A Review of Skin Irritation Protocols for the Evaluation of Skin Irritation Potential

Ward L. Billhimer and James P. Bowman

Hill Top Research Corporation

Introduction

Skin irritation is one of the most common adverse effects in humans. It is an immediate, non-immunological inflammatory response to a physical, chemical, or biological insult. Inflammation produced in the skin can be objectively observed (erythema, edema or heat) or subjectively described (pain, stinging, itching, or burning). The irritant response tends to be universal, unlike most other inflammatory responses which are individual and specific. Irritant contact dermatitis can occur in all individuals provided the irritants are in contact with the skin for a sufficient length of time and in sufficiently high concentration. Testing for potential adverse skin effects is a key component of both the overall safety assessment and the evaluation of potential improvements in mildness of many consumer products. This presentation will provide a review of some key protocols used to assess the irritation potential of a broad base of consumer products and their ingredients as well as provide support for product claims.

Principles of Skin Irritation

An irritant is any agent that, when applied to the skin, has the capacity to produce cell damage with a local inflammatory reaction but without involvement of immunological processes. Virtually any substance, even water, can be an irritant under certain conditions. A multitude of factors are involved in the development and persistence of skin irritation. Strong irritants produce inflammation or cellular damage after a single exposure - primary irritation. However, most irritants produce inflammation after repeated or prolonged exposures - cumulative irritation. While the different types of irritants exert their effects through different mechanisms, their common point of attack is on the stratum corneum.

Strategies for Evaluation of Skin Irritation Potential

Historically, the prophetic patch test has been the method of choice for assessing the irritation potential of consumer products and their ingredients. It originated as a diagnostic tool for identifying allergens and subsequently evolved into a predictive tool for premarket screening for allergens and ranking irritants. Two basic approaches are typically used – single application and repetitive application of the test materials. Due to the diversity of factors influencing the development and intensity of skin irritation, a large number of testing protocols have been developed to evaluate product safety or to substantiate label or advertising claims. These patch tests simulate exaggerated exposure conditions and are used to define the safety of the materials under extreme exposures or "worse case" scenarios.

Single application protocol: Numerous versions of the original 24 hour patch test have been designed to evaluate the primary irritation potential of test materials following a single patch application. It has even been refined and standardized as a 4 hour patch test capable of classifying the irritation hazard of substances with significant irritation potential while avoiding unacceptable strong reactions.

Cumulative irritation patch tests: Kligman and Wooding developed a test designed to evaluate the irritation potential of products that have low irritation potential or tend to be used repetitively. It requires the application of occlusive patches (exaggerated exposure) of the test material to the same skin site for 21 consecutive days. Berger and Bowman showed that cumulative irritation scores for a 14-day test have a positive correlation to the 21-day test and that 95% of the data is effectively generated in the first 14 days.

Traditional patch testing approaches are not always good predictors of the actual consumer experience because they can not simulate or duplicate the almost infinite combination of factors that lead to irritation under use. For this reason, controlled usage tests and consumer home use testing have become a part of the product assessment process.

Controlled usage tests: Exposures in these tests can mimic the actual use or they can be exaggerated to assess the potential impact of misuse/abuse. These tests typically involve rigorous controls such as supervised application and use thus reducing variability and assuring compliance. They can also differentiate between products that are inherently very mild or aid in evaluating perceived skin effects (e.g. pain, burning, itching or stinging) that are not always evident from the results of patch testing. Several test methods will be discussed including immersion tests, the Forearm Controlled Application Test (FCAT), the Leg Controlled Application Test (LCAT), the Antecubital Fossae Test, and the Behind the Knee (BTK) Test.

Consumer home use tests: Home usage tests reveal the irritation potential of the test materials during actual consumer use. It is important in these designs to control as many factors as possible so that any irritation seen, or not seen, can be associated with the test material used. These tests are less robust than the controlled use tests and generally require larger base sizes to compensate for the increased variability.

Study Design and Statistical Considerations

Statistical rigor begins at the design stage of any study. The study design process consists of convening a study design team, defining the study objective, determining the data to collect, choosing the sample size and the appropriate statistical analysis. It is important to enlist all factions that have a vested interest in the research as well as others that may be able to assist. This study design team needs to determine the study objective. Once the objective has been defined, study parameters can be developed. Sample sizes for irritation studies can be derived using statistical calculations as well as from standard requirements from regulatory agencies. Basic test designs for irritation studies are similar, clear objectives are needed, the objective will drive the design and the design will drive the analysis. Parametric as well as non-parametric approaches are used to appropriately evaluate data. Investigator judgment may be just as important as statistical analysis and it is critical to understand the differences between statistical significance and clinical relevance. Inter-individual variability tends to dominate other sources of variability such as seasonality and ethnicity. Consequently, when designing irritation studies to discriminate between ingredients/products, inclusion of positive and negative controls can aid in calibrating the subject panel.

New Endpoint Measures

Visual scoring has been the cornerstone of skin irritation testing. Trained skin graders can accurately and reproducibly score test sites for erythema and dryness. Physiologic changes in blood flow, moisture content, pH, etc. occur early in the irritation process before any visible reactions. These early changes may be critical to distinguishing subtle effects and differentiating between products. Techniques such as TEWL, polarized light, electrical impedance, 2-D & 3-D ultrasound, laser Doppler velocimetry, immunoassay for inflammatory biomarkers, and quantitative measures of transient sensory responses have brought new technical insight into the irritation process and detection of product differences. These non-invasive techniques can improve the quality and relevance of data since they provide objective and quantitative measures.

Summary

All substances possess some potential irritancy which must be fully evaluated to ensure that marketed products are not unacceptably irritating. With the wide range of ingredients and formulations to be evaluated, adequate thought and preparation must go into the evaluation strategy. In patch testing, prolonged contact under occlusion is expected to reveal whether a substance has demonstrable irritancy potential. However, its potential under use conditions may not be evident. By spreading the evaluation over a spectrum of exposure conditions, a truer evaluation of irritation potential is obtained than by merely taking a single reading at an arbitrary point. Rigorous product assessment is critical in the product development process to assure safety and consumer acceptance.