Thoughts on how to improve the quality of multicentre patch test studies

Magnus Bruze

Department of Occupational and Environmental Dermatology, Skåne University Hospital, Lund University, S-205 02 Malmö, Sweden

doi:10.1111/cod.12507

Summary

Multicentre patch test studies (MPTSs) can contribute useful information for diagnostic and preventive measures. The aim of the present paper is to propose how to perform high-quality MPTSs. To this end, factors of significance for the patch test result are discussed with regard to the standardization and calibration of high-quality MPTSs. The 16 factors discussed are scored 0, 1, 2, or 3, depending on the relative importance of a particular factor for the patch test result. The total score of an MPTS allows it to be ranked as having doubtful, acceptable, high, or excellent quality. A total score of 30 is possible. Depending on the total score the MPTSs are grouped into those with a doubtful, acceptable, high, and excellent quality. In conclusion, high-quality MPTSs can be performed and are facilitated if a guideline and check list are followed when the study is being planned. The scoring enables the calculation of a total score, which can be used for quality ranking.

Key words: allergic contact dermatitis; calibration; check list; contact allergy; guideline; monitoring; quality ranking; standardization.

Multicentre patch test studies (MPTSs) are valuable tools with which to establish contact allergy rates in defined populations, to follow trends in contact allergy rates in various geographical areas, and to facilitate preventive measures. Increasing rates signal that actions might have to be taken, for example voluntary actions by the manufacturers/suppliers of products containing the sensitizer, or legislative measures limiting or banning the use of the sensitizer in certain products. When such measures have been taken. MPTSs can help to show the effectiveness of the measures. Furthermore, the results of MPTSs usually constitute the basis for the use of a certain concentration (dose/cm²) of a sensitizer in establishing the threshold of skin irritation at patch testing (1), and decisions on the patch test concentration (dose/cm²) when a sensitizer is included in a baseline test series, or actually any series (2, 3). There are thus different types of MPTS in which at least two patch test clinics participate. One type involves collecting retrospective data (4), another type is represented by networks producing patch test data constantly (5), and a third type is the prospective MPTSon, for example, the concentration (dose/cm²) of a certain sensitizer to be used (6).

MPTSs have been questioned because of inadequate standardization (7, 8). There is thus a need to improve the quality of MPTSs. The more standardized and calibrated the various parts of the patch test procedure, including the test reading, the more reliable and useful the MPTSs are. Recently, the European Contact Dermatitis Society published guidelines on patch testing (9). Table 1 lists factors of significance for the patch test results, and a scoring system based on the performance of these factors. Hopefully, this table can serve as a guideline and a check list on how to perform an optimal MPTS. The list is not intended to be used to evaluate historical data. However, besides being a tool to improve the quality of MPTSs during planning, the
Table 1. Factors of significance for the patch test result and their relative importance for the quality of multicentre patch test studies

<table>
<thead>
<tr>
<th>Factor of significance</th>
<th>Performance</th>
<th>Score</th>
<th>Highest possible score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturer of compound/chemical – same manufacturer and batch</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Manufacturer of patch test preparation – same manufacturer and batch</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Transportation – risk factors minimized</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Storage – refrigerated when not used and replacement when needed</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Patch test system – same</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Application of preparation to unit/chamber – just before application when needed</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Dose – recommended dose used</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Application area – same</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Occlusive tape – same</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Occlusion time – 48 h</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Adhesiveness of patch tests after 48 h – controlled</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reading times – only one option possible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3 + D7 or D4 + D7</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>D3/D4 + D6–D8</td>
<td>Yes</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>D3/D4</td>
<td>Yes</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>Other reading times</td>
<td>Yes</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Classification system – same</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Calibration of reading – performed</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Study period – same/equivalent time period</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Monitoring – performed</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Total possible score</td>
<td></td>
<td></td>
<td>30</td>
</tr>
</tbody>
</table>

Scores: 1 implies low importance, 2 implies average importance, and 3 implies high importance. D, day.

list and scoring used can also, on the rare occasions when measures are needed but there are two or more MPTSs with conflicting results, serve as a tool with which to evaluate the quality of the MPTSs, similarly to the systematic approach suggested by Klimisch for evaluating the quality of experimental toxicological and ecotoxicological data (10). In the following, the discussion will focus on the performance of high-quality MPTSs.

Factors of Significance for the Patch Test Result

Manufacturer of the compound/chemical
Contact sensitizers to be patch tested may represent chemically defined compounds (substances), such as nickel and formaldehyde, or chemicals such as colophonium and p-tert-butylphenol formaldehyde resin, which consist of an unknown number of compounds that, to a large extent, might be unidentified. Both the synthesis of these compounds/chemicals and the extraction processes may vary with the manufacturer, which is why the test compound/chemical used should be from one manufacturer, to minimize the possibility of variation in the numbers, types and concentrations of ‘contaminants (undesired compounds)’.

Although the same synthetic pathways and extraction procedures might have been used by one manufacturer, there might still be batch-dependent variation, with possible significance for the patch test result (11). The same batch of a test sensitizer should therefore be used for an optimal MPTS.
Manufacturer of the test preparation

Studies have shown that the concentrations of test compounds may vary between manufacturers of test preparations (12, 13). The same manufacturer should therefore be used for an optimal MPTS. Again, it is important that all test preparations from the manufacturer originate from the same batch, to minimize the risk of variation.

The vehicle should be chosen to minimize possible reactions resulting in a decreased concentration because of oxidation, polymerization, or degradation. Certain sensitizers, such as formaldehyde releasers, have been tested in both water and petrolatum (14). Pet. is usually a feasible vehicle for most sensitizers, and it seems to prevent oxidation, polymerization, and degradation. The quality and presence of intentionally added substances such as antioxidants should be reported. Pet. should not be used for methylisothiazolinone, as it is not evenly distributed in this vehicle (15). Certain sensitizers require special considerations. Diisocyanates should not be tested in alcoholic vehicles, as they react with such solvents, generating carbamates. Similarly, amines should not be tested in 100% acetone (ketones), as imines will be formed. The combination of acrylates in acetone and an aluminium-based test chamber is also inappropriate, as the acrylate will polymerize because of the formation of aluminium ions acting as a catalyst for the polymerization (16).

Transportation

The test preparations should be shipped to the centres participating in the MPTS in a way that prevents/substantially diminishes the risk of changes to the test compound/chemical. Oxygen, ultraviolet (UV) radiation and high temperature are the major noxious factors that may affect the test preparation. The containers, bottles and syringes should prevent migration of oxygen through the material and themselves, or be packed in such a way that UV irradiation of the test preparation is prevented. The containers should also prevent evaporation of the test compound/chemical and/or vehicle (17, 18).

Storage

The same noxious factors that may affect the test preparations during transportation may affect them during storage. When not being used, the test syringes and bottles should be refrigerated or, occasionally, frozen if there are long periods between patch test occasions. The containers as such and the lids, as well as the caps of syringes, should prevent evaporation of the vehicle and/or the sensitiser. The time for which the bottles and syringes do not have the lids and caps on, respectively, should be minimized (19).

When the test preparations are known to be stable for only limited periods of time that do not cover the whole investigative period of the MPTS, the test preparations have to be replaced regularly, on the basis of the known duration of stability and/or chemical analyses (20, 21). Again, the new test preparations should be from the same manufacturer and batch of test substance/chemical, and the patch test preparation should also be from the same manufacturer.

Certain sensitizers, including elemental mercury (22) and, perhaps, particularly metal salts, will not be dissolved in pet. When delivered from the patch test manufacturers, these substances are (hopefully) evenly distributed in the pet. in the form of ‘miniature droplets’. During storage, these droplets may aggregate. Depending on the density of the test substance, there might be a risk of ‘sinking’ of the droplets through the pet., resulting in accumulation and thus a too high concentration in the lower part of the syringe. This will only be a problem when syringes are kept vertically, provided that, when they are kept horizontally, the opening of the syringe is symmetrical and placed centrally in one of the ends of the syringe. Elemental mercury in pet. is an example where a ‘sinking phenomenon’ may be expected to occur when the patch test syringe containing it is kept vertically for an extended period of time.

Patch test system

In principle, there are two types of system, one having test units preloaded with the sensitizers, and the other with empty units/chambers that have to be loaded before application on the skin of those to be tested. The preloaded system guarantees that the same dose of the sensitizer is applied every time.

There are many test systems in which the test preparation has to be applied to the unit/chamber before application on the skin. They differ concerning shape (round or square), material (aluminium or plastics), and the use of filter paper. Investigations have, on the whole, shown good agreement between various systems (23, 24), but there are differences (25), which is why the same system, whether it is a preloaded or initially unloaded system, should preferably be used for the same sensitizers when an MPTS is performed. Thus, some sensitizers within an MPTS can be tested with a preloaded system, and other sensitizers can be tested with an unloaded system.
Application of the test preparation to the test unit/chamber

For a long time, it has been known that test substances/chemicals in volatile solvents such as acetone may be tested at a too high concentration, owing to evaporation of the solvent before application to the test unit/chamber. Surprisingly, it was around the new millennium when I first realized that many volatile test substances/chemicals may also evaporate not only from solvents but also from pet. preparations. Subsequently, investigations have shown that fragrance substances, acrylates (26, 27) and formaldehyde evaporate from the test preparation and test units/chambers. Therefore, the preparations of these and other volatile substances/chemicals must be applied to the test units/chambers immediately before application to the skin. For other sensitizers for which it is known or has been shown that application hours or days before application to the skin will not change the test preparation applied to the unit/chamber, such application is possible.

Dose

Both induction of sensitization and elicitation are dose-responsive phenomena. The number of molecules per area skin is the decisive factor. Defined doses/cm² have been used for liquid test preparations for the aluminium test technique and the Finn Chambers® technique in both animals and humans since the 1980s (28). For pet., there was no dose/cm² recommendation until 2007, when a 20-mg pet. preparation was recommended for a small Finn Chamber® with a diameter of 0.8 cm, giving a dose of pet. of 40 mg/cm² (29). When other test units/chambers are used, the dose/cm² should be equivalent. Liquid preparations should be applied with a micropipette to ensure that the desired volume is applied (30). With a higher pet. amount or volume of liquid than recommended, there is the risk of a too high dose/cm², possibly resulting in irritant reactions, false-positive reactions, and patch test sensitization as the most serious adverse effect of patch testing. A too low dose/cm² may result in a false-negative reaction.

Application area

Any skin area can be used in an MPTS as long as all centres test on the same area. For practical reasons, the upper back is the preferred area, as it enables many test preparations to be tested, and the skin on the back is among the most sensitive areas concerning elicitation (31).

Occlusion tape

Our knowledge is limited concerning any possible differences between various tapes regarding the occlusive effect and thus the influence on the patch test result (32). Hence, any tape that does not induce sensitization and causes no other adverse reactions, and at the same time gives permanent occlusion for the time intended, may be used.

Occlusion time

The occlusion time is standardized at 2 days (48 h) (33, 34). With the same patch test system, including the tape, the patch test result depends on the dose/cm² and occlusion time (35). The occlusion time can be lowered for sensitizers for which ‘all molecules’ in the applied dose migrate into the skin within the occlusion time used. For other sensitizers, that is, most of them, a shorter time can be compensated for by a higher dose/cm² (36). Thus, 48 h should be used in all MPTSs unless it has been unambiguously proven that another occlusion time can be used, probably by increasing the dose/cm².

Occasionally, the purpose of an MPTS may be to investigate the significance of a shorter or a longer occlusion time. In this situation, participating centres must use the same occlusion time.

Adhesiveness of patch tests

When patch testing, many clinics will see the tested individual on D2 to ensure that the patch tests have been appropriately attached to the skin before removal.

Reading times

Partly for practical reasons, various reading times are used. If only one reading time is used, the best one of those investigated is day (D)4 (96 h after application to the skin) (37). For many sensitizers, an additional reading after 1 week is necessary in order to not miss a substantial number of contact allergies (38–40). The conclusions of an MPTS should be based on the results of readings performed on the same day – preferably D4 if only one reading is performed, and D4 and D7 if two readings are performed. For practical reasons, two readings on D3/D4 and D6–D8 are acceptable. The results from a clinic reading once, for example on D2, should never be compiled and compared with the results from clinics reading more than once, for example on D2 + D4 + D7, unless only the reading on the same day is considered for all clinics.

Occasionally, the purpose of an MPTS may be to investigate the significance of test reading on another day or
on days other than the traditional ones. In this situation, participating centres must use the same reading time(s).

**Classification system**

There are many classification systems (34, 41, 42). The morphological features of an allergic patch test reaction consist of erythema, infiltration (oedema), papules, and vesicles. In almost all systems, except in the most recent system (42), the minimum criterion for a reaction to be classified as an allergic reaction is the presence of erythema and infiltration covering the whole test area. Subclassification with regard to intensity of test reaction is then based on the presence of papules and vesicles, sometimes coalescing and forming a bulla. When the morphological feature or features constituting an allergic reaction are present but do not fulfill the minimum criteria, the reaction should be classified as doubtful. A doubtful reaction can represent either a weak allergic reaction or an irritant reaction. An irritant reaction has a morphology that cannot be classified as allergic or doubtful. For high-quality MPTSs, the same classification system should be used.

**Calibration of reading**

Although patch test readers claim to read according to the same classification system, they do not always do this (43, 44). Differences exist regarding the discrimination of weak allergic reactions from doubtful reactions, and of doubtful reactions from irritant reactions (43). A large variation, particularly in the number of irritant reactions, indicates, as a possible explanation, a lack of calibration of the test readers in the various centres against the classification system used (21). Investigations have shown that the same classification system can be used with significantly less variation after teaching of the necessary requirements for the various subclassifications (43).

**Study period**

An MPTS should be performed during the same time at the participating centres and preferably during 1 year, or multiple years, to eliminate the possible significance of season for the patch test result (45, 46), particularly when centres on both sides of the equator participate.

**Monitoring**

In investigations exploring the effects of pharmaceutical drugs in humans, monitoring is mandatory. For an MPTS, monitoring is desirable but difficult to perform, owing to a lack of financial resources. If possible, the monitoring can look at all factors of significance for an optimal MPTS. Currently, to the best of our knowledge, there is only one MPTS so far in which monitoring has taken place (47–49).

**Quality Ranking**

In Table 1, the factors of significance for the patch test result are listed, but Table 1 also includes an attempt to rank their relative importance for the standardization on the basis of my experience of MPTSs since the 1980s, and the results of investigations on these factors performed at my department and/or reported in the literature. It is emphasized that the scoring is partly subjective and arbitrary, like other classifications (34, 38). When a factor has not been considered, the score is 0. A score of 1 means low importance, 2 means average importance, and 3 means high importance. For most of the factors, it is easy to decide whether the requirement to obtain a score is fulfilled. For other factors, such as test reading days and area of application on the skin, there will most likely be deviations from the protocol used. The patients/subjects tested may become ill and not be able to attend on the scheduled reading days. When a strong contact allergic reaction is suspected, the application of a sensitizer on the upper arm to enable removal of the test unit in advance and treatment with a topical corticosteroid is another reason for deviation. As long as there is < 3% deviation for any factor, the requirements are considered to have been fulfilled. The total possible score is 30. A classification based on the score is suggested in Table 2, with the MPTSs grouped in four categories.

**Conclusion**

Table 1 can be considered as a check list when an MPTS is being planned, to achieve high quality of the MPTS. The list with associated scores for the individual factors of significance for the patch test result can also serve to rank the MPTS with regard to quality. The more clinics that participate in an MPTS, the more likely it is that there will be a deviation in one or more of the factors of significance for the patch test result, which will probably result in a lower quality ranking. Obviously, a prospective MPTS can be planned differently from a retrospective MPTS, which is why a higher total score and, consequently, ranking can be expected. The quality ranking is not intended to be used on MPTSs already performed, unless decisions have to be made on the basis of historical MPTSs with conflicting results. In such a situation, the ranking can help to determine which MPTSs to give precedence to. Again,
it is emphasized that the scoring is partly subjective and arbitrary. Therefore, when new knowledge on the significance of the factors listed in Table 1 or currently non-listed factors appears, preferably based on the results from randomized, controlled and blinded studies, a different scoring system should be considered. A factor of possible significance for the patch test result, as it would substantially diminish the risk of bias (44) when the patch test reactions are read, is randomized application of the test chemicals on the back and subsequent ‘blind’ reading of the reactions. Such a design has been used in patch testing (50), but has been intentionally omitted as a factor in Table 1, as the risk of mistakes is considerable, and almost no patch test clinics currently have the resources to perform such MPTSs on a large scale.

Besides the differences in contact allergy rates among participating clinics in MPTSs depending on the factors listed in Table 1, differences may also depend on other circumstances, for example differences in occupational exposures, variation in referral patterns, and indications for patch testing. Use of the MOAHFLA index allows for stratification of the material and multivariate statistical analysis (51).

### References

3. Hauksson I, Pontén A, Gruvberger B et al. Routine diagnostic patch-testing with formaldehyde 2.0% (0.6 mg/cm²) may be an advantage compared to 1.0%. *Acta Derm Venereol* 2010; 90: 480–484.
20. Pontén A, Aalto-Korte K, Agner T et al. Patch testing with 2.0% (0.60 mg/cm²) formaldehyde instead of 1.0% (0.30 mg/cm²) detects significantly more contact allergy. *Contact Dermatitis* 2013; 68: 50–53.
38 Isaksson M, Brandão F M, Bruze M, Goossens A. Recommendation to include budesonide and tixocortol pivalate in the European standard series. ESCD and EECDRG. European Society of Contact Dermatitis. Contact Dermatitis 2000: 43: 41–42.
44 Uter W, Frosch P J, Becker D et al. Are we biased when reading a doubtful patch test reaction to a ‘clear-cut’ allergen such as the thiuram mix? Contact Dermatitis 2009: 60: 234–235.