hyperkeratotic hand and foot dermatitis
M	19	Clerk	Yes	Hands, arms and legs	Balsam of Peru	+/-	Unknown	Negative		Atopic dermatitis
F	32	Factory worker	No	Face, hands and forearms dermatitis	Thimerosal	+/-	Unknown	Not retested		ACD
F	16	Student	Yes	Hands, neck, earlobes	Carba mix	-/+	Definite (present)	Positive (++)	Nickel (+/+) Thiuram mix (+++/++) Nickel (+++/++)	ACD

dichromate (1), formaldehyde (1) and colophonium (1) (Table 5).

The area of the positive test reaction was confined between 80 and 120% of the test size in 94% of the tests. Tape adhesiveness was complete in 96% of patients. Test strips came partly detached in 20 cases. Tape irritation was observed in 23 patients (4.6%) and was mostly mild. Severe tape irritation was not observed. No severe side effects or late reactions were reported. Clinical relevance for positive reactions was found in 76% of the cases.

16 of the 19 patients with discordant reactions were available for retesting. 10 of the discordant reactions were positive on retesting and 9 were negative (Table 5). Most of the reactions which were negative on retesting were not considered to be clinically relevant.

Discussion

The value of any test depends on its ability to yield the same result when reapplied to stable patients. Our results document that patch testing produces highly reproducible results when methodological variability is decreased. Much has been done to optimize and standardize patch test materials and methodology. However, systematic studies of important aspects are lacking and sources of variability remain. The outcome of an individual patch test not only depends on the existence of delayed hypersensitivity to the tested substance but also on the test concentration and the dose delivered, which will depend on the amount of percutaneous absorption achieved by the method of exposure. Skin absorption varies depending on the patch test system (30). Factors such as conformity to the skin surface and degree of occlusion could be responsible for differences in the kinetics of allergen penetration.

Variations in the amount of material applied can lead to erroneous results (9, 31, 32). Excessive amounts can provoke spillover and irritant reactions, whilst inadequate dosing may, conversely, result in false-negative and doubtful reactions. The amount of allergen applied with the Finn Chamber technique should approximate to 15 μ l (range 12–18 μ l), but, as a manually dispensed system, it varies within a wider range (9, 31). This variation was reported to be higher when testing allergens in solution (32, 33). Pre-prepared patch test systems prevent the inconsistencies in the concentration of the tested allergen, the amount of material applied and the possibility of error in the sequence of allergens.

The use of an appropriate vehicle is crucial as it influences the bioavailability and subsequent percutaneous penetration of the allergen (34–44).

Pet. remains the standard vehicle for most allergens, with the exception of the TRUE TestTM. Earlier studies demonstrated that patch test suspensions in pet. contained undispersed allergen particles, and both the particle size and number differed significantly between different test substances and different manufacturers (41–46). The release of some allergens such as metallic salts from the vehicle is non-homogeneous, resulting in false-positive reactions (45). Additionally, the concentration for some allergens, such as nickel, fluctuated significantly from batch to batch in commercial materials (46). These factors may have influenced the results in the earlier studies on patch-test reproducibility. Improvements are being made in the optimization of patch test preparations and the dispersion of allergens.

Uniform criteria of patient inclusion and characterization of the population being tested are crucial to achieve comparable results, as is standardization of application and reading times (47–54). Paired readings are more accurate than single readings, test adherence can be verified at removal on day 2, and later readings allow the reaction fully to develop (51, 53, 54). As for any test that involves human perception and judgement, patch testing is also bedevilled by variability in reading and interpretation. Inter-observer and intraobserver variability is an inevitable consequence of the use of perceptual tests and constitutes one of the more significant limitations of the method (55).

All the above-mentioned factors may bias the reproducibility rate and should be carefully considered, especially in multicentre studies. Yet, methodological factors do not explain all discordant reactions; additional sources of discrepancies are even more difficult to appraise and overcome. Variability is inherent to all biological processes and several individual sources of unreliability exists, namely, regional and intraregional variations in skin absorption and responsiveness (e.g. upper versus lower back) (56–58), skin hyperreactivity, i.e. excited skin syndrome (28, 59), or hypo-reactivity, i.e. silent back syndrome (60), subclinical sensitization (61), miscellaneous factors, such as cyclical variations throughout the menstrual cycle (62, 63), seasonal variations (65, 66), etc. All the above may account for inconsistency.

The validity of any test system is its intrinsic ability to detect which individuals have the target disease and which do not, and depends on its capability to detect both true-positive and true-negative reactions, while minimizing the number of false-positive and false-negative reactions. The most significant discussion concerns whether a positive test reaction represents an allergic

reaction in contradistinction to an irritant one. A clear distinction between irritant and allergic reactions is not always possible on morphological grounds. The inherent irritancy of some standard allergens, such as cobalt chloride, potassium dichromate, formaldehyde, mixes, etc., probably accounts for the non-reproducibility of testing with these substances. Our results show that discordant reactions are mostly seen with allergens that are marginally irritant, and many of them where not reproducible on retesting, which strengthens the assumption that some of them may be false-positive, irritant reactions.

These findings are coincident with those from previous studies. Belsito et al. (13) tested concomitantly 128 patients with 3 lots of an allergen to which they previously reacted positively. They found that cinnamic aldehyde and formaldehyde, which are tested at concentrations close to the level of irritancy, produced the higher number of non-reproducible reactions (27% and 21%, respectively). The comparative multicentre studies with TRUE TestTM and Finn ChambersTM indicate proximity between limits of weak sensitization and irritancy with nickel sulfate, potassium dichromate, cobalt chloride, balsam of Peru (Myroxylon pereirae resin), fragrance mix, carba mix and thimerosal (19, 20). In addition, some allergens such as lanolin alcohol and cobalt chloride frequently produce doubtful and irrelevant reactions (20, 21, 67, 68).

Confirmation of the accuracy of discordant patch test reactions requires individual, sequential testing of the involved substances. Using serial dilution of the allergen to elicit a dose-response may contribute further to rule out false-positive reactions. The greater concern, however, involves false-negative test reactions, as they may lead to failure to diagnose contact allergy. A false-positive reaction may be reasonably easily uncovered, while a false-negative reaction may lead to diagnostic failure with ongoing dermatitis and continuing distress for the patient. Discordant reaction rates must always be correlated with clinical relevancy; this yields insights into the range of false-positive or false-negative reactions and gives additional weight to the operational definition of allergic contact dermatitis (69). We found that discordant reactions were clinically relevant in only 3.1% of all patients with positive allergic reactions.

Our study indicates that methodological variability can lead to non-reproducibility and false-positive and false-negative results. Taken together, our results demonstrate that patch testing is a reasonably reproducible procedure as long as methodological error is minimized.

Acknowledgements

Bo Kreilgård, PhD, (Pharmacia and Upjohn) kindly provided the allergens.

References

1. Fischer T I, Hansen J, Kreilgård B, Maibach H I. The science of patch test standardization. *Immun Allergy Clin* 1989; 9: 417–434.
2. Fischer T, Maibach H I. Improved, but not perfect, patch testing. *Am J Contact Dermat* 1990; 1: 73–90.
3. Gollhausen R, Przybilla B, Ring J. Reproducibility of patch tests. *J Am Acad Dermatol* 1989; 21: 1196–1202.
4. Gollhausen R, Przybilla B, Ring J. Reproducibility of patch test results. Comparison of TRUE Test and Finn Chamber test results. *J Am Acad Dermatol* 1989; 21: 843–846.
5. Brasch J, Henseler T, Aberer W et al. Reproducibility of patch tests. A multicenter study of synchronous left-versus right-sided patch tests by the German Contact Dermatitis Research Group. *J Am Acad Dermatol* 1994; 31: 584–591.
6. Breit R, Agathos M. Quality control in epicutaneous testing. Reproducibility in the right-left comparison. *Hautarzt* 1992; 43: 417–421.
7. Lindelöf B. A left versus right side comparison study of Finn Chamber patch test in 200 consecutive patients. *Contact Dermatitis* 1990; 22: 288–289.
8. Bousema M T, Geursen A M, van Joost T H. High reproducibility of patch tests. (Letter). *J Am Acad Dermatol* 1991; 24: 322.
9. Lachapelle J-M, Antoine J L. Problems raised by the simultaneous reproducibility of positive allergic patch test reactions in man. *J Am Acad Dermatol* 1989; 21: 850–854.
10. Lachapelle J-M. A left versus right side comparative study of Epiguick patch test results in 100 consecutive patients. *Contact Dermatitis* 1989; 20: 51–56.
11. Machácková J, Seda O. Reproducibility of patch tests. (Letter). *J Am Acad Dermatol* 1991; 25: 732–733.
12. Bourke J F, Batta K, Prais L et al. The reproducibility of patch tests. *Br J Dermatol* 1999; 140: 102–105.
13. Belsito D V, Storrs F J, Taylor J S et al. Reproducibility of patch tests: a United States multicenter study. *Am J Contact Dermat* 1992; 3: 193–200.
14. Fischer T, Maibach H I. The thin layer rapid use epicutaneous test (TRUE-test), a new patch test method with high accuracy. *Br J Dermatol* 1984; 112: 63–68.
15. Kreilgård B, Hansen J. Aspects of pharmaceutical and chemical standardization of patch test materials. *J Am Acad Dermatol* 1989; 21: 836–838.
16. Prue C, Martinson M E, McAnally P M, Stagner W C. Post-marketing survey results of T.R.U.E. TEST, a new allergen patch test. *Am J Contact Dermat* 1998; 9: 6–10.
17. Rietschel R L, Marks J G, Adams R M et al. Preliminary studies of the TRUE Test patch test system in the United States. *J Am Acad Dermatol* 1989; 21: 841–843.
18. Lachapelle J-M, Bruynzeel D P, Ducombs G et al. European multicenter study of the TRUE TestTM. *Contact Dermatitis* 1988; 19: 91–97.
19. Ruheek-Forsbeck M. Comparative multicenter studies with TRUE Test and Finn Chambers in eight Swedish hospitals. *J Am Acad Dermatol* 1989; 4: 846–849.
20. Ruheek-Forsbeck M, Fischer T, Meding B et al. Comparative multicenter study with TRUE TestTM and Finn Chambers[®] in eight Swedish hospitals. *Acta Derm Venereol (Stockh)* 1988; 68: 123–128.
21. Stenberg B, Billberg K, Nordin L. Swedish multicenter study with TRUE Test, Panel 2. *Contact Dermatitis* 1989; 21: 518–523.
22. Fischer T, Kreilgård B, Maibach H I. The True value of the TRUE Test for allergic contact dermatitis. *Curr Allergy Asthma Rep* 2001; 1: 316–322.

23. Wilkinson D S, Fregert S, Magnusson B et al. Terminology of contact dermatitis. *Acta Derm Venereol* 1970; 50: 287–292.
24. Ale S I, Maibach H I. Clinical relevance in Allergic Contact Dermatitis. *Dermatosen Beruf Umwelt* 1995; 43: 119–121.
25. Lachapelle J M. A proposed relevance scoring system for positive allergic patch test reactions: practical implications and limitations. *Contact Dermatitis* 1997; 36: 39–43.
26. De Groot AC. Clinical relevance of positive patch test reactions to preservatives and fragrances. *Contact Dermatitis* 1999; 41: 224–226.
27. Ale S I, Maibach H I. Scientific basis of patch testing. Part III. *Dermatosen Beruf Umwelt* 2002; 50: 131–133.
28. Mitchell J C. The angry back syndrome: Eczema creates eczema. *Contact Dermatitis* 1975; 1: 193–194.
29. Williams H C, Burney P G, Pembroke A C, Hay R J. The U.K. Working Party's diagnostic criteria for atopic dermatitis in a population setting. III. Independent hospital validation. *Br J Dermatol* 1994; 131: 406–416.
30. Kim H O, Wester R C, Mc. Master J A, Bucks D A, Maibach H I. Skin absorption from patch test systems. *Contact Dermatitis* 1987; 17: 178–180.
31. Fischer T, Maibach H I. Amount of nickel applied with a standard patch test. *Contact Dermatitis* 1984; 11: 285–287.
32. Shaw D W, Zhai H, Maibach H I, Niklasson B. Dosage considerations in patch testing with liquid allergens. *Contact Dermatitis* 2002; 47: 86–90.
33. Moffit D L, Sharp L A, Sansom J E. Audit of Finn Chamber® patch test preparation. *Contact Dermatitis* 2002; 47: 334–336.
34. Marzulli F N, Maibach H I. Effect of vehicles and elicitation concentration in contact dermatitis testing (I). Experimental contact sensitization in humans. *Contact Dermatitis* 1976; 2: 325.
35. Marzulli F N, Maibach H I. Further studies of effects of vehicles and elicitation concentration in experimental contact sensitization testing in humans. *Contact Dermatitis* 1980; 6: 131–133.
36. Tanglertsampan C, Maibach H I. The role of vehicles in diagnostic patch testing. A reappraisal. *Contact Dermatitis* 1993; 29: 169–174.
37. Wester R, Maibach H I. Relationship of topical dose and percutaneous absorption in rhesus monkey and man. *J Invest Dermatol* 1976; 67: 518.
38. Atkinson J C, Rodi S B. Effects of vehicles and elicitation concentration in contact dermatitis testing. II. Statistical analysis of data. *Contact Dermatitis* 1976; 2: 330–334.
39. Lee E E, Maibach H I. Role of vehicles in diagnosing contact allergy: an update. *Exog Dermatol* 2002; 1: 107–111.
40. Mendelow A Y, Forsyth A, Florence A T, Baillie A J. Patch testing for nickel allergy. The influence of the vehicle on the response rate to topical nickel sulphate. *Contact Dermatitis* 1985; 3: 29–33.
41. Fischer T, Maibach H I. Patch test allergens in petrolatum: a reappraisal. *Contact Dermatitis* 1984; 11: 224–228.
42. Wahlberg J E. Petrolatum-A reliable vehicle for metal allergens? *Contact Dermatitis* 1980; 6: 134–136.
43. Van Ketel W G. Petrolatum again: an adequate vehicle in cases of metal allergy? *Contact Dermatitis* 1979; 5: 192–193.
44. Vanneste D, Martin P, Lachapelle J-M. Comparative study of the density of particles in suspensions for patch testing. *Contact Dermatitis* 1980; 6: 197–203.
45. Fisher T, Rystedt I. False-positive, follicular and irritant patch test reactions to metal salts. *Contact Dermatitis* 1985; 12: 93–98.
46. Nakada T, Hostýnek J J, Maibach H I. Nickel content of standard patch test materials. *Contact Dermatitis* 1998; 39: 68–70.
47. Kalimo K, Lammintausta K. 24 and 48 h allergen exposure in patch testing. Comparative study with 11 common contact allergens and NiCl₂. *Contact Dermatitis* 1984; 10: 25–29.
48. Skog E, Forsbeck M. Comparison between 24- and 48-hour exposure time in patch testing. *Contact Dermatitis* 1978; 4: 362–364.
49. Goh C L, Wong W K, Ng S K. Comparison between 1-day and 2-day occlusion times in patch testing. *Contact Dermatitis* 1994; 31: 48–50.
50. Manuskianti W, Maibach H I. 1- versus 2- and 3-day diagnostic patch testing. *Contact Dermatitis* 1996; 35: 197–200.
51. Mathias C G T, Maibach H I. When to read the patch test? *Int J Dermatol* 1979; 3: 127–128.
52. Todd J, Handley J, Metwali M. Day 4 is better than day 3 for a single patch test reading. *Contact Dermatitis* 1996; 34: 402–404.
53. Rietschel R L, Adams R M, Maibach H I et al. The case for patch test readings beyond day 2. *J Am Acad Dermatol* 1988; 18: 42–45.
54. Jonker M J, Bruynzeel D P. The outcome of an additional patch-test reading on days 6 or 7. *Contact Dermatitis* 2000; 42: 330–335.
55. Bruze M, Isaksson M, Edman B et al. A study on expert reading of patch test reactions: inter-individual accordance. *Contact Dermatitis* 1995; 32: 331–337.
56. Magnusson B, Hersle K. Patch test methods II. Regional variations of patch test responses. *Acta Derm Venereol* 1965; 45: 257–261.
57. Lindelöf B. Regional variations of patch test response in nickel-sensitive patients. *Contact Dermatitis* 1992; 26: 202–203.
58. Van Strien G A, Korstanje M J. Site variations in patch test responses on the back. *Contact Dermatitis* 1994; 31: 95–96.
59. Maibach H I. The E.S.S. – excited skin syndrome (alias the 'angry back'). In: *New Trends in Allergy*, Ring J, Burg G (eds): Berlin, Heidelberg, New York, Springer-Verlag, 1981.
60. Koehler A M, Maibach H I. Skin hypo-reactivity in relation to patch testing. *Contact Dermatitis* 2000; 42: 1–4.
61. Friedmann P S, Rees J L, White S I, Matthews N S. Low-dose exposure to antigen induces sub-clinical sensitization. *Clin Exp Immunol* 1990; 81: 507–509.
62. Rohold A E, Halkier Sorensen N, Andersen K E et al. Nickel patch test reactivity and the menstrual cycle. *Acta Derm Venereol* 1994; 74: 383–385.
63. McLelland J, Lawrence C M. Premenstrual exacerbation of nickel allergy. *Br J Dermatol* 1991; 125: 83–83.
64. Hindsén M, Bruze M, Christensen O B. Individual variation in nickel patch test reactivity. *Am J Contact Dermatitis* 1999; 10: 62–67.
65. Bruze M. Seasonal influence on routine patch test results. *Contact Dermatitis* 1986; 14: 184.
66. Uter W, Geier J, Land M et al. Another look at seasonal variation in patch test results. *Contact Dermatitis* 2001; 44: 146–152.
67. Carmichael A J, Foulds J S, Bransbury D S. Loss of lanolin patch test positivity. *Br J Dermatol* 1991; 125: 573–576.
68. Rystedt I. Evaluation and relevance of isolated test reactions to cobalt. *Contact Dermatitis* 1979; 5: 233–238.
69. Ale S I, Maibach H I. Operational definition of allergic contact dermatitis. In: *Advances in Dermatotoxicology*, New York, Taylor & Francis, 2000.

Address:
 S. Iris Ale
 Department of Dermatology
 School of Medicine, University of Uruguay
 Arazati 1194, PC 11300, Montevideo
 Uruguay
 Tel: +5982 487 2571
 Fax: +5982 622 0882
 e-mail: irisale@hc.edu.uy