

Increased sensitivity of patch testing by standardized tape stripping beforehand: a multicentre diagnostic accuracy study

HEINRICH DICKEL¹, BURKHARD KREFT², OLIVER KUSS³, MARGITTA WORM⁴, STEPHANIE SOOST⁴, JOCHEN BRASCH⁵, WOLFGANG PFÜTZNER⁶, JÜRGEN GRABBE^{7,8}, IRENA ANGELOVA-FISCHER⁷, PETER ELSNER⁹, JOACHIM FLUHR^{9,10}, PETER ALTMAYER¹ AND JOHANNES GEIER¹¹

¹Department of Dermatology and Allergology, Ruhr University Bochum, Bochum, ²Department of Dermatology and Venereology, ³Institute of Medical Epidemiology, Biostatistics, and Informatics, Martin Luther University of Halle-Wittenberg, Halle (Saale), ⁴Department of Dermatology and Allergy, Allergy-Center-Charité, Universitätsmedizin Berlin, Berlin, ⁵Department of Dermatology, University Hospitals of Schleswig-Holstein, Campus Kiel, Kiel, ⁶Department of Dermatology and Allergology, Philipps University Marburg, Marburg, ⁷Department of Dermatology, University Hospitals of Schleswig-Holstein, Campus Lübeck, Lübeck, Germany, ⁸Department of Dermatology, Kantonsspital, Aarau, Switzerland, ⁹Department of Dermatology and Allergology, Friedrich Schiller University Jena, Jena, ¹⁰Bioskin GmbH, Berlin, ¹¹Information Network of Departments of Dermatology (IVDK), Georg August University Göttingen, Göttingen, Germany

Background: As a modification of patch testing, the strip patch test was established to obtain more sensitive and reliable test results. Comparative data on diagnostic accuracy for both tests are missing.

Objectives: To compare the diagnostic accuracy of strip patch tests and patch tests in detecting sensitizations in patients with suspected allergic contact dermatitis by using patient history as the reference standard.

Patients/methods: In a multicentre, prospective, investigator-blinded study 790 patients were enrolled. The defined reference standard was established prior to patch testing. Patch tests were performed with nickel sulfate, potassium dichromate, and lanolin alcohol. Duplicate tests were simultaneously performed on both sides of the back, of which one randomly chosen side was tape stripped beforehand, according to a standardized procedure. Primary outcome was the difference in sensitivity between strip patch test and patch test.

Results: Seven hundred and eighty-seven patients were included in the analysis. Strip patch tests detected considerably more sensitization to nickel sulfate and potassium dichromate than patch tests: differences of sensitivities were 16.4% (95% CI, 8.7–24.1%) for nickel sulfate and 25.0% (95% CI, 8.9–41.0%) for potassium dichromate, both favouring the strip patch test.

Conclusions: The standardized strip patch test proved to be accurate and clinically safe and is promising to improve diagnosis of allergic contact dermatitis beyond the patch test.

Key words: contact dermatitis; diagnostic accuracy; lanolin alcohol; nickel sulfate; patch test; potassium dichromate; sensitivity and specificity; strip patch test. © John Wiley & Sons A/S, 2010.

Conflict of interests: The authors have declared no conflicts.

Disclosures: The multicentre study was supported with test material (Blenderm™ surgical tapes, Clippers) by 3M™ Medica, 3M Deutschland GmbH, Neuss, Germany.

Author contributions: H.D. developed the study protocol and was the coordinating investigator of the multicentre study; O.K. and J.G. participated in designing the study and the analysis of the data. All authors participated in generating or gathering the data for the study, in writing the manuscript, and have approved the final version of this manuscript. H.D., O.K., and J.G. are guarantors for this manuscript. No prior presentation on this study was done.

Accepted for publication 17 January 2010

Since its inauguration in 1895 by Jadassohn (1), the 'conventional' patch test is the mainstay of diagnosis in allergic contact dermatitis with an assumed sensitivity and specificity of between 70% and 80% (2, 3). There is little information presenting explicit figures on these standard indices of validity (4). The reported sensitivities and specificities for the North American Contact Dermatitis Group (NACDG) and International Contact Dermatitis Research Group (ICDRG) baseline series were 77% and 71% and 68% and 77%, respectively. Since then, the method of patch testing has been repeatedly criticized (5–7) for its limited validity and is therefore by no means accepted uniformly as a reliable test method.

Although a non-*in vivo* test would be desirable, to date, none of the *in vitro* tests for allergic contact dermatitis (e.g. lymphocyte transformation test) fulfils the requirements of an easy, accurate, and safe test with sufficient sensitivity and specificity (8).

As a modification of the patch test, tape stripping the skin prior to patch testing was first mentioned in 1953 by Spier and Natzel (9). Since then, the basic idea of the 'strip' patch test is to increase (i) bioavailability of a test substance in the deeper epidermal cell layers (10) and (ii) skin reactivity to the test substance (11) with the aim of obtaining more sensitive and reliable patch test results. We recently presented a standardized protocol for performing the strip patch test (12). This approach was adopted in 'Allergy diagnostic testing: An updated practice parameter' by the American Academy of Allergy, Asthma and Immunology and the American College of Allergy, Asthma and Immunology (13). Furthermore, we could show a significant increase of test sensitivity in strip patch test versus patch test by using serial dilution tests of nickel sulfate and potassium dichromate in patients with positive routine patch test (14). According to our results, the elicitation threshold concentrations of nickel sulfate and potassium dichromate could be lowered by a factor of 10 and 5, respectively.

This multicentre study presented here aimed to (i) estimate the diagnostic accuracy of strip patch tests and patch tests relative to a predefined 'gold standard' reference, and to (ii) compare the accuracy between these two tests with three common test substances from the baseline series (nickel sulfate 5% petrolatum (pet.), potassium dichromate 0.5% pet., and lanolin alcohol 30% pet.).

Patients and Methods

This multicentre, prospective, randomized, investigator-blinded study took place from 4 December 2006 to 31 March 2009. During this period, patients were consecutively screened in seven different patch

test units of tertiary care university-affiliated dermatological clinics. The study was first approved by the ethics committee of the Medical Faculty of the Ruhr University Bochum in November 2006 (registration no. 2881). The study protocol was then subject to review and approval by the local ethics committees of each of the participating centres. All procedures were explained in detail to all patients, who signed informed consent forms prior to participation.

Patients

Eligibility criteria for enrolment was the clinical suspicion of current or past allergic contact dermatitis because of presenting symptoms or referral by another dermatologist. For inclusion, patients had to be at least 18 years old and should have clinically normal skin. Exclusion criteria were pregnancy, lactation, intake of immunosuppressive or immune modulating drugs such as glucocorticoids or ciclosporin, topical use of glucocorticoids at the back, and intense UV exposure on the back within the last 4 weeks prior to the study enrolment, respectively. Discontinuation criteria for individual patients were severe skin reactions or intense discomfort in the test areas of the index tests, but these criteria were not met in any case.

Reference standard

External criteria for the reference standard 'by history delayed-type sensitization present: yes/no' were established by the decision of the German Contact Dermatitis Research Group (DKG) before the study was initiated (Table 1). A total weighting factor of ≥ 10 for nickel sulfate, potassium dichromate, and lanolin alcohol, respectively, was by definition the cut-off value for the presence or absence of delayed-type sensitization (Fig. 1).

Patients were classified as to whether or not they had a delayed-type sensitization to nickel sulfate, potassium dichromate, and lanolin alcohol, respectively, based on their response to the defined reference standard prior to undergoing the index tests. Reference standard and index tests were performed on the same day.

Index tests

Patients were tested on the upper back using small Finn Chambers® on Scanpor® (inner diameter 8 mm, capacity 20 µl; Epitest Ltd Oy, Tuusula, Finland). Test sites were on normal skin, which showed no signs of dermatitis. The index tests were applied to the left and right sides of the upper back of the patient, with one randomly chosen side tape stripped for the strip patch test. Randomization was

Table 1. Reference standard ($N = 787$)

Question	Yes		No	
	WF	N	WF	n
<i>Nickel sulfate</i>				
Intolerance to fashion jewellery?	5	174	0	613
Jeans button eczema?	5	110	0	677
Hand eczema with electroplating plating processes?	5	4	0	783
Sensitization to nickel sulfate proved by PT in the past?	5	102	0	685
<i>Potassium dichromate</i>				
Hand eczema through contact with cement or products containing cement?	5	29	0	758
Hand eczema through wearing leather gloves?	3	24	0	763
Eczema through wearing a leather bracelet or leather clothing?	5	25	0	762
Foot eczema through wearing leather shoes?	2	29	0	758
Sensitization to potassium dichromate proved by PT in the past?	5	27	0	760
<i>Lanolin alcohol</i>				
Eczema after application of topical products containing lanolin alcohol and other causes could be ruled out?	5	12	0	775
Lower leg eczema with varicose ulcer?	5	50	0	737
Sensitization to lanolin alcohol proved by PT in the past?	5	7	0	780
Contact allergy to lanolin alcohol established by ROAT?	10	0	0	787

WF, weighting factor; PT, patch test; ROAT, repeated open application test.

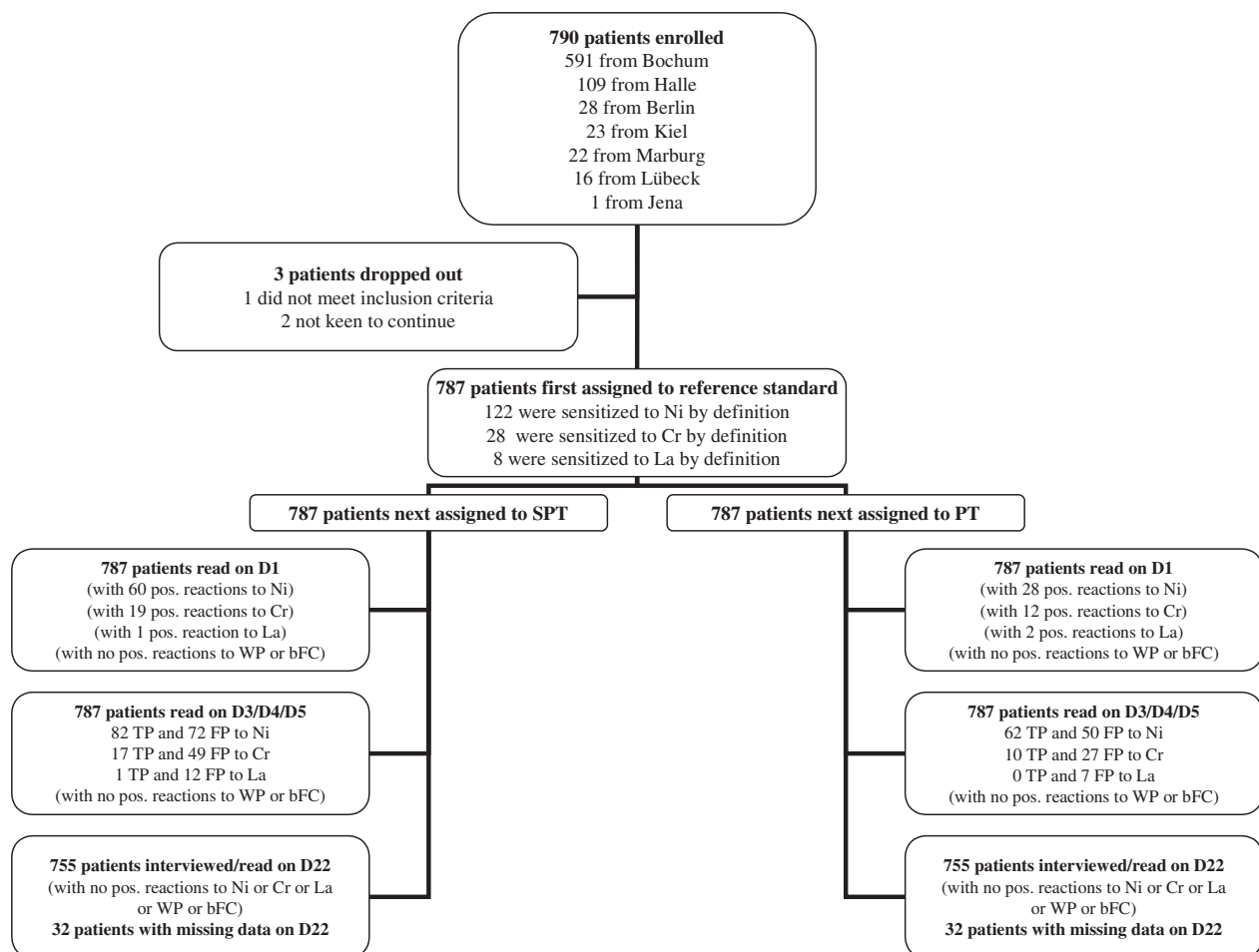


Fig. 1. Study flow chart. Ni, nickel sulfate 5% pet.; Cr, potassium dichromate 0.5% pet.; La, lanolin alcohol 30% pet.; SPT, strip patch test; PT, patch test; WP, white petrolatum pure; bFC, blank Finn Chamber; TP, true positive(s); FP, false positive(s).

conducted according to an individual randomization list which was sent to each of the patch test units at the beginning of the study and was controlled

at the end of the study (supervision by O.K.). Tests were uniformly stopped after 1 day. The first reading was routinely performed at D1 and the

second reading was done between D3 (in the vast majority of patients, 782 out of 787, 99.4%) and D5 (Fig. 1). Positive results are reported based on the second reading. In a telephone interview on D22, the patients were finally asked for persistent and/or new reactions on both test sites. Only in the case of an affirmative answer was the patient called in for an additional reading. The test results were graded using the classification system of the ICDRG (15) and dichotomized for statistical analysis. All '+', '++', and '+++' positive results were summarized as positive responses, whereas doubtful '?/(+)', irritant 'IR', and negative '-' reactions were counted as negative responses.

Patients were tested with nickel sulfate 5% pet., potassium dichromate 0.5% pet., and lanolin alcohol 30% pet. supplied by Hermal, Reinbek, Germany. The selection of these contact allergens was based on a decision of the DKG. Reasons for the decision were (i) substances should be from the baseline series (cp. <http://www.ivdk.gwdg.de/dkg/dkgblo.html#a001>) with a relative frequency of positive reactions $\geq 2\%$ and (ii) substances should cover the range of solubility characteristics from lipophobic/hydrophilic, like both metal salts, to lipophilic/hydrophobic, like lanolin alcohol. Thereby, it was assumed that a slightly damaged stratum corneum barrier by repeated tape stripping increased percutaneous penetration of primarily lipophobic substances, whereas the percutaneous penetration of the lipophilic substances should remain almost unaffected (16). Additionally, the vehicle white pet. (Hermal) and one blank Finn Chamber were tested as negative controls (Fig. 2). Other patch test series were applied as appropriate to the patient's presenting complaints.

Patch testing was performed according to standard methodology (17). Strip patch testing was performed according to our standardized method described in detail elsewhere (12, 14). Tape stripping was carried out with a 3M™ Blenderm™ surgical tape (3M™ Medica, 3M Deutschland GmbH, Neuss, Germany; metric 25 mm $\times$ 4.5 m; CE no. 0493–BSI 0086–class I; batch no. 201110, exp. date October 2011); body hair was removed with a 3M™ Clipper before tape stripping.

Data management

All participating dermatology clinics are active members of the Information Network of Departments of Dermatology (IVDK, cp. www.ivdk.gwdg.de/ivdk/eng/index.html), an epidemiological surveillance system for monitoring contact allergy, where data management of the study was located. Each clinic sent patient's data anonymously to the IVDK

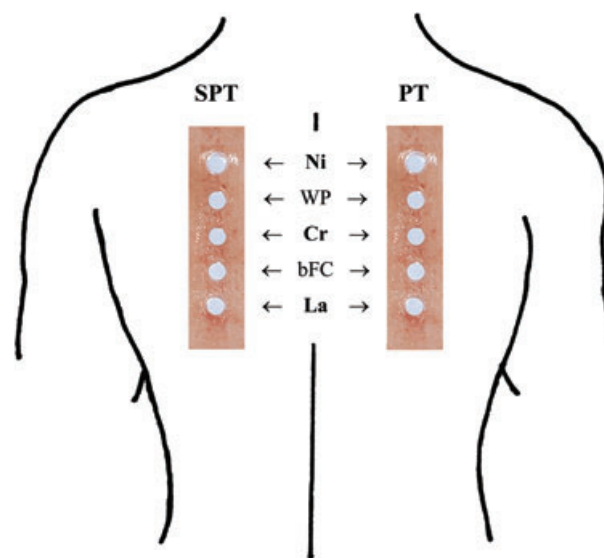


Fig. 2. Position of the test substances and index tests, the latter by randomization. SPT, strip patch test; PT, patch test; Ni, nickel sulfate 5% pet.; WP, white petrolatum pure; Cr, potassium dichromate 0.5% pet.; bFC, blank Finn Chamber; La, lanolin alcohol 30% pet.

for centralized evaluation (performed by J.G.). Recruiting numbers and potential problems of the study had been presented and discussed twice a year on routine basis in scheduled meetings of the DKG.

Blinding

The reading physician of each patch test unit was blinded with respect to (i) the reference standard (i.e. patient history) and (ii) which side of the patient's back was tape stripped. Unblinding was carried out after the second reading, for clinical interpretation of test results.

Statistical methods

Sample size calculation was based on the assumption that the sensitivity of the patch test is 73% (4). If 95 sensitized patients were included in the study, a clinically relevant increase by 10% points (18) to 83% for the strip patch test could be detected with 90% power (McNemar test, $\alpha = 0.05$). As the prevalence of nickel sensitization in our study population was expected to be 15%, $95/0.15 = 634$ patients should have been included in the study. To allow some moderate loss of follow-up (approximately 10%), we aimed to include 700 patients.

In patch testing, we consider the sensitivity (i.e. the probability of detecting a 'truly' sensitized patient) as clinically more important than the specificity. Thus, we defined the difference of sensitivities between the two index tests as our primary outcome. However, specificity and the corresponding predictive values were defined as secondary outcomes.

Patients who completed the two reading visits (D1, D3/D4/D5) were included in the analysis for the primary outcome. All outcomes were assessed within the family of marginal models, parameters were estimated by the generalized estimating equation method (19). Marginal models are a natural extension of the McNemar test to our situation where two index tests are compared. Additionally, these models account perfectly for the repeated testing in the same patient. In deviation to Leisenring et al. (19), we used the identity link function to achieve direct estimates of differences. All parameters are given with the respective 95% confidence interval (CI). CIs for the positivity ratio (PR) (20) are computed as exact intervals for binomial proportions, but those for the reaction index (RI) (21) were calculated by a non-parametric bootstrap experiment. The significance level was set at $\alpha = 0.05$.

We additionally conducted subgroup analyses (by interaction tests) to check if patients' sex, age, Fitzpatrick skin phototype, the various inter-individual numbers of tape strips for the strip patch test, or the participating dermatological clinics modify the observed sensitivity differences.

For descriptive purposes, we report categorical variables as absolute (relative) frequencies, and continuous variable as median (min–max, Q1, Q3). All statistical analyses were conducted with SAS® (SAS Institute, Cary, NC, USA), Version 9.2.

This report is written in close compliance with the Standards for Reporting of Diagnostic Accuracy statement (22).

Results

Figure 1 shows the flow chart of the study. In the 790 patients with suspected allergic contact dermatitis enrolled in the study, there were three dropouts: one female because of being 17 years old, other 81-year-old female because of removing all patches by

Table 2. Demographic and clinical patient characteristics (N = 787)

Characteristic	Value
Age (years), median (min–max, Q1, Q3)	51 (18–92, 39, 62)
Sex, n (%)	
Female	498 (63.3)
Male	289 (36.7)
Fitzpatrick skin phototype, n (%)	
I	10 (1.3)
II	528 (67.1)
III	233 (29.6)
IV	15 (1.9)
V	1 (0.1)
VI	0 (0)
Tape stripped side of the back, n (%)	
Left	370 (47.0)
Right	417 (53.0)
Tape strips down to the SL, median (min–max, Q1, Q3)	37 (9–96, 27, 51)
Tape strips for the SPT, median (min–max, Q1, Q3)	
Calculated*	16 (4–41, 11, 22)
Actually carried out	16 (4–43, 11, 22)

Q1, 25th percentile; Q3, 75th percentile; SL, stratum lucidum; SPT, strip patch test.

*strips down to the SL $\times$ 11/26.

herself on the day of application, and another 22-year-old female because of unexcused absence at the second reading. No serious adverse events from performing the index tests were reported. Hence, all the 787 patients were included in further data analysis.

Patients

Table 2 shows the demographic and clinical characteristics of all those patients who were further examined. Test sides of the upper back where the tape strips for the strip patch test are to be placed were balanced. In strip patch test, appearance of the stratum lucidum required a median number of 37 tape strips. After multiplying by the tape-specific

Table 3. Reactions to test substances at D3/D4/D5 reading (N-tested = 787)

		Reaction						RI	PR
		+	++	+++	?	IR	Neg.	(95% CI)	(%) (95% CI)
Ni	SPT	88	57	9	14	7	612	0.76 (0.66–0.85)	57 (49–65)
	PT	69	38	5	10	3	662	0.79 (0.68–0.89)	62 (52–71)
Cr	SPT	52	10	4	12	5	704	0.59 (0.42–0.76)	79 (67–88)
	PT	25	9	3	14	7	729	0.28 (0.03–0.52)	68 (50–82)
La	SPT	9	4	0	4	3	767	0.30 (–0.10–0.70)	69 (39–91)
	PT	6	1	0	3	3	774	0.08 (–0.38–0.54)	86 (42–100)

Ni, nickel sulfate 5% pet.; Cr, potassium dichromate 0.5% pet.; La, lanolin alcohol 30% pet.; SPT, strip patch test; PT, patch test; RI, reaction index; PR, positivity ratio.

Table 4. Paired index test results at D3/D4/D5 reading (N -tested = 787)

SPT		Ni	
		Neg., ?, IR	pos. (+, ++, +++ +)
PT	Ni neg., ?, IR	625	50
	pos. (+, ++, +++ +)	8	104
SPT		Cr	
		neg., ?, IR	pos. (+, ++, +++ +)
PT	Cr neg., ?, IR	717	33
	pos. (+, ++, +++ +)	4	33
SPT		La	
		neg., ?, IR	pos. (+, ++, +++ +)
PT	La neg., ?, IR	771	9
	pos. (+, ++, +++ +)	3	4

SPT, strip patch test; PT, patch test; Ni, nickel sulfate 5% pet.; Cr, potassium dichromate 0.5% pet.; La, lanolin alcohol 30% pet.

correction factor ($cf = 11/26$) (12), the 'median' patient needed 16 tape strips prior to patch testing. However, in 29 patients (3.7%) we noticed calculation errors which we defined as a divergence of at least two tape strips (data not shown).

Index tests

Table 3 shows the readings of reactions to nickel sulfate, potassium dichromate, and lanolin alcohol at D3/D4/D5 in the total of patients examined. In nickel sulfate, a marginal higher RI favouring the patch test faces a marginal lower PR favouring the strip patch test. The opposite can be observed for potassium dichromate. The too small number of observations for lanolin alcohol makes these parameters not reliable enough to assess the index tests. In both index tests, no positive reactions were observed to any of the negative controls (white pet.; blank Finn Chamber) at any reading time indicative of an irritant test effect (Fig. 1). There were also no persistent or new positive reactions documented to any of the test substances at D22, indicative of an iatrogenic sensitization; documentation of the D22 test reactions was missing in 32 patients (4.1%). Table 4 presents paired index test results obtained with the test substances.

Estimates of diagnostic accuracy

Table 5 summarizes main study outcomes with their 95% CI. Differences between sensitivities of strip patch test and patch test were 16.4% (95% CI, 8.7–24.1%) for nickel sulfate and 25.0% (95% CI,

8.9–41.0%) for potassium dichromate, both favouring strip patch test. In contrast, the resulting advantage of the patch test in terms of specificities was only minor: 3.3% (95% CI, 1.7–5.0%) for nickel sulfate and 2.9% (95% CI, 1.5–4.3%) for potassium dichromate. The positive predictive values (PPV) for nickel sulfate and potassium dichromate did not differ significantly between both index tests, whereas the negative predictive values (NPV) differed, in favour of strip patch test. The too small number of sensitizations to lanolin alcohol hampers a coherent reporting of these standard indices of validity.

Differences of sensitivities were significantly modified by patients' sex (nickel sulfate, $P = 0.03$; potassium dichromate, $P = 0.003$): superiority of the strip patch test was larger in men than in women for nickel sulfate, but vice versa in potassium dichromate (data not shown). However, no effect modification was seen for age (nickel sulfate, $P = 0.06$; potassium dichromate, $P = 0.83$), the Fitzpatrick skin phototype (nickel sulfate, $P = 0.54$; potassium dichromate, $P = 0.18$), the inter-individual different numbers of tape strips for the strip patch test (nickel sulfate, $P = 0.70$; potassium dichromate, $P = 0.052$), or the participating dermatological clinics (nickel sulfate, $P = 0.40$; potassium dichromate, $P = 0.09$).

Discussion

This study is unique in presenting data on diagnostic accuracy not only for our standardized strip patch test procedure (12), but also for the patch test from an independent, blinded comparison with a reference standard in a multicentre setting. We found that the strip patch test showed a better sensitivity than the patch test in detecting 'truly' sensitized patients to nickel sulfate and potassium dichromate, respectively, with only marginal compromises in terms of specificity. Hence, although it is not error-free, the strip patch test balances in a much better way the primary clinical aim of reducing the number of false negatives and the risk of increasing the number of false positives which may result in superior management of patients and improved patient-important outcomes (e.g. morbidity, symptoms, and quality of life) (23).

While sensitivity and specificity provide stable estimates of an index test as they are not dependent on the prevalence of sensitization in the collective tested (2), from a clinical viewpoint it is diagnostically more conclusive to assess the predictive values (3). Our data revealed fair PPVs for nickel sulfate and potassium dichromate with no significant differences between both index tests (Table 5). In all patients tested positive to nickel

Table 5. Main outcomes ($N = 787$)

				Difference	
				SPT-PT	<i>P</i> value
Primary outcome					
Sensitivity, % (95% CI)	Ni	67.2 (58.4–74.9)	50.8 (42.0–59.6)	16.4 (8.7–24.1)	<0.0001
	Cr	60.7 (42.0–76.7)	35.7 (20.4–54.6)	25.0 (8.9–41.0)	0.008
	La	12.5 (1.7–53.7)	0.0 (0.0–31.2)	No convergence	n.c.
Secondary outcomes					
Specificity, % (95% CI)	Ni	89.2 (86.6–91.3)	92.5 (90.2–94.3)	−3.3 (−5.0 to −1.7)	<0.0001
	Cr	93.5 (91.6–95.1)	96.4 (94.9–97.5)	−2.9 (−4.3 to −1.5)	<0.0001
	La	98.5 (97.3–99.1)	99.1 (98.1–99.6)	−0.6 (−1.4–0.02)	0.132
PPV, % (95% CI)	Ni	53.2 (45.4–61.0)	55.4 (46.1–64.3)	−2.1 (−7.9–3.6)	0.471
	Cr	25.8 (16.7–37.6)	27.0 (15.2–43.3)	−1.3 (−11.0–8.5)	0.800
	La	7.7 (1.1–39.1)	0.0 (0.0–34.8)	No convergence	n.c.
NPV, % (95% CI)	Ni	93.7 (91.5–95.3)	91.1 (88.7–93.0)	2.6 (1.2–4.0)	0.0003
	Cr	98.5 (97.3–99.2)	97.6 (96.2–98.5)	0.9 (0.2–1.6)	0.012
	La	99.1 (98.1–99.6)	99.0 (98.0–99.5)	0.1 (−0.1–0.4)	0.340

SPT, strip patch test; PT, patch test; PPV, positive predictive value; NPV, negative predictive value; Ni, nickel sulfate 5% pet.; Cr, potassium dichromate 0.5% pet.; La, lanolin alcohol 30% pet.; n.c., not computed.

sulfate and potassium dichromate, respectively, only approx. 54% and 26% were ‘truly’ sensitized, as determined by one of the index tests. Or, put in absolute numbers, we will have to do two positive patch tests to nickel sulfate (and even four positive tests to potassium dichromate) to see one ‘truly’ sensitized patient. Two explanations may contribute to this unsatisfying situation. First, test substances such as potassium dichromate have a certain irritant potential that is often clinically detectable even in non-sensitized patients (24). Second, there may be truly sensitized patients who do not fulfil our reference standard, for example because of lack of remembrance or hidden allergen sources not covered by our standardized questions. As our study mainly concerned comparing the sensitivity of strip patch tests and patch tests, this second potential shortcoming does not affect the study outcome, because the same reference standard was used for both tests. As the PPV declines with declining prevalence of sensitization in the collective tested, the widespread use of strip patch test as well as patch test can be expected to result in even more false-positive results and thus lower PPV than has been found in our study (5). This outcome may have serious consequences as patients might be diagnosed allergic to environmental agents to which they are not. This emphasizes again that the strip patch test and the patch test can both be only valuable diagnostic tools when additionally based on clues from the patient’s history and in a given case on threshold testing and use tests, respectively (3, 25). On the other hand, NPVs of both index tests turned out to be very good (Table 5).

By tape stripping the skin, prior to patch testing according to our standardized method (12), we

intentionally induce a defined and reproducible damage to the stratum corneum, the principal barrier against the percutaneous penetration of substances (16). This most probably results in the observed increase of sensitivity. The non-existent statistical effect of the inter-individual different numbers of tape strips for the strip patch test on our estimated sensitivities is thereby indicative of an inter-individually equal percutaneous penetration of test substances, thus supporting once more a purposeful standardization (14). However, there may be an additional impact of strip patch test on the (bio)chemical and the immunological barriers of the epidermis which needs to be further investigated (26). In this context, one may also speculate on strip patch test unmasking patch test results which are clinically negative but show immunopathological reactions in the skin (27).

We acknowledge several potential limitations of this study. First, this study would need an independent confirmation with the substances tested here, but more important with other standard test substances. Second, the problem with allergic contact dermatitis is that a perfect ‘gold standard’ of irrefutable delayed-type sensitization does not exist at present (7). Therefore, the best feasible reference standard, that is the patient history (4, 28) in a weight of evidence decision and specified *a priori* (Table 1), was used for comparison. Moreover, documentation of the reference standard, application and reading of the index tests were performed each by different investigators. Thus, reading of the reactions was performed by an investigator blinded for patient history or placement of the index tests. A use test such as the repeated open application test may be an even more specific reference standard

for allergens with elusive exposure scenarios like lanolin alcohol (29). However, study feasibility should be always kept in mind (30). Third, the collective analysed is certainly a non-representative sample in as much as 15.5%, 3.6%, and 1.0%, respectively, of the patients tested were sensitized to nickel sulfate, potassium dichromate, and lanolin alcohol by definition (Fig. 1). These high prevalences are a result of the referral of these patients for assessment by other dermatologists who suspected that the patients had allergic contact dermatitis. If the index test methods were applied to a group of patients in which there were smaller proportions of allergic contact dermatitis and delayed-type sensitization, respectively, more false-positive and less false-negative test results would be expected (4). Finally, we can make no point on the impact of our standardized strip patch test procedure (12) on test substances with different physicochemical properties, that is the principal impact of tape stripping on increasing percutaneous penetration of lipophobic substances versus lipophilic ones (16). To evaluate such test substances like lanolin alcohol, one would need a much bigger sample size.

Studies analysing the standard indices of validity for the patch test are scarce. Nethercott and Holness (4) reported sensitivities, specificities, PPVs and NPVs for the NACDG and ICDRG screening series of 77, 71, 69, and 78% and 68, 77, 66, and 79%, respectively. However, indices of validity for individual test substances were not published and it is not clear whether the index tests were performed blinded for the reference standard. Recently, Fernandes et al. (31) evaluated their alternative strip patch test procedure versus patch testing by applying a standard series consisting of 30 test substances in duplicate on the back of 65 patients. Although sensitivity and specificity were reported to be higher for the strip patch test (83% and 77%) versus patch testing (77% and 71%), no statistically significant difference was seen between both methods in sensitization to nickel sulfate. However, these data are hardly comparable to ours. First, their study potentially suffers from a 'diagnostic review bias' (32) because relevance analysis was performed with knowledge of the results of the index tests. Additionally, 'relevance' has been applied only to positive test results and not to negative ones to determine how many of the latter are false-negative. It is presently acknowledged that data on false-negative patch test results for individual test substances are lacking but are assumed to be high with some allergic contact dermatitis conditions failing to improve (2, 3, 5, 6). Kligman and Gollhausen (33) emphasized that the problem of the 'silent back syndrome' is underestimated. Finally, data quite comparable to ours for the patch test obtained with fragrance mix I

8% pet. (sensitivity = 25.2%; specificity = 96.5%; PPV = 45.1%; NPV = 91.8%) and fragrance mix II 14% pet. (sensitivity = 13.5%; specificity = 98.8%; PPV = 55.6%; NPV = 90.8%), respectively, were published in a study with 1701 patients by Frosch et al. (28). However, a direct comparison is hampered to some extent because of an adversely defined reference standard, although this was also based on patient history and defined prior to patch testing.

In summary, it can be stated that since 1895 (1) various modifications of the technique of patch testing have been proposed in an attempt to obtain a more sensitive and reliable test (34). Of these modifications only the strip patch test (18) has survived in clinical practice, but is used rarely mainly because of its time-consuming limitations (10) and a lack of standardization. In this context the presented data, support that our easy-to-use, standardized strip patch test (12) would have several clinical implications for current patch test practice: (i) testing patients with a suggestive history and negative or questionable patch test results to suspected substances, (ii) testing patients with substances of expected poor percutaneous penetration or low concentration, (iii) testing patients if percutaneous penetration at the test site is lower than that of clinical exposure (e.g. eyelids, axillae) (10), and (iv) testing patients by reproducing the clinical exposure, if damaged stratum corneum is present. Although we acknowledge that each of these indications would require a study of its own, we are confident that our strip patch test would perform well in these settings, which all have significant implications for patient management, for example in medicolegal considerations (35).

Conclusions

In this multicenter diagnostic accuracy study, our standardized strip patch test procedure (12) proved to be safe and particularly valid for obtaining additional information on a patient's delayed-type sensitization status and supporting diagnosis of allergic contact dermatitis. Thus, the strip patch test may play a decisive role in routine patch testing. However, in further studies not only its reproducibility (7) but also its usefulness should be scrutinized; the latter in a randomized study in which patients receive either the strip patch test or the patch test, and patient-important outcomes are measured (36).

Acknowledgements

We thank the patients for their co-operation in this study, the nurses for their outstanding engagement, and the

dermatologists-in-training for their excellent support. The multicentre study was supported by the DKG.

References

- Jadassohn J. Zur Kenntnis der medicamentösen Dermatosen. *Verhandlungen der Deutschen Dermatologischen Gesellschaft, 5. Kongreß, Graz, 1895. Braumüller, Wien* 1896: 103–129.
- Nethercott J R. Sensitivity and specificity of patch tests. *Am J Contact Dermat* 1994; 5: 136–142.
- Diepgen T L, Coenraads P J. Sensitivity, specificity and positive predictive value of patch testing: the more you test, the more you get? *Contact Dermatitis* 2000; 42: 315–317.
- Nethercott J R, Holness D L. Validity of patch test screening trays in the evaluation of patients with allergic contact dermatitis. *J Am Acad Dermatol* 1989; 21: 568.
- Nethercott J R. Practical problems in the use of patch testing in the evaluation of patients with contact dermatitis. *Curr Probl Dermatol* 1990; 2: 97–123.
- Shuster S. Patch-test sensitivity and reproducibility in individuals and populations. *Am J Contact Dermat* 1992; 3: 74–78.
- Belsito D V, Storrs F J, Taylor J S, Marks J G Jr., Adams R M, Rietschel R L, Jordan W P, Emmett E A. Reproducibility of patch tests: a United States multicenter study. *Am J Contact Dermat* 1992; 3: 193–200.
- Traidl-Hoffmann C, Ring J. Is there an in vitro test for type IV allergy discriminating between sensitization and allergic disease? *Clin Exp Allergy* 2008; 38: 1412–1415.
- Spier H W, Natzel R. Chromatallergie und Zementekzem. *Gewerbedermatologischer und analytischer Beitrag. Hautarzt* 1953; 4: 63–65.
- Lachapelle J M, Maibach H I. Patch Testing and Prick Testing – A Practical Guide (Official Publication of the ICDRG), 2nd Auflage, Springer-Verlag, Berlin Heidelberg, 2009.
- Dickel H, Geier J, Kuss O, Altmeyer P. Strip patch test vs. conventional patch test to detect type IV sensitization in patients with allergic contact dermatitis. *J Eur Acad Dermatol Venereol* 2008; 22: 1516–1517.
- Dickel H, Bruckner T M, Erdmann S M et al. The “strip” patch test: results of a multicentre study towards a standardization. *Arch Dermatol Res* 2004; 296: 212–219.
- Bernstein L, Li J T, Bernstein D I et al. Allergy diagnostic testing: An updated practice parameter. *Ann Allergy Asthma Immunol* 2008; 100: S1–S148.
- Dickel H, Kamphow J, Geier J, Altmeyer P, Kuss O. Strip patch test vs. conventional patch test: investigation of dose-dependent test sensitivities in nickel- and chromium-sensitive subjects. *J Eur Acad Dermatol Venereol* 2009; 23: 1018–1025.
- Fregert S. Manual of contact dermatitis, 2nd Auflage, Munksgaard, Copenhagen, 1981.
- Kezic S, Nielsen J B. Absorption of chemicals through compromised skin. *Int Arch Occup Environ Health* 2009; 82: 677–688.
- Schnuch A, Aberer W, Agathos M, Brasch J, Frosch P J, Fuchs T, Richter G. für die Deutsche Kontaktallergie-Gruppe. Leitlinien der Deutschen Dermatologischen Gesellschaft (DDG) zur Durchführung des Epikutantests mit Kontaktallergenen. *Hautarzt* 2001; 52: 864–866.
- Spier H W, Sixt I. Untersuchungen über die Abhängigkeit des Ausfalles der Ekzem-Läppchenproben von der Hornschichtdicke. Quantitativer Abriß-Epikutantest. *Hautarzt* 1955; 6: 152–159.
- Leisenring W, Alonzo T, Pepe M S. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics* 2000; 56: 345–351.
- Geier J, Uter W, Lessmann H, Schnuch A. The positivity ratio – another parameter to assess the diagnostic quality of a patch test preparation. *Contact Dermatitis* 2003; 48: 280–282.
- Brasch J, Henseler T. The reaction index: a parameter to assess the quality of patch test preparations. *Contact Dermatitis* 1992; 27: 203–204.
- Bossuyt P M, Reitsma J B, Bruns D E et al. for the STARD Group. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003; 138: 40–44.
- Basketter D A, Kimber I. Information derived from sensitization test methods: test sensitivity, false positives and false negatives. *Contact Dermatitis* 2007; 56: 1–4.
- Pasche-Koo F, Hauser C. How to better understand the angry back syndrome. *Dermatology* 1992; 184: 237–240.
- Aberer W, Andersen K E, White I R. Should patch testing be restricted to dermatologists only? *Contact Dermatitis* 1993; 28: 1–2.
- Proksch E, Brandner J M, Jensen J M. The skin: an indispensable barrier. *Exp Dermatol* 2008; 17: 1063–1072.
- Buckley C, Poulter L W, Rustin M H. Immunohistological analysis of “negative” patch test sites in atopic dermatitis. *Clin Exp Allergy* 1996; 26: 1057–1063.
- Frosch P J, Pirker C, Rastogi S C et al. Patch testing with a new fragrance mix detects additional patients sensitive to perfumes and missed by the current fragrance mix. *Contact Dermatitis* 2005; 52: 207–215.
- Schnuch A, Uter W, Dickel H et al. Quantitative patch and repeated open application testing in hydroxyisohexyl 3-cyclohexene carboxaldehyde sensitive-patients. *Contact Dermatitis* 2009; 61: 152–162.
- Frosch P J, Rastogi S C, Pirker C et al. Patch testing with a new fragrance mix – reactivity to the individual constituents and chemical detection in relevant cosmetic products. *Contact Dermatitis* 2005; 52: 216–225.
- Fernandes M F M, de Mello J F, Pires M C, Vizeu M C M. Comparative study of patch test sites using traditional method vs. prior skin abrading. *J Eur Acad Dermatol Venereol* 2007; 21: 1351–1359.
- Bossuyt P M, Reitsma J B, Bruns D E et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003; 49: 7–18.
- Kligman A, Gollhausen R. The ‘angry back’: a new concept or old confusion? *Br J Dermatol* 1986; 115: 93–100.
- Anderson W A, Shatin H, Canizares O. Influence of varying physical factors on patch test responses. *J Invest Dermatol* 1958; 30: 77–82.
- Dickel H, Scola N, Altmeyer P. The strip patch test – indication in occupational dermatology demonstrated with a case history. *J Dtsch Dermatol Ges* 2009; 11: 965–967.
- Brozek J L, Akl E A, Jaeschke R et al. for the GRADE Working Group. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. *Allergy* 2009; 64: 1109–1116.

Address:

Heinrich Dickel, MD
 Department of Dermatology and Allergology
 St. Josef-Hospital
 Ruhr University Bochum
 Gudrunstrasse 56
 44791 Bochum
 Germany
 Tel: +49 234 509 3448
 Fax: +49 234 509 3451
 e-mail: h.dickel@klinikum-bochum.de