Practical Aspects of Cosmetic Testing

How to Set up a Scientific Study in Skin Physiology
The idea of “practical guidebook” was born after a hands-on workshop where several participants asked about a recommendation for reading of practical aspects. The current handbooks offer a plethora of scientific overviews and cover the broad spectrum of noninvasive measurement devices for cosmetic skin testing. However, practical aspects of performing cosmetic aspects are not always covered. Also, the published guidelines do not always cover the day-to-day questions arising during the preparation, performance, and evaluation of clinical studies. The aim of the present book is to provide practical guidance for scientists, especially those new in the field or those who face practical problems with their studies. New lab members should have a useful first-to-read source at hand.

I would also like to honor some “corner stones” in the development of modern biophysical instrumentation such as Rony Marks, Harvey Blank, Pierre Agache, Gary Grove, Jorgen Serup, Howard Maibach, Peter Elsner, Enzo Berardesca, Albert Kligman, and the most innovative company in the field, Courage & Khazaka. Some of them have played an important role in the development of my personal career.

I would like to thank all authors of this book. Without their dedicated contributions this project would not have been possible. Special thanks should go to Ms. Blasig, from Springer. She supported this project during its entire process with enthusiasm and dedication.

Albert Kligman had a saying, which I would like to keep in mind when starting and advancing in the field of biophysical assessment of skin functions: “A fool with a tool is still a fool”. Thus, the brain of the scientist should be active when performing and analyzing measurements. Hopefully this book will fill the gap between the detailed scientific textbooks, original and review publications in international journals, and the practical hands-on training that needs to be integrated in the education of young scientists in cosmetic testing.

Berlin, October 2010

Joachim W. Fluhr
# Contents

**Introduction: History and Progress of Objective Skin Analysis**
Joachim W. Fluhr

## Part I  Legal Aspects of Cosmetic Testing

1 Regulatory Aspects ................................................................. 3
   Oliver Wunderlich

2 Ethical Aspects of Cosmetic Testing .................................... 15
   Hristo Dobrev

3 Good Clinical Practice ............................................................ 27
   Betsy Hughes-Formella

4 Guidelines in Dermocosmetic Testing ................................. 33
   Gerald E. Piérard, Valerie Docquier, Anelavre Schreder, Sophre Neuberg,
   and Claudine Piérard-Franchimont

5 Claim Support: How to Create and Substantiate Claims ........ 43
   Sinead Hickey and Steve Barton

## Part II  General Aspects of Cosmetic Testing

6 Testing Laboratory ................................................................. 65
   Oliver Wunderlich

7 Research Staff ........................................................................ 75
   Ragna Williams

8 Testing Population ................................................................. 79
   Ragna Williams

9 Testing Devices and Methods ................................................ 83
   Gabriel Khazaka
Contents

10 Factors Influencing Measurements ......................................................... 89
   Enzo Berardesca

11 Study Design .......................................................................................... 101
   Betsy Hughes-Formella

12 Cosmetic Testing Report ......................................................................... 107
   Hristo Dobrev

Part III Practical Aspects of Testing: Typical Examples of Test Settings

13 Moisturizers and Emollients ................................................................. 123
   Razvigor Darlenski and Joachim W. Fluhr

14 Antiaging and Antiwrinkle Products .................................................... 143
   Razvigor Darlenski, Theresa Callaghan, and Joachim W. Fluhr

15 Products for Impure, Acne-Like Skin ................................................... 155
   Hristo Dobrev

16 Assessment of Hair Morphology .......................................................... 171
   H. Zahouani and M. Fougère

17 Skin Colour and Pigmentation .............................................................. 179
   Elisabeth Ammitzbøll Holm

18 Characterization of Sunscreens: Determination of the SPF ............... 187
   Juergen Lademann and Joachim W. Fluhr

19 Practical Aspects of Shampoo and Conditioner Testing .................... 197
   Trefor Evans

20 Antiperspirants and Deodorants ............................................................ 217
   Razvigor Darlenski and Joachim W. Fluhr

21 Hair Growth ............................................................................................ 227
   Tobias W. Fischer

22 Sensory Perception .................................................................................. 247
   Gregor B.E. Jemec

23 Practical Use and Significance of Transepidermal Water Loss
   Measurements .......................................................................................... 253
   Truus Roelandt, Diane Roseeuw, and Jean Pierre Hachem

24 Compliance Check with On-Site Measurements: Advanced Ways in
   Product Testing ....................................................................................... 261
   Gabriel Khazaka

Index ........................................................................................................... 267
In the era of evidence-based medicine we are witnessing a growing demand for standardization and objective assessment of different physiologic and pathologic conditions as well as for monitoring the efficacy of different therapeutic modalities. The same trend is seen for cosmetic studies. A great number of devices for efficacy testing of cosmetic products have been invented and developed in the past decades. However, a lot of questions remain open when proving efficacy and safety of cosmetic products:

- What parameters and devices are suited to test specific claims?
- How to perform the measurements in practice?
- What should be the test population and what study protocol is most appropriate?
- What environment-, subject-, and instrument-related conditions should be considered?
- What are the regulatory and ethical aspects of cosmetics testing?

This book is intended in the first line for answering the above-stated questions. It will be of practical interest especially to the “new-bees” in the field of cosmetic research, including cosmetic scientists, cosmetic chemists but also dermatological researchers of functional aspects (clinical assessment of disease activity and skin physiology), pharmacists, biologist, and biochemists. Furthermore, professionals working in clinical study centers and contract research organizations (including project managers, study nurses, and physicians) will benefit from the lecture of the present book. The targeted practical point of view together with the “step-by-step” approach when planning and performing cosmetic testing are the main advantages of this manual in comparison to former books in the field.

Skin physiology assessment is moving rapidly from a descriptive approach to a deeper understanding of biophysical and biochemical processes in the epidermis, namely epidermal barrier function, stratum corneum hydration, and the underlying regulated processes. The research with noninvasive biophysical measurements, formerly called bioengineering methods, offer now reliable and reproducible approaches for product testing in the cosmetic industry as well as in basic research. Herein, basic information on technical and legal aspects on cosmetic testing is presented. The authors give insight into very practical aspects of basic skin physiology and the assessment of skin functions in controlled studies. The last (and the broadest) part of the book is dedicated to specific typical examples of test
settings. Of course, a book dedicated to address basic aspects cannot cover the entire spectrum of possible test approaches. The most frequently used endpoints are described.

Research with non or minimal-invasive devices to study skin physiology and the effect of cosmetic products enters in its 5th decade. We are now using the “fourth generation” of instruments. It started in the early 1970s with the assessment of stratum corneum hydration, epidermal barrier function (by measuring transepidermal water loss), and skin mechanical properties. The instruments were often the size of a cupboard or table. Individual instruments were built in the labs and most of the time only prototypes were manufactured that never made their way to wide distribution and general acceptance.

These “first generation” individual instruments were often designed by cosmetic companies for testing their specific product claims. The “second generation” instruments were stand-alone devices, significantly smaller and cheaper. They were the first ones built on larger scales, thus accessible for broader public and academic institutions (Fig. 1).

The “third generation” of measurement devices consisted of instruments attached to a PC via a central-unit allowing direct storage of measurement values. Some manufacturers minimized the measurement technology in such a way that the device was actually in the handheld itself. A skin physiology lab could now fit into a small suitcase. The “fourth generation” is now available with easy to perform calibration check. State of the art instrumentation allows good validation studies and interlaboratory comparisons. The next step (maybe the “fifth generation”) should be to transform the measurement units, today in most cases arbitrary units (AU), to SI units.

The work on standardization led to the publication of several guidelines (see Chap. 2.4. by Pierard et al.) Unfortunately, the working group of European Group for Efficacy Measurements on Cosmetics and other Topical Products (EEMCO) is no longer in place and thus no update on these guidelines will be available in the near future. Maybe, widely accepted standard operation procedures (SOPs) will be implemented or guidelines will be published that are accepted by regulatory authorities. Systematic reviews on the evidence (EBM) for different methods or compounds might help to improve the standards in cosmetic testing. Another step would be to harmonize training courses and maybe to install training certificates based on standardized training sessions. There is still a lot of work to
do, but we have reliable instruments and knowledge available to perform good scientific studies. Rigorous scientific planning together with accurate data analysis will ensure and enhance the credibility of cosmetic testing, especially if some of the standards already in place for testing of topical drugs (e.g., comparing to placebo/control, sample size calculation, and submission to authorized ethic comities) are implemented in the cosmetic study protocols. Today, consumers are well informed not only via the internet, but also due to the easily accessible information especially by the big cosmetic companies. Thus, claims have to be substantiated with good science and controlled clinical studies. Noninvasive instrumentation is a corner stone of standardized clinical testing.

The present book provides basic knowledge on how to plan, perform, and evaluate scientific studies. The authors are recognized experts in the field and describe in comprehensive chapters the practical aspects of noninvasive measurements. The first part of the book is dedicated to regulatory aspects of cosmetic testing including guidelines, ethical aspects, and claim support. The second part deals with the general aspects of cosmetic studies, namely requirements of the testing laboratory, testing staff, populations, study design, and the reporting of the study outcomes. The third part is dedicated to some typical examples of test settings including efficacy claim studies for moisturizers and emollients, antiaging and antiwrinkle products, antiperspirants and deodorants, and cosmetics for impure skin; and assessment of hair morphology, skin color, and others.

Getting acquainted with the good practice in cosmetic testing would be helpful to the reader not only for better practical performance of a study, but also to interpret and evaluate the strong and weak points of other investigators’ research.

The book should guide into good planning, careful performing, and critical interpretation of cosmetic studies. Of course, reading does not replace hands-on training and personal experience.
Part I

Legal Aspects of Cosmetic Testing
In general, the regulatory basics for cosmetics are different in different countries. Even the classification rules defining what a cosmetic is differ between countries. A product that is classified as a cosmetic in one country may be classified as a drug in another country. This has implications for the testing of cosmetics. Depending on which country the testing is

O. Wunderlich
Bioskin, Burchardstrasse 17, 20095 Hamburg, Germany
e-mail: oliver.wunderlich@bioskin.de

J. Fluhr (ed.), Practical Aspects of Cosmetic Testing, DOI: 10.1007/978-3-642-05067-1_1, © Springer-Verlag Berlin Heidelberg 2011
performed and/or the test product is (to be) marketed, different specific testing may be required, desirable and/or allowed. The following chapter gives a brief introduction on the regulatory aspects of clinical testing of cosmetics. However, this chapter serves for informational purposes only and does not provide legal advice.

1.1 Comparison Between General Cosmetic Legislation in Europe and Other Countries

The differences in regulatory basics for cosmetics in different countries are numerous, complex, and may be confusing for the regulatory layman. Fortunately, there are also some universal similarities at least for the major market economies:

- The marketer has full responsibility for the safety of products – the expected manner of use must not be harmful for the health of the consumer. No premarket approval by the authorities is necessary.
- In-market control of the cosmetic products is performed by the respective authorities (in different ways).
- Any distribution channel for the product may be used (shops, mail order etc.).
- Claims and other information given to the consumer regarding the cosmetic product must not be misleading.
- For ingredient declaration on the packaging, the INCI (International Nomenclature of Cosmetic Ingredients) system is widely used and required.

Furthermore, the cosmetics regulation frameworks may be classified into two large groups:

1. Regulation systems with broad definitions of cosmetics. These employ extensive lists with restrictions for specific ingredients as well as positive lists for allowed ingredients and require safety data to be available.

This framework model roughly describes the EU cosmetics regulation. Considering its success in regulating cosmetic product safety on the one hand and allowing innovation of cosmetic products on the other, and keeping in mind the global importance of the European cosmetic product market, it is not surprising that many emerging countries have modeled their cosmetic regulation systems after the European example. In fact, the cosmetic regulations of the ASEAN countries (Indonesia, Malaysia, the Philippines, Singapore, Thailand, Brunei, Burma (Myanmar), Cambodia, Laos, and Vietnam), the Mercosur countries (Brazil, Argentina, Paraguay, Uruguay), the Andean Pact countries (Bolivia, Colombia, Ecuador, and Peru) as well as South Africa are very similar to the respective EU regulations (except regarding the ban on animal testing, which up to now remains a European “specialty”).

2. Regulation systems with narrow definitions of cosmetics. These impose few specific restrictions regarding ingredients and few requirements regarding available safety data for cosmetics. However, depending on the claims made, or depending on contained ingredients for which a therapeutic effect is known, many products that may be classified
as a cosmetic within the former regulation systems may be classified as an over-the-counter (OTC) drug here. This framework model roughly describes the cosmetics regulatory system in the USA.

The regulatory systems of two further major markets, Japan and Canada, fall between the two antipode systems described above. The Japanese system also works with positive and negative ingredient lists, but features a third, intermediate category of products, the quasi-drugs. Quasi-drugs are defined as articles that are used for certain purposes/indications/claims and are restricted to a list issued by the Ministry of Health and Welfare Japan (MHW). The Canadian system is similar to the US system, but with longer restriction lists (“Cosmetic Ingredient Hotlist,” modeled after the negative lists valid in the EU).

The cosmetics regulatory system of China, a market that is of increasing interest, is somewhat different, but currently under review and may change considerably in the future. Currently, the regulation system (which is integrated in the drug regulation) differentiates between “ordinary” and “special cosmetics.” To put it simply, the former category consists only of decorative and cleansing cosmetics, and the latter category of cosmetics that have a function that goes beyond that. Furthermore, cosmetics are also handled differently depending on whether they are produced by a domestic Chinese manufacturer or imported from outside of China: Imported cosmetics must undergo a rather lengthy and complicated premarket registration and approval process, requiring strictly specified safety testing and acquisition of several licenses involving several national and regional government bodies. Furthermore, all testing has to be done in China. These obstacles have been criticized by some representatives of western economies as being discriminatory. It will be interesting to observe if this will change in the future.

Lastly, India is also a large market that was of less interest to major cosmetic manufacturers in the past, probably because of the low average income of the population. However, it is becoming increasingly important considering its economic growth. Its cosmetic regulatory system is completely integrated in its drug regulatory system (its origins dating back to the 1940s), with rather narrow definition of cosmetics similar to the USA system.


1.2 Recent Changes in European Cosmetic Regulation

On March 24, 2009, the European parliament recasted (with a few amendments) the fundamentals of European cosmetics law, the European Cosmetics Directive 76/68 EEC, which originated in 1976 and was updated and amended many times until today. The original directive was a European guideline which had to be translated into national law by each of the EU states. This led to a few differences between the cosmetics regulations in the EU countries, for example, in the “positive lists” containing allowed ingredients (colorants, preservatives, UV filters).

The present recast is a directly effective legal act, eliminating such differences. Furthermore, it integrates all former amendments and includes some clarifications and
definitions, for example, a glossary on legal terms used. For the first time, it is specified
which safety tests and assessments are to be done for the safety dossier that is to be kept on
file at the manufacturer. The specifications of the safety dossier were modeled on the recom-
mandations by the Scientific Committee on Consumer Products (SCCP) on safety testing of
cosmetics, which up to now were required only for new ingredients to be added to the posi-
tive lists. For the first time, the use of nanomaterials is specifically regulated. Furthermore,
the recast stipulates EU-centralized premarket submittal of basic information on the product
in order to facilitate and strengthen postmarketing surveillance and “cosmetovigilance.”
Recast of the European Cosmetics Directive:
For tracking of the legislative status this link may also be checked:

1.3
Important Weblinks

1.3.1
European Union

Consolidated version of the seventh amendment to the European Cosmetics Directive:
SCCP central website:
Important SCCP guidelines/opinions:
Colipa (European Cosmetics Association) central website:
http://www.colipa.eu
German Cosmetic, Toiletry, Perfumery and Detergent Association (IKW):
http://www.ikw.org
General information for marketers:
koerperpflegemittel&page_title
German Federal Institute for Risk Assessment (BfR):
http://www.bfr.bund.de/cd/template/index_en

1.3.2
USA

http://www.cfsan.fda.gov/~dms/cos-toc.html
The Personal Care Products Council (formerly the Cosmetic, Toiletry and Fragrance Association (CTFA)):
http://www.personalcarecouncil.org

1.3.3
Canada

Health Canada, centralized information on cosmetics regulation:
Cosmetic Ingredient Hotlist:

1.3.4
Japan

Ministry of Health, Labour, and Welfare (MHLW):
http://www.mhlw.go.jp/english/topics/cosmetics/index.html
No official translation of the Pharmaceutical Affairs Law (PAL) available, unofficial version:
http://www5.cao.go.jp/otodb/english/houseido/hou/lh_02070.html
Japan Cosmetic Industry Association (JCIA), available only in Japanese:
http://www.jcia.org/

1.3.5
Mercosur Countries (Examples)

Argentinian National Administration of Pharmaceuticals, Food and Medical Technology (ANMAT), search also for Resolución 155/98 del 13/03/98:
http://www.anmat.gov.ar/cosmeticos.asp
Brazilian National Health Surveillance Agency (ANVISA), search also for Resolução nº 79, de 28 de agosto de 2000
http://www.anvisa.gov.br/e-legis/

1.3.6
ASEAN Countries (Example)

Singapore Health Sciences Authority (HSA):
1.3.7 South Africa

South African Government Information, search for “FOODSTUFFS, COSMETICS AND DISINFECTANTS ACT, 1972 (ACT NO. 54 OF 1972)”:  
Cosmetics, Toiletry and Fragrance Association of South Africa (CTFA):  
http://www.ctfa.co.za/

1.3.8 China

State Food and Drug Administration (SFDA):  
http://eng.sfda.gov.cn/eng/  
General Administration of Quality Supervision, Inspection and Quarantine of P.R.C. (aqsiq)  
http://english.aqsiq.gov.cn/

1.3.9 India

Central Drugs Standard Control Organization (CDSCO):  
http://www.cdsco.nic.in/html/law.htm  
http://cdsco.nic.in/html/Copy%20of%201.%20D&CAct121.pdf

1.4 Cosmetic Safety Testing

This section mainly addresses the situation regarding cosmetic safety testing in Europe. However, in the other major markets most aspects apply accordingly. The most important difference may be the different views on ethical aspects regarding the availability of animal data prior to human testing. In Europe, such data may be replaced by alternative (nonanimal) methods, animal data in some cases even being impossible to obtain due to the animal testing ban. In contrast, in other major markets such animal data is required prior to testing in humans.

The most important property of a cosmetic is its safety, which must be checked prior to marketing a cosmetic product (EU Cosmetics Directive).

In the recast of the EU cosmetics directive as well as in guidelines by the SCCP and Colipa, the product safety properties that should be addressed and documented are listed.

Already in 1999, the SCCP issued a “Guideline on the use of human volunteers in compatibility testing of finished cosmetic products” providing the basic principles. Weblink:
These were based on the declaration of Helsinki, Good Clinical Practice principles and national regulations regarding human studies. Helpful in this regard are also the “Notes of guidance for testing of cosmetic ingredients for their safety evaluation” issued by the SCCP, which are updated occasionally. Weblink:

http://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_o_03j.pdf

In these guidelines, it was made clear that “cosmetic compatibility tests on human volunteers cannot be considered as a replacement for animal testing,” and that such tests “can only be performed to confirm...that products do not damage skin and mucous membrane, as already expected from other sources.”

There is no explicit legal requirement that the finished products have to be tested in humans at all prior to marketing. There is also no specific regulation on which tests have to be completed prior to testing a cosmetic product clinically in humans. However, it is obvious that the toxicological profile must be available and that there are no concerns based on the data. For instance, if there are, for example, indications for corrosivity of the test product in the nonclinical test model, the product should not be tested in humans to prove the opposite!

Parameters on which data should be available, either based on own tests or derived from literature data (e.g., of known individual ingredients) include:

- Corrosivity
- Mutagenicity
- Genotoxicity
- Carcinogenicity
- Reproductive toxicity
- Dermal/percutaneous absorption
- Phototoxicity
- Acute and repeated dose toxicity
- Sensitizing/photosensitizing potential

1.5 Responsibility Considerations for Planning and Conduct of a Cosmetic Safety Study

Even for clinical testing under a controlled environment, for example, by a Clinical Research Organization (CRO) specialized in these tests, the responsibility for the safety of the test products ultimately remains at the manufacturer. However, for both ethical and liability reasons, the testing organization should always scrutinize all aspects of a planned clinical study to ensure that the health of testing subjects is not harmed due to the study. Even if a specific subject insurance is contracted, which is rarely the case for cosmetic studies since it is not legally required (other than for drug studies), a testing organization may still be liable for compensation of damages to subjects in cases of negligence, which is never covered by this type of insurance. In this respect, the responsible staff of the
testing organization takes a significant part of the responsibility for the safety of the subjects during clinical testing and should therefore always try to minimize the risk for the volunteers to the best of their knowledge.

Often, the manufacturer wants to keep existing data on the test product (e.g., ingredients) confidential, so no complete “picture” of the product is available to the testing organization. However, the testing organization should at least insist on a confirmation by the manufacturer/its safety assessor that the safety assessment and toxicological profile of the test product was considered and the test product is judged safe under the conditions of the study. A formal signed release of the test product stating this must be issued prior to the study. Furthermore, it should be confirmed by the manufacturer that the product conforms to the local (e.g., European) cosmetics laws, for example, by containing ingredients only in the permitted concentrations. It should also be confirmed whether or not the test product contains ingredients never used before in marketed cosmetics. If it does in fact contain a novel ingredient, the testing organization should insist on more information (e.g., a risk assessment by an expert) to be able to take responsibility for the safety of the volunteers during the study.

In general, it should be kept in mind that the list of ingredients in the test product is very helpful for testing. First, to be able to protect test volunteers by excluding them from the study if they already have a known hypersensitivity to a certain cosmetic ingredient which is contained in the test product. Second, an experienced testing organization can consult on the correct choice of the adequate study design to meet the study objectives, which is often critically dependent on the general characteristics of the product, the formulation type, or certain ingredients.

It is not legally required for cosmetic tests that an independent ethics committee reviews the study documents prior to the study. However, this should always be considered, at least if a residual risk for the volunteers is present, for example, if the test product contains novel ingredients, or if invasive or stressful subject procedures are planned for the study. The ethics committees have a perspective “from outside” which can be helpful to detect safety issues.

1.6 Frequent Cosmetic Safety Study Models

Testing cosmetic safety in humans mainly includes investigation of the acute and/or chronic unwanted effects of application of cosmetics on the skin. For this purpose, a variety of safety study models to simulate or even exaggerate the normal conditions of use are employed. Most frequent models are:

- Dermal irritation patch tests, exaggerating normal use. These test models employ controlled single or repeated application of test products to small test fields using special application systems (open, semiocclusively or occlusively). Skin reactions to the test products such as skin reddening (erythema) or scaling are graded by trained observers using standard or modified clinical assessment grading scales.
• In-use irritation tests, simulating normal use. Here the test products are applied repeatedly in an open manner by the subjects. The skin reactions are assessed by trained observers; in addition, biophysical data of the skin (e.g., skin moisture, transepidermal water loss) may be collected. Subjective assessments by the subjects complete the picture.

• Human Repeat Insult Patch Test to assess the sensitization potential of a test product. The product is repeatedly applied during an induction phase, after which follows a rest period and challenge application. Skin assessments for sensitization reactions (allergic potential) are performed following the challenge application. This kind of test is quite controversial. Only very limited information on the long-term consequences for volunteers who have been sensitized during these tests is available. On the other hand, the sample sizes used for this test are often too small to reliably predict a sensitization potential, rendering the study insufficient to meet the study aim.


• At the moment, there is no validated replacement test method available, only a refined animal test (the Local LymphNode Assay). The best approach is apparently choosing ingredients with known low sensitization potential and avoiding those with a known high sensitization potential.

Over the years, the SCCP has also issued several further guidelines/opinions in the area of clinical testing of the safety of cosmetic products.

Weblinks:
http://ec.europa.eu/health/ph_risk/committees/sccp/docshtml/sccp_out45_en.htm

See also the “Guidelines for Assessment of Skin Tolerance of Potentially Irritant Cosmetic Ingredients” issued by COLIPA in 1997.

Weblink:

1.7 Cosmetic Efficacy Testing

In Europe, scientific data substantiating the claims made on the packaging must be available in the product information file (stipulated already in the sixth amendment of the EU cosmetics directive). This data collection may consist of nonclinical (e.g., derived from cell cultures) and/or clinical study data.

In the United States, the situation of cosmetic claim substantiation is quite complicated; however, here also the efficacy claims must be reasonably substantiated to avoid diverse sanctions. The enforcement of claim substantiation standards is shared mainly between the FDA and the FTC (Federal Trade Commission). A good overview of this topic is given by McEwen and Murphy [1].
In Japan, data substantiating efficacy claims is required only for quasi-drugs (i.e., their specific active ingredients) and not for cosmetics. Only specific, authorized claim wording may be used.

The efficacy of cosmetic products may be tested only if there are no founded concerns regarding safety (see section “Cosmetic safety testing”). This is true regardless of the fact that the safety of the product is almost always “cotested” as a secondary objective in a cosmetic efficacy study (e.g., by observing any adverse reactions).

In 2008, the COLIPA issued a revised “Guideline for evaluation of the efficacy of cosmetic products.” This guideline contains general principles for efficacy tests, requirements for test protocols and reports, as well as some sample human and nonhuman efficacy testing models.


Various test designs have been developed in the past years to address the multifaceted requirements of efficacy claim substantiation, driven by marketing interests as well as progress in the understanding of skin physiology. Progress in the field of biophysical measurement and standardized photodocumentation methods provides new opportunities.

In efficacy testing of cosmetics, there are very few test models that are standardized across the industry (one of the exceptions being the International SPF method).

A large number of diverse cosmetic efficacy studies have been published. However, keeping in mind the many possible aspects involved for the diversity of cosmetic products, a cosmetics manufacturer will still need to rely on an experienced cosmetics testing organization to select the adequate study design.

Considering the diversity of the topic, a description of the types of efficacy tests would go too far. However, there is one basic principle to be kept in mind: To deliver relevant substantiation for efficacy claims, the study must simulate normal use.

---

1.8 Cosmetic Labeling and Packaging

The labeling and packaging of cosmetics put on the market is regulated in detail in the respective laws in force in the major markets. For instance, the EU cosmetics directive stipulates that the labeling must contain:

- Name of the marketer
- Weight or volume
- Date of minimum durability (if less than 30 months, otherwise, period of durability after opening)
- Precautions for use
- Batch number
- Product function

Furthermore, in all major markets, the ingredients must be listed using the INCI nomenclature.
In 2006, an international standard for cosmetics packaging and labeling was proposed (ISO 22715:2006). This standard might be adopted internationally in the future.


An important issue of cosmetic labeling is the efficacy claim which usually appears both on the packaging and advertising material. The claims that are allowed depend on the respective regulatory system (see section above).

A special feature of cosmetics labeling in the USA is that a cosmetic product for which no adequate safety data is available might still be put on the market provided they are labeled “Warning: the safety of this product has not been determined.”

There are no specific regulatory requirements for labeling and packaging of test products for use in a clinical cosmetic study. Apart from common-sense responsibility for the test volunteers, practical aspects are most important here, for example, whether the product is handed out to the volunteers for application at home, storage requirements, blinding, etc. However, the safety of the volunteers should be paramount. The label should include all precaution statements that might be reasonable in the specific setting. Typical labels of test products to be handed out to test volunteers (in-use tests) may include precaution wording such as “for cosmetic study use only,” “keep out of reach of children,” “for external use only,” “store at room temperature in a safe place” etc. Further, brief specific instructions for use may be added on the label. The usage instructions should also be included in an extensive subject information sheet or a treatment diary. Contact information of the investigator may also be added on the label.

References

2. RPA Ltd. for European Commission Directorate General Enterprise: Comparative Study on Cosmetics Legislation in the EU and Other Principal Markets with Special Attention to so-called Borderline Products. Final Report – August 2004
Ethical Aspects of Cosmetic Testing

Hristo Dobrev

Core Message

Cosmetic product safety and claim substantiation have evidently progressed during the past years. A number of skin bioengineering techniques and instrumentation have been developed that are able to prove various cosmetic claims. It is very important that the cosmetic testing on humans is conducted ethically and follow proper scientific design. Compliance with the basic ethical principles originated in the Declaration of Helsinki, and internationally accepted scientific principles of the Good Clinical Practice provides public assurance that the rights, safety, and well-being of participants are protected and that the study data are credible.

2.1 Introduction

Ethical considerations are an essential part of any biomedical research involving human subjects [16, 22].

Medical research is a research conducted to increase the knowledge in the field of medicine. It can be divided into two main categories: basic science (nontherapeutic or nonclinical) medical research and applied (therapeutic or clinical) medical research (clinical trial). The first one predominantly involves healthy persons and is carried out to increase the understanding of fundamental principles and thus to contribute to the applied clinical research. The second one involves sick persons and is intended to evaluate a new diagnostic or therapeutic method for both safety and efficacy.

Studies involving skin measurement methods and testing of cosmetic products on humans are similar to medical research. They involve the use of human beings as research...
subjects and also deal with pure scientific research, whose primary purpose is to contribute to generalized knowledge about the human skin physiology and active substances, and with applied research, aimed to evaluate the safety and efficacy of new cosmetic ingredients and finished products.

In both studies, the ethical considerations are related to the relationship between the physician/the investigator and the human subject/the healthy or sick volunteer and their main objective is the protection of the human being. So, the ethical considerations for cosmetic testing and use of skin measurements are similar to those for medical research on humans, particularly nontherapeutic research. They are subject to the ethical principles of the Declaration of Helsinki and the guidelines for Good Clinical Practice (GCP), and are integrated into the research design.

The aim of this chapter is to outline the ethical aspects of cosmetic testing using non-invasive skin methods.

2.2 Brief History of Research Ethics

Ethics is a set of principles of right human conduct. It deals with moral values such as good or bad, right or wrong, appropriate or inappropriate. Medical ethics is a branch of so-called applied ethics, which explores the application of moral values in medicine. Medical ethics encompasses mainly its practical application in clinical settings and is treated as an applied professional ethics. Research ethics is also a field of applied ethics, which involves the application of fundamental ethical principles to scientific research. It is most developed as a concept in medical research and includes the design and implementation of research involving human experimentation.

Professional medical ethics originates in the Hippocratic Oath written in the fourth century BC by Hippocrates. It is an oath traditionally taken by physicians with which they become obliged to act in conformity with the rules of medical profession and to current best practice for the benefit of the patients. In modern medicine, the significance of the Hippocratic Oath has been reduced to a symbolic right of passage for medical school graduates [23].

The first Code of medical ethics was written by the American Medical Association (AMA) in 1846. It was based upon the guidelines of the English physician Thomas Percival (1740–1804) of Manchester related to physician consultations. This code of ethics dictates the moral authority and independence of professional physicians in service to others and their responsibility towards the sick, as well as the physician’s individual honor [1].

The Nuremberg code (1947) was the first international instrument on the medical research ethics. It was adopted as a consequence of the Nuremberg trial of physicians (the Doctors’ Trial) at the end of the Second World War. The Code was designed to protect the integrity of the research subject and sets out ten conditions for the ethical conduct of research involving human subjects. Among them were such principles as voluntary informed consent, favorable risk–benefit assessment, performance by scientifically qualified persons, termination of the experiment at any stage by subject or scientist either voluntarily or in response to excessive risk, pain, or injury [14, 17, 18].
The Nuremberg code was followed by the Declaration of Geneva and World Medical Association International Code of Medical Ethics. The Declaration of Geneva was adopted by the second General Assembly of the World Medical Association at Geneva in 1948. It was attended as a modern updated revision of the ancient Hippocratic Oath and represents the physicians’ dedication to the humanitarian goals of medicine. The Declaration of Geneva has been revised several times since, most recently in 2006 [27]. The WMA International Code of Medical Ethics was adopted by the third General Assembly of the World Medical Association in London in 1949 and revised in 1968, 1983, and 2006. It indicated the duties of the physicians in general as well as the duties of the physicians to their patients and colleagues [28].

The fundamental document in the field of human research ethics is the Declaration of Helsinki. It was originally adopted at the 1964 World Medical Association General Assembly in Helsinki, Finland, and has undergone six revisions since then (the most recent in October 2008). The Declaration of Helsinki is a comprehensive international statement of the research ethics involving human subjects. It sets out basic ethical guidelines for the medical community regarding the protection of human beings involved in both clinical and nonclinical biomedical research. The first revision of the Declaration of Helsinki (1975) introduced the concept of oversight by an “independent committee” which became a system of Institutional Review Boards (IRBs) in the US, also known as independent ethics committees (IEC) or ethical review boards (ERBs) in other countries, which are empowered to review, approve, and monitor biomedical research involving humans with the aim to protect the rights and welfare of the research subjects. The Declaration of Helsinki was the basis for GCP used today [2, 25, 26, 29].

In 1979, the US National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research published the Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects Research (“The Belmont Report”). It provides guidance for distinguishing therapeutic medicine from research, identifies three fundamental ethical principles for the protection of human subjects (respect for persons, beneficence, and justice), and shows how these ethical principles apply to the conduct of human research (informed consent, assessment of risk and benefits, selection of subjects). These principles continue to provide the ethical foundation for conducting research with human subjects [15, 18].

In 1981, the Department of Health and Human Services (DHHS) issued regulations based on the Belmont Report named Code of Federal Regulation (45 CFR 46). Later, the core of these regulations was formally adopted as “The Federal Policy for the Protection of Human Subjects”, or “Common rule” (1991), which is a rule of medical ethics in the United States [10]. The main elements of the Common Rule include requirements for assuring compliance by research institutions, requirements for researchers obtaining and documenting informed consent, requirements for IRB, additional protections for certain vulnerable research subjects – pregnant women, prisoners, and children [18].

After 1982, the Declaration on Helsinki is not the sole universal guide, since the Council for International Organizations of Medical Sciences (CIOMS) and the World Health Organization (WHO) have developed their own biomedical-research ethical guidelines named International Ethical Guidelines for Biomedical Research Involving Human Subjects. The Guidelines relate mainly to ethical justification and scientific validity of
research, ethical review, informed consent, research involving vulnerable individuals, equity regarding burdens and benefits, choice of control in clinical trials, confidentiality, compensation for injury, strengthening of national or local capacity for ethical review, and obligations of sponsors to provide health-care services. The publication was revised updated in 1993 and 2002. The 2002 CIOMS Guidelines were designed to be of use to countries in defining national policies on the ethics of biomedical research involving human subjects, applying ethical standards in local circumstances, and establishing or improving ethical review mechanisms. ICH guidelines have been adopted as law in several countries, but are only used as guidance for the U.S. Food and Drug Administration [9, 14].

In 1996, the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) published its own Guideline on GCP [13]. It was designed to ensure that data generated from the clinical trials are mutually acceptable to regulatory authorities in the European Union, Japan, and the United States of America, as well as those of Australia, Canada, the Nordic countries, and the WHO. GCP guidelines include ethical and scientific standards for the design, conduct, recording, and reporting of clinical research involving the participation of human subjects and define the roles and responsibilities of clinical trial investigators, sponsors, monitors, and research subjects. Compliance with GCP provides public assurance that the rights, safety, and well-being of research subjects are protected and respected, in accordance with the principles enunciated in the Declaration of Helsinki and other internationally recognized ethical guidelines. It also ensures the integrity of clinical research data. Currently, this guideline is an international quality standard for clinical trials involving human subjects. Any country that adopts this guideline technically follows the same standard.

In 2001, the Council of Ministers of the European Union adopted a Directive on clinical trials (Directive 2001/20/EC) related to the implementation of GCP in the conduct of clinical trials on medicinal products for human use within the European Union [7]. It was intended to simplify and harmonize the administrative provisions governing clinical trials in the European Community, by establishing a clear, transparent procedure. The Articles of the Directive include guidances on protection of clinical trial subjects, ethics committee, conduct of a clinical trial, guidance concerning reports, and many others. The Member States of the European Union were obliged to adopt and publish the laws, regulations, and administrative provisions necessary to comply with this Directive and to apply them from 1 May 2004.

General Medical Council (GMC) in England has also published guidance for Good Practice in Research (Research: The Role and Responsibilities of Doctors) in 2002 [12]. This guidance sets out the general principles and standards expected of all doctors working in research in the National Health Service, universities, and the private sector in England.

In order to assist national regulatory authorities, sponsors, investigators, and ethics committees in implementing GCP for industry-sponsored, government-sponsored, institution-sponsored, or investigator-initiated clinical research, the WHO issued in 2002, Handbook of Good Clinical Research Practice [24]. The handbook is based on current major international guidelines and is organized as a reference and educational tool to facilitate the understanding and implementation of GCP research process.

At present, regulation of medical research is based on the current international ethical standards as well as on a country’s ethics standard codes.
2.3 Ethical Aspects of Cosmetic Testing

According to EU Cosmetics Directive 76/768/EEC, the cosmetic product must not cause damage to human health when applied under normal or reasonably foreseeable conditions of use. According to the sixth Amendment 93/35/EEC, manufacturer shall for control purposes keep ready information concerned at the assessment of the safety for human health of the ingredients and the finished product as well as proof of the effect claimed for the cosmetic product [5]. In order to achieve these requirements, cosmetic active ingredients and finished products must be tested, including on human volunteers, for evaluation of their safety, compatibility, and efficacy. Studies on cosmetic ingredients and products should be carried out in accordance with the principles of “Declaration of Helsinki” and the guidance for “GCP.” As a rule, safety testing on human volunteers should be preceded by animal or in vitro methods, whereas efficacy testing should be performed when there is evidence that the product does not cause local or systemic adverse responses [11, 19, 20].

Since the past years, there has been a tendency for replacement of animal tests with alternative methods. The seventh Amendment to the Cosmetic Directive establishes a prohibition to test cosmetic ingredients and finished cosmetic products on animals (the testing ban) and a prohibition to market in the EU finished cosmetic products and ingredients included in cosmetic products which were tested on animals (the marketing or sales ban). Both the bans are fully applied from 11 Mar 2009 with the exception of the marketing ban for three types of toxicity tests (repeated-dose toxicity, reproductive toxicity, and toxicokinetics) whose deadline is 11 Mar 2013 [6]. The promotion of scientific and regulatory acceptable alternative methods which reduce, refine, or replace the use of laboratory animals is the main goal of the European Center for the Validation of Alternative Methods (ECVAM) [19].

2.4 Ethical Aspects of Noninvasive Skin Measurements

Due to the developments in bioengineering technology during recent years, it is now possible to evaluate many skin morphology and function parameters by noninvasive instrumental measuring techniques. A “noninvasive” technique means “a procedure or instrument causing minimal and only temporary changes to structure or function, and in particular, not involving pain, incision, or loss of blood” [19]. Skin bioengineering techniques can be successfully applied in safety and efficacy assessment of dermato-cosmetic products to quantify and objectivate the results. They can detect even subtle changes in skin structure and function, and those can enhance the study. Noninvasive skin methods pose no real ethical problems, because they are regarded harmless to the human subject and are not connected with unpleasant or fearful procedures. Since the measurements are only one part of the study, the ethical considerations related to the entire research project are not superfluous. Studies involving noninvasive skin measurements should be conducted according to ethical standards for clinical studies on human subjects [16, 19, 21, 22].
2.5 Essential Ethical Requirements for Performing a Study

Cosmetic testing involving human volunteers must comply with the applicable regulatory requirements for medical research involving human subjects. The basic ethical and scientific principles are provided by the Declaration of Helsinki [25, 26, 29] and current international guidelines for Good Clinical and Research Practices [4, 7, 9, 13, 24] as well as by the guidelines for the evaluation of safety, compatibility, and efficacy of cosmetic products [3, 11], guidelines concerning medical devices [6], and national regulations regarding human studies.

The following principles must be taken into consideration:

2.5.1 Principles Related to Study Conduct

- The study should be preceded by a risk–benefit evaluation, which takes into consideration all study elements (including substances tested and measurement techniques). Concern for the interests of the subject must always prevail over the interests of science and society. The study should be initiated only if the anticipated benefits outweigh the risks.
- The study should conform to a well-designed and scientifically valid methodology according to good practices. The good design would minimize any risks to human beings.
- The design and performance of each procedure should be clearly described in a study protocol, which should be submitted for consideration, comment, guidance, and where appropriate, approval/favorable opinion to an independent institutional review board/Independent ethics committee (IRB/IEC).
- The study should be conducted in accordance with the basic ethical principles, which have their origin in the Declaration of Helsinki.
- The study protocol should always indicate that the ethical principles are observed and an informed consent is obtained.

2.5.2 Principles Related to Study Investigator

- It is the duty of the investigator to protect the life, health, privacy, and dignity of the person on whom biomedical research is being carried out. He must conduct research in an ethical manner and one that accords with the best practice.
- The investigator (research team) should be qualified by education, training, and experience to ensure the proper conduct of the study. He should be thoroughly familiar with the appropriate use of the investigational products and measuring devices as described in the protocol.
• The investigator should inform the participants about all aspects of the study, including its risks and benefits. The information must be written, relevant, understandable, and contain all necessary data to make it possible for human subjects to come to a decision to participate. After ensuring that they have understood the information, the investigator should obtain the subject’s freely given competent informed consent which is documented by means of a written, signed, and dated informed consent form.

• The investigator generally assumes responsibility for obtaining IEC/IRB review of the study protocol.

• The investigator should discontinue the research if in his judgment, it may be harmful to the individual, if continued.

• The investigator should ensure adequate compensation and medical services to subjects in case of adverse events or injury related to the study.

2.5.3 Principles Related to Study Subjects/Participants

• The study subjects must be volunteers and informed participants in the research project.

• The study subjects must be adequately informed of the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, anticipated benefits and potential hazards of the study, and the discomfort it may entail.

• The study subjects should freely give their competent informed consent prior to study participation. They must sign an informed consent document that describes the nature of the study, the products to be tested, known or potential risks, the subject’s rights, and whom to contact in case of problems.

• The study subjects should be selected by strict inclusion/exclusion criteria.

• The subjects should be informed of the right to abstain from participation in the study or to withdraw consent to participate at any time for any reason without reprisal.

• At the completion of the study, participants are entitled to be informed about the outcome of the study and to share any benefits that result from it.

• The study subjects are eligible to receive appropriate compensation for their time and any inconveniences during the study participation which should be described in detail in the IRB protocol at the time of initial review.

• The study subjects could be secured with adequate insurance provided by the sponsor against claims for any trial-related injuries.

2.5.4 Principles Related to Investigational Products

• Investigational products, i.e., cosmetic products tested should be manufactured, handled, and stored according to the Good Manufacturing Practice of Cosmetic Products (GMPC) and the manufacturing specifications [8]. They should be used only in accordance with the approved study protocol.
2.5.5 Principles Related to Measuring Techniques/Devices

- The skin measuring devices should be manufactured and handled according to the current regulations on medicinal devices. They should bear the CE mark to indicate their conformity with the EU consumer safety, health or environmental requirements [6]. The measuring devices should be used only by or under the control of a suitable qualified professional in accordance with the approved study protocol.

2.5.6 Principles Related to Institutional Review Board/Independent Ethics Committee (IRB/IEC)

- The IRB/IEC is an independent body constituted of medical, scientific, and nonscientific members, whose responsibility is to ensure the protection of the rights, safety, and well-being of all human subjects involved in a biomedical research.
- The IRB/IEC should obtain the following documents: study protocol, written informed consent form, subject recruitment procedures (e.g., advertisements), written information to be provided to subjects, Investigator’s Brochure (IB), available safety information, information about payments and compensation available to subjects, the investigator’s current curriculum vitae and/or other documentation evidencing qualifications, and any other documents that the IRB/IEC may need to fulfill its responsibilities.
- The IRB/IEC should review the scientific merit and ethical acceptability of the proposed study and make a statement in writing within a reasonable time.
- The IRB/IEC should conduct further reviews as necessary in the course of the research and monitor the progress of the study.

2.5.7 Other Considerations

Occasionally, the investigator may consider that it is not essential to obtain an approval for his study from the IRB/IEC. This mainly refers to efficacy cosmetic studies, including the use of skin bioengineering methods. If the tested cosmetic product has been on the market for a long time, it should be considered as safe for human health. If the skin measuring device used is noninvasive, it should be considered as harmless to humans. The initial assessment of whether or not the investigational product and measuring technique represent nonsignificant risk is made by the investigator. However, the final decision should be made by the IRB/IEC. When the proposed research involves no more than minimal risk for human subjects, it could be reviewed by the IRB through an expedited review procedure. The requirement for obtaining informed consent from the study subjects always remains obligatory.
2.6 Conclusion

Ethical principles originating in the Declaration of Helsinki and the generally accepted scientific principles of the GCP should be applied to all biomedical studies involving human subjects. Properly designed and well-conducted cosmetic study supported by skin measuring techniques does not generate particular ethical problems. However, all volunteers should provide signed informed consent, and an approval of the study protocol should be obtained from local ethics committees or other authorized institutions. Compliance with these requirements provides public assurance that the rights, safety, and well-being of the participants are protected and that the trial data are credible.

2.7 Key Messages for Performing an Ethical Study

- A favorable risk/benefit assessment
- Scientific study design
- Voluntary and informed consent
- Strict selection of research subjects
- Ethical approval by IRB/IEC
- Compliance with national and international regulations and standards

References


Good Clinical Practice (GCP) is an internationally recognized standard governing the ethical and scientific quality for the design, conduct, recording, and reporting of clinical trials involving human subjects. The main principles are set down by the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Good Clinical Practice Guidance (ICH GCP) [1].

While full compliance with each of the principles of ICH GCP is not mandatory for the conduct of cosmetic studies, the spirit of GCP should be complied with insofar as possible as these studies involve testing in humans. Only then can it be ensured that the rights, safety, and well-being of study subjects are protected, and that the results of clinical studies are credible and accurate.

3.1 Principles of GCP

Section 2 of the ICH GCP Guidance sets down the general principles of GCP, all of which also apply to cosmetic studies. These principles with comments to their applicability to cosmetic studies are as follows:

B. Hughes-Formella
bioskin, Burchardstrasse 17, 20095 Hamburg, Germany
e-mail: betsy.hughes@bioskin.de
1. Clinical trials should be conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki, and that are consistent with GCP and the applicable regulatory requirements.

This principle applies in full to cosmetic studies. The Declaration of Helsinki is a cornerstone describing the ethical principles for the medical community regarding human experimentation [3]. Further, the product being tested must conform with cosmetic regulations of the country/region in which the study is being conducted.

2. Before a trial is initiated, foreseeable risks and inconveniences should be weighed against the anticipated benefit for the individual trial subject and society. A trial should be initiated and conducted only if the anticipated benefits justify the risks.

As stated in the proposed recast of the European Cosmetic Directive [2], a risk-benefit reasoning for cosmetics should not justify a risk to human health. A cosmetic study should be undertaken only if a high degree of safety is expected, regarding both dermal tolerability of the test products and procedures related to the study.

3. The rights, safety, and well-being of the trial subjects are the most important considerations and should prevail over interests of science and society.

This principle is a cardinal rule and applies without exception to any testing involving human subjects.

4. The available nonclinical and clinical information on an investigational product should be adequate to support the proposed clinical trial.

Before conducting a clinical study with a cosmetic product, the investigator should ensure that the test product meets cosmetic regulations and that all ingredients are safe or not banned for use in cosmetics. If previously used in a clinical study, the investigator should be informed of any potential adverse reactions to the test product in these studies.

5. Clinical trials should be scientifically sound, and described in a clear, detailed protocol.

Points to be addressed in a clinical protocol are outlined in the section Design and Protocol Standards (see below).

6. A trial should be conducted in compliance with the protocol that has received prior institutional review board (IRB)/independent ethics committee (IEC) approval/favorable opinion.

It is not compulsory to obtain an IEC vote for a cosmetic study. However, there are certain cases where a vote should be obtained, for example, testing of sensitization potential, or in designs with minimally invasive procedures such as abrasive wounds or suction blisters. An IEC should always be study specific, not design specific, as the ingredients of the test product are critical to the committee’s decision. In certain situations, the IEC may not feel that the use of a certain design is justified for a cosmetic product.

7. The medical care given to and medical decisions made on behalf of subjects should always be the responsibility of a qualified physician or, when appropriate, qualified dentist.

While it is generally not necessary for a physician to conduct medical examinations for a cosmetic study, any medical decisions necessary during a study, for example, treatment of adverse local reactions, should be made by the supervising physician.

8. Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective task(s).
Adequately qualified and trained study staff is a must for the conduct of any clinical study.

9. Freely given informed consent should be obtained from every subject prior to clinical trial participation.
   Coercion of any kind is unacceptable in clinical trial conduct and every subject must have the right to withdraw their consent at any time during a study.

10. All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation, and verification.
    Clinical trial data should be handled in a way allowing reconstruction of the sampling procedures at any time.

11. The confidentiality of records that could identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirements.
    All clinical trial data underlies national data protection laws. Matters pertaining to data confidentiality should be explained in the informed consent forms.

12. Investigational products should be manufactured, handled, and stored in accordance with applicable good manufacturing practice (GMP). They should be used in accordance with the approved protocol.
    In contrast to clinical trial materials for drugs, GMP practices are generally not followed in full for cosmetic test samples. However, the correctness of the ingredients of the test materials should be guaranteed, test products should be adequately labeled to allow identification and correct usage during the cosmetic study, and stability and lack of contamination of the materials should be ensured. It should be possible to reconstruct the manufacturing process at any time.

13. Systems with procedures that assure the quality of every aspect of the trial should be implemented.
    Control measures should be in place which ensure conduct of every study-related task in the same way each time it is completed. These control measures should be available in written form as standard operating procedures (SOPs).

3.2 Standard Operating Procedures (SOPs)

Every test laboratory or institute should have a full set of SOPs governing performance of all routine tasks. ICH defines a SOP as “Detailed, written instructions to achieve uniformity of the performance of a specific function” (ICH GCP 1.55). SOPs are used as a reference and guideline regarding the actions to be taken to complete each research related task and to identify the responsible person for each task. A SOP should exist for each key process, for example, preparing a Study Protocol, obtaining Informed Consent from a potential study participant, establishing study files, recording and correcting data in CRFs, and dispensing test products, among others. Further, a device SOP should exist for all bioengineering devices used in cosmetic studies.
3.3 Study Design and Protocol Standards

A well thought through study design and protocol which lays down all of the key points is essential to the scientific integrity and credibility of data for every clinical study.

In the protocol, the study design should be documented, including subject inclusion and exclusion criteria, primary and secondary endpoints, a description of measures to minimize bias (e.g., randomization, blinding), a description of the test products and their application, compliance measures (e.g., subject diaries, weighing of returned product containers), and criteria for discontinuation/withdrawal of individual subjects or the study. In particular, a detailed plan describing the methods and timing for assessing, recording, and analyzing safety/efficacy measurements is essential. A description of the planned statistics, including a justification of the sample size and the level of significance, should be included in the protocol.

3.4 Conduct Standards

Compliance with the protocol, obtaining informed consent and adherence to data protection regulations, medical management and recording of test-related adverse reactions, correct distribution of test products and instruction in their usage, and accountability and documentation of staff qualification and training are all an integral part of conduct standards.

3.5 Recording and Reporting Standards

All study data must be correctly entered into the case record form (CRF) and subsequently into the study database. Records of test product use must be maintained. All information and relevant correspondence should be gathered in a central study file allowing ready retrieval of the information.

In a cosmetic study, essential documents for the study file, among others, include:

- Signed protocol and amendments
- Sample CRF
- Informed consent form and any other written information given to the subjects
- Any advertisements used for subject recruitment
- IEC correspondence and vote, if applicable
- CV of investigator
- Shipping records of test products, test product accountability forms, documentation of unused test product destruction
- Any available release certificates for test products, at the least confirming that the test products are fit for use in humans
- Master randomization list
- List of treatment codes and information about deblinding or date of breaking treatment code
- Any quality control (monitoring) reports
- If on paper, signed, dated, and completed CRFs including corrections
- Correspondence with sponsor
- Subject identification code list and enrolment log (confidential)
- Signature sheet to document signatures and initials of all persons authorized to make entries or corrections into CRFs
- Record of retained tissue samples (if any)
- Study report

References

3. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. Adopted by the 18th General Assembly, Helsinki, Finland, June 1962, with the most recent amendment by the 59th WMA General Assembly, Seoul, October 2008
Core Message

Dermocosmetic testing combines subjective ratings and objective bioinstrumental assessments. With new developments, a convergence of interest is clearly perceptible among dermatologists, bioengineers, cosmetologists, clinical scientists, physiologists and biologists.

4.1 Introduction

In many instances, subjective clinical assessments are open to bias and interobserver variation. Thus, objective assessments and measurements in dermatology and cosmetology are continuously in search of improvements and additional breakthroughs. Dermometrology and bioengineering have been and remain closely associated in improving both the descriptive and quantitative noninvasive assessments. A few decades ago, ingenious researchers pioneered methods with continuous improvements; they may now look crude, time-consuming, and sometimes lacking fine-tuned reproducibility. These early bioengineering times and the initial phase of dermometrology are over. With more recent progress, the noninvasive technology has made great advances in the design of measuring devices. Such an evolution was

G.E. Piérard, V. Docquier, A. Schreder, S. Neuberg, and C. Piérard-Franchimont
Laboratory of Skin Bioengineering and Imaging, Department of Dermatopathology, University Hospital Sart Tilman, CHU Sart Tilman, 4000 Liège, Belgium
e-mail: gerald.pierard@ulg.ac.be

paralleled by a greater knowledge in many aspects of skin biology. The issue is now evident with the development and refinement of meaningful approaches used in cutaneous biology, physiology, and dermocosmetology. Dermometrology has become a full-fledged science rooted in several nonbiological disciplines. In addition, healthy close working relationships were established among a series of professional societies, research institutes, and private commercial companies. A convergence of interest is clearly perceptible among dermatologists, bioengineers, cosmetologists, clinical scientists, physiologists, and biologists.

4.2 Skin Bioengineering Endeavor

Obviously, collecting accurate clinical data in trials first relies on well-qualified investigators. The clinical presentation must be described in clear objective terms for each grading scale. Any ordinal grading category must correspond to a very distinctive clinical description; otherwise, the grading categories are blurred. The situation must not appear confusing to multiple investigators conducting a multicenter clinical trial in order to avoid great 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 integer 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 bioengineering methods that prove to be increasingly 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 without restriction in human trials. They provide objective and quantitative biophysical information on skin reactions linked to product tolerance and efficacy. They often disclose subclinical effects, predicting the onset of overt skin reaction. The correlation between data and clinical readings affords more specific and in-depth information. Bioinstrumentation provides statistical advantages by showing less variation in the measurements. The procedure of multiparametric testing bypass limitations linked to a single device which provides information about a limited range in a given skin function. One pitfall of bioengineering methods is the absence of control thus introducing inherent methodological bias.

Despite the advantages, it remains 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 the possibilities offered by the astonishing number of measuring devices to be limitless. However, the situation is not so bright. At present, there are only few guidelines addressing standardization of techniques, and only rarely recognized quality control procedures are available for ensuring uniformity of data collection and interpretation. This is probably a flaw inherent to some fast-growing innovative developments.
Although bioinstrumentation has proven its relevance in many instances, it remains that human clinical judgments have historically been preferred by many regulatory agencies for the assessments of drug therapies in dermatology.

4.3 Validation of Methods and Instrumentations

Dermometry covers an ever 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, 7, 10]. In addition, they are used for marketing purposes by drug and cosmetic industries. Some of the dermometry techniques are unique and investigational, while others are present both in a number of research units and at the bedside for monitoring patients. For laypeople lacking specific expertise, there are other commercially available techniques. The typical pitfalls indeed reside in the apparent but deceptive easiness in handling these devices.

Overall, there may be some questionable use of dermometry techniques [1]. Sound methods are possibly subverted for mercantile purposes and falsehoods promoted under the guise of scientific information. The created unsubstantiated claims 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, on standardized ones. Lack of expertise of the owner of the dermometry device and some credulity of the observer are the two most unwise facets of unfulfilled dermometry.

In any case, the device must be calibrated and validated. The study aim and design must be supported by wise concepts. The terminology must conform to 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).

4.4 A Plea for 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 dermometry-supported claims. So far, manufacturers of cosmetics have perceived the great value of the continuously growing field of dermometry. In general, the drug industry and physicians have lagged behind for many years, but now show a similar focus of interest.

Optimization of noninvasive biophysical methods benefits from strictly calibrated, controlled, and standardized procedures of measurements. In spite of relentless developments in relevant links between dermometry, skin physiology, and biology, only a handful of works have been devoted to the standardization of measurements of skin properties. The problem is further clouded by some laypeople in the field of cutaneous biology who are owners of biophysical devices and who may speculate on data. This situation blurs the borderline between claim, dogma, axiom, and facts.
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.

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

Another problem is the clever application of methods to skin-related biomedical problems. The basic knowledge of the biological aspect under evaluation is of paramount

---

**Table 4.1 Instrumentation validation and definitions**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>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</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>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</td>
</tr>
<tr>
<td><strong>Repeatability</strong></td>
<td>Expresses the situation under the same conditions, that is, same operator, same apparatus, short time interval, identical samples</td>
</tr>
<tr>
<td><strong>Reproducibility</strong></td>
<td>Expresses the situation under different conditions, that is, different laboratories, different samples, different operators, different days, different instruments from different manufacturers</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>The interval between the upper and lower levels for which the procedure has been applied</td>
</tr>
<tr>
<td><strong>Linearity</strong></td>
<td>Ability of the procedure (within a given range) to obtain test results directly proportional to true values</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td>Capacity of the procedure to record small variations or differences within the defined range</td>
</tr>
<tr>
<td><strong>Limit of detection</strong></td>
<td>Lowest detectable change</td>
</tr>
<tr>
<td><strong>Limit of quantification</strong></td>
<td>Lowest change above zero that can be quantitatively determined (not only detected) with defined precision and accuracy under the defined experimental conditions</td>
</tr>
<tr>
<td><strong>Ruggedness</strong></td>
<td>Evaluates the effects of small changes in the test procedure on measuring performance</td>
</tr>
</tbody>
</table>

Adapted from reference [26]
Table 4.2 Basic guidelines for bioengineering

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the study purpose and what is the corresponding study endpoint?</td>
<td></td>
</tr>
<tr>
<td>Is the study endpoint quantitative in nature, narrow enough for specific study, and suited to support the study purpose?</td>
<td></td>
</tr>
<tr>
<td>Stratification of endpoints into primary, secondary, tertiary, etc…?</td>
<td></td>
</tr>
<tr>
<td>Shall one or several instruments be used (monoinstrumental or multipronged design)?</td>
<td></td>
</tr>
<tr>
<td>Function of the measurements and the instrument in the study: support, description, exclusion, comparison, validation during study, etc…?</td>
<td></td>
</tr>
<tr>
<td>Which structure or function is actually being measured?</td>
<td></td>
</tr>
<tr>
<td>Range, linearity, and expected change of variables during study?</td>
<td></td>
</tr>
<tr>
<td>When should measurements be performed?</td>
<td></td>
</tr>
<tr>
<td>Interindividual, intraindividual, and intralesional variation, and if possible, variability data from normal and healthy skin?</td>
<td></td>
</tr>
<tr>
<td>Influence of gender, age, phototype, and race?</td>
<td></td>
</tr>
<tr>
<td>Statistical evaluation of the design and the size of the sample studied?</td>
<td></td>
</tr>
<tr>
<td>Studies and literature validating the device(s)</td>
<td></td>
</tr>
<tr>
<td>If the target or measured area is small, do measurements need be repeated to overcome local site variation?</td>
<td></td>
</tr>
<tr>
<td>Existing in-house standards or recommendations, standard operating procedure (SOP)?</td>
<td></td>
</tr>
<tr>
<td>Guidelines and legal requirements, including ethical aspects?</td>
<td></td>
</tr>
<tr>
<td>Output from the instrument and source data. Are they handled and stored safely?</td>
<td></td>
</tr>
<tr>
<td>Are the laboratory facilities up to good standard?</td>
<td></td>
</tr>
<tr>
<td>Is a backup situation prepared if unexpected breakdown occurs in the device or in the laboratory?</td>
<td></td>
</tr>
<tr>
<td>Are ambient conditions such as temperature and humidity under control and expected to remain constant during the study period?</td>
<td></td>
</tr>
<tr>
<td>Needs for preconditioning of study subjects?</td>
<td></td>
</tr>
<tr>
<td>Are experimentalists well educated, trained, and prepared for the specific study?</td>
<td></td>
</tr>
<tr>
<td>Are various types of bias identified and, whenever possible, eliminated?</td>
<td></td>
</tr>
<tr>
<td>How is it ensured or monitored that the study develops as planned, and what are the requirements for constancy and the consequences of inconstancy?</td>
<td></td>
</tr>
<tr>
<td>Calibration, maintenance, and control of instruments before, during, and at end of study?</td>
<td></td>
</tr>
<tr>
<td>Events and circumstances that exclude measurements from being performed or invalidate results?</td>
<td></td>
</tr>
<tr>
<td>How to conclude and report the study?</td>
<td></td>
</tr>
<tr>
<td>Timetable for the study – is it realistic and satisfactory?</td>
<td></td>
</tr>
<tr>
<td>Resources involved – are they available from start to end?</td>
<td></td>
</tr>
<tr>
<td>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?</td>
<td></td>
</tr>
<tr>
<td>Is the study documented and prepared for a special situation if some accusation about fraud would come up?</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from reference [12]
importance because it conditions the choice of the method of investigation. Whenever possible, a combination of methods of evaluation 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 device. The association of a series of distinct methods likely provides a better evaluation of complex interrelationships between cutaneous properties.

The experimental conditions have to be correctly settled, controlled, and monitored. In general, this procedure is applicable to both the tested individuals and the environmental conditions. Several characteristics of panelists have to be chosen and recorded. These include, race, gender, age, region of the body under investigation, and any other specific feature of interest for the study. The seasonal, ovarian, and nycthemeral cycles, as well as diseases and previous interventions such as skin preconditioning 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. Exposure to nonionizing irradiations including ultraviolet light, total sunlight spectrum, and near infrared energy strongly affects specific properties and function of skin, with sometimes a long remanence. It is also obvious that a series of drugs and cosmetics influence many measurements. Therefore, the choice of panelists or patients in a study proves to be very important. Their numbers are also critical [8]. 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 prove to be difficult for the scientists even when oversimplified for commercial strategies. It should always include adequate statistical methods in combination with clever criteria of biological relevance. Investigators should select the statistically validated data with regard to their relevance to the biological aspects. The reverse is also true, and in some instances, increasing the number of measurements helps reaching validation using statistical analysis [8].

4.6 Guidelines in Perspective

Evidence-based guidelines bring the best scientific evidence regarding the diagnosis and management of a particular condition. As such, 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 expertises to evidence-based consensual statements. The procedure is managed by experts, but the guidelines are ideally driven by the available unbiased literature. Such an approach is scientifically sounder, 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 dermocosmetology [12, 17–19, 25]. 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
Table 4.3  Published guidelines for noninvasive measurements

<table>
<thead>
<tr>
<th>Subject of the guideline</th>
<th>First author</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transepidermal water loss</td>
<td>Pinnagoda</td>
<td>1990</td>
<td>[22]</td>
</tr>
<tr>
<td>Skin color and erythema</td>
<td>Fullerton</td>
<td>1996</td>
<td>[6]</td>
</tr>
<tr>
<td>Dry skin and xerosis</td>
<td>Piérard</td>
<td>1996</td>
<td>[13]</td>
</tr>
<tr>
<td>SC hydration</td>
<td>Berardesca</td>
<td>1997</td>
<td>[2]</td>
</tr>
<tr>
<td>SLS irritation test</td>
<td>Tupker</td>
<td>1997</td>
<td>[27]</td>
</tr>
<tr>
<td>Skin color</td>
<td>Piérard</td>
<td>1998</td>
<td>[14]</td>
</tr>
<tr>
<td>Skin topography</td>
<td>Leveque</td>
<td>1999</td>
<td>[9]</td>
</tr>
<tr>
<td>Tensile functional properties part I</td>
<td>Piérard</td>
<td>1999</td>
<td>[15]</td>
</tr>
<tr>
<td>Tensile functional properties part II</td>
<td>Rodrigues</td>
<td>2001</td>
<td>[23]</td>
</tr>
<tr>
<td>Skin greasiness</td>
<td>Piérard</td>
<td>2000</td>
<td>[21]</td>
</tr>
<tr>
<td>Transepidermal water loss</td>
<td>Rogiers</td>
<td>2001</td>
<td>[24]</td>
</tr>
<tr>
<td>Skin microcirculation</td>
<td>Berardesca</td>
<td>2002</td>
<td>[4]</td>
</tr>
<tr>
<td>Antiperspirants and deodorants</td>
<td>Piérard</td>
<td>2003</td>
<td>[18]</td>
</tr>
<tr>
<td>Hair shedding and alopecia</td>
<td>Piérard</td>
<td>2004</td>
<td>[20]</td>
</tr>
</tbody>
</table>

Products (EEMCO) group have substantially worked in that field [16] (see Table 4.3). They produced series of guidances and reviews about noninvasive methods applicable to skin investigations. The 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

Dermometry 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 inexperienced beginners. Even skilful investigators are facing a number of problems related to relevance and interpretation of data. In every instance, emphasis should be placed on a strict respect of controlled and standardized conditions. At present, dermometry is still in a developmental phase where the researcher must control every single aspect of measurements. Devices merely provide figures which are relevant for a given biological aspect of skin but may prove to be distractants only in some instances. The skill of the
researcher and conversely the naive handling by a layperson influence the value of the collected data. In addition, the interpretation of the biophysical measurements requires expertise. The question of whether a license is needed to manipulate biometrological devices as we need one to drive a car arises. Some regulatory procedures should probably be introduced to control claims and creative advertisements deceptively offered under the cover of dermometry.

References

This chapter provides an overview of the processes, the decisions required and the options open to a product development team when creating claims for cosmetic products. We shall highlight the essential steps in thinking through this process by reference to recent work in this field. While most claims in most companies and most tests conducted in testing houses can be very straightforward, the chapter aims to bring out some issues that may be overlooked in the process of creating claims. The approaches suggested here arise predominantly from experience in the European Community, and also have application globally.

### 5.1 Introduction: Cosmetic Claims and Cosmetic Product Testing

Claim substantiation has become an even more important part of product development in the last 10 years for a number of reasons. It is very important to substantiate claims made on cosmetic products to protect consumers from being misled. The adoption of the 6th Amendment to the Cosmetics Directive 76/768/EEC within the European Community placed a legal obligation on marketers of cosmetics to have supporting evidence available for any product claims [29]. Advances in cosmetic science resulting from investment in R&D have led to more ambitious claims as these become a source of greater differentiation for cosmetic companies. However, claims substantiation is not only important from a legal point of view. The cosmetics industry has a keen interest in protecting its customers and meeting their needs since unfulfilled claims can lead to consumer scepticism of not only the culprit product, but of other products sold within this segment.

The process of supporting claims has thus gained its own importance and a number of initiatives addressing this are referred to below. However, one of the major areas of discussion has been how to use science to support product claims. It has to be understood that in the majority of cases, there is a difference between a claim and a test to substantiate the
claim. It is the dialogue between technical development teams and marketing teams that is at the centre of the process of creating a claim: Interpretation of technical terms, creating a rationale that withstands scientific scrutiny and also communicates to a potential customer what they can expect from the product. This is discussed further in Sect. 5.2. This approach has been used by most responsible product developers and marketers of cosmetics and they have worked together with national and international bodies to agree how to go about acquiring the information needed to support the claim. The industry and its regulators need to know what defines sound, relevant, proof of a claim; what evidence, and how extensive does this evidence need to be in order to support a cosmetic claim and what type of evidence is required and to what degree. This chapter aims to capture some of the themes from this work.

5.1.1 What Is a Cosmetic Claim?

A claim on a cosmetic product is defined as any public information on the content, the nature, the effect, the properties, or the efficacy of the product. A claim can be made in words but also images, marks or illustrations [8]. Examples of claim wording includes terms such as, “conditioning”, “breath freshening”, “reduces wrinkles”, “skin feels soft”, “moisturises the skin for up to 24 h”, “SPF 15”. A claim can appear on any material related to the product including the packaging of the product, on a label or insert within the packaging, on the internet, or on the advertisement (be it magazine, TV or on other promotional material) [8]. The product name may also be considered as a claim especially where this infers some product benefit. In Sect. 5.3, the differences between the types of product claim are explored further. The regulations covering substantiating claims are covered elsewhere in this book (Chap. 1.1).

One of the biggest challenges is keeping a claim within the definitions of a cosmetic product, especially not appearing to stray into the area of medicinal claims. A cosmetic is defined in Europe as

“any substance or preparation intended to be placed in contact with the various external parts of the human body (epidermis, hair system, nails, lips and external genital organs) or with the teeth and the mucous membranes of the oral cavity with a view exclusively or mainly to cleaning them, perfuming them, changing their appearance and/or correcting body odours and/or protecting them or keeping them in good condition” [10].

The words we have emphasised here in bold imply that secondary actions may be permitted. Words underlined are the permitted aims of cosmetics in Europe and central to any rationale for claims support. The key parts of the definitions of a medicine on the other hand refer to

“any substance… used or administered… with a view to restoring, correcting or modifying physiological functions by exerting a pharmacological, immunological or metabolic action.” …used in the “treatment, cure, prevention, or diagnosis of disease”…

When creating a cosmetic product claim the first decision point is therefore – is the claim consistent with the definition of a cosmetic?
5.1.2
Borderline Claims

Many factors will influence the view of regulatory bodies on this topic in creating the impression of a medicinal claim: Presentation and marketing, including packaging, where marketed, the target consumer, along with wording and images used. Therefore, the decision on whether a product will be defined as a drug or a cosmetic depends on its primary aim, which in turn can be thought of as why the consumer buys it, having been attracted by the claim. If this aim fulfils one or more of the purposes listed in the definition of a cosmetic then it can be marketed as a cosmetic. The product cannot claim that it can be used in the treatment of disease, but can claim beneficial effects such as “soothes the skin”, “keeps the skin in good condition” or “helps prevent harmful damage caused by the environment”.

Recent advances in understanding skin physiology and the impact of cosmetic products upon this physiology have given rise to a great deal of debate. Definitions of medicinal claims often use the terms “physiological” and “pharmacological” effect interchangeably giving rise to the view that any physiological effect brought about by a cosmetic makes it a medicine. Fortunately, clarification has been recently provided for products marketed in Europe\(^1\). The document not only acknowledges the likely physiological effects of cosmetics but clarifies that for medicinal regulations to apply, “modifying physiological functions by exerting a pharmacological immunological or metabolic action…” “… has to be more than insignificant”.

Products that push the boundaries of cosmetic towards medicinal definitions have been referred to as “cosmeceutical”. This phrase was coined by Kligman [28] and refers to products that are a drug, a cosmetic or a combination of both, and are generally sold under the remit of cosmetics [22]. In spite of its frequent and current use, there is no such entity as a cosmeceutical. Conditions that are on the boundary between healthy skin and disease, such as cellulite, broken veins and spot-prone teenage skin are often the targets of “cosmeceuticals”. This highlights the other important decision in creating a cosmetic claim – is the target skin problem actually a disease. While there are guidelines on this, often, the only way of defining this is to approach the appropriate regulator for an opinion.

\(^1\)http://ec.europa.eu/enterprise/cosmetics/doc/guidance_doc_cosm-medicinal.pdf

5.2
The Process of Creating Cosmetic Claims

A number of factors require consideration when creating a claim for a cosmetic product the primary factors being the intended use and intended benefits that the product will bring to the consumer. These are discussed below and summarised in Table 5.1. The hierarchy proposed also provides a process for recording decisions that can be used as part of the whole claims substantiation package.
<table>
<thead>
<tr>
<th>Decision to be made</th>
<th>Key considerations</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the claim cosmetic?</td>
<td>Cosmetic?</td>
<td>“Cosmeceutical” claims do not exist</td>
</tr>
<tr>
<td></td>
<td>Borderline?</td>
<td>Pushing boundaries towards a medicine may influence the views of the regulatory bodies whether your claim is “cosmetic”</td>
</tr>
<tr>
<td></td>
<td>Medicine?</td>
<td>Ensure that the claim reflects cosmetic benefits – especially changing the appearance</td>
</tr>
<tr>
<td></td>
<td>Target market/consumer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Presentation</td>
<td>Primary action should be cosmetic; secondary actions may be medicinal</td>
</tr>
<tr>
<td></td>
<td>Local regulations</td>
<td>If the mode of action (or claim) borders on medicinal, take advice from local medicinal regulator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The potential for physiological action must be balanced against cosmetic benefit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Presentation of product can infer medicinal benefits</td>
</tr>
<tr>
<td>Where will the product be sold?</td>
<td>What are the local regulations in each of the target markets and does this suggest the different approach to claims support?</td>
<td>Regulatory expectations may differ in different geographies</td>
</tr>
<tr>
<td></td>
<td>Geography/market sector</td>
<td>Claims expectations differ between market and product sectors</td>
</tr>
<tr>
<td>How do you intend to communicate the claims?</td>
<td>Pack</td>
<td>Use of medium to qualify claims – for example, back of pack supports front of pack; footnotes to qualify advertising copy; one click away on website</td>
</tr>
<tr>
<td></td>
<td>Point of sale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Broadcast</td>
<td>What is the overall impression communicated in the different media?</td>
</tr>
<tr>
<td></td>
<td>Print</td>
<td>The environment (point of sale; type of publication); type of sounds/images used to deliver message</td>
</tr>
<tr>
<td></td>
<td>Direct mail</td>
<td>Decide early in the process – re-testing to support advertising claims can be expensive and high risk</td>
</tr>
<tr>
<td></td>
<td>Website</td>
<td></td>
</tr>
<tr>
<td>What is the intended use and benefits of the product?</td>
<td>Performance</td>
<td>How will the consumer know that the claims are fulfilled?</td>
</tr>
<tr>
<td></td>
<td>Ingredient</td>
<td>Build a rationale based on the likely consumer experience</td>
</tr>
<tr>
<td></td>
<td>Aesthetic/sensory</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combination</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparison</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aspirational</td>
<td></td>
</tr>
</tbody>
</table>
5.2.1 Global or Regional

Regulations may vary in different parts of the world with regard to the types of claims permitted and also the amount of support required to substantiate a claim. Therefore, global marketers need to be consciously evaluating how their claim will be perceived in a number of different regions. There are three main areas of cosmetic regulation in the world; the United States, Europe and Japan [26, 29, 35]. Any company wishing to sell their cosmetic product in the US must “have in possession reliable and competent substantiating data of the type and quantity appropriate for the representation” and “substantiating data subject to limitation must be made apparent to the consumer” [17]. Product developers have to decide on the best strategy and there are three basic approaches (and probably variants on these):

- A global claim with separate support packages tailored to the local market needs.
- A global claim with a single support package tailored to the most stringent market needs.
- Regional claims with support packages tailored to the local market needs.

Details of requirements in different regions are covered in more detail in Sect. 2 Chap. 1.

5.2.2 Communication Style and Medium

Though claims are a strong product differentiator, a balance must be established between the amount of information provided to the consumer and the intended benefits; too many claims may lead to confusion or even skepticism; too few claims may lead the consumer to select a competitor’s product. The space available is also important in agreeing product
claims with limitations brought about by factors such as small products (e.g. mascaras) or multi-lingual globally marketed products. An often used tool to substantiate claims is where strong front of pack claims are qualified or further described on back of pack wording where font size and space are less restrictive.

Most attention is focused on pack wording, but product claims will be made available and even developed for the consumer at point of sale and through the various media. Therefore, any promotion or advertising should be considered when deciding what to claim. Consistency of message is an important factor, though the different strengths and needs of each medium will influence decisions on how claims are framed and substantiated.

The impression provided to the consumer in the final message may influence how the claims are supported. Images, sounds and the surrounding environment may change the nature of the claim to be made, over and above the pack wording.

5.2.3 Consumer Factors

It is essential to the marketer that the claims fulfil the consumer’s expectations and not simply from a legal point of view. Meeting consumer needs makes good commercial sense – well-supported claims help to realise sale and potential repeat sales of a product. This is an additional driver in proving product performance. Regulatory bodies require clear, sound and relevant evidence supporting a claim. If the marketer has a rationale describing how the consumer relates to the product claims and evidence to support this, the risk of successful regulatory challenge is reduced.

To create motivating claims, the marketer will have to use particular attributes that differentiate their product in the market and appeal to potential consumers. Any or several features and benefits that differentiate its pack, its formulation, its ingredients, its performance benefits, its texture and even its novelty may be targeted. In doing this, technological terms may require translation into terms more easily understood by the consumer. If so, the marketer needs to create a rationale defining how the technological characteristic and the claim are related.

Pre-existence of a claim can influence decision making – more work may be required to substantiate a totally new claim since it will be hard to evaluate consumer expectation. Consumers will have clearer understanding and expectation of claims that have been in the marketplace for a long period. This factor will also influence decisions later – see Sect. 5.3.2.

5.3 Deciding How to Support a Claim

Having a rationale that is accepted by industry and regulators is one of the most difficult aspects of creating a claim. A recent initiative developed by the United Kingdom industry body (Cosmetics Toiletries & Perfumers Association – CTPA) in association with the UK
Advertising regulatory body (Advertising Standards Authority – ASA) attempted to resolve this [12]. While this addresses a specific regional need, the approach offers a useful guide for approaching claims substantiation in asking the following questions:

1. What is the nature of claim (or consumer message) to be made?
2. How will the supporting information provide proof that the copy fulfils reasonable consumer expectation of that claim?
3. What are the requirements for quality of tests within the supporting information?

Step 1 gives a rationale based on the type of claim that is to be made. Step 2 builds the rationale further by defining the specificity, strength and novelty of claim. Having built the rationale, the final step determines what quality of supporting information would be anticipated. The underlying principle at work here is that each marketer will have their own rationale for translating product technological information into consumer language in order to make claims consistent with their brands and products. The approach also avoids the check list approach used by some regulatory bodies, where a specific test (or tests) is required for a specific claim. Check lists fail to recognise the fact that there is more than one way of defining and many more ways of measuring an attribute. For example, moisturised skin could be defined as plumping up the stratum corneum and/or measurably more hydrated and/or feeling smoother and/or looking more even and/or is measurably more flexible etc.

5.3.1 Classifying Claims

The CTPA/ASA approach divides claims into five classes:

- **Performance Claims**
  Performance claims refer to the effect of using the product on the substrate or how the product changes the substrate appearance, protects it, or maintains the substrate in good condition. This type of claim may also refer to the intensity and/or the mode of action of the product, or the duration of the effect. Examples include “exfoliates the skin for a more radiant complexion” “firms the skin”; “reduces the appearance of wrinkles”; “24 h protection” (e.g. deodorant) and “SPF 20”.

- **Ingredient Claims**
  These can be based on a single or combination of ingredient(s) contained within the product. Ingredient claims may imply that the activity of the ingredient or combination of ingredients makes an additional contribution to the product in which it is incorporated. An example of an ingredient claim is “contains retinol to reduce the appearance of wrinkles”.

- **Sensory/Aesthetic Claims**
  Sensory/aesthetic claims refer to any sensorial attributes experienced by the consumer during product use, including olfactory, tactile or visual effects. It may also refer to a particular aesthetic property of the product. Examples include “leaves the skin feeling soft and smooth” (perception in use); “roll-on applicator” (aesthetic property of product).
• Combination Claims
  Combination claims are some combination of the above claims – for example “spray applicator for a more even, longer-lasting coverage”.

• Comparison Claims
  Comparison claims are where the claimed benefits of one product are compared with those made by another product. Claims such as these must comply with the Misleading and Comparative Advertising Directive [14]. Comparison claims may refer to an older product “New Improved” or a benchmark “More effective than a market leading moisturiser”. It can be seen that there are inherent problems with this type of claim, not least of which is the ever-changing number of comparator products.

There are some claims that are out of scope for definition and support. Aspirational claims, sometimes referred to as “puffery” or “hype” cannot be substantiated as they say nothing about the product. Claims such as “For that Special Feeling” or “Pure Indulgence” would fall into this class. These types of claims are not discussed in this chapter; however, it should be noted that use of this type of claim together with other types (especially performance claims) can change the overall impression communicated to the consumer and thus impact on the level of evidence that might be demanded in support.

5.3.2
Level of Evidence

Once the claim has been classified, the next step is to establish a rationale for the level of evidence anticipated in support, once again consistent with reasonable consumer expectations. The supporting evidence can be acquired from a number of different sources. Table 5.2 outlines an approach to defining the different levels of evidence anticipated in support of claims for a product. The decision process for the product developer here is to assess whether the claim to be substantiated requires any specific testing. This is costly in terms of time and money. It is possible that by changing the claim slightly, the consumer will get the same benefit without the need to test.

5.3.2.1
Established and Widely Available Evidence

Such claims include products whose effect is obvious – for example lipsticks (that colour the lips) or shampoos (that cleanse the hair). Evidence that is widely accepted within scientific and cosmetic industry circles will most likely be based on published reports, publicly available information or product formulation details. Supporting evidence for a product may also be obtained from previous studies (market research, sensory panel, and consumer and/or expert grader evaluations). Moreover, it can be obtained from data on an ingredient or combination of ingredients (relevant to the inclusion level in the final product) or from a previously marketed product (once a correlation can be made between the performances of the products).
### Table 5.2 Levels of evidence anticipated in supporting different types of cosmetic claim

<table>
<thead>
<tr>
<th>Claims widely accepted to be established</th>
<th>Claims with established rationale but requiring product or ingredient-specific evidence</th>
<th>Claims that are based upon a significant advance in science or technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Such a claim would be one whose rationale is consistent with the status of current knowledge commonly accepted within the scientific communities, as applicable to cosmetic science and related disciplines</td>
<td>Such a claim would be one with an underlying rationale consistent with the status of current knowledge commonly accepted within the scientific communities. However, specific support or substantiation might be needed, for example, because the claim is:</td>
<td>The claim will be based upon the advances in the science and technology behind it, not simply that the words used in the claim have not been used before. For example, where the claim is based upon:</td>
</tr>
<tr>
<td>Highly dependent on formulation factor(s)</td>
<td>An entirely new consumer benefit</td>
<td>The action of an entirely new type of ingredient</td>
</tr>
<tr>
<td>For a product type or format not normally associated with the established claim</td>
<td>An entirely new sensory property</td>
<td>An entirely new means of qualifying or quantifying the product’s effects on the substrate</td>
</tr>
<tr>
<td>Dependent on factors described in guidelines, industry recommendations or scientific reviews outlining the need for specific test(s)</td>
<td>An enhancement or quantification of an established claim (for example – intensity of benefit, duration of benefit, a percentage claim)</td>
<td>A new insight into the biology of the body part to which the product is applied</td>
</tr>
<tr>
<td>Targeted to particular populations (for example, hair/skin type, age, ethnic origin)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*From Cosmetics Toiletries & Perfumers Association (CTPA) Guidelines Document [12] – reproduced with permission*
An example of a claim supported by established and widely available evidence would be “2-phase refreshing skin toner”. This claim is classed as a sensory/aesthetic claim and it is based on the widely accepted fact that a toner refreshes the skin by toning it (sensory) and the formulation’s aesthetic appearance (2-phase). Another example of an established claim is “rich intensive moisturiser”. This claim could be based on the nature and inclusion level of known oils or emollients and their contribution to the rich texture of the product.

The key factor here is that these types of claim should match consumer expectation and awareness. It should be remembered that the Cosmetic Regulations require the evidence to be available for all claims, so there is still a need to record why the claim is widely established, if only to show that the thought process has been completed prior to marketing the product.

5.3.2.2
Established Rationale but Requiring Specific Evidence

This type of evidence differs from that described above, in that further evidence in addition to that based on established and widely available evidence is required to prove a specific element of the claim. This additional evidence can come from a number of different sources such as clinical studies or consumer evaluations. The reason for the extra support may be due to formulation differences, requirements from guidelines, enhancement of an established claim etc.

An example of a claim requiring this category of evidence would be a product claiming to have a SPF of 15. This is a performance claim – the effect of using this product protects the skin by a defined amount. A product containing UV filters may have some sun protection properties based on the fact that these filters protect the skin from UVB rays (in published information). However, the current state of knowledge suggests that many factors will influence the absolute effectiveness of such a product claim. Thus in order to claim a measured sun protection factor (e.g. SPF 15), the product must be tested in accordance with the current guidelines (e.g. Colipa guidelines [9]).

The same applies for a product claiming to offer “24 h moisturisation”. A formulation containing ingredients that are known by the industry to moisturise the skin should not require further evidence. However, by emphasising or quantifying the duration of moisturisation, a study would be required to support this enhanced benefit.

Another example of a claim requiring specific evidence would be “SPF 15 moisturiser containing retinol to reduce the appearance of wrinkles”. This is an example of a product making a combination of performance and also ingredient claims. As stated previously, the SPF 15 requires a level of evidence specific to the product on top of the established evidence. In addition, “containing retinol to reduce the appearance of wrinkles” could be based on established reports, publications, concerning the effects of retinol on the skin. The marketer may also choose to assess the consumer perception and even measure the effect of the final product in order to assess the likely performance once marketed.

In summary, each marketer will have their own approach to the specific evidence required based on their own experience with the types of formulation, their knowledge of their own
customer’s expectations of the quantified benefits and their own background of knowledge in the area of science impacting on the product formulation, ingredients and tests used.

5.3.2.3 Evidence Based upon Significant Advance in Science or Technology

Communicating such a claim by reference to a revolutionary formulation, ingredient, or claim providing novel consumer benefit would require evidence describing and supporting the significant advance in science or technology. The level of proof required of claims on such a product may well be more complex than simply gathering evidence to support wording of the claim. It may involve description of a number of specific studies that were designed to test the advances in technology as well as how these contributed to the benefits described for the final product. The evidence of efficacy is generally based on the action the new ingredient(s) and/or full product has on the substrate. New scientific evidence must be available or published to support this type of claim and it should target a new consumer benefit and/or a new target in biology. These types of claims are usually based on a significant amount of research and would be the most challenging in terms of study design, reporting and consumer relevance.

For all levels of evidence, there is the likelihood that supporting evidence will best be provided from a number of complementary studies – at the very least using information from (subjective) consumer studies, (objective) in-use measurements and formulation or ingredient information. In the case of “breakthrough” claims described here, it is highly likely that this “body of evidence” approach is used.

5.4 Generating Information to Support Claims

It is the responsibility of the marketer of cosmetic products to ensure the quality of evidence in supporting a claim. In order to do this, the marketer should have information that is clear and relevant to each claim being made and that is related to the product or ingredient or combination of ingredients in the marketed product. As explained in the previous section, there are a number of different means of obtaining evidence to substantiate claims. These range from basing support on well established information using existing evidence for ingredients, or previous evidence gathered for closely related products, to basing support on more complex methods including time consuming studies. The approach chosen will depend on the claims rationale. In all cases, the basis for supporting the claim should be sound; where experimental studies are used as support, the quality of the study design and the data itself will also be an important consideration, and where no studies are to be performed, quality of the information provided in evidence is still important.

The decision to conduct a test to generate supporting evidence may be based on the need to demonstrate what the consumer can perceive (sensory, expert-grading, consumer studies) and/or define under standardised controlled conditions that can be scientifically
analysed (instrumental studies – in vitro, in vivo, ex vivo). In all cases, realistic or reasonably foreseeable usage conditions are important considerations. A single study or a combination of studies can be used on either ingredient or combination of ingredients or the full formulation to gather the evidence required to substantiate the claim.

The quality of evidence required is often driven by the communication planned; advertised claims often attract a higher risk of challenge from competitors, regulators and consumers; greater consideration of the advertising regulator’s expectations may demand additional testing. For similar reasons, closer scrutiny of borderline claims discussed in Sect. 5.1.2 may demand a different approach to assembling supporting evidence.

The main types of studies to obtain evidence of product efficacy are outlined in Table 5.3 along with some pros and cons that will influence if and how these may be appropriate to the claims required.

The sensory properties of a product are fundamental in cosmetic science and can help in the understanding of consumer perception related to consumer needs and claimed benefits. The evaluation method used will depend on the sensory attributes being examined and the claim required [27]. Trained panels of volunteers with high levels of sensory acuity can define the language and descriptors of key performance attributes of products [16]. Alternatively, naïve panels can provide useful spontaneous responses to product concepts. This approach is covered in more detail in Sect. 4 Chap. 1.12.

Consumer studies are primarily used to mimic likely consumer response for the product to be marketed. Information generated may be used to support product claims based on perceived benefit – such as “skin feels softer”; “wrinkles are less visible”; “skin is more resilient and flexible”. One of the frequent challenges is that the desired claim has not been captured exactly within the question choice. It is also important that these studies use self selecting subjects representative of target consumer. Other important considerations include questionnaire design, layout, avoiding leading questions and ensuring balance in scale of response (see Table 5.4).

Expert grading is considered a more robust form of subjective evaluation and often performed as part of a “clinical study”. Grading can be performed on a variety of characteristics from dry skin to wrinkle severity [19, 21, 25]. Photographs of the area of interest are often taken in combination with expert grading and can be graded blindly and randomly at a later date by an expert grader or in some cases, by the volunteers themselves as further support of the visual performance of the product. Figure 5.1a, b demonstrate the use of this method for assessing the water resistance of a mascara or the crease resistance of eye shadows. The level of expertise must be consistent within any given study and training or other validation essential whether conducted by a dermatologist, ophthalmologist or non-clinical scientist conducting the study.

Instrumental evaluation to assess cosmetic benefit in vivo has continued to grow in importance in assessing skin characteristics such as moisturisation, trans-epidermal water loss, firmness [11, 15, 18], as well as skin colour and gloss effect both in vivo and ex vivo [4, 7, 24]. The technology of these instruments is constantly being updated and improved,

---

2This term is commonly applied to controlled studies conducted in quasi-clinical conditions or by clinicians.
### Table 5.3 Considerations when choosing sources of experimental data to support claims

<table>
<thead>
<tr>
<th>Method</th>
<th>Pros</th>
<th>Cons</th>
<th>Data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensory properties</td>
<td>Trained panels</td>
<td>Does not necessarily equate to consumer “preference”</td>
<td>Statistical power of testing not particularly high</td>
</tr>
<tr>
<td></td>
<td>Valuable especially in concept and prototype testing</td>
<td>Regulator confidence low for this type of test</td>
<td>Principal Component Analysis useful for comparing position of product within competitor set</td>
</tr>
<tr>
<td></td>
<td>Validation essential but accepted methods exist</td>
<td></td>
<td>Simple statistical comparisons possible for given sensory attributes</td>
</tr>
<tr>
<td></td>
<td>Good for comparative properties (e.g., triangle testing), for formulations and ingredients</td>
<td></td>
<td>Interval scales usually indicate use of non-parametric statistics</td>
</tr>
<tr>
<td></td>
<td>Useful for market comparison of texture and thus consumer expectations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Naive panels</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Useful leveller, sense-check on formulation and concept development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumer testing</td>
<td>Valuable for generating “market-place” equivalent</td>
<td>Requires large numbers to achieve significance</td>
<td>Statistical power of testing not particularly high</td>
</tr>
<tr>
<td></td>
<td>Useful for targeting customer or skin/hair type</td>
<td>Distrusted by many regulators</td>
<td>Very large numbers of subjects give greater confidence of outcome once marketed</td>
</tr>
<tr>
<td></td>
<td>Closely reflects market use, so a realistic model</td>
<td>“Controlled” test difficult and costly to organise</td>
<td>Highly dependent on methodology since question order, response scales and mode of acquisition can influence overall outcome</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Halo effect” can influence subjective responses</td>
<td>Use of confidence intervals helps to understand outcome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chi-square for comparative tests possible</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interval scales usually indicate use of non-parametric statistics</td>
</tr>
</tbody>
</table>
### Table 5.3 (continued)

<table>
<thead>
<tr>
<th>Method</th>
<th>Pros</th>
<th>Cons</th>
<th>Data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>In vivo “clinical”/expert</td>
<td>Trained experts to grade visual, tactile, odour changes</td>
<td>Expensive and time consuming Results can vary between centres</td>
<td>Similar issues to sensory and consumer testing.</td>
</tr>
<tr>
<td>assessment</td>
<td>Preferred by many regulators</td>
<td>Design criteria can result in being distant from real in-use conditions</td>
<td>Advantage is that subjective responses should be immune from bias, but large numbers of observations essential</td>
</tr>
<tr>
<td></td>
<td>For wrinkle severity tests, the different wrinkles on the face can be graded e.g. peri-orbital wrinkles, naso-labial wrinkles</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Use of photographic scales for reference enhances validity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grading of “clinical” photos, a powerful method of comparing expert and consumer perception and inclusion in claims support package further strengthens the case</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrumental/biochemical</td>
<td>Useful quantitative and qualitative data when used in in vivo study</td>
<td>Relationship to claim and consumer relevance sometimes difficult</td>
<td>Use of controls and validation of methods can help inform and improve statistical power</td>
</tr>
<tr>
<td>tests</td>
<td>Useful to assess poorly perceived physiological changes</td>
<td>Scientific rationale sometimes questioned</td>
<td>Data manipulation/normalisation can be performed</td>
</tr>
<tr>
<td></td>
<td>Rapid and objective</td>
<td></td>
<td>Comparisons to control helpful</td>
</tr>
<tr>
<td></td>
<td>Validation and standardisation possible</td>
<td></td>
<td>Usually provide continuously distributed data for standard statistical analysis</td>
</tr>
<tr>
<td></td>
<td>Good for screening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In vitro/ex vivo tests</td>
<td>As for instrumental/biochemical tests</td>
<td>Criticised for being removed from reality</td>
<td>As for instrumental/biochemical tests</td>
</tr>
</tbody>
</table>
### Table 5.4 Sample questions that can be used to assess consumer perception

<table>
<thead>
<tr>
<th>Type of Question</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dichotomous questions</strong>: a choice between yes and no</td>
<td>E.g. Does this product cleanse the skin effectively? Yes  No</td>
</tr>
<tr>
<td><strong>Multiple choice questions</strong>: a choice between three or more answers</td>
<td>E.g. Which of the following terms would you use to describe the action of the product you are testing? Refreshing Irritating Toning Revitalising</td>
</tr>
<tr>
<td><strong>Scale-based questions</strong>: Q. Did this product remove all traces of make up?</td>
<td>Q. Did this product remove all traces of make up? 1 2 3 4 5 6 7 8 9 9</td>
</tr>
<tr>
<td><strong>Agreement</strong>: Q. After using this product, my skin feels softer. How much do you agree with this?</td>
<td>Q. After using this product, my skin feels softer. How much do you agree with this? 1 2 3 4 5 6 7 8 9 9 9</td>
</tr>
</tbody>
</table>
further improving the accuracy [6, 30, 37]. The area of skin imaging has exploded in recent years with 2D and 3D image analysis devices and software readily available to quantify features such as size of pores; eye bags; facial wrinkles; scalp hair; cellulite [1, 2, 5, 6, 20, 30, 32, 37]. These approaches provide quantitative and sensitive support for claims such as “reduces wrinkles by 30%”, “10X volume mascara” or “24 h moisturisation”.

This type of approach has been criticised on the basis that changes measured can be too small to be perceptible by the consumer.

In *vitro/*ex *vivo* studies are rarely used alone, but backed up with evidence using an additional method in vivo to reflect the effect of the product on human skin. However, they are useful to support the benefits of product ingredients. In the case of broad-spectrum sun protection in vitro can be used to provide final formulation support [3, 13, 31, 33]. Ex *vivo* studies on final products has also been suggested on hair tresses for claims on styling products or determination of hair damage [23, 34].

The methods briefly described in this section are dealt with in more detail in later chapters. However, with regard to claim substantiation, the method of evaluation used to provide evidence to support a claim must be directly related to what is being claimed on the product when it is being sold in the market.

5.5 Presenting Information in Support of a Claim

Other chapters in this book deal with the organisation and presentation of study data and reports. However, by following the process outlined in this chapter, a marketer would have defined what the product is, where and how it is to be marketed, its uses and benefits, what type of consumer and what expectations they may have. Documentation of this “top-line”...
rationale is an important step in presentation of the evidence. In the majority of cases, more than one source of evidence will be used and it is important to document how these contribute to the claims in question. It is also the norm that more than one claim will be made for a product, and it is essential to highlight how each source of support relates to each specific claim. Simply presenting all the supporting evidence and a list of product claims is unlikely to succeed in convincing a regulator.

Information supporting claims is a legal requirement in the EU, so there needs to be some way of having this accessible. There is no prescribed method or format; however, the approach described above should be a useful starting point. These results usually form part of an information package for the product in question (Product Information File – PIF). This package contains all information related to the product including the formulation, claim rationale, product claims report, study documentation including information on where the study was performed, why it was performed, the protocol, report etc.

In the cases of other markets or other authorities demanding claims support information, it is useful to take their requirements into account at the start of the product development process to ensure that all testing requirements are essential and to forestall the need for expensive re-testing.

5.6
State of the Art

A number of challenges lie ahead in the area of claim substantiation. Advances in the understanding of skin physiology and the real potential of new technologies to change this have led to increasingly complex claims especially in skincare products [36]. As new ways of measuring structural and physiological details of skin evolve, the challenge to make these relevant to consumer experience will continue.

Use of such technological messages in advertising claims for cosmetics needs to be communicated to the consumer without appearing to be medicinal. Some brands occupy this “cosmeceutical” space in the market and this will continue to raise challenges from customers, regulators and competitors alike. A good understanding of customer needs will help here, and the notion of the “averagely well-informed person” as an arbiter of how a claim may be perceived is another useful concept to keep in mind for all concerned.

In understanding consumers more, the focus on specific customer needs will continue to grow; especially the needs of sub-groups. Products that target particular ethnic groups already exist, and the growing markets in developing countries will only enhance this trend.

Another challenge is the growing consumer demand for “natural products”. The ongoing debate of what constitutes a “natural” claim has already given rise to a number of “standards” and there will undoubtedly be continued activity in this area.

3The claim “Natural” could be defined as an aesthetic claim under the scheme described in Sect. 5.3.1., though how “natural” itself is defined, together with its validity, is a point of considerable discussion within the cosmetics industry.
5.7 Conclusion

The process of cosmetic claim substantiation will continue to present challenges. It is important to always consider the consumer’s perception when creating new claims – consumers actually use the products and theirs is the final judgement; this will determine the success of your strategy. Taking the needs of the consumer on board when creating and substantiating a claim eases the process of convincing regulatory authorities. This thinking underpins the approach presented here – creating a rationale for the claim using the building blocks approach to classify and support claims will help overcome challenges in claim substantiation.

Acknowledgments  The authors would like to acknowledge the work of the members of the CTPA Claims Working Group in developing the Building Blocks approach to Claims Support. We are also grateful to Ms Claire Byrne of Oriflame for suggestions, discussions and support in preparing this manuscript.

References

Part II

General Aspects of Cosmetic Testing
The characteristics of the laboratory in which the cosmetic products are clinically tested in human volunteers have a large impact on the overall quality of the results and the overall success of the testing business. On the whole, an unsuitable testing laboratory can make good cosmetic study conduct difficult or even impossible. In this chapter, some basic requirements for a laboratory for clinical testing of cosmetic products are described.

O. Wunderlich
Bioskin, Burchardstrasse 17, 20095, Hamburg, Germany
e-mail: oliver.wunderlich@bioskin.de
6.1 Requirements for Accommodation and Arrangement

The laboratory premises should be situated in an area that can be reached well by volunteers, in order to facilitate quick volunteer recruitment. This is especially important in cases where special volunteer collectives (e.g., with certain skin conditions) are required. For this purpose, the laboratory should be situated in the center or a subcenter of a large city, with good access to public transport and sufficient car parking space. Possibilities for shopping should be within close reach so that volunteers can easily fit in the visits to the lab in their normal life.

In general, the laboratory premises should be large enough to comfortably accommodate research staff and test volunteers. The required minimum size and number of rooms depend on the number and type of studies to be performed in parallel. Insufficient size and number of rooms may result in overfilling, leading to unnecessary stress for staff and volunteers, which might in turn compromise the quality of the procedures.

It is beneficial if the premises are used only for the purpose of clinical testing and are clearly separated from rooms with other functions, for example, in a clinic or physician’s practice. Only in premises dedicated to clinical testing, an optimal arrangement of all necessary parts is usually possible. Nevertheless, it is generally beneficial for daily procedures, as well as for client and volunteer confidence, if the premises in general make the impression of a well-run clinic or practice, with the accompanying cleanliness and tidiness.

The rooms that belong to the testing laboratory should be arranged in a compact manner. This facilitates quick access of both staff and volunteers to the required rooms, making procedures more efficient and less time-consuming.

However, the laboratory rooms that contain confidential information should be clearly separated from those that are to be accessed by volunteers. The former rooms should be protected from trespassing by suitable measures (e.g., locked doors) and should ideally be concentrated in one area for easy access control.

A typical testing laboratory consists of the following types of rooms:

- Reception/volunteer recruitment office
- Data base/subject file room
- Waiting area for volunteers
- Room for conducting general subject information sessions
- Investigator’s office(s)/examination room(s)
- Testing lab room(s)
- Staff office(s)
- Rest rooms for volunteers
- Test product/sample storage room

Following are some general requirements for these individual rooms.
6.1.1 Reception/Volunteer Recruitment Office

The reception room should be easy to find and should be designed to be inviting for potential volunteers. The reception staff, often also functioning as volunteer recruitment staff, should have the possibility of privacy to talk to the volunteer freely. Volunteers should, in general, not be able to overhear the names of other volunteers, or other information that might be confidential. In this room, the appointments are made and volunteers are guided to the respective waiting area or directly to an investigator’s office or testing lab. In case the volunteer recruitment also takes place here, direct (but restricted to the authorized staff!) access to the data base/subject file room is convenient. Furthermore, access to the electronic volunteer database and a system for keeping track of the volunteer appointments should be provided here. The reception office should be positioned in such a way that from this place, access to the entire laboratory can be controlled. Attendance logs should be maintained and kept there for this purpose.

6.1.2 Data Base/Subject File Room

This room contains confidential information, including named personal clinical data. Therefore access to this room must be thoroughly restricted using suitable measures (e.g., locked doors with keys held only by authorized staff members). Since the subject files may contain information that is not kept elsewhere and may be impossible to recover, fireproof cabinets for the subject files are beneficial. The room should be situated near the volunteer recruitment office (see above).

6.1.3 Waiting Area for Volunteers

The waiting area for volunteers should be as comfortable as possible. In some cases, the volunteers have to wait for longer periods of time, and the conditions in the waiting area have an impact on the overall volunteer contentment and willingness to participate. Furthermore, it is necessary for some measurement methods that the volunteers are as relaxed as possible, and a comfortable waiting area helps with that. In this area, volunteers should, at least, be provided with a sufficient number of comfortable chairs and drinking water. A television set, magazines etc. help passing the time, as well as toys for children. The area should be quiet and offer pleasant ambient conditions. Access to and departure from this area should be controlled. The rest rooms should be nearby for convenience. Often, the waiting area also serves as a room for conducting general volunteer information sessions (see below).
6.1.4 Room for Conducting General Volunteer Information Sessions

Here, general volunteer information sessions may be performed. “General” in this case means, for example, information sessions regarding the general stipulations, rights, and obligations in clinical studies (not study-specific) as well as the parts of study-specific volunteer information sessions that are nonconfidential and can be held simultaneously in a number of volunteers (a second, face-to-face, confidential part may then be held with individual volunteers in the investigator’s office). The room should be designed in a way that all volunteers in the room may easily follow the explanations by the staff member. Projectors for displaying a short presentation of the study may be employed for this purpose.

6.1.5 Investigator’s Office(s)/Examination Room(s)

The investigator’s office(s)/examination room(s) serve to accommodate for individual volunteers information about a study given by the investigator, for checking in- and exclusion criteria, for performing physical examinations including basic measurements (e.g., vital signs, body height and weight) and study-specific clinical assessments as well as for any other discussion of individual volunteers with the investigator. For this purpose, the office should offer privacy (to allow undressing as well as communication of confidential information) and should be equipped with the necessary furniture and devices (desk, examination bed, scale etc.). Furthermore, lockable cabinets for subject files that are currently in use are necessary. The room should have adequate lighting to allow skin assessments.

6.1.6 Testing Lab Room(s)

The testing lab room(s) is the core component of every clinical cosmetic testing laboratory. Here, most of the clinical procedures, such as measurements, assessments, and treatments take place. Special care should therefore be taken so that these rooms are adequate for those purposes. The rooms should be spacious and as quiet as possible to minimize stress during procedures. The lab rooms should be versatile, so that different kinds of studies can be performed in the same room. The rooms should be dedicated as testing lab rooms and should not serve secondary purposes (e.g., as additional storage rooms). The rooms should be easily accessible from the waiting area. There are special requirements for the ambient conditions (see below), since some assessments can be influenced by room temperature and humidity.

Tables, chairs, cabinets, and examination beds are usual components of the furniture. The cabinets should be large enough to hold any instrumentation and other supplies that are currently not in use, and should be lockable. For acclimatization or other waiting
periods of the volunteers in the lab rooms, as well as for measurements in the sitting position, comfortable but firm chairs should be provided. The chairs should be easily cleaned. The furniture should be arranged in such a way that measurement instruments can be used properly and comfortably for both the staff and the volunteer. A sufficient number of power outlets for measurement and other devices should be provided. Radio-controlled clocks should be present in each testing lab room, since many assessments are time-critical. Lighting should be adequate to allow safe and correct procedures and consistent visual assessments as well as photography. This includes possibilities for shading daylight. Sufficient privacy should be offered (including window shading), since for some studies, volunteers may be required to (partly) undress. Possibilities for entering volunteer data and/or filling Case Report Forms (desk, computer) should be provided in the lab rooms. For special kinds of assessments, further requirements should be met as needed (e.g., for hot rooms, especially sensitive instruments such as Raman spectroscopes, dedicated photography rooms etc.).

6.1.7
Staff Office(s)

The staff office(s) should be equipped as usual office rooms. Volunteers should have no access to this area, since confidential information is handled here.

6.1.8
Rest Rooms for Volunteers

The rest rooms for volunteers should be close to the waiting area and the testing lab rooms.

6.1.9
Test Product/Sample Storage Room

The test product storage room contains test products and biological samples. No further items should be stored here. Access as well as input and output of test products must be thoroughly restricted (locked doors with keys held by authorized staff members only) and documented. Ambient conditions must be kept (suitable heating and air conditioning) and documented (e.g., by data loggers). Some products and samples must be kept in adequate refrigerators or freezers within specified temperature ranges. For backup reasons, it is wise to have at least two of them each. The temperature of these refrigerators and freezers should be documented/logged and the temperature sensors should be connected to an alarm system, with a quick action plan implemented for cases of malfunctions. Some samples may be required to be stored in liquid nitrogen, so enough space and safety equipment (goggles/gloves) should be provided for that.
6.2
Requirements for Ambient Conditions

For many kinds of cosmetic studies, the ambient conditions (room temperature and humidity) must be kept constant within a certain range (typical: 21±1°C, 50±10% relative humidity). Certain measurement methods, most prominent the measurement of transepidermal water loss, are strongly influenced by the ambient conditions, mainly by the physiological reactions of the volunteers to these conditions (sweating). Therefore well-designed, adjustable air conditioning should be provided at least for the testing lab rooms in order to obtain valid measurement data. The actual temperature and humidity should additionally be logged near the place where the actual measurements are performed. Deviations from the stipulated ranges must be documented.

For some kinds of studies (e.g., so-called kinetic studies with several measurement times per day following a strict time schedule), the area where volunteers wait should also be air-conditioned in order to avoid prolonged acclimatization times.

Ideally, the humidity is also controlled by the air condition systems. In cases where this is not possible, separate humidifiers may be used. Hygienic conditions must be maintained (e.g., by using disinfectants in the water used for humidifying and regular cleaning of air conditioning and humidifying devices).

6.3
Laboratory Quality Management System

In the business of cosmetic clinical testing, quality is paramount. Only study data that is collected correctly under strict quality control in a transparent and reproducible manner and that is fully, correctly, and clearly documented and analyzed is valuable for the client. Therefore, suitable quality management systems must be implemented for all procedures performed in a cosmetic testing laboratory.

6.3.1
General Quality Management

An extensive system of standard operating procedures (SOPs) should be implemented, reflecting all procedures performed at the testing laboratory for a cosmetic study, including planning, performance, and documentation of clinical studies as well as staff training. The individual SOPs should be written by persons who know the “external” requirements (such as regulatory requirements and limitations) and methodical aspects of the procedures as well as specifics of the testing laboratory (such as organizational aspects, qualification of the staff who should follow the SOP, equipment and premises which are available).

It has proved effective to model the SOPs to Good Clinical Practice (GCP) requirements for clinical studies with pharmaceutical products, as far as they are applicable for cosmetic

SOPs from the following main areas are typically needed:

- Generation and maintenance of SOPs
- Quality management procedures
- Generation of study documentation
- Responsibilities of different study staff
- Measurement devices
- Proper handling of personal volunteer data
- Staff qualification and training
- Data management and statistical procedures
- Emergencies
- Handling of study and reference products

The SOP system must be “lived,” in the sense that the SOPs are not only written down, but also actually followed in daily work, and that they are kept under permanent internal and/or external review in order to adapt to the actual procedures performed if required. Internal or external audits help to detect deviations.

Staff training sessions (e.g., for specific measurement methods), both theoretical and practical, must be done regularly and documented in order to implement and maintain quality. Incorrect handling of measurement devices might compromise the complete value of studies performed (see below).

6.3.2 Measurement Devices Quality Management

Most of the measurement devices used in cosmetic studies require regular maintenance or checks (see Sect. 9.4), some of which can be performed only by the manufacturer. The various recommendations of the respective device manufactures in this regard should be followed strictly and should be integrated in specific measurement device SOPs. Deviations from these recommendations should be conservative and well justified. Only devices that are in technical order are able to deliver valid data.

The absolute minimum requirement is that the instruments to be used are checked immediately prior to study. However, a system should be implemented that keeps track of the maintenance/check status of all devices in order to avoid study delays due to time consuming check procedures and unavailability of reserve instruments.

Staff members should be assigned to be responsible for the condition and availability of measurement devices. The respective SOPs and logs documenting the check and maintenance status should be kept close to the individual measuring devices to allow quick reference.

When writing SOPs for specific measurement devices, it is a good idea to take into account any advice that can be obtained from the respective device manufacturer, as well
as any guidelines that have been published by experts in the field of skin physiology measurement (such as the *European Expert Group on Efficacy Measurement of Cosmetics and other Topical Products* (EEMCO) [2–9] or the *Standardization Group of the European Society of Contact Dermatitis* (ESCD) [10–12]) and publications in pertinent journals (such as *Skin Research and Technology*, the journal of the *International Society for Biophysics and Imaging of the Skin* (ISBS), or other reputable dermatology-focused journals).

The handling of the devices must of course be trained. All staff handling the measurement devices must know the basics about the device such as measurement principle, general design of the device, and correct manner of measurement to be able to make valid measurements. Furthermore, expected measurement ranges (where applicable, for different skin types or areas) must be known, in order to be able to detect any implausible measurement results quickly that might occur. That way, good quality measurement results can be obtained.

References

1. Guideline for Good Clinical Practice E6 (R1) (ICH harmonized Tripartite Guideline), CMP/ICH/135/95, July 2002; European Medicines Agency
measurement of skin colour and erythema. A report from the Standardization Group of the European Society of Contact Dermatitis. Contact Dermatitis 46(3), 129–140 (2002)

7.1 Requirements for Qualification and Training of Study Nurses and Investigators

7.1.1 Investigator

The investigator is the lead scientist within a cosmetic study and he/she has to have an overview of the entire study. The investigator has to be qualified by education, training, and experience to assume the responsibility for proper planning, conduct, and completion of a project. The qualification has to be confirmed by providing an up-to-date curriculum vitae (CV) and supporting documents. Preferably, he/she should be experienced in conducting cosmetic studies.

The investigator has a pivotal role in developing the study design. He/she has to supervise the actual study conduct with the aim to successfully complete the study performance in the agreed upon time frame.

Core Messages

- The investigator must have thorough knowledge in testing the safety and efficacy of cosmetic products and requires ongoing training regarding
- General regulations/guidelines
- Scientific issues
- Test methods
- Test devices
- Possible risks (side effects, interactions,...) of test product components etc.

R. Williams
Bioskin GmbH, Burchardstrasse 17, 20095 Hamburg, Germany
e-mail: ragna.williams@bioskin.de
He/she has to be familiar with all study-related Standard Operating Procedures (SOPs). As the investigator is the central contact person for the sponsor, the study team, and the subjects, managerial and interpersonal skills have to be excellent for this key role in coordination. He/she has to ensure that the study team follows the required procedures and quality standards and that the study is conducted in accordance with the main principles of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Good Clinical Practice (ICH-GCP), insofar as reasonable, even though ICH-GCP is not considered mandatory in the field of cosmetics (see Sect. 3.1). To this end, the investigator has to ensure that the study staff is adequately informed about the study protocol and other study documents (subject information, case report forms [CRFs], diaries, etc.), correct application of the test products, the test methods, and devices. The investigator should ensure that all the staff are adequately trained for their study-related functions and duties.

The investigator has to prioritize the subject’s safety and well-being over the sponsor’s interests by following ethical requirements as laid down in the Declaration of Helsinki. Moreover, he/she is responsible for the supervision of the practical study performance according to study protocol and the correct handling of the required study documents (CRFs, subject’s records, etc.). The investigator must have an understanding of cosmetic regulations and guidelines in order to ensure that the test product falls under the cosmetic regulations (e.g., in Europe the current version of the EU cosmetics directive). A basic understanding of the ingredients listed in the International Nomenclature of Cosmetic Ingredients (INCI) declaration or toxicological issues for new ingredients is necessary.

Critical trial-related procedures such as special measurements with bioengineering devices and clinical evaluations (e.g., clinical assessment of skin condition, evaluation of product safety) are generally performed by the investigator or by study personnel trained by the investigator. Therefore, the investigator should be well versed in the reading of skin reactions in safety studies and in the grading of skin conditions in efficacy studies, e.g., grading of dry skin, oily skin, dandruff, or lesion counts for impure skin.

Importantly, since the investigator is responsible for all relevant study-related decisions, a sound understanding of diverse areas is required, e.g., evaluation of critical issues such as adverse event assessments which may require medical background or technical issues with devices.

If an investigator has the license to practice medicine, he/she can furthermore bear the medical responsibility and act as a study physician, if necessary. As mandatory for all physicians, he/she is obliged to ensure continuous medical education.

Physicians and other qualified experts who conduct studies are required to comply with applicable statutes and regulations intended to ensure the integrity of data on which product safety and efficacy are based. They have to protect the rights, confidentiality, data (following European Union Data Protection Directive and other national laws), safety, and welfare of subjects involved in research.
7.1.2
Study Assistant (Study Nurse)

Core Messages

- A study assistant or study nurse is involved in various activities during the conduct of a cosmetic study. Among others he/she is responsible for the following tasks:
  - Recruitment and registration of subjects
  - Organization and coordination of study conduct
  - Correct performance of the study according to study protocol
  - Correct handling of measurement devices
  - Informing, instructing, and supervising of subjects during the study
  - Collection and documentation of study-relevant data
  - Handling and shipment of test products and samples

Furthermore, the study assistant is a contact person for the investigator, sponsor, and subject. Since the study assistant plays a central role in the communication circle and the coordination of the study team, he/she has to have excellent interpersonal skills. The study assistant is also involved in the quality procedures during a study. The study assistant must be able to understand and handle study documents (study protocol, subject information/informed consent forms, CRFs, subject diaries, etc.) as well as skin function measurement devices (e.g., Chromameter [Konica Minolta], Evaporimeter [Servomed], Corneometer [Courage and Khazaka], etc.). Additionally, computer skills are required, and among others, a good knowledge of the English language is important, especially for international studies.

To qualify for work in this area, a medical and/or scientific background (nurse, medical, or scientific degree) and/or comparable education, training in Good Clinical Practice (GCP), and/or licenses/certifications as study assistant/study nurse are required. Additional training is required for the correct use of measurement devices as well as regular training updates regarding quality standards. Familiarity with SOPs is a must.
Having the right person in the right place and at the right time is crucial for the performance of a cosmetic study. Recruitment is a critical activity: the required number of test persons has to be achieved within an appointed time period. Extended and/or inefficient recruitment may have negative scientific and economic consequences. Therefore, successful strategies for raising interest in persons for participation in a cosmetic study are an essential aspect in subject recruitment. Information about perspective studies has to reach the respective volunteer pool. Between several recruitment methods, the best suited recruitment method has to be selected. Motivation of a person to participate in a study is the first step to good recruitment. All recruitment material should follow ethical requirements although review and approval by an ethics committee is not mandatory for cosmetic studies in most countries. However, for cosmetic studies which should have ethics committee approval, for example, tests for investigating sensitization potential or studies with invasive procedures, approval of advertisements by the ethics committee is usually necessary.

Core Messages

- A recruitment tool provides information to potential subjects about the study aim and criteria as well as contact information of the researcher
- A recruitment tool should include at least the following information
  - Purpose of the study
  - Main criteria to determine a volunteer’s eligibility
  - All key points related to study design and conduct should be detailed in the study protocol
  - Address, telephone, and e-mail address of the investigator and/or research facility and person or office to contact for further information
  - More details on study procedures, a summary of in- and exclusion criteria, and special commitments may be included. Participation benefits, if any (e.g., remuneration, a no-cost health examination, etc.) may also be included

R. Williams
Bioskin GmbH, Burchardstrasse 17, 20095 Hamburg, Germany
e-mail: ragna.williams@bioskin.de
Recruitment tools may state that subjects will receive a remuneration, but should not over-emphasize the amount of the monetary compensation. Product claims may be indicated, but should not be misleading and should not promise unreasonable benefits. It has to be pointed out that the test product is investigational.

Common recruitment tools in cosmetic testing are the use of advertisements or recruitment over an existing database by telephone recruitment or mailings. Recruitment through third parties is another option for contacting potential subjects.

Regional, national, or international studies may need different recruitment tools. Advertisements in local newspapers may serve better for regional recruitment and web-based recruitment over a third party may be more suitable for international acquisition.

8.1 Advertisements

Advertisements include newspaper or magazine ads, postcards, press releases, brochures, flyers, postings on the internet (announcements placed on the web or sent by e-mail), posters or info screens in public transport, mailings, and/or media (radio and TV).

Before placing an announcement, it has to be assured that the general information which is provided to prospective subjects to determine their eligibility and interest is sufficient, the announcement is readable, and the visual impact is appealing. A good positioning of the advertisement (“eye-catcher,” special pages of high interest) can also be helpful to reach a large number of people. This should be requested when corresponding with the respective advertiser.

An advertisement is an indirect manner of recruitment. The interest of a potential volunteer has to be raised, otherwise they will not follow-up on the advertisement contact. In subsequent telephone dialog, more specific explanations can follow and an appointment can be arranged for those persons who signal continued interest.

A website intended to be used in recruiting subjects should not be too technical or scientific. Therefore, a separate, study-specific site is useful. Recruitment material has to be clearly presented on the website and be easily understood. The invitation to contact the investigator/facility via e-mail should contain all information concerning confidentiality issues associated with e-mail communication.

Radio spots have to be very concise and should only briefly present the core message. Repeated announcement of the contact address is recommended.

Plasma or LCD displays, so-called info screens, have the advantage that the information is addressed on a large screen to people who are waiting and are often not otherwise occupied. The message usually appears within a program of news, entertainment, and advertisements at highly frequented locations (train stations, airports, etc.) or in public transport vehicles. The information and the contact address should be presented in such a way that the viewer is able to capture the content and to memorize it. To catch the attention of the observer, the spot should be designed appealingly.
8.2 Telephone Recruitment

A telephone call is a more direct way of recruitment. The interest of a person has to be raised during the call; however, a person’s privacy must be respected. The person should already be familiar with the test institute or investigator behind the call. This is not telemarketing!

A script should be available for the initial call. This should, in general, include:

- A check to make sure that the right person is on the line; important to ensure confidentiality
- An introduction identifying the caller and the reason for the call
- A question that allows the prospective subject to opt out of the telephone interview, or set a more convenient time (“Are you interested in hearing more about this study? Is this a good time to talk?”)
- A general description of the study
- Questions that screen for the eligibility of the prospective subjects (prescreening); personal and sensitive questions should be asked in a very careful way and may be accompanied by the statement, “You are free not to answer any questions”
- The arrangement of an appointment for those eligible persons who wish to participate
- A closing that includes a contact name, telephone number, e-mail address

If a message is left for persons who are not available to take the call, this should be kept very short and regard confidentiality.

8.3 Recruitment through Third Parties/Intermediaries

Third parties/intermediaries are persons or organizations who have prior contact with prospective subjects and who can provide a link between the potential subject and the investigator. The essential role is to explain the study in neutral terms and to obtain permission from the prospective subject to release his or her name to the investigator. In the subsequent contact between the investigator and the potential subject, the investigator has to ensure that the contact was made without undue prejudice or pressure.

“Subjects recruit subjects” is a common method for acquisition over an intermediary.

8.4 Acquisition Criteria and Screening

Screening is the next step in the selection process. At this juncture, a dialog takes place between an investigator or his/her representative and a potential subject. Participation in a group information session in which general aspects of cosmetic testing
and study-specific procedures are explained may be the most efficient way to explain specifics about the study. In parallel, the potential subjects should always be given written information which provides all relevant details (study background, objective, study design, test products, restrictions, study procedures, activities and measurements, study schedule, potential risks from the study procedures, data protection, contact, remuneration). Study-specific inclusion and exclusion criteria should be checked and the subject must have sufficient time and opportunity to ask questions in a personal interview with the investigator.

Characteristics which serve as eligibility criteria may include: age, weight, height, body mass index, sex, physical activities, occupation, hobbies, smoking and drinking habits, and skin condition.

The subject gives his/her informed consent in writing on the informed consent form during the screening visit.

Depending on study design, the study procedures and distribution of study products may begin immediately following screening or the subject may be asked to return at an agreed time for the first study measures.

8.5 Product-Related Information and Compliance

Any warnings and precautions should be clearly communicated and preferably be written on the product containers. These may include such warnings as “to be stored out of the reach of children,” “not to be used in the eye region.” The subjects must be advised if it is necessary to protect clothes to prevent staining or other damage. Further, any special storage conditions for the test products must be communicated (e.g., storage away from light, storage at room temperature, storage at 2–8°C).

Compliance to instructions for the application of test products at home and adherence to study procedures and restrictions are an important determinant of the outcome of cosmetic studies. Detailed instructions on the manner and time of application of the test products must be given. For example, if bioengineering methods such as corneometry or transepidermal water loss measurements are planned, it is critical that the time window between last application and measurement is adhered to. Subject diaries may be used as an instrument to monitor compliance. In addition, test product containers should be weighed before distribution and on return as a compliance measure.

Acclimatization before measurements may be necessary, particularly if certain bioengineering methods are used. Subjects must be carefully instructed on the time to return for measurements allowing for an acclimatization period.
9.1 Introduction

To perform efficacy testing of cosmetic and pharmaceutical products on the skin surface, noninvasive technical devices are used to measure objectively a variety of skin parameters. Without such instrumentation, it is almost impossible to determine what the product is really doing on the skin.

Subjective and objective evaluation of the products efficacy, however, is mutually dependant. An objective measurement cannot be performed without a subjective interpretation chart for the measured results; a subjective evaluation of the skin without the suitable measurement technique is not able to achieve accurate results either.

With the example of measuring erythema, the skin’s redness (e.g., in evaluating sunscreen products or claims for sensitive skin), this becomes clearer.
Subjectively, the human brain cannot process slight changes in color, especially when viewed at different points in time. Therefore, instrumentation to record even slightest changes in the redness of the skin is indispensable. The achieved results must then be interpreted in context with the expected outcome or the hypothesis. This can be done only with subjective experience and knowledge of the investigator.

This fact is true not only for the evaluation of the skin color, but for most of the measurement parameters which are of essential importance to cosmetic industry, for example, hydration level of the stratum corneum, sebum level, pH of the skin surface, elasticity and biological skin aging, skin thickness, skin structure, and many more.

Here the person designing and performing a trial cannot rely only on his own subjective evaluation.

Since 1982, the use of technical equipment to execute efficacy testing and claim support in the cosmetic industry has become an indispensable and established standard. It was at that time the first commercially available skin measurement instruments were brought to the market at an affordable price. This trend has contributed to the ability of manufacturers of cosmetics to improve their product quality and also the confidence of the consumer in choosing the right product for her/his specific needs.

In industrialized nations, such tests are not only obligatory because of the cosmetic directives, but have also become routine for the manufacture of products. Independent testing institutes offer their services with regard to such tests and the necessary documentation. This also allows small and midsized manufacturers to have the efficacy of their products documented and certified by testing laboratories.

By having these economic and easy to use devices to test skin parameters in their own laboratory, the small and midsized companies can monitor the efficacy of a product in each step of the development and formulation process. This ensures that when it is sent for final certification, for example, by independent testing institute, it really has the effects the manufacturer wishes to claim for this product. In developing countries, such tests are rather rare since there are no respective directives for development and testing of cosmetic products.

9.2 Technical Assessment of the Various Skin Parameters

Before the first skin testing devices entered the market in 1982, some Universities and R&D laboratories of large multinational cosmetic manufacturers tried to measure certain parameters by using in-house developed measurement techniques. These first attempts of measurement principles were only used where they were developed and remained available only to the respective company. Literature shows a variety of examples of such principles [1, 7, 9, 10]. Now, almost 30 years later, there is a large variety of equipment available for a number of different parameters and it is difficult to mention them all. The most important functions of the skin (epidermis, dermis, and subcutis), hair, and nails to measure objectively are:
Skin barrier function (TEWL)
Stratum Corneum water content (SC hydration)
Surface pH
Mechanical properties and biological age (elasticity)
Color and photo protection (melanin and erythema)
Gloss
Sebum excretion
Epidermal and dermal thickness
Temperature and thermoregulation
Blood flow and microcirculation
Surface morphology (e.g., wrinkles, roughness, scaling)
Digital imaging

These techniques have been described in detail in several scientific books [2, 5].

The development of other, new techniques is moving on quickly. In addition to in vivo Raman spectroscopy, multiphoton tomography, and ESR spectroscopy, there are new oscillating spectrographic methods to assess the penetration into the skin and fluorescent laser scan microscopy for the characterization of skin properties. These new techniques enrich the current range of measuring systems and are subject of specific scientific meetings.

9.3 The Measurement Technique and Norms, Standards and Directives

As described above, there is an extensive range of techniques at the disposal of the interested investigator, offering a solution for many questions.

The decision as to which instrument to use for one’s individual application depends on a variety of different factors. Before discussing them in more detail, it has to be stated that there are no international obligatory directives to help the investigator on which instrument to buy. The cosmetic directives only determine that a manufacturer has to produce proof of the efficacy of their cosmetic products. However, they leave the investigator more or less alone with the decision of how to comply with this directive. There are also no agreed standards of how trials with such measurement techniques have to be performed.

The need for objective measurements to evaluate the efficacy of a cosmetic or pharmaceutical product is widely accepted. However, the clear directives on appropriate use of such techniques are missing.

The results of trials of one group are not necessarily comparable to those of another. Another difficulty lies in the different technical background of the devices themselves. The measurement principles differ considerably depending on the parameter and manufacture of the instrument, thus making comparisons of the results close to impossible or very cumbersome.
This difficulty, which has been noticed for quite some time, needs an international solution or agreement to standardize the techniques to be used and the structure and procedure of a study. The EEMCO-group (European group of efficacy measurement of cosmetics and other topical products) has, from the mid 1990 until 2007, published a variety of so-called guidelines for several different skin measurement parameters (e.g., moisture, pH-, microcirculation, and many others) [3, 4, 6, 8].

The only legally determined procedure is the efficacy testing of sun screen products (see Chap. 18). But even here, different countries go different ways. In Europe, tests for the efficacy of sun screen products are determined by Colipa (European Cosmetic, Toiletry and Perfumery Association). Other countries, for example, USA, Australia, and Japan, follow their own national standards.

9.4
The Practical Use of Measurement Devices

The long-standing investigator’s experience of the successful use of skin measurement techniques is summarized in the following. For optimal use of these technical devices, a minimum of the following factors should be considered:

9.4.1
Factors Within the Laboratory

- Location:
  Sufficient space with the appropriate configuration is crucial. If the available space for making measurements is too small and lacks the required facilities the results may not be usable. The measurement of the skin surface is made in vivo, i.e., on a living organ which constantly reacts to the environment. Controlled relative humidity and air temperature are only two of the factors to consider. Furthermore, sufficient time for acclimatization (normally 20–30 min) should be respected for the volunteers.

- User:
  Even though modern skin testing devices are very easy to handle, their use can only be successful if the investigator or technicians have appropriate training and knowledge of the skin’s functions, skin structure and influences on the skin parameters. In addition they should receive intensive training on the operation of the devices, know what they are measuring and understand the considerations that are necessary before and during the measurement process. If certain factors are not considered appropriately, the results are either falsely interpreted or interpretation is not possible. Such training must be repeated regularly and is as important as the purchase of the instrumentation. Optimal results can only be achieved when all possibilities and limitations of a device are known and subsequently time and money can be saved. Training will need to be repeated should the staff be changed.
• Internal technical service of the instruments:
  Like all other precise measurement tools, skin testing devices must be calibrated regularly. Some manufacturers offer check calibration functions in their units which allow the users to quickly and easily check the accuracy of their device and document it accordingly. Recalibration at the time intervals suggested by the manufacturer is also essential. As soon as inconsistencies in the results occur the manufacturer or the service point should be contacted.

9.4.2
Factors Influencing the Choice of the Most Suitable Measurement Device

The acquisition of skin testing equipment does not differ much from other technical acquisitions. The following should be considered:

• A scientifically acknowledged measurement principle. Even though there are no standards and directives for such devices, the level of awareness suggested by the number of frequently appearing publications should be taken into account.
• Quality and quick service are a prerequisite of smooth operation during trials, which are normally taking place in a limited time period. If service is not adequate the whole trial may be jeopardized.
• Since different manufacturers offer different software packages, with their devices, to collect measuring data and save it for statistical analysis, it is recommended to stick to one manufacturer as far as possible. This offers additional advantages for service (e.g., recalibration) and necessary training.
• The purchase of such equipment should stay within a predetermined budget. There are less expensive techniques widely recognized over many years that may be used, depending on the aims of the trial and the application. For example, the skin moisture measurement provided by the Corneometer® could be used in place of the expensive analysis with a Raman spectroscopy system.

This handbook offers an informative overview of currently available measurement systems, highlight possibilities as well as limitations and factors to consider when using these techniques.

References

## Core Message

- Environment and climate
- Room temperature and relative humidity
- Seasonal variations
- Outside temperature and relative humidity
- Instrument related
- Calibration
- Different models
- New vs. old devices
- Measurements
- Area, position, surface
- Probe
- Subject related
- Age, sex, ethnic group, body site
- Cleansing

### 10.1 Introduction

During the last two decades, an increasing number of noninvasive methods have been developed to determine skin properties in an objective way. The subjective, visual or tactile evaluation of skin conditions can now be quantified and numerical values can be obtained. These techniques are particularly useful in cosmetic testing. These types of methods can potentially detect and quantify some subclinical symptoms. However, standardization...
between instruments is at present imperfect and measuring the same skin property with different instruments can give different results. Instruments of different companies, even based on the same principle, use different scales and the transfer of data between laboratories is difficult. Therefore standardization, including calibration of the devices, becomes a key issue in applying these methods in efficacy testing, skin compatibility, mildness assessment, and in particular, in safety testing. Standardization is necessary on four different levels: environmental factors (room temperature, relative humidity, light sources, air circulation,...), instrumental variables (zero setting, calibration, probe properties, probe position,...), volunteer-linked factors (age, sex, race, anatomical site, diurnal rhythm, skin type, cleansing procedures, skin diseases, medication,...), and product-linked variables (galenic form, dilutions, amount per surface unit, frequency and mode of application, inclusion of blanks,...).

10.2 Sources of Error and Associated Variables

Three categories of factors and sources of variation including instrumental, environmental, and individual (person-linked factors) variables may interfere with measurements. A detailed account of these influencing variables is given in this section.

10.2.1 Instrument-Related Variables

10.2.1.1 Instrumental Variability, Start-Up, and Use

Commercially available instruments must be calibrated according to manufacturer’s guidelines. However, differences in output data may exist, and data obtained by these devices based on the same physical principle may be not directly comparable. A good example is transepidermal water loss (TEWL) measurement. This can be measured directly with the open chamber evaporation gradient method. Therefore, most scientific literature on TEWL is referring to these apparatus [30, 47]. Other devices are based on the closed chamber evaporation gradient methods. Only limited literature is yet available for the last two instruments and their performances should be confirmed [21, 22, 34]. The same is for devices measuring skin color based on single wavelengths or calculating skin color based on CIELAB coordinates. Differences may appear also during the “aging process” of the instrument, and therefore, the same device can give different results along the years.

Thus, the calibration of the instrument, if possible, should be checked frequently.

To enable successful and reliable interlaboratory comparison of results overcoming the effect of the interinstrumental variability, an additional calibration procedure incorporating a calibration for an actual gold reference standard can be adopted [44].
10.2.1.2 Measurements

Surface Area

It is advisable to measure on a horizontal plane to avoid skin curvature. If there is a contact between the device and the skin, the pressure of the probe on the skin surface should be held constant [28, 32, 37]. In order to obtain a constant probe pressure, a built-in spring-system could be used (if not provided by the manufacturer). In some instruments, screens and grids can be used. These elevate the probe, thus also the sensors. This will directly affect readings and one should be careful in comparing data with or without these devices [52]. Measuring area should be defined and, in case of devices with small probes, consecutive adjacent measurements in the same skin area are recommended to reduce standard deviation.

Contact Time

The time the probe is applied to the skin should be as short as possible to avoid occlusive effects which may alter skin surface. In case of TEWL measurement, stabilization of the TEWL value is usually reached by 30–45 s after starting the measuring [4, 8]. Disturbances in the microclimate are immediately detected as a fluctuation in TEWL and can alter other measures such as capacitance, friction, stratum corneum cohesion, etc. [1].

10.2.2 Environment-Related Variables

10.2.2.1 Air Convections

This is the main source of disturbance resulting in rapid fluctuations of the measurements [1]. It is commonly produced by disturbances in the room, such as people moving around, opening and closing doors, breathing across the measurement zone, air conditioners, etc. As these disturbances are difficult to avoid, the use of a covering box was proposed to shield undesirable air turbulence as much as possible [32, 37]. This is particularly indicated for TEWL, capacitance, and other hydration measurements, since air flow on skin surface can change moisture content and skin temperature. On the other hand, a shielding box should have an open top, covered with a cotton cloth in order not to build up occlusivity; such shield might also increase the relative humidity in the space around the measured skin area. Therefore temperature and relative humidity in this box should be recorded.
10.2.2.2 Ambient Air Temperature

The most important effect of the temperature of the air is that it influences the skin temperature both directly (by convection) and indirectly (by central thermoregulatory effects) [2, 45].

A distinction must be made between the temperature of the measuring room and the temperature (climate) where the volunteers live. Therefore an adaptation time of 15–30 min is mandatory.

It was found that fluctuations in the temperature of the measuring room affected SC hydration and TEWL [37]. It is advised to control the room temperature below 22°C. It must, however, be mentioned that a room temperature of 18°C as suggested in some articles seems impossible since test persons complain of having cold and usually refuse to continue the study.

Since ambient air temperature affects TEWL measurement, seasonal variations should be avoided. Even if one works in a temperature-controlled room, it has been observed that the TEWL baseline is not stable. No significant difference between summer and winter baseline TEWL, however, have been described by Agner and Serup [42]. In aged people, however, it has been shown that SC lipid levels of different body sites were depleted in winter in comparison with spring and summer. These changes contribute to the increased susceptibility of aged skin to perturbation of barrier function and xerosis, particularly during the winter months [41]. Sweating in the summer and cold feeling in the winter seems to be obvious problems. As a consequence of this, it is evident that geographical variations also may affect measurements.

10.2.2.3 Ambient Air Humidity

The relationship of the TEWL to ambient air humidity is not linear [28, 32]. The message is that ambient relative humidity is a complex but important variable in TEWL measurements.

It is generally advised to work in a temperature-controlled room with additional control of the relative humidity. The latter should be set close but lower than 50%. The remarks made earlier for seasonal and geographical variations also apply here.

10.2.2.4 Light Sources

Any light source close to the test site affecting the ambient air temperature, the probe temperature, and the temperature of the skin surface of the test persons should be avoided [25].

10.2.2.5 Skin Cleansing

Cleansing of the skin with surface active agents and solvents could modify surface microenvironment [14] due to damage of skin barrier function [26, 37]. Changes in skin
hydration and TEWL could also result from the removal of occlusive substances from the skin surface, like cream [18] and possibly also from the removal of sebum. It is also obvious that exposure of the skin to water-containing products could result in elevated water loss from the surface and SC hydration which might interfere with several biophysical parameters ranging from microtexture, cell cohesion, stratum corneum mechanical properties, and friction coefficient [38]. A rapid superficial cleansing of the skin with water-free ether appears not to increase the water loss values [18]. It is also important to keep in mind that agents in cleansers may deposit on the skin surface and modify its chemical composition leading to errors in reading capacitance measurements or TEWL.

10.2.3 Individual-Related Variables

10.2.3.1 Age, Sex, and Race

Age, sex, and race could be important variables influencing skin function and biophysical measurements as well. Therefore, all these variables should be controlled or standardized when planning a product efficacy study. In particular, studies should be designed within the same ethnic group, age range, and possibly, sex, unless the purpose of the trial itself is to highlight these differences. Aging skin is usually characterized by alteration in water content (uneven distribution), reduction of TEWL due to changes in corneocyte size (larger corneocytes in elderly people), increased microrelief, and loss of mechanical function. Furthermore, during certain “specialized” periods of life significant differences may occur. For example, impaired epidermal barrier properties could be demonstrated in immature infants of less than 30 weeks gestation. A remarkable rate of barrier maturation has been seen during the first few days of postnatal life [17]. In fact, based on the almost identical values for the parameters of TEWL, SC hydration, and pH value, the skin physiology of the child differs very little in SC hydration and barrier function from that of adults [53].

From adulthood through senescence, the age dependence of the cutaneous permeability barrier function has been controversially discussed. Wilhelm and Maibach [12] suggest that there is evidence that baseline TEWL is reduced in aged individuals as compared with midadulthood values. Furthermore, with increasing age, significantly decreased levels of all major barrier lipids have been observed contributing to an increased susceptibility of aged skin to perturbation of the barrier function and xerosis [13, 54]. Between black and white human skin some differences have been reported [5–7]. This is also the case for white and Hispanic subjects [55]. Therefore, in cosmetic testing studies the “ethnic” variable should be carefully controlled.

10.2.3.2 Anatomical Sites

Different anatomical sites differ widely from a physiological point of view, being characterized by different anatomical characteristics. For instance, skin thickness is lower on the volar forearm and higher on the dorsal forearm or trunk and face. Indeed, it must be
emphasized that connective tissue varies tremendously according to the site of the body. In this structural organization, the dermis of the face, scalp, back, forearm, legs, palms, and soles differs greatly from site to site. There are also considerable differences in the relative proportions of each of the connective tissue components and epithelial adnexae in different regions of the body. For instance, the volume of the sebaceous gland is larger inside facial skin. Dermal thickness decreases with age and is greater in men than women. It varies with the body site and is susceptible to endocrine influences. If tests are conducted at the same body site in normal individuals of similar age and sex, then the assumption that there is little variation is just acceptable. However, even under these circumstances, there may still be 5–10% normal variation, and because this may well influence the results, it must be taken into account during interpretation.

The regional variations in TEWL are related to the varying skin structure, particularly, the different lipid fractions between individual locations [36], different corneum thickness between anatomical sites [16], and the regional distribution of eccrine sweat glands, which are concentrated on the palms and soles, face, and upper trunk.

TEWL values of different anatomical sites found in the literature from 1977 to 1988 indicate the following ranking: palm > sole > forehead = postauricular skin = nail = dorsum of hand > forearm = upper arm = thigh = chest = abdomen = back [32]. Indeed, skin at different body sites shows distinct patterns of barrier recovery that are likely to be related to structural and physiological differences [50]. Lipid-rich skin areas (e.g., the forehead) are the most vulnerable to barrier disruption [50]. Measurements are often carried out on the volar forearms. However, TEWL on the dominant forearm might be significantly higher than the one on the nondominant forearm [29, 40], but this has not been found by all investigators [37, 46]. Recently, it has been suggested that for the evaluation of cosmetic formulations, facial skin would be more suitable than the volar forearm [35].

### 10.2.3.3 Sweating

Physical, thermal, and emotional sweating are important variables which need to be controlled [23, 48]. If the ambient air temperature is below 20°C and the skin temperature is below 30°C, thermal sweat gland activity is unlikely, provided the skin is not exposed to forced convection and no excessive body heat is produced (as a result of physical exercise) [3, 31, 33]. Therefore, a premeasurement after a 15–30-min rest in a temperature-controlled room of 20–22°C is, in most studies, taken into consideration. Also physical activity is kept to a minimum. It must, however, be mentioned that it is impossible to control the insensible perspiration.

### 10.2.3.4 Skin Surface Temperature

Skin temperature (and all factors influencing it) is an important preconditioning factor of the test persons and a temperature-controlled room is required. This is of particular importance for those instruments which measure blood flow, skin color (erythema is dependent
on blood supply), and, of course, thermography. Sudden changes in skin temperature (i.e., localized heating or cooling some test areas) are used as “stress test” to induce a vascular reaction to monitor some functional aspects of microcirculation.

10.2.3.5

Skin Damage and Diseases

Cosmetic testing should be avoided in subjects or sites affected by skin disease unless specifically required by the study design. Obviously, skin disease induces tremendous changes in biophysical parameters and all these factors should be carefully standardized and monitored during the study. Skin diseases in which the barrier function is significantly altered, including burns, psoriasis, some ichthyotic disorders, contact dermatitis, and atopic dermatitis, are characterized by increased TEWL, erythema, and blood flow values [4, 9, 25, 38, 49]. Changes in barrier function are caused by chemical contacts [10, 51], surfactant damage [15, 20, 43], or diseased states (dermatitis), for example, psoriasis [11] and eczema [49], resulting in an increased water evaporation rate within the range of 20–60 g/m² h. However, more severe damage to the barrier such as burns [13] gives rise to much higher evaporation rates (above 100 g/m² h) that should be considered in the interpretation of the results.

10.2.3.6

Circadian Rhythm

Fluctuations of some skin parameters such as TEWL, capacitance, blood flow, and pH have been described and have been recently reviewed [27]. Fluctuations of the TEWL may be mainly temperature-dependent. TEWL has been reported to undergo a circadian rhythm and to be higher in the evening and during the night than in the morning [24, 39]. More recently, Le Fur et al. have found a bimodal rhythm for TEWL with two peaks located at 8:00 and 16:00 [24].

10.2.3.7

Intra- and Interindividual Variation

For most skin sites, important interindividual variations occur, usually also dependent on the measuring device [19, 28, 29, 32, 37]. It should be taken into account that some skin sites, including some parts of the forehead, the palm of the hand, and the wrist, should be avoided because of their very high interindividual variability. The intraindividual variation per site is usually smaller [19, 37]. A wide interindividual range of variation is particularly seen when individuals undergo certain treatments. A typical example is TEWL measurements after sodium lauryl sulfate damage of the skin.
Conclusions

Several factors influence skin measurements and biophysical skin testing (Table 10.1). When testing efficacy of cosmetic products in human volunteers, the following conditions must be taken into account:

- If possible, measurements should be carried out in a temperature- and relative humidity-controlled room. Usually, it is suggested to keep the temperature between 20–22 °C and the relative humidity lower than 60%. However, depending on the purpose and design of the study conditions may be different.
- The skin temperature of the test persons should be measured on the test site and the measuring probe should then be warmed up to this particular temperature. This can be done on a part of the skin surface where no measurements are carried out. When test products need to be applied or skin damage must be provoked, relevant controls have to be included, for example, corresponding spots on left and right forearms. Only TEWL values from the same anatomical sites are expected to be comparable.
- Measurement should be carried out in a room with limited air circulation. A shielding box with an open top can be used if doubt exists whether undesirable air turbulence is present or not, particularly when measuring TEWL.
- Hydration and TEWL measurements of a single experiment should, whenever possible, be completed within one season. Measurements during hot summer and freezing winter conditions should be avoided.

<table>
<thead>
<tr>
<th>Table 10.1 Factors influencing measurements in cosmetic testing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables influencing measurements</strong></td>
</tr>
<tr>
<td><strong>Environment and climate</strong></td>
</tr>
<tr>
<td>Room temperature and relative humidity</td>
</tr>
<tr>
<td>Seasonal variations</td>
</tr>
<tr>
<td>Outside temperature and relative humidity</td>
</tr>
<tr>
<td><strong>Instrument related</strong></td>
</tr>
<tr>
<td>Calibration</td>
</tr>
<tr>
<td>Different models</td>
</tr>
<tr>
<td>New vs. old devices</td>
</tr>
<tr>
<td><strong>Measurements</strong></td>
</tr>
<tr>
<td>Area, position, surface</td>
</tr>
<tr>
<td>Probe</td>
</tr>
<tr>
<td><strong>Subject related</strong></td>
</tr>
<tr>
<td>Age, sex, ethnic group, body site</td>
</tr>
<tr>
<td>Cleansing</td>
</tr>
</tbody>
</table>
days should be avoided, with the exception, of course, when the aim of the study asks for this kind of environmental conditions.

- Direct and close light sources should be avoided.
- The measuring surface should be placed in a horizontal plane and the probe should be applied perpendicularly to this surface with a constant but light pressure. Measurements within one experiment should preferably be performed by the same operator.
- Contact measurements should be as short as possible in order to avoid occlusion.
- If skin cleansing is carried out before measurements take place its effect should be investigated.
- Long-term or repeated measurements are preferably done at comparable time periods (e.g., same hour per day, same number of hours after skin cleaning, etc.).

References

Study Design

Betsy Hughes-Formella

Core Messages

- Key factors related to study design are method of randomization, blinding, and controls
- Each study should have a clearly defined study objective which is the focus of the study and takes priority in design issues
- All measurement variables should be directly linked to the study objectives
- Statistical plans should take into account the defined study objectives and include a sample size calculation based on the primary objective
- Accessibility of subjects, potential compliance issues and representativeness must be considered in the definition of inclusion/exclusion criteria and study procedures
- Study documents include a detailed study protocol, subject information sheet and informed consent form, case report forms, and subject diaries, if applicable
- Provisions for data entry and processing, as well as a description of the study data to be entered in the study database, should be defined in SOPs or the study protocol

Only a well-designed clinical study can deliver scientifically sound results. A good study design eliminates as many sources of bias or systemic errors as possible, leading to higher acceptance and validity of study results. In the following section, factors impacting cosmetic study design and their handling are briefly described.

B. Hughes-Formella
bioskin, Burchardstrasse 17, 20095, Hamburg, Germany
e-mail: betsy.hughes@bioskin.de

J.W. Fluhr (ed.), Practical Aspects of Cosmetic Testing,
DOI: 10.1007/978-3-642-05067-1_11, © Springer-Verlag Berlin Heidelberg 2011
11.1
Determination of Study Design

The gold standard of clinical study designs is the randomized controlled study. In this design, the treatments or conditions are allocated to test fields or subjects in a random and unpredictable sequence. Whereas treatment controls such as an untreated test field or marketed comparator are generally the main controls in these designs, other factors may need to be balanced such as age or sex and may also be an integral part of these designs.

In order to reduce the variability due to differences between individuals, designs with intraindividual comparison of test products are generally preferred in cosmetic testing if feasible. Multiple test fields with similar anatomical locations are compared in each test subject, for example, for assessment of a moisturizing effect, multiple test fields on the two forearms may be chosen or for assessment of an antiwrinkle effect, the periorbital area around the left and right eye may be examined separately. Untreated fields or other treatment controls can easily be built into these designs. Importantly, parallel application of test products to the same subject reduces the impact of personal factors such as stress or external factors such as weather or sun exposure. A random allocation of the test products or untreated control should be used in these designs.

Blinding is also a central design issue. Ideally, testing is conducted in a double-blind fashion: Neither the investigator/study staff nor the subject is aware of the identity of the allocation of the test products. However, if this is not feasible, then an observer-blind or single-blind design may be sufficient. In these cases, the assessor or subject is unaware of the test product assignment.

11.2
Study Objective(s)

Every study should have a fully formed, clearly stated research question. This objective(s) should be defined in a manner that allows investigation by quantitative assessment of the endpoints. Inadequately or loosely stated objectives may give rise to skepticism about the study results, owing to concerns that the study conclusions were drawn on the basis of definitions created post-hoc with foreknowledge of the study data [1].

If there are multiple study objectives, these should be defined as primary and secondary. The primary objective should be the focus of the study and take priority in design issues. As a help in defining the primary objective, one might ask if only one question could be answered, what would that question be. Secondary objectives allow for investigation of subsidiary questions which, while being interesting and scientifically arresting, do not have the same priority and are not essential.
11.3 Measurement Variables and Endpoints

The relationship between study objectives, measurement variables, and endpoints is central to statistical design and interpretation of study results. The study endpoint is the outcome that the study is designed to evaluate. The measurement variable is the actual instrumental reading or assessment score that is used as the basis for calculations of the endpoint.

To present an example:

<table>
<thead>
<tr>
<th>Study objective</th>
<th>The evaluation of skin moisturizing effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement variable</td>
<td>Corneometric value at any given time point</td>
</tr>
<tr>
<td>Study endpoint</td>
<td>The change to baseline in the corneometric value at the end of the treatment interval</td>
</tr>
</tbody>
</table>

A maxim is that all measurement variables should be directly linked to the study objectives. A measurement variable means nothing within the context of a study if it is poorly related to the study objective.

Measurement variables and endpoints should be defined for primary and secondary objectives. When planning, it is important to keep in mind that multiple endpoints can yield conflicting results which can prevent meaningful study interpretation. Therefore, it is important to limit the number of objectives and measurement methods to those central to the scientific question at hand.

11.4 Statistics

11.4.1 Sample Size Calculation

The sample size calculation should be based on the primary variable. An estimate of the expected variance and effect size are needed for this calculation. Often, it is possible to take these estimates from historical data which were obtained with similar measurement procedures, since expected effects are often within a similar range. In the event that there are no existing data to draw on, a pilot study should be conducted for this purpose.
11.4.2
Statistical Methods

The selection of statistical methods should primarily take into account the study objective. Other considerations include type of data (e.g., ordinal, cardinal) or type of comparison (paired vs. group comparisons). When applying inferential statistics the significance level should be defined in the protocol. Planned comparisons should be formulated as hypotheses and clearly stated. Unnecessary comparisons should be avoided and there should be a clear prioritization as to the primary and secondary objectives. Statistical procedures to maintain the overall significance level (e.g., adjustment according to Bonferroni) are often omitted in cosmetic testing; however, implications for validity of the interpretation should be kept in mind.

There are a number of available statistical software products which are well suited for cosmetic testing. R, SPSS, Statistica, and SAS are examples, just to name a few. Considering the large number of programs offered also via internet, it must be assured that the statistical software is validated and truly designed for a proper implementation of statistical methods.

11.5
Ethical Review and Scientific Validity

The scientific value of a clinical study is evaluated during the process of ethical review. Even though this is generally not legally required for cosmetic studies, there are often reasons for a review. An evaluation of study design is always a central step in this review process. As stated in Fernando et al. [2], scientifically unsound research in human participants is unethical in that it may expose participants to risks or inconvenience to no purpose. In particular, the following points must be satisfied:

- The research has a clear scientific objective
- The research is designed using accepted principles, methods, and reliable practices
- The research has sufficient power to definitively test the objective with the smallest number of research participants
- A plausible data analysis plan is provided
- The researcher possesses the necessary qualifications, experience, and access to facilities to carry out the proposed study

11.6
Provision for Test Products

Any practical requirements regarding dispensing, storage, or labeling of test products should be taken into account when designing a study. For example, in studies with application of the products at home, scheduled visits for product dispensation and return of containers must be included.
11.7
Provision for Test Persons

The study population is defined over inclusion/exclusion criteria. These criteria should be based on the appropriateness of this population for achieving the study objectives. Accessibility, potential compliance issues, and representativeness must be considered. Even the best scientific design is of no value if suitable and willing test persons cannot be recruited. The inclusion/exclusion criteria must be such that it is possible to recruit the volunteers within a reasonable time frame. Factors such as vacation time and season may be critical. Further, the length of time required per visit or the number of visits to a test institute may need to be modified to be acceptable for the test persons.

11.8
Study Documents

11.8.1
Study Protocol

A study protocol must be available which should describe every step of a study from identification of the study objectives to application of the results. Regarding the study design, the protocol should include background information as to why the study is necessary and how the study results will be used; the objectives and research questions; study design and randomization procedures; inclusion/exclusion criteria and methods of recruitment, sample size calculations; items to be measured and methods of assessment; schedule of procedures with test points (flow chart); data handling and processing; and data analysis procedures.

During the development of the protocol, the following questions regarding study objectives and design should be asked:

- Are the study objectives in line with the research question?
- Does the study design achieve the objectives?
- Are the procedures clearly stated in such a way to avoid false interpretation?
- Is the sample size adequate?
- Will all crucial information be collected?
- How will the study results be used?

11.8.2
Subject Information Sheet

In a subject information sheet with informed consent forms, the test participants are instructed about relevant details of the study design including purpose, randomization, and blinding procedures, criteria for selection and lifestyle restrictions (e.g., restricted use of...
other cosmetics, restrictions on sun-bathing, etc.), what will happen during the study, the number of visits and schedule of procedures (flowchart often helpful), and use of results. Other details regarding voluntary participation, benefits of participation, data protection and confidentiality, who is organizing and funding the study, ethical review, and contact names must also be included in these forms.

11.8.3
Case Report Form and Subject Diaries

A data collection form, the case report form (CRF), is developed for each study in accordance with the design details outlined in the protocol. All data obtained from each test person during participation in a study is collected in this form whereby extraneous data collection should be avoided. Test participants are assigned a unique study/randomization number as identification throughout the study which ensures anonymity of the collected data. The main logistic goal of a study is accurate completion of the CRFs.

It may also be necessary to collect information from test participants outside of scheduled visits, for example, regarding time of application of the test products or other subjective assessments such as itching. A subject diary is developed in accordance with the protocol for this purpose.

11.9
Data Processing

Provisions for data entry and processing should be described in SOPs or in the protocol, including the software/hardware used. In contrast to pharmaceutical studies, in general, it is not necessary to use a data management system with audit trail for cosmetic studies. Entry screens for the study database should accurately reflect the information collected in the CRFs and include fields for all items measured during the study. However, if certain information is collected but not entered into the study database, this should be recorded in the appropriate study documents, for example, the study protocol. It should be stated whether electronic data capture or single or double entry into the database will be used and the processes for data queries to clarify questionable entries in the CRFs should be in place.

References

12.1 Introduction

Cosmetic testing on humans is similar to clinical research involving the participation of human subjects and is submitted to the international guidelines for good clinical research practice (GCRP) [9, 15]. GCRP is a set of ethical and scientific principles for designing, conducting, recording and reporting clinical trials involving humans. These principles should consider and generally apply to all types of research on human subjects including investigation of physiological or pathological process, evaluation of new diagnostic or therapeutic approach, safety and efficacy assessment of medicinal and skin care product. The observance of these principles, wherever applicable, ensures the ethical, scientifically sound and accurate conduct of the research.

According to GCRP the results of each clinical study involving a therapeutic or diagnostic investigational agent or device should be summarized and described in an integrated
full clinical study report [9, 15]. Since the study of cosmetic active ingredient or finished product in human subjects is identical with clinical study, the results obtained should be presented in corresponding cosmetic testing report which consolidates all information about the study background, rationale, objectives, design, protocol, ethical and statistical considerations, results, analysis and conclusions. Both reports are subject to identical principles and international guidelines. The compliance with these guidelines enables the compilation of cosmetic testing report acceptable to the regulatory authorities.

The aim of this chapter is to outline the general requirements for preparing final testing report and publishing the results of cosmetic study.

12.2
Report Structure and Contents

Upon completion of the biomedical research (clinical trial), a final report should be prepared and signed by the principal investigator [9, 15]. The basic principles, structure and contents of the report originated from the guideline for structure and content of clinical study reports adopted by the International Conference on Harmonization (ICH) in 1995 [10]. Generally, the compilation of cosmetic study report should consider all of the topics described in this guidance. Full study report must be prepared for all clinical and human pharmacology investigations that evaluate the efficacy and safety of therapeutic agent. However in some instances, the number, sequence, grouping and data content of topics and appendices may change and adapt depending on the specific nature of the particular study and the preparation of an abbreviated CSR is allowed [1, 12]. This assumption could be applied to the cosmetic testing report, which may be less detailed but enough informative for the regulatory authorities. If there is any question regarding the report, it may be discussed with the reviewing committee.

The following information should appear on cosmetic testing reports [4, 5, 10, 13, 15, 16].

12.2.1
Title Page

The title page should include the following information:

- Title of the study
- Identifying number of the study (if any)
- Name and type of the test active ingredient(s) or finished product(s) tested
- Indication studied
- Name and full address of the sponsor(s)/funder(s) of the study
- Name and full address of the study site/center/department/cosmetic testing laboratory involved in the study
- Name, title, affiliation and full address of the principal investigator or the person(s) responsible for testing
- Study initiation date and study completion date
- Statement indicating whether the study was performed in compliance with Good Clinical Practices (GCRP), including the archiving of essential documents
- Date of the report

12.2.2
Synopsis (Study Summary)

The synopsis is a brief description of the entire study. It should include information about the study title, sponsor, design, duration, center, objectives, participants (main inclusion criteria), test products (active ingredients), methods, instrumentation, measures, protocol, statistics, results with numerical data and conclusion.

12.2.3
Table of Contents

The table of contents should include the page number of each section (and appendix if available).

12.2.4
List of Abbreviations (If Applicable)

A list of the abbreviations and definitions of specialized or unusual terms or measurements units used in the test report should be provided. Abbreviated terms should be spelled out and the abbreviation indicated in parentheses at first appearance in the text.

12.2.5
Ethics

The test report should have a description of the ethical considerations related to the study. It should be stated that the study was conducted in accordance with the ethical principles originated from the Declaration of Helsinki and described how the investigator(s) have obtained informed consent from the participants. It should be indicated if the study was reviewed by an Independent Ethics Committee (IEC)/Institutional Review Board (IRB).

12.2.6
Investigators and Study Administrative Structure

Information about the administrative structure of the study should be provided in the report and corresponding appendix. It should include the name, qualifications (curriculum vitae), institutional affiliation and role of each investigator or other persons (physician, nurse,
laboratory assistant) involved in the study, contract research organization, laboratory facilities, author(s) of the report, biostatistician(s), monitoring and evaluation committees.

12.2.7 Introduction

The introduction should provide background information about the nature of study and rationale. It should include a well documented review of the problem that is the basis of the research, describe all available previous data and specify the reasons for conducting the study in light of current knowledge.

12.2.8 Study Objectives

A statement describing the objective of the test should be provided. The objective(s) should be simple, clear and specific. Besides the primary objective, secondary objectives may be mentioned. In cosmetic studies the primary objective usually covers the main claim (e.g., could be safety or efficacy claim) of the product tested.

12.2.9 Investigational Plan (Material and Methods)

(a) Product tested – This section should provide a brief description of tested product(s) (cosmetic or pharmaceutical product) and reference product(s) (placebo or active control/comparator product) if used, preparation, packaging, blinding, receiving, storage, dispensing and return of tested product, method of assigning participants to treatment groups, product application (quantity, frequency, time, areas), subject compliance monitoring, prior and concomitant skin care products/medicines allowed.

(b) Study participants (volunteers) – The study population (healthy volunteers and occasionally patients) and the selection criteria used to enter the participants into the study should be described. The information should include the number and demographic data (age, sex, race), subject recruitment manner, criteria for inclusion, exclusion and early withdrawal and specific criteria linked to the study applied, training of the participants.

(c) Study design and plan – The design and plan of the study should be clearly described. They must be scientifically appropriate and suitable to prove study objectives. The information provided should include:

- General design – the type of the study (method of blinding/masking, control/comparison groups, method of assignment to treatment groups).
- Primary and secondary study endpoints (efficacy and safety endpoints).
- Test schedule – sequence and duration of all study periods (study timetable).
(d) Evaluation parameters (Test variables) – This section should describe the specific variables (e.g., efficacy or safety criteria adopted, undesirable and side effects) to be assessed.

(e) Evaluation methods and equipment – This section should describe the specific methods used for assessment the evaluation parameters:

- Use tests by consumers (method and format)
- Auto-evaluation by users themselves (questionnaires and visual analog scales)
- Sensorial evaluation tests used by trained panelists (notation method and types of scales)
- Scoring done by a suitably qualified health or professional expert (visual, tactile, or other sensorial scores and scales)
- Ex vivo/in vitro tests (the substrate/reagents and methodology)
- Instrumental non-invasive methods and devices (the equipment, usage conditions, operations and measured skin structure and function variables)
- The means of recording, rating and reporting of unanticipated problems and adverse events should be described (checklist, questioning).

(f) Study protocol – The study procedures accomplished at each visit should be described.

(g) Statistics – The information about the statistical analysis applied should describe the sample size determination, variables analyzed, comparison grouping, statistical tests and data processing used.

(h) Data quality assurance – Systems with procedures that assure the control and quality of every aspect of the trial should be described.

(i) Changes in the conduct of the study or planned analyzes – Any change in the conduct of the study or planned analyzes done after the start of the study should be described.

12.2.10

Results

This section should present the study results, their statistical significance and interpretation. The important demographic, efficacy, safety and other results should be presented in summary tables, graphs and figures in the text of report. Individual data and very large tables of raw data should be included in an appendix.

12.2.11

Discussion and Conclusions

In this section the main results of the study should be briefly summarized and their importance and practical benefit should be discussed in the light of other existing data. The discussion and conclusion should identify if the results support the objective (original hypothesis) of the study, are they consistent with those reported by other investigators,
what are the possible explanations in case of unexpected results and if any potential limitations in the study exist, which require further investigations.

12.2.12 Reference List

This is a list of articles from the literature that are actually cited in the report. References should be given in accordance with the internationally accepted standards of the 1979 Vancouver Declaration on “Uniform Requirements for Manuscripts Submitted to Biomedical Journals” [8].

12.2.13 Signatures

The test report should be dated and signed by sponsor, investigator(s) and author of the test report.

12.2.14 Appendices (Attachments)

All relevant documents associated with the management and detailed clarification of the study should be included as appendices. The section should start with a full list of all appendices available for the study report. Some of them could be submitted with the report while the rest may be provided only on request.

12.3 How to Publish the Results

Cosmetic tests generate important information about the skin care products acceptance, safety and efficacy. After a cosmetic study is completed, the investigators normally would like to present their results to other scientists and the community by oral or poster presentation in scientific meetings and by publishing in popular and peer-reviewed journals. Nowadays, the internet gives an opportunity to a revolutionary spreading of scientific information [6, 7].

Publication policy regarding the dissemination of cosmetic study results to the participants, the community and in the scientific media should be specified in the study protocol. It should be clearly indicated who holds the primary responsibility for the publication and who will take the lead in publication. There is a need to first obtain approval from the primary responsible party before any information can be used or passed on to a third party [13, 16].

Both authors and publishers have ethical obligations in publication of the research results. The authors must accurately report and correctly interpret the results. Articles
based on studies performed not in accordance with the principles of the Declaration of Helsinki should not be accepted for publication [15].

Test reports and scientific papers follow basic rules of good writing and research. However, the test report commonly requires less background information and less detailed discussion section. The professional study reports increasingly approach to scientific publication [5].

There are many guidelines for preparation and submission of scientific papers to biomedical journals [2, 3, 8, 14]. The most used guideline is the “Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication” created by the International Committee of Medical Journal Editors (ICMJE) in 1979 and last updated in 2008 [8]. The ICMJE requirements help authors and editors in creating and distributing accurate, clear, easily accessible reports of biomedical studies.

Biomedical journal articles and scientific articles based on cosmetic testing in particular adhere to a basic standard format and usually include the following elements: abstract, introduction, aim(s), material and methods, results, discussion and conclusions. Manuscripts submitted for publication should be composed of the following sections [6–8, 14]

12.3.1 Title Page

The title page should contain:

- Article title, which should be concise but enough informative about the contents of the paper in order to facilitate the electronic search of the article
- Authors’ names, highest academic degree(s) and institutional affiliations
- Name of department(s) and institution(s) to which the work should be attributed
- Name and full address of the authors responsible for correspondence and reprints request
- Source(s) of support in the form of grants, equipment, drugs, or all of these
- A short running head
- Word counts for the text only (excluding abstract, acknowledgments, figure legends, and references)
- The number of figures and tables

12.3.2 Conflict of Interest Notification Page

The authors should always declare any conflicts of interest that may arise. This information should be included on a separate page.

12.3.3 Abstract (Summary)

The abstract must accurately reflect the content of the article. It should summarize the background, the objective(s), the design and methods (selection of study subjects, observational
and analytical methods), the main results and their statistical significance, and the conclusions. Requirements for length and format of the abstract may vary from journal to journal.

12.3.4

Keywords

Below the abstract, 3–5 keywords or short phrases should be provided.

12.3.5

Introduction

This section provides background information about the topic of the manuscript (the nature and significance of the problem, a brief review of what is known to date, the rationale for the study) and a statement of main and secondary study objectives (or hypothesis tested).

12.3.6

Material and Methods

This section should provide enough information to enable another investigator to replicate the study. It consists of several subsections:

(a) Participants (subjects) – describe the study population (healthy volunteers or patients, including controls), their demographic characteristics (age, sex, race or ethnicity) and selection (inclusion and exclusion criteria). When reporting experiments on animals their selection and characteristics should be described

(b) Products tested – describe test products (manufacturer’s name and address, active ingredients)

(c) Methods – identify the assessment methods and devices (manufacturer’s name and address)

(d) Protocol – identify precisely the procedures (preconditioning of test subject, environmental conditions, application of test product, localization of test site, performing of assessment and measurement)

(e) Ethics – indicate that the procedures involving human subjects adhered to the ethical standards of the responsible committee on human experimentation (institutional or regional) and to the Declaration of Helsinki. When reporting experiments on animals, authors should indicate whether the institutional or national guidelines or laws on the care and use of laboratory animals were followed

(f) Statistics – describe in detail the methods used for statistical analysis, randomization and blinding of observations. Specify the computer software used

12.3.7

Results

This section is the core of the paper. It gives a detailed description of the data collected by the researchers and their statistical significance. The data must be presented clearly and in logical sequence in the text, tables and illustrations.
12.3.8 Discussion and Conclusions

This section usually begins by summarizing briefly the main findings. Then the authors interpret the results, discuss whether they are new, unique, similar or different to previously relevant studies, emphasize the new and important aspects, state whether any potential limitations of the study exist and make conclusions and suggestions for further research.

12.3.9 Acknowledgments

This section is optional. All people or institutions who contributed to the completion the study and preparation of the manuscript (administrative, technical, intellectual help, writing assistance, financial and material support) should be listed after obtaining written permission from them.

12.3.10 References

References indicate the original research sources considered in the article. They should be presented and ordered according to the Uniform Requirements style for references [11] and specific requirements of the selected journal.

12.3.11 Tables and Figures

Tables and figures help in more expressive and efficient presentation of the study findings and make it possible to reduce the length of the text. They should fit the requirements of the particular journal.

12.3.12 Tables and Figures Legends

The legends for illustrations should be given on a separate page.

12.3.13 Sending the Manuscript to the Journal

Most journals now accept electronic submission of manuscripts by downloading directly onto the journal’s website. For specific instructions on electronic submission, authors should consult the journal’s Instructions for Authors. If a paper version of the manuscript
is submitted, the required number of copies of the manuscript and figures must be sent to the editorial office [8].

All manuscripts must be accompanied by a cover letter signed by the corresponding author that includes the following information: the name and full address of the corresponding author, a statement that the manuscript has been read and approved by all the authors, a statement about any conflicts of interest and other specific information that is required by the journal. A transfer of copyright may be required after the paper has been accepted for publication. Many journals now provide a presubmission checklist to help the author in the manuscript submission [6, 8] (Table 12.1).

<p>| Table 12.1 Key messages for preparing cosmetic testing report and scientific article |
|-----------------------------------------------|-----------------------------------------------|
| <strong>Cosmetic testing report</strong> |
| Title page | Title page |
| Title of the study | Title of the article |
| Identifying number of the study | Author(s) name and their affiliations |
| Sponsor(s) | Name of department(s) and institution(s) |
| Study site/center/department/laboratory | Name of corresponding author |
| Investigator(s) | Source(s) of support |
| Test product(s) | A short running head |
| Study date | Word counts for the text |
| Date of report | Tables and figures number |
| Conflict of interest notification page |
| <strong>Scientific article</strong> |
| Title page |
| Synopsis (study summary) | Abstract (summary) |
| Study title | Background and aims |
| Study sponsor | Material and methods |
| Study design | Results |
| Study duration | Discussion and conclusions |
| Study center |
| Objectives |
| Participants (main inclusion criteria) |
| Test products (active ingredients) |
| Methods and instrumentation |
| Measures |
| Protocol |
| Statistical methods |
| Results with numerical data |
| Conclusion |</p>
<table>
<thead>
<tr>
<th>Table 12.1 (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cosmetic testing report</strong></td>
</tr>
<tr>
<td><strong>Keywords</strong></td>
</tr>
<tr>
<td><strong>Table of contents</strong></td>
</tr>
<tr>
<td><strong>List of abbreviations</strong></td>
</tr>
<tr>
<td><strong>Ethics</strong></td>
</tr>
<tr>
<td><strong>Investigators and study administrative structure</strong></td>
</tr>
<tr>
<td><strong>Introduction and objectives</strong></td>
</tr>
<tr>
<td><strong>Investigational plan (material and methods)</strong></td>
</tr>
<tr>
<td><strong>Participants (volunteers)</strong></td>
</tr>
<tr>
<td><strong>Product(s) tested</strong></td>
</tr>
<tr>
<td><strong>Study design and plan</strong></td>
</tr>
<tr>
<td><strong>Evaluation parameters (test variables)</strong></td>
</tr>
<tr>
<td><strong>Evaluation methods and equipment</strong></td>
</tr>
<tr>
<td><strong>Study protocol</strong></td>
</tr>
<tr>
<td><strong>Statistics</strong></td>
</tr>
<tr>
<td><strong>Data quality assurance</strong></td>
</tr>
<tr>
<td><strong>Changes in the conduct of the study or planned analyzes</strong></td>
</tr>
<tr>
<td><strong>Results</strong></td>
</tr>
<tr>
<td><strong>Discussion and conclusions</strong></td>
</tr>
<tr>
<td><strong>Acknowledgements</strong></td>
</tr>
<tr>
<td><strong>Reference list</strong></td>
</tr>
<tr>
<td><strong>Signatures of sponsor, investigator(s) and author of the test report</strong></td>
</tr>
<tr>
<td><strong>Appendices (attachments)</strong></td>
</tr>
<tr>
<td><strong>Tables and figures legends</strong></td>
</tr>
<tr>
<td><strong>Units of measurement</strong></td>
</tr>
<tr>
<td><strong>Abbreviations and symbols</strong></td>
</tr>
<tr>
<td><strong>Cover letter</strong></td>
</tr>
<tr>
<td><strong>Statements about any possible conflicts of interest</strong></td>
</tr>
<tr>
<td><strong>Name, address and signature of corresponding author</strong></td>
</tr>
</tbody>
</table>
12.4

Conclusions

Cosmetic testing report and resulted scientific publication should be complete, well organized and easy to review. They should provide a clear explanation and enough information on the rationale, objectives, plan, methods and conduct of the study so that there is no ambiguity in how it was carried out. The compliance with common guidelines enables the compilation of study report and scientific article of any research study acceptable to the regulatory authorities and publishers.

References


Part III

Practical Aspects of Testing: Typical Examples of Test Settings
Core Messages

- The formation of skin hydration is a complex and multifactor process including the natural sources of skin moisturization and the effect of exogenously applied substances on the skin.
- An objective evidence for the claimed effect of a moisturizer/emollient is required.
- A variety of non-invasive methods for the evaluation of skin hydration exist, however none discloses the complete interactions between a moisturizer and the skin.
- A multiparametric approach is useful in the assessment of moisturizers efficacy.
- Evaluation of the cutaneous electrical properties (capacitance, resistance, impedance) is the most commonly used method in proving the efficacy of moisturizers.
- The choice of the proper device should be based on considering its technical bases, the assessed parameters of skin hydration, and the objective of the study.
- Study design, population, anatomical test site, and procedures should be in accordance with the study aims.
- Controlling subject-, instrument-, and environment-related variables is a key issue in performing skin physiology study.

---

R. Darlenski
Department of Dermatology and Venereology, Medical University, Sofia, Bulgaria
e-mail: darlenski@abv.bg

J.W. Fluhr
Department of Dermatology and Allergology Center of Experimental and Cutaneous Physiology Charitéplatz 1, Charité University Clinic, 10117 Berlin, Germany
e-mail: joachim.fluhr@gmx.net

J.W. Fluhr (ed.), Practical Aspects of Cosmetic Testing,
DOI: 10.1007/978-3-642-05067-1_13, © Springer-Verlag Berlin Heidelberg 2011
13.1
Introduction

13.1.1
Moisturizers and Emollients – What Are They?

Moisturizers and emollients are among the most commonly used cosmetic products in the everyday life [11]. However, no unified definition of “moisturizer” exists. In a broad sense, a moisturizer is a product that increases the water content of the skin, although the use of this term is rather related to marketing instead of having a scientific background. On the other hand, emollients are topical products which have a softening and soothing effect. Emollients, together with the humectants (that increase the water content and attract water to the epidermis), and the occlusives (decreasing the evaporation of water from the skin surface) are considered as the three key structural components of moisturizers. The main characteristics of an ideal moisturizer have been proposed [13]:

- Efficacy – raises the hydrating the stratum corneum (SC) and decreases transepidermal water loss (TEWL)
- Smoothens and softens skin (acts as an emollient)
- Helps for the restoration of the lipid barrier, and enhances the skin’s natural moisture retention mechanisms
- Cosmetically elegant and acceptable
- Rapid absorption providing immediate hydration and assuring long-lasting effect
- Properties related to sensitive/irritable/allergic skin: hypo-allergenic, non-sensitizing, fragrance free, non-comedogenic
- Affordable price

13.1.2
Natural Sources for the Skin Hydration

The epidermis and in particular its superficial layer, the SC, accomplishes approximately 90% of the skin barrier function. The epidermal barrier protects the human body against many external stressors: physical stress (e.g., mechanical, thermal, radiation), chemical stress (tensides, solvents, topical xenobiotics), and environmental conditions; as well as it prevents the organism from loss of essential components such as ions, water and serum proteins. The SC with its structural and functional components (Fig. 13.1) is responsible for the retention of water and the hydration balance in the superficial skin layers. Several mechanisms maintain the epidermal barrier-related functions, potentially contributing to cutaneous hydration state: (1) the unique “bricks and mortar” organization of the SC with corneocytes and the cornified envelope (bricks) as well as the adjacent intercellular bila-mellar lipids (mortar); (2) the natural moisturizing factors (NMF) – a highly hygroscopic complex of free amino acids (mainly derived from the enzymatic degradation of filaggrin), salts, urea and other molecules; (3) endogenous glycerol, derived through the aquaporin-3
pathway, or synthesized in the pilosebaceous unit; (4) calcium ion gradient in the epidermis; (5) and the process of desquamation regulated by the activity of proteolytic enzymes (kallikreins 5 and 7) responsible for the SC integrity/cohesion and programmed renewal [5]. Schematic overview on the mechanisms involved in the formation of the skin hydration is presented in Fig. 13.2.

Mammalian skin is exposed to relatively dry surrounding environment. Retaining water, a function predominantly attributed to SC, ensures skin flexibility and elasticity. Water distribution is not homogenous in the epidermis. In vivo confocal Raman microspectroscopy (RCM) studies showed a continuous rise in the water concentration in SC from approximately 15–25% at the skin surface to about 40% at the SC/stratum granulosum border [3, 18]. This is followed by a steep rise to a constant level of about 70% in the viable epidermis. A typical curve on the water distribution in the epidermis representative for healthy skin is demonstrated on Fig. 13.3.

13.1.3
Legislation and the Marketplace of Moisturizing Cosmetics

Moisturizers represent a huge and intensively competitive share of the consumer product market [11]. This is a dynamic field for constant development not only for cosmetic ingredients but for trademarks and brands as well.
In general moisturizers are registered as cosmetic products and/or over-the-counter (OTC) drugs as is the case with some products in the US. The product category ranges from mass-market value brands to boutique and prestige luxurious products even claiming additional therapeutic effect. As for the term “cosmeceutical,” till date it has no separate and defined meaning with regard to legislation and regulatory process.

Moisturizers hold the third place in the ranking of most commonly recommended OTC topical skin products behind hydrocortisone and anti-microbial medications [17]. In such an environment a more strict legislation would be beneficial. Hence, with the sixth Amendment of the European Cosmetics Directive, an evidence for the claimed effect is required from the producer or the importer of a cosmetic product [4].

13.2 Non-Invasive Biophysical Methods in Testing Moisturizers and Emollients

Different methods exist for the proof of efficacy of topical products. They can be subdivided into several groups, i.e., clinical assessment and evaluation (performed either by a trained observer/investigator or by the volunteer him-/herself), non-invasive biophysical measurements (such as the assessment of the electrical properties of the skin and spectroscopic methods), and invasive procedures (e.g., skin biopsy and microdialysis). While the latter are traumatic to the panelists, clinical assessment is subjective and lacks the precision of an instrumental evaluation. On the contrary, skin physiology measurements have the advantage
of being non-invasive, non-traumatizing, causing minimal discomfort and not altering the skin functions per se [5]. In addition, they offer the possibility for objective detection of defined parameters, which in most cases cannot be discriminated by visual scoring.

Efficacy testing of moisturizers involves the application of a number of techniques namely:

- Assessment of the cutaneous electrical properties (capacitance, resistance, impedance) in relation to the water content of the outer skin layers (electrical hygrometry)
- Measurement of the evaporation through the epidermis, i.e., TEWL
- Instrumental evaluation of skin topography: surface/texture and desquamation (digital image analysis, silicone replicas, profilometry and squamometry)
- Analyzing some spectroscopic and optical properties of the skin (near infrared and Raman spectroscopy, nuclear magnetic resonance, optical coherence tomography, and confocal microscopy)
- Evaluation of the mechanical/visco-elastic properties of the skin (dynamometry, measurement of the skin reaction to friction, torsion and suction, and shear wave propagation method)
13.2.1
**Evaluation of the Electrical Properties of the Outermost Skin Layers**

Both skin conductance and capacitance rise with the increase of skin hydration/moisturization. However, the relationship between these parameters and the water content of the skin is not linear but rather more complex, as others factors such as ions, the dipolar structure of proteins as well as the different strength of water binding to keratin, the presence of hairs and cosmetic remnants on the skin surface influence the measurement. In addition it should be noted that skin is a biologic media and its behavior is not as predictable as an electrical conductor for instance. Finally, all electrical probes give to a greater extend integrated information on the hydration over the skin depth depending on the skin site and the instrumental characteristics. Thus, a direct comparison between measurements from different devices and anatomical test sites is not a proper approach.

The most commonly applied devices are based on measuring the electrical conductance, capacitance, resistance or impedance as an indirect indicator for SC water content. Low-frequency skin impedance measurements reflect rather the water content in the viable epidermis, whereas high-frequency conductance detects more selectively the hydration of SC [14].

A comparative in vivo study between five instruments (the capacitance-based Corneometer CM 820 and CM 825, the conductance-based Skicon 200 and DermaLab, and the impedance based Nova DPM) measuring the hydration of SC was performed [9]. CM 820 and its successor differentiate more precisely than the Nova DPM, the Skicon 200 and the DermaLab in dry conditions, while Skicon 200 is more sensitive in well-hydrated skin. Nova DPM, Skicon 200 and CM 825 are suitable for dynamic measurements of SC hydration such as the sorption-desorption test and moisture-accumulation test. A highly significant correlation was revealed between the instruments. However, a substitution of a device with another in the study course as well as simple comparison of the data from different devices is not advised.

A micro-sensor, multi-cell technology, Skin Chip®, is useful in assessing detailed capacitance mapping of the skin surface in vivo, thus providing an extended picture of the SC hydration [16]. The data on the skin capacitance acquired from each single cell of the sensor is processed and in this way a detailed “capacitance” map of the investigated skin site is obtained.

Innovative devices based on the estimation of the skin electrical properties have been proposed such as the multifrequency impedance instruments.

13.2.2
**Transepidermal Water Loss Assessment**

The ability of the SC to prevent the uncontrolled water evaporation from the living epidermal layers, i.e., permeability barrier, is reflected by the parameter TEWL. A low TEWL, therefore, is a characteristic feature of an intact and healthy skin state. In general, there is proportionality in the relation between TEWL and SC hydration (unaffected,
Moisturizers and Emollients

healthy skin). However, this is not the case with a number of conditions such as the hyperhydration state immediately after applying tensides, e.g., sodium lauryl sulfate (SLS), (have generally desiccating effect), the measurement in specific anatomical sites – palmar and plantar skin, and the measurements on a skin site with a heavily perturbed epidermal barrier [7].

A decrease in TEWL, parallel to an increase in the SC hydration, is observed after the application of occlusive substances (oils, petrolatum) on the skin. Elevated TEWL is registered directly (10–15 min) after the application of moisturizing agents on the skin surface. This effect is not due to the increase in the SC hydration, but reflects the evaporation of the water incorporated in the cosmetic product itself. In addition, moisturizing the skin leads to an increase in the SC thickness on the molecular level. As a consequence a hindered evaporation of water can appear. Thus, the use of TEWL as a single direct parameter for the characterization of the skin hydration should be applied cautiously. A combination of TEWL measurement with another non-invasive method is advised as more reliable and accurate approach.

13.2.3
Instrumental Evaluation of the Skin Topography

The examination of the skin surface topography with its roughness and scaling can be performed by several non-invasive techniques. The use of high-quality digital photography has been applied in studying skin surface properties. The use of light filters, allowing only a selected wavelength spectrum to reach the measurement area, contributes to the optical properties of the obtained surface picture.

Scanning microdensitometry was undertaken to estimate the shadows and the highlights of the photographic negative of the skin surface taken under standardized conditions (light, exposure time, camera angle and distance to the surface). This technique was efficient in diseased skin (ichthyosis, psoriasis); however, its discriminative ability was not as obvious in healthy volunteers.

The analysis of silicone replicas of the skin surface is another option for the assessment of skin topography. The microrelief variations of the surface can be followed by further analyzing of the skin replica – using a stylus (mechanical profilometry); auto-focus laser beam (laser profilometry); lateral illumination of the replicas; and the measurement of the thickness of translucent replicas. Limitations of profilometry include the detachment of the scales for the skin surface by the replica itself, the irritating and sensitizing properties of the replica’s materials, and the subjectivity in the reading of the results.

The scaliness of the skin surface generally corresponding to a dry skin state can be assessed by harvesting SC material and its consecutive analysis with spectroscopic, cytological and qualitative/quantitative methods. Skin barrier disruption by adhesive tapes is influenced by different variables, such as pressure, time and anatomical site. The demand to gain reproducible and reliable data resulted in standardization of stripping adhesive tape method [2]. This method is limited to the analysis of the skin surface material (scales) and is not a direct parameter of the hydration state of the skin.
13.2.4 Spectroscopic and Optical Properties of the Skin

A variety of spectroscopic methods have been applied in studying skin hydration namely infrared spectroscopy, photoacoustic spectroscopy, terahertz spectroscopy, millimeter wave reflectivity and in vivo confocal Raman microspectroscopy. These techniques, however, have certain limitations: the strong absorption of the infrared radiation by water (in Fourier transform infrared spectroscopy) limits the penetration depth of the light to a few micrometers (this technique reflects the hydration only in the outermost parts of SC). Exogenously applied substances (moisturizer ingredients) can potentially interfere with the electromagnetic emission and, thus, influence the optical properties of the pre-treated skin.

A novel spectroscopic method, in vivo confocal RCM, offers precise information on the water content in the skin layers with a high axial resolution of 2 µm. The physical basis of the technique is the inelastic light scattering of different molecules. During the measurement different Raman spectra are obtained specific to the chemical structure of the molecules. Considering the intensities of the Raman bands at certain shifts, the water to protein ratio is calculated.

The high axial resolution (maximum 2 µm) and the specificity of RCM are further used in the semi-quantitative measurement of skin components (lipids, lactate, urea, urocanic acid) and exogenously applied substances (dimethyl sulfoxide, trans-retinol, carotenoids) as a function of the depth of the epidermis [3].

Experimental methods such as nuclear magnetic resonance, optical coherence tomography, and ultrasound 20-MHz B-scanning have been applied in assessing epidermal hydration. However, these techniques need to be standardized and compared to the classical methods (e.g., electrical) in order to be implemented into practice.

13.2.5 Evaluation of the Mechanical/Visco-Elastic Properties of the Skin

The response of the skin to mechanical stress with intrinsic or environmental origin depends on tensile, rheological, and biochemical properties of the distinct cutaneous layers and the subcutis. A correlation between epidermal hydration and skin visco-elastic properties in vivo was evidenced [6]. Thus, estimation of the skin visco-elastic properties can be used as an indirect indicator for skin hydration. However, no significant correlation between skin capacitance and mechanical parameters was found and the results of this technique should be interpreted cautiously.

The in vivo mechanical properties of the skin were studied by different methods based on torsion stress, indentation, ballistometric techniques, uniaxial stretching, and suction. When using the measurements of the mechanical skin properties, one should keep in mind that the obtained parameters are indirect. Beyond water content they reflect rather complex skin characteristics such as the organization of the dermal collagen and elastic fibers, and the desquamation process. Thus, assessment of skin mechanical can only be used in addition to the established hydration measurements.
13.3 Practical Aspects of Moisturizers Testing

13.3.1 Selection of the Investigative Method

Selection of the proper study method and technique is a crucial step in the study planning process. It is dependant on a number of factors, namely the aim of the study, the availability of the selected equipment at the laboratory, price of the apparatus, and the requested time for each measurement. For instance, RCM offers the most detailed and complex information on the skin hydration (with an axial resolution of 2 μm) as well as on the topical moisturizer distribution in the skin. On the other hand, the time for each of the measurement exceeds significantly the time elapsed when using classical methods, e.g., capacitance measurement. The overview on the different devices in the previous section of this chapter aims to facilitate the investigator’s choice of proper equipment with regard to their advantages and drawbacks.

The variety of different techniques can make the selection process cumbersome and less efficient. Moreover none of the method discloses the complete interactions between a moisturizer and the skin. Hence, a multiparametric approach is useful in assessment skin physiology [5]. In this way different aspects of the moisturizer effect on the skin and its functional consequences can be revealed.

An important issue that has to be addressed is the irreplaceability of the devices during the whole study. Even using the same measuring principle the data obtained with different instruments cannot be directly compared and/or substituted [9]. Some devices need internal calibration that is valid only for the specified apparatus and this obstructs transferring data between instruments form the same brand.

A calibration procedure must be performed before starting a new study. If the study includes a larger time-span calibration must be performed during the study and according to the instructions provided by the device manufacturer [10].

13.3.2 Study Design

Various study designs exist with regard to testing hydrating cosmetic products. Independent from the study design, the inclusion of a reference moisturizer with the same vehicle (or based on the same emulsion type) is recommended. Petrolatum and paraffin are not suitable for this purpose due to the occlusion caused by these substances. One test field should be left untreated and measured at each time point. In this way each panelists serves as his/her control.

13.3.2.1 Single Application Tests

In the short-time experiments the product(s) is applied on the skin and the biophysical measurements are performed at different intervals up to 4–6 h after the application of the
moisturizer. The instrumental measurements are performed at baseline ($t=0$) at defined time points (e.g., every 30 min). A key issue in this protocol is to perform the first post-application measurement at least 30 min after the product introduction [1]. Earlier performed measures can detect the evaporation of the water in the moisturizer itself. In addition, the application of occlusives (e.g., petrolatum) results in a blockage of the transepidermal evaporation immediately after their application on the skin surface.

Single application tests have the advantage of being faster and cheaper but are far from the simulation of a “real-life” usage of moisturizers. They are significantly influenced by the physical properties of the remains of the cosmetic product.

13.3.2.2
Multiple Application Tests

The design with multiple applications resembles in a greater extent the everyday setting compared to the short-term studies. In this case the product is applied for a longer period – 2–4 weeks, most often twice daily. A baseline measurement ($t=0$) is followed by measurements at defined time points, e.g., every week. The assessment of the skin site should be performed at least 8–12 h after the last product application.

The long term design is more realistic; however, it is difficult to control the effect of environment-related variables (climate) on the panelists. Therefore, more subjects should be included to diminish the statistical effect of the influencing factors. This raises directly the price of the whole study. In addition, the use of multiple application tests can make the discrimination between different moisturizing products difficult.

13.3.2.3
Regression Tests

Since its introduction, the “classic” Kligman regression test (duration of 5 weeks, test site – lower legs) has been modified with regard to the duration, number of subjects and products tested, and test site [12]. In this model the continuous (multiple) application of the product is followed by a regression phase. The aim of this phase is to evaluate the rate of recurrence of the dry skin condition. A prerequisite for this design is to be performed on panelists with cutaneous xerosis. The duration of the regression varies between 6 and 21 days and is a direct function of the duration of the application (treatment) phase. In this period the measures are performed at a regular basis and the time elapsed to reach the baseline values of the skin hydration (before treatment) is accounted. A shorter timeline protocol gives the advantage to decrease the study costs. The regression protocols answer the question on the performance of a moisturizer to sustain the hydration of the healed dry skin.

13.3.2.4
Efficacy on Pre-Irritated Skin

A common cause for the cutaneous xerosis is the frequent use of detergents, water, and other potentially irritating substances. Thus, simulating these conditions can be useful in
moisturizers testing. Preliminary skin irritation can be achieved by different approaches: mechanical irritation – tape stripping; detergent-induced irritancy – SLS (open and occluded application); delipidization of SC – acetone, and in a more realistic way – a number of washing models. The reader is referred to the specified chapter in this book or to the state-of-art review on standardized washing models used in skin physiology studies [8]. One may keep in mind that certain time (30 min–12 h) is needed before the first measurement after the irritation can be performed, in order to allow the drying of the skin, to avoid the false registration of an occlusive effect (chamber application), and the initial “hyperhydrating” effect of some irritants (e.g., corneocyte swelling caused by SLS).

After the initial skin barrier disruption (induction of skin irritation), the moisturizer is applied in a multiple manner (see Sect. 3.2.3) and the dynamics of the skin hydration restoration is registered. With this model different products can be compared regarding their restoration capacity of desiccated skin and how long they can sustain this effect over time.

Remark

A wide range of washing and cleansing products claim to exert hydrating effect on the skin. However, the application of surfactants/tensides on the skin alone or combined with water results almost unexceptionally to exsiccation of the skin surface. Therefore such type of products cannot be accepted as moisturizers and emollients in a broad sense. Some often have mild desiccating effect.

13.3.3 Study Population

13.3.3.1 General Considerations

The adequate recruitment is a major condition for the successful study conduction. In general, an integral part of each study protocol is the inclusion and the exclusion criteria section. Herein, the frame of the study population is set. Thus, the effect of different variables should be considered in the planning phase.

Testing moisturizers requires a homogenous population, e.g., subjects with dry skin. However, when applying some of the previously discussed designs (testing on pre-irritated skin) no such a prerequisite is necessary.

Before entering the study, each of the participants should receive written, detailed information about the nature and the aim of the study and should sign written informed consent form according to the study protocol approved by the local regulatory board. Hence, the written informed consent must be set as one of the inclusion criteria.

Certain specific populations such as pregnant and lactating women, children (age <18), imprisoned and subjects with mental/psychiatric illness have to be excluded. Furthermore, the presence of concomitant systemic disease (diabetes, renal insufficiency, autoimmunity, acute infections, tuberculosis, neoplasm), clinically relevant skin diseases (except skin
dryness), and use of certain systemic and topical medications (e.g., immune modulators, antibiotics, glucocorticosteroids, retinoids, azelaic acid, acne and seborrhea therapeutics) can alter the skin biophysical characteristics. Previously reported hypersensitivity reaction of any type to the product ingredients is a relevant exclusion criterion. The simultaneous participation of a subject in multiple studies is unacceptable and should be stated in the selection criteria.

13.3.3.2
Number of Subjects

Statistical power calculation methods should be addressed when the number of subjects is determined. Larger study population is required when applying long-term design in order to alleviate the influence of external variables. The greater number of panelists can be helpful in revealing obscure differences between groups/products. On the other site, an over exceeding number of volunteers can hinder the study procedures and raise the expenses. Generally, a minimum of 12 subjects to complete the study are necessary for statistical data evaluation.

An over-recruitment (initially performed by a phone talk) can be helpful to replace the do not show up, the drop-outs, and the ineligibility according to the inclusion/exclusion criteria (approximately 5–10%).

13.3.3.3
Individual-Related Variables Influencing the Biophysical Measurements

Subject-related variables (Table 13.1) influence skin physiology measurements. For example, age-related differences in skin hydration have been revealed with hydration decreasing in elderly [1, 5]. Although no explicit and unidirectional discrepancies with regard to race/ethnicity are known, a divergence depended on this factor has been reported. No gender influence on skin hydration exists. A good marketing approach is to pre-select the subjects in a way they would represent best the targeted (by the product) population (e.g., ethnic group, skin phototype, gender, and age range group). If not applicable, the selection process should be based on the already described criteria.

13.3.4
Test Site

Intraindividual (site-to-site) variations in the skin hydration are well documented [5]. Higher hydration values are obtained form regions with higher density of the sweat glands (palms, forehead) while lower values are evidenced on the abdomen, the lower legs and the extensor site of the extremities.
Lower legs and volar forearms are the most widely used test sites. Both locations offer the possibility for contra-lateral comparison (left vs. right) as well as the application of multiple products. A bilateral consistency for the hydration of the skin sites is required at baseline. When multiple products are tested a standard randomization procedure should be followed. The zones of anatomic occlusion (e.g., cubital fossa) and the wrist should be avoided. The minimal distance of the test site border from these locations must be at least 5 cm. Selecting anatomical sites which are hard-to-reach or have an uneven relief can hinder the measuring process. The choice of an “exotic” test filed renders difficulties in the external validation and the comparison of the study data.

Hairs must be removed by gentle clipping (electrical clipper or scissors) prior to the start of the study and at least 30 min before the first measurement. If a razor-blade is used the shaving procedure must be completed 48 h before entering the study. Hairs impede the contact of the measuring probe to the skin surface. The presence of scars in the test filed must be avoided since no physiological conditions are present.

The size of the test site is depended on the selected design, i.e., multiple product testing requires more test fields. The site is delineated by using skin marker and a stencil allowing the precise location of the test field. In general, the size of the field varies from 6 to 12 cm². No overlaying between the separate sites is accepted.

### Table 13.1 Subject- and environment-related variables influencing skin hydration measurement assessed by the evaluation of the electrical properties of the skin and the transepidermal water loss (TEWL) measurement

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SC hydration</td>
</tr>
<tr>
<td>Age</td>
<td>+</td>
</tr>
<tr>
<td>Gender</td>
<td>−</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>±</td>
</tr>
<tr>
<td>Anatomical site</td>
<td>+</td>
</tr>
<tr>
<td>Skin temperature</td>
<td>+</td>
</tr>
<tr>
<td>Sweating</td>
<td>+</td>
</tr>
<tr>
<td>Air convection</td>
<td>+</td>
</tr>
<tr>
<td>Ambient temperature</td>
<td>+</td>
</tr>
<tr>
<td>Humidity</td>
<td>+</td>
</tr>
<tr>
<td>Direct light</td>
<td>+</td>
</tr>
<tr>
<td>Season</td>
<td>+</td>
</tr>
<tr>
<td>Circadian rhythms</td>
<td>+</td>
</tr>
</tbody>
</table>

SC stratum corneum
“+” influencing; “−” no influence; “n.d.” no data; “±” controversial data
13.3.5
The Study Procedures

13.3.5.1
Instructions to the Panelist

Before entering the study each panelist should receive written instructions with detailed description of the study procedures. The information sheets should be written in clear language and the use of specific medical terms is undesirable. The main point is to raise the volunteer’s compliance and thus the language in the instruction sheet should be comprehensive. A schematic overview of the study procedures over time could be beneficial for better visualization and understanding the time points of each product application and visit. The approximate planned time for each visit may be included.

13.3.5.2
The Washout Phase

The panelists should be aware of the restrictions during the study performance. In general, a washout phase should be obeyed prior to the study start. During this phase the washing is standardized in all subjects, e.g., number of bathing per day (most commonly restricted to once daily) and the use of pre-selected cleansers (mild soap either provided by the study team or giving a recommendation on an available trade mark on the market). The use of hydrating milks, lotions, creams and other cosmetics that could potentially alter skin hydration is forbidden during the washout phase as well during the whole study. The same is valid for any topical medication. If the avoidance of topical product application is not possible for the whole body, its use should be restricted outside the test areas.

Alcohol consumption, caffeine intake and some vasoactive medications alter skin microcirculation which indirectly can influence skin hydration measurement. The use of such substances is prohibited at the day of the measurements.

On the other hand, tests on pre-irritated skin request dry skin as a prerequisite which can be accomplished by either restriction of the daily moisturizer use or in an exaggerated mode – by the treatment with model chemical irritants (acetone, SLS) or mechanically (tape stripping).

The duration of the washout phase varies from 48 h to 1 week and depends on the study design.

13.3.5.3
Environment-Related Variables Influencing the Measurements and the Panelist Acclimatization

Environment-related variables influence skin hydration and its quantification (Table 13.1). Thus controlling the microenvironment at the measuring room is critical for skin physiology measurements. In order to avoid the measurement of the sweat gland activity, the ambient air temperature should be below 22°C. Room temperature range from 18 to 22°C (20 ± 2°C) is recommended.
### Table 13.2 Recommendations for skin hydration measurements at a glance

<table>
<thead>
<tr>
<th>Variable related to:</th>
<th>Variable</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panelist</td>
<td>Age</td>
<td>18–60</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>No requirement for gender matching</td>
</tr>
<tr>
<td></td>
<td>Number of subjects</td>
<td>12–30</td>
</tr>
<tr>
<td>Test site</td>
<td>Anatomical location</td>
<td>Lower legs and/or volar forearms (avoid flexural areas)</td>
</tr>
<tr>
<td></td>
<td>Test site size</td>
<td>6–12 cm² (no overlying of the test fields)</td>
</tr>
<tr>
<td></td>
<td>Hair clipping/shaving</td>
<td>30 min/48 h</td>
</tr>
<tr>
<td>Environment</td>
<td>Room temperature</td>
<td>18–22°C</td>
</tr>
<tr>
<td></td>
<td>Ambient air relative humidity</td>
<td>40–60%</td>
</tr>
<tr>
<td></td>
<td>Acclimatization duration</td>
<td>20–30 min</td>
</tr>
<tr>
<td></td>
<td>Circadian rhythms</td>
<td>Perform measurements in the same daytime and season; avoid measurements in summer</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>Avoid direct air flow and direct light</td>
</tr>
<tr>
<td>Product</td>
<td>Product quantity</td>
<td>1–3 mg/cm² (or μL/cm²)</td>
</tr>
<tr>
<td></td>
<td>Application technique</td>
<td>Rub with glove-covered finger or with glass/plastic rod</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td>Use a referent moisturizer; measure at an untreated test site</td>
</tr>
<tr>
<td></td>
<td>Randomization</td>
<td>Double-blinded product allocation to the test sites</td>
</tr>
<tr>
<td></td>
<td>Coding and labeling</td>
<td>Study number, subject number (initials), product number</td>
</tr>
<tr>
<td>Design</td>
<td>General</td>
<td>Always perform a baseline ($t=0$) measurement</td>
</tr>
<tr>
<td></td>
<td>Washout phase</td>
<td>2–7 days</td>
</tr>
<tr>
<td></td>
<td>Duration of a study with a single application design</td>
<td>4–6 h</td>
</tr>
<tr>
<td></td>
<td>Interval between measurements in a single application design</td>
<td>30 min–1 h</td>
</tr>
<tr>
<td></td>
<td>Duration of the treatment phase in multiple application designs</td>
<td>2–4 weeks</td>
</tr>
<tr>
<td></td>
<td>Duration of the regression in regression tests</td>
<td>6–21 days</td>
</tr>
<tr>
<td></td>
<td>Measurement on pre-irritated skin</td>
<td>30 min–12 h from the irritating procedure</td>
</tr>
<tr>
<td>Measurement</td>
<td>Number of consecutive measurements</td>
<td>3–10</td>
</tr>
<tr>
<td></td>
<td>Probe application to the skin surface</td>
<td>Perpendicular</td>
</tr>
<tr>
<td></td>
<td>Avoiding occlusion from the measuring probe</td>
<td>Leave interval of 5 s between measurements; or measure at close-standing but not overlying areas</td>
</tr>
<tr>
<td></td>
<td>Cleansing of the probe head</td>
<td>After each measurement</td>
</tr>
<tr>
<td></td>
<td>Consequence of measurements with multiple devices</td>
<td>Start with the less invasive and the less time consuming</td>
</tr>
<tr>
<td></td>
<td>Interval from the last product application</td>
<td>At least 12 h</td>
</tr>
<tr>
<td></td>
<td>Interval from skin cleansing</td>
<td>2–4 h</td>
</tr>
</tbody>
</table>
A linear relationship between skin hydration assessed by capacitance and the ambient air relative humidity exist [1]. Hence a control in this parameter (room air humidity) must be performed and should be within the range of 50 ± 10%. The presence of air-humidifier is necessary in the measuring room.

Direct air flow can interfere with some of the measurements such as the TEWL assessment [15]. Therefore, the air flow produced by air-conditioning systems should be kept away as far as possible from the test area. Door and window opening as well as the breathing in the direction of the measuring probe must be avoided. The use of a protection shield/chamber to isolate the air flow at the test area is recommended by some authors, especially for older device models.

Direct light exposure of the test site either from natural sources or electrical ones should be diminished as this can alter skin surface temperature and cause heating, respectively sweating. The close contact with the hands of the investigator can also produce a rise in the skin temperature of the panelist. Thus, the use of isolation gloves by the measurer can be helpful. In most of the presently available probes, however, the use of such protective gloves is not needed.

A period of acclimatization of the panelists to the standardized conditions of the laboratory for at least 20 min is necessary. We recommend an optimal period of 30 min acclimatization. A fully air-conditioned room, connected to the measurement room can serve as an acclimatization space. During the acclimatization, the subject should be completely relaxed from physical activity and psycho-emotional stress. The anatomical test site should be left uncovered (from clothing and jewelry for instance) as this can lead to occlusion and mechanical irritation.

The study instruments need to be equilibrated to the laboratory conditions and should be calibrated as recommended by the manufacturer.

Biorhythms affect skin physiology. Seasonal, day-to-day and even diurnal variations have been demonstrated [1]. Therefore, the measurements should be performed in the same time of the day. This should be included in the panelist instructions and considered in the study planning phase. A study should be completed in a single season (if not aiming to prove inter-seasonal variation). In general, summer (July, August in the Northern hemisphere) is avoided when performing skin hydration measurements, due to the great temperature variances and the UV influence on the skin immunology and physiology.

13.3.5.4
The Product Application

The quantity of the applied product is calculated as a function of the test field surface area and ranges from 1 to 3 mg/cm² (or µL/cm²). The tested product and the reference moisturizer are gently rubbed with a glove-covered finger to the test field in order to avoid interference with the investigator’s sebum and sweat secretion or remnants from a different product. Each product is applied after the glove finger has been replaced with a new one (to avoid product mixing). Instead of using latex glove-covered finger, the cosmetic can be applied by a laboratory glass or plastic rod. However, the latter is more expensive and can cause inconvenience in handling the rods by the volunteers.

In multiple application designs, an in-office application must be performed by the study investigators so to demonstrate to the panelist the way of product application. A repetition
of the procedure by the panelist before the investigator guarantees that he/she has accurately understood the application procedure.

When measuring multiple products the containers with the separate products should be coded and labeled with as little product information as possible. The label should contain the study number, the product code and the panelist number or initials. A finger-tip unit may be an appropriate dosing strategy. Otherwise, weighting of the containers is necessary before the beginning of the study and at any time point of the measurements. The separate products are randomized to the distinct test fields under a standard double blinded (for both the investigator and the subject) procedure. Thus, bias with regard to anatomical site or the severity of the skin dryness is avoided.

In the multiple application tests the last product application should be performed the night before the measurement (at least 12 h). Otherwise, the cosmetic remnants on the skin surface might be measured.

When the protocol of the study does not allow a 12 h period free of product application, cleansing of the skin is recommended by a mild, non-alcoholic and non-aggressive cleanser at least 2–3 h before the measurement. Such a procedure should be performed for each test site and at each time point including the baseline measurement ($t=0$). Even though skin hydration can be altered and the cleansing procedure should be taken in consideration itself.

### 13.3.5.5

#### The Measurement

A mean of multiple measurements (at least 3) from each test site can eliminate outlying values. In the case of multiple measurements, each next measurement must be performed with a delay of at least 5 s from the previous to avoid the occlusion effect of the probe. Another approach is to perform the consecutive measurements at areas standing close to each other but not overlying in the test field. Cleansing of the probe head is recommended between the measurements. A constant pressure should be kept at the probe despite that the majority of the devices have inbuilt spring mechanism to control the pressure to the skin surface. The probe head must be applied perpendicularly to the skin surface.

The sequence of measurements using multiple devices should be in the order in which the measurement procedure affects the next one in a minimal way. There is no uniform understanding on the order of measurements. However, it is practical to perform first the measure which takes shortest contact time of the probe with the skin surface. Graduating the procedures according to their invasiveness is also useful, starting with the less invasive. For instance, electrical methods must be performed before harvesting techniques such as tape stripping.

### 13.3.5.6

#### Data Management and Evaluation

Performing the measurement is as critical as handling the obtained results. All data should be managed according to the study protocol. Data management includes the registration of
the measurement values in a clinical research form either in an electronic database or a hard copy (or both). Recording the microenvironment condition (temperature and humidity) is advised. Data must not be freely accessible neither by each member of the study team nor from a third party. Thus, data should be kept in internal servers, not allowing internet access to the database.

Statistical evaluation of the data is performed with software products such as SAS, SPSS and Prism. The proper choice of a statistical method is important in interpreting the data. An experienced statistician should be consulted in the data evaluation process as well as in the planning phase of the study. One should keep in mind that the majority of the methods register indirect parameters for skin hydration. Thus, a interpreting the measurement values directly in terms of a change in SC hydration can be imprecise. A better approach is to give a comparative analysis between the tested products or to compare the effect of the product to the one of the referent moisturizer.

13.4 Conclusion

Human skin is a complex and dynamic biologic system. There is no a single non-invasive method which can fully describe the hydration properties of the skin and the influence of externally applied topical substances. Hence, a multiparametric approach is more useful in interpreting skin moisturization. In addition, the combined use of biophysical methods and clinical (trained observer) evaluation could be the most complex and realistic approach in testing moisturizers. The detailed planning as well as controlling different subject-, instrument-, and environment-related variables is a necessary for performing a successful skin physiology study.

References

Antiaging and Antiwrinkle Products

Razvigor Darlenski, Theresa Callaghan, and Joachim W. Fluhr

Core messages

› The chronological (intrinsic) and extrinsic aging demonstrate typical macroscopic, histological, and functional characteristics
› The relative improvement in different parameters characterizing aging skin can be used in efficacy proof of antiaging and antiwrinkle cosmetic products
› Different approaches to investigate the efficacy of antiaging products exist such as clinical evaluation and objective assessment with noninvasive methods and invasive procedures
› A multiparametric approach is useful in the assessment of antiaging products efficacy
› There is no uniform consensus on the protocol and the design of studies aiming efficacy proof of antiaging cosmetics

Abbreviations

RCM In vivo confocal Raman microspectroscopy
SC Stratum corneum
TEWL Transepidermal water loss
UV Ultraviolet

R. Darlenski (*)
Department of Dermatology and Venereology, Medical University, Sofia, Bulgaria
e-mail: darlenski@abv.bg

T. Callaghan
Witts Park 16, 22587 Hamburg, Germany

J.W. Fluhr
Department of Dermatology, Charité University Clinic, Charitéplatz 1, 10117 Berlin, Germany

DOI: 10.1007/978-3-642-05067-1_14, © Springer-Verlag Berlin Heidelberg 2011
14.1 Introduction

The elderly population is growing in the developed countries and the same trend is expected in the first half of the twenty first century for the developing countries also [8]. The reasons for the global trend with a shift toward an older population profile are complex and include the improvement of health care and the living conditions, the decrease of fertility rate and infant mortality, as well as the effective treatment of infectious diseases.

Cutaneous aging is the irreversible process of the development of specific morphologic, biochemical, and biophysical alterations of skin parallel to the senescence of the whole organism. In contrast to other organs and systems, skin aging is not only determined by the genetic program of the individual (intrinsic aging; chronologic aging) but is also influenced by external factors such as the ultraviolet (UV) radiation, smoking, and lifestyle (extrinsic aging, photo aging). The major clinical, microscopic, and functional characteristics of the aged skin are summarized in Table 14.1.

Proving the efficacy of antiaging products together with the study of the intimate mechanisms of aging and age-related disorders as well as the socio-economic factors and consequences of aging are the major branches of research in the field. The market is abundant in antiaging agents and procedures. Claim support is an essential part in the registration of cosmetic products. Thus, proving the efficacy of such products (antiaging and antiwrinkle products in particular) is a binding procedure not only for marketing purposes, but also for legislation issues.

14.2 Methods Used in the Evaluation of Antiaging Products

Different approaches have been applied in testing antiaging product efficacy namely clinical scoring, visual methods, and functional assessment of biophysical parameters, e.g., surface parameters, transepidermal water loss (TEWL), stratum corneum (SC) hydration, and skin mechanical properties.

| Table 14.1 Clinical, histological, and functional characteristics of intrinsically and photoaged skin |
|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
| Intrinsic aging                  | Extrinsic aging                  |
| Clinical appearance              | Fine wrinkles                    | Deep wrinkles                    |
| Thin and transparent skin        | Thinning and firmness            | Irregular and uneven pigmentation|
| Loss of underlying fat with     | noticeable loss of firmness      | (freckling, lentigines, guttate   |
| Skin sagging and slackness       |                                 | hypomelanosis)                   |
| Xerosis                          |                                 | Decreased elasticity and         |
| Benign neoplasms, e.g., seborrheic keratoses and cherry angiomas | Xerosis                          | pronounced flaccidness          |
|                                  |                                  | Vascular lesions (teleangiectasia, |
|                                  |                                  | purpura senilis, venous lakes)   |
|                                  |                                  | Neoplastic lesions, e.g., actinic keratoses, comedones (Favre-Racouchot disease), nonmelanoma skin cancer, lentigo maligna |
Table 14.1 (continued)

<table>
<thead>
<tr>
<th>Histological features</th>
<th>Intrinsic aging</th>
<th>Extrinsic aging</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epidermis</strong></td>
<td>No alterations in SC</td>
<td>Increased compaction of SC</td>
</tr>
<tr>
<td></td>
<td>Discrete/absent changes in keratinocyte shape</td>
<td>Increased thickness of stratum granulosum</td>
</tr>
<tr>
<td></td>
<td>Fewer melanocytes and Langerhans cells</td>
<td>Reduced epidermal thickness</td>
</tr>
<tr>
<td></td>
<td>Flattened rete ridges at the dermo-epidermal junction</td>
<td>Reduced epidermal mucin content</td>
</tr>
<tr>
<td><strong>Dermis</strong></td>
<td>Atrophy with alteration of the connective tissue</td>
<td>Increased number of melanosomes in the basal keratinocytes</td>
</tr>
<tr>
<td></td>
<td>Fewer fibroblasts and mast cells</td>
<td>Irregularities in cell and nuclear shape and size</td>
</tr>
<tr>
<td></td>
<td>Decrease in the number of dermal blood vessels with shortening of the capillary loops</td>
<td>Flattened rete ridges at the dermo-epidermal junction</td>
</tr>
<tr>
<td></td>
<td>Decrease in the nerve endings and the neuroreceptive apparatus of the skin</td>
<td>Decrease in nerve endings and the neuroreceptive apparatus of the skin</td>
</tr>
<tr>
<td></td>
<td><strong>Epidermis</strong></td>
<td><strong>Dermis</strong></td>
</tr>
<tr>
<td></td>
<td>Increased compaction of SC</td>
<td>Prominent grenz zone in the papillary</td>
</tr>
<tr>
<td></td>
<td>Increased thickness of stratum granulosum</td>
<td>Homogenization of the connective tissue</td>
</tr>
<tr>
<td></td>
<td>Reduced epidermal thickness</td>
<td>Decreased and degraded collagen (collagens I, III, and VII)</td>
</tr>
<tr>
<td></td>
<td>Reduced epidermal mucin content</td>
<td>Deposition of abnormal elastotic material</td>
</tr>
<tr>
<td></td>
<td>Increased number of melanosomes in the basal keratinocytes</td>
<td>Increased matrix-degrading metalloproteinases</td>
</tr>
<tr>
<td></td>
<td>Irregularities in cell and nuclear shape and size</td>
<td>Decreased amount of glucosaminoglycans</td>
</tr>
<tr>
<td></td>
<td>Flattened rete ridges at the dermo-epidermal junction</td>
<td>Vascular dilatation and increased capillary tortuosity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Functional abnormalities</th>
<th>Competent basal epidermal barrier function</th>
<th>Competent basal epidermal barrier function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delayed recovery processes after acute barrier perturbation</td>
<td>No difference in the barrier function in comparison to intrinsically aged skin</td>
<td></td>
</tr>
<tr>
<td>Decreased SC hydration</td>
<td>Decreased SC hydration on the photoexposed vs. protected skin – impairment in SC water-binding capacity (compared to intrinsically aged skin)</td>
<td></td>
</tr>
<tr>
<td>Elevated skin surface pH (not valid for all anatomical sites)</td>
<td>Elevated skin surface pH (not valid for all anatomical sites)</td>
<td></td>
</tr>
<tr>
<td>Reduced vasorelaxant response of the cutaneous microvessels</td>
<td>Reduced luminescence of skin color (in comparison to photoprotected areas)</td>
<td></td>
</tr>
<tr>
<td>Decrease in elasticity and extension but increased fatigability</td>
<td>Elevated values obtained by LDV, corresponding to an increase in the subpapillary vascular plexus</td>
<td></td>
</tr>
</tbody>
</table>

*SC* stratum corneum
14.2.1 Clinical Assessment

A plethora of clinical scoring scales have been implicated in staging the visual parameters of skin aging. One of the most widely used grading systems is the Glogau scoring of photoaged skin according to the degree of wrinkling (Table 14.2) [6]. The reduction of the grade of visual skin aging can be used as a parameter for the assessment of antiwrinkle products efficacy. The evaluation should be performed by the same trained observer as before the treatment as at each time point of the study. However, one should keep in mind that the clinical assessment lacks the precision of the objective methods.

14.2.2 Non-invasive Biophysical Methods

14.2.2.1 Instrumental Evaluation of the Skin Surface Topography

The examination of the skin surface topography with its roughness and its scaling can be performed by several noninvasive techniques. The use of high-quality digital photography has been applied in studying skin surface properties. Using polarized light reduces the shine reflection from the surface and, thus facilitates the greater visualization of the fine lines and wrinkles as well as vascular and pigmentation components. Video microscopy is another noncontact method, which can be used in revealing quantitatively the changes in the cutaneous microrelief [3].

Fringe Projection is a method by means of which a 3-D surface contour of the skin can be obtained. Commercially available devices are PRIMOS (Primos, Teltow, Germany) and DermaTOP (Breuck-Mann GmbH, Meersburg, Germany) [3,5]. The method is used to measure the depth and width of wrinkles.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Skin characteristics</th>
<th>Age group (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>Few wrinkles, no keratoses</td>
<td>28–35</td>
</tr>
<tr>
<td>Moderate</td>
<td>Early wrinkling, sallow complexion with early actinic keratoses</td>
<td>35–50</td>
</tr>
<tr>
<td>Advanced</td>
<td>Persistent wrinkling, discoloration of the skin with telangiectases and actinic keratoses</td>
<td>50–60</td>
</tr>
<tr>
<td>Severe</td>
<td>Severe wrinkling, photoaging, gravitational and dynamic forces affecting the skin, actinic keratoses with or without skin cancer</td>
<td>65–70</td>
</tr>
</tbody>
</table>
Profilometry of skin surface replicas is an approach to assess skin topography. However, a discrepancy in the depth of the wrinkles measured directly compared to the depth obtained by measuring the replicas has been reported [1]. Corneosurfometry and the use of D-squames are useful and standardized in the evaluation of the skin microrelief and roughness. In vivo confocal microscopy reveals the intimate characteristics of the aged skin noninvasively. However, validation studies are required to implement this method as a routine in the evaluation of skin aging.

14.2.2.2

Epidermal Barrier Function by Transepidermal Water Loss Assessment

Estimation of the TEWL values is generally accepted as a reliable parameter reflecting the permeability barrier function of the skin. In elderly, TEWL are either similar or even lower than those in young subjects, although the restoration of acutely damaged barrier is delayed in aged subjects (reviewed in [4]). A lower basal TEWL was shown in the perioral area, the neck, and the forearm in a group of aged vs. young volunteers. On the contrary, higher TEWL values in the older age group were registered in the nasolabial area, the upper eyelid, forehead, chin, and nose. No difference in the TEWL values was found in photo-damaged compared to intrinsically aged skin [7].

14.2.2.3

Evaluation of the SC Hydration by Electrical Methods

A decreased SC hydration and reduced water-binding capacity due to diminished quantities of water binding substances, i.e., the natural moisturizing factors, is observed in aged skin. Additionally, skin dryness is higher on sun-exposed vs. unexposed sites. Thus, reduction in the SC hydration can be used in the assessment of antiaging cosmetics. Different techniques for the assessment of SC hydration have been described: microwave, thermal, spectroscopic, including nuclear magnetic resonance spectroscopy, infrared and Raman spectroscopy. However, the most commonly applied methods are measuring the electrical conductance, capacitance, or impedance as indirect indicators for SC hydration. A microsensor, multicell technology, Skin Chip®, is useful in providing detailed surface capacitance mapping of the skin surface in vivo, thus providing an extended picture of the SC hydration. The data on the skin capacitance acquired from each single cell of the sensor is processed and in this way, a detailed “capacitance” map of the investigated skin site is obtained.

Newly developed techniques such as in vivo confocal Raman microspectroscopy (RCM), offers precise information on the water content in the skin layers with a high axial resolution of 2 μm. The high axial resolution and the specificity of RCM are further used in the semiquantitative measurement of skin components (lipids, lactate, urea, urocanic acid) and exogenously applied substances (dimethyl sulfoxide, transretinol, carotenoids) as a function of the depth of the epidermis.
14.2.2.4
Skin Pigmentation and Color

Visualization techniques such as analysis of high-quality digital photography and video microscopy can be used in studying the pigmentation changes of the aged skin. UV-reflectance photography and in vivo confocal microscopy allow the observation of areas of hyperpigmentation below the surface of the skin.

Tristimulus reflectance colorimetric analysis of light reflected from skin structures (e.g., ChromaMeter-Minolta, Osaka, Japan), reflectance spectrophotometry (e.g., Dermaspectrometer-Cortex Technology, Hadsund, Denmark; Mexameter MX 18- Courage & Khazaka electronic, Cologne, Germany), and chromophore mapping (e.g., SIAscopy™ – Astron Clinica, UK and RBX™ – Canfield Imaging Systems, Fairfield, NJ, USA) have been successfully applied in quantification of the aging skin pigmentation.

14.2.2.5
Skin Surface Acidity (pH)

A decreased buffer capacity of the skin related to increased skin surface pH was revealed in elderly (reviewed in [4]). An elevated pH under basal conditions was evidenced on eight different facial skin sites and on the forearm skin in elderly vs. younger study groups. Statistically significant differences were observed for the forehead, upper eyelid, neck, forearm, and the ankle.

Different methods for the assessment of the skin surface pH exist. Flat glass electrode measurements remain the most commonly used, as being simple, quick, and reproducible. Any commercialized pH meter devices fitted with a flat glass electrode can be used for the measurement of skin surface pH. However, SC is semihydrophobic tissue but not a solution. That is why the conditions in which the pH-metry is performed are not fulfilling the defined demands (i.e., standardized temperature, liquid media) for pH assessment.

14.2.2.6
Assessment of Skin Microcirculation

A significant decrease in the capillary loop density but increased vascular length was reported in elderly [9]. These changes were parallel to elevated values obtained by laser Doppler velocimetry, corresponding to an increase in the subpapillary vascular plexus. Studies on the dynamic properties of the skin microcirculation showed reduced vasorelaxant response of the cutaneous microvessels in senescent subjects [2].

14.2.2.7
Assessment of Skin Surface Lipids by Photometric Methods

Skin surface lipid content decreases with age. The sebum amount present at the skin surface can be assessed noninvasively using one of several methods based on solvent
extraction, cigarette paper pads, photometric assessment, bentonite clay, and lipid-sensitive tapes. By means of these techniques, different quantitative parameters can be assessed, i.e., sebum casual level, sebum excretion rate, sebum replacement time, instant sebum delivery, follicular excretion rate, density in sebum-enriched reservoirs, and sustainable rate of sebum excretion. One of the most widely used techniques is photometry (e.g., Sebumeter SM 815 – Courage & Khazaka electronic, Cologne, Germany) for its high reproducibility, simplicity, and easiness to perform. It should be noted that the excretion rate has to be assessed at least at two different time points after standardized removal of surface lipids, for example, with alcohol.

14.2.2.8
Assessment of the Skin Mechanical Properties

Skin mechanical properties are changing with aging in relation to the alterations of the dermal connective tissue. The in vivo mechanical properties of the skin were studied by different methods based on torsion stress, indentation, ballistometric techniques, unaxial stretching, and suction. The instruments that are commonly used in dermatological and cosmetic studies are: suction chamber-based, i.e., Dermaflex, Dermalab (Cortex Technology, Hadsund, Denmark) and Cutometer (Courage & Khazaka electronic, Cologne, Germany); torsion-based, Torque Meter (Dia-Stron Ltd, Andover, UK); and shear wave propagation method – Reviscometer RVM 600 (Courage & Khazaka electronic, Cologne, Germany).

14.2.3
Invasive Procedures

A number of invasive skin procedures have been applied in studying aging and to prove the efficacy of topical medications, cosmetics, and products. Skin biopsies can be evaluated histologically and immunohistochemically for the hallmarks of aging such as the degenerative changes in collagen and elastin, the reduction in the glucosaminoglycans, and the expression of antitumor protective molecules (e.g., p53). Additionally, electron microscopy can be performed on skin biopsies, for example, after stress tests to document the morphological bases of altered barrier homeostasis. Despite their precision, invasive procedures are time- and cost-ineffective as well as they raise ethical problems and inconvenience to the study subjects.

14.3
Study Design for Testing Antiaging Products

There is no generally accepted guideline on the performance of efficacy testing of antiaging products. An example content and format of a study protocol for cosmetic testing has been proposed [10]. A schematic overview of a study schedule is presented in Fig. 14.1.
Preselection and clear definition of the study population is required. Issues such as inclusion and exclusion criteria, prohibitions and restrictions, withdrawal and dismissal from the study must be addressed (described in detail in Chap. 8 of this book). As the objective in testing antiaging products is the improvement of the different parameters of the aged skin, the study population should include subjects possessing the typical characteristics of skin aging. The study group should be homogenous, i.e., similar stage/grade of aging. There is no defined number of panelists requested to perform an adequate study and the greater study population
facilitates the validation of the study results. The minimal number of participants must not be less than 12 as, if so, the statistical data evaluation may be hindered. An overrecruitment (initially performed by a phone call) should be considered due to the no-shows (minimum of 5–10%), drop-outs, and the ineligibility according to the inclusion/exclusion criteria.

### 14.3.2 Study Duration

As with the number of the participants, the duration of the study is not uniformly agreed. Having in mind the mechanisms and histological characteristics of aging skin, a prolonged period of product application is required. The majority of studies embrace a period of 30–60 days for the product application on a daily basis. However, the duration varies from 15 days to 1 year (as for the topical retinoids).

A washout phase before entering the study is required for the panelists. During this period, the application of any skin care, UV protection products, and topical medications must be avoided to the test sites. Personal hygiene procedures should also be standardized during this phase namely the number of bathing per day (most commonly restricted to once daily) and the use of preselected cleansers (e.g., mild soap provided by the study team). The duration of the washout period varies and 1 week is generally considered as sufficient not to interfere with the study preparations and procedures.

### 14.3.3 Test Site

The choice of anatomical test site is based on the targeted study objectives. If a product is intended to be applied on intrinsically aged skin, the test site should be confined to a generally photoprotected skin area, e.g., volar arm, volar forearm, buttock, and thigh. In contrast, facial area or the dorsal aspects of the arm are suitable sites for testing products targeting photodamaged skin. Crow’s feet is the most often involved anatomical site for antiwrinkle products testing in the facial area.

### 14.3.4 Product Application

The identically packaged containers with the separate products (test product and vehicle) should be coded with as little product information as possible. The label should contain the study number, the product code, and the panelist number or initials. The quantity of the applied product is calculated as a function of the test field surface and ranges from 1 to 3 mg/cm² (or 1–3 μL/cm²). The test product and the reference product (or vehicle-control) are gently rubbed with a glove-covered finger to the test field in order to avoid interference with the investigator’s sebum and sweat secretion. The application frequency (e.g., once or twice daily) is according to the product recommendation.
In the standardized manner, the half face (or one crow’s feet) is treated with the test formulation, and the contralateral (control) site is left untreated or a placebo (vehicle)/referent product is applied simultaneously. Before applying the formulation, a randomization of the treated site (left/right) is performed.

The last product application should be performed the evening before the measurement (at least 12 h). Otherwise, a risk of measuring the cosmetic remnants on the skin surface exists.

When the protocol of the study does not allow a 12-h period free of product application, cleansing of the skin might be useful by a mild, nonalcoholic, and nonaggressive cleanser at least 2–3 h before the measurement. Such a procedure should be performed for each test site and at each time point including the baseline measurement ($t=0$).

### 14.3.5 Efficacy Assessment

Efficacy assessment is performed according to the chosen endpoint, i.e., clinical evaluation, visual methods, biophysical measurements. A baseline measurement ($t=0$) is followed by measurements at defined time points, e.g., every week or every other week or monthly. A mean of multiple measurements (at least three) from each test site can eliminate outlying values. The consequence of measurements when using multiple devices should be in the order in which the measurement procedure affects the following one in a minimal way. For instance, if the protocol includes both SC hydration and skin surface pH measurement, the latter should be performed at the end, because the water ensuring the liquid media for the pH measure can alter SC hydration. Graduating the procedures according to their invasiveness is also useful, starting with the less invasive one. Different subject-, environment-, and instrument-related variables influence the measurements. As they are focused in other sections of this book, no detailed description is provided herein.

### 14.3.6 Data Management and Evaluation

The variable for statistical analysis is the relative reduction of the signs of skin aging (or the change of the investigated objective parameters) in the treated test area compared to the control area (vehicle, reference product, or untreated). Descriptive statistics including mean, median, standard deviation, minimum, maximum, and the 95% confidence interval are calculated. An experienced statistics specialist should be consulted in the data evaluation process as well as in the planning phase of the study.

### 14.4 Conclusion

Efficacy claim support is necessary for all cosmetic products manufactured and marketed in the European Union. Efficacy testing is mandatory not only for legislation reasons but also to demonstrate the superiority of one product to another. Different approaches to gain
an efficacy proof are used such as clinical evaluation, objective assessment with noninvasive methods and invasive procedures. The selection of the proper test method is based on the study objectives and in particular, on the aim of the investigation or the targeted endpoint.

References

Products for Impure, Acne-Like Skin

Hristo Dobrev

15.1 Introduction

Impure, acne-like (acne prone) skin is common for both males and females in their puberty and sometimes in adulthood. It is due to increased sebum secretion under the influence of male sex hormones, abnormal keratinization of the follicle and consecutive increased proliferation of the skin microflora (e.g., Propionibacterium acnes) and inflammatory responses. The skin looks greasy and shiny, rough with enlarged pores and has a greater tendency to develop comedones, pimples, and pustules [15, 21]. Most people feel the skin unpleasant and perceive it as a serious cosmetic problem. That is why the control over the impure skin requires daily application of multifunctional cosmetic products for cleansing and intensive care of the skin. Market products should have a proven effect. Testing on human volunteers using sensorial self- and expert evaluation, instrumental skin bioengineering techniques, and questionnaires for quality of life assessment are the preferred ways to prove product claims.

Core Message

Many people suffer from impure, acne-like skin. This type of skin looks greasy and glossy, rough with enlarged pores, and has a tendency to develop comedones, pimples, and pustules. It feels unpleasant and may be a serious cosmetic problem. The effective control over the impure skin requires daily application of multifunctional cosmetic products for cleansing and intensive care of the skin. Market products should have a proven effect. Testing on human volunteers using sensorial self- and expert evaluation, instrumental skin bioengineering techniques, and questionnaires for quality of life assessment are the preferred ways to prove product claims.
the cosmetic product, *where justified by the nature of the effect or product*” [8]. Testing on human volunteers, conducted in accordance with the rules of good clinical practice (GCP), is the preferred way to obtain proof of efficacy [19].

The aim of this book chapter is to outline the principles of efficacy testing of cosmetic products for impure, oily, and acne-like skin.

### 15.2 Test Settings

#### 15.2.1 Study Background

The cosmetic testing should be preceded by a detailed review of the problem that is the basis of the research. The investigator must examine all available previous data and clarify the study nature and rationale. Reference to literature is also helpful in discussing the findings of the test [31].

#### 15.2.2 Study Objectives

Generally, the study objectives (study end points) are identical with the efficacy claims of the product tested [31]. Cosmetic products for impure skin are aimed to clean and care the skin. Their cosmetic claims may therefore include cleansing, degreasing, controlling shine, restoring the balance of oily skin, cleaning pores, removing excessive oil and unwanted microorganisms as well as clarifying, reducing blemishes, and improving the overall skin health of acne prone skin. Antiacne, anticomedonal, and antibacterial (anti-inflammatory) action are considered drug claims.

Cosmetic testing may have one primary study objective that usually covers the main claim of the product and a number of secondary objectives. The study objectives can be quantitative and qualitative as well as direct or surrogate that provides indirect evidence on the postulated claims.

The combined assessment of study endpoints using ordinal global assessment scale, acne type lesion counts, and instrumental assessments of relevant skin parameters allows for a balanced approach toward the evaluation of impure, acne-like skin severity [14].

#### 15.2.3 Material and Methods

##### 15.2.3.1 Study Products

Depending on the study nature, test products may include investigational product (active product designated to improve impure skin), nonactive control product (placebo), and/or
active control product (comparator). The investigational product must be identical to the final product to be marketed. A placebo formulation without active ingredient(s) is the preferable method for comparison. If this is impossible, a marketed neutral product should be used. The investigational product could also be compared with other active reference product with the same claimed effect. It is very important for all tested products to possess similar physical properties (viscosity, color, fragrance) [31].

Investigational products should be manufactured, handled, and stored according to good manufacturing practice of cosmetic products (GMPC) and used according to the approved protocol [9, 31]. The investigator should provide a brief description of the test products including the name, manufacturer, composition, formulation, quantity, and mode of application. Methods of product management such as packaging, batch number, expiry date, receiving, storage conditions, dispensing and return, and if applicable, product blinding (labels, codes) and assigning subjects to treatment groups (randomization codes, patient identifier) should also be described.

After the test, a sample of the product tested and the reference product should be retained for at least 6 months under suitable conditions by the investigator and/or the promoter [7].

15.2.3.2 Study Subjects

It is very important that the study subjects realistically represent the range of anticipated users of the product tested with respect to sex, age, and skin condition. The study population usually includes healthy volunteers with impure, acne-like (acne-prone), and oily skin and occasionally patients with mild to moderate acne. The information that should be considered includes [10, 31]:

- Subject’s number and demographic data such as age, sex, and race.
- Subject’s recruitment manner. All recruitment methods and materials (press and web advertisements, letters and electronic mail to subjects, visitors of dermatology clinic) should be reviewed and approved by the independent review board (IRB).
- Criteria for inclusion. Specific requirements for subject’s age, skin type, presence and severity of acne type lesions, and duration of preceding washout period may be defined. Generally, the subjects are submitted to preliminary skin type determination by means of specific questionnaire, clinical examination, and noninvasive instrumental measurements.
- Criteria for exclusion. They may include the presence of severe acne, rosacea, pregnancy, breastfeeding, known allergy to constituents of the test products, and systemic disorders.
- Criteria for early withdrawal. If any predetermined reasons for removing subjects from study exist, they should be defined.
- Preexperimental restrictions. They are usually related to skin care regimen allowed and the prohibition for use of particular topical products on the test sites.
- Training of the subjects. If any preliminary training of the study participants has been carried out, it should be specified.
15.2.3.3 Study Design

The experimental design of cosmetic testing studies must be scientifically appropriate and suitable to prove study objectives. The main points that should be considered are [7, 14, 15, 31]:

- Use of control/comparison group (noncontrolled or controlled study; single or comparative study). The comparison can be performed intraindividually (for example half-side test) or interindividually. On the other side, active treated areas can be compared with untreated or nonactive (vehicle or placebo/active (comparator product) treated areas.
- Use of blinding (open-labeled or single-/double-blinded study).
- Method of assignment to treatment groups (nonrandomized or randomized study).

The open-labeled, noncontrolled study with comparison before/after is easiest to perform, but is not considered enough indicative. The randomized controlled trial is the gold-standard design for cosmetic products testing and should be used whenever applicable.

The other aspect of cosmetic testing design is the study duration [10]. Efficacy of cosmetic products can be evaluated either by short-term or long-term studies. Short-term testing with a single application of the product is less sensitive to changes in the environment, while the long-term testing with multiple application of the product is more realistic, reflects daily life situations, and permits a subjective evaluation by the subjects. It is observed that the single application test is often predictive of the data outcome seen in long-term application test. The choice between short-term and long-term cosmetic testing is also dependent on the product claim to be studied. The claim of an impure skin-cleansing effect can be evaluated after a single application of the product tested, whereas the claim of an impure skin care requires evaluation after longer (usually 8–12 weeks) application of study product. In acne treatment studies, a minimum duration of 12 weeks as well as a posttreatment follow-up period is recommended to demonstrate efficacy and to evaluate recurrences following treatment discontinuation [14].

Cosmetic testing may be performed at single (single center study) or several (multi-centre study) testing laboratories.

15.2.3.4 Study Methods

The cosmetic testing protocol must define the evaluation parameters (test variables) and the methods and devices used for their assessment. Evaluation parameters should be related to the characteristics of impure, acne-like, and oily skin, and the postulated effects of the product investigated. Besides the efficacy criteria, product acceptance and unwanted effects can also be assessed. Generally, evaluation methods may be divided into two main categories [7, 31]:

- **Sensorial tests.** They are based on appreciation of skin appearance and product performance made through the senses (sight, touch, and olfaction) of consumers them-
themselves or trained experts. They give information mainly about observed or perceived parameters and are very subjective.

- **Instrumental tests.** They are performed with instrumental methods and devices that can precisely measure given skin structure and function parameters, according to a defined protocol, following the application of a product on human subjects. The measurements are carried out in controlled laboratory conditions by educated and trained operators. It is very important for the test and parameters measured to be relevant to the effect of the product and claim. The introduction of noninvasive biophysical methods offers real progress in the objective and quantitative evaluation of dermato-cosmetic products. These techniques can also be used to substantiate the final product as a stand-alone test.

Lately, more often, **quality of life questionnaires (QoL)** are used to evaluate the influence of product application on the subjects’ quality of life.

**Oily Skin Assessment**

Sebaceous gland function and oily facial skin can be evaluated using subjective and objective, quantitative and qualitative, static and dynamic, morphologic and functional, direct and indirect, convenient and cumbersome methods [3, 12, 13, 26].

**Sensorial Assessment**

Sensorial assessment of oiliness of the skin of the forehead, nose, chin, and cheeks is based on visual and tactile appreciation, how shiny the skin appears, and also by the presence of large (open) pores and follicular plugs in these areas. It can be performed in the form of auto evaluation (self-assessment) and investigator (expert) evaluation using clinical examination [7, 26].

**User-self assessment** gives information about the initial degree of skin oiliness, the improvement in the severity of oily skin condition after treatment, and the cosmetic properties, acceptance, and performance of the product.

Subjects can evaluate the degree of their skin oiliness on a four-point scale (absent, mild, moderate, or severe) or on a visual analog scale of 0 to 10 (0 – normal skin; 10 – very severe oiliness).

Sensorial assessment made through the **use test** evaluates subjects’ perception of product cosmetic properties and efficacy based on parameters that they can observe or feel. Specific questionnaires and scoring systems using numerical or rating scales are applied for auto-evaluation. The use test may be performed as blind use test without providing any product information to consumers or concept use test combined with elements of communication between consumers and investigator [7].

**Investigator assessment** also gives information about the initial degree of skin oiliness and the product efficacy. It should be carried out at least 2 times – before and after the study completion. Generally, a five-point scale (normal skin, slight, easily visible, obvious, and extensive oiliness) is used and the changes in mean scores are determined. Product efficacy can be rated both by subjects and investigator on a four-point scale (0= no change to 3= very good effect or marked improvement) [7, 26].
Instrumental Assessment

Instrumental assessment is performed with noninvasive biophysical methods and devices that can precisely measure parameters related to sebaceous glands function. Measurements are made on the facial skin as the forehead is the mostly used test site. The usually determined variables are [3, 12, 13, 26]:

- **Sebum casual level (SCL).** It represents the spontaneous layer of lipids on the skin surface. SCL is considered to be a constant value for each normal adult and an estimate of the skin greasiness. It has been reported that the time needed to recover CL after removing sebum from the skin surface is about 4 h in subjects with normal sebum excretion. Therefore, SCL is measured after at least 4 h following uncontrolled removal of the sebum film from the skin surface. During this time the skin should remain untouched. The values of SCL are expressed as the amount of sebum (lipids) per unit skin surface area.

- **Sebum excretion rate (SER).** This is the amount of sebum excreted during a defined period of time over a given skin area expressed in \( \mu g/cm^2/h \). SER is usually determined on the forehead during an hour after preliminary skin defatting. It is considered that the SER is related to the delivery of the pool of sebum already secreted and stored in the outer portion of the pilosebaceous duct, that is, it is a measure of sebum excretion rather than sebum secretion. SRT is used as an indicator and quantifier of the activity of the glands.

- **Variables related to active sebaceous glands** (number, pore size, area, density, distribution, level of activity)

- **Skin gloss.** It is an indirect indicator of skin greasiness.

The methods and devices most often used are [3, 12, 13, 26]:

- **Photometric techniques.** The measuring principle is based on the fact that opaque glass (Lipometer, L’Oreal, France; not commercialized) or plastic film (Sebumeter, Courage+Khazaka, Cologne, Germany; www.courage-khazaka.com) become transparent when their surface is covered by lipids. Using these techniques, SCL and SER can be determined depending on the presence or lack of preliminary defatting of the skin test area. The amount of lipids collected for 30 s is displayed as \( \mu g/cm^2 \). It is considered that these techniques collect all the sebum present on the stratum corneum including the follicular reservoir and the interfollicular skin surface.

- **Lipid absorbent tapes.** They are special white microporous, nonadhesive or adhesive, tapes that absorb the sebum originating from the follicular openings, which forms transparent spots easily visible on a dark background. The number and the size of the spots indicate the active sebaceous glands (identifies the sebum-poor and the sebum-rich follicles) and the area covered by them is proportional to the collected volume. It is considered that these techniques are indicative of sebum excretion from the infundibular reservoir, that is, the tapes absorb only the sebum present in the upper part of the infundibulum. The tapes can be evaluated by visual scoring or by image analysis that allows several parameters to be analyzed (the percentage area covered by oily spots, the sebum area in square micrometer, and the number of oily spots). Depending on the absorbent tape used, the skin may be (Sebutape adhesive patches, CuDerm, Dallas, USA; www.cuderm.com) or not defatted (Sebufix F16, Courage+Khazaka, Cologne,
Germany; www.courage-khazaka.com) prior to the tape application and the collection time may be between 30 s (Sebufix F16) and 1 h (Sebutape adhesive patches)

- **Image analysis techniques.** The aspect of oily skin surface and follicular openings can be directly examined, recorded, and analyzed using a regular videomicroscopy and/or video cameras working under ultraviolet light illumination (Visioscan VC98, Courage+Khazaka, Cologne, Germany; www.courage-khazaka.com). Visioscan is also used for image analysis of Sebufix tapes.

- **Skin glossymetry.** Nowadays, it becomes possible to measure skin gloss with noninvasive and convenient bioengineering devices. The first one (Skin-Glossymeter GL 200, Courage+Khazaka Electronic, Cologne, Germany; www.courage-khazaka.de) measures both the portion of directly reflected (which is related to the gloss) and the scattered LED white light from the skin surface. The skin gloss is expressed by two parameters – skin gloss and skin gloss with diffuse scattering correction. The second device (SkinGlossMeter, Delfin technologies Ltd, Kuopio, Finland; www.delfintech.com) measures the specularly reflecting red light (generated by built-in 635 nm red semiconductor diode laser) from skin surface. The gloss values are measured with a photodetector and the total intensity of the reflected beam is then calculated.

**Quality of Life Assessment**

The excessive oily skin is a cosmetic problem that often may cause significant concern for people who have the condition. Recently, two specific instruments for assessment of adverse psychological and social effects of oily skin were developed. The oily skin self-image questionnaire (OSSIQ) is a concise 18-item questionnaire designed to assess perception, behavioral, and emotional consequences associated with oily skin condition and to assess the effects of a skincare treatment for oily skin [30]. The second one is the oily skin impact scale (OSIS) that was introduced by Arbuckle et al. and is designated to measure the impact of oily skin on emotional well-being [4]. The last authors have additionally developed a 26-item questionnaire for a measure of oily skin severity named oily skin self-assessment scale (OSSAS). Both the OSSIQ and OSIS instruments can be used to monitor the benefits of cosmetic skincare treatments.

**Acne Lesions Assessment**

Acne lesions are divided into several types [2, 14]:

- Noninflammatory lesions – open and closed comedones, and uninflamed nodules (sometimes called cysts)
- Inflammatory lesions – papules, pustules, inflamed nodules, and cysts (>5 mm in diameter)
- Secondary lesions – excoriations, erythematous and pigmented macules, scars.

Since cosmetic products applied can be effective in different ways on the particular acne type lesions, it is recommended that they are evaluated separately [16].

Acne lesions occur predominantly on the face and it is the most examined area. To improve the reliability of counting, the evaluated face area may be subdivided into several
areas, for example, entire forehead (or forehead right and forehead left), right cheek and left cheek, chin, nose, and close areas, (or nose and chin) and the upper part of the neck, each being examined separately. Acne lesions may also affect the nonfacial sites such as the chest, back, and shoulders, which may be examined as well. In efficacy studies, if a particular area is not counted at baseline (or is scored 0) that area should not be counted at subsequent evaluations [23].

Sensorial Assessment
Sensorial assessment of acne lesions is based on visual and tactile appreciation and can be performed in the form of autoevaluation (self-assessment) and investigator (physician) evaluation using clinical examination [7].

*User-self assessment* gives information about the initial degree of acne lesions, the improvement in their severity after treatment and the cosmetic properties, acceptability and performance of the product.

Subjects can evaluate the extent of their acne lesions on different scales: four-point scale (none, mild, moderate, and severe), five-point scale (none, minimal, mild, moderate, and severe) or visual analog scale of 0 to 10 (0 – no acne; 10 – very severe acne) [23].

*Sensorial assessment* of product performance is made through the use test that evaluates subjects’ perception of product cosmetic properties and efficacy based on parameters that they can observe or feel. The autoevaluation is made through specific questionnaires and scoring systems using numerical or rating scales. Participants can rate the overall improvement in their acne on a six-point Likert scale with the following categories: worse, no improvement, slight improvement, moderate improvement, excellent improvement, and completely cleared [7].

*Investigator (physician) assessment* also gives information about the initial degree of acne lesions (baseline acne severity) and the efficacy of product applied (change from baseline in acne severity). It should be carried out at least 2 times – before and after the study completion, but generally it is made more often (usually at 2-week intervals). The duration of the study is usually 8 or 12 weeks.

The investigator assessment of the severity of acne is typically based on qualitative global assessment of severity (visually or using photographs), preferred for clinical practice, and lesion counting, preferred for clinical investigations [1, 2, 5, 6, 14, 18, 22, 27, 28, 34]. Table 15.1 summarizes the methods for acne grading. Although plenty of acne grading systems currently exist, there is no universally recognized and standardized system.

The simplest way of grading acne is based on the predominant type of lesion present on the skin, regardless of number. One can therefore grade acne as follows [32]:

- Grade 1: Comedones only
- Grade 2: Inflammatory papules present in addition to the comedones
- Grade 3: Pustules present in addition to any of the above
- Grade 4: Nodules, cysts, conglobate lesions, or ulcers present in addition to any of the above

The most often used methods for acne grading are the Leeds acne grading scale and the combined acne severity classification.
The Leeds Acne Grading Scale [6, 18, 22] is the most widely used qualitative photonic grading system. It is based on a large study of 435 patients with acne of varying severity and is composed of 16 facial, 8 chest, and 8 back severity categories. The patient’s acne is graded separately for the face (including the neck), chest, and back, because at these sites, acne frequently varies in severity and treatment responses. The acne (inflammatory) lesions are inspected visually and palpated to detect nodules or cysts. Then they are compared with a set of standard photographs of patients with acne of varying severity to determine the grade, which may be from 0 to 10 (severe nodulocystic acne) in original version and to 12 (exceptionally severe acne) in revised version. In addition, grades 0 to 2 are subdivided by 0.25 increments while grades 2 to 10 are subdivided by 0.50 increments.

The Combined Acne Severity Classification [20, 34] employs lesion counting combined with some type of global assessment of severity and comprises three categories:

- Mild acne: fewer than comedones, or fewer than 15 inflammatory lesions or a total lesion count lower than 30.
- Moderate acne: 20–100 comedones, or 15–50 inflammatory lesions or a total lesion count of 30–125.
- Severe acne: more than five cysts, or comedone count greater than 100, or a total inflammatory count greater than 50, or a total lesion count greater than 125.

### Table 15.1 Disease-specific instruments for assessment of acne

<table>
<thead>
<tr>
<th>Qualitative methods (global assessment)</th>
<th>Description</th>
</tr>
</thead>
</table>
| Visually scoring | The facial acne scale (Allen and Smith 1982 [1])  
| | Global severity scoring of the AAD (Pochi et al. [27])  
| | Quantitative assessment over the whole acne area (Lucky and Beth 1996 [1])  
| | ECLA (Echelle de Cotation des Lesions d’Acne or Acne lesion score scale) (Dreno et al. 1997 [2])  
| | Global acne grading scale (GAGS) (Doshi et al. 1997 [34])  
| | Combined acne severity classification (Lehmann et al. [20])  
| | Investigator global assessment (IGA) (Dermatologic and ophthalmic drugs advisory committee briefing, FDA, November 2002 [16])  
| | Simple grading of acne (Consensus document of the Global Alliance to improve outcomes in acne 2005 [32])  
| | Comprehensive acne severity scale (CASS) (Tan et al. [35])  
| | ECCA grading scale (Dreno et al. 2007) – for acne scars [34] |

<table>
<thead>
<tr>
<th>Scoring using photographs</th>
<th>Description</th>
</tr>
</thead>
</table>
| Acne grading method using photographs (Cook 1979 [1])  
| Leeds acne grading scale (Burke and Cunliffe [6])  
| The revised Leeds acne grading scale (O’Brien et al. [22])  
| New photographic techniques for clinical evaluation of acne (Rizova and Kligman [29]) |

<table>
<thead>
<tr>
<th>Quantitative methods (lesion count)</th>
<th>Description</th>
</tr>
</thead>
</table>
| Half-face grading (Plewig and Kligman 1974 [2])  
| Michaelson acne severity score (Michaelsson et al. 1977 [1]) |
The investigator evaluation of product efficacy is based on the changes in global severity score and reduction in acne lesions count (separately in the count of inflammatory and noninflammatory lesions as well as in total lesion count) toward the baseline. The subject response to treatment can be rated on a four-point scale (0 = no to 3=very good effect) or on a five-point scale (0 = worse to 4 = marked improvement) [16].

A Physical Global Improvement score can also be determined at the final visit – for example, the subjects can be judged to be almost clear (90–99% clearance), marked improved (75–89% clearance), moderate improved (50–74% clearance), minimal improved (25–49% clearance), or judged to have no change (0–24% clearance) from baseline [17].

Recently, the FDA agency recommended an investigator global assessment (IGA) ordinal scale with five severity grades (0–4) to be used in efficacy studies. Each grade is defined by a distinct and clinically relevant morphologic description that minimizes interobserver variability (Table 15.2). For assessment of efficacy, the success is defined using one of the following two criteria: (1) as “clear” (grade 0) or “almost clear” (grade 1) or (2) as improvement of two grades from baseline score at the prespecified primary time point. In addition, noninflammatory and inflammatory acne lesion counts must be separately obtained. The Agency recommended that each subject’s improvement be verifiable via photographic records of baseline and assessment time points [14].

Table 15.2 IGA scale for acne vulgaris

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Clear skin with no inflammatory or noninflammatory lesions</td>
</tr>
<tr>
<td>1</td>
<td>Almost clear; rare noninflammatory lesions with no more than one small inflammatory lesion</td>
</tr>
<tr>
<td>2</td>
<td>Mild severity; greater than grade 1; some noninflammatory lesions with no more than a few inflammatory lesions (papules/pustules only, no nodular lesions)</td>
</tr>
<tr>
<td>3</td>
<td>Moderate severity; greater than grade 2; up to many noninflammatory lesions and may have some inflammatory lesions, but no more than one small nodular lesion</td>
</tr>
<tr>
<td>4</td>
<td>Severe; greater than grade 3; up to many noninflammatory and inflammatory lesions, but no more than a few nodular lesions</td>
</tr>
</tbody>
</table>

The investigator evaluation of product efficacy is based on the changes in global severity score and reduction in acne lesions count (separately in the count of inflammatory and noninflammatory lesions as well as in total lesion count) toward the baseline. The subject response to treatment can be rated on a four-point scale (0 = no to 3=very good effect) or on a five-point scale (0 = worse to 4 = marked improvement) [16].

A Physical Global Improvement score can also be determined at the final visit – for example, the subjects can be judged to be almost clear (90–99% clearance), marked improved (75–89% clearance), moderate improved (50–74% clearance), minimal improved (25–49% clearance), or judged to have no change (0–24% clearance) from baseline [17].

Recently, the FDA agency recommended an investigator global assessment (IGA) ordinal scale with five severity grades (0–4) to be used in efficacy studies. Each grade is defined by a distinct and clinically relevant morphologic description that minimizes interobserver variability (Table 15.2). For assessment of efficacy, the success is defined using one of the following two criteria: (1) as “clear” (grade 0) or “almost clear” (grade 1) or (2) as improvement of two grades from baseline score at the prespecified primary time point. In addition, noninflammatory and inflammatory acne lesion counts must be separately obtained. The Agency recommended that each subject’s improvement be verifiable via photographic records of baseline and assessment time points [14].

Instrumental Assessment

Besides the measurement of sebum excretion, acne prone skin can be evaluated with other methods such as polarized light photography [25], digital fluorescence photography [24, 33, 38], videomicroscopy [1, 29], and ultraviolet light illumination [37].

Polarized light photography enhances visualization of inflammatory acne lesions and permits accurate evaluation of the extent of disease and the effectiveness of therapy.

For acne patients, the orange-red fluorescence of facial ultraviolet photographs is thought to be due to the porphyrins produced by *P. acnes* found in the pilosebaceous unit. In this connection a commercial device (Visiopor PP34, Courage+Khazaka, Cologne, Germany; www.courage-khazaka.com) was developed. The camera uses a specific UV-light to show the microcomedones and comedones in the skin which become visible as fluorescent orange-red spots in the pores. Studies have shown that the image analysis of porphyrin fluorescence correlates well with the decrease in *P. acnes* density from scrub cultures [24]. Recently, using a
UVA-induced digital facial fluorescent imaging system and image analysis method, Youn et al. [38] have found that the comedonal red fluorescence area shows a stronger correlation with sebum secretion than with the presence of *Propionibacterium acnes*. This finding suggests that the red fluorescence is affected by sebum not just *P. acnes*.

Uhoda et al. [37] have used a videocamera equipped with an internal ultraviolet illumination source (Visioscan, Courage+Khazaka www.courage-khazaka.com, Cologne, Germany) and image analysis software for determination of follicular plugs and product comedolytic activity.

**Quality of Life Assessment**

Impure skin affected by acne has a significant effect on the subject’s quality of life, which can be evaluated by specific questionnaires filled out by the subjects before and after treatment [34, 39].

The first group includes dermatology-specific QoL instruments such as the Dermatology Life Quality Index (DLQI), Children’s Dermatology Life Quality Index (CDLQI), and Skindex-29.

The second group includes acne-specific QoL instruments such as Acne Disability Index (ADI), Cardiff Acne Disability Index (CADI), Acne-Specific QoL Questionnaire (Acne-QoL), four-item condensed version of the Acne-QoL (Acne-Q4) [36], nine-item Acne QoL scale (AQLS) [35], and the most recently introduced instrument assessment of the psychological and social effects of acne (APSEA) [34].

Measuring the impact of acne on quality of life allows us to understand the disease from patient’s point of view. In clinical research, new medications are increasingly being evaluated according to their impact on QoL, which is an addition to the traditional approach of assessing only treatment safety and efficacy. Besides the lack of strong association between acne severity and QoL, several studies have shown improvement in QoL with effective acne treatment [39].

**Product Acceptance Assessment**

Product acceptance is related to physical properties (consistency, color, fragrance) and cosmetic properties (spreadibility, permeability, and fixation) of the product. They are subjectively evaluated using specialized questionnaires filled out by subjects at the end of cosmetic testing. The answers can be rated on a four-point scale (0 = none to 3 = very good) or on a five-point scale (0 = none to 4 = excellent).

To measure subject’s satisfaction with the cosmetic product properties, a five-point (very satisfied, satisfied, neither satisfied nor dissatisfied, dissatisfied, very dissatisfied) or a six-point (extremely satisfied, very satisfied, somewhat satisfied, somewhat dissatisfied, very dissatisfied, extremely dissatisfied) Likert scale or a visual analog scale of 0 to 10 (0 = not satisfied to 10 = very satisfied) can also be used.

Finally, the overall product approval generates the subject’s willingness to purchase or not the product tested.

Recently, some new techniques were introduced in order to improve the reliability of the sensory evaluation of consumer products. The objective emotional assessment (OEA) technique is based on the evaluation of psycho-physiological reactions and parameters and
had been proven to be highly suitable for determining emotional consumer response [11].
With the Sensorimeter SR 100 (Courage+Khazaka, Cologne, Germany; www.courage-
khazaka.com), the experienced properties of the product are determined and combined
with objective measurements of the skin after applying a product.

Assessment of Unwanted Effects

Generally, the efficacy testing is performed when there is already evidence that the tested
product does not cause local or systemic adverse responses. However, due to the existence
of a widespread variation in sensitivity within the human population, undesirable and side
reactions may be sometimes observed. When this happens, it should be recorded and used
as complementary information for the safety assessment of the product [10].

Features of local irritation (erythema, scaling, itching, burning) are evaluated visually
and through the clinical examination. They can be scored by the participants using a
10-mm visual analog scale (from 0 = none to 10 = extremely noticeable) and by investiga-
tors using a four-point scale (absent, mild, moderate, severe).

15.2.3.5
Study Protocol

The study protocol should describe the following points [7, 10]:
- Test schedule (timetable), that is, the sequence and duration of all study periods
- Test procedures that will be accomplished at each visit – examinations and
  measurements
- Product application
  - Amount of product applied (usually 1–2 mg/cm² skin).
  - Frequency of use (usually one to two applications a day).
  - Time of product application (usually morning and/or evening).
  - Application areas (measuring test sites). For identification and marking of the appli-
    cation sites, a template is preferably used.
- Environmental conditions (usually 22–24°C and 50–60% relative humidity). Temperature
  and relative humidity of the room must always be noted.
- Preconditioning. Physiological conditions related to the subject should be respected:
  minimum of 15 min acclimatization to the room temperature, rest, relaxation, thermal
  comfort, absence of sweating.
- Prior and concomitant skin care products/medicines allowed (restriction for use).
  Participants were usually asked to use only regular skin cleansing products during the
  study period. In addition, they are instructed not to use any cosmetic products on the
  study day and not to wash within 3 h of measurements.
- Pretreatment (define and set out the chronology if necessary).
- Subject compliance monitoring.
15.2.3.6 Study Ethics

The cosmetic testing study must be conducted in accordance with the ethical principles originated from the Declaration of Helsinki including written information and signed informed consent from the participants and the approval by an Independent Ethics Committee (IEC)/Institutional Review Board (IRB). The compliance with the ethical standards and the ICH guideline for GCP provides public assurance that the rights, safety, and well-being of trial subjects are protected and that the clinical trial data are credible [31].

15.2.3.7 Statistical Analysis

The management of data obtained from cosmetic testing is submitted to the general rules for statistical analysis. The investigator should carefully consider the following points: sample size (groups larger than 20 are recommended), variables analyzed, randomization and blinding of observations, comparison grouping, expression of results, data processing, statistical tests, and computer software used [7, 10, 14, 31].

15.3 Conclusion

People with impure, acne-like skin do not always consult with dermatologist and usually start with application of cosmetic products. That is why these products should have a proven efficacy. Any product claim that can be verified and measured must be documented in a randomized controlled trial in humans who are a representative sample of the anticipated user group. The new instrumental techniques provide the cosmetic industry with the possibility of demonstrating the actual ability of their products. For the consumers, there is an opportunity to choose the relevant product for their skin on the basis of evidence and not on the basis of promises.

15.4 Key Messages

- Testing of cosmetic products for impure skin is ruled by the principles of GCP.
- A randomized controlled trial on humans is the gold standard to be used whenever applicable.
- Study objectives (study end points) cover the claims of the product tested.
• Test subjects must be a representative of the anticipated user group.
• Test product must be identical to the final product to be marketed.
• Sensory self- and expert evaluation and instrumental skin bioengineering techniques are the preferred ways to prove product claims.
• Effective treatment is associated with significant improvement in the quality of life.

References


Assessment of Hair Morphology

H. Zahouani and M. Fougère

16.1 Introduction

Nowadays, hair is an important part of our appearance and reflects our physical state. Thus, the development of hair care products has greatly increased since 15 years and so has the research. To improve products like shampoo or conditioners and avoid complex sensitive studies, cosmetic companies and scientists designed methods of evaluation based on scientific values.

Mostly, these methods are focused on the cuticle characterization. This outer part of 3 to 5 μm in height and constituted of flat overlapping cells of 50 μm in length protect the inner part, the cortex, from external environment. The cortex is a complex polymer structure for which the main constituent of the matrix is the alpha keratin, a protein with high sulphur content. This part, about 90% of a hair, gives it the longitudinal mechanical properties.

The mechanical and chemical solicitations occurring during daily washing, combing, or drying induce cracks and breaks on the cuticle. Besides the cortex mechanical properties attenuation, it also decreases the brightness and softness. Using hair care or hair styling
products reduce significantly the degradation of brightness and softness acting on topography and surface chemistry of the cuticle. This work will especially focus on the first aspect.

The quantification of the sensitive modifications, connected to the cuticle topography, is an important stake to increase hair care product performances. The measurement of the cuticle topography needs the use of specific instruments with a sufficient resolution to observe elements at nanometric scale. The scanning electron microscope (SEM) and the atomic force microscope (AFM) are commonly used for this application. Swift and Smith developed these two measurement methods [4–6]. The two techniques use different principles of operation and present differences in the results.

The AFM reconstructs a 3D-image of the surface sample based on the x, y, and z coordinates of each point of the surface. These sets of points can be analyzed with dedicated surface topography software to examine the surface with quantitative data. Moreover, AFM can work in hair without any specific preparation of the sample and in non contact mode, even avoid contact between the tip probe and the sample surface. Therefore, any possible contamination or alteration of the surfaces is prevented.

By contrast, the SEM gives a “photographic” image and no quantitative data. It requires the hair sample to be covered with a very thin layer of a conductive material and the measurement done in a vacuum chamber. Moreover, metallization of hair altered and modified the hair surface, increasing the scale height and the roughness in the scale profiles [3].

For the reason enounced previously, AFM seems to have a huge interest to study hair cuticle and especially to examine fibres before and after various chemical and physical treatments [1]. This comparison of the same area is a key point to reduce the number of samples examined: hair fibres show numerous topographical, mechanical, and chemical differences along the length of each hair [5] and the only other method of obtaining adequate sampling is to look at statistical differences, which may be very subtle, over many hair samples.

There is a third technique that could be applied for such samples with comparison before and after treatment which has not been reported in the literature: Optical Interferometry. This non contact method that quickly determines a three dimensional surface shape can replace AFM for the cosmetic applications and be the unique solution for others. It is described in this paper.

16.2 Material and Methods

16.2.1 Optical Interferometry

The interferometer (Veeco Instruments, Wyko, Cambridge, UK), is an optical device designed to measure the topography of surfaces. A white light source is split into two separate beams: one is reflected on the surface studied before touching the CCD detector and the other is reflected from the reference mirror to the CCD detector. The beams recombined create white and dark bands called fringes that make up an interferogram. Fringes can be associated to lines on topographic map. To increase the measurement amplitude, the microscope is connected to a piezo-electric translator for which the displacement is controlled by a micro-calculator.
The data collected with this method are point coordinates on x, y, and z axis. These points permit a three-dimensional reconstruction of the surface. For hair imaging, the Vertical Scanning Interferometry mode is used with a magnification of 73.4 and the image obtained measure 80 µm lengths by 60 µm height. This mode, with a resolution of 2 nm, is particularly suitable for rough surface up to 1 mm amplitude. The image obtained is composed by different wavelength scales: high frequency for roughness, medium for waviness, and low for shape. To study specifically, the roughness modification, a method based on multi scale decomposition has been developed [2, 7]. This approach allows continuously following the evolution of a surface for the different wavelength family, from the roughness to the shape, determining a mean amplitude criteria for each wavelength [8]. Then using a parameter called transfer function based on the comparison of two cuticle states (after and before treatment), it is possible to quantify the impact of each treatment on hair surface. The surface roughness has also an impact on its brightness and on the way for cuticle to reflect the light. When a light beam impacts a perfectly smooth surface, it is completely reflected; for a rough surface, there is a dispersion of the beam: The more the light is reflected in the same direction the more the surface is bright. Thus to translate the relation between the roughness and the brightness of a surface, a quantifying method by a local normal is used. The image of the cuticle is decomposed in elementary areas. Then the slope of each area is calculated in relation with the local normal. A surface presenting a local normal equal to 90° is very smooth and bright. Thus, a decrease of the value of the local normal influences the smoothness and brightness in the same way.

16.2.2 Static and Dynamic Experiments

Two types of experiments are done with the interferometer: static and dynamic. Static tests consist of measuring the same hair area before and after treatment. In the present case, the sample is straightened to evaluate the smoothing performance of a series of straighteners by quantifying the modification on the cuticle. The fibres used for this study are taken from untreated black Caucasian tresses and stored at 20°C and 30% RH. They are cleaned in a 1% aqueous solution of sodium dodecyl sulphate, rinsed with distilled water and then dried at room temperature. The sample is a 20-mm long part from a hair mid-length, fixed on a piece of rigid paper.

In this position, the sample can neither rotate nor translate and is incised at different points. This small incision, perpendicular to the hair axis, allows recognizing the area to be measured. It resists to mechanical, chemical, and thermal solicitations. The hair is then mounted on the interferometer to be observed, then straightened twice at a well-known temperature between 150 and 200°C and observed again after the treatment.

In dynamic mode, a pull-push machine is placed under the interferometer and the hair topography is measured without removing the hair from the machine. This apparatus allows associating the topographic modifications with a pull test. In this case, the same area is measured also between the different operations.

The fibres used in this dynamic mode are identical to the static mode. The main difference is the way in realizing the position marker and the sample size which is about 40 mm.
The incision of sample would yield the sample and cannot be used in this mode. A new solution, using small droplet of silicon has been introduced.

The pull-push machine has been specifically designed for the fibre traction test.

16.3 Results and Discussion

16.3.1 Static Experiments

The image of the cuticle (Fig. 16.1) proves the interest in the interferometry for hair surface measurement. It combines the advantages of AFM, by imaging a surface precisely without contact and SEM, by displaying a large scan area, without metallization. It could be argued that AFM provides a better precision than Optical Interferometry, but the acquisition time and the set time for an image with the same dimensions is more than 10 times faster with the interferometer. So one can speculate how a cosmetic product evolves on the cuticle in time, and this reduces the advantage of the AFM precision. Figure 16.1 shows images of the same area of the hair fibre measured before and after a treatment with a straightener. It is possible, looking at the points 1, 2, 3, and 4, to recognize the same scales on each image. Nevertheless these topographical representations show little differences. The surface has been smoothed and the roughness has decreased. For the same Z axis, the contrast of the picture after the treatment is lower than before, which confirms the reduction of roughness.

The data from the acquisition are used to calculate the variation of the mean roughness. The wavelet transform allows separating the shape, the waviness, and the roughness and then to extract only the information connected to the roughness [8]. The impact of the treatment is finally determined by a following transfer function.

The quantification of the straightener performance is obtained from the previous formula. Thus it is possible to classify these apparatus. A study on ten of them gave a 20% more modification for the best one and less than 5% for the worst. On the same images, we can determine the brightness criteria that gave for the same study, a value higher than 25% for the best and less than 3% for the worst.

In some cases, the treatment that normally reduces the roughness could cause damage. The main damages observed on the image are scratches. A specified study on the scratches has been performed to understand the behaviour on time of such deformations. The sample was left on its holder at room temperature (24°C and 30% RH) and observed during 24 h to know if the surface modifications were reversible or not.

The image and the profile, shown in Fig. 16.2, demonstrate a low evolution between the topography 10 min and 24 h after the treatment. It seems that there is a small viscous recovery of 0.15 μm and a plastic deformation.

The same sample is used to study the influence of water on the recovery of scratches. The sample mounted on the holder is immersed in deionised water by a droplet of 5 μL. A part of the water is absorbed by the hair and other is evaporated. New measurements with interferometer are done after the complete disappearance of the droplet. The image and the
Assessment of Hair Morphology

Fig. 16.1 Comparison between the same area (a) before and (b) after the treatment

profile, Fig. 16.3, show the recovery of the cuticle after 10 min, the water go through the outer cells of the cuticle and after 2 h, water keeps evolving in the cuticle layers penetrating into deeper parts.

This last experiment demonstrates the ability of hairs to recover their surface shape after treatments and that the plastic strain provoked by straightener is recoverable. Moreover, it demonstrates the interest for the interferometric method associated to the repositioning system. This method, simpler than AFM, reduces the acquisition time and could be extended to cosmetics products.

16.3.2 Dynamic Experiments

Dynamic experiments combine the interferometry with a push-pull machine dedicated to fibre studies. It uses the ability of the interferometer to make a topographic image with non
contact and in a short time, less than 1 min. A sample (20 mm long) is mounted on the machine and certain parts of it are located with silicon droplets. An image of each area is taken before the pull test. During the test, images are taken on the areas located, and the corresponding force and strain is recorded. Thus images of the same area can be compared as shown in Fig. 16.4. From the image, it is possible to extract a profile and observe the elongation of the cuticle. The solicitation of the cuticle causes an elongation of the scales and an increase of the roughness (15% in the present case). The aim of this experiment would be to reach the rupture of the fibre and measure the evolution of the topography from the initial state to the final.
16.4 Conclusions

We have presented a new method to evaluate hair cuticle modifications. This new method keeps the advantages of AFM and SEM, and offers new possibilities. It provides a large image until $80 \times 60 \, \mu m$ without contact with the sample and allows measurement of the same area before and after a treatment. Thus the study of topographical modification...
becomes easier. Due to the high speed of the acquisition, this technique can be combined to other measurement devices.

In this paper, we particularly describe straightener treatment for which the interferometry is particularly suitable, but it could be adapted to cosmetic products provoking a modification of the cuticle topography.

References

17.1 Introduction

Quantification of disease severity is a prerequisite for the development of evidence-based therapy. Today, there is no international consensus on guidelines for assessment of skin colour, and the majority of assessment methods are not standardized. Today, patient history and clinical scoring are the main tools for dermatologists when attempting to assess the morbidity of patients with various skin diseases. These methods however have their limitations; as they frequently show poor inter- and intra-observer reproducibility, due to the different ways doctors assess, for example, erythema or dry skin [14]. In addition, many of the scoring systems include assessment of disease extent, which has been shown to be difficult [1]. To provide objective and reproducible methods, instrumental evaluation of skin morphology and functions has been proposed.

Several of these techniques have reached technological maturity which suggests that a more routine use is possible [2, 4, 7, 10, 12, 13, 15–17].

Quantification of erythema is an obviously relevant technique for measurement of physiological changes in a majority of skin diseases. The first quantitative evaluation of skin colour was performed in 1937 by Edwards Duntley, who obtained data specific for different types of pigmentation [4].

The perception of colour by the human eye and brain has a range from 400 to 800 nm with a maximum sensitivity between 500 and 600 nm, corresponding to the colour of blood and hence of redness [11]. However, it is quite difficult to study and record skin colour changes quantitatively, as individual perception of colour is complex and subjective. Several commercial instruments are available: scanning reflectance spectrophotometers, tristimulus colorimetric and narrow-band simple reflectance [3, 9, 18].
**17.2 Skin Colour**

Skin colour depends largely on the properties of the epidermis, which acts as a biologic filter. A band of visible light incident on the skin may be transmitted, absorbed, diffused or reflected in different proportions. The colour that the human eye perceives is the result of these phenomena as determined by the presence of chromophores such as melanin in keratinocytes, carotenoids and haemoglobin in the dermal capillaries.

All optical determinations of skin pigmentation rely on the fact that melanin absorbs light and therefore pigmented skin areas appear to be darker that less pigmented areas, which is also the visual inspection [15].

*Melanin pigment* in the basal epidermis absorbs without peaks all wavelengths of the visible spectrum, increasingly with shorter wavelengths. It imparts colour ranging from brown to black, and when intense, it masks the red of haemoglobin.

*Carotenoids* are yellow pigments of exogenous origin acquired by eating yellow fruits and vegetables. They are stored in the corneous layer, sebaceous glands and subcutaneous fat. Carotenoids have only a minor influence on skin colour, except in rare cases, in which they accumulate excessively [4, 19].

*Haemoglobin* is responsible for the red component of skin colour. Oxyhaemoglobin in the capillaries and arterioles imparts a bright red colouring, whereas reduced haemoglobin in the venules imparts a bluish colour. The colour of oxyhaemoglobin is more accentuated in areas where the stratum corneum is thin or absent as the lips. In areas like the face, where arterial and capillary blood supply predominates, skin colour tends to be reddish. On the other hand, skin colour is more bluish in areas where venous circulation predominates as the trunk and dorsal side of the feet. Under conditions of vasodilatation as seen in inflammatory diseases, the skin becomes redder due to an increase of oxyhaemoglobin, whereas vasoconstriction as in ischaemia causes the skin to be more bluish. Haemoglobin shows specific and high absorption of light in the green spectral range and minimal absorption in the red range. With increase in erythema, a greater amount of green light is absorbed and less is reflected.

**17.3 Individual Differences**

Skin colour varies from race to race, and in all subjects, skin colour is fairer in areas rarely or never exposed to the sun. Visual estimation of erythema in pigmented skin is difficult and may even be misleading, since redness is influenced by the red component of the brown colour.

A measurement of the buttocks area is often used as a reference point for assessment of UV radiation’s influence on pigmentation.
The colour of non-exposed skin is known as constitutional, and that of exposed skin as facultative. The difference between the two colours can be calculated and indicates the pigmentation capacity of the subject.

The constitutional skin colour correlates with Fitzpatrick photo type, with a darkening colour gradient from I to VI. The Fitzpatrick Classification Scale was developed in 1975 by Harvard Medical School dermatologist, Thomas Fitzpatrick, MD, PhD. This scale classifies a person’s complexion and their tolerance of sunlight. It is used by many practitioners to determine how someone will respond or react to facial treatments, and how likely they are to get skin cancer. Constitutional skin colour can be taken as a parameter of photo type.

Aged skin seemed to be less coloured and male skin may be redder than female skin [2].

Clinical scoring of skin colour by visual readings is the most widely used method for assessment of, for example, erythema. Visual readings normally use a grading system of erythema by four levels:

Fitzpatrick Classification Scale:

<table>
<thead>
<tr>
<th>Skin Type</th>
<th>Skin Color</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>White; very fair; red or blond hair; blue eyes; freckles</td>
<td>Always burns, never tans</td>
</tr>
<tr>
<td>II</td>
<td>White; fair; red or blond hair; blue, hazel, or green eyes</td>
<td>Usually burns, tans with difficulty</td>
</tr>
<tr>
<td>III</td>
<td>Cream white; fair with any eye or hair color; very common</td>
<td>Sometimes mild burn, gradually tans</td>
</tr>
<tr>
<td>IV</td>
<td>Brown; typical Mediterranean caucasian skin</td>
<td>Rarely burns, tans with ease</td>
</tr>
<tr>
<td>V</td>
<td>Dark Brown; mid-eastern skin types</td>
<td>Very rarely burns, tans very easily</td>
</tr>
<tr>
<td>VI</td>
<td>Black</td>
<td>Never burns, tans very easily</td>
</tr>
</tbody>
</table>

0 = no visible erythema  
1 = faint, but well-defined erythema  
2 = moderate erythema and  
3 = severe erythema

Similar grading scale for clinical assessment of pigmentation after UV exposure is available:

0 = no reaction  
1 = minimal perceptible pigmentation, faint or no borders  
2 = light brown  
3 = moderately brown  
4 = dark brown
Assessment of pigmentation by diffuse reflectance spectroscopy correlated well with the corresponding clinical evaluation of pigmentation ($R^2 = 0.852$) [16]. Also, study of contact dermatitis compared visual scoring with Laser Doppler flowmetry and found a high correlation ($r = 0.9079$, $p < 0.001$) [17].

17.4
Instruments

17.4.1
Reflectance Spectrophotometric Evaluation

Older spectrophotometric scanning devices had previously little practical use because the broad melanin absorption band overlaps the haemoglobin band. However, new instruments that measure the haemoglobin band specifically and express erythema as an index of haemoglobin relative to melanin are now available. Recording of reflectance are influenced significantly by non-specific optical phenomena of the skin related to scaling and scattering, which is present in many diseases.

Scanning reflectance spectrophotometers are very expensive, cumbersome and not portable for routine clinical work. These instruments are mainly used for fundamental laboratory research. Simpler portable skin reflectance instruments are developed as the Erythema/Melanin Metre (DiaStron Ltd, Andover, Hampshire, UK), DermaSpectrometer (Cortex Technology, Hadsund, Denmark), Mexameter (Courage & Khazaka GmbH, Köln, Germany) and the UV-Optimize (Matic, Naerum, Denmark).

The Dermaspectrometer® (Cortex Technology, Hadsund, Denmark) is a narrow-band reflectance spectrophotometer designed for measuring specific colours of two major chromophores; haemoglobin and melanin in human skin. The Dermaspectrometer light emitting diodes emit light at two defined wavelengths, 568 (green) and 655 nm (red). The photo detector measures the light reflected by the skin at wavelengths in the green and the red for haemoglobin and for melanin, respectively. The melanin Index (M-index) is mainly influenced by the melanin content, whereas the Erythema Index (E-Index) is influenced by the haemoglobin content, assuming that absorbance of haemoglobin and melanin is similar to the wavelengths for green and red light.

Measurement of erythema can also be determined by skin reflectance spectroscopy using UV-Optimize®, Model Matic 555, Matic (Naerum, Denmark). The skin reflectance metre measures skin pigmentation (melanin) and redness (haemoglobin). It is based on the principle that light absorption in specific chromophores is quantified as the reflected light relative to the incident light. The peak wavelengths at 555 and 660 nm are used because of the optimal discrimination between absorption in haemoglobin and melanin [6, 8, 9]. Redness is measured on a scale from 0 to 100. “0” is the colour of an area, where blood circulation has been temporarily stopped and all blood has clinically been drained from the area, and “100” is the most intensive redness seen in dark bluish red nevus flammeus and represents a clinically defined maximum intensity. The value of erythema can be given as single measurements or as a mean of several.
The Minolta Spectrophotometer (Minolta, Osaka, Japan) is based on physical measurements of reflected light at specific wavelengths corresponding to the spectrum of visible light. The skin surface is illuminated by a pulse xenon arc lamp. The light reflected perpendicular to the surface is collected for a tristimulus colour analysis at 450, 560 and 600 nm. A number of light filters are built in. Results are displayed as a graph showing reflectance versus wavelength. Results are then converted and displayed as colorimetric values in the three dimensional L*a*b* system, as determined by the Commission Internationale de l’Eclairage (CIE) (REF). The L* value (luminance) gives the relative lightness ranging from total black (L* = 0) to total white (L* = 100). The a* value represents changes along a red–green axis with changes from −60 for a green surface to +60 for a red surface, whereas the b* value changes from −60 (blue surface) to +60 (yellow surface).

Chromameter does not give any information about the substances that generate the colour in contrast to the other instruments, which take a measure at wavelengths of specific chromophores.

L* is mentioned as the value of the lightness of an object, regardless of its chrominance and is mainly influenced by the green light, which is absorbed not only in melanin but also in haemoglobin. The L* value is therefore interfered by the pigmentation as well as the erythema, whereas the M-index is determined only by the intensity of the red light from the skin. The L* value and the M-index are therefore not equivalent, and the latter is considered to be a more specific marker of pigmentation than L*.

Moderate to high significant linear correlations could be established between the CIE L*a*b* colour parameters and the erythema/melanin indices [3].

17.5 Study Procedure

To say it simple: “A fool with a tool is still a fool”.

Established guidelines for measurements of erythema must be followed in any study [5]. Before the start of any measurement session, an acclimatisation period of at least 20 min is recommended, and the subject should be placed in a horizontal position. This is especially relevant when measuring the extremities to avoid orthostatic effects influencing skin colour. Test sites should be uncovered and motionless at least 5 min before measurements. To reduce the influence of seasonal variation on environment-related variables, room temperature and relative humidity must be controlled and continuously measured. The most optimal setting is assessment of erythema in a study room with constant temperature (19–23°), to avoid variation of skin vasodilatation or vasoconstriction.

Avoid measurements in direct sunlight, because the false light can cause errors into the sensors.

Also the pressure of the probe on the skin affects vasodilatation. As with any non-invasive instruments it is crucial to know:

- The diameter of the skin measuring area
- The instructions for application of the probe on the test area (application is simple by probe weight or by a constant pressure using a spring)
• Calibration instructions
• Take at least three measurements at the same spot and calculate the mean value (lift and reapply between each recording fast)

For repeated measurements in the case of evaluation of treatment efficacy over time, marking or detailed information of the measured area is important. Also repeated measurements should preferably be done at the same time of the day due to diurnal variation in skin colour, with increasing level of $a^*$ during the day. For assessment of skin pigmentation changes, a reference point – as a non-sun exposed area – may be needed.

Physical activity, local heating and alcohol causes cutaneous vasodilatation and thereby influence cutaneous blood flow and colour, whereas smoking, caffeine, mental stress and adrenergic system activation is associated with pale skin colour (vasoconstriction). These factors should also be systematically considered in preparation for a study (Table 17.1.).

Table 17.1 Factors influencing skin colour measurements

<table>
<thead>
<tr>
<th>Variables that influence skin colour</th>
<th>How they influence?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Aged skin may be less coloured</td>
</tr>
<tr>
<td>Sex</td>
<td>Male skin may be redder than female skin</td>
</tr>
<tr>
<td>Race</td>
<td>Different pigmentation among races</td>
</tr>
<tr>
<td></td>
<td>Visual estimation of erythema in pigmented skin is misleading</td>
</tr>
<tr>
<td>Anatomical sites</td>
<td>The face, neck, palm and soles have, in general, the highest index of erythema</td>
</tr>
<tr>
<td>Diurnal variation</td>
<td>Erythema of the skin increases during the day</td>
</tr>
<tr>
<td>Physical activity</td>
<td>Erythema of the skin increases with the level of exercise due to increased blood flow. An acclimatisation period of 20 min is important before assessment</td>
</tr>
<tr>
<td>Mental activity</td>
<td>May activate the adrenergic system resulting in peripheral vasoconstriction making the skin pale</td>
</tr>
<tr>
<td>Orthostatic effects</td>
<td>Erythema of the skin decreases if the measured area rises from the heart level</td>
</tr>
<tr>
<td>Nicotine</td>
<td>Cause vasoconstriction</td>
</tr>
<tr>
<td>Alcohol</td>
<td>Cause vasodilatation</td>
</tr>
<tr>
<td>Temperature</td>
<td>Normal skin surface temperature is 30°C. If temperature rises, vasodilatation and thereby erythema occurs, and reach its maximum at 40°C. Room temperature is best between 19 and 23°C</td>
</tr>
<tr>
<td>Ambient light/sun</td>
<td>May influence measurement of skin colour. The probe must be kept perpendicular to the skin surface to avoid false light</td>
</tr>
<tr>
<td>Seasonal variation</td>
<td>Tan in summer and xerosis in winter influence skin colour</td>
</tr>
</tbody>
</table>
References

18. Wulf, H.C., inventor: Method and apparatus for determining an individual’s ability to stand exposure to UV. US patent 4;882;598 (1989)
Core Message

- High doses of sun radiation can cause sunburn, skin aging, immune-suppression, and in long-term conditions, also skin cancer. The human organism has developed protection strategies against nonphysiological high doses of UV sun radiation in form of sun-tanning effect and hyperkeratosis. The development of these protection mechanisms takes days or even weeks. Therefore, the application of sunscreens, which develop their optimum protection within a few minutes, is an efficient protection strategy. The protection efficacy of a sunscreen is characterized by the sun protection factor (SPF). The SPF has the disadvantage that they are only related to the protection in the UVB part of the spectrum. Therefore, regulation was developed to also determine the criteria for an efficient protection in the UVA spectral range. The methods for the determination of the SPF values and the regulations for the classification of the sunscreen products are described in the following paragraph.
18.1 Introduction

The sun emits a continuous spectrum, which extends from high-energy X-ray radiation to low-energy heat radiation. Only UVB, UVA, visible, and infrared radiation reach the surface of the earth. The other spectral parts are absorbed by the earth’s atmosphere.

On one hand, sun radiation is essential for human life; for instance, it stimulates vitamin D synthesis and is important for human wellbeing [6]. On the other hand, high doses of UV radiation can cause sunburn, immune-suppression, skin aging, and even skin cancer [11, 13]. However, the human body has developed protection strategies against UV radiation by increasing melanin synthesis, which is an efficient UV absorber [12, 19]. Additionally, melanin has strong scattering properties in the visible spectral range, which reduce the penetration of the sunlight into the living tissue [1]. The high melanin concentration in the skin is the base of visible sun-tanning effect. Furthermore, UV radiation may induce an increase in the thickness of the uppermost layer of the human skin – the stratum corneum. This effect is called hyperkeratosis [7]. It represents an additional efficient defence against UV-radiation.

The development of these protection effects takes several weeks. In most cases, in our daily life, when skin is exposed to sun irradiation, for example, during leisure activities or holidays, the skin does not have enough time to adapt to the changes originating from sun exposure.

Under such circumstances, artificial sun protection is necessary. Covering the sun exposed skin with textiles and sunscreens are both effective forms of protection. In the past, the sunscreen efficacy was characterized by the sun protection factor (SPF) [11], which describes the UVB protection. Today, the UVA protection is also declared by the manufacturer on sunscreen products.

18.2 UVB Protection

The protection efficacy of sunscreens in the UVB spectral range (295–320 nm) is determined by the SPF. This factor is based on the formation of an erythema as a biological response of the skin to UVB irradiation [5]. The determination of the SPF is performed in vivo on humans: the unprotected skin is irradiated with a UV lamp, which emits a spectrum similar to that of sun radiation on the surface of the Earth. Subsequently, the minimal UVB dose necessary to produce an erythema is determined (minimal erythema dose – MED). Afterwards, the same procedure is repeated on the sunscreen pretreated skin. This procedure is described in detail by the COLIPA Guidelines [2] of the working group of the European Cosmetic Industry. The characteristics of the UV lamp are resolved in these guidelines, as well as the characterization of sunscreens, which should be applied onto the skin at a concentration of 2 mg/cm².

Taking these conditions into consideration, the SPF can be determined by the ratio:

\[
\text{SPF} = \frac{\text{MED of the Sunscreen Protected Skin}}{\text{MED of the Unprotected Skin}}
\]
In accordance with the COLIPA Guidelines [2], the SPF is determined on the back of 12–20 volunteers with skin types II and III. The detection of the erythema should be performed 16–24 h after irradiation. Subsequently, the average value and standard deviation can be determined.

There are several factors, which can influence the SPF value obtained in individual volunteers. Recently, it was demonstrated that the skin surface structure, i.e., depth and density of the furrows and wrinkles, influences the SPF values. This fact is not surprising as the efficacy of the sunscreen is determined by the absorption properties of the UV filter substances and by the homogeneity of their distribution on human skin. While the absorption properties of the filters are determined by the applied chemical compounds, the homogeneity of their distribution on the skin depends on the type of application (for instance, the application of a massage) and by the skin surface structure [8].

Application on smooth skin surface results in a more homogeneous distribution than that on skin characterized by furrows and wrinkles. Consequently, the selection criteria of volunteers can influence the measured SPF values. Also, the different steps of the UVB dose increase, necessary for the determination of MED may influence the results. During the application of sunscreens, the same concentration of 2 mg/cm² is applied homogeneously on the complete investigated skin surface.

18.3 Example for the SPF Determination

18.3.1 Determination of MED

The following is an example for the SPF determination of a sunscreen: In the first step, the MED of the volunteer is determined on the nonprotected skin on the back (Fig. 18.1). Subsequently, the MED is measured on the sunscreen protected skin.

Fig. 18.1 Determination of the MED on the skin on the back by a sun simulator
### 18.3.2 Sunscreen Application

It is recommended that the skin area to be tested on the back of volunteers should be marked. The calculated amount of sunscreen corresponding to 2 mg/cm² should be applied homogeneously. This can be achieved utilizing syringes: the empty syringe is filled with the sunscreen, taking care that the entrance to the syringe hub is also filled with cream and no air bubbles are visible. The syringe is then emptied and subsequently weighed. Afterwards, the corresponding amount of the sunscreen should be filled into the syringe and weighed again. Small drops of the cream are applied homogeneously onto the skin, over the complete marked area, as demonstrated in Fig. 18.2. The sunscreen applied by means of this procedure is then distributed homogeneously on the skin, with a glove-protected finger. It is necessary to previously store the glove in the cream for 10 min, in order to ensure that it is saturated with the cream (Fig. 18.3). Subsequently, the glove is cleaned with paper towels and used for the distribution of the sunscreen on the skin (Fig. 18.4).

---

**Fig. 18.2** Homogeneous application of the sunscreen onto a marked skin area

**Fig. 18.3** Saturation of the glove in the sunscreen
18 Characterization of Sunscreens: Determination of the SPF

18.4 International Methods for the SPF Determination

Different methods are used for the determination of the SPF. In Europe, the COLIPA method [2] is applied, while in the USA, the method of the Drug and Food Administration (FDA) is utilized. Also Australia and Japan have their own standards for SPF determination. All these methods are similar, however, with minor differences, i.e., the number of volunteers to be tested in the measuring procedure. In accordance with COLIPA Guidelines, ten volunteers as a minimum and 20 as a maximum are proposed, depending on the standard deviation of the individual SPF values determined. In the FDA method, 20 volunteers are proposed as a minimum. Additionally, the recommended UV lamps have slightly different spectra. For this reason, American SPF values determined by the FDA method are slightly higher than the SPF values determined by the COLIPA procedure [2].

On account of the nonlinear relation between the UV filter absorption and the SPF, there are strong differences in the absorption values between sunscreens with different low SPF values and no major differences in the absorption values of sunscreens with a high SPF. For instance, in the case of an SPF of 20, already 95% of the UVB radiation of the sun is absorbed by the sunscreen. However, in the case of a sunscreen SPF = 50, 98% of UVB is absorbed.

18.5 Determination of the UVA Protection Factor

The determination of the UVA protection factor is similar to that of the SPF determination [18]. In contrast to the SPF, where the erythema is analyzed, however, in the case of the UVA protection factor, the changes in the pigmentation (sun-tanning) of the skin are analyzed after irradiation with UV light.

Fig. 18.4 Homogeneous distribution of the sunscreen with the saturated glove
This method can be influenced by various factors:

1. The intensity of the pigmentation fluctuates in different volunteers according to their ethnic background and their life style.
2. A sensitive determination of the changes in the pigmentation is difficult.

The evaluation of the UVA protection efficacy of sunscreens becomes even more complicated, as two methods are proposed. When considering the first method, termed “immediate pigment darkening” (IPD) [10], the changes in the color of the pigmentation are determined 15 min after irradiation. In the case of the second method, called “persistent pigment darkening” (PPD) [9], the changes in the pigmentation are determined 2 h after irradiation. The UVA protection factors obtained with both the methods differ significantly. The IPD values are higher than the PPD values. Many sunscreen producers use the PPD method for the evaluation of the UVA protection in their products. The UVB (SPF) and UVA protection factors are independent of each other. Both methods do not represent internationally accepted standards, but represent only a feature for the determination of an approximate value.

The only valid method for the determination of the UVA protection is the Australian standard [14]. Following this standard, UVA protection of a sunscreen could be declared only in the case of the sunscreen having a transmission of maximally 10% in the UVA spectral range.

### 18.6 Classification of Sunscreens

Following the regulations of the European Commission from September 22nd 2006 [3], a sunscreen is classified by the following criteria:

A sunscreen should have at least a SPF of six and an UVA protection of at least one-third of the SPF, determined by the PPD method.

The different SPF values are summarized in groups and categories of low, medium, high, and very high protection, as shown in Table 18.1.

<table>
<thead>
<tr>
<th>Labeled category</th>
<th>Labeled sun protection factor (SPF)</th>
<th>Efficacy of the product against erythema (SPF)</th>
<th>Recommended minimal UVA protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low protection</td>
<td>6</td>
<td>6–9</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10–14</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td>Medium protection</td>
<td>15</td>
<td>15–19</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>20–24</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>25–29</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td>High protection</td>
<td>30</td>
<td>30–49</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>50–59</td>
<td>1/3 of the SPF</td>
</tr>
<tr>
<td>Very high protection</td>
<td>50+</td>
<td>60+</td>
<td>1/3 of the SPF</td>
</tr>
</tbody>
</table>
18.7 Outlook

The SPF is an important value, which allows the consumer to distinguish between the various sunscreen products. However, there are two main disadvantages. First, the SPF is related only to UVB protection and second, it is determined “invasively” on human volunteers in vivo using UVB light. The first disadvantage is reduced by the regulation of the European Community [15] that a sunscreen should have a UVA protection of at least one-third of the SPF, determined by the PPD method. In this context, the question arises as to whether sunlight at wavelengths <400 nm, i.e., visible or infrared light, has a biological effect and thus damaging human skin. Recently, it was reported that approximately 50% of the free radicals produced by sun radiation in the human skin are related to the visible and near infrared spectral ranges. As a result, a radical protection factor was introduced in analogy to the SPF, characterizing the formation of free radicals, relating to unprotected and sunscreen protected skin. The measurements must be carried out in vitro on skin biopsies representing an invasive method.

Several attempts were undertaken to determine the SPFs in vitro; for instance, plastic or quartz plates with a structured surface were used, on which the sunscreens were applied at specific concentrations [4]. Subsequently, the transmission of the plates was analyzed by means of a spectrometer. However, this method has the disadvantage that the structure of the plates did not correspond to the real structure of the skin surface. This problem can be solved by the method using the determination of a “Universal Sun Protection Factor” (USBF) [16, 17]. In this case, the sunscreen products were applied on the skin in accordance with the COLIPA protocol. After a penetration time of 1 h, the upper layers of the stratum corneum, where the sunscreen had penetrated, were removed by means of tape stripping. With this procedure, single layers of corneocytes were removed together with the sunscreen that had penetrated into these layers. It could be demonstrated that the non-homogeneous distribution of the sunscreen on and in the skin is transferred onto the tape strips, which were analyzed spectroscopically. The sum transmissions of all tape strips containing sunscreen were determined. This sum transmission in the UVB corresponds to the classical SPF values [17]. Additionally, this method can be easily used to determine the UVA protection as well as the protection efficacy in the visible and near infrared spectral ranges.

18.8 Summary

The classical SPF is an important value for the characterization of sunscreen products. It is well-known and accepted by the general public. The SPF determination is an invasive process based on a biological response – the formation of an erythema. Different protocols are used in the USA, Europe, Asia, and Australia for the determination of the SPF values. In Europe, the COLIPA Protocol is an accepted guideline. The SPF is strongly related to
the UVB part of the sun spectrum. For the assessment of the UVA protection, up to now, no generally accepted method is available. Therefore, in Europe, it is now established that a sunscreen should possess an UVA absorption of at least 30% of the UVB absorption. The determination of the influence of visible and infrared light on the skin still remains a topic of intensive research. Depending on the results, further guidelines are to be expected.

References


19

19.1 Introduction

Walking down the hair care aisle of a store can be an intimidating experience as shelves are literally packed with a huge variety of products promising all manner of benefits. Yet, technically, differences between formulations of a given product form are not especially dramatic. A glance at the label ingredient statements is enough to demonstrate that the same ingredients, or class of ingredients, are present in most instances. Of course, these ingredients can be formulated in a variety of ways to bring about differences in the magnitude of certain benefits, while also leading to differing levels of aesthetic acceptance: however, with this said, it is often the manner by which products are advertised and marketed that leaves the deepest image in the mind of the consumer. As such, in any evaluation of hair care products, there is a need to consider the technical benefits, the aesthetics, and the marketing message – where a combination of all three leads to the most successful products.

The differentiation between technical performance and consumer performance is an important one. The functionality of products lies with the underlying surfactant science; yet, with such similarities in formulation, differences in technical performance can often be overridden by aesthetic drivers. For example, an especially appealing fragrance may often be the primary reason that a consumer chooses a specific product. The consumer may make a conscious decision to pick the best smelling product; or may be haloed into believing that the product performs best due to the allure of the odor. Technical evaluation methods are available to quantify how products can change the properties of hair; but, they contain no information about consumer acceptance. By means of illustration, instrumental testing methods can be used to measure the level of lubrication provided by conditioning products to facilitate hair manageability, but the experiments provide no information as to whether this occurs in an aesthetically pleasing manner. On the other hand, consumer evaluation methods evaluate acceptance and liking, but interpretation can often be
complicated by the presence of so many aesthetic and performance drivers. In short, there is the desire to perform both consumer and instrumental testing of products. It must also be remembered that consumer products are marketed using “consumer language,” which often may not match technical understanding. It is the job of market researchers, marketers, and consumer scientists to analyze the ways the consumers talk about their hair – and, in doing so, allow product benefits to be recounted in the same vernacular. As such, this language persists throughout the hair care world, even though there are clear examples of where consumer self-prognosis incorrectly identifies the root cause. Further complicating the issue is the existence of a number of rather ambiguous consumer terms, such as “body” or “conditioning,” which largely defy technical description and measurement. Therefore, there is the need to consider where and when technical or consumer evaluation methods are most appropriate. Logic suggests that technical measures be used to quantify physical changes in the properties of hair; while human subjects are best utilized to assess obvious consumer attributes.

19.2 Hair Science Overview

Before talking about the testing of hair care products, there is the need to briefly overview the science of hair and hair care. This is a large and varied area, which cannot be given full justice in these few pages. Instead, an attempt has been made to provide a succinct summary, while directing the reader to more comprehensive text in specific areas.

Individual hair fibers grow from follicles that cover the scalp. In actuality, such follicles cover our whole body, but coarser so-called terminal hair is produced on the head, while finer vellus hair covers most other regions. Cell growth and division occur deep in the base of the follicle [1, 5, 7] and, when hair emerges onto the scalp, it exists in a biologically inert form (perhaps somewhat overdramatically, it is often heard that hair consists of dead cells). So, the well-being and integrity of the hair structure depends on the habits and practices of the wearer. On average, hair grows at a rate of ½ in. per month; so the tips of shoulder-length hair have been exposed to a year’s worth of wear and tear from washing, grooming, and potentially still more deleterious practices. In particular, so-called chemical treatments of hair (e.g., coloring, bleaching, perming, and relaxing) are recognized to be especially damaging – yet this does not act as much of a deterrent, with a high percentage of female consumers regularly utilizing at least one of these treatments.

Hair itself has an almost unbelievably complex structure. Again, it is not the purpose of this article to comprehensively review this subject, but the interested reader is directed to other text on this topic [2, 4, 7]. For our purposes, it is sufficient to divide this structure into three regions. The outside of hair consists of a series of hard, protective, overlapping tile-like structures – known collectively as the cuticle. In large part, the integrity of this structure dictates the feel properties of hair. Healthy hair will possess a relatively pristine cuticle structure, but the constant wear and tear of everyday grooming gradually results in cracking, chipping, and abrasion – and a commensurate degradation in feel properties. A degrading cuticle structure will also contribute to manageability issues, as increased interfiber
friction leads to higher likelihood of knots, tangles, and more arduous grooming. Furthermore, a more irregular surface does not reflect light as readily – leading to the perception of dullness. A more detailed discussion on the evaluation of these, and other, parameters will be given shortly.

While this outer structure largely dictates visual and tactile properties, it is widely accepted that it provides little or no contribution to the strength of hair. Instead, this property relates to the internal structure of the fiber – collectively termed the cortex. This too is a complex structure consisting of crystalline and amorphous protein regions, all held together within a lipid membrane. The previously mentioned chemical treatments are the main culprit in attacking this structure – perhaps most importantly, depleting cross-linking bonds that are the main contributor to fiber strength. Another assailant on these same bonds is the sun’s UV rays which induce cleavage by photochemical oxidation. These factors measurably diminish the tensile properties of hair leaving it more prone to breakage. Then, with further washing and grooming, the tips of broken fibers will fray and form split ends. This breakdown of the internal structure also leads to greater swelling when immersed in water – which again contributes to wet-state detangling and grooming issues. It is often convenient to categorize the broad topic of “hair damage” in terms of these internal and external structures; but, obviously, both are intimately linked. The cuticle represents a tough, resistant structure that shields the more vulnerable cortex. Hence, its degradation reduces this protective capability, leaving the inner portions more exposed. Meanwhile, repeated swelling and deswelling of the cortex puts pressure on the cuticle scales, ultimately resulting in cracks and further degradation.

Also contained in the cortex are melanin granules that give hair its color. The difference in hair color arises from both the amount and the type of melanin present. Increasing concentrations of eumelanin progressively lead to blond, brown, and ultimately black hair; while the presence of pheomelanin results in red and yellow pigmentation. The absence of melanin leads to grey hair. These pigment granules are believed to constitute less than 2% of the hair’s total composition, and so, despite the obvious influence on the appearance, the presence (or absence) of melanin is generally believed to have no influence on physical properties.

The third component of hair’s structure is the very innermost portion, which is termed the medulla. Little cosmetic relevance is paid to this region, which is generally considered equivalent to hollow space at the center of a fiber and simply serves to increase thickness. A pronounced medulla is present in thick fibers (e.g., Asian hair), while it may be completely absent in fine hair. These three domains of the hair structure can be seen in Fig. 19.1, which shows a high magnification scanning electron microscopy (SEM) image of a fractured hair fiber.

After touching on hair ethnicity, there is the need to further expand on the subject of “hair type.” The vast majority of studies suggest no significant difference in the chemical and structural composition of hair regardless of ethnicity, shape, or color. Therefore, it is convenient to think of all hair types as essentially consisting of the same “stuff” that has been extruded out onto the scalp in different shapes and sizes. This positioning is contrary to consumer desires, where the idea of custom-designed products made especially for exactly their hair type and condition is particularly appealing. However, in reality, differences in the properties of hair generally dictate the need for products with different levels
of performance, rather than a completely different treatment altogether. By means of
illustration, the thickness of Asian hair, coupled with a frequent desire for straight, sleek
styles, leads to the desire for highly conditioning products. These same products will likely
be too heavy-coating for most Caucasian consumers who desire the same conditioning
benefits, but not to the same extreme. The same scenario also exists for different hair types
within a given ethnicity, where differing conditioning levels are needed for long, thick,
curly hair vs. short fine hair. A possible exception to this line of thinking involves Afro
hair, where the highly kinky conformation leads to specific issues that result in quite dif-
ferent habits and practices.

When considering the general properties of hair, it is impossible to ignore the consider-
able influence of water. Many differences in the properties of wet and dry hair are clearly
evident. Technically, this relates to water readily penetrating into the hair structure with
commensurate swelling and plasticization (softening). Therefore, wet fibers are weaker,
and feel distinctly rougher – thus making wet state grooming considerably more arduous.
Many will also be familiar with an inability to maintain hair styles on hot humid summer
days – which arises from high levels of adsorbed water softening fibers to a point that a
style collapses under its own weight. At the other end of the extreme, the buildup of static
flyaway becomes an issue on cold, winter days, when hair has diminished moisture con-
tent. As such, there is always the need to pay attention to the relative humidity when per-
forming hair testing.

### 19.3 Shampoos: An Overview

With so many claims and benefits listed on a typical shampoo product, it may sometimes
be forgotten that the primary function is to clean hair. Contained within the hair follicle is
a sebaceous gland that secretes oily material (sebum) onto the hair and scalp. Therefore,
after a day or so, build-up is likely sufficient to necessitate cleaning. Hair will also become
coated with exogenous soils, where major contributors may be deposits resulting from the use of conditioning and styling products. The use of anionic surfactants (e.g., sodium lauryl sulfate) for detergency is common in many related cleansing products. Such ingredients are formulated in a way that produce dynamic aggregate structures, called micelles, which allow solubilization of oily materials that would not otherwise be removed by water alone. The formation of these structures is aided by the presence of cosurfactants (e.g., cocamidopropyl betaine) and salts (e.g., sodium chloride). However, as anyone who has had shampoo in their eyes can attest, such surfactants tend to be rather irritating. Over the years, manufacturers have partially shifted to ethoxylated anionic surfactants (e.g., sodium laureth sulfate) which are less irritating, but also somewhat less cleansing. As such, a trade-off exists where the use of more aggressive surfactants, and higher concentrations, results in better cleaning, but also leads to more irritation. To deal with this issue, manufacturers produce shampoos with a spectrum of cleansing performance depending on the needs of the consumer. At the high end of this range are highly cleansing “clarifying shampoos,” which generally are not used on a daily basis; but instead are employed periodically as a means of minimizing the build-up of more substantive ingredients. At the other end of the spectrum, are children’s shampoos, where a lower level of sebum production, coupled with little usage of conditioning and styling products, allows for the use of milder formulations. Further following this logic, it can be seen how, and why, formulation differences may be present for different variants within in a product range (e.g., shampoos for normal hair, oily hair, dry hair, etc.).

A quick estimation of surfactant concentration can be obtained by performing simple gravimetric experiments, where these materials represent the major nonvolatile component after thorough drying. If desired, very accurate determinations of these concentrations can be obtained by surfactant titrations. Meanwhile, the aggressiveness of the surfactant system may be predicted by well-known methodologies such as corneosurfametry or Zien testing.

Despite relatively few measures of technical performance, shampoos come with many aesthetic drivers. First, consumer opinion is often swayed by the viscosity of the product – where thicker formulations are generally thought to be “richer” or “more-luxurious.” Technically, a water-thin shampoo can be just as effective at cleaning hair; but improved ease of application, coupled with this consumer mind-set dictates the need for thicker formulations. This said, the product must still apply easily and be readily distributed throughout the hair. Consideration of this seemingly conflicting situation drifts into the realm of “rheology” – that is, the study of how materials flow. In rheological terms, shampoos and conditioners fall into a category of liquids termed “shear thinning” – where the application of shear forces (e.g., rubbing) causes a decrease in viscosity. In this way, products that appear thick and creamy while sitting on the palm of a hand can still be easily massaged into the hair. Once on the hair, the foam quality is another important aesthetic consideration. While not strictly part of the technical functionality, consumers have come to expect a rich creamy foam. Yet, the foam and the product must readily be removed from the hair during rinsing; preferably without leaving any residual feel.

Briefly, there is the need to address the stability of hair care products. As described, such formulations are created using complex mixtures of water, surfactants, cosurfactants, oils, salts, and polymers all of which interact to produce the desired structures and benefits.
However, such interactions can continue over time and may lead to changes in the properties of a formulation. Thus, desirable aesthetic formulations may fall apart with time; or upon exposure to certain external stimuli. Obviously, there is the desire for a product to possess the same structure, properties, and performance throughout its lifetime – a period that covers manufacturing, shipping, warehouse storage, store shelf-time, and ultimately, bathroom shelf-time. As such, it is generally taken that products should be stable for up to 3 years. This said, product launch schedules do not allow for new formulations to undergo such lengthy stability testing. Instead, formulations endure accelerated aging studies; where a 3 month exposure to 50°C is generally taken as a proxy for 3 year shelf-life stability. Furthermore, there is the potential for products to see (or even cycles between) relatively extreme temperatures during shipping and/or warehouse storage. As such, as part of stability testing, it is also common to monitor stability after a series of freeze-thaw cycles.

### 19.4
2-in-1 Shampoo+Conditioners: An Overview

In the mid 1980s, Procter & Gamble introduced shampoo + conditioner products (sometimes called 2-in-1 s) onto the market under the Pert Plus brand. Such products deposit high molecular weight silicone oil on to the hair to improve feel qualities and to help with manageability. Other manufactures quickly followed, and for a while, these new products received top billing and evolved into an off-shoot of the shampoo category. Today, such products have largely been reabsorbed under the shampoo umbrella. Some manufacturers may still distinguish between “traditional” shampoos and 2-in-1 products; while others abandon the distinction and include both within the variants of a product line. The presence of dimethicone in a label ingredient statement indicates the 2-in-1 status of such products.

It is worth highlighting the ability to deposit silicone oils from a product whose primary purpose is to remove oils and dirt from the hair surface. There are two mechanisms by which this can be accomplished. Most basically, the use of relatively large silicone droplets (≥20 μm) leads to deposition via simple entrapment between hair fibers. However, such deposits are constantly under attack from the surfactant system, and consequently, this is not the most efficient method. A second, more elegant approach makes use of the complex interactions between anionic surfactants and certain cationic quaternary amine functionalized polymers [3] (polyquats). At elevated surfactant concentrations, the polymer and surfactant exist as a one-phase, homogeneous system; but, upon dilution, a complex precipitates out which causes flocculation of silicone droplets and facilitates deposition. In addition, as deposition occurs in the presence of reduced surfactant concentration, there is less likelihood of removal. The approach is often termed polymer-assisted or coacervation-assisted deposition.

Formulation stability becomes an even greater concern in such 2-in-1 products, where silicone droplets must be uniformly distributed and maintained throughout the sample for its effective lifetime. Differences between the density of water and silicone result in a
tendency for the oil droplets to cream to the surface – with the rate of movement being directly proportional to the size of the droplets. Stability is improved by the inclusion of polymeric thickeners (e.g., glycol distearate, carbomer), as increased low shear viscosity leads to slower settling rates.

There are many formulation variables that contribute to the amount of silicone deposited on the hair from a 2-in-1 product. These include, for example, the surfactant strength, the silicone oil droplet size, the presence of cationic polymers, and the nature of such polymers. Consequently, there is often the desire to quantify deposition levels of silicone. Two analytical techniques are commonly employed – inductively coupled plasma optical emission spectroscopy (ICP-OES) and X-ray fluorescence (XRF). While ICP-OES may be somewhat more sensitive, it does involve extraction of silicone from the hair surface, while XRF can be performed in situ.

In considering the depositing efficiency of any conditioning formulation, it should be remembered that there will be considerable dependence on the manner of usage. The application of higher product amounts, and/or the use of longer residence times on the hair, will significantly increase deposition. Similarly, the thoroughness with which the hair is rinsed will also play a role – with longer rinsing times and higher water flow rates leading to less deposition. Testing under laboratory (and often salon) conditions allows for control of these variables – for example, by weighing out, or syringe application of a product; by carefully timing the application; and by using showerheads with controlled temperature and flow rate. However, when put in the hands of consumers, this control is lost and formulations will be used in a variety of ways. As such, it is worth emphasizing that consumers may adapt usage conditions to fit their specific needs. For example, if a formulation is not conditioning enough, consumers may compensate by using higher doses, longer residence times, and/or repeated application. Thus, with some user adaptation, a given product can still be applicable for multiple hair types – although this fact is rarely publicized.

There is the need to briefly address aesthetics factors associated with 2-in-1 products. Obviously, such formulas closely resemble conventional shampoos, except that the low shear viscosity is boosted to aid with stability. The inclusion of silicone in the formulation creates an oil-in-water emulsion and consequently results in the formation of a white opaque base. It is possible to make a clear 2-in-1 through the use of microemulsions – where an exceptionally small particle size (<100 nm) allows light to pass through without undergoing scattering. The inclusion of pearlescent ingredients (e.g., ethylene glycol distearate, EGDS) is often employed to improve the visual perception of opaque products. The major aesthetic issue with 2-in-1 s involves another trade-off, this time surrounding the positives and negatives associated with the deposited materials. As already eluded to, individuals with damaged, thick, and/or curly hair, often desire high levels of conditioning to aid with manageability; but these same formulations will weigh down fine, thin hair leading to limpniness, and a lack of volume and body. As possibly anticipated, 2-in-1 s cannot deliver the same level of conditioning as a conventional conditioner product. However, for some, the deposition of silicones, surfactant-polymer complex, and pearlescent ingredients can still be too much; and will leave their hair feeling coated and unclean. This trade-off involving Conditioning and Body/Volume is further discussed in the conditioner section.
19.5 Conditioners: An Overview

The launch of 2-in-1 products that provide both cleansing and conditioning obviously follows on from the introduction of hair conditioner products some years earlier. Often termed “cream rinses” in the early days, such products are usually white, opaque, viscous liquids, whose functionality lies with the ability to lubricate the hair surface. While conditioning products are also plastered with claims promoting a variety of benefits, virtually, all stem from creative extrapolation of this lubricating ability. “Conditioning” is a consumer term, so there is danger in trying to ascribe a technical definition – however, it seems logical to presume that a perceived improvement in hair “condition” is somehow involved. Consequently, it can be seen how the ability to mask a rough cuticle surface can lead to significant improvement in the perceived condition of hair. Furthermore, continuing in the consumer vernacular, if the condition of the hair is improved, the perception may be that some form of “repair” has taken place: or, if the belief was that negative feel qualities were a consequence of “dryness,” then “moisturization” has occurred. Technically, the product has not in any way altered the physical structure of hair, nor had any effect on the water content; yet, consumer perception dictates the continued use of this language in the marketing of products. Continuing still further, it can be seen how the presence of fewer snags and tangles leads to easier grooming and improved “manageability.” A case can also be made for a reduction in surface abrasion leading to “protection” of the hair, which ultimately leaves fibers in a “stronger” state than unconditioned hair – and, as such, results in “less breakage.”

This lubricating ability results from the use of cationic surfactants that are formulated in such a way as to produce a lamellar liquid crystalline structure. In such layered structures, the sheet-like arrangements of molecules easily slide over each other and provide slipperiness (for comparison, the outside of a wet bar of soap consists of the same structure). Under everyday conditions, the surface of hair possesses a slight negative charge which attracts and facilitates deposition of positively charged surfactants. For this reason, such structures are built using quaternary amine functionalized surfactants, commonly termed quats (e.g., cetrimonium chloride, dicetyl dimmonium chloride), together with cosurfactants (e.g., stearyl alcohol, cetyl alcohol). It should be emphasized that these lamellar structures exist only in the wet state, and so the bulk of the lubricating benefit occurs under these conditions. Upon drying, a waxy deposit is left behind which provides some lubrication and feel benefits; although this is generally not as noticeable. Nonetheless, a significant dry state benefit of conditioners involves the ability to retard static flyaway. Under low humidity conditions, grooming of dry hair will be accompanied by considerable build-up of static electricity. This causes repulsion between individual fibers, and results in an undesirable wispy, flyaway appearance. There is some dispute about the mechanism by which the benefit occurs; but there can be no questioning the effectiveness of these products in preventing this condition.

As outlined earlier, conditioner products are also manufactured to give a spectrum in performance as a result of the differing needs of customers. More heavily depositing variants are targeted at individuals with longer, thicker, curlier, and more damaged hair. Typically, these products are marketed with language such as “intensive,” “moisturizing,”
“extra moisturizing,” or possibly with a designation for use with “dry, damaged hair.” Conversely, lighter conditioners are aimed at those with finer, thinner hair – where the intention is to provide conditioning without adversely affecting body or volume. As such, these variants often bear descriptors declaring “extra volume” or “extra body” – although, in fact, they do not contain “extra” anything, and achieve this end goal through using a lower concentration of conditioning ingredients.

Despite considerable formulation differences, conditioners share many of the same aesthetic drivers that were previously described for shampoos. Product rheology is again important, where a thick, rich consistency is desired upon dispensing; although the product must apply easily with a pleasant creamy consistency. Again, the product must be readily rinsed from the hair and should leave the hair without a coated-feel in both the wet and dry state.

19.6 Evaluation Approaches

The preceding text has attempted to introduce the reader to the complex world of hair and hair care products. As part of this introduction, both the technical functionality of products and consumer opinion has been highlighted; while emphasizing the potential for a complete lack of correlation between the two. Now it is time to specifically focus on assessment approaches; and, in particular, to look at some of the major attributes associated with these products. Over the years, these features have been identified by talking to consumers about issues, desires, habits and practices; while paying careful attention to the language being used. After discovering these attributes, there may be a desire to develop an understanding of whether scientific causes and explanations are present to describe these perceptions: in short, to de-code “consumer-speak” into “scientific-speak,” and consequently allow for development of treatments that can alleviate or correct negative symptoms. This is an on-going process, with the scientific literature and technical symposia constantly adding to our learning. The final part of this chapter focuses on current thinking with regard to some of the most important areas. However, first it is necessary to provide some introduction to the areas of consumer and instrumental testing.

19.7 Consumer Evaluation Approaches

In principle, the expression “consumer testing” can cover a broad spectrum of activities, ranging from a small, random group of individuals discussing their liking or disliking for a product; to a very large group, all with a specific hair type, filling in a lengthy questionnaire about the applicability of multiple products to a particular brand or product concept. Obviously these two scenarios represent very different levels of effort and yield very different levels of information. Simple guidance may be obtained from quick and simple evaluations; although new insights likely require more complex and detailed approaches.
In the product development process, the level of consumer testing generally becomes progressively more detailed as products get closer to launch. New product development begins with a skilled formulation chemist, whose experience allows him/her to produce prototype formulas which already have reasonably good performance and aesthetic profiles. However, this is still the opinion of a single individual and lead prototypes will be passed around colleagues for additional opinions. Nevertheless, such assessments are performed by individuals who are intimately familiar with hair care products, and the opinion of these expert evaluators may not be representative of more naive, typical consumers. Nevertheless, this process provides a means for the formulator to begin making iterative refinements to lead formulas.

In many instances, prototype testing may then progress to in-house salons, where experienced hair stylists will use lead formulations on human subjects in a relatively controlled manner. Feedback is obtained from both the stylists and panelists, generally in the form of a questionnaire. Thus, weaker formulas continue to be weeded out, and guidance is provided for further refinement of lead formula. Eventually, it becomes time to put formulas in the hands of volunteer consumers, who will take them home and use them as part of their normal everyday hair care practices. This process is often coupled with the inclusion of a benchmark — generally a leading commercial product — which sets a suitably high target for comparison of performance and aesthetics. Feedback is generally obtained from more detailed and lengthy questionnaires, which now specifically probe a variety of attributes in considerable depth. In addition, participants may be asked to take part in a Focus Group — usually consisting of a roundtable conversation where the merits and shortfalls of the formulations are discussed. Even after developing strong formulations, there may still be the need to check for a good match with the marketing concept for the product. Consequently, there is the potential for differences in consumer opinion from “concept” and “nonconcept” testing of the same products.

The Consumer Science discipline is a highly developed field, which again cannot be given full justice in this short overview. Again, the interested reader is referred to detailed text in this specific area [6].

19.8 Technical Evaluation Approaches

As hair is a biologically inactive material, it can be harvested and stored without any appreciable change in its structure and properties. Therefore, it becomes possible to perform in vitro laboratory experiments on individual fibers or specially prepared bundles — often called tresses, switches, or swatches (see Fig. 19.2). Indeed, such is the demand from companies that produce hair care products (and also the companies that manufacture the ingredients), that businesses have emerged that operate by procuring and selling hair. Most often, such bundles or tresses consist of blended hair that is obtained from an assortment of individuals with a common hair-type. The hair is then carefully mixed in a uniform manner, as shown in Fig. 19.3. However, it is often possible to procure specific hair types (e.g., fine, thick, straight, curly), as well as hair of different ethnicity — or even hair that has been predamaged to specific levels. In laboratory testing, it is often useful to use chemically
damaged hair as a test substrate, as benefits associated with certain products are magnified. Furthermore, given the high incidence with which these treatments are used, it can legitimately be argued that this state is more representative of hair on the head of typical consumers.

The advantage of laboratory experiments is that they can be performed in a highly reproducible manner. Testing can be performed on specific hair types, under controlled conditions, providing a standardized approach to evaluating the efficacy of hair care products.
environmental conditions, using precise application, rinsing, and drying conditions. Of course, such conditions are rather artificial, as consumers will use products in a multitude of different ways. Subsequently, instrumental testing is generally considered to produce fundamental information about product functionality.

When performing such testing, again there is the need to emphasize the effect of humidity on the dry state properties of hair. As such, generation of reproducible data necessitates the need for environmental chambers which not only encase the equipment, but also allow for full equilibration of hair prior to testing. It is also worth noting that hair is a highly variable substrate; so it is imperative that sufficient replicate samples be performed to ensure proper statistical analysis of the results. It is also good laboratory practice to utilize internal controls that act as anchors, while also providing a record of method reproducibility.

The following represent common instrumental testing approaches that are used to investigate consumer-relevant attributes of hair and product performance. It is worth noting that universal standard testing protocols are, to all intents and purposes, absent from the hair care industry. This is frustrating for most; although it can be beneficial for some who exploit the ambiguity this situation causes. Nevertheless, there are a number of methods that have become widely adopted throughout the industry; although each laboratory likely utilizes them in their own slightly different manner.

19.9 Assessing Surface Damage: Microscopy

An old proverb declares that “a picture is worth a thousand words.” Therefore, it is no surprise that high magnification images showing the condition of hair, or the manner by which products are deposited, are very popular. Degradation of the cuticle structure can involve chipping, cracking, uplifting, and/or abrasion; all of which can be visualized under high magnification. The images shown in Figs. 19.4 and 19.5 were generated using SEM at magnification levels of 450× and 750× respectively. Figure 19.4 shows how the cuticle layer has been completely eroded away in the upper portion of the fiber; while, Fig. 19.5 shows radial cracking that has developed as the results of repeated grooming stresses.

It should again be emphasized that shampoos and conditioners do not physically “repair” this degrading cuticle structure; but the lubrication provided by such products helps mask this effect – and, in doing so, provides the illusion of repair. Figure 19.6 shows conditioning deposits smoothing the cuticle structure.

19.10 Lubrication

As already emphasized, the primary function of conditioning products is to lubricate the hair surface and, in doing so, facilitate manageability and improve feel. The most popular, and perhaps the most relevant, approach for quantifying this lubrication involves
**Fig. 19.4** Scanning electron microscopy (SEM) image of a hair fiber showing cuticle abrasion

**Fig. 19.5** SEM image of a hair fiber showing cracking of the cuticle

**Fig. 19.6** SEM image showing conditioner deposits smoothing the hair surface
instrumental combing experiments. The basic concept involves measuring the frictional force associated with a comb passing through a hair tress. Testing requires an appropriate mechanical testing instrument (e.g., an Instron™, Diastron™ or Texture Analyzer™), and may involve the force detector being attached to the comb or the tress. The hair tresses are combed repeatedly to obtain an average value, while multiple tresses are used to ensure appropriate statistical rigor.

Such experiments are commonly performed in both the wet and dry states – although, as already noted, benefits associated with conditioner products are more evident in wet-state testing. That is, there is a considerably larger reduction in combing forces. The need for products with differing levels of conditioning has been described, and so it is seen how an Extra Moisturizing product would generally be expected to lower the combing forces more than an Extra Body variant.

Of course, “lubrication” is a scientific property, and not a term that is generally used by the consumer. However, as already mentioned, this property frequently correlates with (or is used to substantiate claims on) smoothness, softness, conditioning, manageability, protection, and even moisturization.

19.11 Strength

Consumers appear to ascribe a strong correlation between the health of their hair and its strength. We are familiar with the consumer expression “strong, healthy hair” which distinguishes from hair that is brittle, fragile, easy-to-break, and therefore “damaged.” The actual tensile properties of individual hair fibers are often evaluated by generating stress-strain curves. Fibers are stretched on an appropriate mechanical testing instrument (e.g., Diastron Mini Tensile Tester™), while the internal forces are measured by a load cell. Obviously, thicker fibers will have a tendency to be stronger than thinner fibers; so it is common to normalize the force against the cross-sectional area of the fiber to produce a unit of measure that is independent of fiber dimensions (i.e., stress=force/unit area). As already mentioned, chemical treatments, the sun’s UV rays, and everyday wear and tear will take a toll on the tensile properties of hair fibers – and significant effects can be evaluated by this approach. However, it is also likely that negative feel properties (relating to a degrading cuticle structure) may signal “damage” in the consumer’s mind and lead to a presumption of diminished strength. Shampoos and conditioners do not physically change the tensile properties of individual hair fibers; but, again, in consumer-speak, the masking of surface damage may lead to a perception that the condition has been alleviated.

Nevertheless, as already described, it may be argued that the lubrication provided by these products reduces snagging, tangling, and abrasion – thus providing a degree of protection, and ultimately leaving the properties of hair more intact than if the products were not used. A dramatic illustration of this benefit can be obtained from repeated grooming experiments, where hair is repeatedly combed or brushed with the subsequent counting of broken fibers. In principle, such experiments can be performed by hand grooming of tresses; although homemade devices involving brushes or combs on a rotating drum are more reproducible
and less strenuous. The use of conditioning products leads to substantially fewer broken fibers, and this method is often behind the substantiation of “anti-breakage” claims.

## 19.12
### Static Electricity Retardation

When two bodies are rubbed together, a transfer of electrons generally occurs. In the case of grooming hair, individual fibers normally give up electrons and develop a positive electrostatic charge. The resulting repulsion between fibers with the same surface charge leads to the static flyaway condition. The stability of this surface charge is related to the conductivity of the fiber, which in turn is related to the moisture content of hair. Therefore, under low humidity conditions, hair has little conductivity and charges are not readily dissipated. Accordingly, static flyaway is a bigger problem during winter months. Conventional conditioner treatments are highly effective at retarding static build-up; although there is some contention about the mechanism. It appears likely that there is a contribution from lubrication (which reduces the amount of charge build-up), and also an increased surface conductivity that arises from the deposition of cationic surfactants (which facilitates charge dissipation).

In the laboratory, static levels can be quantified using a variety of commercially available sensors. Hair tresses must be equilibrated under low humidity conditions, and are subsequently brushed or combed a specific number of times (preferably using an automated device, although manual grooming can suffice). Scientific measurement of the resulting static charge allows for quantification of product efficacy; although an estimate can be made from simple visual observation of the static “bloom” that occurs.

It was initially noted that this condition relates to the ease with which electrons can be exchanged between hair and the brush or comb. Therefore, results will have a dependence on the material used to manufacture these implements. Combs and brushes may be made from a variety of substances, and the ability to give up or acquire electrons is indicated by their relative positioning in what is termed the triboelectric series.

## 19.13
### Shine

Technically, shine is taken to be a measure of the ability for light to reflect cleanly off a surface. Therefore, as the cuticle structure degrades, or as materials build up on the surface, a dulling of the hair occurs. As such, it is widely accepted that the shiniest state attainable is that of clean healthy hair. It can also be seen how shampoos are inherently considered to improve shine, as dulling deposits are removed — indeed, measurements of increased shine have been used to indicate the cleansing power of different formulations. However, in theory, deposits of conditioning or styling ingredients would be expected to reduce shine, by hindering reflection from hair’s otherwise shiny surface. This said, another major contributor to shine is alignment — where light reflects more cleanly off sleek,
straight styles that involve highly aligned fibers. As such, it can be argued that products which help alignment (i.e., conditioners and styling products) are able to boost shine. This illustrates the ambivalence that can occur in the hair care world – particularly when attempting to generate compelling product claims.

The ability for light to reflect off the surface of a single hair fiber can be measured by a goniophotometer. Cleanly reflected light will bounce off a surface at a 90° angle to the incident light (specular reflection), while light that undergoes some degree of scattering will reach the detector at a variety of other angles (diffuse reflection). A number of equations have been proposed which attempt to quantify technical shine through various ratios involving the relative amounts of specular and diffuse reflection; but how well these technical measures agree with consumer observations is debatable.

While, single fiber measurements provide fundamental information, there is clearly the need to also evaluate hair arrays. This is often performed by an approach that collects images of the shine band while using both polarized and nonpolarized light. With parallel polarizers in place, all reflected light is captured in an image; but perpendicular polarizers eliminate the specular reflection, thus allowing only the diffuse reflection to be observed. Consequently, subtraction of the two images allows the specular reflection to be evaluated, and these two variables can again be utilized in conjunction with the various shine equations to provide quantification. This said, such measurements are usually performed on hair that has been anchored down in an aligned state, and so, this approach still does not properly capture the effect of alignment.

Another major contributor to technical shine is hair color. High levels of pigmentation (i.e., dark hair) prevents light from penetrating into hair, where it becomes scattered, and ultimately reaches the detector at angles other than that of the incident beam. Consequently, darker hair leads to higher proportions of specular reflection and thus higher shine results.

19.14 Color Fade

In recent years, it has become very popular for hair care products to make claims involving protection against color fade. These focus predominantly on chemically dyed hair, where there is a well-know tendency for some degree of color change as a function of washing and/or exposure to the sun’s UV rays. Permanent color products work by initiating complex reactions between dye precursors, which gradually diffuse into the hair. Such reactions produce larger molecules which possess the desired color, and also have a reduced potential for diffusing back out of the hair. Nevertheless, some color “bleeding” still occurs upon washing, which results in a gradual change in the overall color.

One approach to “color protection” involves the use of less aggressive shampoo surfactant systems. That is, milder surfactants are speculated to be less effective at removing residue from the hair, and consequently will be less likely to induce color fade. Another frequently proposed approach involves the deposition of hydrophobic materials – where it is speculated that a barrier is formed on the hair surface, which helps “seal-in” the dye; or, at least, helps slow down any dye diffusion from the hair.
Hair color can be measured using commercially available colorimeters, most often using the CIELAB (L, a, b) system—where “L” refers to the lightness on a scale of 0–100, “a” denotes the red–green color range (positive value denotes higher red) and “b” represents the yellow–blue color range (positive value denotes higher yellow). As such, color is quantified in three-dimensional space, and color change can be evaluated as $\Delta L$, $\Delta a$, and $\Delta b$, or as an overall color change, $\Delta E$.

That is,

$$\Delta E = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2}.$$  

Such experiments are best performed on relatively lightly pigmented hair—as dying produces larger differences between the initial and final states. Consequently, there is the potential for higher levels of fade and thus better differentiation between formulations.

19.15 Moisturization

The ability for water to have a dramatic effect on the properties of hair was mentioned earlier, and is possibly the reason behind why the consumer perception of “dry” hair is taken so literally. In actuality, the moisture content of hair is overwhelmingly related to the relative humidity of the environment; and, at best, has only minimal dependence on habits, practices, and product usage. As already alluded to, hair contains higher moisture content at elevated humidity, and a reduced level at lower humidity. As a result, the water content of hair changes continuously as the wearer goes about their daily activities—passing from room to room, building to building, outdoors to indoors—and constantly encountering different humidity conditions.

It is likely that the word “dry” arises through an analogy with skin care, where coarseness and roughness are associated with “dry skin.” However, as we have seen, hair roughness is related to a degrading cuticle structure and is not in any way related to the presence of any partially desiccated state. Thus, as already described, the cure for this symptom involves lubrication, rather than any regulation of the water content. Indeed, when talking to consumers, one notes that the words “conditioning” and “moisturizing” are used interchangeably.

This serves as a good example of a point that has been emphasized throughout this article—namely, that “consumer language” and “scientific language” are not necessarily the same thing. From a technical stance, this may prompt a desire to reeducate the consumer regarding these misguided assumptions. However, it must be remembered that the ultimate goal is to sell products; and to this end, it is usually easier to stick with “consumer language.” Therefore, when consumers demand a “moisturizing” treatment to help with their “dry, damaged hair,” manufacturers provide lubricating treatments—which eliminates the condition, and leaves the end user reflecting on how their “moisture balance” has been restored.
Shampoos and conditioners are complex mixtures of water, surfactants, cosurfactants, salts, polymers, and oils, whose functionality resides in the scientific discipline of surfactant and colloid science. Skilled formulation chemists have a communal pallet of such ingredients at their disposal, from which functional and aesthetically pleasing products are developed. However, such formulas usually do not stray too far from tried and tested compositions; and so, despite specifically designed differences in the “strength” of certain variants, the benefits and aesthetics associated with a specific product will generally be very comparable.

For this reason, differentiation between the multitudes of products packed onto supermarket shelves must be obtained by alternative means. As has already been touched upon, maybe the most important ingredient in the whole formulation is the fragrance. If a new product is put in front of a consumer, the first inclination will likely be to unscrew the cap and smell it. The fragrance can literally make or break a product; which is why it is often one of the most expensive ingredients. This said, a new product must stand out on a shelf sufficiently well to entice a new customer to pick it up in the first place. As such, attractive, attention-grabbing packaging is also a critical component of a successful product. Indeed, premium brands generally achieve this status by investing in high end fragrances and packaging, while most often utilizing relatively standard formulations.

With all this said, perhaps the main way of differentiating products is still through the marketing position. Successful products generally have found a communication strategy that resonates with consumers. Very often, this involves some aspect of functionality – perhaps specifically touting one of the afore-mentioned consumer attributes. Conversely, some brands prefer to use a holistic approach that pushes the overall pleasurable experience associated with their products (i.e., package, fragrance, and aesthetics). Tied in with this approach, it is common to see exotic-sounding ingredients being included in a formulation as a way to entice consumers, while also providing a level of differentiation. In reality, such ingredients are generally included at very low levels, and should not be expected to provide any functionality. Marketing strategies change continuously with the times, often being influenced by factors such as lifestyle changes and/or social issues. At the time of writing, an ever-increasing awareness of environment issues is carrying over into the cosmetic industry – with the concept of natural products and ingredients becoming very hot.

In summary, while hair styles and social behaviors will continue to change, the recipe for a successful hair care product remains the same – one part science, one part art, all mixed together with a healthy dose of marketing.

Acknowledgments The author would like to acknowledge discussions and conversations with many friends and colleagues throughout the industry in preparing this article. Thanks also to Anthony Ribaudo, TRI Staff Microscopist, for supplying the SEM images; and to Dan Loren at International Hair Importers for permission to use the images of hair tresses and hair blending.
References

Core messages

- Antiperspirants and deodorants are two separate entities with regard to their definition, mechanism of action, and legislation.
- The most important methods used for efficacy proof of antiperspirants include gravimetric, biophysical, and imprint casting assessment as well as visual evaluation.
- Different designs for antiperspirants testing exist depending on the study objectives.
- Standardization of the criteria for selecting study population, measuring conditions, anatomical test site, and timing should be considered when performing antiperspirants testing in practice.
- Multiple products test design is useful in assessing relative differences between products and allows ranking of products and comparison with a standard of known efficacy.
- Clinical evaluation, that is, sniff test, microbiological analyses, and chemical chromatographic analyses are applied in deodorant testing.
- Sensory evaluation of the armpit malodor is still present in the routine procedure for deodorants efficacy testing.

Abbreviation

TEWL       Transepidermal water loss

R. Darlenski
Department of Dermatology and Venereology, Medical University, Sofia, Bulgaria
e-mail: darlenski@abv.bg

J.W. Fluhr
bioskin GmbH, Bergmannstraße 5-7, 10961 Berlin, Germany and
Department of Dermatology, Charité University Clinic, Charitéplatz 1, 10017 Berlin, Germany
20.1 Introduction

Sweating also named perspiration is the physiologic process of the secretion of water together with the diluted sweat compounds through the openings of the sudoriferous glands. Perspiration is a key element in the metabolism and contributes to the body temperature homeostasis. The transformation of sweat constituents by skin microflora plays an important role as a causative agent in the body malodor.

Sweating especially in the underarm region and the unpleasant body odor influence the individual’s self-confidence and thus can result in hindrances in social adaptation. In addition, certain diseases such as hyperhidrosis can be the cause for excessive sweating decreasing the quality of life in the suffering patients. The anatomical skin sites with highest density of eccrine and apocrine sweat glands are shown schematically in Fig. 20.1.

Diminishing and masking the unpleasant smell of the sweat dates since ancient times. Nowadays, antiperspirants and deodorants are used almost ubiquitously and present a necessary feature of the everyday cosmetic care. Beyond developing medications and procedures which decrease sweat gland activity, the market for cosmetics with antiperspirant/deodorant activity is constantly growing.

![Diagram showing anatomical skin sites with higher density of apocrine (axillae, groin) and eccrine (palms, soles, forehead, back) sweat glands](image_url)
Claim support documentation is an essential part in the registration of cosmetic products and for antiperspirants/deodorants in particular. Thus, proving the efficacy of such products is a binding procedure not only for marketing purposes, but also for legislation issues.

20.2 Antiperspirants and Deodorants: What Are They?

There is no a clear-cut differentiation between antiperspirants and deodorants. In general, an antiperspirant is a product designed to reduce the sweat production at the site of its application. In the broad sense, antiperspirants act as deodorants themselves by decreasing the amount of substrate for malodor generation by the skin microflora. Deodorants are cosmetics applied to reduce the body odor by their absorbing, masking (fragrances), and antibacterial properties.

A variety of molecules have been implemented as antiperspirant ingredients. The main applied principle is to block the openings of the sweat gland thus decreasing the evaporation on the skin surface. The most widely used ingredients are the metal salts such as aluminum chloride hexahydrate. However, certain agents namely zirconium salts have been banned due to the increased risk of skin granuloma formation. Other approaches include the application of astringents (glutaraldehyde, formaldehyde), film-forming polymers, and the combinations of surfactants and cosurfactants demonstrating swelling properties (oleic acid and glycerol monolaurate) [8].

In the US, antiperspirants are classified as over-the-counter medications (while deodorants are accepted as cosmetics). A list with allowed active ingredients is implemented [5]. In Europe, both antiperspirants and deodorants are regarded as cosmetics [4]. Nonetheless, efficacy claim support is necessary for all cosmetic products manufactured and marketed in the European Union.

20.3 Testing Efficacy of Antiperspirants

20.3.1 Test Methods

Different approaches have been applied in testing antiperspirant efficacy such as visual assessment, casting methods, gravimetric techniques, and the application of biophysical measurements.

The principle behind optical methods is the visualization of stained spots corresponding to the sites of sweat droplets on the skin surface. A variety of staining agents have been implemented such as bromophenol blue, Prussian blue, and rhodamine (reviewed in [1]). The most popular staining method is the starch-iodine test in which the active sweat glands are visualized as small dark blue spots. Visual techniques offer the possibility to measure
the sweat excretion rate by taking into account the volume of each sweat droplet. However, evaluation of different staining is generally accepted as inaccurate although being the most easily performable. Remnants on the skin surface, accumulation of the single droplets into larger conglomerates, and the anatomical specifics of defined body regions (such as the axilla) influence the method and increase the risk for misinterpretation of the results.

Casting methods are based on the application of a hydrophobic material (e.g., silicone rubber, and plastic solutions) [6]. The secretion of sweat droplets leaves imprints in the covering hydrophobic material. Further analysis of the replica can be achieved by photographic evaluation, the application of specific light sources (UV), and visualization techniques such as scanning microscopy. A key source for mistake in casting methods is the formation of air bubbles in the replica.

Gravimetric methods are the “golden standard” for evaluation of the quantity of the secreted sweat and its reduction by antiperspirants. In its variants, absorbing material (e.g., cotton pads) is weighed before and after the product application. The reduction of sweating is calculated by the difference in the mass of the absorbing pad in comparison to a baseline value and/or a control (untreated corresponding) skin site. A drawback of this method is that a minimum of collected sweat is required for adequate quantification. Gravimetric tests are recommended on skin areas where a large amount of sweating is expected. Precise technical equipment and trained personnel are mandatory for adequate following of the protocol.

Biophysical methods are the most sensitive and with greater discriminative properties [2]. Devices designed for the assessment of transepidermal water loss (TEWL) from the skin can be applied in the evaluation of eccrine sweating. A continuous mode measurement is applicable. However, the test site is demarcated by the dimensions of the probe head. This method is strongly influenced when there is excess sweating and when other factors enhancing TEWL (e.g., epidermal barrier disruption) are present. Assessment of the cutaneous electrical properties (capacitance, conductance, impedance) is applied to evaluate the sweat gland activity, as the electrical conductivity of the skin is proportional to the water content on the skin surface. The drawbacks of these methods overlap with those valid for TEWL assessment.

Further methods such as cyanoacrylate skin surface biopsy and the application of injectable dyes have been implemented. Due to the need of validation and/or ethical considerations, these techniques remain in an experimental phase in efficacy testing of antiperspirants.

20.3.2
Test Design

Different approaches exist for testing the efficacy of antiperspirants. The major objective is to assess the reduction of sweating by the application of cosmetic products.

20.3.2.1
Gravimetric Evaluation of the Armpit Sweating

The most commonly applied design involves the gravimetric evaluation of axillary sweat collected on absorbing pads at defined intervals after the product application. A paired
comparison test design is applied where one axilla is treated with the test product in a standardized manner and the other one is left untreated or placebo/vehicle is applied. When comparing two products, both axillae are treated with the different cosmetics and then a comparison between both sites is performed. A major disadvantage of this design is that only two products/treatments can be compared in a single panelist. The number of subjects required exceeds the one applied in other protocols. In addition, seasonal variations in the antiperspirants’ effect of the same formulations tested on the axillary vault was shown while no such difference was observed in the simultaneous conduction of the test on the back of the panelists [3].

Study Population

The number of panelists in this test design varies from 30 to 60 adult subjects. The gender characteristics of the study population do not affect the study design. In general, the selected subject should correspond to the population to which the product is targeted. Volunteers should not have any dermatoses in the armpit region. A general medical check-up is performed before entering the study. The presence of systemic disease (infections, HIV, AIDS, autoimmune disorders, neoplasms, neurological, hematological, and metabolic disorders) as well as skin conditions (psoriasis, eczema, bullous disorders, tinea, and bacterial infections) should be considered as exclusion criteria in the recruitment phase. The intake of medications that can potentially interfere with sweat secretion (e.g., cholinergic agents) and/or can suppress an eventual reaction to the antiperspirant (e.g., glucocorticosteroids and immune suppressors) is a contraindication for study enrolment. The use of any systemic and topical antibiotics should be discontinued at least 2 weeks prior to the study. Issues such as written informed consent, inclusion/exclusion criteria, and over-recruitment are dealt with in other chapters of this book, but are valid and should be considered when testing antiperspirants.

A minimum secretion range of 600 mg/axilla for 20 min must exist between the lowest and the highest sweat rate among the volunteers. Subjects with sweat secretion lower than 150 mg/20 min per axilla are not enrolled in the treatment phase. Shaving of axillae should not be performed at least 48 h before the study.

The Conditioning Phase

The test phase is preceded by a conditioning period of 17 days. During this phase, the application of axillary antiperspirants is not allowed to the panelists. A mild hygiene product (e.g., soap bar that has no antibacterial properties) is provided for washing during the conditioning. Thus, a potential interference with the volunteer’s own product is avoided.

Test Procedure

Sweat collection can be performed either in ambient conditions (conditions as the daily routine ones) or in a controlled setting such as the hotroom – temperature $37.8 \pm 1.1^\circ C$ and
relative humidity 35–40%. A full recording of the microenvironmental conditions on regular intervals is necessary in this case.

During sweat collection, the subject should not experience psychological stress. Subjects should wear large cotton T-shirts which makes access to the axillary vaults easier. Panelists must be seated with both feet on the floor and arms resting against their sides symmetrically.

A baseline (before treatment) sweat collection is advised although not being compulsory. Cotton pads (cotton fabric) are used to collect the axillary sweat. The sweat collection is performed for 80 min. The first 40 min are considered as acclimation (warm-up) period and the collected sweat in this period is discarded in the data evaluation. In the next 20 min, the sweat is collected by the absorber pads applied by trained personnel. After the end of this period, each pad is placed in capped and marked plastic vial, and then weighed. The procedure is repeated with another pair of cotton pads for the last 20 min. In case the test is performed in ambient environmental conditions, a 3- to 5-h duration of collection is required for a sufficient quantity of sweat to be collected.

A supervised washing of each axilla is performed by using disposable towels soaked with a standardized washing solution, then rinsing and wiping with a dry towel. The procedure is repeated before each application of the product. Bathing and washing by the panelist is forbidden during the test phase.

Before applying the formulation, a randomization of the treated site (left/right) is performed. A controlled amount of the product is applied in a standardized manner depending on the product type (spray, bar, roll-on) and as advised by the manufacturer. The area of application is delineated to the axillary vault and has the size of 6 × 12 cm. After the application of the products, the subjects stay at the laboratory for least 30 min to allow drying of the applied products and the absorption of the ingredients. Four applications are generally performed, most commonly in four consecutive days. In modifications of this protocol, the intervals between applications can be shortened, for example, twice or 4 times daily.

Generally, sweat collection is performed 24 h after the last product application. However, different time points can be selected, for example, 1, 2, 4, and 12 h after the application. Both the pads from the treated and the control axilla (untreated/placebo) are collected, sealed, and weighed.

**Data Evaluation**

At least 20% sweat reduction must be demonstrated for a formulation to be determined as an effective antiperspirant. Sweat reduction by the product is assessed by applying Wilcoxon signed rank test when baseline values have been obtained, and Wilcoxon rank sum test is used when no data for the baseline values exist. When comparing two products, either covariance analysis is performed (available baseline data) or a statistical analysis is conducted on each of the posttreatment evaluations (no baseline data) [8]. In general, an experienced statistics specialist should be consulted in the data evaluation process as well as in the planning phase of the study.
20.3.2.2

Multiple Product Test Design

Time- and cost-effectiveness are major determinants in cosmetics testing. Therefore distinct methods have been proposed for testing multiple antiperspirants. A specific protocol for testing during early product development was introduced [2,9]. The design allows direct comparisons of the efficacy of up to eight test formulations within 1 week or variations of this time period.

Study Population

Each panel includes 20–22 subjects. The general considerations described in the above test are also valid for this study protocol.

The Conditioning Phase

No conditioning phase is required as the test sites are located on the back of the subjects.

Test Procedure

Sixteen test fields (size 4 × 5 cm each) are arranged on the back in a 4 × 4 matrix. Treatments are randomly assigned to the test fields on the left or right side. A defined amount of test product is applied to each of the treated test fields. Even distribution over the test field is performed in a standardized manner depending on the type of product (solution, stick, spray, powder). The test fields are left open for 5 min before covering with occlusive patches. The conditions during occlusion and the length of occlusion vary according to the study protocol. Treatments are usually performed on four consecutive days.

Twenty-four hours after the last application, preweighed pads for absorption of sweat and an occlusive covering to prevent evaporation of sweat are fixed over the test fields. Thermal stimulation is then performed in a sauna at 80°C for 15 min. Immediately after leaving the sauna, the pads are removed and weighed.

Data Evaluation

The variable for statistical analysis is the relative reduction of sweat in the treated test area compared to the corresponding untreated control area. Descriptive statistics including mean, median, standard deviation, minimum, maximum, and the 95% confidence interval are calculated. Effectiveness is proven if the lower 95% confidence limit is greater than zero. Products are ranked according to percent reduction of sweat. Comparisons between products are made using inferential methods defined in the study protocol (e.g., paired t-test, ANOVA).
Variations of the protocol in the evaluation of sweating (e.g., electrical conductance measurement of eluted sweat-saturated cellulose pads and the imprint casting evaluation) and with regard to the application site (e.g., volar forearm) have been described [2,6]. In general, these tests are useful in assessing relative differences between products and allow ranking of products and comparison with a standard of known efficacy. They are suitable for screening for new formulation development. However, there is no sufficient evidence for efficacy extrapolation of the data to the axillary vault.

20.4
Testing Efficacy of Deodorants

20.4.1
Test Methods

Different approaches are used in testing the efficacy of deodorants. As malodor is mainly produced by sweat breakdown by skin microflora, in vitro and in vivo microbiological analyses are performed to assess the inhibition of bacterial growth by the deodorant substance. Gas chromatography can be used in the evaluation of the chemical constituents responsible for the malodor. However, this method does not disclose the complexity of malodor composition as great interindividual variability of the chemical composition of malodor has been demonstrated [1]. Sensory assessment, that is, sniff test is accepted as a standard evaluation procedure for in vivo testing of deodorants.

20.4.2
Test Design

20.4.2.1
Sniff Test

Different designs have been described in performing the sniff test such as single-pair, each vs. control, and round robin. In the standardized manner, one axilla is treated with the test formulation, and the contralateral (control) is left untreated or a placebo is applied.

Study Population

Thirty to sixty volunteers are sufficient for performing the test. The requirements for the panelists overlap with those for antiperspirant testing. In addition, the subjects should not be smokers as this can render difficulties to the judges (sniffers). Each subject must have a minimal score (according to a grading scale) of baseline perceptible odor.
The Conditioning Phase

The duration of the conditioning period varies from 14 to 17 days and, as in antiperspirants testing, the subjects should not use any product in the axillary vault. A mild nonantibacterial soap is provided for cleansing the axillary region.

Test Procedure

A controlled washing similar to the one in antiperspirants testing is performed before entering the study. During the evaluation of the baseline perceptible odor as well as during the whole test period the subjects wear cotton T-shirts. The T-shirts are machine-washed, one cycle with a fragrance-free detergent followed by a cycle without detergent (rinsing). Adhesive pads are fixed to the axillary vault and stay there for a certain period from 4 to 8 h. During this time panelists return to their daily routine. After that, they return to the testing laboratory where the pads are removed by the investigator and are collected in capper containers. Malodor evaluation can be performed directly after the pad removal and 24 h later. After performing the washing the product is applied randomly in a standardized way. Then the pads are attached again to the armpit and the evaluation is performed 4–8 h later. A key issue in the sniff test are the judges also named sniffers. The sniffers’ panel consists of 3–4 trained judges. The selection and the training of the sniffers are crucial for the accuracy and the reproducibility of the study results. A protocol for screening and training for odor testers have been proposed based on the international standards and guidelines [7]. Gender and age should be considered when selecting a sniffers panel, as olfactory perception decreases with age as well as women and men differ in their odor perception. Therefore, at least, the gender distribution should be equal in the panel of judges. The judgment of the odor intensity is generally performed by a grade scale. It is advised to avoid direct contact between judges and panelists as this can influence the sniffers opinion. However, in some protocols, judges evaluate the odor directly from the volunteers’ armpit.

Data Evaluation

The reduction of the odor score obtained by scaling techniques is investigated. The values from the treated armpit are compared to those of the untreated (control) axillary vault.

20.5 Conclusion

The market for antiperspirants and deodorants is constantly expanding. Efficacy testing is mandatory not only for legislation reasons, but also to demonstrate the superiority of one product to another. Hence, comparative and multiple product testing is the field for active
and promising investigations. The selection of the proper test method is based on the study objectives, and in particular, on the aim of the investigation, that is, what is targeted as a proof of concept.

In deodorants testing, both scientific community and industry are still standing far from objectivity and standardization in efficacy proof. Developing noninvasive, cost- and time-effective and precise methods and equipment is the next step in this scope.

References

Core Message

- To perform a clinical hair trial under good clinical practice (GCP) conditions for both cosmetic and medical applications, a placebo-controlled, double-blind, randomized study design with use of appropriate study parameters and methods is most recommended. This chapter gives an overview over the most suitable and common noninvasive, semi-invasive and invasive evaluation methods. Clinical evaluation/grading and hair pull test are basic methods to screen for diagnosis and hair loss dynamics. The daily hair count and hair wash test are methods performed by the patient himself giving a semi-quantitative orientation about hair loss, suitable for home-in-use-tests. Hair weighing is a noninvasive, sensitive, highly reliable, but effortuous method, which gives information about hair volume (i.e., weight), while the trichogram is a semi-invasive method and indicates the activity of hair loss, but shows a high variability. Nowadays, the (contrast-enhanced-) phototrichogram and/or epiluminescence-based digital computer-phototrichogram (TrichoScan) are the most reliable and objective tools to measure quantitative and qualitative parameters of hair growth and loss. However, overall hair appearance is best assessed by standardized global photographs. Optical coherence tomography, electron microscopy and confocal laser scanning microscopy of hair are highly sophisticated methods which are reserved, however, for addressing very specific questions. Thus, the choice of methods for a study should be carefully respect test-hypothesis, practicability, time-consumption, variability and accuracy.
21.1

**Keypoints**

Practical points for performance of a hair study

- Set well-defined clinical hair loss stages for building a study population representative for a specific diagnosis with specific disease dynamics
- Choose study duration between 6 and 12 months
- Choose study parameters sharply adapted to the hypothesis of the study
- Consider invasiveness of methods vs. possible scientific outcome
- Choose a method triplet of clinical, metric and cosmetic quality (e.g., hair pull test, phototrichogram, global photographs)
- Consider statistical aspects for power calculation regarding effect size, number and time needed to treat

21.2

**Introduction**

How to set up a scientific study to assess cosmetic aspects of hair? This is an ever difficult to answer question, since hair cosmetic means hair appearance which is a result of many parameters such as the single hair shaft form, strength, flexibility, hardness and smoothness which in sum of all quality parameters and sum of all hairs on the scalp lead to that positive appearance that is called “beautiful hair.” The line between cosmetic and medical hair aspects is blurred, and in most cases, however, studies are mostly performed not to improve the quality of single hair parameters, but to prevent hair loss and to increase hair density and volume. Therefore, most hair assessment parameters measure hair growth and density and assess different hair types (anagen-, telogen-, vellus-, terminal-hair).

Knowledge of the hair follicle anatomy and the dynamics of hair cycling is substantial for a hair investigator. Recognizing the anagen, catagen and telogen phases clearly characterizes the typical hair chronobiology. Physiological factors of hair growth are metabolic stage, seasonal biorhythms, hormones, neuromediators, biomolecules, micro-inflammation and aging [22]. The influence of these factors has to be taken into consideration when designing a study. Various clinical hair techniques can help in assessing the efficacy of drugs and cosmetics agents on hair growth. Great advances have been made during the recent decades in the methodology of hair growth trials in dermatology and cosmetology [11, 22, 28]. Clinical evaluations benefit from a number of additional specific techniques that enhance the perception of hair growth, shedding and alopecia.

The techniques to evaluate hair growth can be generally classified as:

- **Noninvasive**, e.g., global photographs (GP), daily hair counts (DHC), hair wash test (HWT), hair pull test (HPT), phototrichogram (PTG), epiluminescence microscopy of hair (ELMH), electron microscopy (EM), laser scanning microscopy (LSM)
• *Semi-invasive*, e.g., trichogram (TG), unit area trichogram (UATG)
• *Invasive*, e.g., biopsies

Certain tools are best suited for diagnosis and follow-up in private practice, whereas others are rather applicable for monitoring hair growth under different treatment conditions in clinical studies under GCP conditions.

Methods can be also classified as subjective and objective. For use-test studies as performed e.g., by the dermatologist in private practice, body and scalp hair distribution can be assessed by application of different grading systems, DHC, the HPT and dermoscopy. In addition to these basic diagnostic techniques, hair weighing (HW), TG, computer-assisted PTG or ELMH might be used as state-of-the-art techniques for objective clinical studies [11, 22, 23, 25, 28]. Taking scalp biopsies is not a frequently used technique in clinical studies in Europe, whereas in the USA, several studies have been performed using this method [32]. For research purposes only, optical coherent tomography, EM and confocal LSM are optional tools [10]. How these parameters are tested and how a study for testing hair loss is technically and logistically planned is described in the following.

### 21.3 Study Design

The gold-standard for a hair study – even in a cosmetic context – is a placebo-controlled, randomized and double-blinded study design. The duration of a hair study should be at least 6 months, since hair growth is slow (0.3 mm/day) and first visible and/or measurable changes during treatment with a specific substance usually occur earliest after 3 months. In the majority of studies, a significant and biologically relevant effect has been observed between 6 and 12 months of treatment [16, 19] while others have observed significant effects already at 6 months [2, 5, 18]. Volunteers/patients visits should be scheduled for every 3 months or every 6 months.

### 21.4 Assessment of Clinical Appearance (Investigator's Questionnaire)

A standard clinical questionnaire to assess clinical appearance and quality of hair and scalp skin, performed by the investigator should be included in every cosmetic study about hair growth. The investigator’s questionnaire should contain clinical hair parameters which are assessed optically and by palpation in a five-point scale:

- General appearance of hair
  - Volume: full – medium – small
  - Hair density: dense – thinned/shed
  - Hair reflexion: shiny – blunt
  - Hair plasticity: waved – flat
• Appearance of scalp skin
  – Redness
  – Roughness
  – Scaliness

Basic diagnostic tools are the following:
• Clinical grading systems
• Daily hair count
• Hair wash test
• Hair pull test
• Dermoscopy

21.5
Clinical Grading Systems

Clinical grading systems have been established mainly for androgenetic alopecia (AGA) of men and women. For grading male pattern alopecia androgenetica, Hamilton and Norwood have established the most accepted classification of clinical grades which respects the different grades of temporal and vertex thinning as well as specific balding in the central frontal part and hair thinning grades in the upper-middle head region [8, 21]. In women, the female-pattern AGA is graded in a three-point scale after Ludwig [20]. The Ludwig-classification refers to the different grades of diffuse, but circumscribed thinning in the upper-middle region of the head. The Gan-Sinclair scale (5 grades) and the Savin-scale with eight different grading classes define more specifically the different clinical expressions of hair loss patterns and grades [10]. The grading system should be used before every study to clearly define the patients’ population which should be investigated under treatment with a specific substance, and it is recommended to include the classification stages in the inclusion and exclusion criteria of the study.

21.6
Daily Hair Count

This is a basic test performed by the patient himself after instruction by the investigator. The test consists simply in instructing the patient to count his lost hairs every day over a certain period of time, e.g., 3 or 5 days. The number of lost hairs consists of hairs which are found in the comb and after hair washing or which are laying on the clothes and furniture every day.

The latter two parameters, though, are rather at risk to have a high degree of variation due to inconstant searching and detection of hair on e.g., light and dark textiles or daily variances of searching intensity. Two standards of daily hair count can be differentiated: The first is the counting of lost hair throughout the whole day and at all places where lost
hair can be found. The second is the exclusive counting of hairs which are lost by combing and washing hair at one specific time point of the day, e.g., in the morning. The standardization of the latter is much higher than the first and gives semi-quantitative values which correlate well with hair loss activity. Besides objective techniques this test is a valuable method which can be performed at different time points of a study. Additionally, it supports compliance since patients themselves like to control their hair loss.

### 21.7 Hair Wash Test

The HWT counts hairs which are rinsed off after hair washing. After a standardized 5-days abstinence from shampooing, the patient is asked to wash his hair in a basin with a gauze-covered hole to collect all rinsed off hairs. The collected hairs are grouped in three hair length classes (1) long hair >5 cm, (2) intermediate hair >3 and <5 cm, (3) short telogen or vellus hair <3 cm [24]. This method differentiates between long terminal hair and short telogen or vellus hairs. However, this method can lead to false-high hair counts due to double counting of broken hairs, it cannot be performed in individuals with short or curly hair and it is very time-consuming [10].

### 21.8 Hair Pull-Test

This test allows to obtain a semi-quantitative clinical impression about the epilability of scalp hair, i.e., how active the dynamic of hair loss is at the time point of the investigation: approximately 60 hair shafts are taken between thumb and index finger close to the surface of the scalp skin and pulled firmly, but not forcibly away from the scalp with constant strength along the hair shaft up to the upper hair tip. Hairs which are epilated under this procedure are counted. In contrast to others [10], the author differentiates between “clearly negative,” “slightly positive” and “clearly positive,” since there is an intermediate stage in which there is no massive hair loss, but still more hair loss than normal. Therefore, the pull-test is “negative” (no active hair loss) between none and three epilated hairs, “slightly positive” between three and six epilated hairs and “clearly positive” above six hairs (>10% of tested hairs). The pull-test can be well used to assess the inclusion criteria “active hair loss” for a study and represents also a valuable evaluation parameter to differentiate between the clinical diagnosis of AGA and alopecia diffusa (AD). For active AGA, the pull-test has to be proven positive in the temporal area and/or vertex area (in male pattern AGA) or in the upper-middle head region (in female pattern AGA), but negative in the occipital scalp area. In contrast, for AD the pull-test has to be proven positive in temporal, parietal and occipital area.
21.9
Hair Weighing

For standardized assessment of hair weight, a permanently marked area on the target region of hair loss is assessed by using a plastic template with standardized diameter (e.g., 1 cm). Hair from this area is pulled through the template and the outline of the template drawn on the scalp. Then, the hair is hand clipped and collected carefully at every assessment time point of a study. Subsequently, the hair is weighed and counted in order to estimate cumulative weight of all hairs within the marked area and weight per single hair. Additional parameters such as average hair shaft length and average optical width can be assessed as well. In a study by Price and Menefee, a randomized treatment group (2% minoxidil) and a 50-hair sub-sample was assessed for weight per hair, average length and optical diameter width. The study showed that cumulative hair weight and hair count of the hairs of the assessed area were significantly changing under treatment, whereby average weight, diameter and length from the 50-hair sub-sample showed no significant change [23]. Additionally, it was observed that cumulative hair weight lead to a larger change than total hair count, which implies that factors other than the number of hairs, such as increased growth rate (length) and diameter of hair shaft might contribute to an increased hair weight. Since the cumulative hair weight method is not only easy to perform, but also less susceptible to standard error during sample collection and measurement, this method is a highly reliable and objective method to assess hair growth in clinical trials [23]. However, the investigation process is time-consuming and its use has to be critically considered in balancing the advantages and disadvantages of the method.

21.10
Trichogram

The TG is a standardized light microscopy investigation of the roots of plucked hair with root-typing and counting of anagen-, catagen-, telogen and dystrophic hair. There are generally two areas where a TG should be performed: in AGA, it is performed in the active area, i.e., temporal in male pattern or top of head in female pattern AGA and in the occipital region as a nonaffected reference area or to proof diffuse alopecia (Fig. 21.1a, b). In alopecia areata, one TG is taken at the edge of an affected lesion and one at far distance to it (e.g., contralateral side). To perform the TG, the hairs of the target region are combed in a longitudinal stripe-like way, then taken with a rubber-armed surgical forceps and then plucked with an immediate strong force within half a second, directed away from the patient’s head. Since the patient feels a short, but sharp pain, he/she has to be prepared before the plucking act. After plucking, hairs are placed with their roots into a xylol gel onto a glass slide, covered with a cover glass and then dried for 24 h. After the drying period, hair roots are counted under the light microscope at 40× magnification.
Until the early 2000s, this technique was a possible method to evaluate hair loss activity as percentage ratio of anagen/telogen hairs and to assess treatment effects of hair growth promoting substances in clinical studies [5]. The TG reflects very well the activity of hair loss as represented by percentage of telogen rate which also corresponds well with the results of the clinically performed pull-test. However, there is a slight draw-back of the method, since the hair cannot be plucked at exactly the same place for pre-post comparison, since the hairs including their roots have been removed at the first time point. Therefore, there is a certain degree of variation, which however, can be kept homogenous throughout a study population, when all TGs of a study are analogously taken at a defined neighboring region of the first TG area. Comparable but noninvasive methods to assess anagen/telogen-rate are, the PTG and the epiluminescence microscopy (ELM) of hair which are described in the following. Regarding the above mentioned limitations of TG, a study with 12 volunteers has shown that there are similar assessments of anagen hair in TG and PTG, while there was an underestimation of 181 vs. 237 hairs/cm² of total hair density by the PTG method [25]. It was further found that hair diameter measurements from the PTG were not reliable.
21.11
Unit Area Trichogram

The UATG is slightly different to the above mentioned standard TG, since it assesses different hair parameters from a defined area on the scalp. Hairs are plucked usually from an area of >30 mm² and then counted and measured. Four parameters can be assessed: hair count per mm², anagen rate, hair shaft length and diameter [10]. The latter three parameters are evaluated under the microscope. The UATG is a method to assess hair changes in study populations under treatment (pre-post comparison) to observe hair cycling influences or other effects of topical or systemic substances. While the method is quite accurate, it is time-consuming and therefore not very suitable for larger clinical trials.

21.12
Phototrichogram

The PTG is a further development to classical TG and has been first proposed by Saitoh in 1970 [27] and then been further developed by Bouhanna in 1988 [3] and Van Neste [30]. The PTG is a noninvasive, reproducible method which is based on manual marking of shaved hairs on images taken at close from target areas on the scalp skin which show hair loss (e.g., temporo-frontal and vertex in male pattern AGA). There are several ways to perform PTG such as fully-manual, semi-technical and fully-technical of which the TrichoScan is a fully-technical method which is widespread over dermatology offices and trichological centers in university departments in Europe [6, 11]. The reference area to be assessed is a 0.50-cm² scalp skin area which should be marked with a 0.2-mm wide black micro-tattoo. Nowadays, learning from TrichoScan (see below), it is recommended to use red ink, since red is a more natural color on the scalp skin, resembling e.g., a micro-hemangioma. With this technique, seven quantitative parameters can be assessed in the test area: density of hair follicle implantation, anagen percentage, telogen percentage, growth rate, mean anagen diameter and percentage of fine hair less than 40 μm in diameter (vellus hair) and percentage of thick hairs with more than 40 μm in diameter (terminal hairs). The normal measurement values for a male adult have been measured as follows (vertex): density 204±10 hair/cm², telogen rate 17.8±2.8%, growth rate 0.35±0.03 mm/day, mean anagen diameter 76±5 μm, percentage of fine hair 9.2±1.8% and percentage of thick hair 90.8±2.1% [6].

The great advantage over classical TG is that the assessment of hair can be taken as pre-post comparison at exactly the same location, while the hair in the TG method is removed by plucking and therefore not observable in exactly the same place. However, there is still intra- and inter-individual as well as inter-study variation in the method itself which has been overcome by several modifications. These are: use of a standardized optical equipment with a rigid frame to ensure a fixed distance from the investigated hair follicle bearing scalp skin to the lens and a superimposed glass window with millimetric graduations. When these vari- ances are reduced by standardized measurement procedures, this method represents a very satisfactory qualitative and quantitative technique to study hair growth, diameter, anagen-/telogen-rate and vellus-/terminal-hair-rate in clinical hair loss trials.
21.13
Contrast-Enhanced Phototrichogram

A weakness of the classical PTG method is the difficulty or sometimes inability to detect less pigmented or very thin hair. A study has shown that analysis of individual hair data obtained from differently colored hair revealed light hair to be much more difficult to evaluate than dark hair. Consequently, it was concluded that Caucasian subjects with light hair or individuals with dark skin and dark hair should be excluded from studies employing PTGs [25].

However, since there are the populations of e.g., Middle and Northern Europe as well as North America, in which a considerable number of studies are performed, a refinement of the standard PTG has been made by developing the contrast-enhanced phototrichogram (CE-PTG). In this method, all hairs in the test area are stained black. The reliability of the CE-PTG has been investigated by comparison to classical TG and to histological transverse section of biopsy samples in male subjects with AGA [28]. Compared to PTG, the CE-PTG method offers significantly improved detection not only of thin, but also of thick hair. The number of thick hair (>40 μm diameter) was similar compared to the number of thick hair detected by histology. On the other side, also very thin/miniaturized hair (approximately 8 μm diameter) was detected by CE-PTG in same numbers compared to assessment by histology. The CE-PTG was able to detect all hair growth stages – anagen, catagen and telogen – as well as the empty follicle stage [28].

Compared to standard TG, the CE-PTG has the advantage that the number of hairs in a defined area can also be assessed, thus offering an additional parameter which is relevant for the clinical assessment of hair. These observations revealed such technological advantages compared to classical PTG as well as to histology which requires an invasive biopsy, that the CE-PTG is appreciated as a very attractive method for detailed, reliable analysis of hair cycling and detection of miniaturized hair. These parameters represent essential features which a method has to cover for the assessment of e.g., AGA, which is the most common hair loss and therefore in the focus of most studies performed nowadays in clinical hair science. However, the assessment of the parameters has to be performed by manually counting, and a computerized and automatized method has not yet been developed. Therefore, the method is time consuming and requires thoroughful and accurate evaluation performance which might affect also inter-observer variability [12]. Nevertheless, the method is used and suitable for clinical hair studies.

21.14
Epiluminescence Microscopy of Hair (TrichoScan)

In search for more reliable and inter-observer independent assessment tools for the standardized investigation of hair quantity and quality with the aim to assess hair loss, a method has been developed combining ELM with digital image recording and software based analysis. This method has been established under the name TrichoScan in 2001 [11]. Trichoscan was developed to assess all important parameters of hair growth such as hair
follicle density (n/cm²), hair shaft diameter (µm), cumulative hair diameter, mean daily growth rate (mm/day), anagen-/telogen-rate and vellus- and terminal hair rate (lower part)

To perform a TrichoScan analysis, a transitional area of hair loss between normal hair and the balding area is chosen and a stencil template with a given diameter (1.7 cm) put over the area to clip the hairs in that area. Normal hair dresser scissors are taken for roughly shortening of the hair shafts and then an electronic shaver (e.g., Moser, Wahl GmbH, Unterkirnach, Germany) is used to clip the hairs to the exact length of about 1–2 mm. For reproducibility and pre-post-comparison, it is essential especially for clinical trials to landmark the investigated area with a central red tattoo (<0.2 mm diameter): a small drop of professional red tattoo ink is placed onto the center point of the shaved scalp skin area and then punched through with a syringe tip into the upper layers of the epidermis. The procedure is very superficial and therefore hardly painful as it has been assessed by inter-investigator self-tattooing by the author and a colleague. In the first studies, black ink was used; however, black interfered with the software detection system which was not the case with red ink. As mentioned before, a red ink tattoo is also cosmetically much more acceptable. To ensure sufficient contrast to the scalp skin as well as low variation of hair color for a high-quality digital image, all hair follicles
within the assessment area are dyed with black hair dye, e.g., RefectoCil (Herbert Gschwentner, Siezenheim, Austria) or Wella-Viva, No 28, Saphir noir (Wella/Procter and Gamble, Weybridge, Surrey, UK) or Goldwell black 2N (Goldwell, Darmstadt, Germany). The hair dye is removed after approximately 12 min with alcoholic solution (e.g., Kodan Spray, Schülke & Mayr, Vienna, Austria). Then, a transparent fluid (e.g., Kodan can be used again) is put onto the area, and the transparent plastic cylinder of an ELM system is placed onto this fluid film and a photo taken (ELM: Fotofinder DERMA, TeachScreen Software GmbH, Bad Birnbach, Germany). Two main possibilities exist for taking the images: a digital video dermoscope at 20-fold (analyzed area: 0.62 cm\(^2\)) and 40-fold (analyzed area: 0.23 cm\(^2\)) magnification (Fotofinder DERMA, TeachScreen Software GmbH, Bad Birnbach, Germany) or a digital camera, e.g., Nikon Coolpix 4500/Coolpix 8400 (Nikon Europe BV, Lijnden, The Netherlands) or Canon Powershot A95 (Canon Europe, Amstelveen, The Netherlands) (analyzed area: approximately 1.5 cm\(^2\)). For good quality pictures, it is essential to place the cylinder evenly onto the slightly convex scalp skin surface, to exclude unclipped long hair lying over the assessment field, and to avoid bubble formation of the contact fluid. A digital picture is taken at 20- or 40-fold magnification and automatically downloaded to the computer into the software system (Tricholog GmbH, Freiburg i. Brsg., Germany). A fully automated camera system with a transparent cylinder put onto the original optical objective ensures constant distance between scalp skin and camera. Together with a ring light which ensures constant illumination of the area, digital pictures can be taken at a very high quality, low variation and time-independent high reproducibility. The software scans and registers every hair of the digital picture and automatically counts the number of hair per cm\(^2\), calculates the anagen/telogen hair rate and assesses the rate of vellus and terminal hair. The parameter diameter of single hair shaft and cumulative hair diameter as sum of all assessed hair diameters are only available in the study version software.

The software was validated by use of more than 500 images, which were taken from study participants. The algorithm excludes air bubbles, dust, small hemangiomas, nevi, scales etc. from the calculation without interfering with the number of detectable hairs.

The detection limit of the software depends on the resolution (pixels) of the digital cameras. Using a video system, hairs smaller than 14 \(\mu\)m cannot be analyzed, whereas with higher resolution 7.0 megapixel cameras hairs with 6 \(\mu\)m thickness can be detected.

The interclass correlation lays at approximately 91% within the same TrichoScan operator and the interclass correlation at approximately 97% for different TrichoScan operators \[11\].

A disadvantage is, that TrichoScan needs a hair dye for contrast enhancement and hairs must be clipped. However, this is also the case with PTG and/or CE-PTG. A clear advantage is that it is investigator independent (97% interclass correlation), offers a good response-to-change and is highly reproducible over time during a clinical study \[13, 15\]. It can be used for studying AGA or other forms of diffuse hair loss, and it can be adopted to study the effect of drugs or laser treatment on hypertrichosis or hirsutism. In conclusion, the TrichoScan is an operator- and patient-friendly, validated and reliable hair growth evaluation method and can be used for clinical studies to compare placebo vs. treatment or to compare the relative potencies of different hair growth-promoting substances.
21.15 Global Photographs

Most of the presented methods to assess hair growth or loss are providing rather mathematical values of hair count, hair shaft diameter or ratios of different hair types (e.g., anagen, telogen), but are not direct parameters of global hair appearance. Since the latter is of utmost interest for the patients affected with hair loss and therefore an essential aim of cosmetic/medical studies, standardized GP of scalp hair has been established to assess this hair quality parameter [4].

This method is suitable for home-in-use tests as well as for highly scientific studies to assess hair growth promoting substances. The technique requires a tripod with a carrier arm for a camera which can be swinged in a bow over the hair bearing scalp regions offering standardized fixed positions at 0, 45, and 90° (Fig. 21.3). Such a system has been first established by Canfield Scientific in the USA (Canfield Scientific, Inc., Fairfield, USA) and already been used in several clinical studies, e.g., the study to assess the efficacy of finasteride being the most prominent one [16]. The tripod arm ensures that the distance and the angles to the scalp regions are always the same and a twinflash guarantees standard light conditions which are constant during the time-course of a study. The patient’s neck and shoulders are covered with a black textile to ensure standardized low light reflectance from the surrounding. The evaluation of the photographs is performed independently from the image taking time point by an expert panel of two or three experts which are blinded to the treatment. The examination of pictures with regard to optical hair density from different investigation time points of the study (beginning, after 3 or 6 months, at the end) is rated using a seven-point scale from greatly decreased −3 to greatly increased +3 [4]. These scales have been historically introduced by experts from the USA. A study has been performed to test for interobserver variability between expert panels from the USA and European Union (EU). After having been trained with a training set of GP which were evaluated with a seven-point scale, the comparison between three US and three EU investigators showed a positive correlation of $r=0.795$ in 52 paired images from 35 different subjects. An additional test-retest evaluation of 18 paired images by the EU experts revealed a correlation of $r=0.806$ and identical scores in 78% of cases. This corresponds also well with data from US experts in 119 subjects which had a retest correlation of $r=0.76$ with 75% identical duplicate ratings. This demonstrates the high reproducibility of the method and together with the above mentioned aspects of clinical hair appearance as the final most important outcome of preventing or treating hair loss, the use of this method in clinical trials is not only very justified, but also recommended.

21.16 Optical Coherence Tomography of Hair

Optical coherence tomography of hair (OCT) is originally used to assess skin thickness, edema and other properties of human skin in vivo, but can be also excellently used for the assessment of hair shaft diameter, including the inner hair variation of diameter and
Hair Growth

shape in vivo as well as ex vivo [17]. The pictures of the OCT are comparable with pulse-echo images created by ultrasound, i.e., optical scatter is used to create two-dimensional images of inner tissue microstructures. It can measure not only hair shaft diameter, but also cross section surface and hair shape. This technique is rather a highly scientific method to be used in clinical studies with certain aspects of hair and might be mostly used for studies investigating effects of applied substances to improve hair quality in means of directly or indirectly assess hair shaft surface, consistency, elasticity, shape and ultrastructure.

Fig. 21.3  Global photographs in three positions (0°–45°–90°) to assess the global hair appearance in the fronto-temporal, the upper-head and the vertex area
21.17
Electron Microscopy of Hair

Electron microscopy offers images of tissues in highest resolutions and can be used also for hair. With this technique, the ultra-structure of the hair cuticle surface and optical profile can be assessed which serves to assess hair shaft abnormalities and to obtain longitudinal and/or transversal images of inner hair structure (e.g., medulla, changes/abnormalities). Electron microscopy of hair (EMH) is a technique which does not belong to the standard spectrum of hair assessment methods, since the microscope itself is very expensive and the assessments are time-consuming, therefore high-throughput investigations are not possible. It is therefore only used for highly scientific research purposes, but also for last proof examinations of hair in special clinical cases of genotrichoses, such as pili trianguli et canaliculi [10].

21.18
Confocal Laser Scanning Microscopy of Hair

Confocal LSM is a noninvasive imaging technique providing a three-dimensional image of the whole hair surface and the inner structures (cortex, medulla). Additionally, the emission spectrum of the hair keratin’s autofluorescence or of exogenously added fluorochromes to assess e.g., penetration or distribution of fluorescence-labeled cosmetic or pharmacological substances can be assessed [10]. It has the advantage that the assessment can be performed with the hair in vivo and in situ. Specific structures within the hair can be observed by adding different fluorescent markers specific for the target structures of interest, and following the routes of penetration using fluorescence-labeled substances allow to perform dynamic studies of external hair treatment. This method can be used to test efficiency of cleansing shampoos, to assess the homogeneity of layering polymers, and to evaluate the changes they induce in the optical properties of the hair surface in terms of opacity, transparency, and brilliancy. The technique does not belong to the standard hair study techniques, but has a more interesting and qualitatively different application profile than OCTH and EMH.

21.19
Conclusions

From all presented methods, one has to choose a meaningful combination of several methods to address certain clinical or research questions to be answered in accordance with aspects of feasibility, invasiveness, measurement accuracy and reproducibility, spectrum of parameters to be investigated, sophistication, costs and time-consumption for the patient as well as for the investigator. This chapter is designed to help understanding and using
Hair Growth

criteria for the different aspects while planning a hair study. To facilitate acquiring information about the technical methods, Table 21.1 provides relevant information.

While clinical scoring and HPT are standard methods to screen the study population for hair loss stage and activity with regard to inclusion and exclusion criteria of a hair investigation study, daily hair count, HWT and HW are already partially subjective, partially objective (semi-) quantitative assessment parameters which can be still used as screening parameters, but are already suitable to assess hair growth or loss as target parameters during a study.

HW seems to be a reliable and standardized method, in which the hair is clipped in a defined target area. However, the sample error for different investigators is unknown. This is mainly due to the methodology, because once the hairs are clipped, a second investigator cannot clip the same area again to assess the reproducibility of the method.

The standard TG has long time been the only and exclusive quantitative method to assess hair loss by means of assessing rate of telogen and anagen hair. Nowadays, more refined methods have been developed and are available. One development has been the UATG which allows to additionally assess hair count, hair shaft length and diameter. Both methods, however, have a relatively high variation, longitudinal studies bear the problem that TG assessments cannot be performed exactly at the same place at different time points and the method is invasive therefore tendentially uncomfortable for the patient. Regarding the UATG, a substantial difference between the collected data from different investigators can be observed. In several studies, a significantly larger mean total hair count was reported.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Company (name, address)</th>
<th>Internet-reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trichogram</td>
<td>Galderma Laboratorium GmbH, Georg-Glock-Strasse 8, 40474 Düsseldorf, Germany</td>
<td><a href="http://www.galderma.de">www.galderma.de</a></td>
</tr>
<tr>
<td>Epiluminescence microscopy of hair/TrichoScan</td>
<td>Tricholog GmbH, In den Eschmatten 24, 79117 Freiburg i. Brs., Germany</td>
<td><a href="http://www.tricholog.de">www.tricholog.de</a></td>
</tr>
<tr>
<td></td>
<td>TeachScreen Software GmbH, Aichner-Schmied-Strasse 3, 84364 Bad Birnbach, Germany</td>
<td><a href="http://www.fotofinder.de">www.fotofinder.de</a></td>
</tr>
<tr>
<td></td>
<td>Moser, Wahl GmbH, Roggenbachweg 9, 78089 Unterkirnach, Germany</td>
<td><a href="http://www.moser-online.com">www.moser-online.com</a></td>
</tr>
<tr>
<td>Global photographs</td>
<td>Canfield Scientific, Inc., 253 Passaic Avenue, Fairfield, NJ 07004, USA</td>
<td><a href="http://www.canfieldsci.com">www.canfieldsci.com</a></td>
</tr>
</tbody>
</table>
from experienced vs. inexperienced observers [26].

The PTG/CE-PTG and the ELMH have been further developments of the standard TG with computerized analysis available in the latter method. Both methods have been controversially discussed regarding their accuracy, validity, reproducibility and feasibility [7, 15, 31].

The combination of optical microscopy and computer analysis has been already developed quite early [1, 9], however, an automated process of calculation was not established, and measurements were carried out by assessing the thickness of hairs visually with a cursor on a computer monitor. This method showed a very high variation of 88.4% when images were assessed by different investigators. In contrast, the software-based automated system of the TrichoScan shows a high coefficient of intra-observer and inter-observer correlation of 91 and 97%, respectively [11]. However, these correlations had been shown for the first established digital video dermoscope camera system yielding an assessment area of 0.62 cm² and not, to the best of my knowledge, with the new system using digital cameras such as Nikon Coolpix or Canon Powershot with however a much bigger analyzing area of 1.5 cm² [14].

The conventional PTG method with manual marking of hairs on images is considered a suitable and noninvasive tool to monitor the hair growth phases in situ and sometimes defined as the most precise method of measurement. This technique has been improved by the image analysis [30] and later with the use of immersion oil and digital contrast enhancement [29]. However, although a marked improvement of the images and more accurate quantitative data were collected, a fully automated analysis is to my knowledge not yet possible as this technique still relies on data processing by qualified technicians and computer-assisted image analysis. Nevertheless, the refined form of the classical PTG, the CE-PTG [28], is being continuously optimized, which makes this method a valuable tool in hair quality assessment for clinical studies.

To compare the characteristics of both the PTG and the TrichoScan, several aspects have to be taken into consideration. A study compared CE-PTG and ELMH by TrichoScan and revealed that hair counts were higher in ELMH compared to CE-PTG, whereas hair density in turn was lower with the ELMH compared to CE-PTG, which is somehow contradictory [31]. The anagen hair percentage was lower in ELMH than in CE-PTG except the vertex area of the scalp where it was vice versa. However, most important is the intra-method-variability and within this parameter, CE-PTG had been compared to hair numbers in transverse sections of scalp skin punch biopsies. There, the hair count was considerably higher in the CE-PTG than in the histology transverse sections. Recently, a comparative study between TrichoScan and conventional PTG was performed in ten patients with AGA to evaluate variability, validity and reliability of both methods [7]. The study showed an excellent correlation of parameters evaluated by TrichoScan and manual marking of hairs of 0.894–0.996. A certain variability was noted in the results from repeated measurements assessed manually by PTG (range 2.71–12.95%) compared to TrichoScan (variability 0%). However, this comparative evaluation was based on the results of one investigator [7]. An advantage of TrichoScan was that the results were obtained more quickly and were more reproducible with a smaller margin of operator error compared to manual PTG. It is reported and also our own experience that experienced hands can perform a TrichoScan imaging and analysis with ready results within 8–12 min “hands on” [14]. The consistency in the
TrichoScan data allows statistically significant results to be obtained with a smaller study population size. As a critical point to TrichoScan, it was mentioned that besides edge effects hair fibers escaped the TrichoScan analysis for various reasons including, but not limited to, thickness, pigmentation, closeness and crossing [31]. However, there is an algorithm in the software which counts all hairs that touch the target border of the measurement field by mathematical approximation with 50% since by average, 50% of the hairs touching the border enter the measurement field (with their roots outside) and 50% go out of the area with their roots inside the area [15]. While TrichoScan is criticized to have errors in signal detection, this error, if true, is a constant error, since the variability of repeated measurements is low with 0% [7] or up to 9% [11] and therein also lower compared to 2.71–12.95% in normal PTG [7]. The CE-PTG on the other side has the advantage that catagen hair count/rate as well as linear hair growth rate (mm/d) can be assessed. To conclude, both versions of PTGs are comparable, have a high correlation and should be chosen according to the special needs of an investigator and purpose of a study design.

To conclude, scientists and clinicians have developed and examined methods of measuring scalp hair growth for decades. With the development of drugs that stop or even reverse the miniaturization of AGA, there has been a greater need for reliable, economical and minimally invasive methods of measuring hair growth and, more specifically, response to therapy. Here, the various methods of hair measurements described to date have been reviewed and discussed regarding their limitations and value to the clinician and hair researcher to give valuable information about the correct application and choice of methods for a successful performance of a clinical hair trial.

References


22.1 Introduction

Sensitive skin is a concept with immediate appeal. Drawing on personal experience, most people will acknowledge that the sensory perception of skin irritation (used in the broadest non-dermatological sense of the word) occurs more strongly in some than in others. Usually the patient’s description is that of “allergy” or “hypersensitivity”. The phenomenon varies both between persons and over time, suggesting that a complex set of factors is involved.

The definition of sensitive skin is ephemeral. The lay origins of the concept may well lie at the root of this problem. The concept of sensitive skin is based on some lay persons’ perception of being more intolerant to topical preparations and environmental conditions than the general population [11]. It is easy to imagine the lack of uniformity in this, ranging from persons with established skin diseases, from the depressed who somatise their disease, to persons whose skin genuinely appear to be hyper-reactive in a non-specific manner. Generally, persons who report sensitive skin appear to have a subjective hyper-reactivity to a wide range of non-specific stimuli of their skin, including not only cosmetics, but also physical influences such a temperature changes and air humidity. Nevertheless, sensitive skin constitutes a valid topic for further studies; in as far as it reflects a biological phenomenon that influences the behaviour and life of those it affects.

22.2 Sensitive Skin

No unified definition of sensitive skin exists, but a number of key characteristics have been identified [7, 11]:

G.B.E. Jemec
Department of Dermatology, Health Sciences Faculty, Roskilde Hospital,
University of Copenhagen, Roskilde, Denmark
e-mail: gbj@regionsjaelland.dk
1. Stinging/burning/itching
2. Sensation of tightness of the skin
3. Variability of symptoms in severity and over time
4. None or minimal clinical signs of inflammation
5. Close temporal correlation with exposure to the suspected causative agent

These characteristics imply that sensory testing and exclusion of dermatological disease are core elements in the assessment of sensitive skin. The characteristics imply that sensory testing in a controlled setting are core elements of any investigation.

Sensitive skin appears to occur more often in the face than elsewhere, perhaps because of the more frequent application of cosmetics to the face or due to local anatomical characteristics of facial skin [13]. Epidermal permeability, pore size and neural supply have all been suggested as important parameters in predicting sensitive areas of the skin.

Valid epidemiological studies are naturally difficult to conduct without a more precise definition of the phenomenon studied. A number of studies have been conducted in an attempt to enumerate and better define the magnitude of the phenomenon. The prevalence of sensitive skin is regularly described as being around 40%, with a clear female preponderance. One study found a prevalence of 51% in women and 38% in men [6, 14]. Such a sex-discrepancy is also seen in a number of other health-related self-reported situations, leading some authors to suggest the concept of dermatological non-disease, implying that other factors play a role [7, 9]. Women generally appear to have a higher frequency of health services usage in most countries and to have more symptoms of skin diseases and to be more precise in their reporting of skin complaints. These sex differences may have both biological as well as psychosocial origins, and may therefore be taken to reflect the broad implications of this biologically vague phenomenon. Accordingly, it is increasingly unlikely that a simple, monocausal explanation of sensitive skin will be found, and most likely, there are several pathophysiological paths leading to the same result, complicating systematic investigation and research.

This however also indicates that sensitive skin is a factor affecting not only the well being of a substantial proportion of the population, but may also affect patterns of behaviour, thereby warranting more specific investigation and academic attention in fields not traditionally explored by dermatologists.

### 22.3 Underlying Mechanisms

No single unified concept explains sensitive skin adequately. Subtle abnormalities in the neural supply of the facial skin, in keratinocyte structure, inflammatory mediators and skin barrier function have been proposed.
22.4 Testing Sensory Perception

In agreement with the suspected mechanisms of skin sensitivity, different types of tests have been suggested for the study of sensitive skin. Skin function tests assessing different structural and physiological parameters of the skin non-invasively are often used. Non-invasive biophysical measurements such as transepidermal water loss (TEWL), colourimetry etc. may provide independent information about the functionality of the skin barrier as well as more accurate assessment of skin reactions. These methods are dealt with extensively in other chapters of this book. Guidelines are published for the use of several methods, see Chap. 4.6 for details. Three general types of test can be applied to measure and predict sensitive skin:

1. **Sensory tests** that assess the subjective irritation
   A very large group of sensitive skin patients react only with subjective complaints and no signs of inflammation. A simple test was therefore proposed by Kligman et al., based on the observation that the application of lactic acid to the nasolabial fold causes stinging in some but not in others [8]. The nasolabial fold has been found to be the most sensitive region for this test, followed by the malar region, chin, forehead and upper lip [7, 13]. The lactic acid sting test (LAST) involves the application of lactic acid 5% to the nasolabial fold and facial skin of the test person. The test person should be in a warm humid room, and must be sweating. The stinging is classically scored on a scale from 0 (none) to 3 (severe) after 10 s, 2.5 and 5 min [5]. Lactic Acid-induced stinging is however not predictive of stinging caused by other substances such as capsaicin, menthol and ethanol. In a study, only 4 of 25 panellists were found to react to all four substances [12]. The general interpretation of these tests is that skin sensitivity is inversely proportional to application time and concentration. No sex difference has been shown in reactivity, and although the tactile sensitivity of the skin appears reduced with age, the capacity for pain sensation is not [3], indicating that panellists can have a wide age-range. No reliable data are available for children.

   The test is simple, inexpensive and convenient; it tests the area of the skin most often involved in sensitive reactions, and it does not require the use of any advanced methods. The drawback of these tests is mainly their subjectivity, which requires either very large groups or very well-defined, pre-tested groups of persons with a defined reaction pattern.

2. **Neurological tests** assessing the neural supply of the skin are not traditionally used in the exploration of sensitive skin. They are however occasionally necessary to rule out neurological disease as an underlying factor to sensitive skin. Two types of tests are used: standard clinical tests describing sensitivity to temperature changes, pain, tactile discrimination etc., and more specific tests for peripheral nerve function. These may be electrophysiological or even histopathological studies of nerve appearance, diameter and density in the skin [10].
3. Reactivity tests assessing cutaneous irritation, and generally associated with visible signs of inflammation or stratum corneum disruption. The most widely used test is the sodium lauryl sulphate (SLS) test. SLS damages the stratum corneum chemically inducing barrier damage and inflammation in a predictable manner. The test is applied as a patch test either on the ventral forearm or the upper back for 24 h. A SLS solution of between 0.1 and 2.5% is applied in a closed Finn chamber, and the ensuing reaction can be graded either clinically on a simple four point scale or using TEWL measurements for more continuous data [2]. Other irritants such as retinoids have been used in a similar manner, but the use of SLS is the most widely used substance.

SLS produces an eczema-like reaction, but studies have also been made using other skin reactions as functional markers of skin reactivity. Dimethyl sulfoxide (DMSO) produces an urticarial reaction after diffusion through the stratum corneum. It has therefore also been used to study skin reactivity. Different concentrations of DMSO are placed in wells on the skin for 5 min, removed and the ensuing wheals scored after 10 min. The response is assessed by either noting the lowest concentration of DMSO to produce a response or by quantifying the response using bioengineering techniques [1].

These tests provide an assessment of the stratum corneum integrity, and since a thinner and weaker skin barrier has been associated with increased sensory perception to non-specific stimuli, these tests provide a surrogate measure of the combined epidermal integrity and hence are used by some as measures of sensitive skin.

22.5 Setting Up a Test

At least two approaches can be made to the problem of testing if a given product or substance causes reactivity in persons with sensitive skin. Testing can be made either in very large groups or in very well-defined groups.

The large group approach involves an open application of the test substance to a larger number of healthy volunteers. Some of these volunteers will have sensitive skin, so they need to be classified according to self-reported skin sensitivity and the results interpreted in view of this parameter. The test substance is applied to the nasolabial fold, and any sensory reaction occurring within the first 8 min is recorded and scored on an anchored four point scale ranging from none [0] to severe [3]. The test persons should be allowed to wash off the test substance after 8 min.

Since large numbers of healthy test persons are readily available, this is a method based less on the measurable physical effect of any substance, than on the integrated psychological-biological reactivity of the test persons. In as far as this mimics the real life situation closely; it involves a number of logistic challenges in order to provide useful data. It may however be the most appropriate method for screening the potential of a given substance to cause skin reactivity in the general population [4].
Because sensitive skin is not uniformly present for a screening procedure, another approach is to identify a suitable group of test persons from a relevant population where more specific testing can be done. Ideally, the screening procedure should make use of a combination of tests. In order to identify a suitable background population, test persons should initially be recruited based on self-reported skin reactivity in accordance with the initial definition of sensitive skin, i.e. “Do you have sensitive skin?”

Following broad recruitment, a selection of suitable candidates should be made, in which manifest allergies and skin diseases, depression, infections and neurological disorders should be ruled out. This may require supplementary testing and examination by a dermatologist. This group can then be subjected to the screening tests described above, and the final test-panel selected based on objectively reproducible reactivity. The test persons are classified into three groups: “stinger” “inconsistent stingers” and “non-stingers”. A panel of 24 (eight from each category) are needed providing an even sex ratio in each group [4].

The actual testing is then done in a contralateral comparison design, by applying the positive control (lactic acid 10%) and test substance (contralaterally) to the nasolabial fold with cotton wool buds. The test material is put on the right nasolabial fold in half of the patients and in the left in the other half in a random manner. The test material is left in situ for 8 min during which the test persons are interviewed regarding their experience of stinging on either side, and the reaction graded from 0 (no stinging) to 3 (severe stinging). After 8 min, the face is washed, and any objective changes recorded. The test persons are re-examined after 24 h to establish if any longer-term objective changes have occurred. Using a similar technique, comparisons can also be made to similar products, providing comparative data not against a positive control, but against competing or similar products to establish their relative risk profile for stinging.

This method also involves considerable challenges before the test group is identified, but ensures that subsequent testing can be made very effectively using the more sensitive biological parameters which allow quantification of the response. Therefore, in tests of products especially aimed at patients with sensitive skin this method appears more appropriate.

22.6 Interpretation

Because of the variation in sensitive skin, test results must be interpreted strictly in agreement with the process by which they have been obtained. Exclusion of other diseases is mandatory, and only well-defined test substances with a known low-risk toxicological profile should be used. If objective reactions such as dryness/eczema occur, additional examination and testing must be done to rule out either precipitation/flare of a hitherto unrecognised skin disease or development of allergic contact dermatitis.
References

Transepidermal water loss (TEWL) is universally recognized to be a measure of skin barrier function, either at baseline, after experimentally induced barrier abrogation or following topical treatments. In mammals, it is also known as “insensible water loss” as it is a process over which organisms have little physiological control. Measurements of TEWL (grams per square meter per hour) is useful for identifying skin damage caused by certain chemicals, physical insult (such as “tape stripping”), or pathological conditions such as eczema [11], as rates of TEWL increase in proportion to the level of damage even before the damage is clinically visible. It may thus be considered as the tool that evaluates the water barrier function of the epidermis.

23.1 The Bricks and the Mortar Model

An intact stratum corneum (SC) is crucial to maintain a barrier that prevents the loss of fluids, electrolytes, and other molecules from within the body and prevents penetration of xenobiotics (microorganisms, allergens, toxic materials etc.) [3, 15]. The epidermis constantly replenishes itself by a process of homeostasis. Keratinocytes detach from the basal layer to generate transitional keratinocytes that follow the route of programmed differentiation with differential expression of keratins, the formation of keratohyalin granules, and lamellar body (LB) assembly. The last step in physiologic cell death and SC formation occurs during terminal differentiation, which requires LB dumping, nuclear defragmentation, and cornified envelope (CE) formation [4]. At leading edge of the epidermis, the SC generates an efficient barrier that protects the internal milieu of the organism from the external desiccating environment [20]. The SC is currently viewed as a layer of protein-enriched...
corneocytes embedded in a lipid-enriched intercellular matrix, the so called “bricks-and-mortar model” [6]. The anucleated corneocytes (“bricks”) contain keratin filaments bound to a peripheral CE composed of cross-linked proteins (i.e., loricrin, involucrin, filaggrin). The intercellular spaces are filled with lipid lamellae (“mortar”) built from a mixture of ceramides, free sterols, and free fatty acids [7] made by the secretion of LB at the level of the SC/SG junction [8]. The signaling process of LB secretion and terminal differentiation is tightly regulated and orchestrated through a multitude of signaling mechanism allowing a constant cross-talk between the SC and the underlying SG [4, 12, 18].

23.2
Instrumental Assessment of Epidermal Barrier Status and Repair

23.2.1
Industrial vs. Legislation Needs to Measure Barrier Status and Repair

EU cosmetic legislation requires providing a technical file for claims made for cosmetic products. The proof of efficacy has become an integral part of the technical dossier and can be subject for discussion and reclamation between concerned parties: consumers, industry, and legislator. Private initiatives, resulting in the foundation of EEMCO (European Group for Efficacy Measurements on Cosmetics and Other Topical Products) or Colipa (European Cosmetic, Toiletry and Perfumery Association) regrouping experts have established guidance documents concerning the efficacy of claim measurement [14, 19].

23.2.2
How to Assess the Positive Impact of Body Lotions and Protective Creams on Permeability Barrier Function?

The hydrating effect of a product could be easily accessed by the simple measurement in changes of SC electrical properties (capacitance) under both short- and long-term protocols. When claims regarding the barrier function need to be substantiated, quantification of the TEWL prior and after the application of the tested formulation is applicable. For cosmetic testing, measurements are performed in a group of adult healthy human volunteers with no history of allergies or inflammatory skin disorders. Depending on the targeted effect of the emollient, the gender and the age of the tested population should be chosen as homogeneous as possible. Prior to testing, the volunteers are advised to use gentle soaps and avoid exposure to UVs or irritant products in order to ensure an ideal condition for skin testing by avoiding external factors that might influence result reproducibility.

Additional guidelines need to be taken into consideration for adequate measurement of TEWL because external factors can significantly impact the measurements. Ideally, an ambient room temperature of 19–23°C and a relative humidity (RH) of 40–60% are advised. Preferentially, measurements are performed by the same operator on a horizontal surface, and a low and constant contact pressure between the probe and the skin surface is maintained [17].
The most important issue to consider when TEWL measurements are carried out includes air convection, room temperature, and ambient humidity. When measurements are carried out on forearm skin, a plastic box or draught shield, covered by a cotton cloth may be useful to reduce air convection. The volunteers are put in a relaxed state for at least 15–30 min under climate-controlled room conditions with a RH of 45–55% and a room temperature of 20–22°C. Direct light sources and sun light irradiation must be avoided in order to keep away from local increase in temperature as well as unwanted sweating at the testing sites. Finally and just prior to the assessment, the temperature of the measuring probe should equal the cutaneous temperature of the tested area (Table 23.1 summarizes the steps and recommendations for TEWL measurements).

TEWL measurements immediately after the application of the product are absolutely inadequate. Such protocols will only allow the evaluation of the occlusive properties of the product and do not assess the improvements/impairments in barrier function resulting from the effects of the active ingredients within the applied emollient. Long-term testing (at least 12 h after the first application) will determine the real positive or negative effect that the formulation may have on SC properties.

Generally, the application of a tested product on normal skin is considered as well formulated when its application significantly increases SC hydration and decreases TEWL. However, the claim “improves skin barrier” could only be used when a tested formulation is able to improve the recovery of damaged SC in comparison to physiological barrier repair. Epidermal damage could be performed chemically or physically by the application of an anionic surfactant such as sodium lauryl sulfate (SLS) or cellophane tape stripping of the SC, respectively. The damage caused in the permeability barrier due to SC insult correlates with a significant increase in TEWL. If the repeated application of the tested formulation is able to improve TEWL beyond the physiologic recovery (both untreated and vehicle treated-areas), the claim “improves skin barrier” may be substantiated. The calculation of the barrier recovery is generally expressed as percent of decrease in TEWL following SC abrogation.

### Table 23.1 TEWL measuring steps and recommendations

<table>
<thead>
<tr>
<th>Acclimatize</th>
</tr>
</thead>
<tbody>
<tr>
<td>The temperature (20–22°C) and humidity (45–55% RH) of the experimental environment</td>
</tr>
<tr>
<td>The tested volunteer and the measuring instrument to the climate of the experimental space</td>
</tr>
<tr>
<td>15–30 min prior to TEWL assessment</td>
</tr>
</tbody>
</table>

| Body sites areas are chosen symmetrically to avoid body sites-related differences |

| Carry out measurements with careful handling of the TEWL probe |
| Warm-up the probe to body site temperature of the subject |
| Apply constant light pressure of the probe to the skin |
| Maintain a horizontal contact of the probe with the skin |
| Measure TEWL of the tested site for 30 s or till the TEWL is stabilized (e.g. plateau levels of TEWL are reached) |
| Take a 1-min break between two measurements |
| Measure immediately prior to product application and at least 3 h (>6 h is recommended) after application |
| It is recommended to perform multiple testing (e.g. capacitance, pH) to back-up TEWL data |
Time–zero-TEWL (TEWL₀) is measured immediately after causing the required damage to the epidermis and just prior to the application of the tested product(s). Recovery measurements are carried out at the desired time points (TEWL₁, TEWL₂, TEWL₃, etc....) and formulated as follows: \%barrier recovery = 100 × (TEWL₁₋ₙ − TEWL₀)/TEWL₀

23.2.3
Finding the Right Formula!

The concept of “shielding the skin” has long been outdated. Emollients are now formulated to approach the reestablishment of the proper lipid components of the SC. Modern preparations include the disturbed and missing lipid substances able to restore the physiological bilayer structures. Products may also contain active ingredients that will directly impact epidermal differentiation and/or LB secretion by directly influencing signaling effectors within the viable epidermis at the SG level, for example. The pH of the formulation should be given due importance. Since SC pH is acidic at the surface of the SC approaching neutrality at the SG/SC interface, pH changes can directly influence barrier status [10]. Neutral pH is able to block bilayer formation by inhibiting the appropriate functioning of lipid processing enzymes or enhance desquamation by activating SC serine proteases [13].

23.2.4
How Valid Is TEWL to Measure Skin Barrier?

Permeability barrier function is currently assessed using TEWL instruments either with closed- or open-loop systems (see below). Attempts to question the validity of TEWL as an acceptable method to measure barrier status has been made. Using ex vivo models, no correlation was found between basal TEWL rates and the permeability of either human or pig epidermal membranes [2]. Yet, the validity of TEWL as a tool that reflects barrier status has been more than once validated in experimental designs. Most recently, Fluhr et al found that both open- and closed-loop systems under different experimental in vivo conditions correlated with absolute water loss rates, determined gravimetrically [9]. Measurements with both closed and open systems were also found to correlate not only with each other, but also each method detected different degrees of barrier dysfunction. Moreover, the authors found that all tested instruments are reliable tools for the assessment of experimentally induced variations in permeability barrier.

23.2.5
Few Examples of Currently Used Instruments for Measuring Water Loss

In vivo TEWL can be measured by two different water sampling techniques.

The open chamber method utilizes a skin probe open to the atmosphere. TEWL is calculated from the slope provided by two hygrosensors precisely oriented in the chamber. The diffusion principle in an open chamber can be calculated:

\[ \frac{dm}{dt} = D \times A \times \frac{dp}{dx} \]
23 Practical Use and Significance of Transepidermal Water Loss Measurements

where: \( A = \text{surface (m}^2\) ), \( m = \text{water transported (in g)} \), \( t = \text{time (h)} \), \( D = \text{diffusion constant (=0.0877 g/m.h(mmHg))} \), \( p = \text{vapor pressure of the atmosphere (mmHg)} \), \( x = \text{distance from skin surface to point of measurement (m)} \).

The TEWA meter is manufactured by Courage and Khazaka Electronic, Köln, Germany. The water vapor pressure gradient is indirectly measured by two pairs of a combined thermometer and hygrosensor, present at two different heights inside a hollow cylinder (height 2 cm, diameter 1 cm). The probe head is placed horizontally on the skin at a constant pressure and its small size minimizes the influence of air turbulences inside the probe. The density gradient is measured indirectly by the two pairs of sensors (temperature and RH) inside the hollow cylinder and is analyzed by a microprocessor. The small size of the probe head minimizes the influence of air turbulences inside the probe. Also, the low weight of the probe has no influence on the skin surface structure and allows easy handling.

The closed chamber is formed after touching the skin and the RH inside the capsule is measured with an electronic hygrosensor. Two types of devices are commercially available; the VapoMeter and the Aquaflux. The VapoMeter, produced by Delfin Technologies Ltd (Kuopio, Finland), is a portable and battery-operated device containing a Honeywell humidity sensor HIH 3605-B and uses the unventilated-chamber method of measurement. The cylinder-shaped chamber is equipped with sensors for RH and temperature (T). Water vapor from the skin surface collects in the chamber and causes the humidity to rise with time, slowly at first but linearly thereafter. The flux density is calculated from the slope of the linearly rising part of the curve. After the measurement is complete, the chamber needs to be lifted from the flux source to allow the accumulated water vapor to escape. The AquaFlux (Biox, London, UK) uses the condenser-chamber method of measurement. One end of the chamber is closed by a condenser maintained at a controlled temperature several degrees below the freezing point of water.

23.3 Limits and Pitfalls of TEWL Measurement

“What are meters measuring?” [5]. TEWL data are frequently used to support the claims of products from the cosmetic industry. The claim, “improves skin barrier,” used by the cosmetic production is based upon TEWL measurement experiments performed on humans subjects. The improvement in skin barrier is defined as the improvement of TEWL following the application of the ingredients on the skin of healthy volunteer subjects. Yet, TEWL reflects the passage of water through the lipid mortar since corneocytes are water proof (see Fig. 23.1). A decrease in TEWL in this case will depend on either the content and/or the structure of the lipid bilayers between the bricks. However, improved TEWL could also be linked to hypersecretion of LB at the SC/SG interface, a feature that may precede the development of epidermal hyperplasia. Two examples from our experimental work sustain this hypothesis [12, 18]. Both protease activated receptor (PAR)-2 [12] and caveolin-1 knockout mice [18] show improvement in recovery of TEWL following barrier disruption which correlates with hypersecretion of LB at the SG/SC interface with absolutely no change in the bilayer structure. Yet, absence of PAR-2 and caveolin-1 in mice sensitizes
the epidermis to epidermal hyperplasia and even skin cancer [1, 16, 18]. If TEWL reflects the lipid status, it completely underestimates the status of corneocytes, the “silent” feature of the brick and mortar. Formation of corneocytes largely depends upon the correct terminal differentiation of keratinocytes from the upper SG layer. Even though lipid secretion and corneocyte formation are tightly linked, deficient formation of new corneocytes, coupled to the phenomenon of LB hypersecretion, could clearly reflect an abnormal structure and function of the barrier, while measurement of TEWL remains favorable.

Finally, improvement of TEWL should be carefully interpreted. Ideally, complementary evaluations of SC hydration, pH, and both immuno-histochemical and electron microscopy studies should be performed to sustain the claim “skin barrier improving skin care product.”

References


24

Introduction: Classical Laboratory Test

Efficacy tests of cosmetic products are carried out worldwide according to specific, recognized scientific procedures and criteria. For a typical long-term test regarding the increase of stratum corneum hydration after application of a specific product, volunteers are invited to visit the testing laboratory. Before starting the tests, volunteers have to acclimatize to the ambient conditions for a certain time period in order to have their skin conditioned to the relative humidity and room temperature of the measurement environment. Subsequently (normally after 30 min) the stratum corneum hydration will be measured on the skin area to which the product will be applied and on an untreated control site, comparing treated and untreated skin. This is called baseline or initial value at time point T0 [2].

Each volunteer will be instructed to apply the test products to the defined skin area according to a detailed protocol throughout the long-term test. Normally, such protocols contain information about the skin area where the product has to be applied, the amount of the product, frequency and time of application as well as other important factors which might affect the efficacy of the product (e.g., skin cleansing, food and liquid intake, sleeping habits, sun exposure etc.).

G. Khazaka
Courage + Khazaka electronic GmbH, Mathias-Brüggen-Str. 91, 50829 Köln, Germany
e-mail: gkhazaka@courage-khazaka.de
The volunteers will be invited to the testing laboratory at regular time intervals during the long-term test (e.g., weekly); latest at the end of the trial. There, again, they will acclimatize to the lab conditions and stratum corneum hydration will be measured. The results T0, T1 to Tx are statistically evaluated and a data regarding the efficacy of the product will be generated. From the scientific point of view, this procedure is considered important and indispensable as it controls the influencing factors during the measurements as rigorously as possible.

Considering the test design, certain noncontrollable factors have to be accepted allowing only limited interpretation of the measurement findings. It is of an immense importance that the volunteer follows exactly the product’s application protocol during the test course. In general, the time span between the visits and the single measurements in the laboratory can be considered a “black box,” not accessible for control of the testing lab. The way, how, when and if the volunteer is following the instructions cannot be controlled by the laboratory. Few studies have been published on the compliance in application of sunscreens and protective creams [1, 6, 7].

The product performance during the course of the day under the normal real life conditions of the volunteer; e.g., at home, at work, during sports, are not taken into consideration or investigated in clinical trials. This information, however, is quite important as the product is meant to be effective under normal conditions and daily life situations of the consumer or patient.

In addition, volunteers taking part in clinical tests normally live in the geographic region where the test laboratory is located. This factor limits testing, especially for products developed and intended for the international market. As the product will be offered in different climate zones, tests might be performed simultaneously in several countries and even on different continents. Only very few manufacturers and testing laboratories are able to carry out multicentre tests on a worldwide basis due to the extremely high costs.

Thus it is suggestive to search for new ways of collecting additional information about the efficacy of cosmetic and dermatological products. These new approaches are not intended to replace the current standard tests but they should be considered as complementary to established testing methods and trial design.

### 24.2 Efficacy Tests in the Real Environment of the Consumer (Field Studies)

Efficacy tests in the laboratory as described above are standard procedures that are recognized worldwide. Guidelines have been developed for many methods (see Chap. 9). Efficacy tests of products during the course of a typical day of a volunteer, considering different geographical regions have not yet been reported.

Besides the high cost, one of the most important reasons is the lack of measuring devices allowing the consumer to carry out tests by himself in her/his normal environment. Today, technical devices are a part of everyday living; not only at work, but also at home (cell phones, computer etc.). They are most of the time easy to handle, small, secure and stable so that consumers can easily learn to use them without any specific training.

Due to this fact, manufacturers of skin testing devices are able to produce reasonably priced measuring instruments that, following a short introduction and instruction, can be
handled by the volunteers of clinical trials. The use of such measuring devices will provide additional interesting information about the properties of the skin surface after application of products and during their normal use (e.g., work, at home, at sports).

Another area where field studies can be performed are families with specific (and most of the time rare) skin diseases. Rare skin diseases with an altered barrier function and/or stratum corneum hydration are often seen in families. Researchers might study these families in their home region.

24.3
The Device Corneometer® Mobile Data Collector DC 3000

The Corneometer® is the most commonly used device for measuring stratum corneum hydration worldwide. It is used in cosmetic testing and dermatological trials since more than 25 years. Studies with this instrument are published in all major dermatological journals. Due to its capacitance measuring principle, the measurement is not only quick and easy, but also only marginally influenced by residues on the skin, e.g., salt or products. Mainly the water content of the stratum corneum layer is taken into account; the stratum corneum being the skin layer of interest for cosmetic efficacy testing [3–5].

A small Corneometer® has been developed especially for field studies of stratum corneum hydration. The volunteer number can easily be programmed into this device. The device is lightweight; battery operated and is delivered in a small protective case. Thus, the device can be dispatched easily and operated anywhere, which might contribute to the geographic expansion of recruiting volunteers. The device can easily be sent all over the world together with the testing products and an operation manual for subject training (Figs. 24.1 and 24.2).

![Application of the Corneometer® Mobile Data Collector DC 3000 at the working place of a volunteer on the cheek. Note that an ambient measurement device (temperature, relative humidity) is also attached to the Mobile Data Collector](image)
The Corneometer® Mobile Data Collector DC 3000 is easy to be operated by the volunteers. With each measurement the assessed value, date and time of the specific measurement and the ambient conditions (temperature, relative humidity) using the connected sensor are stored in the device. According to the test protocol, the measurements are carried out several times per day (e.g., at different time points; before and after application of the product etc.). Up to 3,000 measurements can be collected and stored during a long-term study.

After finishing the study, the device is sent back to the laboratory (or brought back for the final visit) where all measuring values are downloaded to a central PC and statistically evaluated. The data can also be sent by the volunteer to the laboratory during the study via internet by connecting the device to a computer using USB. The saved data cannot be manipulated by either the volunteer or the laboratory.

24.4 Conclusion/Perspective

Due to the newly developed technology, it is now feasible and economical to gain a lot of additional information concerning the daily application of products alongside established laboratory studies. The compliance of individuals and the entire test populations can be monitored. The Corneometer® Mobile Data Collector DC 3000 is supplying cosmetic manufacturers with new convincing data for their product claims.

The radius for recruiting volunteers can be enlarged up to a world-wide scale. The devices should of course only be handed over/sent to reliable volunteers who will return them after the end of the test.

It is also relevant to mention, that such field studies with frequent regular measurements exert a certain control of the volunteers. They will be more careful and disciplined in acting according to the protocol and in applying the product knowing that the skin measurement values are recorded on a regular base by the volunteer. They might feel as though they are being supervised.

In the future, other skin measurement parameters could be included in this concept (e.g., melanin-, erythema-, sebum measurements and others).
References

A

Acne lesions assessment. See also Impure, acne-like skin
  - facial sites, 161–162
  - instrumental assessment, 164–165
  - lesion types, 161
  - nonfacial sites, 162
  - quality of life assessment, 165
sensorial assessment
  - disease-specific instrument, 163
  - investigator (physician) assessment, 162–164
  - user-self assessment, 162
Anatomical skin sites, sweat glands, 218

Antaging and antiwrinkle products
non-invasive biophysical methods
  - SC hydration, electrical methods, 147
  - skin mechanical properties, 149
  - skin microcirculation, 148
  - skin pigmentation and color, 148
  - skin surface acidity (pH), 148
  - skin surface lipid, photometric assessment, 148–149
  - skin surface topography, 146–147
  - TEWL values, epidermal barrier function, 147
study design
  - anatomical test site, 151
  - data management and evaluation, 152
  - efficacy assessment, 152
  - product application, 151–152
  - study duration, 151
  - study population, 150–151
  - study schedule, 149–150
Antiperspirants
  - axillary sweat, gravimetric evaluation, 220–221
  - conditioning period, 221
  - data evaluation, 222
randomization, 222
  - study population, 221
  - sweat collection, 221–222
biophysical methods, 220
  - casting methods, 220
  - gravimetric methods, 220
  - multiple product test design, 223–224
visual techniques, 219–220

B
Barrier permeability assessment. See Transepidermal water loss (TEWL)

Bricks model, 253–254, 257–258

C
Carotenoids, 180
Case report form (CRF), 8, 30–31, 106
Central drugs standard control organization (CDSCO), 8
Claim support, 43–44
  - borderline claims, 45
challenge, 59
claim creation
  - communication style and medium, 47–48
  - consumer factors, 48
  - factors involved, 45–47
  - global/regional, 47
classification, CTPA/ASA approach, 49–50
definition, 44
documentation, 58–59
level of evidence
  - established and widely available evidence, 50, 52
  - science/technology, 53
  - specific evidence, 52–53
quality of evidence
consumer perception assessment, 57
consumer studies, 54
expert grading, 54, 58
instrumental evaluation, 54, 58
sensory properties, 54
in vitro/ex vivo studies, 58
Combination claims, 50
Common Rule, 17
Comparison claims, 50
Conditioners. See also Hair; Shampoos
cationic surfactants, formulation, 204
chemical characteristics, 204
deposition variants, 204–205
product rheology, 205
Confocal laser scanning microscopy (CLSM), 240
Contrast-enhanced phototrichogram (CE-PTG), 235, 242–243
Corneometer® Mobile Data Collector DC 3000, 263–264
Cosmetic testing report
good clinical research practice (GCRP), 107–108
result publication, 112–113
abstract, 113–114
acknowledgments, 115
background information, 114
discussion and conclusions, 115
keywords, 114
manuscript submission, 115–117
material and methods, 114
references, 115
tables and figures, 115
title page, 113
structure and contents
abbreviations and definition, 109
appendices, 112
discussion and conclusions, 111–112
ethical considerations, 109
introduction, 110
investigational plan, 110–111
investigators and study administrative structure, 109–110
reference list, 112
results, 111
signatures, 112
synopsis, 109
table of contents, 109
test objectives, 110
title page, 108–109
Council for International Organizations of Medical Sciences (CIOMS) guidelines, 17–18
Cutaneous aging, 144. See also Antiaging and antiwrinkle products

D
Daily hair count, 230–231
Declaration of Helsinki, 17
Deodorants
Sniff test
conditioning period, 225
data evaluation, 225
procedure, 225
study population, 224
test methods, 224
Dermal irritation patch tests, 8
DermaSpectrometer, 182
Dermometry, 33–34, 39–40
evidence-based guidelines, 38–39
good biometrological practice, 36, 38
skin bioengineering, 34–35
standardization and quality controls, 35
validation, methods and instrumentations, 35–37

E
Emollients, 118. See also Moisturizers
Epidermal barrier status and repair, TEWL
industrial vs. legislation needs, 254
lotions and creams, barrier permeability
barrier recovery, 255
emollient formulation, 256
external factors, 254–255
TEWL measuring steps and recommendations, 255
Time–zero-TEWL (TEWL₀), 256
Ethical aspects, 15–16
cosmetic testing, 19
history
Belmont Report, 17
CIOMS guidelines, 17–18
Common Rule, 17
Declaration of Helsinki, 17
Good Clinical Practice (GCP), 18
medical ethics, 16
Nuremberg code, 16–17
research ethics, 16
noninvasive skin measurements, 19
principles
conduct, 20
Institutional Review Board/Independent Ethics Committee (IRB/IEC), 22
investigational products, 21
investigator, 20–21
subjects/participants, 21

techniques/devices, maesurement, 22

Extrinsic aging, 144–145

F

Fitzpatrick classification ccale, 181

G

Glogau scoring, 146

Good Clinical Practice (GCP), 18

case record form (CRF), 30–31

conduct standards, 30

principles, 27–29

protocol standards, 30

standard operating procedures (SOPs), 29

study design, 30

H

Haemoglobin, 180

Hair, 200. See also Conditioners; Shampoos
color, 199

growth, 198

hair type, 199–200

influence of water, 200

internal structure, 199

medulla, 199

outer structure, 198–199

shampoos and conditioners, 214

color fade, 212–213

consumer evaluation approaches, 205–206

hair shine, 211–212

lubrication, 208, 210

moisturization, 213

static electricity retardation, 211

strength, 210–211

technical evaluation approaches, 206–208

technical functionality and consumer opinion, 205

surface damage assessment, SEM, 208, 209

Hair growth assessment

clinical appearance, 229–230

clinical grading systems, 230

confocal laser scanning microscopy, 240

contrast-enhanced phototrichogram, 235
daily hair count, 230–231
electron microscopy, 240

epiluminescence microscopy (see TrichoScan)
global photographs, 238, 239, 241

hair pull-test, 231

hair wash test (HWT), 231

hair weighing, 232

optical coherence tomography (OCT), 238–239

phototrichogram, 234

physiological factors, 228

study design, 229

trichogram, 232–233

Unit Area Trichogram, 234

Hair morphology assessment

atomic force microscope (AFM), 172
cuticle characterization, 171
dynamic experiments, 173–177,
optical interferometry (see also Optical interferometry, hair morphology)

roughness modification, 173

vertical scanning interferometry mode, 173

white light source, 173

scanning electron microscope (SEM), 172

static experiments, 173–175

Hair pull-test, 231

Hair wash test (HWT), 231, 241

Human repeat insult patch test, 9

Hyperkeratosis, 188

I

Impure, acne-like skin

acne lesions (see Acne lesions assessment)
oily skin (see Oily skin assessment)

product acceptance, 165–166

study background, 156

study design, 158

study objectives, 156

study products, 156–157

study subjects, 157

unwanted effect assessment

ethical principles, 167

statistical analysis, 167

study protocol, 166

Ingredient claims, 49

Institutional Review Board/Independent Ethics Committee (IRB/IEC), 22

Instrumental assessment

acne lesions assessment, 164–165

oily skin assessment, 160

Intrinsic aging, 144–145

In-use irritation tests, 125

Investigator global assessment (IGA) score, 164

J

Japan Cosmetic Industry Association (JCIA), 7
Laboratory testing
accommodation and arrangement
data base/subject file room, 67
general volunteer information sessions, 68
investigator’s office/examination room, 68
lab room, 68–69
reception/volunteer recruitment office, 67
rest rooms and waiting area, volunteers, 69
room types, 66
staff office, 69
test product/sample storage room, 69
ambient conditions, 70
quality management system
measurement devices, 71–72
standard operating procedures (SOPs), 70–71
Leeds Acne Grading Scale, 163

Measurement devices and methods
cost issues, 87
laboratory, 86–87
manufacturers and software packages, 87
practical use, 86
quality and quick service, 87
technical assessment
skin parameters, 84–85
standards and directives, 85–86
Melanin, 180
Mild acne, 163
Ministry of Health, Labour, and Welfare (MHLW), 7
Minolta Spectrophometer, 175
Moderate acne, 163
Moisturizers. See also Skin hydration characteristics, 124
investigative method selection, 131
legislation and marketplace, 125–126
non-invasive biophysical methods
mechanical/visco-elastic properties, skin, 130
outermost skin layers, electrical properties, 128
skin surface topography examination, 129
spectroscopic and optical properties, skin, 130
techniques, 127
transepidermal water loss assessment, 128–129
study designs
multiple application tests, 132
preliminary skin irritation test, 132–133
regression tests, 132
single application tests, 131–132
study population
general considerations, 133–134
number of subjects, 134
subject and environment-related variables, 134, 135
test site, 134–135
study procedure
data management and evaluation, 139–140
instructions, panelist, 136
measurement procedure, 139
product application, 138–139
washout phase, 136
Mortar model, 253–254, 257–258

Noninvasive skin measurements, 19
Nuremberg code, 16–17

Oily skin assessment
image analysis techniques, 161
instrumental assessment, 160
lipid absorbent tapes, 160–161
photometric techniques, 160
quality of life assessment, 161
sensorial assessment, 159
skin glossymetry, 161
Optical coherence tomography (OCT), 238–239
Optical interferometry, hair morphology
dynamic experiment
cuticle evolution, tensile stress, 176, 177
fibres used, 173
roughness modification, 173
static experiment
moisturizing, topographical modifications, 176
roughness reduction, 174, 176
sampling, 173
scratch evolution, 174–176
topographical representations, cuticle, 174–175
vertical scanning interferometry mode, 173
white light source, 173

Performance claims, 49
Permeability assessment. See Transepidermal water loss (TEWL)
Photoaged skin. See also Antiaging and antiwrinkle products
clinical, histological, and functional characteristics, 144–145
Glogau scoring, 146
Phototrichogram (PTG), 234, 242
Physical Global Improvement score, 164

Product testing
classical laboratory test
location, 262
noncontrollable factors, 262
product performance, 262
protocol information, 261
stratum corneum hydration, 263
Corneometer® Mobile Data Collector DC 3000, 263–264
field studies, 262–263

Regulatory aspects
ASEAN countries, 7
Canada, 7
China, 8
cosmetic legislation
broad definitions, 4
Canadian system, 5
China, 5
Europe, 4–6
India, 5
Japanese system, 5
narrow definitions, 4
universal similarities, market economies, 4
cosmetics
efficacy testing, 11–12
labeling and packaging, 12–13
safety study, 8–11
European Union, 6
India, 8
Japan, 7
Mercosur countries, 7
South Africa, 8

USA, 6–7
Research staff
investigator, 75–76
study nurse, 77

Sensorial assessment, 158–159
acne lesions assessment
disease-specific instrument, 163
investigator (physician) assessment, 162–164
user-self assessment, 162
oily skin assessment, 159
Sensory/aesthetic claims, 49
Sensory perception
sensitive skin
contralateral comparison design, 251
eye characteristics, 247–248
large group approach, 250
mechanisms, 248
neurological tests, 249
prevalence, 248
reactivity tests, 250
screening tests, 251
sensory tests, 249
sex-discrepancy, 248
very well-defined groups, 251
skin irritation, 247
Severe acne, 163
Shampoos. See also Hair
anionic surfactants, 201
conditioning and styling products, 201
cosurfactants, 201
2-in-1 shampoo+conditioners
aesthetics factors, 203
formulation stability, 202–203
manner of usage, 203
silicone oil, 202
stability and formulation, 201–202
surfactant concentration, 201
viscosity, 201
Singapore Health Sciences Authority (HSA), 5
Skin colour and pigmentation
aged skin, 181
carotenoids, 180
constitutional skin colour, 181
factors influencing, 184
grading system, 181
haemoglobin, 180
melanin, 180
reflectance spectrophotometric evaluation
cromameter, 183
Skin hydration. **See also** Moisturizers

- complex formation, 125, 126
- environment-related variables
  - air flow, air-conditioning system, 138
  - air humidity, 138
  - biorhythms, 138
  - direct light exposure, 138
  - panelists acclimatization period, 138
  - room temperature, 136
- epidermal barrier
  - external stressors, 124
  - mechanisms, 124–125
  - stratum corneum (SC), 124, 125
  - water distribution, 125, 127

**Skin testing and measurements, 85**

- environment-related variables
  - air convections, 91
  - ambient air humidity, 92
  - ambient air temperature, 92
  - light sources, 92
  - skin cleansing, 92–93
- individual-related variables
  - age, 93
  - anatomical sites, 93–94
  - circadian rhythm, 95
  - intra and interindividual variation, 95
  - race and sex, 93
  - skin damage and diseases, 95
  - skin surface temperature, 94–95
  - sweating, 94
- instrument-related variables
  - contact time, 91
  - surface area, 91
  - variability, start-up, and use, 90

**Sniff test. See Deodorants**

- Sodium laurel sulphate (SLS) test, 250
- Standard operating procedures (SOPs), 29
- State food and drug administration (SFDA), 8

**Study design**

- blinding, 102
- data entry and processing, 106
- ethical review and scientific validity, 104
- inclusion/exclusion criteria, 105
- measurement variables and endpoints, 103
- randomized controlled study, 102
- statistics
  - sample size calculation, 103
- statistical methods, 104
- study documents
  - case report form (CRF), 106
  - study protocol, 105
  - subject diary, 106
  - subject information sheet, 105–106
- study objectives, 102
- test products, 104

**Sunscreens**

- hyperkeratosis, 188
- sun protection factor (SPF), 188–189
  - classification, 192
  - factors influencing, 189
  - international methods, 191
  - MED determination, 189
  - sunscreen application, 190–191
  - UVB protection, 188–189, 193
  - in vitro determination, 193
- UVA protection factor, 191–192

**Sweating, 218**

**T**

- Testing population recruitment, 79–80
  - acquisition criteria and screening, 81–82
  - advertisements, 80
  - product-related information and compliance, 82
  - telephone recruitment, 81
  - third parties/intermediaries, 81
- Transepidermal water loss (TEWL)
  - Bricks and Mortar model, 253–254
  - closed chamber method, 257
  - epidermal barrier status and repair (see Epidermal barrier status and repair, TEWL)
  - limits and pitfalls, 257–258
  - open chamber method, 256–257
  - validity, 256
- Trichogram, 232–233, 241
- TrichoScan
  - black hair dye, 236–237
  - digital image and algorithmic analysis, 236
  - disadvantage, 237
  - hair growth parameters, 235–236
  - landmarks, red tattoo ink, 236
  - software validation, 237

**U**

- Unit Area Trichogram (UATG), 234, 241