

Practical Aspects of Cosmetic Testing

How to Set up a Scientific
Study in Skin Physiology

Joachim W. Fluhr
Editor

Second Edition



Springer

Practical Aspects of Cosmetic Testing

Joachim W. Fluhr
Editor

Practical Aspects of Cosmetic Testing

How to Set up a Scientific Study in Skin
Physiology

Second Edition

 Springer

Editor

Joachim W. Fluhr
Charité – Universitätsmedizin, Berlin
Berlin, Germany

ISBN 978-3-030-44966-7 ISBN 978-3-030-44967-4 (eBook)
<https://doi.org/10.1007/978-3-030-44967-4>

© Springer-Verlag Berlin Heidelberg 2011

© Springer Nature Switzerland AG 2020

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface to the Second Edition

Skin physiology assessment is moving from a descriptive approach to a deeper understanding of biophysical functions and their underlying molecular mechanisms in the stratum corneum, e.g., on epidermal barrier function, as well as on stratum corneum hydration. The research with noninvasive instruments now offers reliable and reproducible approaches for basic research and product testing in the pharmaceutical and cosmetic industry. The focus of standard instruments is the functional aspect. Innovative imaging devices offer a deeper understanding of underlying mechanisms.

The present book provides basic information regarding assessment of skin physiology, skin functions in controlled studies using noninvasive biophysical instruments. The framework of correctly planning, performing, and evaluating scientific studies is described in this edition. Almost all chapters from the first edition are updated and two new chapters are added. At this point, I would like to thank again all authors for their careful work to make this book happen. I also would like to thank Dr. Juliette R. Kleemann and Mr. Karthik Periyasamy from Springer Nature Publishing. Without their patient help this new edition would not have been possible. I trust this book finds its way again to young and eager scientists at the beginning of their research career and helps them to answer some of the practical questions that all scientists should have when using new technologies. Furthermore, I hope that the practical work will lead to exciting results, new ideas, and interesting hypothesis.

Berlin, Germany
February 2020

Joachim W. Fluhr

Preface to the First Edition

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 “cornerstones” 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, Germany
October 2010

Joachim W. Fluhr

Contents

Part I Legal Aspects of Cosmetic Testing

- 1 Regulatory Aspects** 3
Oliver Wunderlich
- 2 Ethical Aspects of Cosmetic Testing** 15
Hristo Petrov Dobrev
- 3 Good Clinical Practice** 27
Betsy Hughes-Formella, Nicole Braun, and Ulrike Heinrich
- 4 Guidelines in Skin Testing** 33
Joachim W. Fluhr, G. E. Piérard, and C. Piérard-Franchimont
- 5 Claim Support: How to Create and Substantiate Claims** 43
Sinéad Hickey and Stephen 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** 85
Gabriel Khazaka
- 10 Factors Influencing Measurements** 91
Enzo Berardesca and Norma Cameli
- 11 Study Design** 103
Betsy Hughes-Formella, Nicole Braun, Ulrike Heinrich,
and Carmen Theek

12	Cosmetic Testing Report	111
	Hristo Dobrev	
Part III Practical Aspects of Testing: Typical Examples of Test Settings		
13	Moisturizers and Emollients	127
	Razvigor Darlenski and Joachim W. Fluhr	
14	Anti-aging and Anti-wrinkle Products	147
	Razvigor Darlenski, Theresa Callaghan, and Joachim W. Fluhr	
15	Products for Impure, Acne-like Skin	159
	Hristo Dobrev	
16	Assessment of Hair Morphology	177
	C. Thieulin, R. Vargiolu, and Hassan Zahouani	
17	Skin Color and Pigmentation	189
	Ammitzboell Elisabeth	
18	Characterization of Sunscreens: Determination of the SPF	197
	Juergen Lademann, Joachim W. Fluhr, and Martina C. Meinke	
19	Practical Aspects of Shampoo and Conditioner Testing	207
	Trefor Evans	
20	Antiperspirants and Deodorants	229
	Razvigor Darlenski and Joachim W. Fluhr	
21	Hair Growth	239
	Tobias W. Fischer	
22	Sensory Perception	259
	Gregor B. E. Jemec	
23	Novel Methods for In Vivo Skin Structure Visualization	265
	Giovanni Pellacani, Stefania Guida, and Silvana Ciardo	
24	Assessment of Pruritus and Sensitive Skin	289
	Flavien Huet and Laurent Misery	
25	Practical Use and Significance of Transepidermal Water Loss Measurements	297
	Truus Roelandt and Jean-Pierre Hachem	
26	Compliance Check with On-Site Measurements: Advanced Ways in Product Testing	305
	Gabriel Khazaka	
	Index	311

Part I
Legal Aspects of Cosmetic Testing

Chapter 1

Regulatory Aspects



Oliver Wunderlich

Core Messages

- Cosmetics legislation frameworks in the major markets are based on either broad or narrow definitions of cosmetics.
- Legal frameworks with broad definitions of cosmetics (e.g., in Europe) work with extensive restriction lists for ingredients.
- Legal frameworks with narrow definitions of cosmetics (e.g., in the USA) do not feature restriction lists for ingredients; however, based on the claims made or known therapeutic effects of certain ingredients, products may be classified as a drug.
- Cosmetic safety testing in humans may only be used to confirm a good safety of a test product as already indicated by other available information, not to disprove indications of bad safety properties.
- Even in a clinical testing setting, the responsibility for the safety of the test products ultimately remains at the manufacturer; however, the testing organization should always scrutinize all safety aspects, for both ethical and liability reasons.
- Cosmetic efficacy tests should be designed to be able to substantiate claims that are feasible for cosmetics in the specific regulatory framework.
- No specific regulations for labeling of cosmetic test products exist—general responsibility for the safety of the volunteers and practical requirements lead the way.

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

O. Wunderlich (✉)
bioskin GmbH, Hamburg, Germany
e-mail: oliver.wunderlich@bioskinCRO.com

another country. This has implications for the testing of cosmetics. Depending on which country the testing is 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.

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 acquirement 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 at the time. 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.

A comprehensive study comparing the cosmetics legislation in the EU and other principal markets was prepared for the European Commission Directorate General in 2004 [1]. More recently, a comparison of cosmetics legislations was published which focuses on the United States, European Union, Japan, and India and highlights product safety topics [2].

1.2 Recent Changes in European Cosmetic Regulation

On 11 July 2013, the Regulation (EC) No 1223/2009 on cosmetic products, the legal fundamentals for cosmetic products in the EU, came into force. It repealed the EU cosmetics Directive 76/768/EC, which originated in 1976 and had been updated and amended many times. This 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).

In contrast, the present regulation 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 recommendations by the Scientific Committee on Consumer Products (SCCP, now called Scientific Committee on Consumer Safety (SCCS)) on safety testing of cosmetics, which up to now were required only for new ingredients to be added to the positive 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.”

1.3 Important Weblinks

1.3.1 European Union

Cosmetics Regulation:

Original version: <http://data.europa.eu/eli/reg/2009/1223/oj>

Latest consolidated version: <http://data.europa.eu/eli/reg/2009/1223/2019-08-13>

SCCS (formerly SCCNFP and SCCP) central website: https://ec.europa.eu/health/scientific_committees/consumer_safety_en

SCCP archive: http://ec.europa.eu/health/ph_risk/committees/04_sccp/04_sccp_en.htm

Important SCCP/SCCNFP/SCCS guidelines/opinions: https://ec.europa.eu/health/sites/health/files/scientific_committees/consumer_safety/docs/sccs_o_224.pdf

http://ec.europa.eu/health/ph_risk/committees/sccp/documents/out101_en.pdf

Cosmetics Europe (European Cosmetics Association, formerly called Colipa) central website: <https://cosmeticseurope.eu/>

German Cosmetic, Toiletry, Perfumery and Detergent Association (IKW): <http://www.ikw.org>

General information for manufacturers of cosmetics: <https://www.ikw.org/ikw-english/beauty-care-topics/> German Federal Institute for Risk Assessment (BfR): https://www.bfr.bund.de/en/health_assessment_of_cosmetics-570.html

1.3.2 USA

Food and Drug Administration (FDA): <https://www.fda.gov/cosmetics/cosmetics-guidance-regulation>

The Personal Care Products Council (formerly the Cosmetic, Toiletry and Fragrance Association (CTFA)): <http://www.personalcarecouncil.org>

1.3.3 Canada

Health Canada, cosmetics regulation: https://laws.justice.gc.ca/eng/regulations/C.R.C.,_c._869/index.html

Cosmetic Ingredient Hotlist: <https://www.canada.ca/en/health-canada/services/consumer-product-safety/cosmetics/cosmetic-ingredient-hotlist-prohibited-restricted-ingredients/hotlist.html>

1.3.4 Japan

Ministry of Health, Labour, and Welfare (MHLW), standards for cosmetics: <https://www.mhlw.go.jp/english/dl/cosmetics.pdf>

No official translation of the Pharmaceutical Affairs Law (PAL) available, unofficial version: <http://www.japaneselawtranslation.go.jp/law/detail/?id=3213&vm=04&re=02>

Japan Cosmetic Industry Association (JCIA): <https://www.jcia.org/en/>

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/Cosmeticos_en.asp
Brazilian

National Health Surveillance Agency (ANVISA), search also for Resolução n° 79, de 28 de agosto de 2000: <http://portal.anvisa.gov.br/cosmetics>

1.3.6 ASEAN Countries (Example)

Singapore Health Sciences Authority (HSA):
<https://www.hsa.gov.sg/cosmetic-products>

1.3.7 South Africa

South African Government Information, “FOODSTUFFS, COSMETICS AND DISINFECTANTS ACT, 1972 (ACT NO. 54 OF 1972)”: https://www.gov.za/sites/default/files/gcis_document/201409/32975146.pdf

Cosmetics, Toiletry and Fragrance Association of South Africa (CTFA): <https://ctfa.co.za/>

1.3.8 China

https://www.chemsafetypro.com/Topics/Cosmetics/How_to_Comply_with_Cosmetics_Regulations_in_China.html

National Medical Products Administration (NMPA): <http://english.nmpa.gov.cn/cosmetics.html>

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): <https://cdsco.gov.in/opencms/opencms/en/Cosmetics/cosmetics/>

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 some other markets such animal data is still required prior to testing in humans. However, in recent years many markets followed Europe in the trend for banning animal testing for cosmetics (with a ban either already in force or in the process of phasing them out), so there are currently changes ongoing depending on the specific market and test required.

The most important property of a cosmetic is its safety, which must be checked prior to marketing a cosmetic product (EU Cosmetics Regulation EC No 1223/2009).

In the EU cosmetics regulation as well as in guidelines by the SCCP (now SCCS) and Colipa (now Cosmetics Europe), 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: http://ec.europa.eu/health/ph_risk/committees/sccp/docs/html/sccp_out87_en.htm

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/SCCS, which are updated occasionally.

Weblink:

https://ec.europa.eu/health/sites/health/files/scientific_committees/consumer_safety/docs/sccs_o_224.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, 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, semioclusively 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.

The SCCP issued an opinion [SCCNFP/0120/99] about sensitization tests. Weblink: http://ec.europa.eu/health/ph_risk/committees/sccp/docshtml/sccp_out102_en.htm

At the moment, there is no validated test method available which can replace animal use completely, only a refined animal test which minimizes the use of animals (the Local LymphNode Assay). However, intense research and development towards an animal-free test is ongoing [3], promoted by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) as well as the industry association Cosmetics Europe.

It may be reasoned that the best approach is choosing ingredients with known low sensitization potential in the first place, and avoiding those with a known high sensitization potential.

The European Society of Contact Dermatitis published a guideline for diagnostic patch testing [4], providing recommendations which are also helpful in a cosmetic study setting.

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

Weblinks:

https://ec.europa.eu/health/archive/ph_risk/committees/sccp/documents/out101_en.pdf

https://ec.europa.eu/health/sites/health/files/scientific_committees/consumer_safety/docs/sccs_o_224.pdf

https://ec.europa.eu/health/scientific_committees/consumer_safety/opinions/sccnfp_opinions_97_04/sccp_out45_en.htm

See also the “Guidelines for Assessment of Skin Tolerance of Potentially Irritant Cosmetic Ingredients” issued by COLIPA (now: Cosmetics Europe) in 1997.

Weblink:

https://cosmeticseurope.eu/files/2314/6407/8977/Guidelines_for_Assessment_of_Skin_Tolerance_of_Potentially_Irritant_Cosmetic_Ingredients_-_1997.pdf

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 [5].

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 Sect. 1.4). 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).

Already in 2008, the COLIPA (now: Cosmetics Europe) had 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.

Weblink:

https://cosmeticseurope.eu/files/4214/6407/6830/Guidelines_for_the_Evaluation_of_the_Efficacy_of_Cosmetic_Products_-_2008.pdf

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 regulation 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
- Declaration of presence of specific allergens

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) and last confirmed in 2016. This standard might be adopted internationally in the future.

Weblink:

<https://www.iso.org/obp/ui/#iso:std:iso:22715:ed-1:v1:en>

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,” and “store at room temperature in a safe place”. 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

1. 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. https://rpaltd.co.uk/uploads/report_files/j457-final-report-cosmetics.pdf
2. Suhag J, Dureja H. Cosmetic regulations: a comparative study. *Skinmed*. 2015;13(3):191–4.

3. Hoffmann S, et al. Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database. *Crit Rev Toxicol.* 2018;48(5):344–58.
4. Johansen JD, et al. European society of contact dermatitis guideline for diagnostic patch testing—recommendations on best practice. *Contact Dermatitis.* 2015;73(4):195–221.
5. McEwen GN, Murphy EG. Cosmetic claim substantiation in the United States: legal considerations. In: Elsner P, Merk HF, Maibach HI, editors. *Cosmetics: controlled efficacy studies and regulation.* Berlin: Springer; 1999.

Chapter 2

Ethical Aspects of Cosmetic Testing



Hristo Petrov Dobrev

2.1 Introduction

Ethical considerations are an essential part of any biomedical research involving human subjects [1–3].

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 involves predominantly healthy persons and is carried out to increase 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. Because of that, the ethical considerations connected to 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.

H. P. Dobrev (✉)

Department of Dermatology and Venereology, Medical University, Plovdiv, Bulgaria

The aim of this chapter is to outline the ethical aspects of cosmetic testing using noninvasive 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 [4].

The first *Code of medical ethics* was written by the American Medical Association (AMA) in 1846. It was based upon the guidelines of 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 [5].

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 [6–9].

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 [10]. 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 [11].

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](#) (IRB) in the USA, also known as independent ethics committees (IEC) or ethical review boards (ERB) 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 Good Clinical Practice used today [12–15].

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 [9, 16, 17].

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 [18]. 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 Institutional Review Board (IRB), and additional protections for certain vulnerable research subjects—pregnant women, prisoners, and children [9].

After 1982, the Declaration of 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 healthcare 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](#) [7, 19].

In 1996, the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) published its own *Guideline on Good Clinical Practice* (GCP) [16, 20]. 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 World Health Organisation (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 this same standard.

In 2001, the Council of Ministers of the European Union adopted a *Directive on clinical trials* (Directive 2001/20/EC) related to implementation of GCP in the conduct of clinical trials on medicinal products for human use within the European Union [21]. 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 [6, 22]. 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* [23]. The handbook is based on current major international guidelines and is organized as a reference and educational tool to facilitate 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 6th 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 [24, 25]. 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 "Good Clinical Practice." As a rule, safety testing on human volunteers should be preceded by animal or in vitro methods, whereas efficacy testing should be preceded by safety assessment of each cosmetic ingredient and finished product using in vivo methods (e.g., allergy and irritation skin tests) and should be performed if no evidence for local or systemic adverse responses exists. In addition, safety assessment should always be included as a secondary objective in cosmetic efficacy studies (monitoring for adverse reactions) [26–29].

2.4 Ethical Aspects of Noninvasive Skin Measurements

Due to the developments in bioengineering technology during recent years, it is possible now 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" [28]. 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 [2, 3, 28, 30].

2.5 Essential Ethical Requirements for Performing a Study

Emanuel et al. (2000) propose seven requirements which must be fulfilled if the study is to be ethical [16, 31]:

- A study must be socially valuable.
- A study must be conducted in a methodologically rigorous manner.
- The selection of subjects must be fair.
- The risk-benefit ratio must be favorable.
- There must be an independent review.
- Subjects must freely give their informed consent to participate.
- Subjects must be treated with respect throughout the study and after its completion.

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 [13–15] and current international guidelines for Good Clinical and Research Practices [19–21, 23, 25, 32] as well as by the guidelines for the evaluation of safety, compatibility, and efficacy of cosmetic products [26, 33], guidelines concerning medical devices [34], and national regulations regarding human studies.

The following principles must be taken into consideration:

Principles Related to Study Conduct

- The study should be ethically and scientifically justifiable.
- The study should be preceded by a risk-benefit evaluation, which take 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 related to product use/application and to protocol procedures. The choice of cosmetic study design and type should take into account the nature of the claim (e.g., ingredients claims, performance claims, sensory/aesthetic claims, combination claims, comparison claims) and the strength of the study design as related to consumer expectations [27]. 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.

Principles Related to Study Investigator

- The investigators and other research personnel must be qualified and competent to perform their roles [27].
- 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.

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 give freely their competent informed consent before the study begins. They must sign a written informed consent document that describes the nature of the study, the products to be tested, known or potential risks, the subject's rights, and who to contact in case of problems. In case of substantial changes in the study conditions or procedures, and also periodically in long-term studies, the investigator should once again seek informed consent from the subjects [27].
- The sample size of study subjects should be adequate. Studies with a limited and/or inadequate sample size may produce misleading conclusions, are time and money consuming and unethical [27].
- The study subjects should be selected by strict inclusion/exclusion criteria that should be clearly defined in the study protocol. The study outcomes is strictly related to the right identifying of the population in which it is relevant to assess the impact of the cosmetic product [27].

- 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.
- Study participants may be reimbursed for their transport expenses and may also receive appropriate compensation for their time and any inconveniences due to study participation which should be in details described in the IRB protocol at the time of initial review. Payments in money should not be so large as to influence subjects to agree to participate in the study for them [27].
- Investigators should ensure that research subjects who suffer non-expected/ adverse skin reaction or injury as a result of their participation are entitled to free medical treatment for such injury and to such financial or other assistance as would compensate them equitably for any resultant impairment. Before the research begins, the sponsor should agree to provide an adequate medical treatment or a compensation for any physical injury for which subjects are entitled to compensation [27].
- Investigators should protect the confidentiality of all the information about study participants. During the process of obtaining informed consent, the investigator should inform the prospective subjects about the precautions that will be taken to protect confidentiality [27].

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 [35]. They should be used only in accordance with the approved study protocol.

Principles Related to Measuring Techniques/Devices

A vast instrumental bioengineering techniques are widely used to assess skin characteristics. The devices and their measurement accuracy are constantly being updated and improved. Some researchers criticize these methods because they think the measured changes may be too small to be perceived by the consumer. That is why, in order to improve the relationship between the instrumental measurement and the claim and its relevance for the consumer, expert grading or consumer testing (e.g., self-assessment questionnaire) should be always co-applied [27, 34].

- The skin-measuring devices should be manufactured and handled according to the current regulations on medicinal devices [34]. They should be used only by or under the control of a suitable qualified professional in accordance with the approved study protocol.
- The methods to be used should be appropriate to the objectives of the research and the field of study [27].

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

- The IRB/IEC is an independent body constituted of medical, scientific, and non-scientific 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. Scientific review and ethical review cannot be separated: scientifically unsound research involving humans as subjects is ipso facto unethical in that it may expose them to risk or inconvenience to no purpose; even if there is no risk of injury, wasting of subjects' and researchers' time in unproductive activities represents loss of a valuable resource [27].
- The IRB/IEC should conduct further reviews as necessary in the course of the research and monitor the progress of the study.

2.6 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 for 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 the 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 remains always obligatory.

2.7 Conclusion

Ethical principles originating in the Declaration of Helsinki and the generally accepted scientifically principles of the Good Clinical Practice should be applied to all biomedical studies involving human subjects. Properly designed and well-conducted cosmetic study supported by skin-measuring techniques do 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 participants are protected and that the trial data are credible.

Key Messages for Performing an Ethical Study

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

References

1. Fakruddin M, Bin Mannan K, Chowdhury A, et al. Research involving human subjects—ethical perspective. *Bangladesh J Bioethics*. 2013;4(2):41–8.
2. Noerreslet M, Jemec GBE. Ethical considerations. In: Serup J, Jemec GBE, Grove G, editors. *Handbook of non-invasive methods and the skin*. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2006. p. 73–82.
3. Sohl P, Jemec GBE. Ethical considerations. In: Serup J, Jemec GBE, editors. *Handbook of non-invasive methods and the skin*. Boca Raton, FL: CRC Press; 1995. p. 667–70.
4. The Hippocratic Oath: Classical Version. http://www.pbs.org/wgbh/nova/doctors/oath_classical.html. Accessed 22 Dec 2019.
5. American Medical Association. Code of medical ethics. 1847. <http://ethics.iit.edu/ecodes/sites/default/files/American%20Medical%20Association%20Code%20of%20Medical%20Ethics%20%281847%29.pdf>. Accessed 22 Dec 2019.
6. Hope T, McMillan J. Challenge studies of human volunteers: ethical issues. *Med Ethics*. 2004;30:110–6.
7. Macrae DJ. The Council for International Organizations and Medical Sciences (CIOMS) guidelines on ethics of clinical trials. *Proc Am Thorac Soc*. 2007;4:176–9.
8. Nuremberg Doctors' Trial. The nuremberg code. 1947. *BMJ*. 1996;313:1448. <http://www.bmj.com/cgi/content/full/313/7070/1448>. Accessed 22 Dec 2019.
9. Rice TW. The historical, ethical, and legal background of human-subjects research. *Respir Care*. 2008;53(10):1325–9.
10. World Medical Association. Declaration of Geneva. 2017. http://dl.med.or.jp/dl-med/wma/geneva_e.pdf. Accessed 22 Dec 2019.
11. World Medical Association. International Code of Medical Ethics. 2006. <https://www.wma.net/wp-content/uploads/2006/09/International-Code-of-Medical-Ethics-2006.pdf>. Accessed 22 Dec 2019.
12. Carlson RV, Boyd KM, Webb DJ. The revision of the declaration of Helsinki: past, present and future. *Br J Clin Pharmacol*. 2004;57(6):695–713.
13. World Medical Association. Declaration of Helsinki. *Br Med J*. 1964;2:177.
14. World Medical Association. Declaration of Helsinki. 1975. <https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/doh-oct1975/>. Accessed 22 Dec 2019.
15. World Medical Association. Declaration of Helsinki. 2008. <https://www.wma.net/wp-content/uploads/2016/11/DoH-Oct2008.pdf>. Accessed 22 Dec 2019.
16. Kind C, Salathe M, editors. *Research with human subjects. A manual for practitioners*. 2nd ed. Switzerland: Swiss Clinical Trial Organization. Swiss Academy of Medical Sciences (SAMS); 2015.
17. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Belmont report: ethical principles and guidelines for the protection of human subjects of research. Washington, DC: Federal Register Document; 1979; April 18, 1978: 79-12065. https://globalhealthtrainingcentre.tghn.org/site_media/media/medialibrary/2011/04/Belmont_Report_1979.pdf. Accessed 22 Dec 2019.

18. Department of Health and Human Services. Code of federal regulation, title 45: public welfare, part 46: protection of human subjects (the common rule). 2005. <http://www.hhs.gov/ohrp/documents/OHRPRegulations.pdf>. Assessed 06 May 2009.
19. Council for International Organisations and Medical Sciences, World Health Organization. International ethical guidelines for biomedical research involving human subjects. Geneva, Switzerland: WHO; 2002. https://globalhealthtrainingcentre.tghn.org/site_media/media/medialibrary/2011/04/CIOMS_International_Ethical_Guidelines_for_Biomedical_Research_Involving_Human_Subjects.pdf. Assessed 22 Dec 2009.
20. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. ICH harmonised tripartite guideline: guideline for good clinical practice (E6R1). 2002. <http://www.emea.europa.eu/pdfs/human/ich/013595en.pdf>. Assessed 06 May 2009.
21. Council Directive 2001/20/EC on the approximation of the laws, regulations and administrative provisions of the member states relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. 2001. <http://www.eortc.be/Services/Doc/clinical-EU-directive-04-April-01.pdf>. Assessed 22 Dec 2019.
22. General Medical Council. Research: the role and responsibilities of doctors. 2002. <http://www.gmc-uk.org/guidance/current/library/research.asp>. Assessed 06 May 2009.
23. World Health Organisation. Handbook for Good Clinical Research Practice (GCP). 2002. http://www.who.int/medicines/areas/quality_safety/safety_efficacy/gcp1.pdf. Assessed 22 Dec 2019.
24. Council Directive 76/768/EEC on the approximation of the laws of the member states relating to cosmetic products (consolidated version 2008). <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1976L0768:20100301:en:PDF>. Accessed 22 Dec 2019.
25. Opinion concerning guidelines on the use of human volunteers in compatibility testing of finished cosmetic products—adopted by the Scientific Committee on cosmetics and non-food products intended for consumers during the plenary session of 23 June 1999. https://ec.europa.eu/health/scientific_committees/consumer_safety/opinions/sccnfp_opinions_97_04/sccp_out87_en.htm. Assessed 22 Dec 2019.
26. European Commission. The rules governing cosmetic products in the European Union, volume 3. Guidelines. Cosmetic products. Notes of guidance for testing of cosmetic ingredients for their safety evaluation. 1999. <https://op.europa.eu/en/publication-detail/-/publication/86ca0a09-4aa7-40e3-bc40-d8f846f613e0/language-en/format-PDF/source-search>. Assessed 22 Dec 2019.
27. Nobile V. Guidelines on cosmetic efficacy testing on humans. Ethical, technological, and regulatory requirements in main cosmetics markets. *J Cosmetol Trichol*. 2016;2(1):1000107.
28. Rogiers V, Balls M, Basketter D, et al. The potential use of non-invasive methods in the safety assessment of cosmetic products. *ATLA*. 1999;27:515–37.
29. Salter D. Non-invasive cosmetic efficacy testing in human volunteers: some general principles. *Skin Res Technol*. 1996;2(2):59–63.
30. Serup J. Efficacy testing of cosmetic products. *Skin Res Technol*. 2001;7(3):141–51.
31. Grady C, Wendler D, Emanuel EJ. What Makes Clinical Research Ethical?. *JAMA*. 2000;283(20):2701–11. <https://doi.org/10.1001/jama.283.20.2701>. PMID: 10819955.
32. Committee of Ministries/Council of Europe. Recommendation No R (90) 3 concerning Medical Research on Human Beings. 1990. <http://hrlibrary.umn.edu/instreet/coecrecr90-3.html>. Assessed 22 Dec 2019.
33. Colipa. Guidelines for the evaluation of the efficacy of cosmetic products. Brussels, Belgium. 2008. <http://www.canipeec.org.mx/woo/xtras/bibliotecavirtual/Colipa.pdf>. Assessed 22 Dec 2019.
34. Council Directive 93/42/EEC concerning medical devices (consolidated version 2007). <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993L0042:20071011:en:PDF>. Accessed 22 Dec 2019.
35. Council of Europe. Guidelines for good manufacturing practice of cosmetic products (GMPC). Strasbourg: Council of Europe Publishing; 1995.

Chapter 3

Good Clinical Practice



Betsy Hughes-Formella, Nicole Braun, and Ulrike Heinrich

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 Harmonisation 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:

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.*

B. Hughes-Formella
Strategic and Clinical Consulting, Dermatology Product Development, Schoenningstedter
Strasse 128b, Reinbek, Germany
e-mail: betsy.hughes@derm-consult.com

N. Braun (✉) · U. Heinrich
DermaTronnier, Institute for Experimental Dermatology at Witten/Herdecke University,
Witten, Germany
e-mail: nicole.braun@uni-wh.de; ulrike.heinrich@uni-wh.de

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 [2]. 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 continued only if the anticipated benefits justify the risks.*

As stated in the proposed recast of the European Cosmetic Directive [3], a risk-benefit reasoning for cosmetics should not justify a risk to human health. A cosmetic study should only be undertaken 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/favourable 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, e.g., testing of sensitization potential, or in designs with minimally invasive procedures such as abrasive wounds or suction blisters. An IEC should in general 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,

e.g., 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 with a cosmetic product.

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. This is also required in the data protection regulations.

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 labelled 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 as to 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, e.g., 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 include, among others:

- 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 blinding 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)
- Signed study report

References

1. Guideline for Good Clinical Practice E6 (R2) (ICH harmonized tripartite guideline), CMP/ICH/135/95, last updated December 2016.
2. 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 64th WMA general assembly, Fortaleza, Brasil, October 2013.
3. Recast of the European Cosmetic Directive. http://ec.europa.eu/enterprise/cosmetics/html/cosm_simpl_dir_en.htm

Chapter 4

Guidelines in Skin Testing



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

4.1 Introduction

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

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

J. W. Fluhr (✉)

Department of Dermatology and Allergology, Charité Universitätsmedizin, Berlin, Germany
e-mail: joachim.fluhr@charite.de

G. E. Piérard · C. Piérard-Franchimont

Laboratory of Skin Bioengineering and Imaging, Department of Dermatopathology,
University Hospital Sart Tilman, Liège, Belgium

4.2 Skin Bioengineering Endeavour

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

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

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

4.3 Validation of Methods and Instrumentations

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

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

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

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

Table 4.1 Instrumentation validation and definitions

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

Adapted from Ref. [6]

Table 4.2 Basic guidelines for bioengineering

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

Developed from Ref. [6]

4.4 Standardization and Quality Controls

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

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

4.5 Search for “Good Biometrological Practice”

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

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

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

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

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

4.6 Guidelines in Perspective

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

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

Table 4.3 Published guidelines for noninvasive measurements

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

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

4.7 Conclusion

Noninvasive measurements of skin physiology or 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 unexperienced beginners. Even skillful investigators are facing a number of problems related to relevance and interpretation of data. In every instances, emphasis should be placed on a strict respect of controlled and standardized conditions. At present, dermometry is still in a developmental field where the researcher must control every single aspect of the measurements. Measuring devices provide pure numbers which are only relevant with a specific research question and defined hypothesis. The skill of the researcher and conversely the naive

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

References

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

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

Chapter 5

Claim Support: How to Create and Substantiate Claims



Sinéad Hickey and Stephen Barton

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 Union, but 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 two decades for many reasons, but primarily as a regulatory requirement for cosmetic products, to protect consumers from unsafe or misleading products. Regulation (EC) N° 1223/2009 on cosmetic products of the European Union places a legal obligation on marketers of cosmetics to have supporting evidence available for any product claims [1]. This is supported by Regulation (EU) No 655/2013 which lays down common criteria for the justification of claims (EU Common Claims Criteria) [2]. These criteria particularly concern Truthfulness, Evidential Support, Honesty, and Fairness. The regulations covering substantiating claims are covered further elsewhere in this book (Chap. 2, Sect. 2.1).

S. Hickey
Formerly Health Products Regulatory Authority, Dublin, Ireland

S. Barton (✉)
Skin Thinking Ltd., Nottingham, UK
e-mail: stevebarton@skinthinking.com

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, since unfulfilled claims can lead to consumer skepticism not only of the culprit product, but also of other products sold within this segment. These factors together emphasize the need for claims to focus on real consumers' needs and, importantly, ensuring that these are fulfilled by marketed products.

The process of supporting claims has thus gained its importance and a number of initiatives addressing this will be 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 center of the process of creating a claim. Interpretation of technical terms, creating a rationale that withstands scientific scrutiny, also needs to communicate to a potential customer what they can expect from the product. This will be 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 on how to go about acquiring the information needed to support the claim. The industry and its regulators need to know what defines a sound, relevant, proof of a claim; what evidence and how extensive does this evidence need to be to support a cosmetic claim? 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 text, names, trademarks, pictures and figurative or other signs [1]. Examples of claim wording include such terms as “conditioning,” “breath-freshening,” “reduces the appearance of wrinkles,” “skin feels soft,” “moisturizes the skin for up to 24 h,” and “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 a brand’s internet site, on the advertisement (be it magazine, TV or on other promotional material) [3]. The increasing use of digital platforms such as YouTube, Instagram, and Facebook to advertise and target consumers directly has become a new area of attention from regulators. 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 will be explored further.

One of the biggest challenges faced by cosmetics companies 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”^[1].

The words we have emphasized here in bold imply that *secondary* actions may be permitted but such actions cannot imply medicinal or pharmacological action.

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” ...

The situation is similar in the USA where “cleansing,” “enhancing beauty,” “promoting attractiveness,” and “altering appearance” are permitted claims, but treating a disease or altering the biological structure (“increasing collagen,” “reviving cells”) cannot be made.

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*

The view of regulatory bodies on cosmetics creating the impression of being a medicinal claim will be influenced by many factors. Presentation and marketing, including packaging, where marketed, the target consumer, along with wording and images used can change the impression of any claim, especially with borderline claims. Therefore, the decision on whether a product will be defined as a medicine 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 fulfills 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 on claims. Definitions of medicinal claims 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 provided for products marketed in Europe¹. The document not only acknowledges

¹http://ec.europa.eu/enterprise/cosmetics/doc/guidance_doc_cosm-medicinal.pdf

the likely physiological effects of cosmetics but also clarifies that for medicinal regulations to apply, “modifying physiological functions by exerting a pharmacological, immunological, or metabolic action...”... “has to be more than insignificant.” If this statement seems difficult to interpret, perhaps it acts as a warning to adopt extreme care in attempting to make claims of this type. Indeed, advertising regulators in the UK anticipate the most intense and robust evidence in support of physiological claims.

Products that push the boundaries of cosmetic towards medicinal definitions have been referred to as “cosmeceutical.” This phrase, often attributed to the leading US dermatologist Al Kligman, was actually coined by US cosmetic scientist Raymond Reed [4] and refers to products that appear to be a medicine, a cosmetic, or a combination of both, and are generally sold under the remit of cosmetics [5]. In spite of its frequent and current use, there is no such thing 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 a disease?

It should also be noted that there are borderlines between cosmetics and other product categories—antimicrobial and medical devices are two such classifications that sometimes require thought and care before commencing product development.

While there are guidelines on these matters, often the safest course of action is to approach the appropriate regulator for an opinion.

5.2 The Process of Creating Cosmetic Claims

Some 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 summarized 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.

5.2.1 *Global or Regional*

Regulations may vary in different parts of the world concerning 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. Apart from the EU regulatory area already mentioned, the most common areas of cosmetic regulation that will need consideration before deciding on a marketing approach are Japan, United States, the ASEAN countries, and China [6–9]. Any company wishing to sell their cosmetic product in

Table 5.1 Steps in creating a cosmetic claim

Decision to be made	Key considerations	Comments
Is the claim cosmetic?	Cosmetic? Borderline? Medicine? Target Market/ Consumer Presentation Local regulations	“Cosmeceutical” claims do not exist, Pushing boundaries towards a medicine may influence the views of the regulatory bodies whether your claim is “cosmetic” Ensure the claim reflects cosmetic benefits—especially changing the appearance. Primary action should be cosmetic; secondary actions may be physiological and thus potentially medicinal If the mode of action (or claim) borders on medicinal take advice from local medicinal regulator. The potential for physiological action must be balanced against cosmetic benefit Presentation of product can infer medicinal benefits
Where will the product be sold?	What are the local regulations in each of the target markets and does this suggest the different approach to claims support? Geography/Market sector	Regulatory expectations may differ in different geographies Claims expectations differ between market and product sectors
How do you intend to communicate the claims?	Pack Point of sale Broadcast Print Direct Mail Website	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. What is the overall impression communicated in the different media? The environment (point of sale; type of publication); type of sounds/images used to deliver message. Decide early in the process—re-testing to support advertising claims can be expensive and high risk.
What is the intended use and benefits of the product?	Performance Ingredient Aesthetic/Sensory Combination Comparison Aspirational	How will the consumer know that the claims are fulfilled? Build a supporting rationale based on the likely consumer experience
What are the consumer expectations of the claim?	Is the claim currently marketed? Is the claim new for the product format and type?	Decide on the level of evidence anticipated Beware of implicit claims Different skin types/ consumer populations may have different expectations
What tests (if any) are required?	How does support fit consumer expectation?	Using more than one source of support can be very powerful and rationales built on this principle answer criticism of instrumental measures
What are the success criteria	What quality and quantity of data and information will satisfy your rationale?	See Table 5.3 for pros and cons of different types of testing

the USA, for example, 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” [10]. Product developers have to decide on the best strategy and there are three basic approaches (and probable 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 Chap. 2, Sect. 2.1.

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 pack label space available is important in agreeing on product claims with limitations brought about by factors such as small products (for example, mascaras) or multi-lingual globally marketed products. An often-used tool to substantiate claims is where the strong front of pack claims are qualified or further described on the back of the 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 the 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 fulfill 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 realize sale and potential repeat sales of a product. 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 and the probability of repeat sale increased.

To create motivating claims, the marketer will want to use particular attributes that differentiate its product in the market and appeal to potential consumers. Any number of 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 completely new claim since it will be hard to evaluate consumer expectation. Consumers will have a clearer understanding and expectation of claims that have been in the marketplace for a long period of time. This factor will also influence decisions later (see Sect. 5.3.2 ‘Level of Evidence’).

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. One 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 [11]. 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 the claim* (or consumer message) to be made?
2. How will the *supporting information* provide proof that the copy fulfills reasonable *consumer expectation* of that claim?
3. What are the requirements for the *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 to make claims consistent with their brands and products. The approach also avoids the checklist approach used by some regulatory bodies, where a specific test (or tests) is required for a specific claim. Checklists fail to recognize the fact that there is more than one way of defining, and many more ways of measuring, any given cosmetic attribute. For example, moisturized skin could be defined as plumped up 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 [12] and with the EU Common Claims Criteria, must avoid denigrating another brand. Comparison claims may refer to an older product “New Improved” or a benchmark “More effective than a market leading moisturizer.” 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.

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 will not be 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.

Table 5.2 Levels of evidence anticipated in supporting different types of cosmetic claim

Claims widely accepted to be established	Claims with established rationale but requiring product- or ingredient-specific evidence	Claims that are based upon a significant advance in science or technology
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	<p>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:</p> <ul style="list-style-type: none"> • Highly dependent on formulation factor(s) • For a product type or format not normally associated with the established claim • Dependent on factors described in guidelines, industry recommendations or scientific reviews outlining the need for specific test(s) • An enhancement or quantification of an established claim (for example—intensity of benefit, duration of benefit, a percentage claim) • Targeted to particular populations (for example, hair/skin type, age, ethnic origin). 	<p>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:</p> <ul style="list-style-type: none"> • The action of an entirely new type of ingredient • An entirely new consumer benefit • An entirely new sensory property • An entirely new means of qualifying or quantifying the product's effects on the substrate • A new insight into the biology of the body part to which the product is applied.

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.² 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 color 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

²From CTPA Guidelines Document [11]—reproduced with permission.

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).

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 moisturizer.” 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 before marketing the product.

Another key factor is that obvious effects, hence claims, are related to obvious formats—a wax lipstick containing pigments will color and condition lips; a liquid detergent based shampoo will clean hair. A powder-based lip color or an oil-based detergent-free shampoo would not possess the obvious effects simply by naming the product as lipstick or shampoo—in fact, the opposite is true (see below).

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 many different sources such as clinical studies or consumer evaluations. The reason for the extra support may be due to formulation differences, requirements from regulations or guidelines, enhancement of an established claim, etc.

An example of a claim requiring this category of evidence would be a product claiming SPF 15 sun protection. 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, to claim a measured sun protection factor (for example, SPF 15), the product must be tested in accordance with the current ISO guidelines [13].

There are many other examples, driven by similar considerations, where emphasizing or quantifying the duration or intensity of a benefit would require additional

evidence from tests. For example, a moisturizing emulsion containing a combination of well-established ingredients known to moisturize the skin, for example, glycerol, urea, lanolin, and shea butter, should require further evidence if claiming to offer “24 h moisturization.”

Another example of a claim requiring specific evidence would be “SPF 15 moisturizer 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 consumer perception and even measure the effect of the final product to assess the likely performance once marketed.

Yet another specific case of generating evidence arises when a claim is attributed to a particular ingredient. In many cases ingredient benefits are made in passing allowing the consumer to link—“contains soothing aloe vera” being a common example of this practice. The EU Common Claims Criteria has emphasized that adoption of this approach must ensure that the consumer experiences the benefit. In other cases, brands will state that a given benefit is directly attributable to one ingredient “contains aloe vera to soothe the skin.” The expectation from this type of claim is that evidence of the ingredient in the product formulation will be available—in this case soothing action of aloe vera. Such evidence will usually require studies of the product with and without aloe vera to demonstrate this direct incremental effect.

In summary, each marketer will have their 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 background of knowledge in the area of science impacting on the product formulation, ingredients, and tests used.

5.3.2.3 Evidence Based Upon a Significant Advance in Science or Technology

Communicating such a claim by reference to a revolutionary formulation, ingredient, or claim providing novel consumer benefit, or an existing benefit from an unexpected format 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 the wording of the claim. It may involve a description of several 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. 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 many 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 standardized controlled conditions what can be scientifically analyzed (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 ‘Borderline Claims’ 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 to understand 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 [14]. Trained panels of volunteers with high levels of sensory acuity can define the language and descriptors of key performance attributes of products [15]. Alternatively, naïve panels can provide useful spontaneous responses to product concepts. This approach is covered in more detail elsewhere in this book.

Table 5.3 Considerations when choosing sources of experimental data to support claims

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

(continued)

Table 5.3 (continued)

Method	PROS	CONS	Data analysis
In vivo “Clinical”/ expert assessment	Trained experts to grade visual, tactile, odor changes Preferred by many regulators For wrinkle severity tests, the different wrinkles on the face can be graded, e.g., periorbital wrinkles, nasolabial wrinkles Use of photographic scales for reference enhances validity Grading of “clinical” photos a powerful method of comparing expert and consumer perception and inclusion in claims support package further strengthens case	Expensive and time consuming Results can vary between centers Design criteria can result in being distant from real in-use conditions	Similar issues to Sensory and Consumer testing Advantage that subjective responses should be immune from bias, but large numbers of observations essential
Instrumental/ biochemical tests	Useful quantitative and qualitative data when used in vivo study Useful to assess poorly- perceived physiological changes Rapid and objective Validation and standardization possible Good for screening	Relationship to claim and consumer relevance sometimes difficult Scientific rationale sometimes questioned	Use of controls and validation of methods can help inform and improve statistical power Data manipulation/ normalization can be performed Comparisons to control helpful Usually provide continuously distributed data for standard statistical analysis
In vitro/ ex vivo tests	As for instrumental /biochemical tests	Criticized for being removed from reality	As for instrumental /biochemical tests

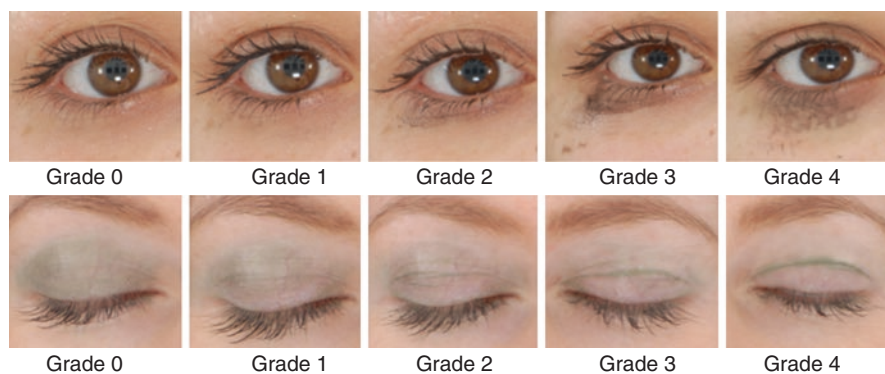
Consumer studies are primarily used to mimic likely consumer response for the product to be marketed. The information generated may be used to support product claims based on perceived benefit—such as “skin feels softer”; “wrinkles are less visible”; and “skin is more resilient and flexible.” One of the frequent challenges and potential flaws is that the desired claim has not been captured exactly within the question choice. It is also important that these studies use sufficient self-selecting subjects representative of the 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.”³ Grading can be performed on a variety of characteristics from dry skin or wrinkle severity [16–18]. Photographs of the area of

³This term is commonly applied to controlled studies conducted in quasi-clinical conditions or by clinicians.

Table 5.4 Sample questions that can be used to assess consumer perception

Dichotomous questions: a choice between yes and no									
E.g., Does this product cleanse the skin effectively?									
		Yes			No				
Multiple choice Questions: a choice between three or more answers									
E.g., Which of the following terms would you use to describe the action of the product you are testing?									
Refreshing		invigorating		Toning		Revitalizing			
Scale based questions:									
Q, Did this product remove all traces of make up?									
1	2	3	4	5	6	7	8	9	
Definitely		Probably		Unsure		Probably not		Definitely not	
Agreement:									
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	
Strongly Disagree		Disagree		Neither agree or disagree		Agree		Strongly Agree	

**Fig. 5.1** Examples of photographic grading scales (reproduced with permission from Oriflame Cosmetics). **(a)** Evaluation of the water resistance of mascara. **(b)** Evaluation of the water resistance of eye shadow

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 demonstrates the use of this method for assessing the water resistance of mascara or the crease resistance of eye shadow. The level of expertise must be consistent within any given study and training or other validation essential whether conducted by a dermatologist, ophthalmologist or nonclinical scientist conducting the study.

Instrumental evaluation to assess cosmetic benefit, *in vivo*, has continued to grow in importance in assessing skin characteristics such as moisturization, trans-epidermal water loss, firmness [19–21], as well as skin color and gloss effect both *in vivo* and *ex vivo* [22–24]. The technology of these instruments is constantly

being updated and improved, further improving the accuracy [25–27]. The area of skin imaging has exploded in recent years with 2-D and 3-D image analysis devices and software readily available to quantify features such as size from pores; eye bags; facial wrinkles; scalp hair; and cellulite [25–32]. These approaches provide quantitative and sensitive support for claims such as “reduces wrinkles by 30%,” “10× volume mascara,” or “24 h moisturization.”

This type of approach has been criticized on the basis that changes measured can be too small to be perceptible by the consumer. One response to this is to perform parallel studies in consumers to confirm the relevance of the claims.

In vitro/ex vivo studies are rarely used alone but, backed up with evidence using an additional method *in vivo*, may help 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, the *in vitro* method can be used to provide final formulation support [33–36]. *Ex vivo* studies on final products have also been suggested on hair tresses for claims on styling products or determination of hair damage [37–39].

The methods briefly described in this section will be dealt with in more detail in later chapters. However, concerning 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.

In summary, any given claim will be supportable from one or more sources of evidence—formulation information, ingredient information, specific test results, and published books and scientific papers. Ultimately, in the case of EU regulations, a “Responsible Person” will decide what evidence is sufficient and what quality and quantity of evidence is to be used in supporting the claim. This evidence requires marshalling into a convincing body of evidence. One important aspect of the updated Common Claims Criteria in EU markets is that ALL evidence will be required; selecting one study that provides the answer you want to hear is not acceptable as this is neither honest, truthful, nor fair.

5.5 Presenting Information in Support of a Claim

Other chapters in this book deal with the organization and presentation of study data and reports. However, by following the process outlined in this chapter, a marketer will have defined what the product is, where and how it is to be marketed, its uses and benefits, what type of consumer, what expectations they may have. Documentation of this “top-line” rationale is an important step in the presentation of the evidence, but as will be obvious from the previous section, any information will require assembling into some kind of dossier. 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 pleasing 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 a Product Information File (PIF) for the product in question. This file 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

Some 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 [40]. 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 subgroups. 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” which perform complex functions. The ongoing debate of what constitutes a “natural” claim⁴ has already given rise to a number of “standards” and there will undoubtedly be continued activity in this area. In the last few years, there has been a huge increased interest in and use of “Free-from” claims. Though this is not the place to go into detail on these contentious and ever-growing claims, the European Commission provided further guidance to the EU Common Claims Criteria in 2017 stating clearly that the only justification for such claims they envisage is to aid

⁴The claim “Natural” could be defined as an aesthetic claim under the scheme described in Sect. 5.3.1 ‘Classifying Claims’ though how “natural” itself is defined, together with its validity, is a point of considerable discussion within the cosmetics industry.

consumer choice, for example, stating a product is “free from alcohol” to assist consumers with religious or cultural reasons to avoid alcohol is acceptable [41]. However, claiming the product to be free from permitted and safe ingredients—for example, “paraben-free”—to infer enhanced safety is not acceptable.

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 interests of the consumer on board when creating and substantiating a claim aligns the marketer well with the regulatory criteria of Truthfulness, Evidential Support, Honesty, and Fairness. This thinking underpins the approach presented here in creating a rationale for the claim using the building blocks approach to classify and support claims. This will help overcome challenges in claim substantiation.

Acknowledgements 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 our former Oriflame colleague Ms Claire Byrne for suggestions, discussions, and support in preparing the original manuscript.

References

1. Regulation (EC) No 1223/2009 of the European parliament and of the council of 30 November 2009 on cosmetic products. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02009R1223-20180801>
2. Commission Regulation (EU) No 655/2013 laying down common criteria for the justification of claims used in relation to cosmetic products. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:190:0031:0034:EN:PDF>
3. COLIPA. Guidelines for the evaluation of the efficacy of cosmetic products, Revised version. 2008.
4. Reed RE. The definiton of cosmeceuticals. *J Soc Cosmet Chem.* 1962;13(2):103.
5. Kligman AM. Cosmeceuticals: do we need a new category? In: Elsner P, Maibach HI, editors. *Cosmeceuticals; drugs vs. cosmetics.* New York: Marcel Dekker; 2000.
6. Masuda M, Harada F. Legislation in Japan. In: Paye M, Barel AO, Maibach HI, editors. *Handbook of cosmetic science and technology.* Boca Raton, FL: Taylor & Francis Group; 2006.
7. Milstein SR, Harper AR, Katz LM. Regulatory requirements for the marketing of cosmetics in the United States. In: Paye M, Barel AO, Maibach H, editors. *Handbook of cosmetic science and technology.* Boca Raton, FL: Taylor & Francis Group; 2006.
8. https://www.hsa.gov.sg/content/dam/HSA/HPRG/Cosmetic_Products/Appendix%20III%2010%20September%202007A.pdf.
9. China Cosmetic Regulations. <https://msc.ul.com/en/resources/article/china-compliance-cosmetic-products/>
10. Federal Food, Drug and Cosmetic Act. 1938.

11. CTPA, ASA. Advertising claims, CTPA guide. 2009.
12. Directive 2006/114/EC of the European parliament and of the council concerning misleading and comparative advertising. 2006 Dec 12.
13. ISO 24444:2010 Cosmetics—sun protection test methods—in vivo determination of the sun protection factor (SPF)
14. Meilgaard M, Civille GV, Carr CM. Sensory evaluation techniques. Boca Raton, FL: CRC Press; 1999.
15. Dus CA, Rudolf M, Civille GV. Sensory testing methods used for claim support. In: Aust LB, editor. Cosmetic claim substantiation. Boca Raton, FL: CRC Press; 1997.
16. Griffiths CE, Wang TS, Hamilton TA, et al. A photonumeric scale for the assessment of cutaneous photodamage. *Arch Dermatol*. 1992;128(3):347–51.
17. Humbert P. Classification of facial wrinkling. In: Agache P, Humbert P, editors. Measuring the skin. Heidelberg: Springer; 2006.
18. Lemperle G, Holmes RE, Cohen SR, et al. A classification of facial wrinkles. *Plast Reconstr Surg*. 2001;108:1735–50.
19. Couteau C, Coiffard LJM, Bille-Rivain V. Influence of excipients on moisturizing effect of urea. *Drug Dev Ind Pharm*. 2006;32(2):239–42.
20. Dobrev H. Use of Cutometer to assess epidermal hydration. *Skin Res Technol*. 2000;6:239–44.
21. Fujimura T, Haketa K, Hotta M, et al. Loss of skin elasticity precedes to rapid increase of wrinkle levels. *J Dermatol Sci*. 2007;47:233–9, Corrected Proof.
22. Ambroisine L, Ezzedine K, Elfakir A, et al. Relationships between visual and tactile features and biophysical parameters in human facial skin. *Skin Res Technol*. 2007;13(2):176–83.
23. Chilcott RP, Farrar R. Biophysical measurements of human forearm skin in vivo: effects of site, gender, chirality and time. *Skin Res Technol*. 2000;6(2):64–9.
24. Latreille J, Gardinier S, Ambroisine L, et al. Influence of skin colour on the detection of cutaneous erythema and tanning phenomena using reflectance spectrophotometry. *Skin Res Technol*. 2007;13(3):236–41.
25. Caspers PJ, Lucassen GW, Puppels GJ. Combined in vivo confocal Raman spectroscopy and confocal microscopy of human skin. *Biophys J*. 2003;85(1):572–80.
26. Pudney PDA, Melot M, Caspers PJ, et al. An in vivo confocal raman study of the delivery of trans-retinol to the skin. *Appl Spectrosc*. 2007;61(8):804–11.
27. Xiao P, Packham H, Zheng X, et al. Opto-thermal radiometry and condenser-chamber method for stratum corneum water concentration measurements. *Appl Phys B*. 2007;86(4):715–9.
28. Abella ML. Evaluation of anti-wrinkle efficacy of adenosine-containing products using the FOITS technique. *Int J Cosmet Sci*. 2006;28(6):447–51.
29. Akazaki S, Imokawa G. Mechanical methods for evaluating skin surface architecture in relation to wrinkling. *J Dermatol Sci*. 2001;27(Suppl 1):5–10.
30. Baret M, Bensimon N, Coronel S, et al. Characterization and quantification of the skin radiance through new digital image analysis. *Skin Res Technol*. 2006;12(4):254–60.
31. Hawkins SS, Perrett DI, Burt DM, et al. Prototypes of facial attributes developed through image averaging techniques. *Int J Cosmet Sci*. 1999;21(3):159–65.
32. Rosen BG, Blunt L, Thomas TR. On in-vivo skin topography metrology and replication techniques. *J Phys Conf Ser*. 2005;13:325–9.
33. Alencastre JB, Bentley M, Garcia FS, et al. A study of the characteristics and in vitro permeation properties of CMC/chitosan microparticles as a skin delivery system for vitamin E. *Rev Bras Cienc Farm*. 2006;42:71–6.
34. Diffey BL. In vitro assessment of the broad-spectrum ultraviolet protection of sunscreen products. *J Am Acad Dermatol*. 2000;43:1024–35.
35. Roguet R, Cohen C, Leclaire J, et al. Use of a standardized reconstructed epidermis kit to assess in vitro the tolerance and the efficacy of cosmetics. *Int J Cosmet Sci*. 2000;22(6):409–19.
36. Schlotmann K, Kaeten M, al BAE. Cosmetic efficacy claims in vitro using a three-dimensional human skin model. *Int J Cosmet Sci*. 2001;23:309–18.
37. Eisfeld W. “Messmethoden zur Bestimmung der Wirksamkeit kosmetischer Mittel – Haar”, chapter 6.3 in: Umbach W. editor, “Kosmetik und Hygiene von Kopf bis Fuß”, 3., completely revised and extended edition, Wiley-VCH GmbH & Co. KgaA, Weinheim, 2004;99–111.

38. Kon R, Nakamura A, Takeuchi K. Artificially damaged hairs: preparation and application for the study of preventive ingredients. *Int J Cosmet Sci.* 1998;20(6):369–80.
39. Hornby et al. Cyclic testing: demonstrating conditioner benefits on damaged hair. *Cosmet. & Toilet.* 2011;116(12):35–39.
40. Watson RE, et al. Repair of photoaged dermal matrix by topical application of a cosmetic ‘antiageing’ product. *Br J Dermatol.* 2008;158:472–7.
41. [Technical document on cosmetic claims](https://ec.europa.eu/docsroom/documents/24847) (version of 03 July 2017). <https://ec.europa.eu/docsroom/documents/24847>

Part II
General Aspects of Cosmetic Testing

Chapter 6

Testing Laboratory



Oliver Wunderlich

Core Messages

Efficient procedures, reliable study results, and overall success of a laboratory for clinical testing of cosmetic products critically depend on the characteristics of its premises:

- Sufficient rooms dedicated to special functions should be available, arranged in a favorable manner.
- Privacy is an important issue leading to different requirements in different areas of the laboratory.
- Special care should be taken for the design and equipment of the testing lab rooms, where most of the clinical procedures take place.
- Ambient conditions must be controllable (at least in some key areas) by air-conditioning.
- A Quality Management System should be implemented, with Standard Operating Procedures (SOPs) reflecting all procedures correctly.
- The Quality Management System must be “lived,” actual adherence to the SOPs must be checked, and SOPs must be under permanent review.
- Staff training is vital for good quality of study results.
- Measurement devices must be maintained and checked properly.

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 GmbH, Hamburg, Germany
e-mail: oliver.wunderlich@bioskinCRO.com

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 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)
- Restrooms 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. It should also be ensured that the room and files are protected from water and humid conditions, e.g., by keeping windows closed at all times. 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. Access to wireless internet is also beneficial. 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), e.g., by use of doors instead of drapes, and blinds for windows, and should be equipped with the necessary furniture and devices (desk, examination bed, calibrated 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. If several rooms are used simultaneously for an individual study, it must be ensured that the time displayed on the clocks is identical. 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 spectrometers, dedicated photography rooms, etc.). A space should be provided for volunteers to place their personal belongings when they are requested to undress, e.g., shelf, chair, and coat rack.

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 Restrooms for Volunteers

The restrooms for volunteers should be close to the waiting area and the testing lab rooms. It should be ensured that a key is easily available to unlock the restroom doors from the outside in case of a medical emergency.

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, $50 \pm 10\%$ 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 studies. A link to the current version of the GCP guideline [1] can be found below (European medicines agency [ema]): <https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice>

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
- Medical 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, following current standards. 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) [2–10] or the *Standardization Group of the European Society of Contact Dermatitis* (ESCD) [1–3]) 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) [11–13]).

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 (R2), EMA/CHMP/ICH/135/1995. December 2016. European Medicines Agency.
2. Monteiro Rodrigues L, Fluhr JW, the EEMCO Group. EEMCO guidance for the in vivo assessment of biomechanical properties of the human skin and its annexes: revisiting instrumentation and test modes. *Skin Pharmacol Physiol*. 2018;20:1–16.
3. Piérard GE. EEMCO guidance to the in vivo assessment of tensile functional properties of the skin. Part 1: relevance to the structures and ageing of the skin and subcutaneous tissues. *Skin Pharmacol Appl Skin Physiol*. 1999;12(6):352–62.
4. Rodrigues L, EEMCO. EEMCO guidance to the in vivo assessment of tensile functional properties of the skin. Part 2: instrumentation and test modes. *Skin Pharmacol Appl Skin Physiol*. 2001;14(1):52–67; Review.
5. Piérard GE, Piérard-Franchimont C, Marks R, Paye M, Rogiers V. EEMCO guidance for the in vivo assessment of skin greasiness. The EEMCO Group. *Skin Pharmacol Appl Skin Physiol*. 2000;13(6):372–89; Review.
6. Berardesca E, Loden M, Serup J, Masson P, Rodrigues LM. The revised EEMCO guidance for the in vivo measurement of water in the skin. *Skin Res Technol*. 2018;24(3):351–8.
7. Berardesca E, Lévêque JL, Masson P, EEMCO Group. EEMCO guidance for the measurement of skin microcirculation. *Skin Pharmacol Appl Skin Physiol*. 2002;15(6):442–56; Review.
8. Parra JL, Paye M, EEMCO Group. EEMCO guidance for the in vivo assessment of skin surface pH. *Skin Pharmacol Appl Skin Physiol*. 2003;16(3):188–202.

9. Piérard GE, Elsner P, Marks R, Masson P, Paye M, EEMCO Group. EEMCO guidance for the efficacy assessment of antiperspirants and deodorants. *Skin Pharmacol Appl Skin Physiol.* 2003;16(5):324–42.
10. Piérard GE, Piérard-Franchimont C, Marks R, Elsner P, EEMCO Group. EEMCO guidance for the assessment of hair shedding and alopecia. *Skin Pharmacol Physiol.* 2004;17(2):98–110.
11. Pinnagoda J, Tupker RA, Agner T, Serup J. Guidelines for transepidermal water loss (TEWL) measurement. A report from the Standardization Group of the European Society of Contact Dermatitis. *Contact Dermatitis.* 1996;35(1):1–10.
12. Fullerton A, Fischer T, Lahti A, Wilhelm KP, Takiwaki H, Serup J. Department of Dermatological Research, Leo Pharmaceutical Products, Ballerup, Denmark.: Guidelines for measurement of skin colour and erythema. A report from the Standardization Group of the European Society of Contact Dermatitis. *Contact Dermatitis.* 2002;46(3):129–40.
13. Fullerton A, Stücker M, Wilhelm KP, Wårdell K, Anderson C, Fischer T, Nilsson GE, Serup J. European Society of Contact Dermatitis Standardization Group.: Guidelines for visualization of cutaneous blood flow by laser Doppler perfusion imaging. A report from the Standardization Group of the European Society of Contact Dermatitis based upon the HIRELADO European community project. *Contact Dermatitis.* 2002;46(3):129–40.

Chapter 7

Research Staff



Ragna Williams

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 to supervise the actual study conduct with the aim to successfully complete the study performance in the agreed upon time frame and may be involved in the development of the study design.

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, Hamburg, Germany
e-mail: ragna.williams@bioskinCRO.com

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, any amendments, etc.), correct application of the test products, the test methods, and devices. The investigator should ensure that all the staff is 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 as well as any allergological 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. In case several investigators are involved, training to align grading assessments among all involved investigators is essential to minimize inter-rater differences.

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

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 current SOPs is a must.

Suggested Readings Cosmetics [Regulation \(EC\) No 1223/2009](#) of the European Parliament and of the Council.

Cosmetics [Directive 76/768/EEC](#) (cosmetics directive), as amended.

[Glossary of common ingredient names](#) for the purpose of labelling cosmetic products placed on the market (as established by [Decision \(EU\) 2019/701](#) of 5 April 2019).

Opinions on cosmetic ingredients of the Scientific Committee for Consumer Safety ([list of SCCS opinions](#)).

[Regulation \(EU\) 2016/679](#) on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.

Chapter 8

Testing Population



Ragna Williams

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 prospective 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.

R. Williams (✉)
bioskin GmbH, Hamburg, Germany
e-mail: ragna.williams@bioskinCRO.com

- 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.

Recruitment tools may state that subjects will receive a remuneration, but should not overemphasize 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/social networks 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 potential volunteers 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. It is essential to arrange an appointment within a relatively brief period after first contact, so that the potential volunteer does not lose 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, the 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 photograph, note, or 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 and have previously consented to being contacted by the test institute. 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” or vendors providing internet advertising (in some cases including pre-screening) are a common methods 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, wash-out periods, 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, gender, 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°).

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 amount, 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.

Suggested Reading https://ec.europa.eu/health/scientific_committees/consumer_safety/opinions/sccnfp_opinions_97_04/sccp_out87_en.htm (Opinion concerning guidelines on the use of human volunteers in compatibility testing of finished cosmetic products - adopted by the scientific committee on cosmetics and non-food products intended for consumers during the plenary session of 23 June 1999).

Chapter 9

Testing Devices and Methods



Gabriel Khazaka

Core Messages

- Context of objective measurement methods
- Claim support and efficacy testing in formulating cosmetic products
- Assessment of functional skin parameters
- Norms, Standard, and Directives explaining the use of the methods
- Facts to consider when using the measurement devices in the laboratory

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 functional skin parameters. Without such instrumentation, it is almost impossible to determine the effect of specific products on skin functions.

Subjective and objective evaluation of the products efficacy, however, are mutually dependent. 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 (e.g., in evaluating inflammation, sun screen products or claims for sensitive skin), this becomes clearer.

Subjectively, the human brain cannot process subclinical changes in color, especially when viewed at different points in time or on pigmented skin. Therefore, instrumentation to record even minimal changes in the redness of the skin is necessary. The achieved results must then be interpreted in context with the expected

G. Khazaka (✉)
Courage + Khazaka Electronic GmbH, Köln, Germany
e-mail: gkhazaka@courage-khazaka.de

outcome or the hypothesis. This interpretation is based on the personal experience and knowledge of the investigator.

This fact is not only true for the evaluation of skin color, but also for most measurement parameters, e.g., 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.

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 that the first commercially available skin measurement instruments were brought to the market at an affordable price. This possibility has contributed to the ability of cosmetic manufacturers to improve their product quality and also has increased 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, they 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 acquiring these economic and easy-to-use devices to test skin parameters in their own laboratory, small and midsized companies can also 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 an 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.

Ethics approval should be obtained even for preliminary studies.

9.2 Technical Assessment of the Various Skin Parameters

Before the first skin testing instruments were introduced to 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 at 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 developments [1–4]. Now, almost 40 years later, there is a large variety of equipment available for assessing a number of different parameters and it is difficult to review 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 (elasticity)
Color and photo protection (melanin and erythema)
Gloss
Scaliness
Molecular profiles (some specific compounds)
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 [5, 6].

The development of other, new techniques is moving on quickly. In addition to *in vivo* Raman spectroscopy, multiphoton tomography, optical coherence 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, 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 on his own 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 the appropriate use of such techniques are missing.

The results of trials of one group are not necessarily comparable to those of a second group. Another difficulty lies in the different technical background of the devices themselves. The measurement principles differ considerably depending on the parameter and manufacturer of the instrument, thus making comparisons of the results close to impossible or very cumbersome.

This challenge 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 today, published a variety of guidelines for different skin measurement parameters (e.g., moisture, pH-, microcirculation, and many others) [7–12]; see also Chap. 4.

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, the 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 as follows. For optimal use of these technical devices, a minimum of these 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 that 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 have to be calibrated regularly. Some manufactures 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.

References

1. Agache P, Caperan A, Barrand C, Laurent R, Dagrass G. Determination of cutaneous surface lipids by the photometric method. *Ann Dermatol Syphiligr (Paris)*. 1974;101:285–7.
2. Peck S, Glick AW. A new method for measuring the hardness of keratin. *J Soc Cosmet Chem*. 1956;7:530–40.
3. Schaefer H, Kuhn-Bussius H. A method for the quantitative determination of human sebum secretion. *Arch Klin Exp Dermatol*. 1970;238:429–35.

4. Tronnier H, Schule D. An experimental and clinical investigation concerning the antipruritic effect of pharmaceutical preparations. *Arztl Forsch.* 1968;22:203–12.
5. Agache P, Humbert P. *Measuring the skin.* Berlin: Springer; 2004.
6. Fluhr J, Elsner P, Berardesca E, Maibach HI, editors. *Bioengineering of the skin: water and the stratum corneum.* 2nd ed. CRC: Boca Raton, FL; 2005.
7. Berardesca E. EEMCO guidance for the assessment of stratum corneum hydration: electrical methods. *Skin Res Technol.* 1997;3:126–32.
8. Berardesca E, Leveque JL, Masson P. EEMCO guidance for the measurement of skin micro-circulation. *Skin Pharmacol Appl Skin Physiol.* 2002;15:442–56.
9. Parra JL, Paye M. EEMCO guidance for the in vivo assessment of skin surface pH. *Skin Pharmacol Appl Skin Physiol.* 2003;16:188–202.
10. Rogiers V. EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences. *Skin Pharmacol Appl Skin Physiol.* 2001;14:117–28.
11. Berardesca E, Loden M, Serup J, Masson P, Rodrigues LM. The revised EEMCO guidance for the in vivo measurement of water in the skin. *Skin Res Technol.* 2018;24:351–8.
12. Monteiro Rodrigues L, Fluhr JW, the EEMCO Group. EEMCO guidance for the in vivo assessment of biomechanical properties of the human skin and its annexes: revisiting instrumentation and test modes. *Skin Pharmacol Physiol.* 2020;33:44–60.

Chapter 10

Factors Influencing Measurements



Enzo Berardesca and Norma Cameli

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 potentially can 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, ...).

E. Berardesca (✉)
Phillip Frost Department of Dermatology, Miller School of Medicine, University of Miami,
Miami, FL, USA
e-mail: berardesca@berardesca.it

N. Cameli
San Gallicano Dermatological Institute IRCCS, Rome, Italy

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 [1]. 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 [2–4]. 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 inter-instrumental variability, an additional calibration procedure incorporating a calibration for an actual gold reference standard can be adopted [5].

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 [6–8]. 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 [9]. 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 [10, 11]. Disturbances in the microclimate are immediately detected as a fluctuation in TEWL and can alter other measures such as capacitance, friction, and stratum corneum cohesion [12].

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 [12]. 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 as much as possible undesirable air turbulence [7, 8]. 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. Closed chamber techniques can be utilized to minimize these possible errors.

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) [13, 14].

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 [8]. 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 [15]. 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 [16]. 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 [6, 7]. 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 [17].

10.2.2.5 Skin Cleansing

Cleansing of the skin with surface active agents and solvents could modify surface microenvironment [18] due to damage of skin barrier function [8, 19]. Changes in skin hydration and TEWL could also result from the removal of occlusive substances from the skin surface, like cream [20] 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 s.c. hydration which might interfere with several biophysical parameters ranging from microtexture, cell cohesion, stratum corneum mechanical properties, and friction coefficient [21]. A rapid superficial cleansing of the skin with water-free ether appears not to increase the water loss values [20]. 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 [22]. In fact, based on the almost identical values for the parameters of TEWL, s.c. hydration and pH value, the skin physiology of the child differs very little in s.c. hydration and barrier function from that of adults [23].

From adulthood through senescence the age dependence of the cutaneous permeability barrier function have been controversially discussed. Wilhelm and Maibach [24] suggest that there is evidence that baseline TEWL is reduced in aged individuals as compared with mid-adulthood values. Furthermore, with increasing age, significant 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 [25, 26]. Between black and white human skin some differences have been reported [27–29]. This is also the case for white and Hispanic subjects [30]. 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 [31], different corneum thickness between anatomical sites [32], 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 [7]. Indeed, skin at different body sites shows distinct patterns of barrier recovery that are likely to be related to structural and physiological differences [33]. Lipid-rich skin areas (e.g., the forehead) are the most vulnerable to barrier disruption [33]. 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 [34, 35] but this has not been found by all investigators [8, 36]. Recently, it has been suggested that for the evaluation of cosmetic formulations facial skin would be more suitable than the volar forearm [37].

10.2.3.3 Sweating

Physical, thermal, and emotional sweating are important variables which need to be controlled [38, 39]. If the ambient air temperature is below 20 °C and the skin temperature is below 30 °C, thermal sweat gland activity is unlikely, provided that the skin is not exposed to forced convection and no excessive body heat is produced (as a result of physical exercise) [40–42]. Therefore, a pre-measurement 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 of 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 [10, 17, 21, 43, 44]. Changes in barrier function are caused by chemical contacts [45, 46], surfactant damage [47, 48], or diseased states (dermatitis), e.g., psoriasis [49] and eczema [44], resulting in an increased water evaporation rate within the range of 20–60 g/m²h. However, more severe damage to the barrier such as burns gives rise to much higher evaporation rates (above 100 g/m²h) 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 [50]. 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 [51, 52]. More recently, Le Fur et al. have found a bimodal rhythm for TEWL with two peaks located at 8:00 and 16:00 [51].

10.2.3.7 Menstrual Cycle and Hormones

It has been reported that differences in skin reactivity may occur in the premenstrual phase [53] since an increased stratum corneum hydration can occur and could potentially affect transcutaneous penetration and other physiological responses (skin mechanical properties, skin texture, etc.). In long-term studies, where repeated control measurements are taken, it is important to take record of the day of cycle.

10.2.3.8 Intra- and Interindividual Variation

For most skin sites important interindividual variations occur, usually also dependent on the measuring device [6–8, 34, 54]. 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 usually is smaller [8, 54]. A wide interindividual range of variation is particularly seen when individuals undergo certain treatments. A typical example are TEWL measurements after sodium lauryl sulfate damage of the skin.

Table 10.1 Factors influencing measurements in cosmetic testing

Variables influencing measurements	What to do
<i>Environment and climate</i>	
Room temperature and relative humidity	Keep between 20 and 21° and 40–60%
Seasonal variations	Prefer short term studies
Outside temperature and relative humidity	Make sure volunteers rest and adjust at least 20 min
<i>Instrument related</i>	
Calibration	Calibrate frequently
Different models	Use gold standard reference
New vs. old devices	Use gold standard reference
<i>Measurements</i>	
Area, position, surface	Standardize as much as possible
Probe	Do multiple measures in the same area
	Reduce contact time with skin
<i>Subject related</i>	
Age, sex, ethnic group, body site	Carefully control and standardize
Cleansing	Standardize and make measures after 2 h from the last procedure

10.3 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 and 22 °C \pm 1 °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, e.g., 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 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 preferably should 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).
- Other factors such as hormones, replacement therapy, and other systemic treatments should be monitored and taken in account eventually.

References

1. Schultz A, Elsner P, Burg G. Quantification of irritant contact dermatitis in vivo: comparison of the Dermatest system with the Evaporimeter. *Contact Dermatitis*. 1991;24:235–7.
2. Grove G, Grove MJ, Zerweck C, Pierce E. Comparative metrology of the Evaporimeter and the DermaLab TEWL Probe. *Skin Res Technol*. 1999;5:1–8.
3. Grove G, Grove MJ, Zerweck C, Pierce E. Computerised evaporimetry using the DermaLab TEWL Probe. *Skin Res Technol*. 1999;5:9–13.
4. Pinnagoda J, Tupker RA, Coenraads PJ, Nater JP. Comparability and reproducibility of results of water loss measurements: a study of four evaporimeters. *Contact Dermatitis*. 1989;20:241–6.
5. Rogiers V. EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences. *Skin Pharmacol Appl Skin Physiol*. 2001;14:117–28.
6. Nilsson GE. Measurement of water exchange through skin. *Med Biol Eng Comput*. 1977;15:209–13.
7. Pinnagoda J, Tupker RA, Agner T, Serup J. Guidelines for transepidermal water loss (TEWL) measurement. *Contact Dermatitis*. 1990;22:164–8.
8. Potts RO. Stratum corneum hydration: experimental techniques and interpretations of results. *J Soc Cosmet Chem*. 1986;37:9–33.
9. Wheldon AE, Monteith JL. Performance of a skin Evaporimeter. *Med Biol Eng Comput*. 1980;18(201):204.
10. Batt MD, Fairhurst E. Hydration of the stratum corneum. *Int J Cosmet Sci*. 1986;8:253–64.
11. Blichmann CW, Serup J. Reproducibility and variability of transepidermal water loss measurements. *Acta Derm Venereol*. 1987;67:206–109.
12. *A Guide to Water Evaporation Rate Measurement*, ServoMed, Vallingby, Stockholm, Sweden, 1981.
13. Agner T, Serup J. Seasonal variation in skin resistance to irritants. *Brit J Dermatol*. 1989;121:323–8.
14. Rothman S. The role of the skin in thermoregulation: factors influencing skin surface temperature. In: Rothman S, editor. *Physiology and Biochemistry of the skin*. Chicago: The University of Chicago Press; 1954. p. 258.
15. Rogers J, Harding C, Mayo A, Banks J, Rawlings A. Stratum corneum lipids: the effect of ageing and the seasons. *Arch Dermatol Res*. 1996;288:765–70.
16. Rodrigues L, Jaco I, Melo M, Silva R, Pereira LM, Catorze N, Barata E, Ribeiro H, Morais J. About claims substantiation for topical formulations: an objective approach to skin care products biological efficacy. *J Appl Cosmetol*. 1996;14:93–8.

17. Lévêque JL. Measurement of transepidermal water loss. In: Lévêque JL, editor. *Contaneous investigation in health and disease. Noninvasive methods and instrumentation.* Marcel Dekker Inc: New York; 1989. p. 135–53.
18. Fartasch M. Human barrier formation and reaction to irritation. *Curr Prob Dermatol.* 1995;23:95–103.
19. Loden M. The increase in skin hydration after application of emollients with different amounts of lipids. *Acta Derm Venereol (Stockholm).* 1992;72:327–30.
20. Giusti F, Martella A, Bertoni L, Seidenari S. Skin barrier, hydration, and pH of the skin of infants under 2 years of age. *Pediatr Dermatol.* 2001;18:93–6.
21. Proksch E, Brasch J, Sterry W. Integrity of the permeability barrier regulates epidermal Langerhans cell density. *Br J Dermatol.* 1996;134:630–8.
22. Fluhr JW, Pfisterer S, Gloor M. Direct comparison of skin physiology in children and adults with bioengineering methods. *Pediatr Dermatol.* 2000;17:436–9.
23. Wilhelm KP, Maibach HI. Transepidermal water loss and barrier function of aging human skin. In: Elsner P, Berardesca E, Maibach HI, editors. *Bioengineering of the skin, water and the stratum corneum.* Boca Raton, FL: CRC Press; 1994. p. 133–45.
24. De Paepe K, Vandamme P, Derde MP, Roseeuw D, Rogiers V. Ceramide/cholesterol/FFA-containing body lotions: effects on the TEWL of aged and SDS-damaged skin; in Ziolkowsky H GmbH (ed): *Conference proceedings, Verlag für Chemische Industrie, Active Ingredients Conference, Paris, pp. 97–111, 1996.*
25. De Paepe K, Vandamme P, Roseeuw D, Rogiers V. Ceramides/cholesterol/FFA-containing cosmetics: the effect on barrier function. *SOFW J.* 1996;122:199–204.
26. Wilson D, Berardesca E, Maibach HI. In vivo transepidermal water loss and skin surface hydration in assessment of moisturization and soap effects. *Int J Cosmet Sci.* 1988;10:201–11.
27. Berardesca E, Maibach HI. Racial differences in sodium-lauryl induced cutaneous irritation. Comparison of black and white subjects. *Contact Dermatitis.* 1988;18:65–70.
28. Berardesca E, Maibach HI. sodium-lauryl-sulphate-induced cutaneous irritation. Comparison of black and white subjects. *Contact Dermatitis.* 1988;19:136–40.
29. Berardesca E, Pirot F, Singh M, Maibach H. Differences in stratum corneum pH gradient when comparing white Caucasian and black African-American skin. *Br J Dermatol.* 1988;139:855–7.
30. Yoshikawa N, Imokawa G, Akimoto K, et al. Regional analysis of ceramides within the stratum corneum in relation to seasonal required for normal barrier homeostasis. *J Invest Dermatol.* 2000;115:459–66.
31. Plewig G, Marples RR. Regional differences of cell sizes in the human stratum corneum. *J Invest Dermatol.* 1970;54:13–8.
32. Fluhr JW, Dickel H, Kuss O, Weyher I, Diepgen TL, Berardesca E. Impact of anatomical location on barrier recovery, surface pH and stratum corneum hydration after acute barrier disruption. *Br J Dermatol.* 2002;146:770–6.
33. Treffel P, Panisset F, Faivre B, Agache P. Hydration, transepidermal water loss, pH and skin surface parameters: correlations and variations between dominant and non-dominant forearms. *Brit J Dermatol.* 1994;130:325–8.
34. Oestmann E, Lavrijsen APM, Hermans J, Ponc M. Skin barrier function in healthy volunteers as assessed by transepidermal water loss and vascular response to hexyl nicotinate: intra- and inter-individual variability. *Brit J Dermatol.* 1993;128:130–6.
35. Rodrigues L, Pereira LM. Basal transepidermal water loss: right/left forearm difference and motoric dominance. *Skin Res Technol.* 1998;4:135–7.
36. Schnetz E, Kuss O, Schmitt J, Diepgen TL, Kuhn M, Fartasch M. Intra- and inter-individual variations in transepidermal water loss on the face; facial locations for bioengineering studies. *Contact Dermatitis.* 1999;40:243–7.
37. Pinnagoda J, Tupker RA, Smit JA, Coenraads PJ, Nater JP. The intra- and inter-individual variability and reliability of transepidermal water loss measurements. *Contact Dermatitis.* 1989;21:255–9.
38. Holbrook KA, Odland GF. Regional differences in the thickness (cell layers) of the human stratum corneum: an ultrastructural analysis. *J Invest Dermatol.* 1974;62:415–22.

39. Shahidullah M, Raffle EJ, Rimmer AR, Frain-Bell W. Transepidermal water loss in patients with dermatitis. *Br J Dermatol*. 1969;81:722–5.
40. Baker H, Kligman AM. Measurement of transepidermal water loss by electrical hygrometry. Instrumentation and responses of physical and chemical insults. *Arch Dermatol*. 1967;96:441–4.
41. Piérard GE, Goffin V, Hermanns-Lê T, Arrese JE, Piérard-Franchimont C. Surfactant induced dermatitis. A comparison of corneosurfametry with predictive testing on human and reconstructed skin. *J Acad Dermatol*. 1995;33:462–9.
42. Pinnagoda J, Tupker RA, Coenraads PJ, Nater JP. Transepidermal water loss: with and without sweat gland inactivation. *Contact Dermatitis*. 1989;21:16–22.
43. Chao KN, Eislely JG, Jei YW. Heat and water losses from burnt skin. *Med Biol Eng Comput*. 1977;15:598–603.
44. Tagami H, Kobayashi H, Zhen XS, Kikuchi K. Environmental effects on the functions of the stratum corneum. *J Investig Dermatol Symp Proc*. 2001;6:87–94.
45. Chilcott RP, Brown RF, Rice P. Non-invasive quantification of skin injury resulting from exposure to sulphur mustard and Lewisite vapours. *Burns*. 2000;26:245–50.
46. Tupker RA, Pinnagoda J, Coenraads PJ, Nater JP. The influence of repeated exposure to surfactants on the human skin as determined by transepidermal water loss and visual scoring. *Contact Dermatitis*. 1989;20:108–11.
47. Fartasch M. Ultrastructure of the epidermal barrier after irritation. *Microsc Res Tech*. 1997;37:193–9.
48. Grice KA. Transepidermal water loss in pathologic skin. In: Jarret, editor. *The physiology and pathophysiology of the skin*, vol. 6. London: Academic Press; 1980. p. 2147.
49. Chilcott RP, Farrar R. Biophysical measurements of human forearm skin in vivo: effects of site, gender, chirality and time. *Skin Res Technol*. 2000;6:64–9.
50. Mehling A, Fluhr J. Chronobiology: biological clocks and rhythms of the skin. *Skin Pharmacol Physiol*. 2006;19:182–9.
51. Le Fuhr I, Reinberg A, Lopez S, et al. Analysis of circadian and ultradian rhythms of skin surface properties of face and forearm of healthy women. *J Invest Dermatol*. 2001;111:718–24.
52. Reinberg A, Le Fur I, Tschachler E. Problem related to circadian rhythms in human skin and their validation. *J Invest Dermatol*. 1998;111:708–9.
53. Agner T, Damm P, Skouby SO. Menstrual cycle and skin reactivity. *J Am Acad Dermatol*. 1991;24(4):566–70.
54. Goffin V, Piérard-Franchimont C, Piérard GE. Passive sustainable hydration of the stratum corneum following surfactant challenge. *Clin Exp Dermatol*. 1999;24:308–11.

Chapter 11

Study Design



Betsy Hughes-Formella, Nicole Braun, Ulrike Heinrich, and Carmen Theek

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.

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, e.g., for assessment of a moisturizing effect multiple

B. Hughes-Formella

Strategic and Clinical Consulting, Dermatology Product Development, Schoenningstedter
Strasse 128b, Reinbek, Germany

e-mail: betsy.hughes@derm-consult.com

N. Braun (✉) · U. Heinrich

DermaTronnier, Institute for Experimental Dermatology at Witten/Herdecke University,
Witten, Germany

e-mail: nicole.braun@uni-wh.de; ulrike.heinrich@uni-wh.de

C. Theek

CTS Statistikservice, Herdecke, Germany

e-mail: carmen.theek@ct-statistikservice.de

test fields on the two forearms may be chosen or for assessment of an anti-wrinkle 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 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 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:

Study objective:	The evaluation of skin moisturizing effect
Measurement variable:	Corneometric value at any given time point
Study endpoint:	The change to baseline in the corneometric value at the end of the treatment interval.

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 variables for confirmatory testing. Usually, there is only one primary variable. Dependent on the type of variable that should be investigated, e.g., a mean or a rate, different underlying test procedures are to be applied. Consequently, different estimates for key input parameters are needed. For example, in case of a two-sample *t*-test for comparison of mean values estimates of the expected standard deviation and mean effect are required for the sample size 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, either a pilot study should be conducted or conservative assumptions should be used for this purpose. Additionally, the significance level α (e.g., 5% two-sided) and the power $1-\beta$ (e.g., 80%) have to be fixed to perform the sample size calculation.

In case the analysis for confirmatory testing is more complex, like considering more than one primary endpoint, a more comprehensive approach for the sample size calculation has to be applied accordingly.

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., continuous, ordinal, cardinal) or type of comparison (paired versus group comparisons). When applying inferential statistics, the significance level α 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 objective to be investigated by confirmatory testing and secondary objectives for exploratory evaluation. 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.

In case data do not follow an underlying normal distribution (e.g., in case of skewness or outliers) required by a parametric approach, a nonparametric test procedure might be the preferred method.

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 also offered 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 a formal independent review is generally not legally required for noninvasive, nontherapeutic cosmetic studies with healthy test participants. Nevertheless, cosmetic studies must comply with international guidelines. In this context, different definitions regarding “healthy volunteer” are described by the WHO or the Royal College of Physicians and parameters based on which the status “healthy” can be ascertained are described by Breithaupt-Groegler et al. [2]. In studies with patients (e.g., diabetics), even if only a part of the collective, or if invasive measurements are used, a positive vote of an ethics committee is mandatory. In general, it also makes sense to have an umbrella positive vote from an ethics committee for noninvasive measurements of substances that are not medical drugs. Whichever the case, there are often reasons for a review and an evaluation of the study design is always a central step in this review process. As stated in Fernando et al. [3], 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:

- Study is safe for the participants.
- 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 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); benefit-risk assessment; expense allowance; insurance coverage of the test subject; data handling and processing; and data analysis procedures.

During 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 *Data Protection*

In May 2018, the European Union adopted the new General Data Protection Regulation 2016/679 regarding processing of personal data and on the free movement of such data.

According to GDPR Article 4 (15) “data concerning health” means personal data related to the physical or mental health of a natural person. These health data belong to the special categories of personal data and there is a special need for protection.

This results in a fundamental ban on the processing of this data. The processing is only possible under certain conditions (Art. 9 (2a-j) GDPR) including, for example, that the subject consents to the processing of his/her data.

According to Art. 37 exp. 1 of the GDPR, a data protection representative must be appointed in each case with the following three constellations:

- It is part of an authority or public office.
- The main activity is to perform data processing which, by virtue of its nature, scope, and/or purpose, requires extensive, regular, and systematic patient monitoring.
- The core activity consists in the extensive processing of special categories of data as defined in Art. 9 GDPR (in particular health data).

In terms of test subject rights the main points are that the test subjects have the right of access to their own personal data, the right to have their personal data deleted, the right to rectify inaccurate personal data, the right to appeal and to effective remedy, and the right to compensation. According to Article 7 (1) GDPR, a controller shall be able to demonstrate that the subject has consented to processing his or her personal data.

In the event of a breach of personal data, the controller (natural persons or entities collecting and processing personal data) shall inform the supervisory authority no later than 72 h after becoming aware of it [4].

With regard to data protection in any subject database, e.g., for recruitment, all test subjects must be anonymized. In addition, test subjects can request deletion from the database at any time.

11.8.3 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), what will happen during the study, the number of visits and schedule of procedures (flowchart often helpful), and use of results using clear and plain language. 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.

Test participants have also to be informed how to withdraw consent (Art. 7, (3) GDPR). Children below 16 need parental consent for the collecting and analyzing of data to be lawful (Art. 8, (1) GDPR).

11.8.4 Case Report Form (CRF) and Subject Diaries

A data collection form, the 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, e.g., 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 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, e.g., 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

1. Bebski V, Marschner I, Keech AC. Specifying objectives and outcomes for clinical trials. *MJA*. 2002;176:491–2.
2. Breithaupt-Groegler K, Coch C, Coenen M, et al. Who is a ‘healthy subject’?-consensus results on pivotal eligibility criteria for clinical trials. *Eur J Clin Pharmacol*. 2017;73(4):409–16.
3. Fernando M, Dissanayake VHW, Corea E, editors. Ethics Review Committee guidelines: a guide for developing standard operating procedures for committees that review biomedical research proposals. Forum of Ethics Review Committees, Sri Lanka, 2007.
4. Regulation (EU) 2016/679 of the European Parliament and of the Council 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

Chapter 12

Cosmetic Testing Report



Hristo Dobrev

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 (GCP) [1, 2]. GCP 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 GCP, 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 [1, 2]. 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 CONSORT Statement guidance could also be adapted and used when preparing cosmetic testing report and publication. The CONSORT Statement is an evidence-based, minimum set of recommendations for reporting randomized trials. It comprises a 25-item checklist and a flow diagram. The checklist items focus on reporting how

H. Dobrev (✉)

Department of Dermatology and Venereology, Medical University, Plovdiv, Bulgaria

the trial was designed, analyzed, and interpreted while the flow diagram displays the progress of all participants through the trial. The CONSORT Statement is to be used in conjunction with the CONSORT “Explanation and Elaboration” document that explains and illustrates all underlying principles. CONSORT 2010 is the current version of the guideline and supersedes the 2001 and 1996 versions [3–5].

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 the completion of the biomedical research (clinical trial), a final report should be prepared and signed by the principal investigator [1, 2]. 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 Harmonisation (ICH) in 1995 [6]. 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 [7, 8]. 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 [2, 6, 7, 9–12]:

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 (GCP), including the archiving of essential documents
- Date of the report (Date of initial version and amendments, if applicable)

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

An 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 person (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 study nature 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 the 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)

12.2.9.1 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 for assigning participants to treatment groups, product application (quantity, frequency, time, areas), subject compliance monitoring and prior and concomitant skin care products/medicines allowed.

12.2.9.2 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.

12.2.9.3 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: 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).

12.2.9.4 Evaluation Parameters (Test Variables)

This section should describe the specific variables (e.g., efficacy or safety criteria adopted, categorical and qualitative variables, undesirable and side effects) to be assessed.

12.2.9.5 Evaluation Methods and Equipment

This section should describe the specific methods used for assessment of the evaluation parameters:

- Use tests by consumers (method and format)
- Auto-evaluation by users (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 noninvasive 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).

12.2.9.6 Study Protocol

The study procedures accomplished at each visit should be described.

12.2.9.7 Statistics

The information about the statistical analysis applied should describe the sample size calculation, subject allocation (and randomization), comparison grouping, variables analyzed, statistical tests, and data processing used.

12.2.9.8 Data Quality Assurance

Systems with procedures that assure the control and quality of every aspect of the trial should be described.

12.2.9.9 Changes in the Conduct of the Study or Planned Analyses

Any change in the conduct of the study or planned analyses 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” [13].

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 by 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 an 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 orally or poster presentation in scientific meetings and by publishing in popular and peer-reviewed journals. Nowadays, the internet gives an opportunity to revolutionary spreading of scientific information [14, 15].

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 is used or passed on to a third party [11, 12].

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 [2].

The test report and scientific paper 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.

There are many guidelines for preparation and submission of the scientific paper to biomedical journals [13, 16–18]. 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 [13]. 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 [13–15, 18].

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 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 provide 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 details 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 clearly and in logical sequence presented 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 of 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 [19] 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 for 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 web site. 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 [13].

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 [13, 14].

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 the way it has been 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.

Key Messages for Preparing Cosmetic Testing Report and Scientific Article

Cosmeting testing report	Scientific article
Title page <ul style="list-style-type: none"> • Title of the study • Identifying number of the study • Sponsor(s) • Study site/center/department/laboratory • Investigator(s) • Test product(s) • Study date • Date of report 	Title page <ul style="list-style-type: none"> • Title of the article • Author(s) name and their affiliations • Name of department(s) and institution(s) • Name of corresponding author • Source(s) of support • A short running head • Word counts for the text • Tables and figures number

Cosmeting testing report	Scientific article
	Conflict of Interest Notification Page
Synopsis (Study summary) <ul style="list-style-type: none"> • Study title • Study sponsor • Study design • Study duration • Study center • Objectives • Participants (main inclusion/exclusion criteria) • Test products (active ingredients) • Methods and instrumentation • Measures • Protocol • Statistical methods • Results with numerical data • Conclusion 	Abstract (Summary) <ul style="list-style-type: none"> • Background and aims • Material and methods • Results • Discussion and conclusions
	Keywords
Table of contents	
List of abbreviations	
Ethics	
Investigators and study administrative structure	
Introduction and objectives	Introduction and objectives
Investigational plan (Material and Methods) <ul style="list-style-type: none"> • Participants (Volunteers) • Product(s) tested • Study design and plan • Evaluation parameters (test variables) • Evaluation methods and equipment • Study protocol • Statistics • Data quality assurance • Changes in the conduct of the study or planned analyses 	Material and Methods <ul style="list-style-type: none"> • Subjects (Participants) including ethical considerations • Product(s) tested • Methods • Protocol • Ethics • Statistics
Results	Results
Discussion and Conclusions	Discussion and conclusions
	Acknowledgments
Reference list	References
Signatures of sponsor, investigator(s)) and author of the test report	
Appendices (Attachments)	Tables and Figures
	Tables and Figures legends
	Units of measurements
	Abbreviations and symbols
	Cover letter <ul style="list-style-type: none"> • Statements about any possible conflicts of interest • Name, address and signature of corresponding author

References

1. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Guideline for Good Clinical Practice E6(R1), 2002. <http://www.jpgmonline.com/article.asp?issn=0022-3859;year=2001;volume=47;issue=1;spage=45;epage=50;aulast=>. Assessed 21 Dec 2019.
2. World Health Organisation. Handbook for good clinical research practice (GCP), 2002. http://www.who.int/medicines/areas/quality_safety/safety_efficacy/gcp1.pdf. Assessed 22 Dec 2019.
3. Campbell M, Piaggio G, Elbourne D, Altman D. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
4. CONSORT transparent reporting of trials. CONSORT 2010 guidelines, 2010. <http://www.consort-statement.org>. Assessed 18 Nov 2019.
5. Moher D, Hopewell S, Schulz K, Montori V, Gøtzsche P, Devereaux P, Elbourne D, Egger M, Altman D. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
6. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports (E3), 1995. <https://www.ema.europa.eu/en/ich-e3-structure-content-clinical-study-reports>. Assessed 21 Dec 2019.
7. Alfaro V, Cullell-Young M, Tanovic A. Abbreviated clinical study reports with investigational medicinal products for human use: current guidelines and recommendations. *Croat Med J*. 2007;48:871–7.
8. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Guidance for Industry: submission of abbreviated reports and synopses in support of marketing applications. Rockville, MD: Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 1999. <https://www.fda.gov/media/71125/download>. Assessed 22 Dec 2019.
9. Bagatin E, Mior H. How to design and write a clinical research protocol in cosmetic dermatology. *An Bras Dermatol*. 2013;88(1):69–75.
10. Colipa. Colipa Guidelines. Guidelines for the evaluation of the efficacy of cosmetic products, 2008. <http://www.canippec.org.mx/woo/xtras/bibliotecavirtual/Colipa.pdf>. Assessed 21 Dec 2019.
11. University of Pennsylvania. Clinical research protocol, 2009. http://www.med.upenn.edu/ohrobjects/PM/2_ProtocolTemp_guidelines.doc. Assessed 22 Dec 2019.
12. World Health Organization. Recommended format for a research protocol, 2009. <https://www.who.int/ethics/review-committee/format-research-protocol/en/>. Assessed 22 Dec 2019.
13. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication, 2008. http://www.icmje.org/recommendations/archives/2008_urm.pdf. Assessed 21 Dec 2019.
14. Fathalla MF, Fathalla MMF. A practical guide for health researchers. WHO Regional Publications Eastern Mediterranean Series 30, Cairo, 2004. <https://apps.who.int/iris/bitstream/handle/10665/119703/dsa237.pdf?sequence=1&isAllowed=y>. Assessed 22 Dec 2019.
15. Highleyman L. A guide to clinical trials. Part II: interpreting medical research. *BETA*. 2006;18(2):41–7.
16. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663–94.
17. Bossuyt P, Reitsma J, Bruns D, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem*. 2003;49(1):1–6.

18. Word-Medex Pty Ltd. Word-Medex Guide. How to write good manuscripts, 2009. <http://www.word-medex.com.au/formatting/manuscript.htm>. Assessed 14 June 2009.
19. Patrias K. Citing medicine: the NLM style guide for authors, editors, and publishers. 2nd ed. Bethesda: National Library of Medicine. Wendling DL (technical editor), [updated 2015]. http://werken.ubiobio.cl/html/downloads/Vancouver/Bookshelf_NBK7256.pdf. Assessed 22 Dec 2019.

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

Chapter 13

Moisturizers and Emollients



Razvigor Darlenski and Joachim W. Fluhr

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 noninvasive 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, Trakia University, Stara Zagora, Bulgaria

J. W. Fluhr

Department of Dermatology and Allergology Charitéplatz 1, Charité Universitätsmedizin, Berlin, Germany

e-mail: joachim.fluhr@charite.de

© Springer Nature Switzerland AG 2020

J. W. Fluhr (ed.), *Practical Aspects of Cosmetic Testing*,
https://doi.org/10.1007/978-3-030-44967-4_13

127

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 [1]. 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 [2]:

- 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, noncomedogenic
- 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 multiple external stressors: physical stress (e.g., mechanical, thermal, radiation), chemical stress (tensides, solvents, topical xenobiotics), and environmental conditions. Furthermore, this barrier 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 bilamellar 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

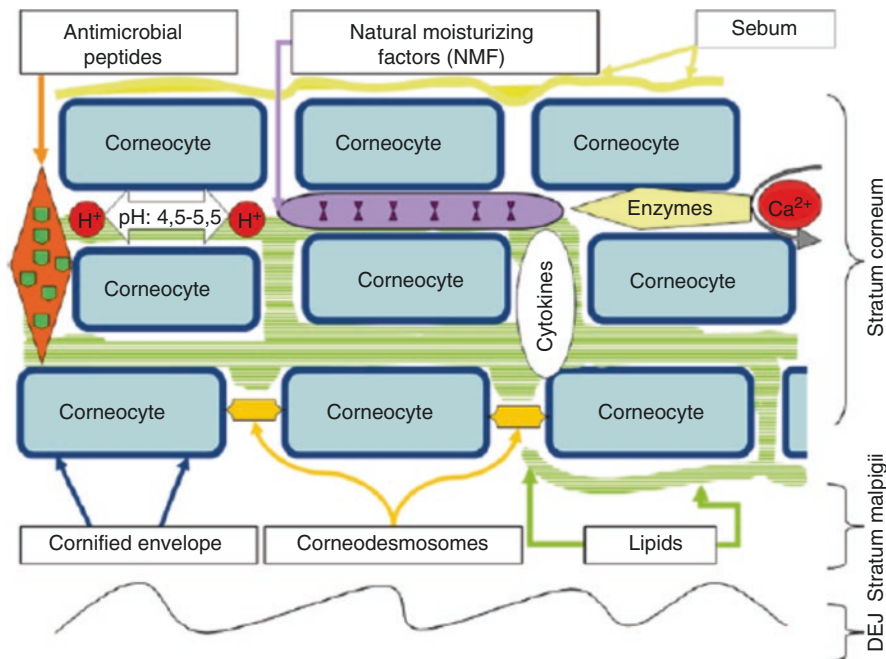


Fig. 13.1 Schematic overview on the complex structure of the stratum corneum (SC). *DEJ* Dermo-epidermal junction, Ca^{2+} Calcium ion, H^+ Hydrogen ion

SC integrity/cohesion and programmed renewal [3]. Schematic overview on the endogenous and exogenous mechanisms involved in the formation of the skin hydration is presented in Fig. 13.2.

Mammalian skin is exposed to a 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 [1]. 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 [4, 5]. This is followed by a steep rise to a constant level of approximately 70% in the viable epidermis. A typical curve on the water distribution in the epidermis assessed by in vivo confocal Raman spectroscopy, representative for healthy skin, is demonstrated in 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 [2]. This is a dynamic field for constant development not only for cosmetic ingredients but for trademarks and brands as well.

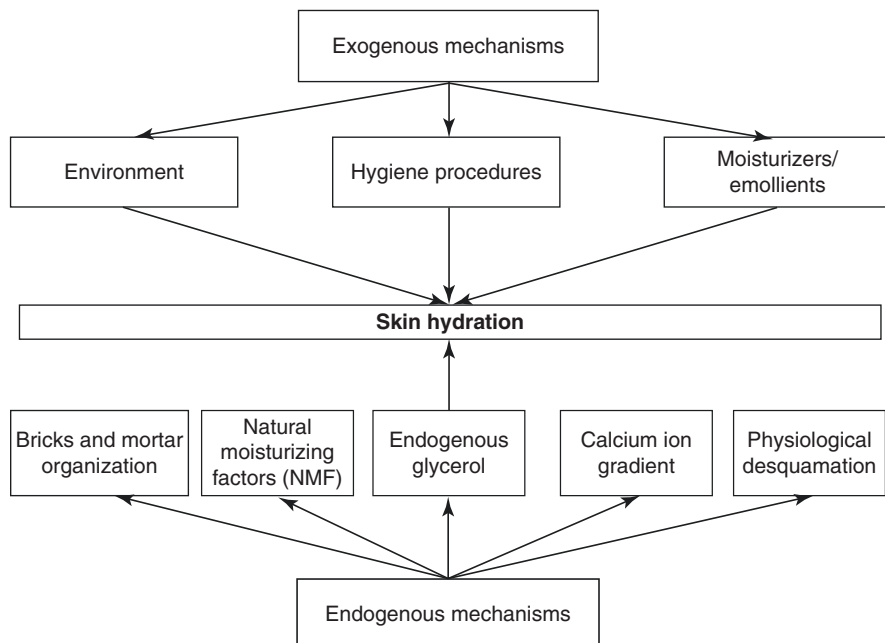


Fig. 13.2 Complex formation of the skin hydration

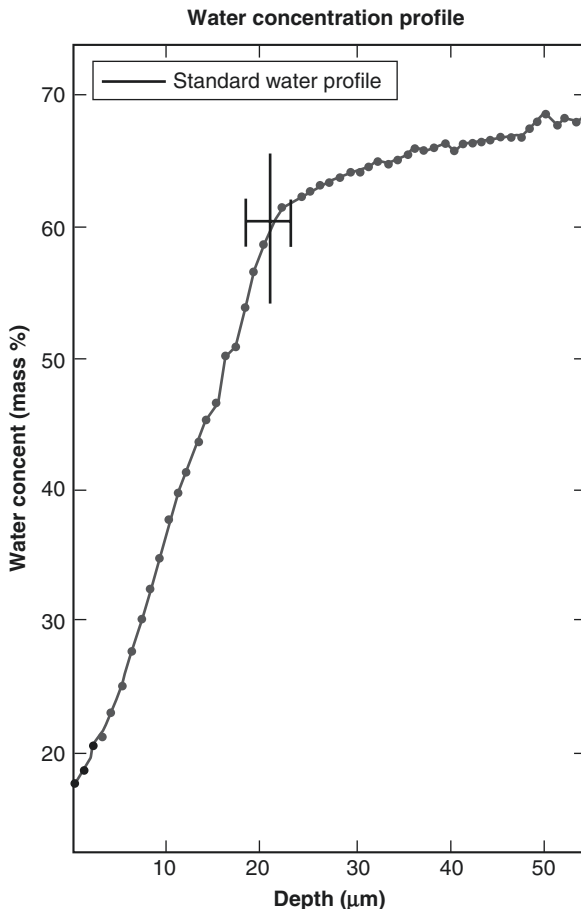
In general, moisturizers are registered as cosmetic products and/or over-the-counter (OTC) drugs as is the case with some products in the USA. The product category ranges from mass-market value brands to boutique and prestige luxurious products even claiming additional therapeutic effect. As for the 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 antimicrobial medications [6]. 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 [7].

13.2 Noninvasive 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 himself/herself), noninvasive biophysical measurements (such as the assessment of the electrical properties

Fig. 13.3 A typical curve representing the water profile of SC obtained by in vivo confocal Raman microspectroscopy (RCM). The line at approximately 20 μm represents the stratum corneum/stratum granulosum interface (\pm error bars). The present curve is representative for healthy skin on the volar forearm.



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 noninvasive, non-traumatizing, causing minimal discomfort and not altering the skin functions per se [3]. 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 and others) 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. In addition, the presence of hairs and cosmetic remnants on the skin surface influence the measurement. Furthermore, 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 extent 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 recommended.

The most commonly applied devices are based on measuring the electrical conductance, capacitance, resistance, or impedance as an indirect indicator for SC water content [8]. 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 [9].

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 [10]. 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 [11]. 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, 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 [12].

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 noninvasive method is advised as more reliable and accurate approach, if included in assessing SC hydration [13, 14].

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 noninvasive 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 accurate 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 tape stripping 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. 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 pretreated skin.

The 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 semiquantitative measurement of skin components (lipids, lactate, urea, urocanic acid, and others) and exogenously applied substances (dimethyl sulfoxide, trans-retinol, carotenoids) as a function of the depth of the epidermis [4].

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 [15]. 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 dependent 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 of skin physiology [3]. 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 [10].

Some devices need internal calibration that is valid only for the specified apparatus and this obstructs transferring data between instruments from 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 [16].

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 [17]. 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 [18]. 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-the-art review on standardized washing models used in skin physiology studies [19]. 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. 13.3.2.3) 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 [17, 20]. Although no explicit and unidirectional discrepancies

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

Variable	Parameter	
	SC hydration	TEWL
Age	+	+
Gender	–	–
Race/ethnicity	±	±
Pigmentation	+	+
Anatomical site	+	+
Skin temperature	+	+
Sweating	+	+
Air convection	+	+
Ambient temperature	+	+
Humidity	+	+
Direct light	+	n.d.
Season	+	+
Circadian rhythms	+	±

SC Stratum corneum

+ Influencing, – No influence, *n.d.* No data, ± Controversial data

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 preselect 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 [17, 21]. Higher hydration values are obtained from 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 contralateral 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 boarder 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 field 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.

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 preselected 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.

A linear relationship between skin hydration assessed by capacitance and the ambient air relative humidity exist [17]. Hence a control in this parameter (room air humidity) must be performed and should be within the range of $50 \pm 10\%$. 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 [22]. 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 [17]. 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, weighing 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, 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 ($t = 0$). 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, 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 noninvasive 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 necessary for performing a successful skin physiology study.

The recommendations for skin hydration measurements are summarized in Table 13.2.

Table 13.2 Recommendations for skin hydration measurements at a glance

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

References

1. Rawlings AV. Molecular basis for stratum corneum maturation and moisturization. *Br J Dermatol.* 2014;171(Suppl 3):19–28.
2. Lynde CW. Moisturizers: what they are and how they work. *Skin Therapy Lett.* 2001;6:3–5.
3. Darlenski R, Sassning S, Tsankov N, et al. Non-invasive in vivo methods for investigation of the skin barrier physical properties. *Eur J Pharm Biopharm.* 2009;72:295–303.
4. Darlenski R, Fluhr JW. In vivo Raman confocal spectroscopy in the investigation of the skin barrier. *Curr Probl Dermatol.* 2016;49:71–9.
5. Warner RR, Myers MC, Taylor DA. Electron probe analysis of human skin: determination of the water concentration profile. *J Invest Dermatol.* 1988;90:218–24.
6. Vogel CA, Balkrishnan R, Fleischer AB, et al. Over-the-counter topical skin products—a common component of skin disease management. *Cutis.* 2004;74:55–67.
7. Council Directive, 1993. In: Nr 93/35/CEE. EU.
8. Tagami H. Electrical measurement of the hydration state of the skin surface in vivo. *Br J Dermatol.* 2014;171(Suppl 3):29–33.
9. Obata M, Tagami H. Electrical determination of water content and concentration profile in a simulation model of in vivo stratum corneum. *J Invest Dermatol.* 1989;92:854–9.
10. Fluhr J, Gloor M, Lazzarini S, et al. Comparative study of five instruments measuring stratum corneum hydration (Corneometer CM 820 and CM 825, Skicon 200, Nova DPM 9003, DermaLab). Part II. In vivo. *Skin Res Technol.* 1999;5:171–8.
11. Uhoda E, Leveque JL, Pierard GE. Silicon image sensor technology for in vivo detection of surfactant-induced corneocyte swelling and drying. *Dermatology.* 2005;210:184–8.
12. Fluhr JW, Darlenski R, Angelova-Fischer I, et al. Skin irritation and sensitization: mechanisms and new approaches for risk assessment. 1. Skin irritation. *Skin Pharmacol Physiol.* 2008;21:124–35.
13. Snatchfold J, Targett D. Exploratory study to evaluate two clinical methods for assessing moisturizing effect on skin barrier repair. *Skin Res Technol.* 2019;25:251–7.
14. Voegeli R, Rawlings AV, Seroul P, et al. A novel continuous colour mapping approach for visualization of facial skin hydration and transepidermal water loss for four ethnic groups. *Int J Cosmet Sci.* 2015;37:595–605.
15. Dobrev H. Use of Cutometer to assess epidermal hydration. *Skin Res Technol.* 2000;6:239–44.
16. Heinrich U, Koop U, Leneveu-Duchemin MC, et al. Multicentre comparison of skin hydration in terms of physical-, physiological- and product-dependent parameters by the capacitive method (Corneometer CM 825). *Int J Cosmet Sci.* 2003;25:45–53.
17. Berardesca E, Loden M, Serup J, et al. The revised EEMCO guidance for the in vivo measurement of water in the skin. *Skin Res Technol.* 2018;24:351–8.
18. Kligman A. Regression method for assessing the efficacy of moisturizers. *Cosmet Toilet.* 1978;93:27–35.
19. Fluhr JW, Ennen J. Standardized washing models: facts and requirements. *Skin Res Technol.* 2004;10:141–3.
20. Darlenski R, Fluhr JW. Influence of skin type, race, sex, and anatomic location on epidermal barrier function. *Clin Dermatol.* 2012;30:269–73.
21. Kleesz P, Darlenski R, Fluhr JW. Full-body skin mapping for six biophysical parameters: baseline values at 16 anatomical sites in 125 human subjects. *Skin Pharmacol Physiol.* 2012;25:25–33.
22. Pinnagoda J, Tupker RA, Agner T, et al. Guidelines for transepidermal water loss (TEWL) measurement. A report from the Standardization Group of the European Society of Contact Dermatitis. *Contact Dermatitis.* 1990;22:164–78.

Chapter 14

Anti-aging and Anti-wrinkle Products



Razvigor Darlenski, Theresa Callaghan, and Joachim W. Fluhr

Abbreviations

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

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 anti-wrinkle 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.

R. Darlenski (✉)

Department of Dermatology and Venereology, Trakia University, Stara Zagora, Bulgaria

T. Callaghan

Callaghan Consulting International, Hamburg, Germany

J. W. Fluhr

Department of Dermatology, Charité Universitätsmedizin Berlin, Berlin, Germany

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 [1]. 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; chronological aging) but is also influenced by external factors such as the ultraviolet (UV) radiation, smoking, and lifestyle (extrinsic aging, photoaging). 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 socioeconomic 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 anti-wrinkle 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.

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) [2]. The reduction of the grade of visual skin aging can be used as a parameter for the assessment of anti-wrinkle 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.

Table 14.1 Clinical, histological, and functional characteristics of intrinsically and photoaged skin

	Intrinsic aging	Extrinsic aging
Clinical appearance	<p>Fine wrinkles Thin and transparent skin Loss of underlying fat with noticeable loss of firmness Skin sagging and slackness Xerosis Benign neoplasms, e.g., seborrheic keratoses and cherry angiomas</p>	<p>Deep wrinkles Irregular and uneven pigmentation (freckling, lentiginos, guttate hypomelanosis) Decreased elasticity and pronounced flaccidness Xerosis Vascular lesions (telangiectasia, purpura senilis, venous lakes) Neoplastic lesions, e.g., actinic keratoses, comedones (Favre-Racouchot disease), nonmelanoma skin cancer, lentigo maligna</p>
Histological features	<p><i>Epidermis</i> No alterations in SC Discrete/absent changes in epidermal thickness and keratinocyte shape Fewer melanocytes and Langerhans cells Flattened rete ridges at the dermo-epidermal junction</p>	<p><i>Epidermis</i> Increased compaction of SC Increased thickness of stratum granulosum Reduced epidermal thickness Reduced epidermal mucin content Increased number of melanosomes in the basal keratinocytes Irregularities in cell and nuclear shape and size Flattened rete ridges at the dermo-epidermal junction</p>
	<p><i>Dermis</i> Atrophy with alteration of the connective tissue Fewer fibroblasts and mast cells Decrease in the number of dermal blood vessels with shortening of the capillary loops Decrease in the nerve endings and the neuroreceptive apparatus of the skin</p>	<p><i>Dermis</i> Prominent grenz zone in the papillary Homogenization of the connective tissue Decreased and degraded collagen (collagens I, III, and VII) Deposition of abnormal elastotic material Increased matrix-degrading metalloproteinases Decreased amount of glucosaminoglycans Vascular dilatation and increased capillary tortuosity</p>
Functional abnormalities	<p>Competent basal epidermal barrier function Delayed recovery processes after acute barrier perturbation Decreased SC hydration Elevated skin surface pH (not valid for all anatomical sites) Reduced vasorelaxant response of the cutaneous microvessels Decrease in elasticity and extension but increased fatigability</p>	<p>Competent basal epidermal barrier function No difference in the barrier function in comparison to intrinsically aged skin Decreased SC hydration on the photoexposed vs. protected skin—impairment in SC water-binding capacity (compared to intrinsically aged skin) Elevated skin surface pH (not valid for all anatomical sites) Reduced luminescence of skin color (in comparison to photoprotected areas) Elevated values obtained by LDV, corresponding to an increase in the subpapillary vascular plexus Decrease in skin elasticity with predominance in the photodamaged areas</p>

SC stratum corneum

Table 14.2 Glogau score of photoaged skin [2]

Grade	Skin characteristics	Age group (years)
Mild	Few wrinkles, no keratosis	28–35
Moderate	Early wrinkling, sallow complexion with early actinic keratosis	35–50
Advanced	Persistent wrinkling, discoloration of the skin with telangiectases and actinic keratosis	50–60
Severe	Severe wrinkling, photoaging, gravitational and dynamic forces affecting the skin, actinic keratoses with or without skin cancer	65–70

14.2.2 *Noninvasive Biophysical Methods*

14.2.2.1 Instrumental Evaluation of the Skin Surface Topography

The examination of the skin surface topography with regard to roughness and 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 or cross-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 that can be used in revealing quantitatively the changes in the cutaneous microrelief [3].

Visioscan (Courage & Khazaka electronic, Cologne, Germany) is an UVA-light video camera with high resolution to study the skin surface directly. SELS® parameters (Surface Evaluation of the Living Skin) have been developed especially for this camera. The images show the structure of the skin and the level of dryness very impressively. The camera can be used on pigmented spots and lesions and on hair.

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 and DermaTOP [3–6]. The method is used to measure the depth and width of wrinkles.

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 [7]. 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.

Visioline® VL 650 (Courage & Khazaka electronic, Cologne, Germany) is used to assess skin replicas in order to visualize deeper lines and macro wrinkles such as “crow’s feet.” skin replica made of Silflo® (a white, non-glossy silicone material) illuminated in the Visioline® device by oblique lighting at an angle of 35° (placed in the unit with 90° of the main wrinkle direction toward the light source). Thus, the mountains in the replica representing the wrinkles of the skin produce measurable shadows.

Another method to evaluate the topography (micro relief) of the skin surface is Skin-Visiometer® SV 700 USB (Courage & Khazaka electronic, Cologne, Germany). It uses light transmission of a very thin, special blue dyed silicone replica.

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 [8]. 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 photodamaged compared to intrinsically aged skin [9].

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® (L'Oreal, France) 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 [4]. MoistureMap MM 100 is a device featuring a capacitance-based sensor comparable to SkinChip®. The sensor gives graphical information on the near-surface hydration distribution and the microtopography of skin. Rather than absolute moisture figures, the MoistureMap indicates the distribution of hydration on the skin surface.

Newly developed techniques such as in vivo confocal Raman microspectroscopy (RCM) offer 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 [10].

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 [8]. Ultimately, a system is needed that detects the evenness of pigmentation or the degree of spotted pigmentation on a defined UV-exposed area.

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 [11]. 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 semi-hydrophobic 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 [12]. 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 [13].

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].

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, uniaxial 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) [8, 14].

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) [5]. 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 [15]. A schematic overview of a study schedule is presented in Fig. 14.1.

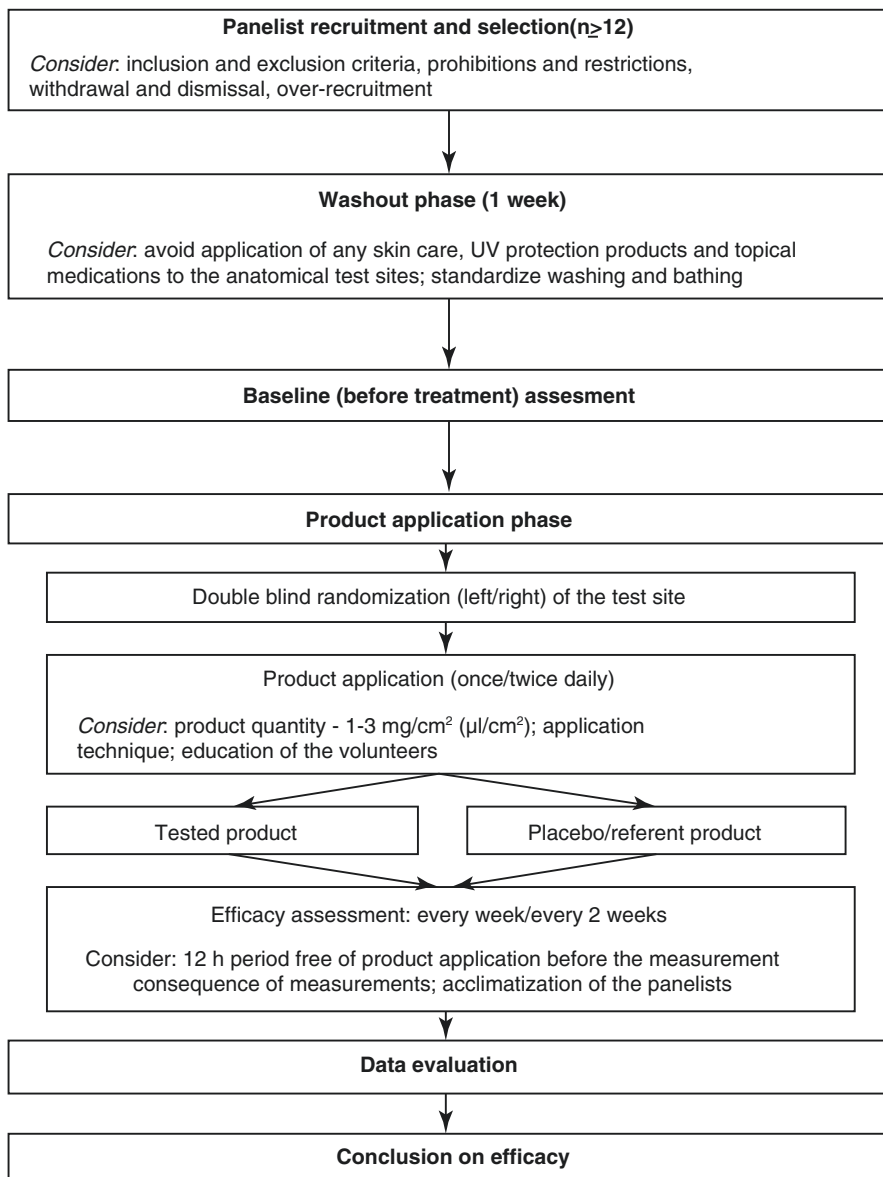


Fig. 14.1 A schematic overview of a typical study schedule for antiaging products

14.3.1 Study Population

Preselection and clear definition of the study population is required. Issues such as inclusion and exclusion criteria, prohibitions and restrictions, and 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). If a real reversal of aging-symptoms is to be revealed, a long study period should be selected.

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 anti-wrinkle 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 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 ($t = 0$).

14.3.5 Efficacy Assessment

Efficacy assessment is performed according to the chosen endpoint, i.e., clinical evaluation, visual methods, and 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

1. Krämer U, Schikowski T. Recent demographic changes and consequences for dermatology. In: Gilchrest BA, Krutmann J, editors. *Skin aging*. Berlin: Springer; 2006.
2. Glogau RG. Aesthetic and anatomic analysis of the aging skin. *Semin Cutan Med Surg*. 1996;15:134–8.
3. Callaghan TM, Wilhelm KP. A review of ageing and an examination of clinical methods in the assessment of ageing skin. Part 2: clinical perspectives and clinical methods in the evaluation of ageing skin. *Int J Cosmet Sci*. 2008;30:323–32.
4. Berardesca E, Loden M, Serup J, et al. The revised EEMCO guidance for the in vivo measurement of water in the skin. *Skin Res Technol*. 2018;24:351–8.
5. Bonte F, Girard D, Archambault JC, et al. Skin changes during ageing. *Subcell Biochem*. 2019;91:249–80.
6. Darlenski R, Sassning S, Tsankov N, et al. Non-invasive in vivo methods for investigation of the skin barrier physical properties. *Eur J Pharm Biopharm*. 2009;72:295–303.
7. Akazaki S, Nakagawa H, Kazama H, et al. Age-related changes in skin wrinkles assessed by a novel three-dimensional morphometric analysis. *Br J Dermatol*. 2002;147:689–95.
8. Bielfeldt S, Springmann G, Seise M, et al. An updated review of clinical methods in the assessment of ageing skin—New perspectives and evaluation for claims support. *Int J Cosmet Sci*. 2018;40:348–55.
9. Kikuchi-Numagami K, Suetake T, Yanai M, et al. Functional and morphological studies of photodamaged skin on the hands of middle-aged Japanese golfers. *Eur J Dermatol*. 2000;10:277–81.

10. Darlenski R, Fluhr JW. In vivo Raman confocal spectroscopy in the investigation of the skin barrier. *Curr Probl Dermatol*. 2016;49:71–9.
11. Zouboulis CC, Elewa R, Ottaviani M, et al. Age influences the skin reaction pattern to mechanical stress and its repair level through skin care products. *Mech Ageing Dev*. 2018;170:98–105.
12. Li L, Mac-Mary S, Marsaut D, et al. Age-related changes in skin topography and microcirculation. *Arch Dermatol Res*. 2006;297:412–6.
13. Andersson SE, Edvinsson ML, Edvinsson L. Cutaneous vascular reactivity is reduced in aging and in heart failure: association with inflammation. *Clin Sci (Lond)*. 2003;105:699–707.
14. Hameed A, Akhtar N, Khan HMS, et al. Skin sebum and skin elasticity: major influencing factors for facial pores. *J Cosmet Dermatol*. 2019;18:1968–74.
15. Salter D. Non-invasive cosmetic efficacy testing in human volunteers: some general principles. *Skin Res Technol*. 1996;2:59–63.

Chapter 15

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 dilated pores and has a greater tendency to develop comedones, pimples, and pustules [1, 2]. Most people feel the skin unpleasant and perceive it as a serious cosmetic problem. That is why the control over the impure, acne-like skin is very important.

Currently there are many topical formulations for cleansing and care of the skin that claimed to normalize sebum production, remove excess oil, reduce pore size, and clean skin imperfections. According to Article 7a.1 of the EU Consolidated Cosmetic Directive, the manufacturer or his agent should hold information about “proof of the effect claimed for the cosmetic product, *where justified by the nature of the effect or product*” [3]. Testing on human volunteers, conducted in accordance with the rules of good clinical practice (GCP), is the preferred way to obtain proof of efficacy [4].

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

H. Dobrev (✉)

Department of Dermatology and Venereology, Medical University, Plovdiv, Bulgaria

15.2 Test Settings

15.2.1 Study Background

The cosmetic testing should be preceded by detailed review of the problem that is the basis of the research. The investigator must examine throughout all available previous data and clarify the study nature and rationale. Reference to literature is also helpful in discussing the findings of the test [5].

15.2.2 Study Objectives

Generally, the study objectives (study end points) are identical with the efficacy claims of the product tested [5]. 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. Anti-acne, anti-comedonal, 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 [6].

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), non-active 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) [5].

Investigational products should be manufactured, handled, and stored according to the good manufacturing practice of cosmetic products (GCPC) and used according to the approved protocol [5, 7]. The investigator should provide a brief description of the test products including 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 [8].

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 [5, 9]:

- 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, 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 non-invasive instrumental measurements.
- Criteria for exclusion. They may include the presence of severe acne, rosacea, pregnancy, breastfeeding, known allergy to constituents of the test products, systemic disorders.
- Criteria for early withdrawal. If any predetermined reasons for removing subjects from study exist, they should be defined.
- Pre-experimental 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 [1, 5, 6, 8]:

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

The open-labeled study without placebo control and comparison before/after treatment applied is easiest to perform but is considered that its design is not strict enough. The randomized controlled trial is the gold-standard design for cosmetic products testing and should be used whenever applicable.

Other aspect of cosmetic testing design is the study duration [9]. 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 for 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 the 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 [6].

Cosmetic testing may be performed at single (single center study) or several (multicenter 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 [5, 8]:

- **Sensorial tests.** They are based on appreciation of skin appearance and product performance made through the senses (sight, touch, and olfaction) of consumers 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, accord-

ing 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 that the test and parameters measured are 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, morphological and functional, direct and indirect, and convenient and cumbersome methods [10–13].

Sensorial Assessment

Sensorial assessment of the oiliness of the skin of the forehead, nose, chin, and cheeks is based on visual and tactile appreciation of how shiny the skin appears and also on 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 [8, 13].

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 4-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 [8].

Investigator assessment gives also information about the initial degree of skin oiliness and the product efficacy. It should be carried out at least two times—before and after the study completion. Generally, a 5-point scale (normal skin, slightly, easily visible, obviously, and extensive oiliness) is used and the changes in mean scores are determined. Product efficacy can be rated by both subjects and investigator on a 4-point scale (0 = no change to 3 very good effect or marked improvement) [6, 14].

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 mostly used test site. The usually determined variables are [10–13, 15, 16]:

- **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 $\mu\text{g}/\text{cm}^2/\text{h}$. 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, i.e., 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 [10–13]:

- **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 collected for 30 s lipids is displayed as $\mu\text{g}/\text{cm}^2$. 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 for sebum excretion from the infundibular reservoir, i.e., the tapes only absorb 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 defatted (Sebutape adhesive patches, CuDerm, Dallas, USA; www.cuderm.com) or not defatted (Sebifix F16, Courage+Khazaka, Cologne, Germany; www.courage-khazaka.com) prior to the tape application and the collection time may be between 30 s (Sebifix 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 Sebifix tapes.
- **Skin glossymetry.** Nowadays it becomes possible to measure skin gloss with noninvasive and convenient bioengineering devices. The first one (SkinGlossmeter 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.
- **Ultraviolet red fluorescence.** It can be used for indirect assessment of skin oiliness. Studies have confirmed the positive correlation between casual sebum level and red fluorescence in subjects with normal skin and patients with oily, acne-prone skin [17–19].

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 [20]. 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 [21]. The last authors have additionally developed a 26-item questionnaire for measurement 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 [6, 22]:

- Non-inflammatory 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 [23].

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 non-facial 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 [24].

Sensorial Assessment

Sensorial assessment of acne lesions is based on visual and tactile appreciation and can be performed in the form of auto evaluation (self-assessment) and investigator (physician) evaluation using clinical examination [8].

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: 4-point scale (none, mild, moderate, and severe), 5-point scale (none, minimal, mild, moderate, and severe), or visual analog scale of 0 to 10 (0—no acne; 10—very severe acne) [24].

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 auto-evaluation 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 [8].

Investigator (physician) assessment gives also information about the initial degree of acne lesions (baseline acne severity) and the efficacy of the product applied (change from baseline in acne severity). It should be carried out at least two 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 the acne is typically based on a qualitative global assessment of severity (visually or using photographs), preferred for clinical practice, and lesion counting, preferred for clinical investigations [6, 14, 22, 25–37]. 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 [38]:

- 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

Table 15.1 Disease-specific instruments for assessment of acne

<i>Qualitative methods (global assessment)</i>	
Visually scoring	The facial acne scale (Allen and Smith 1982)
	Global severity scoring of the AAD (Pochi et al. 1991)
	Classification system for acne (Harrison-Atlas et al. 1996)
	Quantitative assessment over the whole acne area (Lucky and Beth 1996)
	ECLA (Echelle de Cotation des Lésions d'Acne or Acne lesion score scale) (Dreno et al. 1997)
	Global acne grading scale (GAGS) (Doshi et al. 1997)
	Combined acne severity classification (Lehmann et al. 2002)
	Investigator global assessment (IGA) (Dermatologic and ophthalmic drugs advisory committee briefing, FDA, November 2002)
	Simple grading of acne (Consensus document of the Global Alliance to improve outcomes in acne, 2005)
	Comprehensive acne severity scale (CASS) (Tan et al. 2007)
	ECCA grading scale (Dreno et al. 2007)—for acne scars
	Scale for Acne Scar Severity (SCAR-S) (Tan et al. 2010)—for acne scars
	Postacne hyperpigmentation index (PAHPI) (Savory et al. 2014)—for postinflammatory hyperpigmentation
	Artificial intelligence for the objective evaluation of acne Investigator global assessment (Melina et al. 2018)
Scoring using photographs	Acne grading method using photographs (Cook 1979)
	Leeds acne grading scale (Burke and Cunliffe 1984)
	The revised Leeds acne grading scale (O'Brien et al. 1998)
	New photographic techniques for clinical evaluation of acne (Rizova and Kligman 2001)
	Global Acne Severity Scale (GEA Scale) (Dreno and GEA Group 2011)
<i>Quantitative methods (lesion count)</i>	
	Half-face grading (Plewig and Kligman 1974)
	Michaelson acne severity score (Michaelsson et al. 1977)

The methods for acne grading that are used most often are the Leeds acne grading scale and the combined acne severity classification.

The Leeds Acne Grading Scale [28, 31, 32] 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, the 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 vary 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 [35, 39] 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 5 cysts, or comedone count greater than 100, or a total inflammatory count greater than 50, or a total lesion count greater than 125.

The investigator evaluation of product efficacy is based on the changes in the global severity score and reduction in the acne lesions count (separately in the count of inflammatory and non-inflammatory lesions as well as in total lesion count) toward the baseline. The subject response to treatment can be rated on a 4-point scale (0 = no to 3 very good effect) or on a 5-point scale (0 = worse to 4 = marked improvement) [23].

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 the baseline [40].

Recently, the FDA agency recommended an **Investigator Global Assessment (IGA)** ordinal scale with 5 severity grades (0 to 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 the 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 [6].

Table 15.2 Investigator global assessment scale for acne vulgaris

Grade	Description
0	Clear skin with no inflammatory or noninflammatory lesions
1	Almost clear; rare noninflammatory lesions with no more than one small inflammatory lesion
2	Mild severity; greater than Grade 1; some noninflammatory lesions with no more than a few inflammatory lesions (papules/pustules only, no nodular lesions)
3	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
4	Severe; greater than Grade 3; up to many noninflammatory and inflammatory lesions, but no more than a few nodular lesions

The post effects of acne can also be evaluated with specially created instruments such as the postacne hyperpigmentation index (PAHPI) for measurement of postinflammatory hyperpigmentation, ECCA grading scale, Scale for Acne Scar Severity (SCAR-S), and Self-assessment of clinical acne-related scars (SCARS) for assessment of acne scarring [41–45].

It is considered that the evaluation of acne using ordinal scales reflects the clinical perception of severity but has shown low reproducibility both intra- and inter-rater. That is why the latest development in the assessment of acne severity is the application of Artificial Intelligence on images to automatically measure acne severity in the 0–4 numerical range of the Investigator Global Assessment (IGA). Results obtained showed that the Artificial Intelligence was able to directly classify acne patients according to an IGA ordinal scale with high accuracy, no human intervention, and no need to count lesions [46].

Instrumental Assessment

Besides the measurement of sebum excretion, acne-prone skin can be evaluated with other methods such as polarized light photography [14, 47], digital fluorescence photography [48–50], videomicroscopy [25, 51], and ultraviolet light illumination [52]. Multispectral imaging, multimodal digital imaging, 3-D imaging, and computer algorithms for auto-classification also are used to analyze primary and secondary acne lesions [14, 26].

Polarized light photography enhances the visualization of inflammatory acne lesions, helps distinguish them from non-inflammatory lesions, and allows accurate assessment of the extent of disease and the effectiveness of therapy.

For acne patients, the orange-red fluorescence of facial ultraviolet photographs is 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 [48]. The orange-red fluorescence showed stronger correlation with the presence of non-inflammatory acne lesions (comedones) and high sebum amount than the presence of inflammatory acne lesions (pustules) and low sebum amount.

Recently, using a UVA-induced digital facial fluorescent imaging system and image analysis method, Youn et al. [50] 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. [52] have used a videocamera equipped with an internal ultraviolet illumination source (Visioscan, C+K Electronic, Cologne, Germany) and image analysis software for determination of follicular plugs and product comedolytic activity.

Recently, a new camera for measurement of fluorescence and acne inflammation (VISIA-CR, Canfield Scientific, Parsippany, NJ, USA) was introduced. Specific spectral filters in front of the camera are used to separate the Protoporphyrin IX (PpIX) and Coproporphyrin III (CpIII) fluorescence signals. It is considered that *P. acnes* produced CpIII with greenish-orange (570–630 nm) fluorescence. In addition, RBX transformation is performed in order to separate the hemoglobin (Red) and melanin (Brown) absorption from the cross-polarized images. Thus, the inflamed acne lesions are detected and measured from the RBX-Red images. Results obtained demonstrate that CpIII fluorescence spots significantly correlate with the investigators' comedonal lesion counts while the detected acne lesion-specific inflammation significantly correlate with the investigators' papulopustular lesion counts [53].

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 [35, 54].

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), 4-item condensed version of the Acne-QoL (Acne-Q4) [55], 9-item Acne QoL scale (AQLS) [56], Assessment of the Psychological and Social Effects of Acne (APSEA) [35], and Facial acne scar quality of life (FASQoL) [43].

In recent years, new tools have been developed to assess the impact of acne on patients' quality of life. These include a patient-reported outcome measure for facial acne: Acne Symptom and Impact Scale (ASIS) and the Comprehensive Quality-of-Life Measure for Facial and Torso Acne (CompAQ) [57, 58].

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 in 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 [54].

Product Acceptance Assessment

Product acceptance is related to physical properties (consistency, color, fragrance) and cosmetic properties (spreadability, 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 4-point scale (0 = none to 3 = very good) or on a 5-point scale (0 = none to 4 = excellent).

To measure subject's satisfaction with the cosmetic product properties a 5-point (very satisfied, satisfied, neither satisfied nor dissatisfied, dissatisfied, very dissatisfied) or a 6-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 product overall approval generates the subject's willingness to purchase or not the product tested.

Recently, some new techniques were introduced 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 [59]. 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 [9].

Features of local irritation (erythema, scaling, itching, burning) are evaluated visually and through 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 the investigator using a 4-point scale (absent, mild, moderate, severe).

15.2.3.5 Study Protocol

The study protocol should describe the following points [8, 9]:

- Test schedule (timetable), i.e., 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 mg/cm² skin)
 - Frequency of use (usually one to two application a day)
 - Time of products' application (usually morning and/or evening)
 - Application areas (measuring test sites). For identification and marking of the application sites, a template is preferably used.
 - Product application manner
- 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 were 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 good clinical practice provides public assurance that the rights, safety, and well-being of trial subjects are protected and that the clinical trial data are credible [5].

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 [5, 6, 8, 9].

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 abilities 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.

Key Messages

- Testing of cosmetic products for impure skin is ruled by the principles of good clinical practice.
- 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.
- Sensorial self- and expert evaluation and instrumental skin bioengineering techniques are the preferred way to proof product claims.
- Effective treatment is associated with significant improvement in the quality of life.

References

1. Fluhr J, Reinhardt H, Rippke F. Guideline of the Society for Dermopharmacy. Dermocosmetics for the cleansing and care of skin prone to acne. Dermotopics issue 1. 2005. http://www.dermotopics.de/english/issue_1_05_e/leitniedermokosmetika_1_05_e.htm. Accessed 21 Dec 2019.
2. Mehling A, Buchwald-Werner S. Natural actives for impure skin. SOFW J. 2004;130:2–6.
3. Council Directive 76/768/EEC on the approximation of the laws of the member states relating to cosmetic products (consolidated version 2008). <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1976L0768:20080424:EN:PDF>. Accessed 21 Dec 2019.
4. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Guideline for Good Clinical Practice (E6R1). 2002. <https://www.imim.cat/media/upload/arxius/emea.pdf>. Accessed 21 Dec 2019.
5. Serup J. Efficacy testing of cosmetic products. A proposal to the European Community by the Danish Environmental Protection Agency, Ministry of Environment and Energy. Skin Res Technol. 2001;7(3):141–51.
6. FDA Guidance for Industry—acne vulgaris: developing drugs for treatment. 2005. <https://www.federalregister.gov/documents/2005/09/19/05-18512/draft-guidance-for-industry-on-acne-vulgaris-developing-drugs-for-treatment-availability>. Accessed 21 Dec 2019.
7. Council of Europe. Guidelines for Good Manufacturing Practice of Cosmetic Products (GMPC). 1995. <https://book.coe.int/en/health-protection-of-the-consumer/32-pdf-guidelines-for-good-manufacturing-practice-of-cosmetic-products-gmpc.html>. Accessed 21 Dec 2019.

8. COLIPA. Colipa guidelines. Guidelines for the evaluation of the efficacy of cosmetic products. 2008. <http://www.canippec.org.mx/woo/xtras/bibliotecavirtual/Colipa.pdf>. Accessed 21 Dec 2019.
9. De Paep K. Evaluation of the efficacy of dermato-cosmetic products. 2001. (Doctoral Thesis. Vrije University, Brussel).
10. Agache P. Sebaceous function assessment. In: Agache P, Humbert P, Maibach HI, editors. *Measuring the skin*. Basel: Springer; 2004. p. 281–9.
11. el-Gammal C, el-Gammal S. Sebum-absorbent tape and image analysis. In: Serup J, Jemec GBE, editors. *Handbook of non-invasive methods and the skin*. Boca Raton: CRC Press; 1995. p. 517–22.
12. Elsner P. Sebum. In: Berardesca E, Elsner P, Wilhelm K-P, Maibach H, editors. *Bioengineering of the skin: methods and instrumentation*. Boca Raton: CRC Press; 1995. p. 81–9.
13. Pierard GE, Pierard-Franchimont C, Marks R, Pay M, Rogiers V, the EEMCO Group. EEMCO guidance for the in vivo assessment of skin greasiness. *Skin Pharmacol Appl Ski Physiol*. 2000;13:372–89.
14. Becker M, Wild T, Zouboulis C. Objective assessment of acne. *Clin Dermatol*. 2017;35:147–55.
15. Dobrev H. Clinical and instrumental study of the efficacy of a new sebum control cream. *J Cosmet Dermatol*. 2007;6(2):113–8.
16. Dobrev H. Clinical and instrumental study of the efficacy of a new multiaction topical product in acneic skin. *Household and Personal Care Today*. 2009;1:16–22.
17. Choi C, Choi J, Park K, Youn S. Ultraviolet-induced red fluorescence of patients with acne reflects regional casual sebum level and acne lesion distribution: qualitative and quantitative analyses of facial fluorescence. *Br J Dermatol*. 2012;166(1):59–66.
18. Dobrev H. Fluorescence diagnostic imaging in patients with acne. *Photodermatol Photoimmunol Photomed*. 2010;26:285–9.
19. Songsantiphap C, Asawanonda P. The correlations between follicular fluorescence and casual sebum levels in subjects with normal skin. *J Clin Aesthet Dermatol*. 2019;12(8):24–7.
20. Segot-Chicq E, Compan-Zaouati D, Wolkenstein P, et al. Development and validation of a questionnaire to evaluate how a cosmetic product for oily skin is able to improve well-being in women. *J Eur Acad Dermatol Venereol*. 2007;21(9):1181–6.
21. Arbuckle R, Atkinson M, Clark M, et al. Patient experiences with oily skin: the qualitative development of content for two new patient reported outcome questionnaires. *Health Qual Life Outcomes*. 2008;6:80.
22. Agache P. Scoring acne vulgaris. In: Agache P, Humbert P, Maibach HI, editors. *Measuring the skin*. Basel: Springer; 2004. p. 691–7.
23. Food and Drug Administration Center For Drug Evaluation and Research Meeting of the Dermatologic and Ophthalmic Drugs Advisory Committee 8:10 a.m. Tuesday, November 5, 2002 Versailles Ballroom Holiday Inn—Bethesda 8120 Wisconsin Avenue Bethesda, Maryland.
24. Ozolins M, Eady EA, Avery A, et al. Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne. *Health Technol Assess*. 2005;9(1):iii-212.
25. Adityan B, Kumari R, Thappa D. Scoring systems in acne vulgaris. *Indian J Dermatol Venereol Leprol*. 2009;75(3):323–6.
26. Agnew T, Furber G, Leach M, Segal L. A comprehensive critique and review of published measures of acne severity. *J Clin Aesthet Dermatol*. 2016;9(7):40–52.
27. Bhor U, Pande S. Scoring systems in dermatology. *Indian J Dermatol Vereol Leprol*. 2006;72(4):315–21.
28. Burke B, Cunliffe W. The assessment of acne vulgaris: the Leeds technique. *Br J Dermatol*. 1984;111(1):83–92.
29. Dreno B, Bodokh I, Chivot M, et al. ECLA grading: a system of acne classification for every day dermatological practice. *Ann Dermatol Venereol*. 1999;126(2):136–41.

30. Dréno B, Poli F, Pawin H, et al. Development and evaluation of a global acne severity scale (GEA scale) suitable for France and Europe. *J Eur Acad Dermatol Venereol*. 2011;25(1):43–8.
31. Holm E, Jemec GBE. General guidelines for assessment of skin diseases. In: Serup J, Jemec GBE, editors. *Handbook of non-invasive methods and the skin*. 2nd ed. London: Informa Healthcare; 2006. p. 931–44.
32. O'Brien S, Lewis J, Cunliffe W. The Leeds revised acne grading system. *J Dermatolog Treat*. 1998;9:215–20.
33. Pochi PE, Shalita AR, Strauss JS, Webster SB, Cunliffe WJ, Katz HI, et al. Report of the consensus conference on acne classification. Washington, D.C., March 24 and 25, 1990. *J Am Acad Dermatol*. 1991;24:495–500.
34. Richard Thomas D, Sapra S, Lynde C, Poulin Y, Gulliver W, Sebaldt R. Development and validation of a comprehensive acne severity scale. *J Cutan Med Surg*. 2007;11(6):211–6.
35. Tan J. Current measures for the evaluation of acne severity. *Expert Rev Dermatol*. 2008;3(5):595–603.
36. Tan J, Jones E, Allen E, et al. Evaluation of essential clinical components and features of current acne global grading scales. *J Am Acad Dermatol*. 2013;69:754–61.
37. Witkowski J, Parish L. The assessment of acne: an evaluation of grading and lesion counting in the measurement of acne. *Clin Dermatol*. 2004;22(5):394–7.
38. Sinclair W, Jordaan HF. Acne guideline 2005 update. Compiled by Werner Sinclair and H Francois Jordaan and largely based on a consensus document of the Global Alliance to Improve Outcomes in Acne. *S Afr Med J*. 2005;95:883–92.
39. Lehmann H, Robinson K, Andrews J. Acne therapy: a methodological review. *J Am Acad Dermatol*. 2002;47:231–40.
40. Heffernan M, Nelson M, Anadkat M. A pilot study of the safety and efficacy of picolinic acid gel in the treatment of acne vulgaris. *Br J Dermatol*. 2007;156(3):548–52.
41. Clark A, Saric S, Sivamani R. Acne scars: how do we grade them? *Am J Clin Dermatol*. 2018;19(2):139–44.
42. Dreno B, Khammari A, Orain N, et al. ECCA grading scale: an original validated acne scar grading scale for clinical practice in dermatology. *Dermatology*. 2007;214(1):46–51.
43. Layton A, Dreno B, Finlay A, et al. New patient-oriented tools for assessing atrophic acne scarring. *Dermatol Ther (Heidelb)*. 2016;6:219–33.
44. Savory S, Agim N, Mao R, et al. Reliability assessment and validation of the postacne hyperpigmentation index (PAHPI), a new instrument to measure postinflammatory hyperpigmentation from acne vulgaris. *J Am Acad Dermatol*. 2014;70:108–14.
45. Tan J, Tang J, Fung K, et al. Development and validation of a Scale for Acne Scar Severity (SCAR-S) of the face and trunk. *J Cutan Med Surg*. 2010;14(4):156–60.
46. Melina A, Dinh N, Tafuri B, et al. Artificial intelligence for the objective evaluation of acne investigator global assessment. *J Drugs Dermatol*. 2018;17(9):1006–9.
47. Phillips S, Kollias N, Gillies R, Muccini JA, Drake LA. Polarized light photography enhances visualization of inflammatory lesions of acne vulgaris. *J Am Acad Dermatol*. 1997;37(6):948–52.
48. Pagnoni A, Kligman AM, Kollias N, Goldberg S, Stouemayer T. Digital fluorescence photography can assess the suppressive effect of benzoyl peroxide on *Propionibacterium acnes*. *J Am Acad Dermatol*. 1999;41(5 Pt 1):710–6.
49. Son T, Han B, Jung B, Stuart NJ. Fluorescent image analysis for evaluating the condition of facial sebaceous follicles. *Skin Res Technol*. 2008;14(2):201–7.
50. Youn S, Kim J, Lee J, et al. The facial red fluorescence of ultraviolet photography: is this color due to *Propionibacterium acnes* or the unknown content of secreted sebum? *Skin Res Technol*. 2009;15(2):230–6.
51. Rizova E, Kligman A. New photographic techniques for clinical evaluation of acne. *J Eur Acad Dermatol Venereol*. 2001;15(Suppl 3):13–8.
52. Uhoda E, Pierard-Franchimont C, Pierard G. Comedolysis by a lipohydroxyacid formulation in acne-prone subjects. *EJD*. 2003;13(1):65–8.

53. Patwardhan S, Richter C, Vogt A, Blume-Peytavi U, Canfield D, Kottner J. Measuring acne using Coproporphyrin III, Protoporphyrin IX, and lesion-specific inflammation: an exploratory study. *Arch Dermatol Res.* 2017;309:159–67.
54. Zip C. The impact of acne on quality of life. *Skin Therapy Lett.* 2007;12(10):7–9.
55. Tan J, Fung K, Khan S. Condensation and validation of a 4-item index of the acne-QoL. *Qual Life Res.* 2006;15(7):1203–10.
56. Tan J, Fung K, Gupta A, et al. Development and validation of a comprehensive acne severity scale. *J Cutan Med Surg.* 2007;11(6):211–6.
57. Alexis A, Daniels S, Johnson N, et al. Development of a new patient-reported outcome measure for facial acne: the acne symptom and impact scale (ASIS). *J Drugs Dermatol.* 2014;13(3):333–40.
58. McLellan C, Frey M, Thiboutot D, et al. Development of a comprehensive quality-of-life measure for facial and torso acne. *J Cutan Med Surg.* 2018;22(3):304–11.
59. Eisfeld W, Schaefer F, Boucsein W, Stolz C. Tracking intersensory properties of cosmetic products via psycho-physiological assessment. *Int J Cosmet Sci.* 2005;27(5):292–292.

Chapter 16

Assessment of Hair Morphology



C. Thieulin, R. Vargiolu, and Hassan Zahouani

Core Messages

- A new method using wavelet decomposition of a surface is presented. The advantage is that it allows evaluating damage to hair at different scales.
- A vibration analysis combined with friction test is also provided. It provides information on the sensorial properties of the hair surface.

16.1 Introduction

Everybody wants beautiful hair that is pleasant to touch. Indeed, hair has great social significance for [human beings](#) as it reflects our physical state. Thus, for many years, especially in the second half of the twentieth century, scientists have focused on the physical and chemical properties of hair and developed methods of evaluation based on scientific values.

Mostly, these methods are focused on the cuticle characterization. This outer part consists of flat overlapping cells (scales) of 50 μm in length that 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 sulfur content. This part, about 90% of a hair, gives it the longitudinal mechanical properties.

Mechanical processes such as combing, smoothing, and blow-drying serve to style the hair. Chemical products and processes such as chemical dye, colorant, and bleach treatments improve the appearance and hue of the hair. However, while giving a desirable look or style to hair, these treatments can also cause significant damage to the hair fibers, especially to the cuticle. In general, hair fiber damage occurs

C. Thieulin · R. Vargiolu · H. Zahouani (✉)

Laboratory of Tribology and Dynamic System, Ecole Centrale de Lyon, Ecully, France

e-mail: c.thieulin@ece.fr; hassan.zahouani@ec-lyon.fr

due to mechanical or chemical solicitations, or a combination of both (chemo-mechanical). In particular, the mechanical and chemical solicitations occurring during daily washing, combing, drying, smoothing, coloring, or bleaching induce cracks and breaks in the cuticle [1]. This attenuates the mechanical properties of the cortex [2]. The attenuation of its mechanical properties considerably affects the tribological (adhesion, friction, and wear) and morphological properties of hair (aspect of scales, surface roughness) [3]. This work will especially focus on the tribological and morphological properties of hair.

The measurement of the hair morphology needs specific instruments to evaluate hair damage at different scales. The scanning electron microscope and the atomic force microscope are commonly used for this application. A third technique that can be applied for hair samples with comparison before and after treatment is optical interferometry. This noncontact method permits to quickly determine a three-dimensional surface shape. To study the tribological behavior of hair, scientists developed macroscale and micro/nanoscale experimental methods. Macroscale characterization was widely used by the hair care industry, which investigated the friction between a skin replica and a hair swatch [4]. More recently, certain tribological works focused on the frictional behavior of hair using atomic force microscopy (AFM) and friction force microscopy. In particular, AFM studies were carried out to understand damage and the effects of hair care products on hair topography [5, 6] and tribology [5, 7–9].

In this paper, we propose to explore the tribological and morphological hair changes due to chemical and thermomechanical treatments. Optical interferometry combined with wavelet decomposition is used to assess hair morphology. Then, friction measurements are performed with a haptic tribometer. The haptic tribometer allows measuring the acoustic vibratory level generated during the friction test.

16.2 Material and Methods

16.2.1 Hair Samples

Caucasian hair samples were used in this study. We studied the effect of chemical and thermomechanical treatments on the sensorial properties of hair. To investigate the effect of chemical treatments, hair samples were selected from people who had treated their hair by bleaching or coloring.

Then, the effect of smoothing was studied on samples of hair previously damaged. Two smoothing temperatures were chosen: 140 °C (T140) and 200 °C (T200). To obtain damaged hair samples, we applied 25 cycles of dipping in distilled water (10s at 25 °C) and drying (10s) using a hair dryer at 78 °C [10].

Samples of 2 cm length were cut at a distance of around 5 cm away from the root area to be tested. All the measurements were performed in this region. Before any measurement, hair samples were stored in a room at a temperature of 22 °C and 40% RH.

16.2.2 *Optical Interferometry*

The interferometer (Bruker Contour GT-I) is an optical device designed to measure the topography of surfaces. A white light source is split into two separate beams of which one is reflected on the surface studied before touching the charge coupled device (CCD) detector and the other is reflected from the reference mirror to the CCD detector. When recombined, the two beams create a series of white and dark bands called fringes that make up an interference pattern. These fringes can be associated with, and translated into, lines on a topographic map. The microscope is connected to a piezo-electric transducer to increase the measurement amplitude, with displacements controlled by a micro-calculator. The data collected by this method consists of point coordinates in the x , y , and z directions, which, when reconstructed, create a three-dimensional representation of the surface.

Vertical scanning interferometry (VSI) mode is used when imaging hair. A magnification of 115X was used to obtain hair images. With a vertical resolution of 3 nm, VSI mode is particularly suitable for rough surfaces with amplitudes ranging between 160 nm and 2 mm.

The image obtained is composed of different wavelength scales; high frequencies for roughness, medium for waviness, and low for form. To study specifically, the roughness modification, a method based on multi-scale decomposition has been developed [2, 11]. 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, called SMA. The SMA or roughness spectrum is the calculation of the topography coefficient S_a (arithmetic mean) at each resolution level of the signal. 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. To describe this phenomenon, the local normal deviation NAD is used. Thus, a decrease in value of the NAD is related to the smoothness and brightness in the same way.

16.2.3 *Tribohaptic Measurement*

A haptic tribometer system was used to perform the friction tests on hair. The experimental test consists in applying a normal load F_z on the hair through a slider, which allows controlling the contact pressure, after which the hair is dragged away by a

linear displacement motor in order to simultaneously measure the friction force, F_x , and the vibrations between the probe and the hair surface.

The slider is equivalent to a half-spherical probe (radius $R = 10$ mm), composed of an elastomer material whose mechanical properties are very close to the properties of the human finger (the Young's modulus is equal to 55 kPa and its average surface roughness R_a is 2.6 μm) [12]. The normal load applied on the hair by the haptic tribometer is controlled with a normal load sensor. The maximum load during the loading can reach 3 N with a resolution of 1mN. During the test, the hair is fixed on a linear displacement table with a defined preload. The displacement table drags the hair away at constant speed. A tangential sensor (the same characteristics as the normal sensor) and a piezoelectric accelerometer, respectively, record the friction force and the vibrations generated by the sliding between the slider and the hair, fixed on the slider.

For each test, the slider rubbed onto the hair five times, in uniform back-and-forth translations. The friction length was set to 20 mm. In order to represent realistic frictional conditions for people, the sliding speed was set to 20 $\text{mm}\cdot\text{s}^{-1}$ and the normal force applied was equal to 0.3 N. These conditions have been frequently observed in the literature and correspond to classical human handling conditions [13].

To give a quantitative value to the vibrations, the average acoustic vibratory level was defined as:

$$L_a = 20 \log \left(\frac{A_{\text{RMS}}}{A_{\text{ref}}} \right)$$

where L_a is the acoustic vibratory level, A_{RMS} the root mean square value of the acceleration in m/s^2 , and $A_{\text{ref}} = 10^{-6}$ m/s^2 the smallest acceleration that can be detected by the sensor.

16.3 Results and Discussion

16.3.1 Effect of Chemical Treatments

16.3.1.1 Hair Topography

Figure 16.1 shows typical topographies obtained for virgin, colored, and bleached hair. The images show that the cuticle scales are relatively intact for virgin hair. By contrast, an increase of roughness is evident for colored and bleached hair samples. The colored hair sample contains poorly defined cuticles with toothed edges and the surface seems to be rough. In the case of bleached hair, no well-defined cuticles can be observed, the surface is very rough with a lot of asperities. It is suggested that bleaching process has removed almost all the cuticle, leaving the underlying cortex structure. Bleaching is known to remove cuticular layers [14, 15].

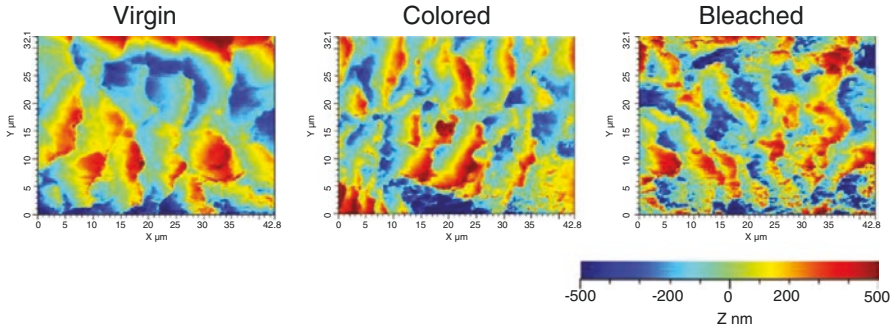


Fig. 16.1 Topography of virgin, colored and bleached hair

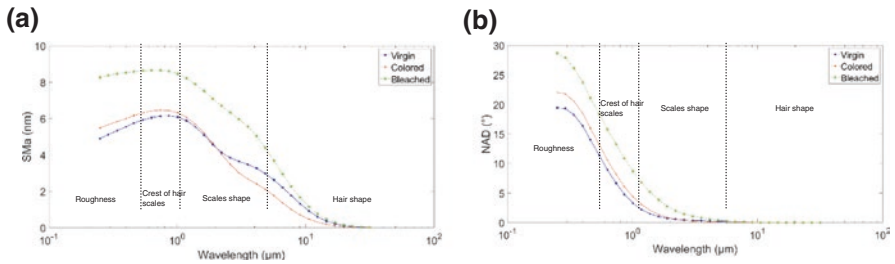


Fig. 16.2 Sma (a) and NAD (b) spectra obtained for virgin, colored, and bleached hair samples. For clarity, the figure represents the mean spectrum of each type of hair

Figure 16.2 presents the SMA and NAD spectra obtained for the different types of hair samples. The spectra can be divided into several parts according to the values of the wavelengths. It can be noticed that the value of SMA and NAD is near zero for the wavelengths higher than 10 μm. This is due to the fact that the cylindrical shape of the hair fiber has been removed for the topographical analysis.

The results show that the differences between the topography of the hair samples is localized at the small and medium wavelengths. Thus, the chemical treatments induce topographical modifications at the level of the hair scales (small wavelengths). Indeed, the SMA value is higher for treated hair samples, especially for bleached hair. This phenomenon is explained by the formation of surface defects due to chemical treatments, which increases the slight roughness at the hair surface.

Regarding the NAD spectrum presented in 2(b), we observed that the difference between virgin and chemically treated hair samples is localized at the level of the hair scales too. Indeed, there is an increase of the NAD. Our hypothesis is that the chemical treatments open and lift the hair scales to allow the diffusion and reaction of colorants inside hair [16]. Thus, the hair surface becomes rougher and less bright because the scales are less flat than for virgin hair.

The SMA and NAD spectra corroborate the previous observations and suggest that chemical treatments have an impact on the friction and wear of hair, which may result in a decrease of its sensorial properties.

16.3.1.2 Friction and Wear of Hair

We are interested in the effect of hair bleaching and coloring on friction and wear. The acoustic vibratory levels measured are presented in Fig. 16.3.

The acoustic vibratory level is expressed in decibels (dB). It should be noted that a difference of 3 dB corresponds to a multiplication of the sound signal by two.

Along the cuticle, the data show that the acoustic vibratory level is higher for colored and bleached hair samples (respectively, 97.65 dB and 98.26 dB) than for virgin hair (96.33 dB). Previous team results have shown that vibrations are related to the softness of the surface [17]. In this work, the authors rubbed a tribo-acoustical probe on the skin surface and measured the acoustic level of vibrations induced by the friction. They showed that the level of acoustic vibrations gave a signature of skin softness. The surface is softer for a lower acoustic vibratory level. Thus, virgin hair samples are the softest. This proves that coloring and bleaching damage hair fibers and make them less soft.

Thus, as with the friction coefficient, we observed a directionality effect for all the hair samples. The acoustic vibratory level was always higher for the tip-to-root friction tests. This means that the hair fibers are felt less soft when touching against the direction of the cuticle scales.

Spectral analyses of friction-induced vibration signals were carried out using Matlab R2018a to assess the frequency content and corresponding amplitude from the power spectral densities (PSD) of the signals [18, 19].

Figure 16.4 presents the PSD signals obtained for the different hair samples. Two main peaks can be identified in the figure for each sample. To compare the virgin

Fig. 16.3 Effect of chemical treatments and directionality dependence of hair focus on acoustic vibratory levels. All the values correspond to the mean \pm standard deviation values

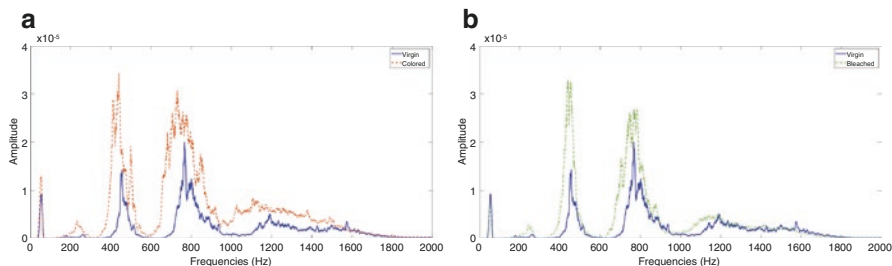
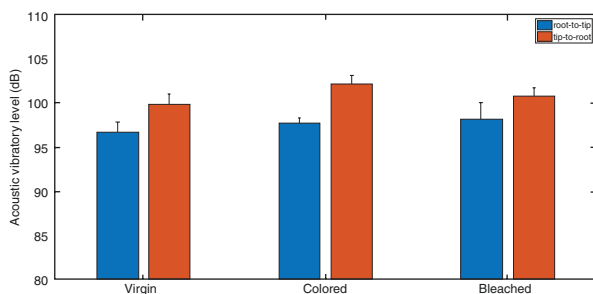


Fig. 16.4 PSD signal (a) virgin vs. colored hair; (b) virgin vs. bleached hair for the root-to-tip direction

and chemically treated hair, Table 16.1 gives the deviations of amplitude and frequency of the treated hair as a function of virgin hair values. These spectra show that the peaks are higher in amplitude for colored and bleached hair compared with virgin hair. There is also a small frequency offset. The peak height is linked to the magnitude of vibrations (due to the contact between the probe and the surface of the hair). Softness and lower friction are believed to be responsible for the lower amplitude of the signal. This was consistent with our previous observations: a higher amplitude corresponds to a higher acoustic vibratory level, thus a rougher surface. Cattaneo and Vecchi explained that smooth surfaces induce virtually no vibration, while rougher surfaces induce more vibrations [20].

To conclude this part, bleaching and coloring processes are responsible for damaging hair fiber, especially at the cuticle level. This damage impacts the sensorial properties of hair, making it rougher, dull, and less soft.

16.3.2 Effect of Thermomechanical Treatments

16.3.2.1 Hair Topography

The topography of virgin, damaged, and smoothed hair is presented in Fig. 16.5. The results show that the cuticle of the virgin hair is complete whereas the cuticles are not well defined and several scales are worn away from the damaged hair. Furthermore, the surface of the damaged hair seems to be rougher with more apparent structure and particulate matter. This result is consistent with previous works [5, 7, 13]. We assume that this damage impacts the sensorial properties of hair.

Table 16.1 Deviations of amplitudes and frequencies of treated hair as a function of virgin hair values

Sample vs. virgin	$\Delta\omega$ (Hz)	Δ Amplitude ($\times 10^{-5}$)
Colored	-16	2 (+142%)
	-37	1.6 (+80%)
Bleached	-6	1.8 (+128.7%)
	-2	1.3 (+65%)

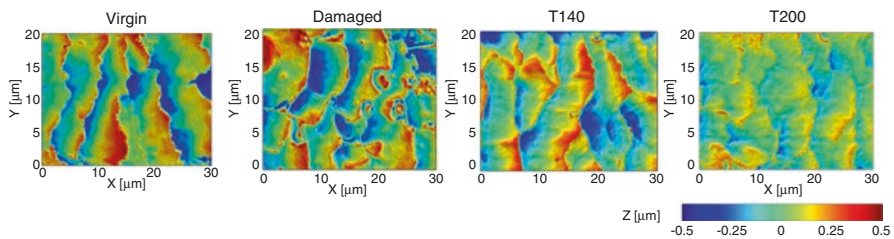


Fig. 16.5 Topography of virgin, damaged, and smoothed hair at 140 °C and 200 °C

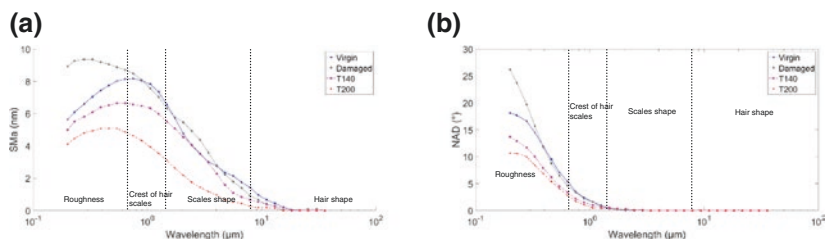


Fig. 16.6 SMA (a) and NAD (b) spectra obtained for virgin, damaged, and smoothed hair samples at 140 °C and 200 °C. For the sake of clarity, the figure represents the mean spectrum of each type of hair

If we focus on smoothed hair, we can observe that there are fewer asperities than for the damaged hair. The morphology of smoothed hair at 140 °C is similar to that of virgin hair.

The scales are well-defined whereas the edges are less marked. Thus, the hair fiber appears smoother. The observation is amplified at a temperature of 200 °C: the hair surface appears very smooth. Our hypothesis is that the temperature could melt the cuticle which is then easily deformed by applying pressure to the hair surface via the smoothing process.

Figure 16.6 presents the SMA and NAD spectra obtained. The differences between the topography of the hair samples are localized at the small and medium wavelengths, particularly at the level of the hair scales.

The damaged hair presents the highest values of SMA and NAD in the band of small and medium wavelengths. This is due to the appearance of asperities, markedly inclined with respect to the hair surface. Thus, the surface is perceived rougher.

Consequently, the results show that the smoothed hair has lower SMA and NAD values than for virgin hair that increase if the temperature is high. This is consistent with the previous topographical observations: the surface becomes smoother. Furthermore, the surface is perceived as being bright due to the lower NAD.

Finally, we assumed that thermomechanical treatments on damaged hair could improve its sensorial properties by erasing asperities and flattening the hair surface. Nevertheless, these treatments did not permit preserving the hair's morphology: the hair cuticles became flatter and the scales were less defined.

16.3.2.2 Friction and Wear of Hair

In this part, we study the effect of smoothing at different temperatures on the sensorial properties of the hair fibers. The objective is to ensure whether thermomechanical treatments improve the sensorial quality of hair fibers damaged previously.

Figure 16.7 presents the acoustic vibratory levels obtained. The highest acoustic vibratory level equal to 101.87 dB. As for coloring and bleaching, the results show that damaged hair is less soft and rougher than virgin hair. This is in accordance with the topographical observations.

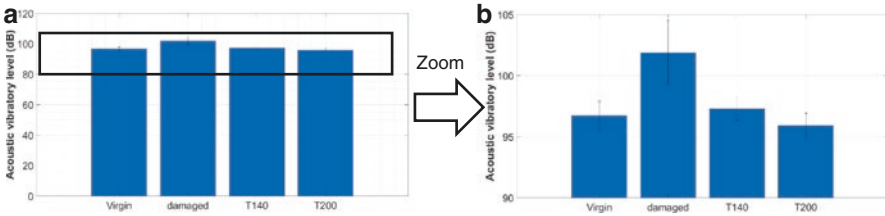


Fig. 16.7 Effect of thermomechanical treatments of hair on (a) acoustic vibratory level, (b) focus on acoustic vibratory levels. All the values correspond to the mean \pm standard deviation values

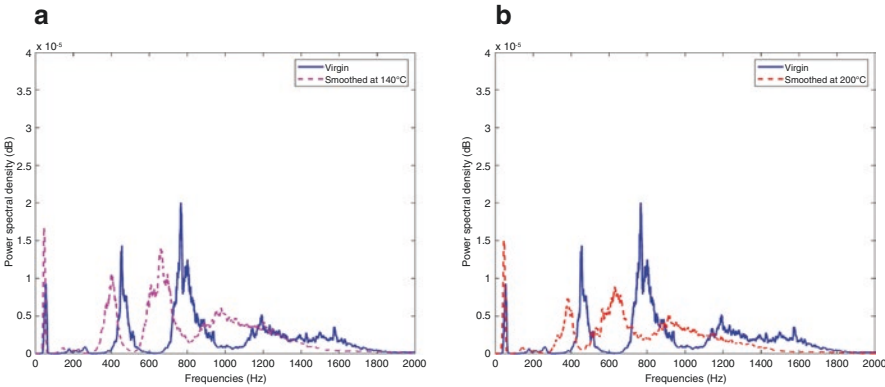


Fig. 16.8 PSD signal (a) virgin vs. smoothed hair at 140 °C; (b) virgin vs. smoothed hair at 200 °C

Table 16.2 Deviations of amplitudes and frequencies of smoothed hair as a function of virgin hair values

Sample/Virgin	$\Delta\omega$ (Hz)	Δ Amplitude ($\times 10^{-5}$)
T140	-56	-0.4 (-28%)
	-105	-0.6 (-70%)
T200	-71	-0.68 (-48.6%)
	-138	-1.12 (-56%)

Smoothing damaged hair induces a decrease in acoustic vibratory level. Indeed, after smoothing treatment, the acoustic vibratory level returns to the base level. For a temperature of 200 °C, it reaches a value lower than that for the virgin hair.

All these results demonstrate that the smoothing treatment improves the softness of hair fiber, making it less rough.

The PSD spectra of virgin and smoothed hair are presented in Fig. 16.8. We again observe two main peaks. Table 16.2 gives the deviations of amplitude and frequency of smoothed hair as a function of virgin hair values. The amplitudes of the peaks are lower for smoothed hair, and there is a significant shift of the peaks towards the low frequencies. The decrease in amplitude reflects the smoother aspect of hair after the thermomechanical treatments. Indeed, a smooth surface reduces the magnitude of vibrations due to the contact between the probe and the hair surface.

Therefore, the decrease in frequency corresponds to an increase in wavelength. Thus, the period of roughness is higher for smoothed hair. This may be due to the flattening of the scales after smoothing, which erases the natural periodicity of the hair surface. This phenomenon becomes more pronounced when increasing temperature.

In conclusion to this part, smoothing can improve the sensorial properties of hair, making it smooth, bright, and soft.

16.4 Conclusion

The present study concerned the modifications of morphology and biotribology of hair due to chemical and thermomechanical treatments. The conclusions of this study are as follows:

- A new method using wavelet decomposition of a surface was used to quantitatively evaluate hair cuticle modifications. The advantage is that it allows evaluating damage to hair at different scales. The SMA and NAD parameters were found to be good indicators of the sensorial properties of hair (softness and brightness).
- We introduced vibration analysis during the friction test as a new parameter to study the tribological behavior of the hair surface. The acoustic vibratory level was found to be a good indicator of the hair surface's softness.
- The dependence of the frequency peak of the induced vibrations during the friction tests, with respect to the roughness of the hair surface, was demonstrated. An increase in roughness corresponds to a peak of higher amplitude.
- It appeared that coloring and bleaching hair damaged it and impacted its sensorial properties: the hair fiber became less soft, less bright, and rougher.
- Applying thermomechanical treatments to damaged hair improved sensorial properties by smoothing the surface and erasing the asperities. Thus, the hair fiber appears brighter and smoother.

Finally, these results are essential for understanding the impact of different solicitations on the physical and sensorial properties of hair. These techniques could be adapted to other cosmetic products that modify the cuticle's topography.

References

1. Garcia M, Epps J, Yare R, Hunter L. Normal cuticle-wear patterns in human hair. *J Cosmet Chem.* 1978;29:155–75.
2. Fougere M, Vargiolu R, Pailler-Mattei C, Zahouani H. Study of hair topography modification by interferometry. *Wear.* 2009;266(5–6):600–4.
3. Rabinowicz E. *Friction and wear of materials.* New York: Wiley-Interscience; 1995.
4. Robbins C. *Chemical and physical behavior of human hair.* New York: Springer; 1994.

5. La Torre C, Bhushan B. Nanotribological characterization of human hair and skin using atomic force microscopy. *Ultramicroscopy*. 2005;105:155–75.
6. Smith J, Swift J. Lamellar subcomponents of the cuticular cell membrane complex of mammalian keratin fibres show friction and hardness contrasts by AFM. *J Microsc*. 2002;202:182–93.
7. LaTorre C, Bushan B. Nanotribological effects of hair care products and environment on human hair using atomic force microscopy. *J Vac Sci Technol A*. 2005;23:1034–45.
8. LaTorre C, Bhushan B. Investigation of scale effects and directionality dependence on adhesion and friction of human hair using AFM and macroscale friction test apparatus. *Ultramicroscopy*. 2006;106:720–34.
9. LaTorre C, Bhushan B, Yang J, Torgerson P. Nanotribological effects of silicone type, silicone deposition level, and surfactant type on human hair using atomic force microscopy. *J Cosmet Sci*. 2006;57:37–56.
10. Gamez-Garcia M. Cuticle decementation and cuticle buckling produced by Poisson contraction on the cuticular envelope of human hair. *J Soc Cosmet Chem*. 1986;153:141–53.
11. Feughelman A. Mechanical properties and structure of alpha-keratin fibres: wool, human hair and related fibres. Sydney: University of South Wales; 1997.
12. Cornuault P-H, Carpentier L, Bueno M-A, Cote J-M, Monteil G. Influence of physico-chemical, mechanical and morphological fingerpad properties on the frictional distinction of sticky/slippery surfaces. *J R Soc Interface*. 2015;12(110):20150495.
13. Bhushan B, Wei G, Haddad P. Friction and wear studies of human hair and skin. *Wear*. 2005;259(7–12):1012–21.
14. Garcia ML, Epps JA, Yare RS, Hunter LD. Normal cuticle-wear patterns in human hair. *J Soc Cosmet Chem*. 1978;29(3):155–75.
15. Jeong MS, Lee CM, Jeong WJ, Lee KY. Significant damage of the skin and hair following hair bleaching. *J Dermatol*. 2010;37:882–7.
16. Luo C, Zhou L, Chiou K, Huang J. Multifunctional graphene hair dye. *Chem*. 2018;4(4):784–94.
17. Zahouani H, Vargiolu R, Boyer G, Pailler Mattei C, Laquière L, Mavon A. Friction noise of human skin in vivo. *Wear*. 2009;267(5–8):1274–80.
18. Thomson W, Dahleh M. Theory of vibration with applications. Upper Saddle River: Prentice Hall; 1998.
19. Ding S, Bhushan B. Tactile perception of skin and skin cream by friction induced vibrations. *J Colloid Interface Sci*. 2016;481:131–43.
20. Cattaneo Z, Vecchi T. Blind vision: the neuroscience of visual impairment. Cambridge: MIT Press; 2011.

Chapter 17

Skin Color and Pigmentation



Ammitzboell Elisabeth

17.1 Introduction

Why do people from different parts of the world have different colored skin? One theory is:

Variations in human skin color are adaptive traits that correlate closely with geography and the sun's ultraviolet (UV) radiation.

Since strong sun exposure damages the body, the solution was to evolve skin that was permanently dark so as to protect against the sun's more damaging rays.

Melanin, the skin's brown pigment, is a natural sunscreen that protects tropical peoples from the harmful effects of ultraviolet (UV) rays. UV rays can, for example, strip away folic acid, a nutrient essential to the development of healthy fetuses. Yet, when a certain amount of UV rays penetrates the skin, it helps the human body to produce vitamin D to absorb the calcium necessary for strong bones. This delicate balancing act explains why people that migrated to colder geographic zones with less sunlight developed lighter skin color. As people moved to areas farther from the equator with lower UV levels, natural selection favored lighter skin which allowed UV rays to penetrate and produce essential vitamin D. The darker skin of people who lived closer to the equator was important in preventing folate deficiency. Measures of skin reflectance, a way to quantify skin color by measuring the amount of light it reflects, in different ethnicities support this idea. It is discussed that the darker skin was developed to protect the person from UV rays that cause skin cancer. But it likely had little effect on the evolution of skin color because evolution favors changes that improve reproductive success and skin cancer usually affects people after they have had children.

A. Elisabeth (✉)
Skin Clinic Hareskov, Værloese, Denmark
e-mail: eaholm@dadlnet.dk

Quantification of disease severity is a prerequisite for the development of evidence-based therapy. Today there is no international consensus on guidelines for clinical assessment of skin color, 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 way doctors assess, e.g., erythema or dry skin [1]. In addition, many of the scoring systems include assessment of disease extent, which has been shown to be difficult [2]. To provide objective and reproducible methods instrumental evaluation of skin morphology and functions have been proposed.

Several of these techniques have reached a technological maturity which suggests that a more routine use is possible.

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 color was performed in 1937 by Edwards Duntley, who obtained data specific for different types of pigmentation [3].

The perception of color by the human eye and brain has a range from 400 to 800 nm, with a maximum of sensitivity between 500 and 600 nm, corresponding to the color of blood and hence of redness [4]. However, it is quite difficult to study and record skin color changes quantitatively, as individual perception of color is complex and subjective. Several commercial instruments are available: scanning reflectance spectrophotometers, tristimulus colorimetric, and narrow-band simple reflectance [5–9].

17.2 Skin Color

Skin color 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 color 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 hemoglobin 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 than less pigmented areas, which is also the visual inspection [4].

17.3 Melanin Pigment

Skin and hair color depend on **melanin** production in the skin. Melanocytes produce pigment at the dermal–epidermal junction and in the outer root sheath of the hair follicle. The ratio is one melanocyte for every 10–20 keratinocytes. Melanin

pigments have a wide range of colors including black-brown eumelanins, yellow-red pheomelanins, and a range of intermediate pigments. Special tyrosinase-rich organelles called melanosomes, which are found in the cytoplasm of melanocytes, produce melanin pigment.

Carotenoids are yellow pigments of exogenous origin acquired by eating yellow fruits and vegetables, e.g., carrots. They are stored in the stratum corneum, sebaceous glands, and subcutaneous fat. Carotenoids have only a minor influence on skin color, except in rare cases in which they accumulate excessively [3, 10].

Hemoglobin is responsible for the red component of skin color. Oxyhemoglobin in the capillaries and arterioles imparts a bright red coloring, whereas reduced hemoglobin in the venules imparts a bluish color. The color of oxyhemoglobin 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 color tends to be reddish. On the other hand, skin color 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 oxyhemoglobin, where vasoconstriction as in ischemia cause the skin to be more bluish. Hemoglobin 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.

17.4 Individual Differences

Skin color varies from race to race, and in all subjects, skin color 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 color.

A measurement of the buttocks area is often used as a reference point for assessment of UV radiation's influence of pigmentation.

The color of nonexposed skin is known as constitutional, and that of exposed skin as facultative. The difference between the two colors can be calculated and indicates the pigmentation capacity of the subject.

The constitutional skin color correlates with Fitzpatrick photo type, with a darkening color gradient from I to IV. Constitutional skin color can be taken as a parameter of photo type.

Aged skin seemed to be less colored and male skin may be more redder than female skin [11].

Numerous drugs have been implicated in pigmentation disorders, including nonsteroidal anti-inflammatory drugs (NSAIDs), antimalarial agents (e.g., chloroquine, quinine), antipsychotic drugs (e.g., chlorpromazine, phenothiazine), anticonvulsants (e.g., phenytoin), amiodarone, tetracyclines, cytotoxic drugs (e.g., cyclophosphamide, busulfan, bleomycin, adriamycin), and heavy metals. The mechanisms involved in pigmentation disorders vary with the type of agent involved.

For example, some drugs may react with melanin to form drug–pigment complexes, which may be exacerbated by sunlight stimulation of melanin synthesis.

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

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

A 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

The Melasma Area and Severity Index (MASI), an outcome measure developed to provide a more accurate quantification of the severity of melasma and changes during therapy, was developed by Kimbrough-Green et al., who based it on a similar scoring system devised for psoriasis. The MASI score is calculated by subjective assessment of 3 factors: area (A) of involvement, darkness (D), and homogeneity (H), with the forehead (f), right malar region (rm), left malar region (lm), and chin (c), corresponding to 30%, 30%, 30%, and 10% of the total face, respectively. The area of involvement in each of these four areas is given a numeric value of 0–6 (0 = no involvement; 1 = 0–10%; 2 = 10–29%; 3 = 30–49%; 4 = 50–69%; 5 = 70–89%; and 6 = 90–100%). Darkness and homogeneity are rated on a scale from 0 to 4 (0 = absent; 1 = slight; 2 = mild; 3 = marked; and 4 = maximum) [12].

Assessment of pigmentation by diffuse reflectance spectroscopy correlated well with the corresponding clinical evaluation of pigmentation ($R^2 = 0.852$). Also a study of contact dermatitis compared visual scoring with Laser Doppler flowmetry and found a high correlation ($r = 0.9079$, $p < 0.001$) [4].

17.5 Instruments

17.5.1 Reflectance Spectrophotometric Evaluation

Older spectrophotometric scanning devices had previously little practical use because the broad melanin absorption band overlaps the hemoglobin band. However new instruments that measure the hemoglobin band specifically and express erythema as an index of hemoglobin relative to melanin are now available.

Recording of reflectance is significantly influenced by nonspecific 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 Meter (DiaStron Ltd., Andover, Hampshire, UK), DermaSpectrometer (Cortex Technology, Hadsund, Denmark), Mexameter (Courage & Khazaka GmbH, Köln, Germany), and the UV-Optimize (Matic, Naerum, Denmark).

Tele-spectroradiometers (TSRs) and spectrophotometers (SPs) are two types of instruments that are most widely used for measuring skin color. For example, TSRs are used for measuring color appearance such as in cosmetic industry for developing skin color charts and evaluating skin products. SPs are mainly used for diagnosing skin disease symptom, such as erythema and irritation [5–8].

The Dermaspectrometer® (Cortex Technology, Hadsund, Denmark) is a narrow-band reflectance spectrophotometer designed for measuring specific colors of two major chromophores: hemoglobin and melanin in human skin. The Dermaspectrometer light emitting diodes emit light at two defined wavelengths: 568 nm (green) and 655 nm (red). The photodetector measures the light reflected by the skin at wavelengths, respectively, in the green and the red for hemoglobin and for melanin. The melanin Index (M-index) is mainly influenced by the melanin content, where the Erythema Index (E-Index) is influenced by the hemoglobin content, assuming that absorbance of hemoglobin 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 meter measures skin pigmentation (melanin) and redness (hemoglobin). 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 hemoglobin and melanin [13–15]. Redness is measured on a scale from 0 to 100. “0” is the color 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 be given as a mean of several.

The Minolta Spectrophotometer (Minolta, Osaka, Japan) (Chromameter) 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 color 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 color in contrast to the other instruments, which take a measure at wavelengths of specific chromophores.

L^* is, as mentioned, 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 hemoglobin. 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 is 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 $L^*a^*b^*$ color parameters and the erythema/melanin indices [16].

Visi-Chroma VC-100[®] is a newer imaging tristimulus colorimeter using same process [9].

17.6 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 [17]. Before the start of any measurement session an acclimatization 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 color. 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 noninvasive 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 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 color, 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, alcohol, and other factors cause cutaneous vasodilatation and thereby influence cutaneous blood flow and color, whereas smoking, caffeine, mental stress, and adrenergic system activation are associated with pale skin color (vasoconstriction). These factors should also be systematically considered in preparation for a study.

The factors influencing skin color measurements are summarized in Table 17.1.

Table 17.1 Factors influencing skin color measurements

Variables that influence skin color	How they influence?
Age	Aged skin may be less colored
Sex	Male skin may be redder than female skin
Race	Different pigmentation among races Visual estimation of erythema in pigmented skin is misleading
Anatomical sites	The face, neck, palm, and soles have in general the highest index of erythema
Diurnal variation	Erythema of the skin increases during the day
Physical activity	Erythema of the skin increases with the level of exercise due to increased blood flow. An acclimatization period of 20 min is important before assessment
Mental activity	May activate the adrenergic system resulting in peripheral vasoconstriction making the skin pale
Orthostatic effects	Erythema of the skin decreases if the measured area rises from the heart level
Nicotine	Cause vasoconstriction
Alcohol	Cause vasodilatation
Temperature	Normal skin surface temperature is 30 °C. If temperature rises, vasodilatation and thereby erythema occurs, and reach its maximum at 40 °C. Room temperature best between 19 and 23 °C
Ambient light/sun	May influence measurement of skin color. The probe must be kept perpendicular to the skin surface to avoid false light
Seasonal variation	Tan in summer and xerosis in winter influence skin color
Drugs	Nonsteroidal anti-inflammatory drugs (NSAIDS), antimalarial agents (e.g., chloroquine, quinine), antipsychotic drugs (e.g., chlorpromazine, phenothiazine), anticonvulsants (e.g., phenytoin), amiodarone, tetracyclines, cytotoxic drugs (e.g., cyclophosphamide, busulfan, bleomycin, adriamycin), and heavy metals

References

1. Sprickelman AB, Tupker RA, Burgerhof H, Schouten JP, Brand PL, Heymans HS, et al. Severity scoring of atopic dermatitis: a comparison of three scoring systems. *Allergy*. 1997;52(9):944–9.
2. Charman CR, Venn AJ, Williams HC. Measurement of body surface area involvement in atopic eczema: an impossible task? *Br J Dermatol*. 1999;140(1):109–11.
3. Edwards E, Duntley S. Pigment and color of living human skin. *Am J Anat*. 1939;65:1–7.
4. Stamatias GN, Zmudzka BZ, Kollias N, Beer JZ. Non-invasive measurements of skin pigmentation in situ. *Pigment Cell Res*. 2004;17:618–26.
5. Serup J, Agner T. Colourimetric quantification of erythema—a comparison of two colourimeters (Lange microcolourand Minolta chroma meter CR-200) with a clinical scoring scheme and laser-Doppler flowmetry. *Clin Exp Dermatol*. 1990;15:267–72.
6. Baquié M, Kasraee B. Discrimination between cutaneous pigmentation and erythema: comparison of the skin colourimeters Dermacatch and Mexameter. *Skin Res Technol*. 2014;20(2):218–27.
7. Stamatias GN, Zmudzka B, Kollias N, Beer JZ. In vivo measurement of skin erythema and pigmentation: new means of implementation of diffuse reflectance. *Br J Dermatol*. 2008;159(3):683–90.
8. Wang M, Xiao K, Luo M, Pointer M, Cheung V, Wuerger S. An investigation into the variability of skin colour measurements. *Color Res Appl*. 2010;43(4):458–70.
9. Barel AO, Clarys P, Alewaeters K, Duez C, Hubinon J-L, Mommaerts M. The Visi-Chroma VC-100: a new imaging colourimeter for dermatocosmetic research. *Skin Res Technol*. 2001;7:24–31.
10. Wutrich B, Kagi MK, Joller-Jemelka H. Soluble CD14 but not interleukin-6 is a new marker for clinical activity in atopic dermatitis. *Arch Dermatol Res*. 1992;284(6):339–42.
11. Chilcott RP, Farrar R. Biophysical measurements of human forearm skin in vivo: effects of site, gender, chirality and time. *Skin Res Technol*. 2000;6:64–9.
12. Pandya A, Hynan L, Bhore R, et al. Reliability assessment and validation of the Melasma Area and Severity Index (MASI) and a new modified MASI scoring method. *Am J Dermatol*. 2011;64(1):78–83.
13. Na R, Stender IM, Henriksen M, Wulf HC. Autofluorescence of human skin is age-related after correction for skin pigmentation and redness. *J Invest Dermatol*. 2001;116(4):536–40.
14. Jemec GB, Na R, Wulf HC. The inherent capacitance of moisturising creams: a source of false positive results? *Skin Pharmacol Appl Ski Physiol*. 2000;13(3–4):182–7.
15. Lock-Andersen J, Gniadecka M, De Fine OF, Dahlstrom K, Wulf HC. UV induced erythema evaluated 24 h post-exposure by skin reflectance and laser Doppler flowmetry is identical in healthy persons and patients with cutaneous malignant melanoma and basal cell cancer. *J Photochem Photobiol B*. 1997;41(1–2):30–5.
16. Clarys P, Alewaeters K, Lambrecht R, Barel AO. Skin color measurements: comparison between three instruments: the Chromameter(R), the DermaSpectrometer(R) and the Mexameter(R). *Skin Res Technol*. 2000;6(4):230–8.
17. Fullerton A, Fischer T, Lahti A, Wilhelm KP, Takiwaki H, Serup J. Guidelines for measurement of skin colour and erythema. A report from the Standardization Group of the European Society of Contact Dermatitis. *Contact Dermatitis*. 1996;35(1):1–10.

Chapter 18

Characterization of Sunscreens: Determination of the SPF



Juergen Lademann, Joachim W. Fluhr, and Martina C. Meinke

Core Messages

- 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. A high SPF leads to extended sun exposure and an increased risk of skin damage by radical formation in the visible spectral range.

J. Lademann (✉) · M. C. Meinke

Department of Dermatology and Allergology, Center of Experimental and Cutaneous Physiology (CCP), Universitätsmedizin – Charité, Berlin, Germany
e-mail: Juergen.lademann@charite.de; martina.meinke@charite.de

J. W. Fluhr

Department of Dermatology and Allergology, Universitätsmedizin – Charité, Berlin, Germany
e-mail: joachim.fluhr@charite.de

© Springer Nature Switzerland AG 2020

J. W. Fluhr (ed.), *Practical Aspects of Cosmetic Testing*,
https://doi.org/10.1007/978-3-030-44967-4_18

18.1 Introduction

The sun spectrum at the surface of the earth extends from UVB, UVA, visible, to infrared radiation. On the one hand, sun radiation is essential for human life; for instance, it stimulates vitamin D synthesis and is important for human well-being [1]. On the other hand, high doses of UV radiation can cause sunburn, immune-suppression, skin aging, and even skin cancer [2, 3]. However, the human body has developed protection strategies against UV radiation by increasing melanin synthesis, which is an efficient UV absorber [4, 5]. Additionally, melanin has strong scattering properties in the visible spectral range, which reduce the penetration of the sunlight into the living tissue [6]. 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 defense 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) [2], which describes the UVB protection. Today, the UVA protection is also declared by the manufacturer on sunscreen products. Nowadays sunscreens contain not only UV filters but also antioxidants, anti-inflammatory agents, and other ingredients whose risks have not been fully evaluated [8]. Furthermore, up to 50% of the radicals are formed in the visible and near-infrared spectral regions [9] which are not covered by UV filters but can be reduced by antioxidants and pigments [10, 11].

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 [12]. 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 ISO24444 [13]. 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:

$$\text{SPF} = \text{MED of the sunscreen protected skin} / \text{MED of the unprotected skin.}$$

In accordance with ISO24444 [13], the SPF is determined on the back of 12–20 volunteers with skin types I, 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 on the skin surface structure [14].

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. Additional agents, such as anti-inflammatory agents [15], antioxidants [11], and DNA repairing enzymes [16], could influence the SPF.

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.

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. 18.1 Determination of the MED on the skin on the back by a sun simulator. <http://www.schrader-institute.de/home/pruefung/haut-pruefung/sonnenschutz/>



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

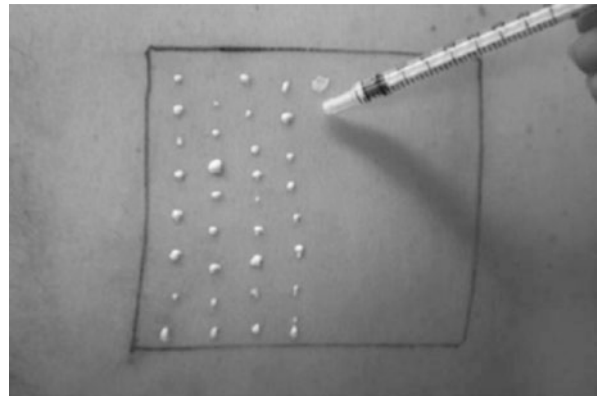
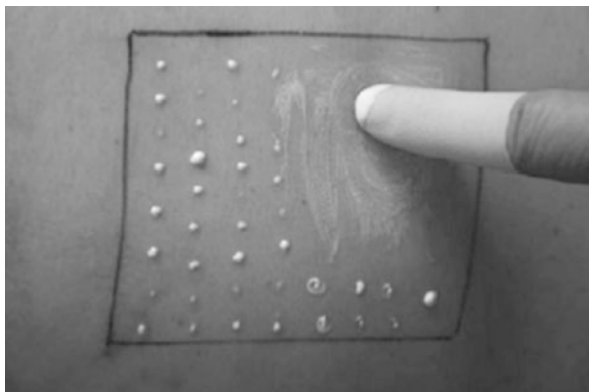


Fig. 18.3 Saturation of the glove in the sunscreen



Fig. 18.4 Homogeneous distribution of the sunscreen with the saturated glove



18.4 International Methods for the SPF Determination

Different methods are used for the determination of the SPF. In Europe, ISO24444 [13] 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 ISO24444, 10 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 ISO24444 procedure [13].

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 [17]. 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.

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 lifestyle.
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) [18], 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) [19], 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 [20]. 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. Nowadays the in vitro method according to DIN EN ISO 24443 [21] is applied in most European countries. It determines the UVAPF (“ultraviolet-A protection factor”) and correlates with the PPD. For determining the UVAPF, a thin film of a sunscreen is applied on a PMMA plate and its transmittance is measured in the UV spectral region. The in vitro SPF is normalized to the in vivo SPF taking into account the photostability.

18.6 Classification of Sunscreens

Following the regulations of the European Commission from September 22, 2006 [22], 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.

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 [23] that a sunscreen should have a UVA protection of at least one-third of the SPF, determined by the PPD method. In this

Table 18.1 The different SPF values are categorized in four groups: low, medium, high, and very high protection

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

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 [9]. 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. Porcine skin can also be used for these measurements [11].

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 [24, 25]. 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 *in vivo* diffuse reflectance spectroscopy (DRS) [26] or Hybrid DRS (HDRS) [27] which is a combination of *in vivo* and *in vitro* method. The diffuse reflectance in the UV is measured *in vivo* with and without an applied sunscreen product and the transmittance is calculated. The subsequent calculation is in accordance to the *in vitro* SPF calculation. For HDRS the *in vivo* measurement is performed in the UVA only and the UV spectrum *in vitro* is adjusted to the *in vivo* UVA spectrum. The SPF corresponds to the classical *in vivo* SPF values [27, 28]. Additionally, this method can be easily used to determine the UVAPF.

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 ISO24444 is an accepted guideline. The SPF is strongly related to the UVB part of the sun spectrum. New noninvasive spectroscopy-based *in vivo* methods are under development. 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

1. Jablonski NG, Chaplin G. The evolution of human skin coloration. *J Hum Evol.* 2000;39(1):57–106.
2. Pathak MA. Sunscreens and their use in the preventive treatment of sunlight-induced skin damage. *J Dermatol Surg Oncol.* 1987;13(7):739–50.
3. Phillips TJ, et al. Effect of daily versus intermittent sunscreen application on solar simulated UV radiation-induced skin response in humans. *J Am Acad Dermatol.* 2000;43(4):610–8.
4. Pathak MA, et al. Melanin formation in human skin induced by long-wave ultra-violet and visible light. *Nature.* 1962;193:148–50.
5. Zonios G, et al. Melanin absorption spectroscopy: new method for noninvasive skin investigation and melanoma detection. *J Biomed Opt.* 2008;13(1):014017.
6. Alaluf S, et al. The impact of epidermal melanin on objective measurements of human skin colour. *Pigment Cell Res.* 2002;15(2):119–26.
7. Jiang R, et al. Absorption of sunscreens across human skin: an evaluation of commercial products for children and adults. *Br J Clin Pharmacol.* 1999;48(4):635–7.
8. Paul SP. Ensuring the safety of sunscreens, and their efficacy in preventing skin cancers: challenges and controversies for clinicians, formulators, and regulators. *Front Med.* 2019;6:195.
9. Zastrow L, et al. From UV protection to protection in the whole spectral range of the solar radiation: new aspects of sunscreen development. *Adv Exp Med Biol.* 2017;996:311–8.
10. Meinke MC, et al. Radical protection by differently composed creams in the UV/VIS and IR spectral ranges. *Photochem Photobiol.* 2013;89(5):1079–84.
11. Souza C, et al. Radical-scavenging activity of a sunscreen enriched by antioxidants providing protection in the whole solar spectral range. *Skin Pharmacol Physiol.* 2017;30(2):81–9.
12. Hadgraft J. Modulation of the barrier function of the skin. *Skin Pharmacol Appl Skin Physiol.* 2001;14(Suppl 1):72–81.
13. ISO, 24444 Cosmetics—Sun Protection Test methods—*In vivo* determination of the sun protection factor (SPF); 2010.
14. Lademann J, et al. Influence of nonhomogeneous distribution of topically applied UV filters on sun protection factors. *J Biomed Opt.* 2004;9(6):1358–62.
15. Werner M, et al. Determination of the influence of the anti-inflammatory ingredients panthenol and bisabolol on the SPF value *in vivo*. *Skin Pharmacol Physiol.* 2017;30(6):284–91.
16. Carducci M, et al. Comparative effects of sunscreens alone vs sunscreens plus DNA repair enzymes in patients with actinic keratosis: clinical and molecular findings from a 6-month, randomized, clinical study. *J Drug Dermatol.* 2015;14(9):986–90.
17. Yamashita T, et al. *In vivo* assessment of pigmentary and vascular compartments changes in UVA exposed skin by reflectance-mode confocal microscopy. *Exp Dermatol.* 2007;16(11):905–11.
18. Moyal D, et al. *In vivo* persistent pigment darkening method: proposal of a new standard product for UVA protection factor determination. *Int J Cosmet Sci.* 2007;29(6):443–9.

19. Moyal D, Chardon A, Kollias N. Determination of UVA protection factors using the persistent pigment darkening (PPD) as the end point (part 1) calibration of the method. *Photodermatol Photoimmunol Photomed.* 2000;16(6):245–9.
20. Roy CR, Gies PH, McLennan A. Sun protective clothing: 5 years of experience in Australia. *Recent Results Cancer Res.* 2002;160:26–34.
21. ISO, 24443:2013–05; Determination of sunscreen UVA photoprotection in vitro (ISO 24443:2012). German version EN ISO 24443; 2012.
22. Empfehlung der Europäischen Kommission vom 22. September 2006 über die Wirkung von Sonnenschutzmitteln und diesbezügliche Herstellerangaben. Aktenzeichen K; 2006. p. 4089.
23. Schrader K, Schrader A. Die Bestimmung der Sonnenschutzfaktoren nach COLIPA. *Aktuelle Dermatologie.* 1994;20:5.
24. Gers-Barlag H, et al. In vitro testing to assess the UVA protection performance of sun care products. *Int J Cosmet Sci.* 2001;23(1):3–14.
25. Pissavini M, et al. Validation of an in vitro sun protection factor (SPF) method in blinded ring-testing. *Int J Cosmet Sci.* 2018.
26. Reble C, et al. Evaluation of detection distance-dependent reflectance spectroscopy for the determination of the sun protection factor using pig ear skin. *J Biophotonics.* 2018;11(1).
27. Rohr M, Ernst N, Schrader A. Hybrid diffuse reflectance spectroscopy: non-erythema in vivo testing of sun protection factor. *Skin Pharmacol Physiol.* 2018;31(4):220–8.
28. Reble C, Meinke MC, Rass J. No more sunburn non-invasive LED-based measurement of the sun protection factor. *Optik Photonik.* 2018;1:4.

Chapter 19

Practical Aspects of Shampoo and Conditioner Testing



Trefor Evans

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 over-ridden 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 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

T. Evans (✉)
TRI-Princeton, Princeton, NJ, USA
e-mail: TEvans@triprinceton.org

facilitate hair manageability, but the experiments provide no information as to whether this occurs in an aesthetically pleasing manner. On the other hand, consumer testing 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 that 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 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 bodies, but coarser so-called *terminal hair* is produced on the head, while finer *vellus hair* covers most other regions. Cell growth and division occurs deep in the base of the follicle [1–4] and, when hair emerges onto the scalp, it exists in a biologically inert form (perhaps somewhat over-dramatically, it is often heard that hair consists of dead cells). For this point on, 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 inch (13 mm) 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 [1, 4–6]. For our purposes, it is sufficient to divide this structure into three regions. The outside of hair consists of a series of

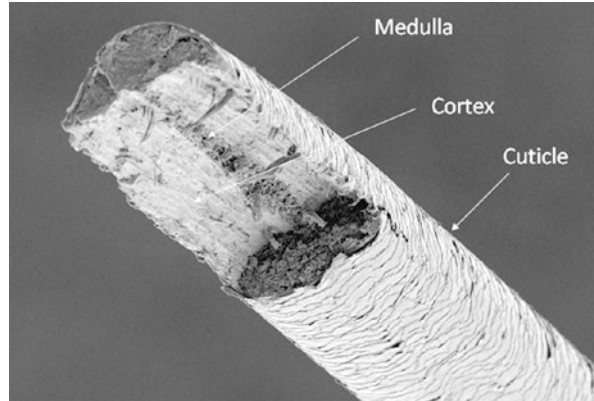
hard, protective, overlapping tile-like structures—known collectively as the cuticle. In large part, the integrity of this structure dictates the feel properties of hair. Healthy hair will possess a relatively pristine cuticle structure, but the constant wear and tear of everyday grooming gradually results 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 de-swelling 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 gray 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 has historically been paid to this region, which is generally considered equivalent to hollow space at the center of a fiber and simply serves to increase thickness. This said, recent research [7] suggests a high lipid concentration in this region and possibly leads to a rethink of this belief. A pronounced medulla tends to be 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

Fig. 19.1 Structural features of a hair fiber



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 beliefs, 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 all together. 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 heavily 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 versus 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 different habits and practices.

When considering the general properties of hair it is impossible to ignore the considerable 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 considerable 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, low humidity, winter days, when hair has diminished moisture content. As such, there is always the need to pay attention to the relative humidity when performing hair testing.

19.3 Shampoos: An Overview

With so many benefits and claims 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, buildup 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 shifted towards ethoxylated anionic surfactants (e.g., sodium laureth sulfate) which are less irritating, but also somewhat less cleansing. Today, among some, there is a belief that these sulfated surfactants are too harsh and sulfate-free shampoos have become popular (although there is little credible technical evidence to substantiate this conviction). 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 buildup 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 a product range (e.g., shampoos for normal hair, oily hair, dry hair).

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 corneorosurfametry or Zein testing.

Despite relatively few measures of technical performance, shampoos come with many aesthetic drivers. Firstly, 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 mindset 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 easily be 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 which all 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, Proctor and Gamble introduced shampoo + conditioner products (sometimes called 2-in-1s) 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 re-adsorbed 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.

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 ($\geq 20 \mu\text{m}$) 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 [8] (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, carboxer), 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 Couple Plasma Optical Emission Spectroscopy (ICP-OES) and X-Ray Fluorescence (XRF). While ICP-OES may be a 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 the low shear viscosity may be 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. An initial strategy to make clear 2-in-1s

involved the use of microemulsions—where an exceptionally small particle size (<100 nm) allows light to pass through without undergoing scattering. Today, lubricating polyquats may be used in place of silicone. 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-1s 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 limpness, and a lack of volume and body. As possibly anticipated, 2-in-1s 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 will be discussed further in the Sect. 19.5.

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, degrading 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 this 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 only exist 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 buildup 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 not leave a coated feel in either the wet or dry states.

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 have 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 ongoing 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.

19.6.1 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 panellists, 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 “non-concept” 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 [9].

19.6.2 *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

Fig. 19.2 Hair tresses of varying shape and size can be procured for testing hair care products. (Picture courtesy of International Hair Importers, New York)



Fig. 19.3 Blending hair to produce uniform tresses of fibers from a range of individuals. (Picture courtesy of International Hair Importers, New York)



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 pre-damaged 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 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.

19.7 Assessing Surface Damage: Microscopy

An old adage suggests “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 scanning electron microscopy (SEM) at magnification levels of 450 \times and 750 \times , 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.

Fig. 19.4 Scanning electron microscopy (SEM) image of a hair fiber showing cuticle abrasion

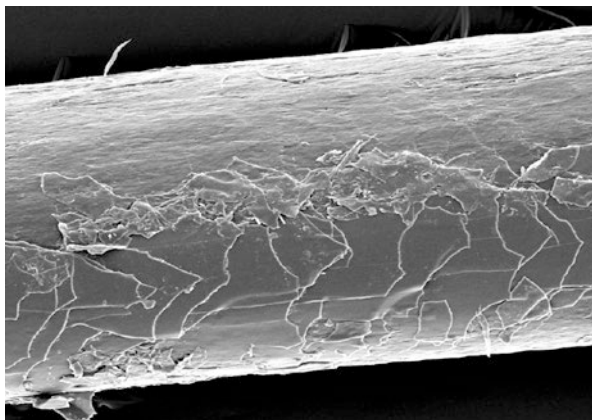
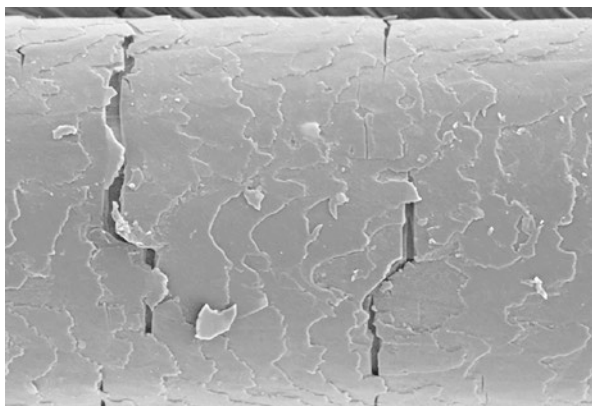


Fig. 19.5 Scanning electron microscopy (SEM) image of a hair fiber showing cracking of the cuticle



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.

A caveat to the use of this technique involves the microscopically small regions of hair fibers under examination. Therefore, within a single test specimen, it may be eminently possible to find both pristine and highly damaged regions. Responsible use of this approach must involve surveying multiple regions on a variety of fibers to identify and record representative images.

There is often the desire to ascribe numerical values to the hair surface integrity. To an extent, this can be performed by implementing a grading scale to describe the degree of surface damage. An example of a five-point grading scale is given in Table 19.1.

Fig. 19.6 Scanning electron microscopy (SEM) image showing conditioner deposits smoothing the hair surface

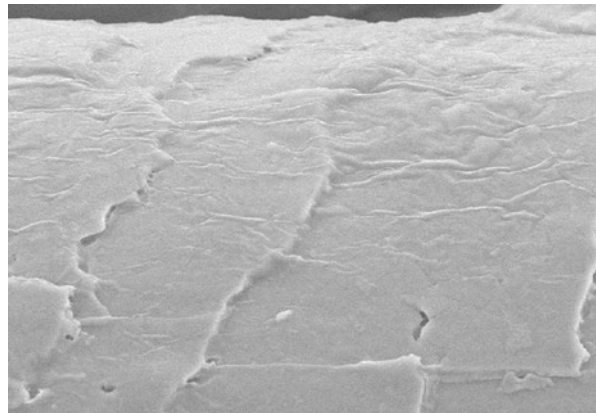


Table 19.1 Grading scale for putting a numeric value on surface damage

Ranking	Description
1	Cuticle is highly compromised. Considerable loss of cuticle layer and/or severely uplifted cuticle scales. Large “blocks” of dislodged scales visible
2	Strong signs of cuticle damage. Individual scales show significant uplift, cracking, chipping, and other signs of wear. Considerable surface debris from fragmented and dislodged cuticle material
3	Medium wear and tear of individual cuticle scales. Some uplift, cracking, and fragmentation
4	Relatively mild wear and tear at cuticle edges. Only small indication of uplift and cracking at cuticle ends
5	Largely pristine cuticle appearance. Minimal wear and tear at cuticle edges

19.8 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 instrumental combing experiments [10, 11]. 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.

19.9 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 [12]. 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 [13]. 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.

19.10 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 buildup, 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 noted initially 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*.

19.11 Shine

Technically, shine is taken to be a measure of the ability for light to reflect cleanly off a surface [14]. 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 non-polarized 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 consequently higher shines results.

19.12 Color Fade

Around 15 years ago, it become very popular for hair care products to make claims involving protection against color fade. These focus predominantly on permanently dyed hair, where there is a well-known 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 ΔL , Δa , and Δb , or as an overall color change, ΔE .

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 dyeing 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 [15].

19.13 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 eluded 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 [16].

It is likely that the word “dry” arises through an analogy with skin care, where coarseness and roughness is 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

re-educate 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.

19.14 A Note on Statistics

Human hair is a remarkably variable material. Most obviously, we note that hair comes in a broad range of sizes, shapes, and colors—but, as outlined, fundamental properties can be further impacted sizably by consumer habits and practices. These elementary properties primarily contribute to the concept of “*hair type*”—wherein it is commonplace to see hair being compartmentalized into a variety of pigeonholes (thick/thin, blonde/brunette, straight/curly, healthy/damaged, etc.).

Yet, there are also sizable differences in the properties of hair on any given head—where, for example, individual fibers exhibit considerable variability in dimensions and shape (ellipticity) [17]. Furthermore, differences in hair integrity are likely upon progressing from the root to the tip end of any given fiber due to increased wear and tear associated with advancing age. In short, any test procedure that is performed using single hair fibers is likely to yield a rather high standard deviation. Similarly, despite best efforts, the blending of hairs to produce tresses cannot yield identical entities—and again a degree of variability will result in test procedures that utilize such arrays. This all necessitates the use of a suitable number of replicate samples in any test to allow for appropriate statistical analysis to be employed.

In general, the use of 8–10 replicate samples in methods involving tress testing is usually sufficient to yield an appropriate standard error (SE)—where $SE = \text{Standard deviation} / \sqrt{\# \text{ replicates}}$. However, this number likely increases to 30–50 (or maybe more) replicates in single fiber testing. This repetition clearly adds considerable time, effort, and expense and might prompt some to cut corners and run more economical versions of the same tests. This practice, nonetheless, results in considerably lesser confidence in conclusions.

19.15 Summary

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 aforementioned 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 Ribaud, 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

1. Robbins CR. Chemical and physical behavior of human hair. 4th ed. Berlin: Springer; 2002.
2. Bouillon C, Wilkinson J. The science of hair care. 2nd ed. Boca Raton: CRC Press; 2005.
3. Jolles P, Zahn H, Hocker H. Formation and structure of human hair. Basel: Birkhauser Verlag; 1997.

4. Evans TE, Wickett RWW. Practical modern hair science. Carol Stream: Allured Books; 2012.
5. Feughelman M. Mechanical properties and structure of alpha-keratin fibers. Randwick: University of New South Wales Press; 1997.
6. Johnson DH. Hair and hair care. New York: Marcel Dekker Inc; 1997.
7. Zhang G, Senak L, Moore DJ. Measuring changes in chemistry, composition, and molecular structure within hair fibers by infrared and Raman spectroscopic imaging. *J Biomed Opt.* 2011;16(5):056009.
8. Goddard ED, Gruber JV. Principles of polymer science and technology in cosmetics and personal care. New York: Marcel Dekker Inc; 1999.
9. Meilgaard MC, Civille GV, Carr BT. Sensory evaluation techniques. 4th ed. Boca Raton: CRC Press; 2006.
10. Garcia ML, Diaz J. Combability measurements on hair. *J Soc Cosmet Chem.* 1976;27:379–98.
11. Evans TA. Evaluating hair conditioning with instrumental combing. *Cosmet Toilet.* 2011;126(8):558–63.
12. Evans TA. Measuring hair strength, part 1: stress-strain curves. *Cosmet Toilet.* 2013;128(8):590–4.
13. Evans TA. Measuring hair strength, part 2: fiber breakage. *Cosmet Toilet.* 2013;128(12):854–9.
14. Evans TA. Equating the measurement of hair shine. *Cosmet Toilet.* 2016;131(1):28–34.
15. Evans TA. Quantifying hair color fading. *Cosmet Toilet.* 2015;130(1):30–5.
16. Evans TA. Measuring the water content of hair. *Cosmet Toilet.* 2014;129(2):64–9.
17. Evans TA. New focus for ethnic hair. *Cosmet Toilet.* 2019;134(6):35–41.

Chapter 20

Antiperspirants and Deodorants



Razvigor Darlenski and Joachim W. Fluhr

Abbreviations

TEWL Transepidermal water loss

Core Messages

- Antiperspirants and deodorants are two separate entities with regard to their definition, mechanism of action, and legislation.
- The major methods used for efficacy proof of antiperspirants include gravimetric, biophysical, and imprint casting assessment 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, i.e., 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.

R. Darlenski (✉)

Department of Dermatology and Venereology, Faculty of Medicine, Medical University, Sofia, Bulgaria

J. W. Fluhr

Department of Dermatology, Charité University Clinic, 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 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 the social adaptation. In addition, certain diseases such as the hyperhidrosis can be the cause for excessive sweating decreasing the quality of life in the suffering patients.

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 everyday cosmetic care. Beyond developing medications and procedures which decrease sweat gland activity, the market for cosmetics with antiperspirant/deodorant activity is constantly growing.

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 inhibit 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) [1]. Currently, aluminum as one of the active ingredients of deodorants/antiperspirants are under discussion due to potential induction of breast cancer. Two consecutive studies found no elevated risk of antiperspirants use in the pooled risk for breast cancer (odds ratio 0.40, 95% confidence interval 0.35–0.46) [2].

In the USA, antiperspirants are classified as over-the-counter medications (while deodorants are accepted as cosmetics) and a lists of allowed active ingredients exist

[3]. 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 visual methods is the visualization of the 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 scratch-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 increases the risk for misinterpretation of the results (Fig. 20.1).

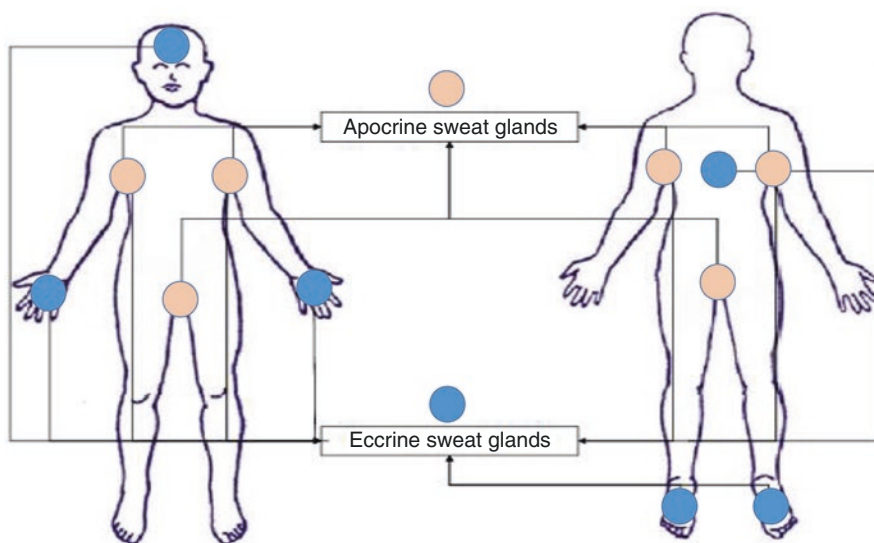


Fig. 20.1 Anatomical skin sites with higher density of apocrine (axillae, groin—orange color) and eccrine (palms, soles, forehead, back—blue color) sweat glands

Casting methods are based on the application of a hydrophobic material (e.g., silicone rubber and plastic solutions) [5]. 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 the casting methods is the formation of air bubbles in the replica and should be taken into account.

Gravimetric methods are the “golden standard” for evaluation of the quality of the secreted sweat and its reduction by the antiperspirant. 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 a skin area 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 the greater discriminative properties [6]. 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.

Microbiological assessment of the treated site vs. untreated control is also applied with respect to the antibacterial activity of the antiperspirant [7].

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 the cosmetic product.

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 maximally two products can be tested in a single panelist. The number of subjects requested 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 [8].

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 screening general medical checkup is performed before entering the study. The presence of systemic disease (infections, HIV, AIDS, autoimmune disorders, neoplasms, and neurological, hematological, and metabolic disorders) as well as skin conditions (psoriasis, eczema, bullous disorders, tinea, and bacterial infections) should be considered as exclusion factors 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 the study. Issues such as the written informed consent, inclusion/exclusion criteria, and overrecruitment focused in other chapters of this book are valid and should be considered when testing antiperspirants.

A minimum 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 ± 1.1 °C and relative humidity 35–40%. A full recording of the microenvironmental conditions on regular intervals is necessary in the latter.

During the sweat collection the subject should not experience psychological stress. Subjects wear cotton T-shirts which makes the 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 (Webril 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 the trained personnel. After the end of this period each pad is placed in capped and marked plastic vial, and then is weighed. The procedure is repeated with another pair of cotton pads for the rest 20 min. When the test is performed in ambient environmental conditions a 3- to 5-h duration of the 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 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, e.g., twice or four times daily.

Sweat collection is performed generally 24 h after the last product application. However, different time points can be selected, e.g., 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) [1]. 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. In 1998, a protocol for testing during early product development was introduced [6, 9]. The design allows direct comparisons of the efficacy of up to eight test formulations within 1 week.

Study Population

Each panel includes 20–22 subjects. The general considerations described 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 to air for 5 min before covering with occlusive patches. The conditions during occlusion and the length of occlusion varied according to the study protocol. Treatments are 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 [5, 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 formulations development. However, there is no sufficient evidence for efficient 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 is not disclosing the complexity of malodor composition as great interindividual variability of the chemical composition of malodor has been demonstrated [10]. Sensory assessment, i.e., sniff test is accepted as a standard evaluation procedure for deodorants *in vivo* testing.

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 the one axilla is treated with the test formulation, and the contralateral (control) is left untreated or a placebo is applied simultaneously to the test product application.

Study Population

Thirty to 60 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 non-antibacterial soap is provided for cleansing the axillary region.

Test Procedure

A controlled washing similar to the one performed 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 consist 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 [11]. 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 the 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 these 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 demonstrated 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, i.e., 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 of noninvasive, cost- and time-effective and precise methods and equipment is the next step in this scope.

References

1. Pierard GE, Elsner P, Marks R, Masson P, Paye M. EEMCO guidance for the efficacy assessment of antiperspirants and deodorants. *Skin Pharmacol Appl Skin Physiol.* 2003;16(5):324–42.
2. Allam MF. Breast cancer and deodorants/antiperspirants: a systematic review. *Cent Eur J Public Health.* 2016;24(3):245–7.
3. Department of Health and Human Services FaDA. Antiperspirant drug products for over-the-counter human use: tentative final monographs. vol 47, p 36491–36505. Fed Reg 1982.
4. Council Directive Nr 93/35/CEE. 1993.
5. Keyhani R, Scheede S, Thielecke I, Wenck H, Schmucker R, Schreiner V, et al. Qualification of a precise and easy-to-handle sweat casting imprint method for the prediction and quantification of anti-perspirant efficacy. *Int J Cosmet Sci.* 2009;31(3):183–92.
6. Bielefeldt S, Frase T, Gassmueller J, Bielefeldt S. Efficacy testing of antiperspirants during early product development. *SÖFW J.* 2000;126(8):12–6.
7. Celleno L, Mastropietro F, Tolaini MV, Pigatto PD. Clinical evaluation of an antiperspirant for hyperhidrosis. *G Ital Dermatol Venereol.* 2019;154(3):338–41.
8. Brandt M, Bielfeldt S, Springmann G, Wilhelm KP. Influence of climatic conditions on antiperspirant efficacy determined at different test areas. *Skin Res Technol.* 2008;14(2):213–9.
9. Wunderlich O, Frase T, Hughes-Formella B. A screening technique for antiperspirant testing. *Cosmet Toilet.* 2003;118(12):59–62.
10. Baxter PM, Reed JV. The evaluation of underarm deodorants. *Int J Cosmet Sci.* 1983;5(3):85–95.
11. Maxeiner B, Ennen J, Rützel-Grünberg S, Traupe B, Wittern K-P, Schmucker R, et al. Design and application of a screening and training protocol for odour testers in the field of personal care products. *Int J Cosmet Sci.* 2009;31(3):193–9.

Chapter 21

Hair Growth



Tobias W. Fischer

Core Messages

- 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 non-invasive, 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 non-invasive, 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 trichoscopy using folliscope 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 carefully respect test hypothesis, practicability, time consumption, variability, and accuracy.

T. W. Fischer (✉)

Department of Dermatology and Venerology, Kepler University Hospital, Linz, Austria
e-mail: Tobias.Fischer@kepleruniklinikum.at

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.

21.1 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 [1]. 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 cosmetic agents on hair growth. Great advances have been made during the recent decades in the methodology of hair growth trials in dermatology and cosmetology [1–3]. 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:

- *Non-invasive*, e.g., global photographs (GP), daily hair counts (DHC), hair wash test (HWT), hair pull test (HPT), hair weighing (HW), phototrichogram (PTG), folliscope (FS) 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 also be 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, HW, TG, computer-assisted PTG, or ELMH might be used as state-of-the-art techniques for objective clinical studies [1–5]. 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 [6]. For research purposes only, optical coherent tomography, EM, and confocal LSM are optional tools [7]. How these methods are used and how a study for testing hair loss is technically and logistically planned is described in the following.

21.2 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 [8, 9] while others have observed significant effects already at 6 months [10–12]. Volunteers/patients visits should be scheduled for every 3 months or every 6 months.

21.3 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 reflection: 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 (DHC)
- Hair wash test (HWT)
- Hair pull test (HPT)
- Dermoscopy

21.4 Clinical Grading Systems

Clinical grading systems have been established mainly for androgenetic alopecia (AGA) of men and women. For grading male pattern androgenetic alopecia, 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 [13, 14]. In women, the female-pattern AGA is graded in a three-point scale after Ludwig [15]. 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 [7]. The grading system should be applied before every study to clearly characterize the study population which should be investigated under treatment, and it is recommended to include the classification stages in the inclusion and exclusion criteria of a study.

21.5 Daily Hair Count

This is a basic test performed by the patient him- or herself 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 semiquantitative 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.

21.6 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 or her 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 [16]. 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 [7].

21.7 Hair Pull Test

This test allows to obtain a semiquantitative 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 [7], 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). A quite recent validation of the hair pull test performed with 181 participants revealed that a lower rate of pulled hair is found: in 97.2% of the participants, two hairs or less could be pulled. Also a still often practiced recommendation of not having washed or brushed the hair for 5 days before performance of pull test was shown not to be relevant. Hair pull tests were also similar in Caucasian-, Asian-, and Afro-textured hair [17]. This study quite much confirms the author’s interpretation that in a range of none to three hairs the pull test is “negative,” i.e., normal/physiologic, while everything above three pulled hairs is considered “positive,” i.e., pathologic.

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 and parietal scalp area. For AD, in contrast, the pull test most often is positive in temporal, parietal, and occipital area.

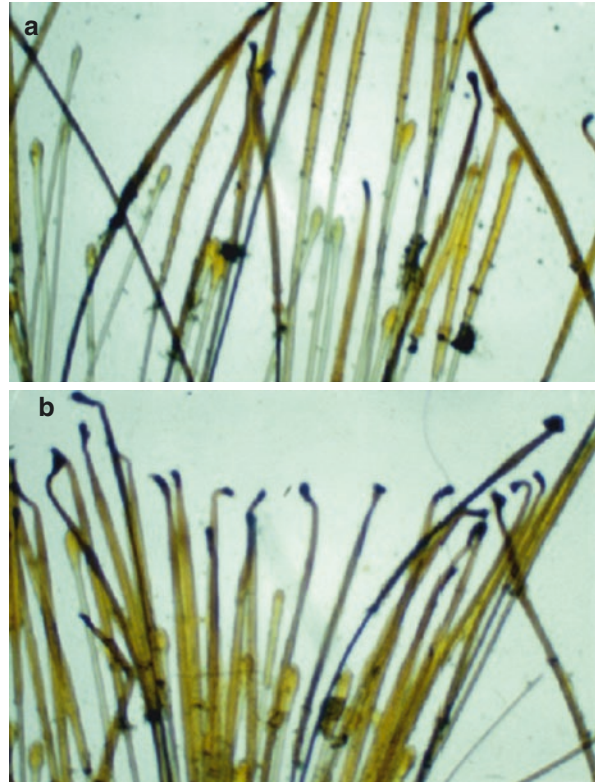
21.8 Hair Weighing

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 subsample 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 subsample showed no significant change [4]. Additionally, it was observed that cumulative hair weight leads 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 [4]. However, the investigation process is time-consuming and its use has to be critically considered in balancing the advantages and disadvantages of the method.

21.9 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 non-affected reference area or to proof diffuse alopecia (Fig. 21.1a, b). In alopecia areata, one TG is taken at the edge of an

Fig. 21.1 Trichogram results from fronto-temporal (a) and occipital (b) regions in a man with androgenetic alopecia with clearly visible high rate of transparent club-like telogen hairs in the fronto-temporal area while showing normal anagen-/telogen-ratio in the occipital part



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 one hand with an immediate strong force within half a second, directed away from the patient's head, while holding with the other hand his or her 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 [11]. 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 drawback 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 non-invasive methods to assess anagen/telogen rate are the PTG and the epiluminescence microscopy (ELM) of hair which are described in the following.

21.10 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 $>30 \text{ mm}^2$ and then counted and measured. Four parameters can be assessed: hair count per mm^2 , anagen rate, hair shaft length, and diameter [7]. 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.

21.11 Phototrichogram

The PTG is a further development to classical TG and has been first proposed by Saitoh in 1970 [18] and then been further developed by Bouhanna in 1988 [19] and Van Neste [20]. The PTG is a non-invasive, 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 [2, 21]. The reference area to be assessed is a 0.50 cm^2 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 $40 \mu\text{m}$ in diameter (vellus hair) and percentage of thick hairs with more than $40 \mu\text{m}$ in diameter (terminal hairs). The normal measurement values for a male adult have been measured as follows (vertex): density $204 \pm 10 \text{ hair/cm}^2$, telogen rate $17.8 \pm 2.8\%$, growth rate $0.35 \pm 0.03 \text{ mm/day}$, mean anagen diameter $76 \pm 5 \mu\text{m}$, percentage of fine hair $9.2 \pm 1.8\%$, and percentage of thick hair $90.8 \pm 2.1\%$ [21].

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 interindividual 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 variances 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.

Regarding the two abovementioned methods (UATG and PTG), a study with 12 volunteers has shown that there are similar assessments of anagen hair in UATG and PTG, while there was an underestimation of 181 vs. 237 hairs/cm² of total hair density by the PTG method [5]. It was further found that hair diameter measurements from the PTG were not reliable.

21.12 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 [5].

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 [3]. 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 [3].

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 research. 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 interobserver variability [22]. Nevertheless, the method is used and suitable for clinical hair studies.

21.13 Trichoscopy

Trichoscopy is generally the investigation with a magnifying dermatoscope used for hair shafts, hair ostiae, and peri- and interfollicular scalp skin. It can be a simple handheld manual trichoscope or a computerized digital- or video-trichoscope. The magnification of the first is usually $\times 10$ or $\times 20$, while the latter can be up to $\times 1000$ [23]. The trichoscope in a simple handheld manner can help to judge hair shaft texture, dryness, and scaling and give a rough estimate of hair follicle density. It further is a useful tool to differentiate between diagnoses, e.g., scaring vs. non-scaring alopecia, androgenic alopecia, alopecia areata, telogen effluvium, trichotillomania, and tinea capitis. Trichoscope in connection with a computer and software can be used to assess real hair density per cm^2 , anagen/telogen hair ratio, vellus/terminal hair ratio, hair diameter, and cumulative hair diameter. It is the basic technical principal for the methods described in the following, i.e., folliscope and epiluminescence microscopy of hair.

21.14 Folliscope

The folliscope is a quite simple, small USB-based handheld device in combination with a computer and a screen to assess hair density and thickness [24]. Technically, it is an automated digital phototrichogram system with a high definition microscopic camera offering a magnification of up to $\times 400$. It visualizes the hair and instantly automatically measures hair density, hair shaft diameter, and proportions of different hair types such as terminal and vellus hairs. The advantage over classical trichogram is that the hair does not need to be plucked or clipped and it is used on dry scalp. An immersion fluid or hair dye is not needed. Since it is connected to a computer, the values can be stored and compared over time with previous and consecutive ones. It has shown to be reliable in assessing hair density and diameter in a study with 32 Caucasian and 13 Asian women carrying female pattern hair loss. A positive response to treatment with minoxidil in 16 of them was also documented [24].

21.15 Epiluminescence Microscopy of Hair (TrichoScan)

In search for more reliable and interobserver-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 [2]. Trichoscan was developed to assess all important parameters of hair growth such as hair follicle density (n/cm^2), hair shaft diameter (μm), cumulative hair diameter, mean daily growth rate (mm/day), anagen and telogen hair rate (%), as well as vellus and terminal hair density (%). The parameters are registered *in situ* and analyzed independently from the patient's presence by a computer-assisted digital image scan and algorithmic analysis (Fig. 21.2) [2].

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

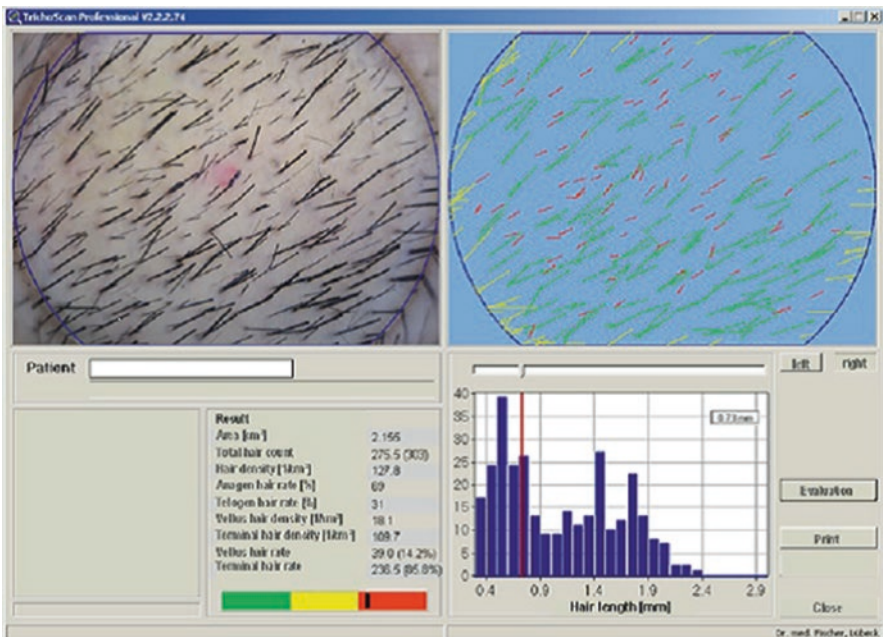


Fig. 21.2 Software-window of the TrichoScan showing the digitalized image of clipped and dyed hair (upper left), the respective hair scan (upper right), and results of measurements for hair density (n/cm^2), anagen/telogen rate, and vellus and terminal hair rate (lower part)

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 needle 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 cm²) and 40-fold (analyzed area: 0.23 cm²) magnification (Fotofinder DERMA, TeachScreen Software GmbH, Bad Birnbach, Germany) or a digital camera. 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 cm², 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 µm cannot be analyzed, whereas with higher resolution 7.0 megapixel cameras hairs with 6 µm 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 [2].

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 [25, 26]. It can be used for studying AGA or 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.

21.16 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 [27].

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 swung 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 used in many clinical studies, e.g., the study to assess the efficacy of finasteride being the most prominent one [8]. 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 process 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 (e.g. beginning, after 3 or 6 months, at the end) is rated using a seven-point scale from greatly decreased -3 to greatly increased +3 [27]. 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

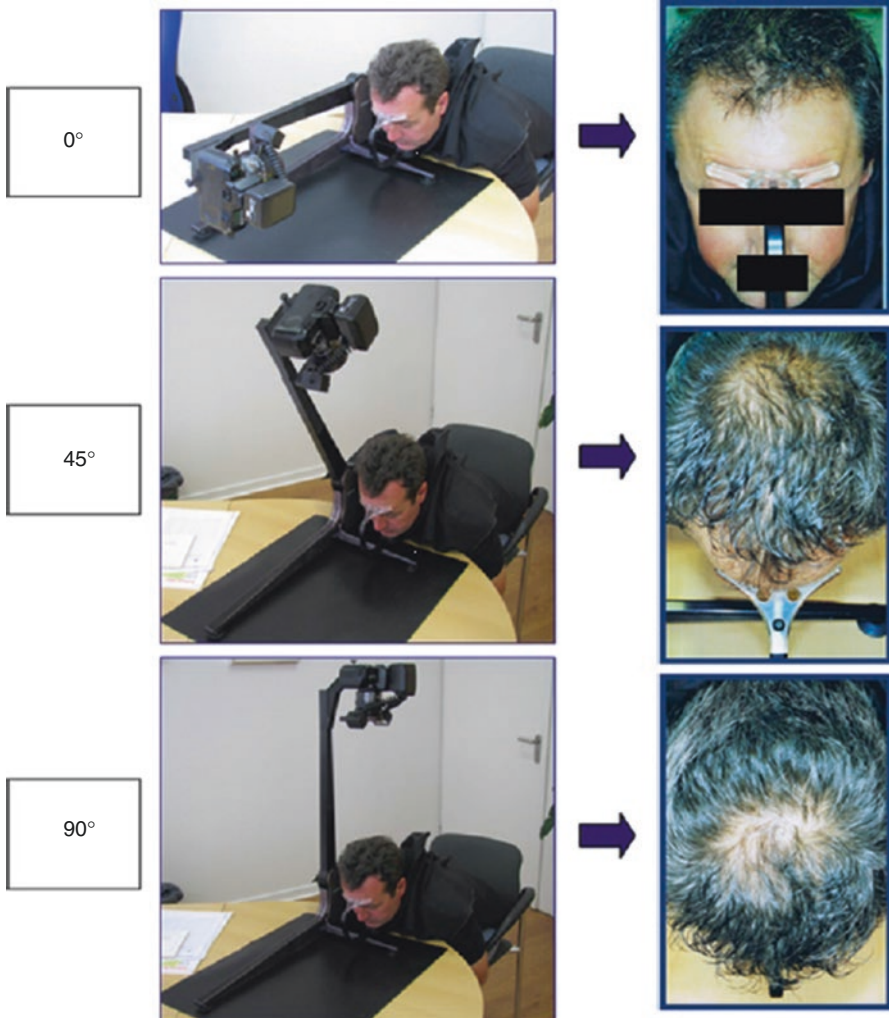


Fig. 21.3 Global photographs in three positions (0°–45°–90°) to assess the global hair appearance in the fronto-temporal, the upper-head and the vertex area

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 abovementioned 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.

21.17 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 also be excellently used for the assessment of hair shaft diameter, including the inner hair variation of diameter and shape *in vivo* as well as *ex vivo* [28]. The images 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 assessing hair shaft surface, consistency, elasticity, shape, and ultrastructure.

21.18 Electron Microscopy of Hair

Electron microscopy offers images of tissues in highest resolutions and can be used also for hair. With this technique, the ultrastructure 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, and 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 [7].

21.19 Confocal Laser Scanning Microscopy of Hair

Confocal LSM is a non-invasive 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 [7]. 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. Following the routes of penetration using fluorescence-labeled substances allow to perform also 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 OCH and EMH.

21.20 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 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.

Table 21.1 List of companies providing hair evaluation techniques

Technique	Company (name, address)	Internet-reference
Trichogram	Galderma Laboratorium GmbH, Georg-Glock-Strasse 8, 40474 Düsseldorf, Germany	www.galderma.de
Trichoscope	Canfield Scientific, 4 Wood Hollow Road, Parsippany, NJ 07054, USA	www.canfieldsci.com
	Firefly Global, 464 Common Street, Suite 281, Belmont, MA 02478, USA	https://fireflyglobal.com
Folliscope	IDCP, B.V., Manuscriptstraat 12–14, 1321 NN Almere, The Netherlands	www.idcp.eu
	C Cube Advanced Technologies, 5/1, Survey No. 16/1, Chowdeshwari Road, Arehalli, 4th Cross, Hosakerehalli, Bengaluru—560061, Karnataka, India	www.indiamart.com/ ccubeadvancedtechnologies/
Epiluminescence microscopy of hair/ TrichoScan	Tricholog GmbH, In den Eschmatten 24, 79117 Freiburg i. Brsg., Germany	www.tricholog.de
	TeachScreen Software GmbH, Industriestr. 12, 84364 Bad Birnbach, Germany	www.fotofinder.de
	Moser, Wahl GmbH, Roggenbachweg 9, 78089 Unterkirnach, Germany	www.moser-online.com
Global photographs	Canfield Scientific, 4 Wood Hollow Road, Parsippany, NJ 07054, USA	www.canfieldsci.com

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 and 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 from experienced vs. inexperienced observers [29].

The PTG/CE-PTG and the ELMH are 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 [26, 30, 31].

The combination of optical microscopy and computer analysis has been already developed quite early [32, 33]; 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 interobserver correlation of 91 and 97%, respectively [2]. 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 the author's knowledge, with new systems using digital cameras of higher pixel size [34].

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 [29] and later with the use of immersion oil and digital contrast enhancement [35]. 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 [3], 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 considerable higher in the CE-PTG than in the histology transverse sections. A comparative study between TrichoScan and conventional PTG was performed in ten patients with AGA to evaluate variability, validity, and reliability of both methods [30]. The study showed an excellent correlation coefficient 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 [30]. 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 the author's own experience that experienced hands can perform a TrichoScan imaging and analysis with ready results within 8–12 min "hands on" [34]. 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 [26]. 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% [30] or up to 9% [2] and therein also lower compared to 2.71–12.95% in normal PTG [30]. 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.

Another comparison was performed between epiluminescence microscopy of hair (ELMH) (TrichoScan), reflectance confocal microscopy (RCM), and optical coherence tomography (OCT). The study was performed with measurement of hair growth after removal of the hair shaft from the lower leg of six female volunteers with a shaving system. For 5 days, hair follicle infundibula and the length of the growing hairs were measured. All three systems showed reliable results, while ELMH was best suited for hair measurement, and RCM in the analysis of hair follicles. The OCT proved to be suitable as additional tool for the evaluation of the hair follicle and length [36].

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

1. Pierard GE, Pierard-Franchimont C, Marks R, Elsner P. EEMCO guidance for the assessment of hair shedding and alopecia. *Skin Pharmacol Physiol*. 2004;17:98–110.
2. Hoffmann R. TrichoScan: combining epiluminescence microscopy with digital image analysis for the measurement of hair growth in vivo. *Eur J Dermatol*. 2001;11:362–8.
3. Van Neste DJ. Contrast enhanced phototrichogram (CE-PTG): an improved non-invasive technique for measurement of scalp hair dynamics in androgenetic alopecia—validation study with histology after transverse sectioning of scalp biopsies. *Eur J Dermatol*. 2001;11:326–31.
4. Price VH, Menefee E. Quantitative estimation of hair growth. I. Androgenetic alopecia in women: effect of minoxidil. *J Invest Dermatol*. 1990;95:683–7.
5. Rushton DH, de Brouwer B, de Coster W, van Neste DJ. Comparative evaluation of scalp hair by phototrichogram and unit area trichogram analysis within the same subjects. *Acta Derm Venereol*. 1993;73:150–3.
6. Whiting DA, Waldstreicher J, Sanchez M, Kaufman KD. Measuring reversal of hair miniaturization in androgenetic alopecia by follicular counts in horizontal sections of serial scalp biopsies: results of finasteride 1 mg treatment of men and postmenopausal women. *J Invest Dermatol Symp Proc*. 1999;4:282–4.
7. Hillmann K, Blume-Peytavi U. Diagnosis of hair disorders. *Semin Cutan Med Surg*. 2009;28:33–8.
8. Kaufman KD, Olsen EA, Whiting D, Savin R, DeVillez R, Bergfeld W, Price VH, Van Neste D, Roberts JL, Hordinsky M, Shapiro J, Binkowitz B, Gormley GJ. Finasteride in the treatment of men with androgenetic alopecia. Finasteride Male Pattern Hair Loss Study Group. *J Am Acad Dermatol*. 1998;39:578–89.
9. Lucky AW, Piacquadro DJ, Ditre CM, Dunlap F, Kantor I, Pandya AG, Savin RC, Tharp MD. A randomized, placebo-controlled trial of 5% and 2% topical minoxidil solutions in the treatment of female pattern hair loss. *J Am Acad Dermatol*. 2004;50:541–53.
10. Blume-Peytavi U, Kunte C, Krisp A, Garcia Bartels N, Ellwanger U, Hoffmann R. Comparison of the efficacy and safety of topical minoxidil and topical alfatradiol in the treatment of androgenetic alopecia in women. *J Dtsch Dermatol Ges*. 2007;5:391–5.
11. Fischer TW, Burmeister G, Schmidt HW, Elsner P. Melatonin increases anagen hair rate in women with androgenetic alopecia or diffuse alopecia: results of a pilot randomized controlled trial. *Br J Dermatol*. 2004;150:341–5.
12. Lengg N, Heidecker B, Seifert B, Trueb RM. Dietary supplement increases anagen hair rate in women with telogen effluvium: results of a double-blind, placebo-controlled trial. *Therapy*. 2007;4:59–65.
13. Hamilton JB. Patterned loss of hair in man; types and incidence. *Ann N Y Acad Sci*. 1951;53:708–28.
14. Norwood OT. Male-pattern baldness. Classification and incidence. *South Med J*. 1975;68:1359–70.
15. Ludwig E. Classification of the types of androgenetic alopecia (common baldness) occurring in the female sex. *Br J Dermatol*. 1977;97:247–54.

16. Rebora A, Guarrera M, Baldari M, Vecchio F. Distinguishing androgenetic alopecia from chronic telogen effluvium when associated in the same patient: a simple non-invasive method. *Arch Dermatol*. 2005;141:1243–5.
17. McDonald KA, Shelley AJ, Colantonio S, Beecker J. Hair pull test: evidence-based update and revision of guidelines. *J Am Acad Dermatol*. 2017;76:472–7.
18. Saitoh M, Uzuka M, Sakamoto M. Human hair cycle. *J Invest Dermatol*. 1970;54:65–81.
19. Bouhanna P. The tractiophototrichogram, an objective method for evaluating hair loss. *Ann Dermatol Venereol*. 1988;115:759–64.
20. Van Neste DJJ, Dumortier M, De Coster W. Phototrichogram analysis: technical aspects and problems in relation with automated quantitative evaluation of hair growth by computer-assisted image analysis. In: Van Neste DJJ, Lachapelle JM, Antoine JL, editors. *Trends in human hair growth and alopecia research*. Dordrecht: Kluwer; 1989. p. 155–65.
21. Friedel J, Will F, Grosshans E. Phototrichogram. Adaptation, standardization and applications. *Ann Dermatol Venereol*. 1989;116:629–36.
22. Hoffmann R. TrichoScan. A new instrument for digital hair analysis. *Hautarzt*. 2002;53:798–804.
23. Lacarrubba F, Micali G, Tosti A. Scalp dermoscopy or trichoscopy. *Curr Probl Dermatol*. 2015;47:21–32.
24. Lee BSL, Chan JL, Monselise A, McElwee K, Shapiro J. Assessment of hair density and caliber in Caucasian and Asian female subjects with female pattern hair loss by using the Folliscope. *J Am Acad Dermatol*. 2012;66:166–7.
25. Hoffmann R. TrichoScan: a novel tool for the analysis of hair growth in vivo. *J Investig Dermatol Symp Proc*. 2003;8:109–15.
26. Hoffmann R. TrichoScan, a GCP-validated tool to measure hair growth. *J Eur Acad Dermatol Venereol*. 2008;22:132–4; author reply 134–135.
27. Canfield D. Photographic documentation of hair growth in androgenetic alopecia. *Dermatol Clin*. 1996;14:713–21.
28. Lademann J, Shevtsova J, Patzelt A, Richter H, Gladkova ND, Gelikonov VM, Gonchukov SA, Sterry W, Sergeev AM, Blume-Peytavi U. Optical coherent tomography for in vivo determination of changes in hair cross section and diameter during treatment with glucocorticosteroids—a simple method to screen for doping substances? *Skin Pharmacol Physiol*. 2008;21:312–7.
29. Rushton DH, Unger WP, Cotterill PC, Kingsley P, James KC. Quantitative assessment of 2% topical minoxidil in the treatment of male pattern baldness. *Clin Exp Dermatol*. 1989;14:40–6.
30. Gassmueller J, Rowold E, Frase T, Hughes-Formella B. Validation of TrichoScan technology as a fully-automated tool for evaluation of hair growth parameters. *Eur J Dermatol*. 2009;19:224–31.
31. Van Neste D, Trueb RM. Critical study of hair growth analysis with computer-assisted methods. *J Eur Acad Dermatol Venereol*. 2006;20:578–83.
32. Barman JM, Pecoraro V, Astore I. Method, technic and computations in the study of the trophic state of the human scalp hair. *J Invest Dermatol*. 1964;42:421–5.
33. Hayashi S, Miyamoto I, Takeda K. Measurement of human hair growth by optical microscopy and image analysis. *Br J Dermatol*. 1991;125:123–9.
34. Hoffmann R. Trichoscan: what is new? *Dermatology*. 2005;211:54–62.
35. Van Neste DJJ, Dumortier M, de Brouwer B, de Coster W. Scalp immersion proxigraphy (SIP): an improved imaging technique for phototrichogram analysis. *J Eur Acad Derm Venereol*. 1992;1:187–91.
36. Kuck M, Schanzer S, Ulrich M, Garcia Bartels N, Meinke MC, Fluhr J, Krah M, Blume-Peytavi U, Stockfleth E, Lademann J. Analysis of the efficiency of hair removal by different optical methods: comparison of Trichoscan, reflectance confocal microscopy, and optical coherence tomography. *J Biomed Opt*. 2012;17:101504.

Chapter 22

Sensory Perception



Gregor B. E. Jemec

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 [1]. It is easy to imagine the lack of uniformity in this, ranging from persons with established skin diseases, from the depressed who somatize their disease, to persons whose skin genuinely appear to be hyperreactive in a nonspecific manner. Generally, persons who report sensitive skin appear to have a subjective hyperreactivity to a wide range of nonspecific stimuli of their skin, including not only cosmetics but also physical influences such as 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 behavior and life of those it affects.

This chapter was not updated from the first edition. The overall content is still relevant, and thus it was decided to keep the chapter in its current form.

G. B. E. Jemec (✉)

Department of Dermatology, Health Sciences Faculty, Roskilde Hospital, University of Copenhagen, Roskilde, Denmark
e-mail: gbej@regionsjaelland.dk

22.2 Sensitive Skin

No unified definition of sensitive skin exists, but a number of key characteristics have been identified [1, 2]:

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 [3]. 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 [4, 5]. 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 [2, 6]. 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 and 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 behavior, 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 noninvasively are often used. Noninvasive biophysical measurements such as transepidermal water loss (TEWL) and colourimetry 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 [7]. 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 [2, 3]. 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 [8]. 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 panelists were found to react to all four substances [9]. 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 [10], indicating that panelists 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, pretested 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 [11].
- ### 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 sulfate (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 [12]. 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 [13].

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 nonspecific 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 [14].

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 [14].

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 unrecognized skin disease or development of allergic contact dermatitis.

References

1. Kligman AM, Sadiq I, Zhen X, Brosby M. Experimental studies on the nature of sensitive skin. *Skin Res Technol.* 2006;12:217–22.
2. Farage MA, Katsarou A, Maibach HI. Sensory, clinical and physiological factors in sensitive skin: a review. *Contact Dermatitis.* 2006;455:1–14.

3. Marriott M, Whittle E, Basketter DA. Facial variations in sensory responses. *Contact Dermatitis*. 2003;49:227–31.
4. Farage M. Perceptions of sensitive skin: changes in perceived severity and associations with environmental causes. *Contact Dermatitis*. 2008;59:226–32.
5. Willis CM, Shaw S, de Lacharriere O, Baveral M, Reiche L, Jourdain R, Bastien P, Wilkinson JS. Sensitive skin, and epidemiological study. *Br J Dermatol*. 2001;145:258–61.
6. Holm EA, Esmann S, Jemec GBE. Does visible atopic dermatitis affect quality of life more in women than in men? *Gend Med*. 2004;1:125–30.
7. Frosch P, Kligman AM. Methods for appraising the sting capacity of topically applied substances. *J Soc Cosmet Chem*. 1977;28:197–209.
8. Daugthers HL, Chew AL, Maibach HI. Sensitive skin. In: Loden M, Maibach HI, editors. *Dry skin and moisturizers*. Boca Raton: CRC Press; 2006.
9. Marriott M, Holmes J, Peters L, Cooper K, Rowson M, Basketter DA. The complex problem of sensitive skin. *Contact Dermatitis*. 2005;53:93–9.
10. Besne DC, Breton L. Effect of age and anatomical site on sensivity of sensory innervation of human epidermis. *Arch Dermatol*. 2002;138:1445–50.
11. Kim SJ, Lim SU, Won YH, An SS, Lee EY, Moon SJ, Kim J. The perception of threshold measurement can be a useful tool for evaluating sensitive skin. *Int J Cosmet Sci*. 2008;30:333–7.
12. Agner T, Serup J. Sodium lauryl sulphate for irritant patch testing—a dose-response study using bioengineering methods for determination of skin irritation. *J Invest Dermatol*. 1990;95:543–7.
13. Agner T, Serup J. Quantification of the DMSO-response—a test for assessment of sensitive skin. *Clin Exp Dermatol*. 1989;14:214–7.
14. Cooper K, Marriott M, Peters L, Basketter D. Stinging and irritating substances: their identification and assessment. In: Loden M, Maibach HI, editors. *Dry skin and moisturizers*. Boca Raton: CRC Press; 2006.

Chapter 23

Novel Methods for In Vivo Skin Structure Visualization



Giovanni Pellacani, Stefania Guida, and Silvana Ciardo

Core Message

Nowadays, exploring the structure of the skin is possible through noninvasive imaging techniques that can be used for diagnostic purposes not only in dermatology but also in the cosmetic field. Interestingly, structural information of the skin may give important insights into skin physiology, diagnosis, and data concerning the outcome of different cosmetic conditions under study. Techniques discussed in the chapter are:

- Reflectance confocal microscopy (RCM) that enables in vivo exploration of horizontal layers of the skin at a quasi-histological resolution, reaching a depth of 250 μm , and can provide information concerning pigmentation, hydration, irritation, aging, and skin rejuvenation techniques.
- Optical coherence tomography (OCT), providing horizontal and vertical sections of the skin reaching a depth of up to 1–2 mm, with a lower resolution as compared to RCM, providing information on epidermis, collagen, and vessels density.
- Multiphoton microscopy (MPM) produces high-resolution images of the skin structure and may provide additional functional information. However, it is limited by the long time to acquire images. Keratinocytes, pigments, collagen, and elastin can be explored with this technique.

G. Pellacani (✉) · S. Guida · S. Ciardo
Dermatology Department, University of Modena and Reggio Emilia, Modena, Italy

23.1 Introduction

Skin is the largest organ in humans. It is a sophisticated and dynamic organ, acting as a barrier between the internal environment and the world outside as well as being involved in many functions: protection, regulation, and sensation [1, 2].

Exploring functions of the skin indirectly is possible through some noninvasive devices, such as corneometry, trans-epidermal water loss, and sebometry. However, new tools are now available in order to explore skin structure, with an impact on skin function, at different levels.

New diagnostic tools such as *in vivo* reflectance confocal microscopy (RCM) [3, 4], optical coherence tomography (OCT) [5], and multiphoton microscopy (MPM) [6] represent noninvasive methods that offer a quasi-histologic resolution of the skin in real time. The main advantage is the possibility to explore healthy skin without any skin biopsy, therefore avoiding scars. This is of peculiar importance when exploring lesions in aesthetic sites such as the face.

In this chapter we describe all the mentioned techniques (RCM, OCT, MPM) and how the skin is visualized and provide insights for the applications of these tools to different skin conditions with a cosmetic impact.

23.2 Reflectance Confocal Microscopy

Reflectance confocal microscopy (RCM) is a noninvasive technique enabling horizontal visualization of the skin with good contrast and high resolution, providing cytologic and architectural details [3].

It has been proven as an excellent add-on tool for diagnostic purposes in dermatology as well as for the analysis of healthy skin compartments, since RCM provides an optical “histological” biopsy of the living tissue in a totally noninvasive manner. Furthermore, the examination of a given skin area can be repeated over time, permitting the dynamic evaluation of skin variations over time, i.e., treatment monitoring.

23.2.1 RCM Technical Notes

There are two commercially available confocal microscopes (VivaScope® 1500, wide probe, and 3000, handheld, Mavig, Germany). A confocal microscope consists of a point source of light, condenser, objective lenses, and a point detector. The pinhole collects light emanating only from the “in focus” plane. The mechanism of bright contrast in RCM is backscattering. In gray-scale confocal images, bright (white) structures have components with high refractive index compared with their surroundings and are similar in size to the wavelength of light. Backscattering is

primarily governed by the structures' refractive index compared to surrounding medium.

Melanin, collagen, and keratin and other elements such as cytoplasmic organelles show the highest reflectivity. The confocal scanning produces high-resolution black-and-white horizontal images (0.5×0.5 mm) with a lateral resolution of $1.0 \mu\text{m}$ and axial resolution of $3\text{--}5 \mu\text{m}$, depth of exploration up to $250 \mu\text{m}$, reaching the upper dermis.

The wide-probe RCM is composed of a probe (the head of the microscope) that is attached to the skin through a disposable plastic window that is in turn taped to a metal ring. A sequence of full-resolution individual images at a given depth is acquired and "stitched" together to create a mosaic ranging in size from 2×2 to 8×8 mm. For inflammatory or physiologic skin conditions, a 3×3 mm VivaCube® composed of four mosaics with a $25\text{-}\mu\text{m}$ step is usually acquired. Besides the horizontal mosaic, a vertical VivaStack® can be imaged. It consists of single high-resolution images acquired from the top skin surface up to $200 \mu\text{m}$ to obtain a sort of "optic biopsy." The VivaStack modality is useful for the assessment of the epidermal thickness.

The handheld RCM consists of a smaller and flexible device that is quite useful in difficult-to-access areas (skin folds, ears). This tool has an on-instrument control for laser power, imaging depth and capture but it does not allow to scan a large field of view [3, 4].

23.2.2 RCM and Skin Structure in Healthy Skin

Several aspects of the epidermis, dermo-epidermal junction (DEJ), and upper dermis may be analyzed with RCM (Table 23.1).

At epidermal level, the first image of the top surface of the skin is represented by stratum corneum, where large, bright, and polygonal-shaped anucleated corneocytes are observed. Corneocytes occupy $10\text{--}30 \mu\text{m}$ in the confocal images and are organized into large "islands" of corneocytes, separated by dark outlines or furrows. Of note, the shape and arrangement of the skin folds is related to body site (being almost absent on the forehead and well represented on the abdomen) and age [7].

Between 15 and $100 \mu\text{m}$ below the stratum corneum, keratinocytes of the granulosum-spinosum level appear as $15\text{--}30 \mu\text{m}$ cells with dark oval nuclei and bright cytoplasm. "Grainy" structure due to organelles may be observed in cells at granulosum level. Keratinocytes are cohesively assembled, forming a honeycombed pattern (Fig. 23.1a) because of the similarity with the honeycomb of bees. The interkeratinocyte contour is usually brighter than the cytoplasm and perfectly outlined [3, 8].

Conversely, dark skin types show pigmented keratinocytes, appearing as bright cells, small in size and always polygonal, separated by a darker contour (cobblestone pattern), thus resembling the negative of the honeycombed pattern [9].

Table 23.1 RCM feature definitions

<i>RCM features</i>	
<i>Epidermis</i>	
Regular honeycombed	Polygonal keratinocytes with the same size and shape and well-outlined and preserved cell border
Irregular honeycombed	Polygonal keratinocytes with variable size and shape; cell border is not always well outlined and preserved
Mottled pigmentation	Clustered bright keratinocytes in context of honeycombed pattern
Epidermal thickness	It is calculated in microns according to the number of stacks starting from the first image revealing a honeycomb pattern to the first appearance of the collagen structures
Furrow aspects	Dark folds between groups of keratinocytes, described for their orientation and intersections of the lines as “rhomboidal”
Detached corneocytes	Polygonal structures with medium to high reflectivity with sharp borders clearly detectable, corresponding to single or small clusters of corneocytes, within the stratum corneum layer
Targetoid keratinocytes	Keratinocytes within the stratum granulosum–spinosum with a bright broadened interkeratinocytic border and a central bright spot
Bright polygonal structures	Keratinocytes at the basal layer with bright cytoplasm and sharply demarcated contours
Exocytosis	Bright small dots usually located between keratinocytes
<i>Dermo-epidermal junction</i>	
Polycyclic papillary contours	Bulbous projections and cords, with variably convoluted arrangement
Sebaceous glands	Roundish to oval-shaped annular structures, corresponding to glandular parenchyma and centered by hair follicles
<i>Dermis</i>	
Thin reticular collagen	Bright thin fibrillar structures forming delicate weblike pattern; this structure can be detected around follicular openings
Coarse collagen	Coarse filamentous thick structures with tendency to be packed; weblike pattern is still observable but with larger and irregularly spaced meshes
Huddle collagen	Large hyporefractive blotches of amorphous and hyporefractive material; individual collagen fibers are no longer visible
Elastosis	Highly refractive thick and short undulated fibers, sometimes forming compact masses when severe solar elastosis is present
Melanophages	Large irregularly shaped bright cells with ill-defined borders and usually no visible nucleus which can be visible within dermal papillae
<i>Pilosebaceous unit</i>	
Size	Main axis of the infundibulum/lesion, measured in μm , at 15 μm depth from the skin surface
Border of the infundibulum	The border may be: regular/non-reflecting, bright and thin, bright and thick, not defined/dysmorphous border
Content	<ul style="list-style-type: none"> – Amorphous material within the infundibulum, visualized as a grayish mass with undefined shape and contours – Organized hyper-reflecting inflammatory infiltrate inside the comedo appearing as a roundish/ovoidal compact hyper-reflecting aggregate

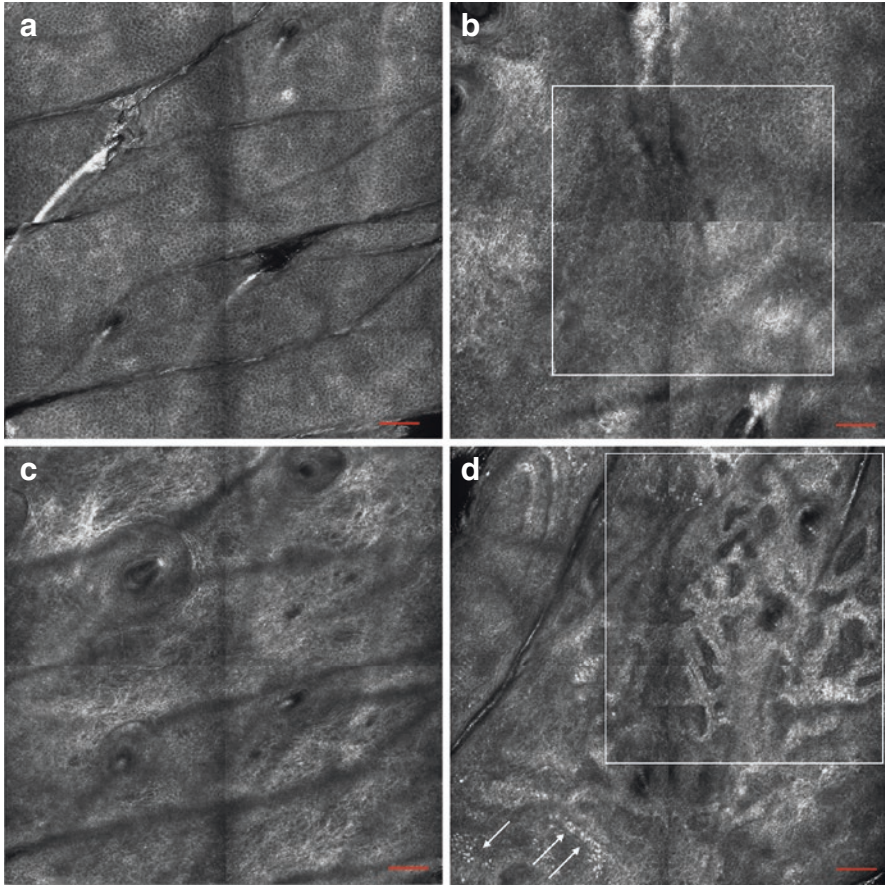


Fig. 23.1 Reflectance confocal microscopy aspect of epidermis and dermo-epidermal junction (DEJ) showing: (a) regular honeycombed, (b) irregular honeycombed (white square), (c) dermo-epidermal junction, typical of healthy skin of light-skinned subjects, (d) polycyclic papillary contours (white square) at DEJ and mottled pigmentation (white arrows). Bar scale 100 μm

An irregular honeycombed pattern is characterized by polygonal keratinocytes with variable size and shape and not always well outlined and preserved cell borders (Fig. 23.1b). Epidermal thickness may be estimated according to the number of stacks starting from the first image revealing a honeycomb pattern to the first appearance of the collagen structures [10].

Going deeper, below the granulosum–spinosum level, basal cells (7–12 μm in size) can be observed that are uniform in size and shape but are smaller and more refractive than spinous KCs because of the melanin caps forming bright disks on top of the nuclei.

At DEJ, in the presence of regular rete-ridges, basal cells form round or oval rings of bright keratinocytes surrounding dark dermal papillae, giving rise to the

ring pattern [3, 7]. Intuitively, the brightness of keratinocytes correlates with melanin content and skin phototype: the darker the skin phototype, the brighter are the basal cells and rings (Fig. 23.1c, d).

Furthermore, in normal conditions, melanocytes cannot be distinguished from keratinocytes since a melanocyte is about the same size of a keratinocyte and has a similar amount of melanin content, and, under normal metabolic conditions, dendrites are not detectable through RCM. Melanocytes can be distinguished in skin tumors (i.e., melanocytic nevi and melanoma) or in inflammatory skin conditions (i.e., melasma). The indirect effect of pigmentation in healthy skin can be observed with RCM as the presence of the so-called mottled pigmentation [10], defined as small clusters of bright keratinocytes in the context of the honeycombed pattern [7].

When the pigmentation is mainly located at the DEJ level, it appears as brightness corresponding to melanin-rich cells associated to the epidermal hyperplasia, generating the polycyclic papillary contours, corresponding to elongated epidermis and the bridging of club-shaped cristae. Skin pigmentation may also be located deeper, appearing as melanophages in the dermal papillae [10–13].

Below the DEJ, blood flow in the capillary loops within each papilla, with circulating blood cells visible at higher resolution and magnification, and collagen fibers and bundles are observed [3].

Collagen fibers may appear in different patterns, usually related with subject's age and photodamage: thin reticular collagen fibers, coarse collagen, huddles of collagen, and curled fibers. Thin reticular collagen fibers correspond to bright thin fibrillar structures forming delicate weblike pattern and are typically observed in young subjects. With aging, coarse collagen, described as coarse network of large hyporefecting collagen, and/or huddle collagen, appearing as large hyporefractive blotches of amorphous and hyporefective material, may be observed. Additionally, curled bright structures are organized in compact masses when severe solar elastosis is present [10] (Fig. 23.2).

Interestingly, a different genetic background has been associated to different collagen types in middle-aged women [14].

Other structures that are visualized are faint spiral-shaped or crescent-shaped substructures that are either in the center or along the periphery within nuclei (presumed to be nucleoli), sebaceous glands, hair shafts within hair follicles, and sweat ducts spiraling into the dermis [3].

The exploration of structures related to the pilosebaceous unit may also be performed with RCM.

In detail, size, border, and content of the follicles can be described (Table 23.1) [15].

On the face, the presence of numerous hair follicles gives rise to dark round areas with a central roundish area from which arises the hair shaft [7].

23.2.3 RCM and Skin Pigmentation

Considering the high reflectivity of melanin, the evaluation of skin pigmentation (as described in the previous section) and related disorders, such as vitiligo and melasma, represents one of the best applications for RCM.

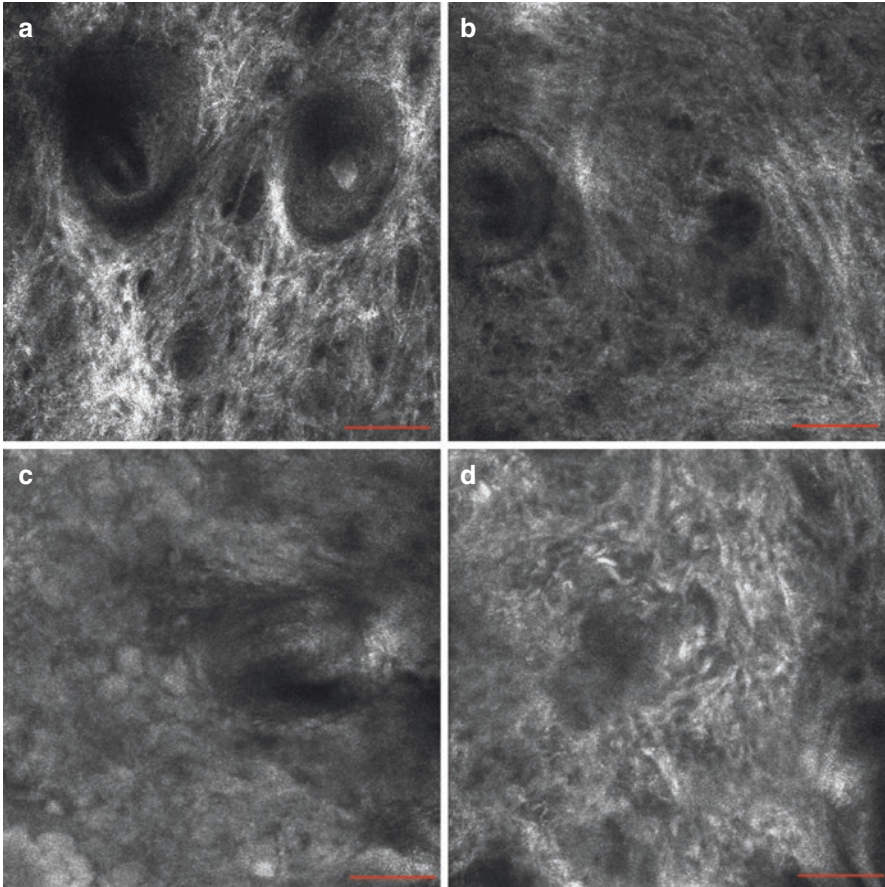


Fig. 23.2 Reflectance confocal microscopy aspect of dermis showing: (a) reticular collagen of a young subject, (b) coarse collagen, (c) huddle collagen, (d) curled fibers observed in a subject with solar elastosis. Bar scale 100 μ m

23.2.3.1 Vitiligo

Vitiligo is an acquired pigmentary disorder resulting from melanocytes destruction with consequent localized to diffused achromic patches on the skin, often associated to a great cosmetic impact, affecting the quality of life [16].

Several methods are currently available for diagnosis and therapeutic monitoring such as vitiligo Area Scoring Index (VASI), Wood's lamp, and, more recently, dermoscopy. However, none of these tools, but RCM, can give information on the microscopic disease substrate [17, 18].

Preliminary studies concerning the application of RCM in vitiligo highlighted different patterns of affected areas, as compared to non-affected, and their histopathologic correspondence.

In achromic areas, an absence of bright papillary rings at DEJ while, in apparently normal skin, an abnormal distribution of pigment, associated with an overall preserved skin anatomy, can be observed with RCM [19].

However, papillary contours remain detectable, though a bright ring is not visible. This peculiar aspect could be defined as “rimmed not-ringed papillae.” Another study, comparing apparently normal skin of vitiligo patients to healthy controls’ skin, confirmed the absence of bright rings in affected areas but not the pigment distribution irregularity [20].

So, absence or strong reduction of epidermal keratinocytes pigmentation is considered as a secondary RCM criterium for vitiligo. A limitation of RCM in vitiligo is the impossibility to study and quantify single melanocytes, in order to evaluate melanocyte depletion versus melanocyte inactivation. In fact, as previously mentioned, it is not possible to distinguish between pigmented keratinocytes and melanocytes in normal human skin.

Considering the high resolution of RCM, in achromic lesions and perilesional skin of active vitiligo, it is possible to observe inflammatory cells, melanophages [21], while in stable vitiligo these cells are very few or absent.

Additionally, the presence of melanophages, inflammatory cells, and remnants of pigment around hair follicles, without any pigment around dermal papilla or other adnexal structures openings, seems to corroborate that the melanocyte destruction occurs from the interfollicular to the perifollicular area and that the repigmentation process frequently may initiate from perifollicular sites.

These perifollicular pigmentary structures may have indirect implications on repigmentation chances, providing a reasonable idea of disease prognosis. However, further studies with long-term follow-up of vitiligo are required to confirm this hypothesis.

A very interesting application of RCM in vitiligo is disease monitoring. Accordingly, serial examinations in the same area can be performed, in order to evaluate the time-course therapy response at a cellular level. RCM can be repeatedly performed without any limitation due to the noninvasive nature of the technique and to its rapidity.

In recently re-pigmented spots, dendritic and activated melanocytes, with their intense brightness and peculiar shape, can be identified. These cells, usually located around hair follicles, look like large polygonal to stellate structures with a dark nucleus and a grainy reflective and well-outlined cytoplasm with evident and sometimes long dendritic branches budding from the main cell body.

In fact, a high amount of these cells can be observed in patients undergoing UVB-NB and PUVA therapy, enabling an early evaluation of clinical response to treatment, before the evidence of clinical improvement.

23.2.3.2 Melasma

Melasma is an acquired disorder of light-to-dark-brown hyperpigmentation, affecting millions of people worldwide. It occurs most frequently in women with Fitzpatrick skin phototypes III through V, with a localization on the face and occasionally on other sun-exposed areas (neck, forearms, etc.) [22, 23].

Treating melasma can be challenging and often a multimodal approach is required [24].

In order to plan the appropriate treatment strategy for each melasma patient, an adequate clinical evaluation should be performed, considering personal data, cutaneous localization, and medical history, such as drug assumption. However, a key role in treatment planning is represented by the assessment of the distribution and depth of melanin of the lesion [13].

Conventionally, three clinical types of melasma have been described: epidermal (melanin present within epidermis layers), dermal (melanin and melanophages in the dermis), and mixed type (combination of the first two types) [25].

Unfortunately, widely used tools for melasma evaluation, such as Wood's lamp and dermoscopy, have demonstrated a poor correlation with histology in terms of pigment localization [26, 27].

Conversely, a good correlation of histopathology and RCM in melasma has been reported [12, 28].

Several studies reported a constantly represented epidermal portion in all melasma lesions investigated with RCM, therefore leading to the hypotheses that two main categories of melasma exist: epidermal and mixed type. Thus, a new classification of melasma is proposed with two categories according to the depth of melanin in a given melasma lesion: epidermal and mixed type [12, 13, 28].

Epidermal hyperpigmentation, a hallmark of melasma in histological studies [12], appears as a mottled pigmentation and strongly visible papillary rings around the dermal papillae [11, 13].

At dermal level, pigment corresponds to melanophages [12].

RCM may enable differential diagnosis between melasma and post-inflammatory hyperpigmentation. Accordingly, this latter differs from melasma in terms of less epidermal involvement while exhibiting strongly bright rims at DEJ. Interestingly, melanophages seem not to be observed, perhaps for their deep localization at deep papillary or/and reticular dermis where RCM has a limited penetration [29].

Additionally, RCM seems to play a role in melasma treatment monitoring. Sequential relapses after melasma treatment may be observed, sometimes in relation to activated melanocytes [11–13, 28].

That can be visualized in perifollicular areas in cases with an early relapse of melasma after laser treatment, suggesting their influence in the therapeutic outcome [11].

Furthermore, the presence of melanophages in the dermis can be associated with a poor therapeutic outcome [12].

In conclusion, RCM seems to offer a reliable “mapping” of melanin distribution in melasma, being an excellent guide not just for selection, but also for monitoring of treatment, since RCM analysis is a highly sensitive instrument for detecting even clinically undetectable changes during treatment of melasma [11, 13, 30, 31].

23.2.4 *RCM and Skin Hydration*

RCM variations of the epidermal pattern are observed in subjects with clinically mild-moderate dry skin, presenting irregular surface, with an uneven reflectivity, mostly due to the presence of variable sized scales appearing as hyper-reflective areas of bright corneocytes detached from the background. At granulosum–spinosum level, the honeycombed pattern shows less evident interkeratinocyte contours and low reflecting borders with no alteration of keratinocyte shapes [8].

RCM may also be applied to treatment monitoring. A previous study on subjects with dry skin showed that the application of moisturizers reduced significantly the amount of scales and furrow size, with an increased regularization of skin surface. A delayed mechanism induced by moisturizer was the influence of the composition of interkeratinocyte space. In addition, water-in-oil moisturizers were found to rapidly reaching the lower epidermal layers but they are characterized by a reduced persistence. Overall, a reinforcement in skin cohesion, skin barrier/integrity, due to potential variations in cell metabolism can be observed after treatment with RCM [8].

23.2.5 *RCM and Skin Irritation*

Some studies have focused on microscopic aspects of skin irritation. The experimental model planned for the study of irritant contact dermatitis included the application of a patch test containing SLS (sodium lauryl sulfate) 5% aq. for 24 h. Then, three sites were compared: one site on which a product with vitamin E was applied before SLS treatment, one site on which the same product was applied after SLS treatment, and one control site (SLS only).

In case of irritant contact dermatitis, enlargement of furrows, detached corneocytes, uneven bright surface, irregular keratinocyte architecture in the granulosum–spinosum layers, targetoid keratinocytes, bright keratinocytes in the basal layer (these two latter entities corresponding to necrotic/apoptotic cells), and papillary-shaped irregularity can be observed (Table 23.1). No significant exocytosis was found. Taken together, these data suggest that irritation caused by SLS seems to reach the deepest layer of epidermis and minimally the rete-ridge architecture [32].

Vitamin E was found to limit skin barrier damage and to provide more rapid repair from epidermal damage. In fact, in the pretreated area, vitamin E application acts as a skin barrier agent, preventing stratum corneum disruption and cytological damage of keratinocytes. Furthermore, application of vitamin E after SLS stimulates skin repair mechanisms, reducing the effects at both epidermal and DEJ level.

Therefore, RCM enables the observation of damage at cellular level induced by irritation, due to the high resolution of the tool. Furthermore, protection and restoration from the skin damage induced by irritants can be analyzed and tested with different agents involved in the skin barrier and/or repair [32].

Additionally, in vivo RCM features have been reported after patch testing to induce a skin reaction in subjects with a previous history of an allergic skin reaction. In these subjects, a disruption of stratum corneum, spongiosis with vesicle formation, exocytosis and infiltrating inflammatory cells within the epidermis, and vasodilation were observed (Table 23.1). Interestingly, while dendritic cells, corresponding to Langerhans cells may be observed at 24 h after patch test, a disruption of the stratum corneum usually occurs at 48 and 72 h after removal of patch test [33].

23.2.6 RCM and Skin Aging

RCM has been employed in order to visualize variations occurring with passing years. Typical features of healthy skin have been previously described. In detail, young healthy skin usually shows a regular honeycombed pattern while at dermal level thin reflective reticulated fibers that form a weblike structure can be observed. However, variations with aging involve all the skin layers.

In order to measure the presence and extent of these variations, different scores have been developed: epidermal disarray score, epidermal hyperplasia score, and collagen alteration score.

- Epidermal disarray score (ranging from 0 to 9) = (irregular honeycombed pattern + epidermal thickness + furrow pattern).
- Epidermal hyperplasia score (ranging from 0 to 9) = (mottled pigmentation + extent of polycyclic papillary + epidermal thickness).
- Collagen alteration score (ranging from 0 to 12) = (3 for curled fibers, 2 for huddles of collagen, 1 for coarse collagen structures, and 0 for thin reticulated collagen).

In a population of 50 women from 24 to 88 years old, these scores have been analyzed on RCM images of the face (left malar eminence). Results showed that the epidermal disarray score was stable up to 65 years old, with a dramatic increase in subjects older than 65. Hyperplasia and total collagen scores were characterized by an ascending trend with increasing age, with hyperplasia increasing from young to middle age subjects. In addition, a different proportion of collagen types was observed [34].

Besides facial skin, RCM features of non-photoexposed skin of the volar lower forearm were analyzed and compared with those of photoexposed skin of the dorsal aspect of lower forearm, in a different study. This study involved 75 individuals with skin Fitzpatrick phototypes I–III, divided in two age groups (20–30 years and 50–60 years). Results of this study highlighted changes that were correlated with chronological aging and which were aggravated on the photoexposed dorsal forearm. Among these changes: loss of small skin furrows resulting in wider and less intersecting furrows; irregular epidermal honeycomb pattern; irregular mottled pigmentation; irregular papillary ring pattern; and loss of thin collagen fibers and presence of collagen huddles [35].

23.2.7 RCM and Skin Rejuvenation

Several studies investigated morphologic changes due to topical and laser treatments by using skin biopsies or animal models. Nowadays, skin changes following topical and laser treatments have been estimated with RCM in some studies.

In detail, RCM offers the opportunity to determine in vivo variations of the skin after the application of topical products, allowing the real-time assessment of changes occurring into the epidermis and superficial dermis at cellular level [8, 36–40].

Recently, skin changes occurring after laser treatments, as compared to basal conditions, have been analyzed. Mainly, RCM assessments were used after fractional laser resurfacing. This technology, when used in the micro-ablative or non-ablative modality, aims at the treatment of several skin conditions such as skin rejuvenation, scars, and striae distensae and it enables the visualization, after treatment, of collagen remodeling, appearing as long, bright, and straight fibers [41–44].

Additionally, the effects of other laser sources have been explored, including those of the innovative picosecond laser treatment [45–48].

Interesting findings were observed in subjects with melasma, as reported in the melasma section.

23.2.8 RCM and Acne-Prone Skin

RCM has been applied to acne skin, and corresponding features of the disease such as comedo, papules, and pustules have been described [15]. However, in the same study, apparently healthy skin of acne patients has been examined and compared to the skin of healthy subjects. Compared with normal skin, bright concentric material surrounding the pilosebaceous follicles (infundibula hyper-keratinization), sometimes organized in onion-like fashion, corresponding to keratin accumulation, was frequently observed around pilosebaceous follicles in acne-prone skin. Infundibula hyper-keratinization may represent the early event leading to microcomedo development, progressive accumulation of sebaceous and keratin material, thus contributing to acne pathogenesis [49]. The progressive increase in altered pilosebaceous units is followed over time by inflammatory reactions and microvascular changes. RCM seems able to identify clinically not detectable feature characterizing the background condition of acne-prone skin [15, 50, 51].

In addition, in vivo RCM may represent a useful tool for the objective assessment of treatment efficacy and individual response evaluation after medical treatment in acne. Accordingly, this tool enables the evaluation of progressive reduction of confocal inflammatory and hyperkeratotic components, in concordance with clinical scores improvement [39].

23.3 Optical Coherence Tomography

Optical coherence tomography (OCT) is an optical imaging method for the in vivo visualization of tissues, which was first applied in ophthalmology [52, 53]. However, due to its technical features such as the high resolution of 3–15 μm and the penetration depth of 1–2 mm, the OCT may also be applied to a noninvasive view of vertical and horizontal (en face) sections of the morphological features of the epidermal, DEJ, and dermal layers [54–56] (Fig. 23.3).

23.3.1 OCT Technical Notes

OCT works similarly to ultrasound. The difference is that OCT measures the intensity of reflection/backscatter of infrared light, instead of acoustical waves employed with ultrasound systems, from the skin. Due to the high velocity of light, optical echoes cannot be measured directly; therefore the integration with low-coherence interferometry correlates reflected/backscattered light from tissue with light that has traveled a known reference path. Data are displayed as images, according to the measured signal strength, by assigning gray scales to each reflection [57, 58].

Early OCT machines, known as time-domain OCT (TD-OCT), provide penetration depths of up to 2 mm at resolutions of approximately 10–15 μm , ultra-high-resolution or high-definition OCT (HD-OCT) scanners give significantly higher resolution (1–3 μm) with a lower penetration depth compared with conventional OCT [59, 60]. In addition, frequency or Fourier-domain OCT (FD-OCT) systems, including spectral domain OCT (SD-OCT) and swept source OCT combine the best aspects of TD-OCT and HD-OCT [60]. Moreover, functional OCTs, such as polarization sensitive OCT (PSOCT; [61–63]), have been developed.

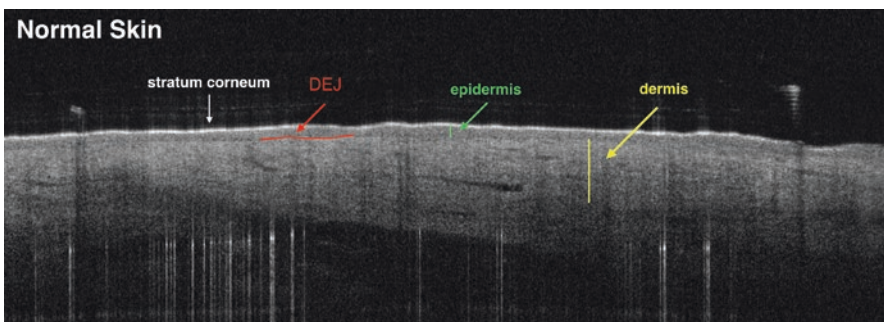


Fig. 23.3 Optical coherence tomography aspect of normal healthy skin showing stratum corneum, epidermis, dermo-epidermal junction, and dermal layer

Progressive technical developments of OCT such as dynamic OCT (D-OCT), based on speckle variance OCT, are of special interest as D-OCT allows the in vivo evaluation of blood vessels and their distribution within specific skin areas. OCT angiography may also be performed with a swept-source OCT [64, 65].

23.3.2 OCT and Skin Layers

Some key applications of OCT in dermatology are the evaluation of epidermis, collagen, and vessels. In fact, numerous OCT tools have been applied to the evaluation of the structural features of the skin, some enabling an accurate evaluation of morphologic features of collagen, others with integrated software enabling to explore vessels [14, 50, 65, 66].

OCT images reflect the well-known layered architecture of the skin, and a demarcation between epidermis and dermis, corresponding to DEJ, is evident in most of the images. Some variations may be observed in normal skin, according to skin site. The epidermis is less reflective/backscattering than the dermis and the speed of signal decay through the skin from the epidermis to the dermis is called attenuation [67].

At epidermal level, vertical sections of the skin show a bright band, due to scattering light at the skin surface, corresponding to the stratum corneum [68]. Epidermal thickness is a mean of at least three predefined measurement sites (center of the image and two lateral sites with a fixed distance to the center), vertically, between the stratum corneum and the DEJ [5]. The epidermal thickness estimated with OCT has been proven to correlate with the thickness measured with histology [69].

Several studies have confirmed that, with passing years OCT detects a decrease in epidermal thickness and a progressive flattening of the DEJ are typically associated with chronological aging of skin [68, 70–72].

Photoaging findings on OCT imaging of photoaged skin include: epidermal atrophy, uneven epidermal surface, and thickening of the stratum corneum [73]. In detail, a significantly thinner epidermis was observed in the 60s age group as compared to 20s and 40s age groups [65].

In addition, similarly to RCM, OCT may enable the evaluation of different furrow patterns on horizontal sections of the skin at epidermal level [72]. In association with progressive epidermal changes and dermal degeneration processes during aging are considered to promote increased skin roughness [35, 74].

At dermal level, OCT enables the evaluation of collagen and vessels. Depending on the specific tool employed and on skin site analyzed, different collagen features can be observed with OCT. In particular, the swept-source OCT provides structural information on skin collagen according to the degree of close packing and density of collagen bundles while polarization-sensitive OCT can collect three-dimensional collagen imaging through the measurement of polarization state of reflected light alterations [54, 75]. In addition, architecture (fibers/bundles), integrity or fragmentation and orientation (parallel/disorganized) of collagen fibers may also be

described [14]. The application of the speckle decorrelation method to OCT tomograms has enabled the evaluation of vessel density [14, 65].

23.3.3 OCT Applications

Variations of the skin occurring with passing years may be noninvasively revealed by means of OCT. Together with a progressive thinning of the epidermis, in the presence of a decreased collagen content, such as in photoaged or photodamaged skin, there is a slow change in the polarization state of light; thus, lower phase retardation rates are observed [76]. Considering different skin sites, the mean phase retardation rate is higher in the lower back skin while very low levels are observed in the temple area [54, 66, 75–77]. With intrinsic skin aging, collagen fibers become thicker, longer, straighter, and progressively aggregated in a lace-like network [72]. Tissues with increased levels of collagen content, such as scars and fibrosis, rapidly change the polarization state of light, showing higher phase retardation rates [76].

Quantitative data from the en face, or horizontal, images focusing on vessels evaluation confirmed that the vascular density and mean vessel diameter of the 60s age group were lower than those of the 20s and 40s age groups [65] (Fig. 23.4). In addition, D-OCT findings revealed that an increased vessel density of the face can be associated to skin photoaging in subjects with a particular genetic background [14].

Similarly to RCM findings, pathologic lesions of acne have correlating features with OCT. However, apparently normal skin of acne patients does not reveal significant differences compared with the one of healthy subjects not affected by acne. This may be related to the lower resolution of OCT as compared to RCM, despite

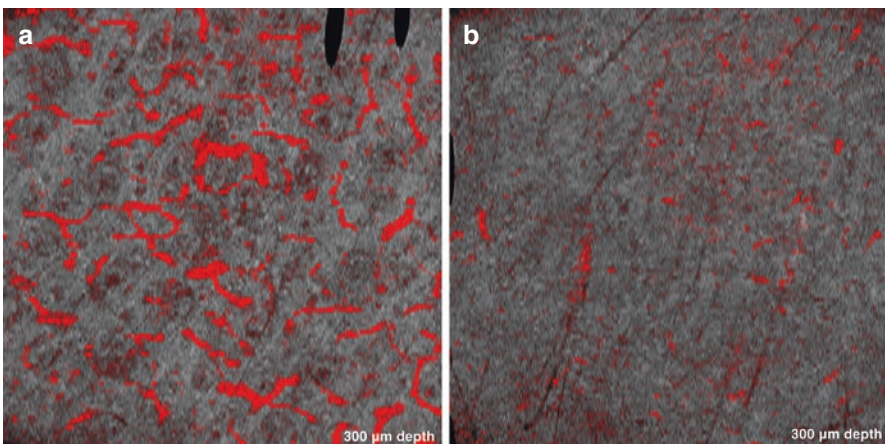


Fig. 23.4 Dynamic optical coherence tomography (Vivosight, Michelson Diagnostics, UK) en face view of normal facial skin (size 6×6) showing different vessels density in (a) a 25-year-old subject, (b) a 61-year-old subject

the increased depth of penetration of OCT, enabling the visualization of inflammation, thanks to D-OCT [50].

Interestingly, OCT has been applied to estimate treatment efficacy after the application of topical products and laser treatment [78]. In particular, a clinical trial highlighted the role of OCT in estimating the reduction of skin roughness after the application of an antiaging cream [79].

A recent study demonstrated that OCT is a good modality to objectively assess the results of Er:YAG laser treatment in a subset of patients treated for improvement of rhytids. Specifically, OCT was used to objectively visualize new collagen fiber formation that correlated with clinical improvement at 4 weeks posttreatment [80].

23.4 Multiphoton Laser Microscopy

Multiphoton laser microscopy (MPM) is a noninvasive *in vivo* technique, providing instant imaging of sections of living skin at cellular and subcellular level. MPM imaging is based on autofluorescence of intrinsic tissue fluorophores and nonlinear harmonic generation from tissue matrix components such as collagen, enabling functional and structural imaging of unstained biological tissue [81–90].

A good correlation between MPM findings and histopathology has been described [89]. However, one of the main disadvantages of the technique is that it is time-consuming since the acquisition of information requires a long time.

23.4.1 MPM Technical Notes

With MPM, images are acquired and correspond to optical sectioning parallel to the tissue surface (MPTflex, JenLab, Germany). Pictures obtained can reach a resolution depth of 200 μm . The estimated resolution is 0.5 μm in lateral direction and 1–2 μm in the axial direction [83, 85, 89].

In conventional confocal fluorescence microscopy, fluorophores are excited by the absorption of individual photons in the visible or ultraviolet spectrum while for MPM excitation is related to the simultaneous absorption of two or more photons of longer wavelength, usually in the near-infrared spectrum. The advantage of the longer wavelength infrared radiation undergoes less scattering than visible light, therefore enabling high-resolution imaging deeper into biological tissue. Additionally, for an efficient MPM excitation, ultrashort femtosecond laser pulses are required. They produce the nonlinear effect of second harmonic generation (SHG), which can be observed in periodic structures such as collagen [84, 89]. Overall, the combination of autofluorescence imaging and SHG in MPT can provide morphology and structure of both cells and extracellular matrix of the skin [81–90].

Autofluorescence is based on the excitation of some intra- or extracellular components, called fluorophores, exhibiting fluorescence. Once fluorophores absorb

energy, they can then emit energy in turn, generating a visible signal [89]. The advantage of autofluorescence is the possibility to avoid using a contrast agent or exogenous marker, simplifying both examination and patient preparation. Endogenous fluorescent biomolecules present in human skin include NADH (reduced nicotinamide adenine dinucleotide), NADPH (reduced nicotinamide adenine dinucleotide phosphate), collagen, keratin, melanin, elastin, flavines, porphyrin, tryptophan, cholecalciferol, and lipofuscin [81–90].

Besides autofluorescence, emission in the visible range also comprises SHG signals. The SHG signal comes from noncentrosymmetric molecules such as collagen and myosin and is characterized by an emission wavelength corresponding to half of that of the incident photon; this particular signal allows the visualization of dermal collagen bundles and their distinction from cellular components and elastin fibers [91].

Enhanced information concerning different states of tissue related to fluorescence decay rates can be obtained with fluorescence lifetime imaging (FLIM). Since FLIM is immune to intensity artifacts, it enables a more robust numerical description of the images than intensity imaging [92]. Different colors, corresponding to a certain fluorescent lifetime at each image pixel, permit the immediate visual identification of cellular, subcellular, or extracellular structures according to their color in the FLIM image. For instance, keratinocytes exhibit long-medium values of fluorescence decay and visually correspond to the blue-green range, whereas melanocytes have a medium to short fluorescence decay time coded in the yellow-red range [93].

Melanoma cells, selectively excited at 800 nm, are highly fluorescent and show atypical aspects and pleomorphism.

23.4.2 MPM Skin Layers/Applications

The MPM/FLIM technique has been applied to the study of physiological and pathological conditions of the skin on ex vivo and in vivo samples. Cell cultures can also be studied [92, 94].

The effects on metabolic activity, morphology, and oxidative stress of mesenchymal stem cells undergoing various differentiation stimuli can be monitored in a noninvasive manner [94].

Using MPM and FLIM, a precise and rapid assessment of the morphologic and metabolic changes which undergo fibroblasts cultures after exposure to various environmental factors or drugs can be achieved without the need of cell processing and staining [92].

Similarly to what has been previously reported for RCM, in healthy epidermis, it has been shown that cell and nucleus diameters, cell density, and MPM/FLIM characteristics vary not only according to epidermal cell depth, but also for different skin sites.

In elderly subjects, epidermal cells show morphologic alterations in shape and size; moreover, both cell and nucleus diameters are smaller at the upper layers.

Finally, the number of basal cells is decreased, whereas FLIM values at both the upper and lower layers increase during the aging process [93]. These data can be used for the comparison with MPM/FLIM aspects of epidermal cells in pathological conditions.

Going deeper into the dermis, the morphology of collagen and elastic fibers can be observed.

23.4.3 MPM Applications

Different findings may be observed with skin aging and pathological skin conditions of the dermis. During the aging process, fiber tension and morphology, network pattern, and clot formation, as assessed by MPM, can vary [95, 96]. Investigating facial skin specimens from patients of different ages by MPM imaging, Lin et al. found age-dependent changes in the dermal extracellular matrix [88].

In the case of 20-year-old individuals, AF and SHG signals were interspersed in the papillary dermis, whereas, in 70-year-old individuals, SHG signals could only be detected in a very thin zone beneath the basement membrane, and large amounts of fluorescent elastic fibers, corresponding to solar elastosis, were found in the dermis. The trend of decreasing SHG signals and increasing autofluorescence signals was well correlated with the histological findings of the decrease of collagen fibers and the increase of elastosis with increasing age.

Furthermore, in order to differentiate chronoaging from photoaging, MPM was performed on the buttocks and forearm. The latter showed a lower density and intensity of the SHG signal.

A further wavelet analysis revealed that the collagen network presents a lower diversity of textures on the forearm compared to the buttock. This can also be explained by the higher reticulation level, which decreases the fibers' motility. On the other hand, a lower level of reticulation on the buttocks leads to a more intense SHG signal as well as to greater collagen density, as the collagen network will occupy more space in the dermis. Additional differences in elastin content may contribute to these differential signals [97].

Another interesting application is represented by melasma diagnosis. MPM confirmed previous RCM findings, revealing a dermal component visualized as the presence of melanophages in patients diagnosed with epidermal melasma. The melanin volume fraction values measured in epidermal melasma ($14 \pm 4\%$) were significantly higher ($p < 0.05$) than the values measured in perilesional skin ($11 \pm 3\%$). Furthermore, elastosis was predominantly more severe in lesions than in perilesional areas and was associated with changes in melanin distribution of the basal keratinocytes [98].

23.5 Conclusions

Noninvasive skin imaging is commonly employed in the diagnostic field in dermatology. However, techniques such as RCM, OCT, and MPM provide important skin structural information that can be correlated to different functions of the skin, enabling the application of these techniques to cosmetology.

References

1. Klaassen CD, editor. Casarett and Doull's toxicology: the basic science of poisons. 5th ed. New York: McGraw-Hill; 1996. p. 529–46.
2. Monteiro-Riviere NA. Introduction to histological aspects of dermatotoxicology. *Microsc Res Tech.* 1997;37:171.
3. Rajadhyaksha M, Gonzalez S, Avislan JM, Anderson RR, Webb RH. In vivo confocal laser scanning microscopy of human skin II: advances in instrumentation and comparison with histology. *J Invest Dermatol.* 1999;113:292–303.
4. Branzan AL, Landthaler M, Szeimies RM. In vivo confocal laser scanning microscopy in dermatology. *Lasers Med Sci.* 2007;22:73–82.
5. Trojahn C, Dobos G, Richter C, Blume-Peytavi U, Kottner J. Measuring skin aging using optical coherence tomography in vivo: a validation study. *J Biomed Opt.* 2015;20:045003.
6. Seidenari S, Arginelli F, Bassoli S, Cautela J, French PM, Guanti M, Guardoli D, König K, Talbot C, Dunsby C. Multiphoton laser microscopy and fluorescence lifetime imaging for the evaluation of the skin. *Dermatol Res Pract.* 2012;2012:810749.
7. Longo C. Well-aging: early detection of skin aging signs. *Dermatol Clin.* 2016;34:513–8.
8. Manfredini M, Mazzaglia G, Ciardo S, Simonazzi S, Farnetani F, Longo C, Pellacani G. Does skin hydration influence keratinocyte biology? In vivo evaluation of microscopic skin changes induced by moisturizers by means of reflectance confocal microscopy. *Skin Res Technol.* 2013;19:299–307.
9. Longo C, Zalaudek I, Argenziano G, et al. New directions in dermatopathology: in vivo confocal microscopy in clinical practice. *Dermatol Clin.* 2012;30:799–814.
10. Longo C, Casari A, Beretti F, Cesinaro AM, Pellacani G. Skin aging: in vivo microscopic assessment of epidermal and dermal changes by means of confocal microscopy. *J Am Acad Dermatol.* 2013;68:e73–82.
11. Longo C, Pellacani G, Tourlaki A, Galimberti M, Bencini PL. Melasma and low-energy Q-switched laser: treatment assessment by means of in vivo confocal microscopy. *Lasers Med Sci.* 2014;29:1159–63.
12. Kang HY, Bahadoran P, Suzuki I, Zugaj D, Khemis A, Passeron T, Andres P, Ortonne JP. In vivo reflectance confocal microscopy detects pigmentary changes in melasma at a cellular level resolution. *Exp Dermatol.* 2010;19:e228–33.
13. Ardigo M, Cameli N, Berardesca E, Gonzalez S. Characterization and evaluation of pigment distribution and response to therapy in melasma using in vivo reflectance confocal microscopy: a preliminary study. *J Eur Acad Dermatol Venereol.* 2010;24:1296–303.
14. Guida S, Ciardo S, De Pace B, De Carvalho N, Peccerillo F, Manfredini M, Farnetani F, Chester J, Kaleci S, Manganelli M, Guida G, Pellacani G. The influence of MC1R on dermal morphological features of photo-exposed skin in women revealed by reflectance confocal microscopy and optical coherence tomography. *Exp Dermatol.* 2019;28(11):1321–7.
15. Manfredini M, Mazzaglia G, Ciardo S, Farnetani F, Mandel VD, Longo C, Zauli S, Bettoli V, Virgili A, Pellacani G. Acne: in vivo morphologic study of lesions and surrounding skin by means of reflectance confocal microscopy. *J Eur Acad Dermatol Venereol.* 2015;29:933–9.

16. Borderé AC, Lambert J, van Geel N. Clinical and emerging therapy for management of vitiligo. *Clin Cosmet Invest Dermatol*. 2009;2:15–25.
17. Ducharme EE, Silverberg NB. Selected applications of technology in the pediatric dermatology office. *Semin Cutan Med Surg*. 2008;27:94–100.
18. Chuh AA, Zawar V. Demonstration of residual perifollicular pigmentation in localized vitiligo—a reverse and novel application of digital epiluminescence dermoscopy. *Comput Med Imaging Graph*. 2004;28:213–7.
19. Ardigò M, Malizewsky I, Dell’anna ML, Berardesca E, Picardo M. Preliminary evaluation of vitiligo using in vivo reflectance confocal microscopy. *J Eur Acad Dermatol Venereol*. 2007;21:1344–50.
20. Kang HY, le Duff F, Passeron T, Lacour JP, Ortonne JP, Bahadoran P. A noninvasive technique, reflectance confocal microscopy, for the characterization of melanocyte loss in untreated and treated vitiligo lesions. *J Am Acad Dermatol*. 2010;63:e97–100.
21. Costa MC, Abraham LS, Pacifico A, Leone G, Picardo M, Ardigò M. New reflectance confocal microscopy features in vitiligo: beyond the papillary rings. *XXIst international pigment cell conference (IPCC)*. *Pigment Cell Melanoma Res*. 24:832–3.
22. Rigopoulos D, Gregoriou S, Katsambas A. Hyperpigmentation and melasma. *J Cosmet Dermatol*. 2007;6:195–202.
23. Sheth VM, Pandya AG. Melasma: a comprehensive update: part I. *J Am Acad Dermatol*. 2011;65:689–97.
24. Sheth VM, Pandya AG. Melasma: a comprehensive update: part II. *J Am Acad Dermatol*. 2011;65:699–714.
25. Sanchez NP, Pathak MA, Sato S, Fitzpatrick TB, Sanchez JL, Mihm MC Jr. Melasma: a clinical, light microscopic, ultrastructural, and immunofluorescence study. *J Am Acad Dermatol*. 1981;4:698–710.
26. Grimes PE, Yamada N, Bhawan J. Light microscopic, immunohistochemical and ultrastructural alterations in patients with melasma. *Am J Dermatopathol*. 2005;27:96–101.
27. Sarvjit V, Sharma S, Mishra S, Shing A. Melasma: a clinicopathological study of 43 cases. *Indian J Pathol Microbiol*. 2009;52:357–9.
28. Liu H, Lin Y, Nie X, Chen S, Chen X, Shi B, Tian H, Shi Z, Yu M, Zhang D, Yang B, Wang G, Wu M, Zhang F. Histological classification of melasma with reflectance confocal microscopy: a pilot study in Chinese patients. *Skin Res Technol*. 2011;17:398–403.
29. Hofmann-Wellenhof R, Pellacani G, Malvehy J, Soyer HP. *Reflectance confocal microscopy for skin diseases*. Heidelberg: Springer; 2012.
30. Goberdhan LT, Mehta RC, Aguilar C, Makino ET, Colvan L. Assessment of a superficial chemical peel combined with a multimodal, hydroquinone-free skin brightener using in vivo reflectance confocal microscopy. *J Drugs Dermatol*. 2013;12:S38–41.
31. Tsilika K, Levy JL, Kang HY, Duteil L, Khemis A, Hughes R, Passeron T, Ortonne JP, Bahadoran P. A pilot study using reflectance confocal microscopy (RCM) in the assessment of a novel formulation for the treatment of melasma. *J Drugs Dermatol*. 2011;10:1260–4.
32. Casari A, Farnetani F, De Pace B, Losi A, Pittet JC, Pellacani G, Longo C. In vivo assessment of cytological changes by means of reflectance confocal microscopy—demonstration of the effect of topical vitamin E on skin irritation caused by sodium lauryl sulfate. *Contact Dermatitis*. 2017;76:131–7.
33. González S, González E, White WM, Rajadhyaksha M, Anderson RR. Allergic contact dermatitis: correlation of in vivo confocal imaging to routine histology. *J Am Acad Dermatol*. 1999;40:708–13.
34. Longo C, Casari A, De Pace B, Simonazzi S, Mazzaglia G, Pellacani G. Proposal for an in vivo histopathologic scoring system for skin aging by means of confocal microscopy. *Skin Res Technol*. 2013;19:e167–73.
35. Wurm EM, Longo C, Curchin C, Soyer HP, Prow TW, Pellacani G. In vivo assessment of chronological ageing and photoageing in forearm skin using reflectance confocal microscopy. *Br J Dermatol*. 2012;167:270–9.

36. Addor FAS. Topical effects of SCA® (Cryptomphalus aspersa secretion) associated with regenerative and antioxidant ingredients on aged skin: evaluation by confocal and clinical microscopy. *Clin Cosmet Investig Dermatol*. 2019;12:133–40.
37. Diluvio L, Dattola A, Cannizzaro MV, Franceschini C, Bianchi L. Clinical and confocal evaluation of avenanthramides-based daily cleansing and emollient cream in pediatric population affected by atopic dermatitis and xerosis. *G Ital Dermatol Venereol*. 2019;154:32–6.
38. Bağcı IS, Ruini C, Niesert AC, Horváth ON, Berking C, Ruzicka T, von Braunmühl T. Effects of short-term moisturizer application in different ethnic skin types: noninvasive assessment with optical coherence tomography and reflectance confocal microscopy. *Skin Pharmacol Physiol*. 2018;31:125–33.
39. Manfredini M, Greco M, Farnetani F, Mazzaglia G, Ciardo S, Bettoli V, Virgili A, Pellacani G. In vivo monitoring of topical therapy for acne with reflectance confocal microscopy. *Skin Res Technol*. 2017;23:36–40.
40. Meinke MC, Richter H, Kleemann A, Lademann J, Tschersch K, Rohn S, Schempp CM. Characterization of atopic skin and the effect of a hyperforin-rich cream by laser scanning microscopy. *J Biomed Opt*. 2015;20:051013.
41. Guida S, Galimberti MG, Bencini M, Pellacani G, Bencini PL. Treatment of striae distensae with non-ablative fractional laser: clinical and in vivo microscopic documentation of treatment efficacy. *Lasers Med Sci*. 2018;33:75–8.
42. Bencini PL, Turlaki A, Galimberti M, Longo C, Pellacani G, De Giorgi V, Guerriero G. Nonablative fractional photothermolysis for acne scars: clinical and in vivo microscopic documentation of treatment efficacy. *Dermatol Ther*. 2012;25:463–7.
43. Grönemeyer LL, Thoms KM, Bertsch HP, Hofmann L, Schön MP, Haenssle HA. Reflectance confocal microscopy and Hailey-Hailey disease: assessment of response to treatment after CO2 laser ablation. *J Dtsch Dermatol Ges*. 2014;12:1135–7.
44. Longo C, Galimberti M, De Pace B, Pellacani G, Bencini PL. Laser skin rejuvenation: epidermal changes and collagen remodeling evaluated by in vivo confocal microscopy. *Lasers Med Sci*. 2013;28:769–76.
45. Guida S, Pellacani G, Bencini PL. Picosecond laser treatment of atrophic and hypertrophic surgical scars: in vivo monitoring of results by means of 3D imaging and reflectance confocal microscopy. *Skin Res Technol*. 2019;25:896–902.
46. Guida S, Bencini PL, Pellacani G. Picosecond laser for atrophic surgical scars treatment: in vivo monitoring of results by means of reflectance confocal microscopy. *J Eur Acad Dermatol Venereol*. 2019;33:e114–6.
47. Jo DJ, Kang IH, Baek JH, Gwak MJ, Lee SJ, Shin MK. Using reflectance confocal microscopy to observe in vivo melanolysis after treatment with the picosecond alexandrite laser and Q-switched Nd:YAG laser in melasma. *Lasers Surg Med*. 2018. [Epub ahead of print].
48. Fu Z, Huang J, Xiang Y, Huang J, Tang Z, Chen J, Nelson JS, Tan W, Lu J. Characterization of laser-resistant port wine stain blood vessels using in vivo reflectance confocal microscopy. *Lasers Surg Med*. 2019;51(10):841–9.
49. Holmes RL, Williams M, Cunliffe WJ. Pilo-sebaceous duct obstruction and acne. *Br J Dermatol*. 1972;87:327–32.
50. Manfredini M, Greco M, Farnetani F, Ciardo S, De Carvalho N, Mandel VD, Starace M, Pellacani G. Acne: morphologic and vascular study of lesions and surrounding skin by means of optical coherence tomography. *J Eur Acad Dermatol Venereol*. 2017;31:1541–6.
51. Manfredini M, Bettoli V, Sacripanti G, Farnetani F, Bigi L, Puviani M, Corazza M, Pellacani G. The evolution of healthy skin to acne lesions: a longitudinal, in vivo evaluation with reflectance confocal microscopy and optical coherence tomography. *J Eur Acad Dermatol Venereol*. 2019;33:1768–74.
52. Fercher AF, Mengedoht K, Werner W. Eye-length measurement by interferometry with partially coherent light. *Opt Lett*. 1988;13:186–8.
53. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991;254:1178–81.

54. Sattler EC, Kastle R, Welzel J. Optical coherence tomography in dermatology. *J Biomed Opt.* 2013;18:061224.
55. Ring HC, Mogensen M, Hussain AA, Steadman N, Banzhaf C, Themstrup L, et al. Imaging of collagen deposition disorders using optical coherence tomography. *J Eur Acad Dermatol Venereol.* 2015;29:890–8.
56. Ulrich M, Themstrup L, de Carvalho N, Ciardo S, Holmes J, Whitehead R, et al. Dynamic optical coherence tomography of skin blood vessels—proposed terminology and practical guidelines. *J Eur Acad Dermatol Venereol.* 2018;32:152–5.
57. Fujimoto JG. Optical coherence tomography for ultrahigh resolution in vivo imaging. *Nat Biotechnol.* 2003;21:1361–7.
58. Gladkova ND, Petrova GA, Nikulin NK, et al. In vivo optical coherence tomography imaging of human skin: norm and pathology. *Skin Res Technol.* 2000;6:6–16.
59. Gambichler T, Schmid-Wendtner M, Plura I, Kampilafkos P, Stücker M, Berking C, et al. A multicentre pilot study investigating high-definition optical coherence tomography in the differentiation of cutaneous melanoma and melanocytic naevi. *J Eur Acad Dermatol Venereol.* 2015;29:537–41.
60. Reddy N, Nguyen BT. The utility of optical coherence tomography for diagnosis of basal cell carcinoma: a quantitative review. *Br J Dermatol.* 2019;180:475–83.
61. De Boer JF, Milner TE, Martin JCG, Nelson JS. Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence tomography. *Opt Lett.* 1997;22:934–6.
62. Hee MR, Huang D, Swanson EA, Fujimoto JG. Polarization-sensitive low-coherence reflectometer for birefringence characterization and ranging. *J Opt Soc Am B.* 1992;9:903–8.
63. Yasuno Y, Makita S, Sutoh Y, Itoh M, Yatagai T. Birefringence imaging of human skin by polarization-sensitive spectral interferometric optical coherence tomography. *Opt Lett.* 2002;27:1803–5.
64. Ulrich M, Themstrup L, de Carvalho N, Manfredi M, Grana C, Ciardo S, et al. Dynamic optical coherence tomography in dermatology. *Dermatology.* 2016;232:298–311.
65. Hara Y, Yamashita T, Kikuchi K, Kubo Y, Katagiri C, Kajiya K, Saeki S. Visualization of age-related vascular alterations in facial skin using optical coherence tomography-based angiography. *J Dermatol Sci.* 2018;90:96–8.
66. Mamalis A, Ho D, Jagdeo J. Optical coherence tomography imaging of normal, chronologically aged, photoaged and photodamaged skin: a systematic review. *Dermatol Surg.* 2015;41:993–1005.
67. Mogensen M, Morsy HA, Thrane L, Jemec GB. Morphology and epidermal thickness of normal skin imaged by optical coherence tomography. *Dermatology.* 2008;217:14–20.
68. Neerken S, Lucassen GW, Bisschop MA, Lenderink E, Nuijs TA. Characterization of age-related effects in human skin: a comparative study that applies confocal laser scanning microscopy and optical coherence tomography. *J Biomed Opt.* 2004;9:274–81.
69. Gambichler T, Moussa G, Regeniter P, Kasseck C, Hofmann MR, Bechara FG, Sand M, Altmeyer P, Hoffmann K. Validation of optical coherence tomography in vivo using cryostat histology. *Phys Med Biol.* 2007;52:N75–85.
70. Wu S, Li H, Zhang X, Li Z. Optical features for chronological aging and photoaging skin by optical coherence tomography. *Lasers Med Sci.* 2013;28:445–50.
71. Florence P, Cornillon C, D'Arras MF, Flament F, Panhard S, Diridollou S, Loussouarn G. Functional and structural age-related changes in the scalp skin of Caucasian women. *Skin Res Technol.* 2013;19:384–93.
72. Boone MA, Suppa M, Marneffe A, Miyamoto M, Jemec GB, Del Marmol V. High-definition optical coherence tomography intrinsic skin ageing assessment in women: a pilot study. *Arch Dermatol Res.* 2015;307:705–20.
73. Barton JK, Gossage KW, Xu W, Ranger-Moore JR, Saboda K, Brooks CA, Duckett LD, Salasche SJ, Warneke JA, Alberts DS. Investigating sun-damaged skin and actinic keratosis with optical coherence tomography: a pilot study. *Technol Cancer Res Treat.* 2003;2:525–35.

74. Lagarrigue SG, George J, Questel E, Lauze C, Meyer N, Lagarde JM, Simon M, Schmitt AM, Serre G, Paul C. In vivo quantification of epidermis pigmentation and dermis papilla density with reflectance confocal microscopy: variations with age and skin phototype. *Exp Dermatol*. 2012;21:281–6.
75. Phillips KG, Wang Y, Levitz D, et al. Dermal reflectivity determined by optical coherence tomography is an indicator of epidermal hyperplasia and dermal edema within inflamed skin. *J Biomed Opt*. 2011;16:040503.
76. Babalola O, Mamalis A, Lev-Tov H, Jagdeo J. Optical coherence tomography (OCT) of collagen in normal skin and skin fibrosis. *Arch Dermatol Res*. 2014;306:1–9.
77. Pierce MC, Strasswimmer J, Hyle Park B, Cense B, De Boer JF. Birefringence measurements in human skin using polarization-sensitive optical coherence tomography. *J Biomed Opt*. 2004;9:287–91.
78. Hara Y, Masuda Y, Hirao T, Yoshikawa N. The relationship between the Young's modulus of the stratum corneum and age: a pilot study. *Skin Res Technol*. 2013;19:339–45.
79. Vasquez-Pinto LM, Maldonado EP, Raelle MP, Amaral MM, de Freitas AZ. Optical coherence tomography applied to tests of skin care products in humans—a case study. *Skin Res Technol*. 2015;21:90–3.
80. Kunzi-Rapp K, Dierickx CC, Cambier B, Drosner M. Minimally invasive skin rejuvenation with erbium: YAG laser used in thermal mode. *Lasers Surg Med*. 2006;38:899–907.
81. Denk W, Strickler JH, Webb WW. Two-photon laser scanning fluorescence microscopy. *Science*. 1990;248:73–6.
82. Kollias N, Zonios G, Stamatas GN. Fluorescence spectroscopy of skin. *Vib Spectrosc*. 2002;28:17–23.
83. König K, Riemann I. High-resolution multiphoton tomography of human skin with subcellular spatial resolution and picosecond time resolution. *J Biomed Opt*. 2003;8:432–9.
84. Zipfel WR, Williams RM, Christiet R, Nikitin AY, Hyman BT, Webb WW. Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. *Proc Natl Acad Sci U S A*. 2003;100:7075–80.
85. Riemann I, Dimitrow E, Fischer P. High resolution multiphoton tomography of human skin in vivo and in vitro in proceedings of the femtosecond laser applications in biology, vol. 5312 of Proceedings of SPIE, Strasbourg, France, April 2004, pp. 21–28.
86. Schenke-Layland K, Riemann I, Damour O, Stock UA, König K. Two-photon microscopes and in vivo multiphoton tomographs—powerful diagnostic tools for tissue engineering and drug delivery. *Adv Drug Deliv Rev*. 2006;58:878–96.
87. Becker W, Bergmann A, Biskup C. Multispectral fluorescence lifetime imaging by TCSPC. *Microsc Res Tech*. 2007;70:403–9.
88. Lin SJ, Jee SH, Dong CY. Multiphoton microscopy: a new paradigm in dermatological imaging. *Eur J Dermatol*. 2007;17:361–6.
89. König K. Clinical multiphoton tomography. *J Biophotonics*. 2008;1:13–23.
90. Tsai TH, Jee SH, Dong CY, Lin SJ. Multiphoton microscopy in dermatological imaging. *J Dermatol Sci*. 2009;56:1–8.
91. Zhao J, Chen J, Yang Y, Zhuo S, Jiang X, Tian W, Ye X, Lin L, Xie S. Jadassohn-Pellizzari anetoderma: study of multiphoton microscopy based on two-photon excited fluorescence and second harmonic generation. *Eur J Dermatol*. 2009;19:570–5.
92. Stefania S, Simona S, Paola A, Luisa B, Stefania B, Jennifer C, Chiara F, Paul F, Stefania G, Karsten K, Cristina M, Clifford T, Christopher D. High-resolution multiphoton tomography and fluorescence lifetime imaging of UVB-induced cellular damage on cultured fibroblasts producing fibres. *Skin Res Technol*. 2013;19:251–7.
93. Benati E, Bellini V, Borsari S, Dunsby C, Ferrari C, French P, Guanti M, Guardoli D, Koenig K, Pellacani G, Ponti G, Schianchi S, Talbot C, Seidenari S. Quantitative evaluation of healthy epidermis by means of multiphoton microscopy and fluorescence lifetime imaging microscopy. *Skin Res Technol*. 2011;17:295–303.

94. Rice WL, Kaplan DL, Georgakoudi I. Two-photon microscopy for non-invasive, quantitative monitoring of stem cell differentiation. *PLoS One*. 2010;5:e10075.
95. Koehler MJ, Preller A, Kindler N, Elsner P, König K, Bückle R, Kaatz M. Intrinsic, solar and sunbed-induced skin aging measured in vivo by multiphoton laser tomography and biophysical methods. *Skin Res Technol*. 2009;15:357–63.
96. Koehler MJ, Hahn S, Preller A, Elsner P, Ziemer M, Bauer A, König K, Bückle R, Fluhr JW, Kaatz M. Morphological skin ageing criteria by multiphoton laser scanning tomography: non-invasive in vivo scoring of the dermal fibre network. *Exp Dermatol*. 2008;17:519–23.
97. Le Digabel J, Houriez-Gombaudo-Saintonge S, Filiol J, Lauze C, Josse G. Dermal fiber structures and photoaging. *J Biomed Opt*. 2018;23:1–12.
98. Lentsch G, Balu M, Williams J, Lee S, Harris RM, König K, Ganesan A, Tromberg BJ, Nair N, Santhanam U, Misra M. In vivo multiphoton microscopy of melasma. *Pigment Cell Melanoma Res*. 2019;32:403–11.

Chapter 24

Assessment of Pruritus and Sensitive Skin



Flavien Huet and Laurent Misery

24.1 Pruritus

24.1.1 Introduction

Pruritus is defined as an unpleasant sensation that provokes the need to scratch [1]. Pruritus is the most frequent symptom in dermatology and is responsible for a huge psychosocial burden: pruritus is listed by the Global Burden of Disease project as one of the 50 most common interdisciplinary symptoms leading to high burden levels [2]. Pruritus is also frequent in the general population. In a recent study, pruritus was reported in 32% of the general population [3]. Furthermore, chronic pruritus, i.e., pruritus lasting longer than 6 weeks, deeply alters the quality of life and sleep of patients [4, 5]. In a recent cross-sectional study, pruritus had a negative effect on daily life in two out of three patients, and sleeplessness was observed in more than 88% of pruritus patients [6]. It should be investigated, because not only dermatological conditions cause chronic itch but also systemic, neurological, or psychiatric diseases. Since pruritus is a subjective symptom, assessing pruritus remains a challenge because an objective method of measuring symptoms has yet to be designed [7]. The International Forum for the Study of Itch (IFSI) established a classification dividing chronic pruritus into three categories upon physical examination:

1. Chronic pruritus on inflamed skin
2. Chronic pruritus on normal skin
3. Chronic pruritus with severe scratch lesions [8]

F. Huet (✉) · L. Misery

Department of Dermatology, University Hospital of Brest, Brest, France

Univ Brest, LIEN, Brest, France

e-mail: flavien.huet@chu-brest.fr

Measuring pruritus is essential for patient assessment but also to raise awareness of the impact of chronic pruritus in health care. Furthermore, a standardized manner of assessment of pruritus to allow comparisons between studies and ensure the highest medical and professional standards is necessary in randomized clinical trials (RCT) [9]. Different types of scales for measuring pruritus intensity are available. Sleep disturbance, anxiety, and depression are other factors that should be taken into account.

24.1.2 Pruritus Intensity

Mono- and multidimensional scales are available for measuring pruritus. The intensity can be quickly measured with mono-dimensional scales that are routinely used in clinical care and RCTs.

The *numerical rating scale* (NRS) is used to measure subjective symptom intensity. Patients can be asked to rate their itch intensity from 0 (“absence of symptoms”) to 10 (“the worst imaginable itch”).

Another mono-dimensional scale, the *visual analog scale* (VAS), offers a rapid method for appraising intensity, thus making them especially useful in RCTs. The VAS is currently the most popular instrument. VAS provides patients the opportunity to indicate the intensity of their itch by marking on a 10-cm-long, ruler-shaped scale [10, 11]. Both endpoints are marked with a number corresponding to the intensity, the left endpoint with 0 representing “no itch” and the right endpoint with 10 representing the “worst imaginable itch. Scores below 3.0 VAS/NRS points are generally associated with mild itch, whereas scores higher than 6.9 illustrate severe itch. Scores above 9.0 represent very severe itch [11].

The *verbal rating scale* (VRS) is a further mono-dimensional scale that provides patients the option of choosing their itch intensity by means of gradually rising adjectives (0-no itch, 4-very severe itch) [10].

The NRS, VAS, and VRS have been validated in large-scale studies consisting of chronic pruritus patients with pruritic dermatoses or pruritus of various origins. These instruments have high reproducibility and a high correlation between different scales [10–12]. NRS, VAS, and VRS are used in RCTs but there is still a debate on the optimum recall period (12 h, 24 h, 3 days, or 7 days). Usually, the past 24 h are taken into consideration. It is important to keep in mind that these scales only provide the itch intensity at a specific point in time, being susceptible to be perturbed by factors such as stress, anxiety, and comorbidities.

Multidimensional scales, such as 5-D itch scale, can be alternatively used. The *5-D itch scale* is a reliable, multidimensional measure of itching that has been validated in patients with chronic pruritus to detect changes over time [13]. The five dimensions are degree, duration, direction, disability, and distribution. The 5-D itch scale is useful as an outcome measure in clinical trials.

The *Pruritus Severity Scale*, a self-reported seven-item scale, combines itch intensity, sleep disturbance, and burden by the symptom (pruritus) aiming to give a comprehensive rating of the itch impact [14].

The *12-Item Pruritus Severity Scale* (12-PSS) has been developed as a simple, multidimensional method for assessing pruritus intensity [15]. This scale assesses pruritus intensity (two questions), pruritus extent (one question) and duration (one question), influence of pruritus on concentration and patient psyche (four questions), and scratching as a response to pruritus stimuli (four questions). 12-PSS has a significant correlation with the VAS and quality of life level, and a high reproducibility.

Other scales such as the Dynamic Pruritus Score (DPS) (measuring the change in pruritus intensity, compared with a defined earlier time-point, enabling a more precise interpretation of the course of pruritus), Itch Severity Scale, ItchyQuant, and the Eppendorf Pruritus Questionnaire exist to evaluate pruritus [16]. Scratch lesions can be also considered with two additional scales: Scratch Symptom Score or Prurigo Activity Score [17].

24.1.3 Quality of Life

Chronic pruritus alters the quality of life and has a huge psychosocial burden. Especially, pruritus impacts negatively sleep of patients and increases the risk of anxiety or depression.

The *DLQI* (Dermatology Life Quality Index) is a health-related quality of life scale specific for dermatologic disorders and is widely used [18]. It comprises ten items focusing on the following eight dimensions: “symptoms,” “daily activities,” “leisure,” “work,” “personal relationships,” “feeling,” “school,” and “treatment.” A total score (between 0 and 30) is calculated and can be expressed as a percentage. The higher the score is, the greater the impairment in quality of life is. Generally, quality of life is considered moderately impaired in patients who score between 6 and 11, very impaired in those scoring between 11 and 21, and extremely impaired in those with a score of 21 or greater [19].

The *36-item short form (SF-36)* may be used to measure the effect of chronic pruritus on patient’s quality of life. The SF-12 is a short version of the SF-36 that is composed of 12 items. It is a health-related generic quality of life questionnaire that can be used in the general population to assess its health status [20]. Responses to questions are dichotomous (yes/no) or ordinal (excellent to poor), or they express a frequency (always to never). The following two scores can be calculated from these 12 questions: a Physical Component Summary (PCS-12) and a Mental Component Summary (MCS-12). There is no overall score. If there was no response to a question from one of these subscales, the score cannot be calculated. Thus, a given subject can have a PCS-12 score but no MCS-12 score. After the abnormal responses and reversed items were processed, each response was assigned a coefficient.

PCS-12 and MCS-12 scores were obtained by summation. The higher the score, the better the quality of life [21].

The *ItchyQoL* has been developed to assess the impairment to health-related quality of life (HRQoL) in patients with CP. It is a pruritus-specific instrument that measures the impact of CP on patients' HRQoL. This 22-item questionnaire describes certain areas of focus such as "symptoms," "functioning," and "emotions" and is strongly correlated with DLQI [22].

Less specific scales of skin diseases can also be used in addition, for example, the EuroQol 5-dimensions (EQ-5D) questionnaire, which is known to be a valuable instrument to assess the impairment of health-related quality of life in skin diseases [23, 24]. EQ-5D is a standardized generic instrument consisting of two parts [25]. The EQ visual analog scale (EQ-VAS) records respondents' self-rated health on a visual analog scale, from "0" to "100" (i.e., worst to best imaginable health state). This information can be used as a quantitative measure of health outcome as judged by the individual respondent. The EQ-5D further assesses the health status across dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. EQ-5D reflects the current general health status.

To complete the evaluation of the impact of pruritus, *specific scales of sleep impairment or psychiatric comorbidities* (anxiety, depression) can be used.

24.2 Sensitive Skin

24.2.1 Introduction

Sensitive skin syndrome (SSS) is a frequent complaint in the general population and can impair the quality of life. A consensus definition has been recently established by the International Forum for the Study of Itch (IFSI) through its special interest group (SIG) on sensitive skin [26]. Sensitive skin is a syndrome defined by the occurrence of unpleasant sensations (stinging, burning, pain, pruritus, and tingling sensations) in response to stimuli that should not normally cause such sensations. The triggering factors can be physical (UV, hot, cold, and wind), chemical (mainly cosmetics and water), or occasionally psychological or hormonal. In addition, these unpleasant sensations cannot be attributed to other skin diseases. Sensitive skin can affect any part of the body, but it affects mainly the face (85%) and the hands (58%).

24.2.2 Assessment

The diagnosis and evaluation of SSS can be performed by many sensory testing methods, from stinging tests with lactic acid (or capsaicin, dimethylsulfoxide or heat probe), occlusion tests, behind-the-knee tests, washing and exaggerated

immersion tests, to the evaluation of itching and quantitative sensory testing (QST) [27, 28]. However, there is still a lack of international consensus on the preferred method. In all cases, there is a need to assess the subjective opinions of the patients. Because sensitive skin is defined as being a subjective symptom with abnormal sensations in response to a variety of factors, the best method to diagnose sensitive skin is utilizing patient-reported scales [29].

The easiest way is to interview people by asking them an open question concerning whether or not sensitive skin was present. Another possibility is to propose to people to choose between four answers: “no sensitive skin,” “slightly sensitive skin,” “sensitive skin,” or “very sensitive skin.”

The Sensitive Scale, with 14-item and 10-item versions, is a good tool to assess SSS and is noninvasive [30]. This scale was tested in 11 countries in different languages on 2966 participants and its correlations with the DLQI was found. Using the 10-item version appeared to be preferable because it was quicker and easier to complete, with the same internal consistency. Among localizations of sensitive skin, the scalp is known to be one of the most commonly affected areas. A specific questionnaire to assess sensitive scalp was validated and named the *3S Questionnaire* [31]. The final score was obtained by multiplying score severity of abnormal sensations by the total number of sensations. These two scales for the evaluation of SSS may be a useful tool for diagnosis as well as the evaluation of the severity of SSS and treatment effectivity.

The *lactic acid sting test* is often used in study of SSS. Lactic acid is applied to the nasolabial fold, followed by evaluation of the intensity of the subjective symptoms. Lactic acid solution is applied to the right nasolabial fold, whereas the equal volume of saline is applied to the contralateral site. The participants are asked to grade the intensity of sting using a 4-point scale (0 = no stinging; 1 = slight stinging; 2 = moderate stinging; and 3 = strong stinging) at 0 s, and 2.5 and 5 min after application of lactic acid. Cumulative scores at 2.5 and 5 min ≥ 3 were considered as the positive. A study on 292 subjects published in 2019 showed that lactic acid sting test scores could be used to identify subjects with sensitive skin characterized mainly by stinging and itching, but not those mainly by burning and erythema [32].

Scale assessments show that there are large variations between subjects and that there is no clear delineation between subjects without sensitive skin and patients with sensitive skin. A defined cutoff based on empirical data is required [30, 31].

24.2.3 Burden

Many skin disorders have been demonstrated to have a high impact on quality of life and, more recently, to have a heavy burden, including alterations of quality of life in sensitive skin [2, 33, 34]. The burden of a disorder can be defined as its effects on overall health [35]. As most skin diseases are not associated with substantial mortality, measures of morbidity assume greater importance in the measurement of the burden of disease [35]. Because there are many specific aspects of skin disorders,

the global tools for the measurement of the burden are not sufficiently relevant or adapted. Individual burden accounts for the broadest aspects of disease-related disability, including psychological, physical, social, and economic factors, and all of these aspects must be integrated in specifically dedicated questionnaires [36].

The *BoSS* (“*Burden of Sensitive Skin*”) is a specific and validated assessment tool related to the individual burden of sensitive skin syndrome [36]. This questionnaire includes 14 items and is correlated with both components of the SF-12 and with the DLQI. The BoSS questionnaire can be used as a total score (range 0–70). In detail, 0 means that there is no impact and 70 means that there is a maximal impact. The total score is calculated by summing the scores obtained for each of the 14 items that composed the questionnaire, which were rated as follows: 0 for never, 1 for rarely, 2 for sometimes, 3 for often, and 4 for constantly.

24.3 Conclusions

The evaluation of pruritus and sensitive skin are based on questionnaires since they are subjective symptoms.

Conflicts of Interest None declared.

References

1. Misery L, Ständer S, éditeurs. Pruritus [Internet]. 2^e éd. Springer International Publishing; 2016.
2. Hay RJ, Johns NE, Williams HC, Bolliger IW, Dellavalle RP, Margolis DJ, et al. The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions. *J Invest Dermatol*. 2014;134(6):1527–34.
3. Huet F, Taieb C, Corgibet F, Brenaut E, Richard MA, Misery L. Pruritus, pain and depression associated with the most common skin diseases: data from the French study “Objectif Peau”. In process.
4. Gowda S, Goldblum OM, McCall WV, Feldman SR. Factors affecting sleep quality in patients with psoriasis. *J Am Acad Dermatol*. 2010;63(1):114–23.
5. Kaaz K, Szepietowski JC, Matusiak Ł. Influence of itch and pain on sleep quality in atopic dermatitis and psoriasis. *Acta Derm Venereol*. 2019;99(2):175–80.
6. Kopyciok MER, Ständer HF, Osada N, Steinke S, Ständer S. Prevalence and characteristics of pruritus: a one-week cross-sectional study in a German dermatology practice. *Acta Derm Venereol*. 2016;96(1):50–5.
7. Ständer S, Augustin M, Reich A, Blome C, Ebata T, Phan NQ, et al. Pruritus assessment in clinical trials: consensus recommendations from the International Forum for the Study of Itch (IFSI) special interest group scoring itch in clinical trials. *Acta Derm Venereol*. 2013;93(5):509–14.
8. Ständer S, Weisshaar E, Mettang T, Szepietowski JC, Carstens E, Ikoma A, et al. Clinical classification of itch: a position paper of the international forum for the study of itch. *Acta Derm Venereol*. 2007;87(4):291–4.
9. Pereira MP, Ständer S. Assessment of severity and burden of pruritus. *Allergol Int*. 2017;66(1):3–7.

10. Phan NQ, Blome C, Fritz F, Gerss J, Reich A, Ebata T, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritus. *Acta Derm Venereol.* 2012;92(5):502–7.
11. Reich A, Heisig M, Phan NQ, Taneda K, Takamori K, Takeuchi S, et al. Visual analogue scale: evaluation of the instrument for the assessment of pruritus. *Acta Derm Venereol.* 2012;92(5):497–501.
12. Furue M, Ebata T, Ikoma A, Takeuchi S, Kataoka Y, Takamori K, et al. Verbalizing extremes of the visual analogue scale for pruritus: a consensus statement. *Acta Derm Venereol.* 2013;93(2):214–5.
13. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. *Br J Dermatol.* 2010;162(3):587–93.
14. Majeski CJ, Johnson JA, Davison SN, Lauzon CJ. Itch severity scale: a self-report instrument for the measurement of pruritus severity. *Br J Dermatol.* 2007;156(4):667–73.
15. Reich A, Božek A, Janiszewska K, Szepietowski JC. 12-Item Pruritus Severity Scale: development and validation of new itch severity questionnaire. *Biomed Res Int.* 2017;2017:3896423.
16. Ständer S, Blome C, Anastasiadou Z, Zeidler C, Jung KA, Tsianakas A, et al. Dynamic pruritus score: evaluation of the validity and reliability of a new instrument to assess the course of pruritus. *Acta Derm Venereol.* 2017;97(2):230–4.
17. Pölking J, Zeidler C, Schedel F, Osada N, Augustin M, Metze D, et al. Prurigo activity score (PAS): validity and reliability of a new instrument to monitor chronic prurigo. *J Eur Acad Dermatol Venereol.* 2018;32(10):1754–60.
18. Finlay AY. Quality of life assessments in dermatology. *Semin Cutan Med Surg.* 1998;17(4):291–6.
19. Lewis V, Finlay AY. 10 years experience of the dermatology life quality index (DLQI). *J Investig Dermatol Symp Proc.* 2004;9(2):169–80.
20. Ware J, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996;34(3):220–33.
21. Lim LL, Fisher JD. Use of the 12-Item Short-Form (SF-12) Health Survey in an Australian heart and stroke population. *Qual Life Res.* 1998;8(1–2):1–8.
22. Zeidler C, Steinke S, Riepe C, Bruland P, Soto-Rey I, Storck M, et al. Cross-European validation of the ItchyQoL in pruritic dermatoses. *J Eur Acad Dermatol Venereol.* 2019;33(2):391–7.
23. Balieva F, Kupfer J, Lien L, Gieler U, Finlay AY, Tomás-Aragónés L, et al. The burden of common skin diseases assessed with the EQ5D™: a European multicentre study in 13 countries. *Br J Dermatol.* 2017;176(5):1170–8.
24. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *Eur J Health Econ.* 2015;16(9):927–39.
25. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199–208.
26. Misery L, Ständer S, Szepietowski JC, Reich A, Wallengren J, Evers AWM, et al. Definition of sensitive skin: an expert position paper from the Special Interest Group on sensitive skin of the International Forum for the Study of Itch. *Acta Derm Venereol.* 2017;97(1):4–6.
27. Berardesca E, Fluhr JW, Maibach HI. Sensitive skin syndrome. New York: Taylor & Francis. 2006; 281 p.
28. Misery L, Loser K, Ständer S. Sensitive skin. *J Eur Acad Dermatol Venereol.* 2016;30 Suppl 1:2–8.
29. Misery L. Sensitive skin. *Expert Rev Dermatol.* 2013;8(6):631–7.
30. Misery L, Jean-Decoster C, Mery S, Georgescu V, Sibaud V. A new ten-item questionnaire for assessing sensitive skin: the sensitive Scale-10. *Acta Derm Venereol.* 2014;94(6):635–9.
31. Misery L, Rahhali N, Ambonati M, Black D, Saint-Martory C, Schmitt A-M, et al. Evaluation of sensitive scalp severity and symptomatology by using a new score. *J Eur Acad Dermatol Venereol.* 2011;25(11):1295–8.
32. Ding D-M, Tu Y, Man M-Q, Wu W-J, Lu F-Y, Li X, et al. Association between lactic acid sting test scores, self-assessed sensitive skin scores and biophysical properties in Chinese females. *Int J Cosmet Sci.* 2019;41(4):398–404.

33. Misery L, Jourdan E, Huet F, Brenaut E, Cadars B, Virassamynaïk S, et al. Sensitive skin in France: a study on prevalence, relationship with age and skin type and impact on quality of life. *J Eur Acad Dermatol Venereol*. 2018;32(5):791–5.
34. Dalgard FJ, Gieler U, Tomas-Aragones L, Lien L, Poot F, Jemec GBE, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. *J Invest Dermatol*. 2015;135(4):984–91.
35. Chren M-M, Weinstock MA. Conceptual issues in measuring the burden of skin diseases. *J Investig Dermatol Symp Proc*. 2004;9(2):97–100.
36. Misery L, Jourdan E, Abadie S, Ezzedine K, Brenaut E, Huet F, et al. Development and validation of a new tool to assess the Burden of Sensitive Skin (BoSS). *J Eur Acad Dermatol Venereol*. 2018;32(12):2217–23.

Chapter 25

Practical Use and Significance of Transepidermal Water Loss Measurements



Truus Roelandt and Jean-Pierre Hachem

25.1 Introduction

The purpose of the skin is to operate as physical barrier preventing water loss body and avoiding penetration of toxic substances as well as infectious agents [1]. Permeability barrier is primarily made possible by the integrity, structure, and function of the stratum corneum (SC) [2]. Therefore measuring transepidermal water loss (TEWL) and skin hydration have been widely used to evaluate skin barrier function [3]. TEWL is the quantity of water that evaporates through skin due to a water vapor pressure gradient from the inner space of the body to the environment. It typically illustrates skin barrier function as well as integrity. The average loss of TEWL in humans is thought to be around 300–400 mL/day, but it also can be influenced by environmental and/or intrinsic factors. Consequently, measuring TEWL allows to unravel disturbances in the skin barrier in an early stage prior to visible signs of diseases. While normal skin conditions allow water evaporation in small amounts, atopic skin type is characterized by much TEWL event at clinically non-eczematous skin. By this means TEWL is a valuable tool to investigate the skin status under basic, diseased, and experimental conditions: i.e., allergic tests, occupational medicine, monitoring of barrier healing processes in damaged skin condition (experimentally induced or active disease), and testing for the effectiveness of cosmetic products [4].

T. Roelandt
AZ Nikolaas Campus Beveren, Beveren, Belgium

J.-P. Hachem (✉)
Centre Hospitalier Emile Mayrisch, Esch Sur Alzette, Luxembourg
e-mail: jphachem@pt.lu

25.2 The Bricks-and-Mortar Model

An intact stratum corneum (SC) is crucial to maintain an epidermal permeability barrier that prevents the loss of fluids, electrolytes, and other molecules from within the body as well as the penetration of xenobiotics (microorganisms, allergens, toxic materials, etc.) [5, 6]. It is currently viewed as a layer of protein-enriched corneocytes embedded in a lipid-enriched intercellular matrix, the so-called “bricks-and-mortar model” [7]. 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 [8] made by the secretion of LB at the level of the SC/SG junction [9]. All these elements lead to the formation of a competent barrier that controls TEWL. In order to implement a competent barrier, the epidermis constantly replenishes itself by a process of homeostasis. First, keratinocytes detach from the basal layer to generate transitional keratinocytes. Second, these keratinocytes follow the route of programmed differentiation: the differential expression of keratins, the formation of keratohyalin granules, and the assembly of lamellar body (LB). Finally and as a last step, cell death (i.e., physiologic apoptosis) occurs in the uppermost granular layer of the epidermis leading to SC formation. This final process of terminal differentiation requires LB dumping, nuclear defragmentation, and cornified envelope (CE) formation [10]. At the leading edge of the epidermis, the SC generates an efficient barrier that protects the internal milieu of the organism from the external desiccating environment and controls TEWL [11].

25.3 Instrumental Assessment of Epidermal Barrier Status and Repair

25.3.1 *Industrial Vs. Legislation Needs to Measure Barrier Status and Repair*

European Union (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 record and can be subject to 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 [12, 13].

25.3.2 *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 [14]. 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.

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 draft 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 25.1. 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

Table 25.1 Guidelines for TEWL measuring

1.	Set the temperature (20–22 °C) and humidity (45–55% RH) of the experimental environment
2.	Acclimatize the subject during 15 min before measuring TEWL
3.	Acclimatize the subject during 15 min before measuring TEWL Body sites areas are chosen symmetrically to avoid body site-related differences
4.	Carry out measurements with careful handling of the TEWL probe
5.	Warm-up the probe to body site temperature of the subject
6.	Apply constant light pressure of the probe to the skin
7.	Maintain a horizontal contact of the probe with the skin
8.	Measure TEWL of the tested site for 30 s or till the TEWL is stabilized (e.g., plateau levels of TEWL are reached)
9.	Take a 1-min break between two measurements
10.	Measure immediately prior to product application and at least 3 h (>6 h is recommended) after application
11.	It is recommended to perform multiple testing (e.g., capacitance, pH) to backup TEWL data

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. 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 – TEWL₀)/TEWL₀**

25.3.3 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 [15]. 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 [16].

25.3.4 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 [17]. 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 [3]. 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.

25.3.5 Few Examples of Currently Used Instruments for Measuring Water Loss

In vivo TEWL can be measured by two different water sampling techniques: (1) open-chambers are open to the surrounding atmosphere and thus are easily influenced by external air convection and turbulence; (2) closed-chambers are enclosed from the surrounding atmosphere and measurements are thus not influenced by external air convection and turbulence [18] (Table 25.2).

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:

$$dm / dt = D \times A \times dp / dx,$$

where: A = surface (m^2), m = water transported (in g), t = time (h), D = diffusion constant ($=0.0877 \text{ g/m.h(mmHg)}$), p = vapor pressure of the atmosphere (mmHg), x = distance from skin surface to point of measurement (m).

Table 25.2 Commercially available TEWL measurement instruments

Instrument	Open/Closed	Manufacturer
Aquaflux	-/+	Biox Systems Ltd., United Kingdom
AS-CT1	-/+	Asahi Biomed Company Ltd., Japan
DermaLab	-/+	Cortex Technology, Denmark
H4300*	-/+	Nikkiso-YSI, Japan
Tewameter	-/+	Courage & Khazaka, Germany
VapoMeter SWL3	-/+	Delfin Technologies, Finland

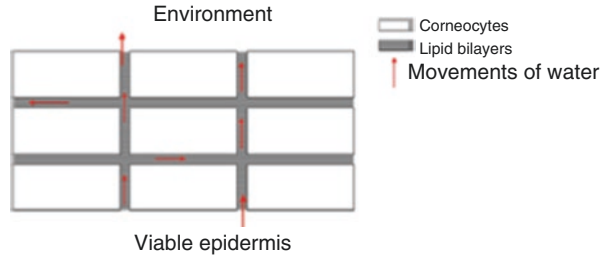
The TEWA meter is manufactured by Courage and Khazaka Electronic, Köln, Germany. The water vapor pressure gradient is indirectly measured by two pairs of a combined thermistor and hygrosensor, present at two different heights inside a hollow cylinder (height 2 cm, diameter 1 cm). The probe head is placed 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.

25.4 Limits and Pitfalls of TEWL Measurement

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. 25.1). 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 (i.e., a pathological condition). Two examples from our experimental work sustain this hypothesis [19]. Both protease-activated receptor (PAR)-2 [20] and caveolin-1 knockout mice [19] 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

Fig. 25.1 The bricks-and-mortar model. *Arrows* indicate movement of “insensible water loss” from the inside to the outside



[19, 21]. 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

1. Hachem JP. The two compartment model of the stratum corneum: biochemical aspects and pathophysiological implications. *Verh K Acad Geneesk Belg.* 2006;68(5-6):287-317.
2. Elias PM, Menon GK. Structural and lipid biochemical correlates of the epidermal permeability barrier. *Adv Lipid Res.* 1991;24:1-26.
3. Fluhr JW, Feingold KR, Elias PM. Transepidermal water loss reflects permeability barrier status: validation in human and rodent in vivo and ex vivo models. *Exp Dermatol.* 2006;15(7):483-92.
4. Schmuth M, Feingold KR, Elias PM. Stress test of the skin: the cutaneous permeability barrier treadmill. *Exp Dermatol.* 2020;29(1):112-3.
5. Corcuff P, Fiat F, Minondo AM. Ultrastructure of the human stratum corneum. *Skin Pharmacol Appl Skin Physiol.* 2001;14(Suppl 1):4-9.
6. Ponc M, et al. Barrier function in reconstructed epidermis and its resemblance to native human skin. *Skin Pharmacol Appl Skin Physiol.* 2001;14(Suppl 1):63-71.
7. Elias PM. Epidermal lipids, barrier function, and desquamation. *J Invest Dermatol.* 1983;80(1 Suppl):44s-9s.
8. Elias PM. Stratum corneum architecture, metabolic activity and interactivity with subjacent cell layers. *Exp Dermatol.* 1996;5(4):191-201.
9. Elias PM, et al. The secretory granular cell: the outermost granular cell as a specialized secretory cell. *J Invest Dermatol Symp Proc.* 1998;3(2):87-100.
10. Demerjian M, et al. Acute modulations in permeability barrier function regulate epidermal cornification: role of caspase-14 and the protease-activated receptor type 2. *Am J Pathol.* 2008;172(1):86-97.
11. Roop D. Defects in the barrier. *Science.* 1995;267(5197):474-5.

12. Meloni M, Berardesca E. The impact of COLIPA guidelines for assessment of skin compatibility on the development of cosmetic products. *Am J Clin Dermatol*. 2001;2(2):65–8.
13. Rogiers V, E. Group. EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences. *Skin Pharmacol Appl Skin Physiol*. 2001;14(2):117–28.
14. de Paepe K, et al. Claim substantiation and efficiency of hydrating body lotions and protective creams. *Contact Dermatitis*. 2000;42(4):227–34.
15. Hachem JP, et al. pH directly regulates epidermal permeability barrier homeostasis, and stratum corneum integrity/cohesion. *J Invest Dermatol*. 2003;121(2):345–53.
16. Hachem JP, et al. Sustained serine proteases activity by prolonged increase in pH leads to degradation of lipid processing enzymes and profound alterations of barrier function and stratum corneum integrity. *J Invest Dermatol*. 2005;125(3):510–20.
17. Chilcott RP, et al. Transepidermal water loss does not correlate with skin barrier function in vitro. *J Invest Dermatol*. 2002;118(5):871–5.
18. du Plessis J, et al. International guidelines for the in vivo assessment of skin properties in non-clinical settings: part 2. Transepidermal water loss and skin hydration. *Skin Res Technol*. 2013;19(3):265–78.
19. Roelandt T, et al. The “caveolae brake hypothesis” and the epidermal barrier. *J Invest Dermatol*. 2009;129(4):927–36.
20. Hachem JP, et al. Serine protease signaling of epidermal permeability barrier homeostasis. *J Invest Dermatol*. 2006;126(9):2074–86.
21. Rattenholl A, et al. Proteinase-activated receptor-2 (PAR2): a tumor suppressor in skin carcinogenesis. *J Invest Dermatol*. 2007;127(9):2245–52.

Chapter 26

Compliance Check with On-Site Measurements: Advanced Ways in Product Testing



Gabriel Khazaka

Core Messages

- Addition to classical laboratory test.
- New approach by gathering measurement data at the volunteers' everyday environment.
- Method to overcome geographical limitations of the acquired volunteers.
- Interesting information on the products' properties in their normal use environment.

26.1 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 at least 30 min) the stratum corneum hydration or other skin parameters will be measured on the skin area where 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 [1].

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

G. Khazaka (✉)
Courage + Khazaka Electronic GmbH, Köln, Germany
e-mail: gkhazaka@courage-khazaka.de

other important factors which might affect the efficacy of the product (e.g., skin cleansing, food and liquid intake, sleeping habits, sun exposure).

The volunteers will be invited to a study visit at the testing laboratory at predefined 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 the study parameters, e.g., stratum corneum hydration, will be measured. The results T0, T1 to Tx are statistically evaluated and 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 great 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 verified by the laboratory. Few studies have been published on the compliance in application of sun screens and protective creams [2–4]. In general, the longer the time period between the different follow-up tests, the lower the compliance that can be assumed.

The product performance during the course of the day under the normal real-life conditions of the volunteer, e.g., at home, at work, and during sports, is 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 close to 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 multicenter tests on a worldwide basis due to the extremely high costs and the availability of test sites and volunteers.

Thus, it could be worthwhile 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 might be considered as complementary to established testing methods and trial design with real-life data.

26.2 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. 4). 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 cost factor, one of the most important reasons is the lack of measuring devices allowing the consumer to carry out tests by themselves in her/his daily routine and life environment. Today, technical devices are a part of everyday living, not only at work but also at home (fitness watches, smartphones, computers, 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.

Taking this into account, manufacturers of skin testing devices have started designing reasonably priced measuring instruments that, following a short training (introduction and instruction), can be handled by the volunteers of clinical trials. The use of such measuring devices will provide additional information about the properties of the skin surface after application of products and during the 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 have often a genetic background and are accumulated in specific families, e.g., atopic dermatitis [5]. Researchers might study these families in their home region.

26.3 The Device Corneometer® Mobile Data Collector DC 3000

The Corneometer® is the most commonly used device for measuring stratum corneum hydration worldwide. It has been used in cosmetic testing and dermatological trials since more than 30 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 [6–8].

A small portable Corneometer® has been developed for field studies of stratum corneum hydration. The volunteer code 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 by the lab and operated at all conceivable locations, which might contribute to the geographic expansion of recruiting volunteers. It can be sent all over the world together with the testing products and an operation manual for subject training (Figs. 26.1 and 26.2).

The Corneometer® Mobile Data Collector DC 3000 can be easily used by the volunteers. With each measurement, the assessed hydration value, date and time of the specific measurement and the ambient conditions (temperature, relative humidity) retrieved from the connected sensor are stored in the device. Depending on 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). Up to 3000 measurements can be collected and stored during a long-term study.



Fig. 26.1 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



Fig. 26.2 Corneometer® Mobile Data Collector DC 3000 in the protected transportation case, ready to be shipped to the volunteers or be handed-out at the test site

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 transmitted by the volunteer to the laboratory during the study via internet by connecting the device to a computer. The saved data cannot be manipulated by either the volunteer or the laboratory.

26.4 Conclusion/Perspective

The combination of new technologies and the above-described ideas for field measurements enable new developments to gain 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 radius for recruiting volunteers can be enlarged up to a worldwide scale. The devices should of course only be handed over/sent to reliable and trained 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, thus increasing compliance.

New device generations could feature a connection via Bluetooth to smartphones, using a dedicated APP that will not only transfer the data and additional information in real time to a central IT system comparable to a sports watch. It might also be possible that such an APP will remind the subject to apply the products, to measure the skin, and include subjective questions about the product performance. One important aspect of such developments is to implement a secure data transfer protocol from the measuring device via smartphone to the central data storage. The legislation has clear rules for such data safety procedures [9–11].

In the future, other skin measurement parameters could be included in these concepts (e.g., melanin, erythema, and sebum measurements or measurement of mechanical properties) delivering additional information about the performance of the tested products.

References

1. Bornkessel A, Flach M, Arens-Corell M, Elsner P, Fluhr JW. Functional assessment of a washing emulsion for sensitive skin: mild impairment of stratum corneum hydration, pH, barrier function, lipid content, integrity and cohesion in a controlled washing test. *Skin Res Technol.* 2005;11:53–60.
2. Armstrong AW, Watson AJ, Makredes M, Frangos JE, Kimball AB, Kvedar JC. Text-message reminders to improve sunscreen use: a randomized, controlled trial using electronic monitoring. *Arch Dermatol.* 2009;145:1230–6.
3. Neale R, Williams G, Green A. Application patterns among participants randomized to daily sunscreen use in a skin cancer prevention trial. *Arch Dermatol.* 2002;138:1319–25.
4. Wigger-Alberti W, Caduff L, Burg G, Elsner P. Experimentally induced chronic irritant contact dermatitis to evaluate the efficacy of protective creams in vivo. *J Am Acad Dermatol.* 1999;40:590–6.
5. Lindh JD, Bradley M. Clinical effectiveness of moisturizers in atopic dermatitis and related disorders: a systematic review. *Am J Clin Dermatol.* 2015;16:341–59.

6. Breternitz M, Kowatzki D, Langenauer M, Elsner P, Fluhr JW. Placebo-controlled, double-blind, randomized, prospective study of a glycerol-based emollient on eczematous skin in atopic dermatitis: biophysical and clinical evaluation. *Skin Pharmacol Physiol*. 2008;21:39–45.
7. Crowther JM, Sieg A, Blenkinsop P, Marcott C, Matts PJ, Kaczvinsky JR, Rawlings AV. Measuring the effects of topical moisturizers on changes in stratum corneum thickness, water gradients and hydration in vivo. *Br J Dermatol*. 2008;159:567–77.
8. Fluhr JW, Darlenski R, Surber C. Glycerol and the skin: holistic approach to its origin and functions. *Br J Dermatol*. 2008;159:23–34.
9. BSI TR-02102-1. (Online; Last accessed January 19, 2020). <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TG02102/BSI-TR-02102-1.pdf>.
10. BSI TR-02102-2. (Online; Last accessed January 19, 2020). <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TG02102/BSI-TR-02102-2.pdf>.
11. General Data Protection Regulation (Online; Last accessed January 19, 2020). <https://gdpr-info.eu>.

Index

A

- Acne lesions assessment, 166–171
 - facial sites, 166
 - instrumental assessment, 169–170
 - lesion types, 166
 - nonfacial sites, 166
 - non-inflammatory lesions, 166
 - quality of life assessment, 165, 170–171
 - sensorial assessment, 166–169
 - disease-specific instrument, 167
 - investigator (physician)
 - assessment, 166
 - user-self assessment, 166
 - See also* Impure, acne-like skin
- Antiaging and antiwrinkle products, 150–156
 - non-invasive biophysical methods
 - SC hydration, electrical methods, 151
 - skin mechanical properties, 153
 - skin microcirculation, 152
 - skin pigmentation and color, 152
 - skin surface acidity (pH), 152
 - skin surface lipid, photometric assessment, 152, 153
 - skin surface topography, 150, 151
 - TEWL values, epidermal barrier function, 151
 - study design
 - anatomical test site, 155
 - data management and evaluation, 157
 - efficacy assessment, 156
 - product application, 156
 - study duration, 155
 - study population, 155
 - study schedule, 153, 154
- Antiperspirants
 - axillary sweat, gravimetric evaluation

- conditioning period, 233
- data evaluation, 234–235
- randomization, 234
- study population, 233
- sweat collection, 234
- biophysical methods, 232
- casting methods, 232
- gravimetric methods, 232
- multiple product test design, 235–236
- visual techniques, 231

B

- Barrier permeability assessment, *see* Transepidermal water loss (TEWL)
- Bricks model, 298, 303

C

- Carotenoids, 134, 151, 190, 191
- Case report form (CRF), 30, 31, 69, 76, 77, 109
- Central drugs standard control organization (CDSCO), 8
- Claim support
 - borderline claims, 45–46
 - challenges, 54, 56, 60
 - claim creation
 - communication style and medium, 48
 - consumer factors, 48–49
 - factors involved, 46, 48, 49
 - global/regional, 46–48
 - classification, CTPA/ASA approach, 50
 - definition, 44, 45
 - documentation, 58, 59, 230

- Claim support (*Cont.*)
 level of evidence
 established and widely available evidence, 51–52
 science/technology, 53
 specific evidence, 52–53
 quality of evidence
 consumer perception assessment, 53, 56, 57
 consumer studies, 54–56
 expert grading, 56
 instrumental evaluation, 57
 in vitro/ex vivo studies, 58
 sensory properties, 55
- Combination claims, 20, 50
- Common rule, 17
- Comparison claims, 20, 50
- Conditioners, 212, 214, 215, 220–223, 225
 cationic surfactants, formulation, 214
 characteristics, 212, 214, 215
 depositing variants, 215
 product rheology, 215
 See also Hair; Shampoos
- Confocal laser scanning microscopy (CLSM), 239, 253–254
- Consumer studies, 56
- Contrast-enhanced phototrichogram (CE-PTG), 239, 247–248, 251, 255, 256
- Corneometer® Mobile Data Collector DC 3000, 307–308
- Cosmetic testing report
 good clinical research practice (GCRP), 111
 result publication
 abstract, 118
 acknowledgments, 119
 background information, 118
 discussion and conclusions, 119
 keywords, 118
 manuscript submission, 120
 material and methods, 118–119
 references, 119
 tables and figures, 120
 title page, 117–118
 structure and contents
 abbreviations and definition, 113
 appendices, 116
 discussion and conclusions, 116
 ethical considerations, 113
 introduction, 114
 investigational plan, 114–116
 investigators and study administrative structure, 113
 reference list, 116
 results, 116
 signatures, 116
 synopsis, 113
 table of contents, 113
 test objectives, 114
 title page, 112
- Council for International Organizations of Medical Sciences (CIOMS) guidelines, 17
- Cutaneous aging, 148
 See also Antiaging and antiwrinkle products, 148
- D**
- Daily hair count (DHC), 239, 241–243, 254
- Declaration of Helsinki, 8, 15, 17–20, 23, 27, 28, 76, 113, 117, 119, 172
- Deodorants
 Sniff test
 conditioning period, 237
 data evaluation, 237
 procedure, 237
 study population, 236–237
 test methods, 236
- Dermal irritation patch tests, 10
- DermaSpectrometer, 152, 193
- Dermometry
 evidence-based guidelines, 38
 good biometrological practice, 37–38
 skin bioengineering, 34
 standardization and quality controls, 37
 validation, methods and instrumentations, 34
- E**
- Emollients, 52, 127–145, 299
 See also Moisturizers
- Epidermal barrier status and repair, TEWL
 industrial vs. legislation needs, 298
 lotions and creams, barrier permeability barrier recovery, 299–300
 barrier recovery, 300
 emollient formulation, 300
 external factors, 299
 TEWL measuring steps and recommendations, 301
 time-zero-TEWL (TEWL0), 300

- Ethical aspects
 cosmetic testing, 19
 history
 Belmont report, 17
 CIOMS guidelines, 17, 18
 common rule, 17
 Declaration of Helsinki, 17
 medical ethics, 16
 Nuremberg code, 16
 research ethics, 16–19
 noninvasive skin measurements, 19
 principles
 conduct, 20
 Institutional Review Board/Independent
 Ethics Committee (IRB/IEC), 22
 investigational products, 22
 investigator, 21
 subjects/participants, 21
 techniques/devices, measurement, 22
 Extrinsic aging, 147–149
- F**
 Fitzpatrick classification scale, 191, 272, 275
- G**
 Glogau scoring, 148
 Good Clinical Practice (GCP)
 case report form (CRF), 30
 conduct standards, 30
 principles, 27–29
 protocol standards, 30
 standard operating procedures
 (SOPs), 29, 30
 study design, 30
- H**
 Haemoglobin, 170, 190–194
 Hair
 color, 209, 223, 224
 growth, 239–257
 influence of water, 210
 internal structure, 209
 medulla, 209, 253
 outer structure, 209
 shampoos and conditioners
 color fade, 223–224
 consumer evaluation
 approaches, 216–217
 hair shine, 222–223
 lubrication, 221
 moisturization, 224–225
 static electricity retardation, 222
 strength, 221–222
 technical evaluation
 approaches, 217–219
 technical functionality and consumer
 opinion, 215
 surface damage assessment,
 SEM, 219–221
 type, 55, 210, 213, 216, 218, 225, 240,
 248, 251
 See also Conditioners; Shampoo
 Hair growth assessment
 clinical appearance, 241–242
 clinical grading systems, 242
 confocal laser scanning
 microscopy, 253–254
 contrast-enhanced
 phototrichogram, 247–248
 daily hair count, 242–243
 electron microscopy, 253
 epiluminescence microscopy (*see*
 TrichoScan)
 global photographs, 251–253
 hair pull-test, 243–244
 hair wash test (HWT), 243
 hair weighing, 244
 optical coherence tomography (OCT),
 253
 phototrichogram, 246–247
 physiological factors, 240
 study design, 241
 trichogram, 244–246
 Unit Area Trichogram, 246
 Hair morphology assessment
 atomic force microscope (AFM), 178
 cuticle characterization, 177
 dynamic experiments
 optical interferometry
 roughness modification, 179
 vertical scanning interferometry
 mode, 179
 white light source, 179
 See also Optical interferometry, hair
 morphology
 scanning electron microscope (SEM), 178
 static experiments
 Hair pull test (HPT), 239, 240, 242–244
 Hair wash test (HWT), 239, 242, 243, 254
 Human repeat insult patch test, 11
 Hyperkeratosis, 198

I

- Impure, acne-like skin
 - acne lesions (*see* Acne lesions assessment)
 - oily skin (*see* Oily skin assessment)
 - product acceptance, 171
 - study background, 160
 - study design, 161–162
 - study objectives, 160
 - study products, 160–161
 - study subjects, 161
 - unwanted effect assessment
 - ethical principles, 172
 - statistical analysis, 172
 - study protocol, 172
- Ingredient claims, 50, 53
- Institutional Review Board/Independent Ethics Committee (IRB/IEC), 20, 22, 23
- Instrumental assessment
 - acne lesions assessment, 163–171
 - oily skin assessment, 163
- Intrinsic aging, 148, 149
- In-use irritation tests, 11
- Investigator global assessment (IGA) score, 168, 169

J

- Japan Cosmetic Industry Association (JCIA), 7

L

- Laboratory testing
 - accommodation and arrangement
 - data base/subject file room, 66, 67
 - general volunteer information sessions, 66, 68
 - investigator's office/examination room, 66, 68
 - lab room, 66, 68–69
 - reception/volunteer recruitment office, 66, 67
 - rest rooms and waiting area, volunteers, 66, 67, 69–70
 - room types, 66
 - staff office, 66, 69
 - test product/sample storage room, 66
 - ambient conditions, 70
 - quality management system
 - measurement devices, 71–72
 - standard operating procedures (SOPs), 65, 70
- Leeds Acne Grading Scale, 167, 168

M

- Measurement devices and methods
 - cost issues, 137
 - laboratory, 85, 86, 88–89
 - manufacturers and software packages, 89
 - practical use, 88–89
 - quality and quick service, 89
 - technical assessment
 - skin parameters, 86–87
 - standards and directives, 87–88
- Melanin, 87, 170, 189–194, 209, 267, 269, 270, 273, 281, 282, 309
- Mild acne, 168
- Ministry of Health, Labour, and Welfare (MHLW), 7
- Minolta spectrophotometer, 193
- Moderate acne, 161, 168
- Moisturizers, 50, 127–145, 274
 - characteristics, 128, 132
 - investigative method selection, 135–136
 - legislation and marketplace, 129–130
 - non-invasive biophysical methods
 - mechanical/visco-elastic properties, skin, 132, 135
 - outermost skin layers, electrical properties, 132–133
 - skin surface topography examination, 133–134
 - spectroscopic and optical properties, skin, 134
 - techniques, 131, 133, 134
 - transepidermal water loss assessment, 133
 - study designs
 - multiple application tests, 136–137
 - preliminary skin irritation test, 137
 - regression tests, 137
 - single application tests, 136
 - study population
 - general considerations, 138
 - number of subjects, 138
 - subject and environment-related variables, 138, 139
 - test site, 139–140
 - study procedure
 - data management and evaluation, 143
 - instructions, panelist, 140
 - measurement procedure, 142–143
 - product application, 142
 - washout phase, 140–141
- See also* Skin hydration
- Mortar model, 298, 303

N

- Non-invasive skin measurements, 19
- Nuremberg code, 16

O

- Oily skin assessment
 - image analysis techniques, 165, 170
 - instrumental assessment, 160, 164–165, 298–303
 - lipid absorbent tapes, 164
 - photometric techniques, 164
 - quality of life assessment, 165
 - sensorial assessment, 163, 166
 - skin glossymetry, 165
- Optical coherence tomography (OCT), 265, 266, 277–280
- Optical interferometry, hair morphology, 178
 - dynamic experiment
 - cuticle evolution, tensile stress, 178, 179
 - fibres used, 282
 - roughness modification, 179
 - static experiment
 - moisturizing, topographical modifications, 181
 - roughness reduction, 179
 - sampling, 29, 35, 301
 - scratch evolution
 - topographical representations, cuticle
 - vertical scanning interferometry mode, 179
 - white light source, 179

P

- Performance claims, 20, 50
- Permeability assessment, *see* Transepidermal water loss (TEWL)
- Photoaged skin, 148–150, 278
 - clinical, histological, and functional characteristics, 149
 - Glogau scoring, 148
 - See also* Antiaging and antiwrinkle products
- Phototrichogram (PTG), 240
- Physical global improvement score, 168
- Product testing, 43–46, 140
 - classical laboratory test
 - location, 88, 140, 144, 246
 - noncontrollable factors, 306
 - product performance, 162, 166, 306
 - protocol information, 305
 - stratum corneum hydration, 39, 97, 305–307

- Corneometer® Mobile Data Collector DC 3000, 307, 308
- field studies, 306–307, 309

R

- Regulatory aspects, 4–13
 - ASEAN countries, 4, 7, 46
 - Canada, 5, 7, 18
 - China, 5, 8, 46
 - cosmetic legislation, 4–5, 298
 - broad definitions, 4
 - Canadian system, 5, 254
 - Europe, 3–6, 8, 11, 12, 44, 45, 76, 88, 231, 241, 246, 247
 - India, 5, 8
 - Japanese system, 5
 - narrow definitions, 3, 4
 - universal similarities, market economies, 4
 - cosmetics, 3–8, 10–13
 - efficacy testing, 12, 19, 85, 86, 88, 91, 131, 153, 159, 171, 229, 232, 307
 - labeling and packaging, 12–13
 - safety study, 9–11
 - cosmetics legislation, 3, 5
 - European Union, 5, 6, 11, 18, 43, 76, 107, 231, 251
 - India, 5, 8, 254
 - Japan, 5, 7, 12, 18, 46, 88, 152, 193, 301
 - Mercosur countries, 4, 7
 - South Africa, 4, 7–8
 - USA, 3, 5–7, 13, 45, 88, 152, 165, 170, 241, 251, 254
- Research staff, 66, 75–77
 - investigator, 13, 15, 18, 21–23, 28, 31, 34, 68, 75–77, 80–82, 86–88, 104, 112, 113, 118, 120, 121, 130, 141, 142, 160, 161, 163, 166–168, 171, 172, 237, 240–242, 250, 251, 254–256
 - study nurse, 77

S

- Sensorial assessment
 - acne lesions assessment
 - disease-specific instrument, 167
 - investigator (physician) assessment, 166, 167
 - user-self assessment, 166
 - oily skin assessment, 163
- Sensory/aesthetic claims, 20, 50, 52

- Sensory perception
- sensitive skin, 263
 - contralateral comparison design, 263
 - eye characteristics, 260
 - large group approach, 262
 - mechanisms, 260, 261
 - neurological tests, 261
 - prevalence, 260
 - reactivity tests, 261, 262
 - screening tests, 263
 - sensory tests, 261
 - sex-discrepancy, 260
 - very well-defined groups, 262
 - skin irritation, 259
- Severe acne, 161, 166, 168
- Shampoos, 51, 211–213, 215, 220, 221, 223, 225, 253
- anionic surfactants, 213
 - conditioning and styling products, 211
 - cosurfactants, 211, 212, 215
 - 2-in-1 shampoo+conditioners
 - aesthetics factors, 213
 - formulation stability, 213
 - manner of usage, 213
 - silicone oil, 212, 213
 - stability and formulation, 212, 213
 - surfactant concentration, 211, 213
 - viscosity, 211–213
- See also* Hair
- Singapore Health Sciences Authority (HSA), 7
- Skin colour and pigmentation
- aged skin, 152
 - carotenoids, 190, 191
 - constitutional skin colour, 191
 - factors influencing, 195
 - grading system, 192
 - hemoglobin, 190, 191
 - melanin, 190–191
 - reflectance spectrophotometric
 - evaluation, 192–194
 - chromameter, 193, 194
 - DermaSpectrometer, 152, 193
 - Minolta Spectrophotometer, 193
 - skin reflectance spectroscopy, 193
 - study procedure, 194–196
- Skin hydration, 39, 94, 127–130, 133–135, 137–144, 274, 297
- complex formation, 130
 - environment-related variables
 - air flow, air-conditioning system, 141
 - air humidity, 141
 - biorhythms, 141
 - direct light exposure, 141
 - panelists acclimatization
 - period, 141–142
 - room temperature, 141
 - epidermal barrier
 - external stressors, 128
 - mechanisms, 128, 129
 - stratum corneum (SC), 128, 129
 - water distribution, 129
- See also* Moisturizers
- Skin testing and measurements
- environment-related variables
 - air convections, 93
 - ambient air humidity, 93–94
 - ambient air temperature, 93, 94, 141
 - light sources, 94
 - skin cleansing, 94
 - individual-related variables
 - age, 95
 - anatomical sites, 95, 96
 - circadian rhythm, 97
 - intra and interindividual variation, 97
 - race and sex, 95
 - skin damage and diseases, 97
 - skin surface temperature, 96
 - sweating, 96
 - instrument-related variables
 - contact time, 93
 - surface area, 92
 - variability, start-up, and use, 92
- Sniff test, *see* Deodorants
- Sodium lauryl sulphate (SLS) test, 39, 97, 133, 137, 141, 261, 262, 274, 300
- Standard operating procedures (SOPs), 29–30, 65, 70–72, 76, 77, 109
- State food and drug administration (SFDA)
- Study design
- blinding, 104, 108
 - data entry and processing, 109
 - ethical review and scientific validity, 106
 - inclusion/exclusion criteria, 107
 - measurement variables and
 - endpoints, 104–105
 - randomized controlled study, 103
 - statistics
 - sample size calculation, 105
 - statistical methods, 105
 - study documents
 - case report form (CRF), 109
 - study protocol, 107
 - subject diary, 109
 - subject information sheet, 108
 - study objectives, 104
 - test products, 106

Sunscreens

- hyperkeratosis, 198
- sun protection factor (SPF)
 - classification, 202
 - factors influencing, 201, 202
 - international methods, 201
 - in vitro determination, 202
 - MED determination, 199
 - sunscreen application, 199–201
 - UVB protection, 198, 199
 - UVA protection factor, 201, 202

T

- Testing population recruitment
 - acquisition criteria and screening, 82
 - advertisements, 80–81
 - product-related information and compliance, 82
 - telephone recruitment, 81
 - third parties/intermediaries, 81
- Transepidermal water loss (TEWL), 133, 151

- Bricks and Mortar model, 298
 - closed chamber method, 301, 302
 - epidermal barrier status and repair (*see* Epidermal barrier status and repair, TEWL)
 - limits and pitfalls, 302–303
 - open chamber method, 301
 - validity, 301
 - Trichogram (TG), 239–241, 244–248, 254, 255
 - TrichoScan
 - black hair dye, 250
 - digital image and algorithmic analysis, 249
 - disadvantage, 251
 - hair growth parameters, 249
 - landmarks, red tattoo ink, 250
 - software validation, 250
- U**
- Unit Area Trichogram (UATG), 240, 246, 247, 255