Chapter 4
Guidelines in Skin Testing

Joachim W. Fluhr, G. E. Piérard, and C. Piérard-Franchimont

4.1 Introduction

The current chapter is a revision made by Joachim W. Fluhr of the original titled "Guidelines in Dermocosmetic Testing" written by Gerald Piérard and Claudine Franchimont in the first edition of the book.

Objective assessments and measurements of skin physiology parameters in dermatology and cosmetology are subject to interobserver variations and biases. Noninvasive biophysical measurements are improving both the descriptive and quantitative assessment conditions. A few decades ago, progressive researchers pioneered methods with continuous improvements; they may now look crude, time-consuming, and sometimes lacking fine-tuned reproducibility. With more recent progress, the noninvasive technology has made great advances both in the design and accuracy as well in reproducibility of the data and reliability of the measuring devices. This evolution was paralleled by an increase in knowledge in skin biology and physiology. Noninvasive devices are now used in translational research involving dermatologists, bioengineers, cosmetologists, clinical scientists, physiologists, and biologists.

J. W. Fluhr (✉)
Department of Dermatology and Allergology, Charité Universitätsmedizin, Berlin, Germany
e-mail: joachim.fluhr@charite.de

G. E. Piérard · C. Piérard-Franchimont
Laboratory of Skin Bioengineering and Imaging, Department of Dermatopathology, University Hospital Sart Tilman, Liège, Belgium

J. W. Fluhr (ed.), *Practical Aspects of Cosmetic Testing*,
https://doi.org/10.1007/978-3-030-44967-4_4

4.2 Skin Bioengineering Endeavour

Collecting accurate clinical data in clinical trials relies on qualified investigators. Clinical relevant parameters must be described in clear objective terms, e.g., for each grading scale. Any ordinal grading category must correspond to a very distinctive clinical description. The situation should be clearly defined for multiple investigators conducting a multicenter clinical trials in order to avoid unnecessary variations in clinical scoring. In addition, the indistinct gradient categories are further clouded when 0.5 ratings are permitted when the clinical sign severity appears to fall between two consecutive defined categories. The clinical trials are negatively influenced when using semiquantitative (ordinal) grading scales corresponding to vague and overlapping definitions. In sum, grading scales aiming to collect accurate data should strive to provide clinically distinct categories for the investigator.

The situation is less confusing for controlled noninvasive biophysical methods that prove to be calibrated, accurate, sensitive, specific, and reproducible. The power of some of these methods is superior to subjective scoring and clinical grading [1]. They have gained popularity in experimental dermatology and cosmetic science by reducing subjectivity of the clinical observations. Noninvasive assessments are ethical and applicable with only few restrictions in human trials. They provide objective and quantitative biophysical information on skin reactions linked to product efficacy on specific aims and allow reliable safety assessment, e.g., inflammatory reactions. They often disclose subclinical effects predicting the onset of overt clinical skin reactions. The correlation between data and clinical readings assures more specific and in-depth information. Bioinstrumentation provides statistical advantages by showing less interobserver and intraindividual variation in the measurements. The procedure of multi-parametric testing bypass limitations linked to a single device which provides information about a limited range in a given skin function [2].

Despite the advantages, it is clear that the mere beginner may believe that the understanding of skin physiopathology is well established in all circumstances using noninvasive dermometrology. In addition, one might expect that the possibilities offered by the variety of measuring devices can be used without restrictions. At present, there are only limited numbers of guidelines addressing standardization of techniques. Furthermore, only rare recognized quality control procedures are available for ensuring uniformity of data collection and interpretation. Although bioinstrumentation has proven its relevance in different clinical situations, still clinical judgments are preferred by regulatory agencies for the assessments of drug therapies in dermatology.

4.3 Validation of Methods and Instrumentations

Noninvasive measurements cover a growing field of skin biology, medicine, and cosmetology. The related developments are, for instance, useful for assessing both the efficacy and safety of topical products [3–5]. In addition, they are used for marketing purposes, e.g., in claim support by drug and cosmetic industries. Some of the

dermometrology techniques are unique and investigational, while others are present both in research units and at the bedside for monitoring patients. Still other commercially available techniques are available to laypeople lacking specific expertise, e.g., in cosmetic studios or pharmacies. The typical pitfalls indeed reside in the apparent but deceptive easiness in handling these devices.

There are some questionable use of noninvasive measurement techniques [1]. Sound methods are possibly subverted for mercantile purposes and falsehoods promoted under the guise of scientific information. They might create unsubstantiated claims and are at risk to be misinterpreted as worthy ones. In fact, the real value of measurements lies in the strict application of controlled procedures and, when available, standardized and with calibrated instruments. Lack of expertise of the measurement performer and some credulity of the observer are the two most nonscientific facets of unfulfilled noninvasive measurement of the skin.

In any case, the device must be calibrated and validated. The study aim and design must be supported by pre-discussed concepts and, at best, hypothesis-driven. The terminology must be conformed to consented and defined well-formulated definitions (Table 4.1). Before initiating a study, a series of questions should be raised when relevant to the purpose of the experimentation (Table 4.2).

Table 4.1 Instrumentation validation and definitions

Accuracy	Degree of similarity between the value that is accepted either as a conventional true value (in-house or local standard) or as an accepted reference value (international standard) and the mean value of repeated measurements. Provides an indication of systemic error.
Precision	Degree of similarity (scatter dispersion) among a series of measurements obtained from multiple sampling of the same homogenous sample under controlled conditions, expressed as a repeatability and reproducibility parameter.
Repeatability	Expresses the situation under the same conditions, that is, same operator, same apparatus, short time interval, identical samples.
Reproducibility	Expresses the situation under different conditions, that is, different laboratories, different samples, different operators, different days, different instruments from different manufacturers.
Range	The interval between the upper and lower levels for which the procedure has been applied.
Linearity	Ability of the procedure (within a given range) to obtain test results directly proportional to true values.
Sensitivity	Capacity of the procedure to record small variations or differences within the defined range.
Limit of detection	Lowest detectable change.
Limit of quantification	Lowest change above zero that can be quantitatively determined (not only detected) with defined precision and accuracy under the defined experimental conditions.
Ruggedness	Evaluates the effects of small changes in the test procedure on measuring performance.

Adapted from Ref. [6]

Table 4.2 Basic guidelines for bioengineering

1	What is the study purpose and what is the corresponding study endpoint?
2	Is the study endpoint of quantitative nature, narrow enough for a specific study, and suited to support the study purpose?
3	Stratification of endpoints into primary, secondary, tertiary, etc.…?
4	Shall one or several instruments be used (mono-instrumental or multipronged design)?
5	Function of the measurements and the instrument in the study: support, description, exclusion, comparison, validation during study, etc.…?
6	Which structure or function is actually being measured?
7	Range, linearity, and expected change of variables during study?
8	When should measurements be performed?
9	Inter-individual, intraindividual and intra-lesional variation, and if possible, variability data from normal and healthy skin?
10	Influence of gender, age, phototype and race?
11	Which season(s) will the study be performed (avoid more than 2 season; especially very warm and sunny season)?
12	Is a meaningful (ideally independent) control included?
13	Statistical evaluation of the design and the size of the sample studied?
14	Studies and literature validating the device(s).
15	If the target or measured area is small, do measurements need be repeated to overcome local site variation?
16	Existing in-house standards or recommendations, standard operating procedure (SOP)?
17	Guidelines and legal requirements, including ethical aspects?
18	Output from the instrument and source data. Are they handled and stored safely?
19	Are the laboratory facilities up to good standard?
20	Is a backup situation prepared if unexpected breakdown occurs in the device or in the laboratory?
21	Are ambient conditions such as temperature and humidity under control and expected to remain constant during the study period?
22	Needs for preconditioning of study subjects?
23	Are experimentalists well educated, trained, and prepared for the specific study; e.g., GCP training?
24	Are various types of bias identified and, whenever possible, eliminated?
25	How is it ensured or monitored that the study develops as planned, and what are the requirements for constancy and the consequences of inconstancy?
26	Calibration, maintenance, and control of instruments before, during, and at end of study?
27	Events and circumstances that exclude measurements from being performed or invalidate results?
28	How to conclude and report the study?
29	Timetable for the study—is it realistic and satisfactory?
30	Resources involved—are they available from start to end?
31	Is the study at an academic level where conclusion and interpretation are independent of economic interest, even if the study is supported by the industry?
32	Is the study documented and prepared for a special situation if some accusation about fraud would come up?
33	Are all study documents stored in an organized fashion to be re-visited over the next 10 (–15) years?

Developed from Ref. [6]

4.4 Standardization and Quality Controls

The scientific production is usually expected to be controlled at the level of peer review journals. The regulatory procedures must be applied before launching dermometrology-supported claims. Manufacturers of cosmetic products have perceived the great value of the continuously growing field of noninvasive biophysical assessment of skin physiology parameters. In general, drug industry and physicians have lagged behind for many years, but now show an increasing interest.

Optimization of noninvasive biophysical methods benefits from strictly calibrated devices as well as from controlled and standardized procedures of measurements. In spite of new developments in translational research between dermometrology, skin physiology, and biology, only a handful of publications have been devoted to standardization of measurements of skin properties. The situation is further clouded by some laypeople in the field of cutaneous biology who are owners of biophysical devices and who may speculate on data instead of scientific interpretation. This situation blurs the borderline between claim, dogma, axiom, and facts.

4.5 Search for "Good Biometrological Practice"

Two distinct features are important in establishing a good biometrological practice, namely the knowledge of the main physical characteristics of the devices and their modalities of application for measuring specific biophysical properties and functions of the skin. Ideally the measured effect can be quantified in SI units and correlates with a defined clinical feature, e.g., epidermal barrier permeability function measured in g/m^2 h.

The adequate conditions for measurement reproducibility must be strictly followed. Calibrations must be performed frequently and documented accordingly. The measuring procedure should be identical or similar in the laboratories using the same device. Each researcher may, however, adhere to local standard operation procedures to guarantee at least reproducibility of data in the laboratory setting over time.

Another problem deals with the relevant application of methods to skin-related biomedical problems. The basic knowledge of the biological aspect under evaluation is of paramount importance because it conditions the choice of the investigational method. Whenever possible, a combination of evaluation methods should be used instead of one single type of measurement. This is crucial for the validity of data interpretation. In addition, a given device is usually designed to measure one single biophysical parameter or function, but the collected data may be influenced by surrogate variables which are not evaluated by that specific device. The association of a series of distinct methods provides a better evaluation of complex interrelationships between cutaneous properties.

The experimental conditions have to be correctly predefined, controlled, and monitored. In general, this procedure is applicable to both the tested individuals and the environmental conditions. The characteristics of each panelist have to be chosen and recorded. This includes race, gender, age, exact anatomical region under investigation, and any other specific feature of interest for the study or potential influencing factor. The seasonal, ovarian, and nycthemeral cycles as well as diseases and previous interventions such as skin pre-conditioning clearly influence some measurements. The environmental conditions significantly alter a number of biophysical properties and functions of the skin. Thus, every single biometrological evaluation benefits from a controlled environment where temperature and relative humidity are monitored and carefully documented. Exposure to non-ionizing irradiations including ultraviolet light, total sunlight spectrum, and near-infrared energy strongly affect specific properties and function of skin, with sometimes long-term effects. It is also obvious that a series of drugs and cosmetics influence many measurements. Therefore, the choice of panelists or patients in a study is crucial. Their numbers are also critical [7]. The same is true for the control groups which should ideally comprise both positive and negative comparators.

The data interpretation in terms of biology may be difficult for the scientists even when oversimplified for commercial strategies. It should always include adequate statistical methods in combination with meaningful criteria of biological/clinical relevance. The reverse is also true, and, in some instances, increasing the number of measurements helps reaching validation using statistical analysis [7]. However, in a prospective study, the numbers of included subjects should be estimated when designing the study using power calculation.

4.6 Guidelines in Perspective

Evidence-based guidelines bring the best scientific evidence regarding diagnosis and management of a particular condition. They play an important role in educating the researcher about the state of science in a particular field. In recent times, guidelines have evolved from opinion-based expertise to evidence-based, consensual statements. The procedure is managed by experts, but the guidelines are ideally driven by the available unbiased and evaluated literature. Such an approach is scientifically concordant, and the guidelines developed using this evidence-based procedure represent improved educational tools.

Over time, several groups of experts have launched guidelines helpful in the field of noninvasive assessment in dermocosmetology [8–12]. The Standardization Group of the European Society for Contact Dermatitis, the European Cosmetic Toiletry and Perfumery Association (COLIPA), and the European Group for Efficacy Measurements on Cosmetic and Other Topical Products (EEMCO) group have substantially worked in that field [13] (see Table 4.3). They published a series of guidelines and reviews about noninvasive methods applicable to skin investigations. The

Table 4.3 Published guidelines for noninvasive measurements

Subject of the guideline	First author	Year	Reference
Transepidermal water loss	Pinnagoda	1990	[14]
Cutaneous blood flow	Bircher	1994	[15]
Skin color and erythema	Fullerton	1996	[16]
Dry skin and xerosis	Piérard	1996	[17]
SC hydration	Berardesca	1997	[18]
SLS irritation test	Tupker	1997	[19]
Skin color	Piérard	1998	[20]
Skin topography	Leveque	1999	[21]
Tensile functional properties part I	Piérard	1999	[22]
Tensile functional properties part II	Rodrigues	2001	[23]
Skin greasiness.	Piérard	2000	[24]
Transepidermal water loss	Rogiers	2001	[25]
Skin microcirculation	Berardesca	2002	[26]
Antiperspirants and deodorants	Piérard	2003	[11]
Skin surface pH	Parra	2003	[27]
Hair shedding and alopecia	Piérard	2004	[28]
Skin surface pH	Stefaniak	2013	[29]
Transepidermal water loss, skin hydration	du Plessis	2013	[30]
Stratum corneum hydration	Berardesca	2018	[31]
Transepidermal water loss, stratum corneum hydration, skin surface pH	van Rensburg	2019	[32]
Mechanical properties of skin and	Monteiro Rodrigues	2020	[33]

current landscape in the field of noninvasive skin methods includes a vast array of information ranging from educational to very sophisticated procedures.

4.7 Conclusion

Noninvasive measurements of skin physiology or dermometrology is a fascinating discipline looking for the contribution of many researchers coming from diverse scientific horizons. Fundamental researches, applied researches, and translational investigations have explored many facets of biophysical properties of skin. Routine use of bioengineering devices may look simple but proves to be a field with multiple pitfalls for the unexperienced beginners. Even skillful investigators are facing a number of problems related to relevance and interpretation of data. In every instances, emphasis should be placed on a strict respect of controlled and standardized conditions. At present, dermometrology is still in a developmental field where the researcher must control every single aspect of the measurements. Measuring devices provide pure numbers which are only relevant with a specific research question and defined hypothesis. The skill of the researcher and conversely the naive

handling by a non-educated person influence the value of the collected data. In addition, the interpretation of the biophysical measurements requires expertise. Should we need a license to manipulate biometrological devices as we need one to drive a car? Some regulatory procedures should probably be introduced to control claims and creative advertisements deceptively offered under the cover of dermometrology.

References

1. Agache P. Measurements: the why and how. In: Agache P, Humbert P, editors. Measuring the skin non invasive investigations, physiology and normal constants. Berlin: Springer; 2004. p. 6–15.
2. Fluhr JW, Lazzerini S, Distante F, Gloor M, Berardesca E. Effects of prolonged occlusion on stratum corneum barrier function and water holding capacity. Skin Pharmacol Appl Skin Physiol. 1999;12:193–8.
3. Berardesca E, Elsner P, Wilhelm KP, Maibach HI. Bioengineering of the skin: methods and instrumentation. Boca Raton, FL: CRC Press; 1995.
4. Hermanns JF, Pierard-Franchimont C, Pierard GE. Skin colour assessment in safety testing of cosmetics. an overview. Int J Cosmet Sci. 2000;22:67–71.
5. Ormerod AD, Dwyer CM, Weller R, Cox DH, Price R. A comparison of subjective and objective measures of reduction of psoriasis with the use of ultrasound, reflectance colorimetry, computerized video image analysis, and nitric oxide production. J Am Acad Dermatol. 1997;37:51–7.
6. Serup J. How to choose and use non-invasive methods. In: Serup J, Jemec GBE, Grove GL, editors. Handbook of non-invasive methods and the skin. Boca Raton, FL: Tailor and Francis; 2006. p. 9–13.
7. Kraemer HC, Thiemann S. How many subjects? Statistical power analysis in research. Newbury Park, CA: Sage Publications; 1987. p. 37–52.
8. Pierard GE. Relevance, comparison and validation of techniques. In: Serup J, Jemec GBE, editors. Handbook of non-invasive methods and the skin. Boca Raton, FL: CRC Press; 1995. p. 9–14.
9. Pierard GE. Instrumental non-invasive assessments of cosmetic efficacy. J Cosmet Dermatol. 2002;1:57–8.
10. Pierard GE. Streamlining cosmetology by standardized testing? J Cosmet Dermatol. 2002;1:97–8; author reply 98
11. Pierard GE, Elsner P, Marks R, Masson P, Paye M. EEMCO guidance for the efficacy assessment of antiperspirants and deodorants. Skin Pharmacol Appl Skin Physiol. 2003;16:324–42.
12. Serup J. Bioengineering and the skin: standardization. Clin Dermatol. 1995;13:293–7.
13. Pierard GE. EEMCO onward and upward. Streamlining its endeavour at the European venture in cosmetic efficacy testing. Int J Cosmet Sci. 2000;22:163–6.
14. Pinnagoda J, Tupker RA, Agner T, Serup J. Guidelines for transepidermal water loss (TEWL) measurement. A report from the standardization group of the European society of contact dermatitis. Contact Dermatitis. 1990;22:164–78.
15. Bircher A, de Boer EM, Agner T, Wahlberg JE, Serup J. Guidelines for measurement of cutaneous blood flow by laser Doppler flowmetry. A report from the standardization group of the European society of contact dermatitis. Contact Dermatitis. 1994;30:65–72.
16. Fullerton A, Fischer T, Lahti A, Wilhelm KP, Takiwaki H, Serup J. Guidelines for measurement of skin colour and erythema. A report from the standardization group of the European society of contact dermatitis. Contact Dermatitis. 1996;35:1–10.

17. Piérard GE. EEMCO guidance for the assessment of dry skin (xerosis) and ichthyosis: evaluation by stratum corneum strippings. Skin Res Technol. 1996;2:3–11.
18. Berardesca E. EEMCO guidance for the assessment of stratum corneum hydration: electrical methods. Skin Res Technol. 1997;3:126–32.
19. Tupker RA, Willis C, Berardesca E, Lee CH, Fartasch M, Agner T, Serup J. Guidelines on sodium lauryl sulfate (SLS) exposure tests. A report from the standardization group of the European society of contact dermatitis. Contact Dermatitis. 1997;37:53–69.
20. Piérard GE. EEMCO guidance for the assessment of skin colour. J Eur Acad Dermatol Venereol. 1998;10:1–11.
21. Leveque JL. EEMCO guidance for the assessment of skin topography. The European expert group on efficacy measurement of cosmetics and other topical products. J Eur Acad Dermatol Venereol. 1999;12:103–14.
22. Piérard GE. EEMCO guidance to the in vivo assessment of tensile functional properties of the skin. Part 1: relevance to the structures and ageing of the skin and subcutaneous tissues. Skin Pharmacol Appl Skin Physiol. 1999;12:352–62.
23. Rodrigues L. EEMCO guidance to the in vivo assessment of tensile functional properties of the skin. Part 2: instrumentation and test modes. Skin Pharmacol Appl Skin Physiol. 2001;14:52–67.
24. Piérard GE, Pierard-Franchimont C, Marks R, Paye M, Rogiers V. EEMCO guidance for the in vivo assessment of skin greasiness. The EEMCO group. Skin Pharmacol Appl Skin Physiol. 2000;13:372–89.
25. Rogiers V. EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences. Skin Pharmacol Appl Skin Physiol. 2001;14:117–28.
26. Berardesca E, Leveque JL, Masson P. EEMCO guidance for the measurement of skin microcirculation. Skin Pharmacol Appl Skin Physiol. 2002;15:442–56.
27. Parra JL, Paye M. EEMCO guidance for the in vivo assessment of skin surface pH. Skin Pharmacol Appl Skin Physiol. 2003;16:188–202.
28. Piérard GE, Pierard-Franchimont C, Marks R, Elsner P. EEMCO guidance for the assessment of hair shedding and alopecia. Skin Pharmacol Physiol. 2004;17:98–110.
29. Stefaniak AB, du Plessis J, John SM, Eloff F, Agner T, Chou T-C, Nixon R, Steiner MFC, Kudla I, Holness DL. International guidelines for the in vivo assessment of skin properties in non-clinical settings: part 1. Skin Res Technol. 2013;19:59–68.
30. du Plessis J, Stefaniak A, Eloff F, John S, Agner T, Chou T-C, Nixon R, Steiner M, Franken A, Kudla I, Holness L. International guidelines for the in vivo assessment of skin properties in non-clinical settings: part 2. Transepidermal water loss and skin hydration. Skin Res Technol. 2013;19:265–78.
31. Berardesca E, Loden M, Serup J, Masson P, Rodrigues LM. The revised EEMCO guidance for the in vivo measurement of water in the skin. Skin Res Technol. 2018;24:351–8.
32. van Rensburg SJ, Franken A, Du Plessis JL. Measurement of transepidermal water loss, stratum corneum hydration and skin surface pH in occupational settings: a review. Skin Res Technol. 2019;25:595–605.
33. Monteiro Rodrigues L, Fluhr JW, the EEMCO Group. EEMCO Guidance for the in vivo assessment of biomechanical properties of the human skin and its annexes: revisiting instrumentation and test modes. Skin Pharmacol Physiol. 2020;33:44–60.