Variability in the Rabbit Skin Irritation Assay

J.P. Rooney[1], N. Choksi[1], R. Rai[1], D.G. Allen[1], N. Kleinstreuer[2], W. Casey[2]

[1]ILS, RTP, NC, United States; [2]NIH/NIEHS/DNTP/NICEATM, RTP, NC, United States

For decades, the gold standard for skin irritation testing has been the in vivo rabbit skin test. As a result, it has been the benchmark against which new approach methodologies have been compared. In vitro methods following the Organisation for Economic Co-operation and Development's Test Guideline 439 are accepted as partial replacements for the in vivo test because they are capable of separating irritant from non-irritant chemicals. However, because none are capable of classifying moderate and mild irritants, there are currently no in vitro methods that are considered full replacements for the in vivo assay. A limiting factor in identifying a full replacement for the in vivo method could be the variability inherent to the subjective scoring of erythema and edema responses in the rabbit test. This is particularly relevant for mild and moderate irritants, which alternative methods have had difficulty differentiating with high concordance. To better characterize the reproducibility of the in vivo assay, in this study we assessed variability in results from chemicals tested multiple times. We compiled and curated a dataset of 3056 test records, representing 727 unique chemicals for which GHS classifications were available. Each chemical was tested at least twice in the in vivo assay. Where possible, primary dermal irritation indexes (PDII) were estimated from the available data and used to classify chemicals according to the U.S. Environmental Protection Agency skin irritation classification criteria. Conditional probabilities were used to evaluate the reproducibility of the in vivo method in identification of severe, moderate, and mild irritants, and non-irritants, according to both GHS and EPA categorization methods. Chemicals classified as moderate irritants at least once were classified as mild irritants or non-irritants over 40% of the time when tested repeatedly. The level of variability found was greatest between mild and moderate irritants. This analysis indicates that the level of variability present in the rabbit skin irritation test should be taken into consideration when evaluating the performance of nonanimal alternative methods. This project was funded in whole or in part with federal funds from the NIEHS, NIH under Contract No. HHSN273201500010C.

Keywords: alternatives to animal testing, cutaneous or skin toxicity, regulatory/policy