Ocular irritation testing has historically been performed using the Draize rabbit eye test (Draize et al.). However, the Cosmetics Testing Ban enacted by the EU and REACH has prompted the development of in vitro ocular irritation alternatives. Recently a number of in vitro testing strategies have been validated as replacements for the Draize rabbit eye test. These include the Bovine Cornea Opacity and Permeability (BCOP) test method (OECD TG 437), Isolated Chicken Eye (ICE) test method (OECD TG 438), the Short Time Exposure (STE) test method (OECD TG 439) and the reconstructed human Cornea-Epithelium (RHE) test method (OECD TG 402), among others. However, there are many factors that may influence the decision whether or not to use a specific method (e.g. in vitro corneal opacity and permeability test (Stephenson et al.))

The objective of this study was to evaluate the reliability of the Short Time Exposure (STE) test method (OECD TG 402) to properly identify the ocular irritation categorization of a set of 80 blinded test chemicals. In addition, we set out to define the applicability domain, strengths and limitations of the STE test system.

All of the in vitro No Cat liquids resulted in a viability >70% at both concentrations and were therefore correctly identified.

Several of the in vivo Cat 2 liquids resulted in a viability >70% at a 0.05% concentration and a viability >70% at 5% concentration and were identified as ‘No prediction can be made’.

The in vivo Cat 2 liquids that were under or over-predicted were not specifically related to the driver of in vivo classification.

When used to identify chemicals inducing serious eye damage (UN GHS Cat 1), the STE test method had an overall accuracy of 61.3% (49/80), a false positive rate of 4.8% (2/42) and a false negative rate of 76.3% (29/38).

When used to identify chemicals not requiring classification (UN GHS No Cat), the STE test method had a sensitivity of 72.5% (50/69), a specificity of 97% (1/10) and a false negative rate of 33.8% (22/65).

All of the in vivo No Cat liquids resulted in a viability >70% at both concentrations and were therefore correctly identified.

Several of the in vivo Cat 2 liquids resulted in a viability >70% at a 0.05% concentration and a viability >70% at 5% concentration and were identified as ‘No prediction can be made’.

The in vivo Cat 2 liquids that were under or over-predicted were not specifically related to the driver of in vivo classification.

When used to identify chemicals inducing serious eye damage (UN GHS Cat 1), the STE test method had an overall accuracy of 61.3% (49/80), a false positive rate of 4.8% (2/42) and a false negative rate of 76.3% (29/38).

When used to identify chemicals not requiring classification (UN GHS No Cat), the STE test method had a sensitivity of 72.5% (50/69), a specificity of 97% (1/10) and a false negative rate of 33.8% (22/65).

The authors would like to acknowledge Cosmetics Europe for their contribution in the chemical selection.

**REFERENCES**
