

## Guidelines on Cosmetic Efficacy Testing on Humans. Ethical, Technical, and Regulatory Requirements in the Main Cosmetics Markets

Vincenzo Nobile\*

*Farcoderm s.r.l., Member of Complife Group, Via Mons Angelini, 21, 27028 San Martino Siccomario, Pavia, Italy*

\*Corresponding author: Vincenzo Nobile, Farcoderm s.r.l., Member of Complife Group, Via Mons Angelini, 21, 27028 San Martino Siccomario, Pavia, Italy, Tel: +39-0382-25504; E-mail: [vincenzo.nobile@farcoderm.com](mailto:vincenzo.nobile@farcoderm.com)

Received date: January 18, 2016; Accepted date: February 05, 2016; Published date: February 07, 2016

Copyright: © 2016 Nobile V. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

The main regulatory frameworks governing the cosmetic industry dates back to 1938 in the United States (US) and 40 years later in Europe (EU). Since then, both the US and EU cosmetic legislations have inspired the regulatory framework of a number of countries working toward harmonisation of cosmetics legislation. During the years the requirements for the efficacy of cosmetic products have been implemented to adapt them to the state of the art; however no clear guidelines for efficacy testing on cosmetic products exist. The lack of guidelines and/or shared position on cosmetic testing represents the missing link between the regulatory requirements and the consumer protection from misleading claims in the real life. On the other side a regulatory claimed prerequisite is ineffective if clear and specific testing methodologies are not available to the cosmetic industry. This introduces a discretionary element decreasing the strength of the original regulatory requirement, having an impact on consumer protection from misleading claims, and sometimes decreasing the credibility of the cosmetic product in the marketplace. Initiatives aimed at developing clear, specific for cosmetic products, and effective guidelines should arrive from the academy, industry, and professional associations. This manuscript is aimed to give an overview of the main ethical, technical and regulatory concerns affecting the design of an efficacy study carried out on humans.

**Keywords:** Cosmetics; Efficacy testing; Regulatory framework; Ethics; Technical requirements

### Introduction

The protection of Consumers from misleading claims concerning efficacy and other characteristics of cosmetic products are the central core of the worldwide regulatory framework. Cosmetic products are required to be effective when used by Consumers under normal, labelled or foreseeable conditions of use [1]. In Europe, the regulation CE 655/2013 [2] clearly states "Claims for cosmetic products, whether explicit or implicit, shall be supported by adequate and verifiable evidence regardless of the types of evidential support used to substantiate them, including where appropriate expert assessments". The evidential support for cosmetic claims should take into account the state of the art practices, studies should be relevant to the product and to the benefit claimed, shall follow well-designed and well-conducted methodologies (valid, reliable and reproducible), and shall respect ethical considerations.

However, despite the increase of regulatory requirements only a few of standards and guidelines exist (Table 1). This difficulty, which has been noticed for quite some time, needs an international solution or agreement to standardize the technical requirements of cosmetic efficacy testing studies on humans. The EEMCO-group (European group of efficacy measurement of cosmetics and other topical products) has, from the mid 1990 published a variety of "so called" guidelines for several different skin measurement parameters [8-18]. This manuscript will introduce the ethical, technical and regulatory requirements to which cosmetic efficacy testing should be inspired.

### Ethical requirements

Ethical considerations are an essential part of any biomedical research involving human subjects. Studies involving skin measurement methods and testing of cosmetic products on humans are similar to biomedical research. They involve the use of human subjects as research subjects and deals with both basic (e.g. investigation on skin physiology or active ingredients) and applied (safety and efficacy testing of ingredients or finished products) scientific issues.

The fundamental document in the field of human research ethics is the Declaration of Helsinki, originally adopted at the 1964 World Medical Association (WMA) General Assembly in Helsinki (Finland) [19].

In the late 1970s, the Council for International Organizations of Medical Sciences (CIOMS) in association with WHO, undertook its work on ethics in relation to biomedical research in order to "indicate how the ethical principles that should guide the conduct of biomedical research involving human subjects, as set forth in the Declaration of Helsinki, could be effectively applied". The outcome of the CIOMS/WHO undertaking was, in 1982, the "Proposed International Ethical Guidelines for Biomedical Research Involving Human Subjects" guidelines. This document was revised many times until the new text, the 2002 text, superseding 1993 text. The 2002 guidelines, "International Ethical Guidelines for Biomedical Research Involving Human Subjects", consists of a statement of general ethical principles, a preamble and 21 guidelines, with an introduction and a brief account of earlier instruments and guidelines [20].

Standard/guideline	Description	Reference
Sensory analysis -- Methodology -- Initiation and training of assessors in the detection and recognition of odours"	ISO 5496:2006 describes several types of method for determining the aptitude of assessors and for training assessors to identify and describe odoriferous products.	[3]
Cosmetics -- Sun protection test methods -- Review and evaluation of methods to assess the photoprotection of sun protection products.	ISO/TR 26369:2009 reviews and evaluates the methods which are currently used to assess, for regulatory or self-regulatory purposes, the photoprotection of sun protection products applied on human body.	[4]
Cosmetics -- Sun protection test methods -- <i>In vivo</i> determination of the sun protection factor (SPF)	ISO 24444:2010 provides a basis for the evaluation of sunscreen products for the protection of human skin against erythema induced by solar ultraviolet rays	[5]
Cosmetics -- Sun protection test methods -- <i>In vivo</i> determination of sunscreen UVA protection.	ISO 24442:2011 provides a basis for the evaluation of sunscreen products for the protection of human skin against UVA radiation from solar or other light sources.	[6]
Final Rule 39: Labeling and Effectiveness Testing: Sunscreen Drug Products for Over-the-Counter Human Use	76FR35619 provides a basis for the evaluation of sunscreen products for the protection of human skin against erythema induced by solar ultraviolet rays.	[7]
EEMCO guidance for the assessment of skin colour.	The manuscript introduces the origins of skin colours and the basic principles of their measurement. Methods of colour reading are reviewed with particular insight into practical procedures, pitfalls and correct interpretation of data.	[8]
EEMCO guidance for the assessment of skin topography.	The manuscript reviews the advantages and drawbacks of the currently used methods to assess skin microrelief/wrinkles and lists some experimental recommendations to make the results of studies using these methods, more reliable.	[9]
EEMCO guidance to the <i>in vivo</i> assessment of tensile functional properties of the skin. Part 1: relevance to the structures and ageing of the skin and subcutaneous tissues.	The manuscript reviews the physiological variables, ageing and skin diseases in altering the tensile functions of the skin.	[10]
EEMCO guidance for the <i>in vivo</i> assessment of skin greasiness.	The manuscript reviews the advantages and drawbacks of the currently used methods to assess skin greasiness.	[11]
EEMCO guidance to the <i>in vivo</i> assessment of tensile functional properties of the skin. Part 2: instrumentation and test modes.	The manuscript reviews the advantages and drawbacks of the currently used methods (tensile, torsional, indentation, impact and elevation modes) to assess the tensile functions of the skin.	[12]
EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences.	The manuscript reviews the advantages and drawbacks of the currently used methods to assess transepidermal water loss.	[13]
EEMCO guidance for the measurement of skin microcirculation.	The manuscript reviews the advantages and drawbacks of the currently used methods to assess skin microcirculation.	[14]
EEMCO guidance for the <i>in vivo</i> assessment of skin surface pH.	The manuscript reviews the conditions to assess the pH of the skin.	[15]
EEMCO guidance for the efficacy assessment of antiperspirants and deodorants.	The manuscript reviews the advantages and drawbacks of the currently used methods to assess the efficacy of deodorants and antiperspirants.	[16]
EEMCO guidance for the assessment of hair shedding and alopecia.	The manuscript reviews the advantages and drawbacks of the currently used methods to assess skin hair shedding and alopecia.	[17]
Inter-instrumental variation of skin capacitance measured with the Corneometer	The manuscript review the advantages and drawbacks of the currently used methods to assess skin moisturization	[18]

**Table 1:** Standards and guidelines for efficacy testing of cosmetics on humans.

American regulations governing the conduct of biomedical research involving human participants were published in 1981 by the federal Department of Health and Human Services (DHHS, formerly the Department of Health, Education and Welfare). Most US health care institutions conducting research have now agreed to apply these federal regulations to all of their research protocols regardless of the funding source for a particular study. Additionally, numerous other federal agencies have adopted the DHHS regulations, as subsequently amended, as a Common Rule to protect human participants in any research protocol. Research involving the testing of investigational drugs or medical devices is regulated concurrently by the federal Food

and Drug Administration (FDA); the Common Rule and FDA requirements overlap considerably but are not completely identical [21]. The Food and Drug Administration's (FDA's) regulations for the conduct of clinical trials, which have been in effect since the 1970s, address both GCP and human subjects protection (HSP).

Studies on cosmetic active ingredients and finished cosmetic products should be carried out in accordance with the principles of the "Declaration of Helsinki" [19] and the "Good Clinical Practice" guidelines [22]; national regulations regarding human studies should also be considered. Efficacy testing should be carried out only when

there is evidence that the product is safe for its use on human volunteers and safety should always “co-tested” as a secondary objective in a cosmetic efficacy study (e.g., by observing any adverse reaction). As a consequence efficacy testing should be preceded by an extensive safety assessment of each cosmetic ingredient in the product formula and by safety testing (e.g. patch testing, photoirritation testing, etc.) in order to exclude local or systemic adverse reactions. The design of the study protocol should minimize the risks related to product(s) use/application and to protocol procedures. A risk/benefit ratio

analysis should be carried out during the design of the study protocol. However, differently of what happens for study carried with therapeutic interventions, the study should be carried out only if risks are absent or minimal. In Table 2 are summarized the minimum set of ethical requirements to which cosmetic efficacy testing should be inspired. Implementation of the ethical requirements reported in Table 2 is recommended in order to take into account specific needs or requirements of the study protocol.

<p>The study should be ethically and scientifically justifiable: Among the essential features of ethically justified research involving human subjects are that the research offers a means of developing information not otherwise obtainable, that the design of the research is scientifically sound, and that the investigators and other research personnel are competent. The methods to be used should be appropriate to the objectives of the research and the field of study. Investigators and sponsors must also ensure that all who participate in the conduct of the research are qualified by virtue of their education and experience to perform competently in their roles.</p>
<p>The study protocol should be submitted for review of its scientific merit and ethical acceptability to one or more Scientific Review and ethical review committees. Scientific review According to the Declaration of Helsinki, the research involving humans must conform to generally accepted scientific principles, and be based on a thorough knowledge of the scientific literature and other relevant sources. Scientific review must consider, inter alia, the study design, including the provisions for avoiding or minimizing risk and for monitoring safety. Ethical Review The ethical review committee is responsible for safeguarding the rights, safety, and well-being of the research subjects. 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.</p>
<p>All study participants should sign a written informed consent before the study begins. Informed consent is a decision to participate in research, taken by a competent individual who has received the necessary information; who has adequately understood the information; and who, after considering the information, has arrived at a decision without having been subjected to coercion, undue influence or inducement, or intimidation. Process obtaining informed consent is a process that is begun when initial contact is made with a prospective subject and continues throughout the course of the study. By informing the prospective subjects, by repetition and explanation, by answering their questions as they arise, and by ensuring that each individual understands each procedure, investigators elicit their informed consent and in so doing manifest respect for their dignity and autonomy. Language informing the individual subject must not be simply a ritual recitation of the contents of a written document. Rather, the investigator must convey the information, whether orally or in writing, in language that suits the individual's level of understanding comprehension. The investigator must then ensure that the prospective subject has adequately understood the information. The investigator should give each one full opportunity to ask questions and should answer them honestly, promptly and completely. Documentation of consent. The subject sign a consent form. Waiver of the consent requirement. Investigators should never initiate research involving human subjects without obtaining each subject's informed consent. Renewing consent. When material changes occur in the conditions or the procedures of a study, and also periodically in long-term studies, the investigator should once again seek informed consent from the subjects.</p>
<p>Participation in the study should be free. Subjects may be reimbursed for lost earnings, travel costs and other expenses incurred in taking part in a study. The payments should not be so large to induce prospective subjects to consent to participate in the research against their better judgment (“undue inducement”). Acceptable recompense. Study participants may be reimbursed for their transport and other expenses associated with their participation in the study. Those who receive no direct benefit from the study may also receive a small sum of money for inconvenience due to their participation in the study. Payments in money or in kind should not be so large as to persuade eligible subjects to participate in the study. Unacceptable recompense. Payments in money or in kind to research subjects should not be so large as to persuade them to take undue risks or volunteer against their better judgment. Payments or rewards that undermine a person's capacity to exercise free choice invalidate consent. Withdrawal from a study. A subject who withdraws from research for reasons related to the study, such as unacceptable side-effects of the tested product or who is withdrawn on health grounds, should be paid or recompensed as if full participation had taken place. A subject who withdraws for any other reason should be paid in proportion to the amount of participation.</p>
<p>Before any participant is exposed to the test product, all safety information regarding the product and its individual ingredients are assessed. All the precautions should be taken in order to avoid adverse skin reactions occurrence. The Declaration of Helsinki in several paragraphs deals with the wellbeing of research subjects and the avoidance of risk. Thus, considerations related to the well-being of the human subject should take precedence over the interests of science and society. Clinical testing must be preceded by a safety assessment (according to Annex 1 of Regulation CE 1223/2009), by adequate laboratory experimentation (when applicable), or screening tests (e.g. patch testing) to demonstrate a reasonable probability of success without undue risk. Every project should be preceded by careful assessment of predictable risks and burdens in comparison with foreseeable benefits to the subject or to others; physician-researchers must be confident that the risks involved have been adequately assessed and can be satisfactorily managed.</p>
<p>Individuals who are not capable of giving informed consent or research in populations and communities with limited resources should not participate in cosmetic products testing. Certain individuals or groups may have limited capacity to give informed consent either because, as in the case of prisoners, their autonomy is limited, because they have limited cognitive capacity or their resources are limited. All of these conditions limit the capacity to make an informed choice, are not meeting the standard of informed consent, and then are not ethically justifiable.</p>
<p>In case of non-expected/adverse skin reaction occurrence the medical investigating specialist should assess the severity of the reaction (reporting it in the data collecting sheet of the volunteer) and as a consequence start the appropriate therapy. Investigators should ensure that research subjects who suffer 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.</p>
<p>All the information about the study participant should be safeguarded confidentially. Research relating to individuals and groups may involve the collection and storage of information that, if disclosed to third parties, could cause harm or distress. Investigators should arrange to protect the confidentiality of such information by, for example, omitting information that might lead to the identification of individual subjects, limiting access to the information, anonymizing data, or other means. During the process of obtaining informed consent the investigator should inform the prospective subjects about the precautions that will be taken to protect confidentiality.</p>

Table 2: Ethical requirements for cosmetic efficacy testing Adapted from [20].

Even for clinical testing under a controlled environment, for example, by a cosmetic testing organization, the responsibility for the safety of the test product ultimately remains at the manufacturer. Often, a manufacturer wants to keep confidential the existing data on the test product (e.g., ingredients). However, the testing organization should at least insist on a confirmation by the manufacturer or by its safety assessor that the safety assessment and toxicological profile of the test product was considered and the test product was judged safe under the normal conditions or foreseeable conditions of use. Furthermore, it should be confirmed by the manufacturer that the test product conforms to the local cosmetics laws. In general, it should be kept in mind that the ingredients list 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 this is contained in the test product. Second, an experienced testing organization can give recommendation 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.

Even if, an independent safety review is not legally required for cosmetic tests, this should always be considered, at least if a residual risk for the volunteers is suspected, for example, if the test product contains novel ingredients, or if invasive or stressful subject procedures are planned for the study.

### Technical requirements

In order to promote sound scientific design of research, each cosmetic efficacy testing study protocol should include a series of detailed information on how to carry out the study (Table 3) [23].

The saying “the claim dictates the test” is a good starting point when considering whether and what type of clinical study is needed. Different study designs provide different strength and limitation. The gold standard of clinical study designs is the randomized controlled study (RCT). According to this study design, treatments (e.g. active ingredient and placebo or active ingredient and benchmark) are allocated to subjects in a random and unpredictable sequence. 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. In order to reduce the variability due to differences between individuals, designs with intraindividual comparison (e.g. half-face/half-body application, multiple applications on forearm, etc.) of test products are generally preferred in cosmetic testing, if feasible. A random allocation of the test products or untreated control/placebo product should be used in these designs. Blinding is also one of the major concerns in study designing. In addition to RCTs there is a number of study types that the investigator can choose in the clinical trial field that are also applicable to cosmetic efficacy studies. The choice of the 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. The main types of studies used in the cosmetic efficacy testing field include: i) sensory properties studies, ii) consumer studies, iii) expert grading studies, and iv) instrumental measurement studies.

The sensory properties of a product are fundamental in cosmetic science and can help in the understanding of consumer perception related to consumer needs and claimed benefits. A sensory property contributes substantially to whether a product is liked and thus used by

consumers. For example, if a cosmetic product is perceived as unpleasant to touch, it is unlikely to be used voluntarily even if it is potentially beneficial to skin health. Additionally, some skin care products are designed such that their primary benefit is perceptual as opposed to tangible, benefits to the skin. The assessment method used will depend on the sensory attributes being examined and the claim required [24]. Trained panels of volunteers with high levels of sensory acuity can define the language and descriptors of key performance attributes of products [25]. Trained panels are usually valuable in prototype testing, comparative properties, and market comparison. However, trained panels assessment does not necessarily equate to consumer preferences. Alternatively, naïve panels can provide useful spontaneous responses to product concepts. In both cases regulator’s confidence for this type of testing is from low to very low.

Consumer studies are primarily used to mimic the consumer’s response for the cosmetic product (e.g. self-perceived efficacy, sensory properties, consumer’s attitudes toward buying products, and consumers purchase intentions for the products). Consumer’s studies are performed in real life conditions on a representative panel of the target population. The test product should be supplied in an anonymous pack in order to avoid any bias related to the brand strength (“halo” effects). Consumer preferences are reported then, after a variable period of use, on a self-assessment questionnaire; online surveys, can also be carried out. One of the frequent challenges of consumer’s studies is that the desired claim has not been captured exactly within the question choice. Sample size is another major concern. Other important considerations include questionnaire design/layout, avoiding leading questions, and ensuring balance in scale of response. Consumer studies are distrusted by many regulators and should be not used alone for efficacy testing but as a part of an expert grading or an instrumental study.

Expert grading is carried out by a professional (dermatologist, make-up artist, hairdresser, etc.) on a variety of characteristics [26-28]. Digital pictures of the test area should be taken under standard and reproducible light conditions. Picture scoring should be carried out by the expert randomly and under blind conditions. The level of expertise must be consistent within any given study and training or other validation essential whether conducted by a dermatologist, ophthalmologist or non-clinical scientist conducting the study. One of the frequent challenges of expert grading studies is the reference scale used for scoring. As a general rule, scoring scales should take into account consumer perception and the visibility of the effect. Well-designed studies based on expert grading are accepted by regulators.

Instrumental measurements were made possible by the birth of bioengineering techniques and has continued to grow in importance in assessing skin characteristics [8,10-15,18]. The technology of these instruments is constantly being updated and their accuracy is further improving. The area of skin imaging has exploded in recent years with 2D and 3D image analysis devices and software readily available to quantify features such as size of pores; eye bags; facial wrinkles; scalp hair; cellulite [29-36]. All of these approaches provide quantitative results that can be further exploited as percentage variation (e.g. increased moisturization by 20%, reduction of wrinkles by 30%, etc.). This type of approach has been criticized on the basis that changes measured can be too small to be perceptible by the consumer. 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-tested.

The choice of the “right” study population is critical to the study. Investigators should be inclusive in selecting participants and inclusion and non-inclusion criteria should be clearly defined in the study protocol. Inclusion and non-inclusion criteria have the joint goal of identifying a population in which it is relevant to assess the impact of the cosmetic product use on outcomes. The design of the inclusion criteria should take into account: i) the target population (those who were intended to use the cosmetic product), ii) the maximization of the generalizability of the study finding, iii) the complexity and cost of recruitment. For example, if the outcome of interest is related to the assessment of the moisturizing efficacy, it is necessary to enrol subjects showing the clinical signs related to skin dryness or to skin tendency to be dry. Borderline, altered and/or pathological skin conditions (e.g. xerosis) should be avoided since this is not the field of application of cosmetics. On the same way, the design of the non-inclusion criteria should be parsimonious because unnecessary exclusions may diminish the generalizability of the results, make more difficult to recruit subjects and increase the complexity and cost of recruitment [37]. Particular attention should be paid to stratification by particular characteristics (e.g. age, sex, cosmetic preferences, etc.). Stratification could be desired in certain study designs (e.g. consumer studies) and undesired in other study designs. When stratification is not properly

addressed study findings are related to the extent of the stratification limiting the generalizability of the study findings. At baseline, investigators should collect enough information (e.g. age, gender, ethnicity, skin conditions) in order to describe the study participant and to help other to judge the generalizability of the findings. On the other side a well-documented baseline description of the study participants allows the comparability of the study groups (for parallel study design) or between studies.

Another important concern related to study population is the sample size. Elements of sample size calculation are: i) the estimated outcome, ii) the  $\alpha$  (type I) error level, iii) the statistical power ( $\beta$  or type II error level), and iv) the standard deviation of the measurement [38]. Often, estimated outcomes and standard deviation of the measurement can be obtained from historical data obtained with similar measurement procedures, since expected effects are often similar within a similar range. When no existing data are available, a pilot study should be conducted. Type I and Type II errors are usually set at 0.05 (5% chance of a “false positive”) and  $\leq 0.2$  (20% chance of a “false negative”; at least 80% statistical power) respectively. Studies with a limited and/or inadequate sample size may produce misleading conclusions, are time and money consuming and unethical.

Item	Details
<b>General Information</b>	General information should include: i) protocol title, ii) protocol identifying number (and date), iii) name and address of the sponsor, iv) name, title, and address(es) of the investigator(s) who is (are) responsible for conducting the study, v) the address and telephone number(s) of the study site(s), and the vi) name(s) and address(es) of the clinical laboratory(ies) and other medical and/or technical department(s) and/or institutions involved in the study.
<b>Study objective and purposes</b>	The study protocol should give as more as possible a detailed description of the study objective and purposes.
<b>Study design</b>	The description of the study design should include the following information: i) a description of the primary and secondary endpoints to be measured during the study, ii) a description of the type/design of the study (e.g. placebo-controlled, uncontrolled, double blind, etc.), iii) a detailed diagram of the study flow, iv) a chart reporting the schedule of the assessments, v) the expected duration of the study, vi) a description of the “stopping rule” or “discontinuation criteria” for individual subjects, and vii) the maintenance of study treatment randomization codes and/or the measures taken in order to safeguard the study blindness (when applicable).
<b>Selection and withdrawal of Subjects</b>	Inclusion, non-inclusion and withdrawal criteria should be clearly stated in the study protocol.
<b>Treatment of Subjects</b>	The description of the treatment of subjects should include the following information: i) the name (including the reference and the batch numbers), ii) the qualitative INCI* formula, iii) a statement that the product formula is compliant to Regulation EC 1223/2009, iv) a statement that the product safety was addressed according to Regulation EC 1223/2009, v) a statement that the product is safe for its use on humans, and vi) the way of use (quantity, frequency, etc.) of the test product.
<b>Assessment of Efficacy</b>	The specification, the methods for assessing, recording, and analyzing of efficacy parameters should be clearly described in the study protocol.
<b>Assessment of Safety</b>	The specification, the methods for assessing, recording, and analyzing of safety parameters should be clearly described in the study protocol.
<b>Statistics</b>	A description of the descriptive and inferential statistical methods to be employed, the level of significance to be used, the procedures for accounting for missing, unused, and spurious data, the selection of subjects to be included in the analyses (e.g., all randomized subjects, all dosed subjects, all eligible subjects, evaluable subjects) should be clearly described in the study protocol.
* INCI, International Nomenclature of Cosmetic Ingredients	

**Table 3:** Study protocol minimum information.

The outcome measurements should take into account the main efficacy claim or safety issue to be addressed by the study. The choice of the outcome measurements has an impact on the feasibility of the study in answering the question as well as on cost. Studies should include several measurements to increase the study robustness and the opportunity for secondary analysis. However one or two outcomes, or

“primary endpoint(s)”, must be chosen to assess the extent of the product effect and to give proof of the claimed effect. For example, the efficacy of a moisturizing product is properly assessed by the measurement of the stratum corneum water content; while the measurement of the product effect on skin barrier, measured by means

of evaporimetry, can be used to assess the mechanism of action of the product.

Products can be applied by the investigator (controlled, short-term tests under the expert supervision) or can be used at home (long-term “use test”). Products are applied by the investigator to assess products efficacy under standardized/controlled conditions. The quantity of product applied should be calculated based on the estimated daily exposure for the cosmetic [39-41] and skin surface area [41] in order to calculate the rate of product application (in mg/cm<sup>2</sup>). Usually in short-term tests products are applied to a 2 mg/cm<sup>2</sup> application rate based on the application rate indicated by an ISO standard to assess the sun protection factor [5]. In the long-term use test, products should be used at home by subjects under real-life conditions of use. Products way of use and frequency of use should be shared with the Sponsor of the study and close to the real and normal condition of use. The compliance of subjects to treatment should be carefully monitored. Techniques to assess products vary from very simple techniques (e.g.

product weight assessment at the end of the study period, questionnaire) to very complicate and expensive monitoring systems (e.g. recoding cameras during product application).

The characteristics of the laboratory in which the cosmetic products are clinically tested have a large impact on the overall quality of the results. The choice of the testing laboratory is then crucial to have a robust study. Table 4 reports the minimum requirements to be checked during an official audit to the testing laboratory facility(ies). Beyond ISO 9001 certification, particular attention should be paid on the existence of a well-documented list of standard operative procedures. Standardization of the measurement technique is fundamental in order to have reproducible and reliable data. Instrument-, environmental-, and individual-related variability can affect the reproducibility of the measurement and then the accuracy of the obtained data. A detailed description of testing laboratory requirements is discussed by Wunderlich [42].

Requirement	Description
<b>Accommodation and arrangement of volunteers</b>	In order to facilitate volunteers enrolment the laboratory facility(ies) should be located in an area that can be easily reached (e.g. access to public transport and sufficient car parking) by subjects. The laboratory premises should be large enough to comfortably accommodate research staff and test volunteers. Restricted/insufficient spaces lead to investigator, staff and volunteers stress, which might in turn compromise the quality of the procedures. It is beneficial that the premises are used only for the purpose of clinical testing and are clearly separated from rooms with other functions. A typical testing laboratory should consists of the following rooms: i) reception, ii) waiting room for volunteers, iii) investigator office, iv) staff office(s), v) testing laboratory room(s), vi) test products storage room, and vii) rest room(s).
<b>Volunteers database and recruitment procedures</b>	In order to speed up the enrolment process, information on candidate subjects should be organized in a centralized database. Recruitment procedures (e.g. advertisements, telephone, third parties/intermediaries) should be clear and well documented.
<b>Waiting area for volunteers</b>	A waiting room is required for subjects acclimation before any measure take place. In some cases, volunteers have to wait for long time periods (e.g. during short-term test). The conditions of the waiting area have then an impact on the overall volunteer willingness to participate in the study. Access to and departure from this area should be controlled as well as the ambient conditions.
<b>Testing laboratory rooms</b>	The testing laboratory 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. The laboratory rooms should be versatile, so that different kinds of studies can be performed in the same room. The rooms should be dedicated as testing laboratory rooms and should not serve secondary purposes. The ambient conditions of the testing laboratory rooms should be controlled and kept within a certain range (typically: 22±2°C temperature and 50±10% relative humidity).
<b>Staff office rooms</b>	Staff office rooms are dedicated to staff and should be not accessible to volunteers since confidential information is handled here. Staff office rooms serve as office for the study management.
<b>Samples storage room(s)</b>	The test product storage room contains test products. Access as well as input and output of test products must be restricted. Ambient conditions should be controlled in order to avoid samples deterioration.
<b>Quality Management System</b>	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. SOPs from the following main areas are typically needed: i) quality management procedures, ii) generation and maintenance of study documentation, iii) responsibilities of different study staff, iv) management of measurement devices, v) handling of confidential data, vi) staff qualification and training, vii) data management and statistical procedures, viii) change control procedures and emergencies management, and ix) handling of study and reference products. SOPs on instruments use, maintenance and logs should be implemented.

**Table 4:** Testing laboratory requirements.

## Regulatory requirements

In general, the regulatory requirements for cosmetics are different in different countries (Table 5). Even the classification rules defining what a cosmetic is differ between countries. A product classified as a cosmetic in one country may be classified as a drug in another country. This has implications for the testing of cosmetics products. Depending on which country the testing is performed and/or the test product is marketed, different specific testing may be required, desirable and/or allowed.

Current regulatory frameworks have developed over a considerable period of time, and reflect cultural differences between markets as well as legislative traditions. Countries with OTC/quasi-drug categories, for example, are reluctant to move towards a wider definition of cosmetics. Cosmetic regulations can be classified in two large groups, as follows: i) regulations with broad definitions of cosmetics, and ii) regulations with narrow definitions of cosmetics. The first regulatory frameworks 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. A number of regions working towards harmonisation of cosmetics legislation have used the EU Directive as a model, in particular: Mercosur, Comunidad Andina (Andean Pact), and ASEAN. Other countries have reproduced some features set out in the EU Cosmetics Directive in their national legislation, mainly the definition of the cosmetic product and/or the lists of regulated ingredients. The two most important markets in which this has taken place are Japan and Canada. In Europe, scientific data substantiating the claims made on the packaging must be available in the product information file. In Japan, data substantiating efficacy claims is required only for quasi-drugs (e.g. their specific active ingredients) and not for cosmetics. Only specific, authorized claim wording may be used. The second regulatory framework imposes 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 classified as a cosmetic within the first regulation framework may be classified as an over-the counter (OTC) drug according the second regulation framework. This framework model roughly describes the cosmetics regulatory system in the United States of America (USA). In the United States, the situation of cosmetic claim substantiation is quite complicated; however, also here 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 Federal Trade Commission (FTC). If the FTC determines that the advertisements are not supported by competent and reliable scientific data, constitute unfair or deceptive acts or practices, or are simply false, the FTC can separately pursue enforcement action against the manufacturer. FTC enforcement action can result in consent decrees that require strict controls and oversight for promotional messaging and, in some instance, a money judgment in the form of redress or disgorgement.

The cosmetics regulatory framework 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 differentiates between “non-special-purpose cosmetics” and “special-purpose cosmetics”. In the last two decades, China’s cosmetics industry has undergone tremendous change. To make the cosmetic regulatory framework consistent with the current industry situation, in the second half of 2013, China Food and Drug Administration (CFDA) launched a far reaching campaign set on thoroughly amending the overarching cosmetics regulation. The amendment is not simply a refinement of its predecessor but represents a complete overhaul, which will see the definition, classification and registration requirements of cosmetics completely changed. Once effective, the new regulation will necessitate that both domestic and overseas cosmetic companies adopt new compliance strategies. Cosmetic companies cannot claim functions or features that the products do not possess. Medicinal or therapeutic function claims and any misleading wordings are prohibited. Local expertise will be necessary for balancing compliance with Chinese claim requirements and the addressing the interests of manufacturers who rely on the added value derived from

inclusion of certain claims in their marketing strategies. Testing of imported cosmetic products, domestic special use cosmetics, and new cosmetic ingredients need to be carried out by CFDA-approved laboratories. Laboratories are divided in two categories due to different testing items, as follows: i) hygiene safety testing institutions performing all microbiological, hygienic chemical and toxicological tests and ii) human safety testing institutions performing human safety tests (including human skin patch test and human trials) and cosmetic sunscreens efficacy test.

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

Some products meet the definitions of both cosmetics and drugs. This may happen when a product has two intended uses. For example, according to FDA a shampoo is a cosmetic because its intended use is to cleanse the hair. An antidandruff treatment is a drug because its intended use is to treat dandruff. Consequently, an antidandruff shampoo is both a cosmetic and a drug. Among other cosmetic/drug combinations are toothpastes that contain fluoride, deodorants that are also antiperspirants, and moisturizers and makeup products marketed with sun-protection claims. Such products must comply with the requirements for both cosmetics and drugs. Products that push the boundaries of cosmetic towards medicinal definitions have been referred to as “cosmeceutical” by Kligman [43] and refer to products that are a drug, a cosmetic or a combination of both, and are generally sold under the remit of cosmetics [44]. Many factors will influence the view of regulatory bodies on this topic in creating the impression of a medicinal claim (e.g. presentation of the product, the target consumer, and the mechanism of action). In general, if these products are classified as cosmetic products, they cannot claim that they can be used in the treatment of a disease, but they can claim beneficial effects such as “soothes the skin”, “keeps the skin in good condition” or “helps prevent harmful damage caused by the environment”. Cosmeceuticals are not regulated as such in the European Union, United States or Japan. In the EU, most are considered cosmetics; in the United States, most are seen as drugs that probably have not been approved by the U.S. Food and Drug Administration (FDA). In Japan, cosmeceuticals are regulated as quasi-drugs [45].

Cosmeceuticals are not regulated as such in the European Union, United States or Japan. In the EU, most are considered cosmetics; in the United States, most are seen as drugs that probably have not been approved by the U.S. Food and Drug Administration (FDA). In Japan, cosmeceuticals are regulated as quasi-drugs [45]. Cosmeceuticals are not regulated as such in the European Union, United States or Japan. In the EU, most are considered cosmetics; in the United States, most are seen as drugs that probably have not been approved by the U.S. Food and Drug Administration (FDA).

Country	Law	Remark	Status
Europe	Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products	Main regulatory framework	In force

	Commission regulation (EU) No 655/2013 of 10 July 2013 laying down common criteria for the justification of claims used in relation to cosmetic products	Defines common criteria for claims justification	In force
<b>USA</b>	United state code title 21, chapter IX, subchapter VI. Federal Food, Drug, and Cosmetic Act	Main regulatory framework	In force
	Subchapter VI, section 361 - Adulterated cosmetics		In force
	Subchapter VI, section 361 - Adulterated cosmetics		In force
	Subchapter VI, section 362 - Misbranded cosmetics		In force
	Subchapter VI, section 363 - Regulations making exemptions		In force
	Sec. 364 - Repealed. Pub. L. 86-618, title I, §103(a)(3), July 12, 1960, 74 Stat. 398		In force
	United state code title 21, chapter I, subchapter G, part 700, subpart B Fair Packaging and Labeling Act (FPLA)	Requirements for Specific Cosmetic Products Consumers protection from health hazard and deceptive practice and consumers help to make informed decisions	
<b>China</b>	Regulations concerning the Hygiene Supervision over Cosmetics 1989 “化	Overarching: main regulatory framework promulgated by MOH	In force, but being revised
	Hygienic Standard for Cosmetics 2007 “化	Hygiene standard: GMP-like guidance for products	In force, but being revised
	Hygienic Standard for Production Enterprises of Cosmetics 2007 “化	Hygiene standard: GMP-like guidance for Manufacturers	In force
	Requirements for Application and Acceptance of Administrative Licensing for Cosmetics 2009 “化	Licensing: specifies registration dossiers for imported cosmetics, domestic special use cosmetics and new cosmetic ingredient.	In force
	Requirements for Filing of Domestic Non-special Use Cosmetic Products “国	Licensing: guidance for filing of domestic non-special use cosmetics	In force
	Guidance on Application and Review of New Cosmetic Ingredient 2011”化	Licensing: registration of new cosmetic ingredients	In force
	Guidance on Application and Review of Children's Cosmetics 2012 “儿童化	Licensing: registration of cosmetic products for children	In force
	Requirements for Cosmetic Administrative Licensing Testing 2010 “化	Testing: testing requirements for imported cosmetics, domestic special use cosmetics and new cosmetic ingredient	In force
	Provisions on the Administration of Cosmetics Labeling and instruction “化	Labeling: standardizes required information in label and instruction of cosmetic products	Not published yet
	GB 5296.3-2008 General Labeling for Cosmetics “GB 5296.3-2008 消	Labeling: details required information on cosmetic labels	In force
	Guide to the Naming of Cosmetics 2010”化	Naming: details requirements for naming	In force
	Inventory of Existing Cosmetic Ingredients in China 2014”已使用化	Ingredients: list of ingredients that does need not need registration	Draft
	Administrative Measures on Inspection and Quarantine of Import and Export Cosmetics “	Import & export: regulates the import and export of cosmetics	In force
<b>Japan</b>	Pharmaceutical Affairs Law (PAL) Ministry of Health and Welfare Notification N.331 of 2000	Main regulatory framework	In force
	Director General of the Pharmaceutical Affairs Bureau Notification N.1339	Standard for cosmetic products	In force
	Ministry of Health and Welfare, October 9, 1980 for the Standards for Fair Advertising Practices of Drugs, Quasi-drugs, Cosmetics and Medical Devices	Adversting practice	In force

**Table 5:** Main regulatory frameworks.



## Conclusion

At present, the harmonization of the main regulatory frameworks is incomplete and barriers to innovation are still in place. The misalignment of the international rules and regulation governing the cosmetic industry is challenging for the industry and represent a limit for testing and innovation. For manufacturers in particular, differences in the regulatory frameworks impose significant financial and other costs and prevent the free movement of cosmetic products. Even if it is unrealistic to assume that complete harmonisation of legislation is possible (certainly not in the short or medium term) there is a considerable potential for further alignment. A careful analysis of the main regulatory frameworks reveals that the legislators approach is fundamentally the same. In fact, the following points can be found recurring in each cosmetic legislation: i) the responsibility for the cosmetic product remains at the manufacturer, ii) a limited number of specific ingredients requires prior review and approval, iii) a limited number of ingredients is banned or restricted, iv) independent expert bodies are engaged to review test data on ingredients, v) promotional claims and or advertisement campaigns need to be scientifically substantiated.

Despite the highly regulated cosmetic framework, no clear and specific guidelines, scientific positions or statements exist for efficacy testing on humans. The lack of communication between regulators and scientists produce a “grey area” where there is enough space for interpretation of the underlying regulatory framework. These areas have an impact on the quality of the studies carried out on cosmetic products, on consumer protection from misleading claims, and on cosmetic products credibility in the marketplace. It is auspicious then, in the next years, an international solution or agreement to standardize the technical requirements for cosmetic efficacy testing study on humans.

## References

1. Regulation (EC) No 1223/2009 of the European Parliament and of the council of 30 November 2009 on cosmetic products (recast) (Text with EEA relevance).
2. Commission Regulation (EU) No 655/2013 of 10 July 2013 laying down common criteria for the justification of claims used in relation to cosmetic products.
3. ISO 5496:2006. Sensory analysis -- Methodology -- Initiation and training of assessors in the detection and recognition of odours.
4. ISO/TR 26369:2009. Cosmetics -- Sun protection test methods -- Review and evaluation of methods to assess the photoprotection of sun protection products.
5. ISO 24444:2010. Cosmetics -- Sun protection test methods -- *In vivo* determination of the sun protection factor (SPF).
6. ISO 24442:2011. Cosmetics -- Sun protection test methods -- *In vivo* determination of sunscreen UVA protection.
7. Food and Drug Administration (FDA). Labeling and Effectiveness Testing: Sunscreen Drug Products for Over-the-Counter Human Use.
8. Piérard GE (1998) EEMCO guidance for the assessment of skin colour. J Eur Acad Dermatol Venereol 10: 1-11.
9. Lévêque JL (1999) 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 12: 103-114.
10. Piérard GE (1999) 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 12: 352-362.
11. Piérard GE, Piérard-Franchimont C, Marks R, Paye M, Rogiers V (2000) EEMCO guidance for the *in vivo* assessment of skin greasiness. The EEMCO Group. Skin Pharmacol Appl Skin Physiol 13: 372-389.
12. Rodrigues L, EEMCO (2001) 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 14: 52-67.
13. Rogiers V, EEMCO Group (2001) EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences. Skin Pharmacol Appl Skin Physiol 14: 117-128.
14. Berardesca E, Lévêque JL, Masson P, European Group for Efficacy Measurements on Cosmetics and Other Topical Products (EEMCO Group) (2002) EEMCO guidance for the measurement of skin microcirculation. Skin Pharmacol Appl Skin Physiol 15: 442-456.
15. Parra JL, Paye M, EEMCO Group (2003) EEMCO guidance for the *in vivo* assessment of skin surface P<sup>H</sup>. Skin Pharmacol Appl Skin Physiol 16: 188-202.
16. Piérard GE, Elsner P, Marks R, Masson P, Paye M, EEMCO Group (2003) EEMCO guidance for the efficacy assessment of antiperspirants and deodorants. Skin Pharmacol Appl Skin Physiol 16: 324-342.
17. Piérard GE, Piérard-Franchimont C, Marks R, Elsner P, EEMCO group (European Expert Group on Efficacy Measurement of Cosmetics and other Topical Products) (2004) EEMCO guidance for the assessment of hair shedding and alopecia. Skin Pharmacol Physiol 17: 98-110.
18. O'goshi K, Serup J (2005) Inter-instrumental variation of skin capacitance measured with the Corneometer. Skin Res Technol 11: 107-109.
19. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects.
20. International Ethical Guidelines for Biomedical Research Involving Human Subjects (2002) Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO), Geneva.
21. Code of Federal Regulations, Parts 50, 56, 312, and 812.
22. Guideline for good clinical practice E6 (R1) (1996) ICH harmonised tripartite guideline.
23. Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice E6 (R2) (2015).
24. Meilgaard M, Civille GV, Carr CM (1999) Sensory Evaluation Techniques. CRC Press, Boca Raton, FL.
25. Dus CA, Rudolf M, Civille GV (1997) Sensory testing methods used for claim support. In Aust LB (edn.) Cosmetic Claim Substantiation. Marcel Dekker, New York.
26. Griffiths CE, Wang TS, Hamilton TA, Voorhees JJ, Ellis CN (1992) A photometric scale for the assessment of cutaneous photodamage. Arch Dermatol 128: 347-351.
27. Humbert P (2006) Classification of facial wrinkling. In: Pierre Agache, Philippe Humbert (edn.) Measuring the Skin. Springer, Heidelberg.
28. Lemperle G, Holmes RE, Cohen SR, Lemperle SM (2001) A classification of facial wrinkles. Plast Reconstr Surg 108: 1735-1750.
29. Abella ML (2006) Evaluation of anti-wrinkle efficacy of adenosine-containing products using the FOITS technique. Int J Cosmetic Sci 28: 447-451.
30. Akazaki S, Imokawa G (2001) Mechanical methods for evaluating skin surface architecture in relation to wrinkling. J Dermatol Sci 27: 5-10.
31. Baret M, Bensimon N, Coronel S (2006) Characterization and quantification of the skin radiance through new digital image analysis. Skin Res Technol 12: 254-260.
32. Caspers PJ, Lucassen GW, Puppels GJ (2003) Combined *in vivo* confocal Raman spectroscopy and confocal microscopy of human skin. Biophys J 85: 572-580.
33. Hawkins SS, Perrett DI, Burt DM, Rowland DA, Murahata RI (1999) Prototypes of facial attributes developed through image averaging techniques. Int J Cosmet Sci 21: 159-165.
34. Nobile V, Buonocore D, Michelotti A, Marzatico F (2014) Anti-aging and filling efficacy of six types hyaluronic acid based dermo-cosmetic

- treatment: double blind, randomized clinical trial of efficacy and safety. *J Cosmet Dermatol* 13: 277-287.
35. Rosen BG, Blunt L, Thomas TR (2005) On in-vivo skin topography metrology and replication techniques. *J Phys Conf Ser* 13: 325-329.
36. Xiao P, Wong W, Cottenden AM, Imhof RE (2007) Opto-thermal radiometry and condenser-chamber method for stratum corneum water concentration measurements. *Appl Phys B* 86: 715-719.
37. Cummings S (2007) In *Designing Clinical Research*. Lippincott, William and Wilkins, Philadelphia, USA.
38. Campbell MJ, Julious SA, Altman DG (1995) Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ* 311: 1145-1148.
39. Hall B, Tozer S, Safford B, Coroama M, Steiling W et al., (2007) European consumer exposure to cosmetic products, a framework for conducting population exposure assessments. *Food and Chemical Toxicology* 45: 2097-2108.
40. Hall B., Steiling W., Safford B., Coroama M., Tozer S., et al., (2007) European consumer exposure to cosmetic products, a framework for conducting population exposure assessments, Part 2. *Food Chem Toxicol* 49: 407-421.
41. Bremmer HJ, Prud'Homme de Lodder LCH, van Engelen JGM (2006) Cosmetics Fact Sheet to assess the risks for the consumer, ConsExpo4, RIVM Report 320104.
42. Oliver Wunderlich (2011) Testing laboratory. In *Practical Aspects of Cosmetic Testing: How to Set up a Scientific Study in Skin Physiology*. Joachim Fluhr, Springer-Verlag, Berlin.
43. Kligman AM (2000) Cosmeceuticals: do we need a new category? In: Elsner, P, Maibach HI (edn.) *Cosmeceuticals, Drugs vs Cosmetics*. Marcel Dekker, New York.
44. Milstein SR, Halper AR, Katz, LM (2006) Definition of cosmetics. In: Paye, M, Barel AO, Maibach HI (edn.) *Handbook of Cosmetic Science and Technology*. Taylor and Francis Group, Boca Raton, FL.
45. Priyank S (2011) Cosmeceuticals: regulatory scenario in US, Europe and India. *IJPT* 3: 1512-1535.