

Validating Alerts in Derek for Windows

Williams, Richard V.; Marchant, Carol A.; Covey-Crump, Elizabeth M.; Langton, Kate; Naven, Russell T.
Lhasa Limited, 22-23 Blenheim Terrace, Woodhouse Lane, Leeds LS2 9HD, UK

Introduction

Derek for Windows is a knowledge-based expert system designed to predict the toxicity of a chemical from its structure. The knowledge base is composed of structural alerts (or "alerts"), example compounds and rules each of which contributes to predictions made by the system (Figure 1).

Recently, a new feature has been introduced into Derek for Windows which allows validation data for an alert to be displayed when it is activated by a query compound. Such data assist the user in understanding the reliability of an alert and contribute to compliance with the OECD Principles for (Q)SAR validation (1).

The first step in the development of this feature has been the addition of validation data for mutagenicity using four data sets, including a collection of proprietary Ames test data arising from a data sharing initiative.

This work describes preliminary results for the extension of the feature to the chromosome damage endpoint (used in Derek for Windows to describe structural or numerical chromosomal aberrations) for a data set of *in vitro* chromosome aberration test data.

Method

Mutagenicity datasets: The CGX data set and a data set of marketed pharmaceuticals were compiled in Accord for Access database format (2,3). National Toxicology Program (NTP) and proprietary data collated as part of a data sharing initiative were already available in the Vitic database format (4).

Chromosome damage dataset: Data for 721 chemicals in the Sofuni data book were compiled in ISIS/Base format (5).

Structures from each database were exported as SDF files and processed through Derek for Windows. Predictions of activity were compared with experimental data and used to determine the positive predictivity for each alert, for each data set. The positive predictivity describes the proportion of chemicals predicted to be active which are found to be so experimentally.

The results of the mutagenicity analysis were added to the new validation comments display feature introduced in Derek for Windows version 10. The results of the chromosome damage analysis are intended for future inclusion in Derek for Windows version 11.

Mutagenicity Results

The Derek for Windows knowledge base currently contains eighty-five alerts for the prediction of Ames test mutagenicity. In total, approximately 85% of the alerts were activated by at least one chemical in the data sets examined. This may reflect both the size and the composition of the data sets, which contain both drug-like and non-drug-like chemicals. Figure 3 summarises the positive predictivities for the alerts activated.

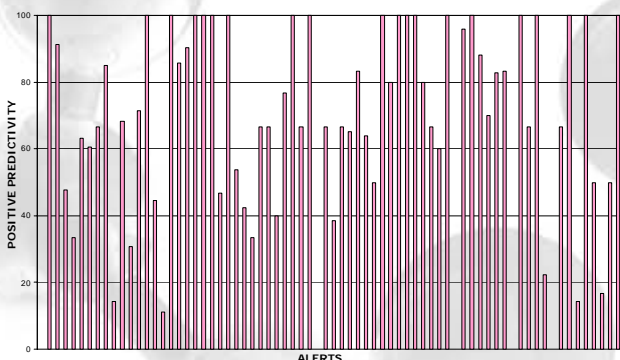


Figure 3 – Positive predictivity for each Ames test mutagenicity alert

Twenty alerts were associated with predictivities of 100%. Alerts describing the mutagenicity of aromatic amines also all had predictivities of 70% or higher following recent improvements in their coverage. Alerts activated on multiple occasions and associated with low predictivity were highlighted for further review. These include alerts for N-haloamines, peroxides and furans. Low predictivity was also observed for alerts based on chemical classes which typically display mutagenicity in strains of *Salmonella typhimurium* which may not have been included in the standard test battery.

The implementation of the results of this work in Derek for Windows is illustrated in Figure 4.

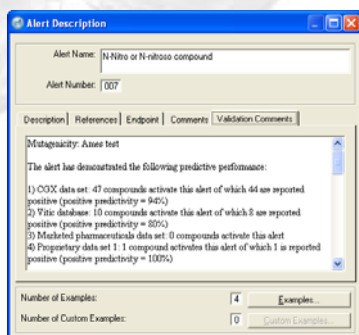


Figure 4 – Example validation comments in Derek for Windows

Conclusions

The results of the work on validation comments for mutagenicity have demonstrated that this approach may be a useful tool in evaluating the significance of toxicity predictions based upon alerts. Preliminary results for the chromosome damage endpoint are also promising, but highlight the value of using, wherever possible, multiple data sets from diverse sources that have not been used previously in alert development. In this respect, future plans include the introduction of results from additional analyses of experimental test data sets held in-house by Derek for Windows member organisations.

References

(1) OECD (2004). *The report from the expert group on (quantitative) structure-activity relationships [(Q)SARs] on the principles for the validation of (Q)SARs*; (2) Kirkland D et al. (2005). *Mutation Research* 584 1-256; (3) Snyder RD and Green JW (2001). *Mutation Research* 488 151-159; Snyder RD et al. (2004) *Environmental and Molecular Mutagenesis* 43 143-158; (4) Judson PN et al. (2005). *Toxicology* 213 117-128; (5) Sofuni T (editor) (1999). *Revised Edition 1998 Data Book of Chromosomal Aberration Test In Vitro*.

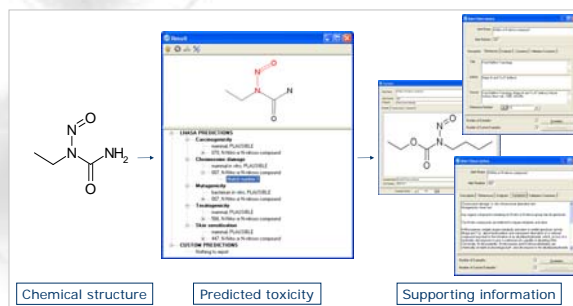


Figure 1 – Derek for Windows results display for N-ethyl-N-nitrosourea and information supporting the predictions of mutagenicity and chromosome damage based on the presence of the same alert

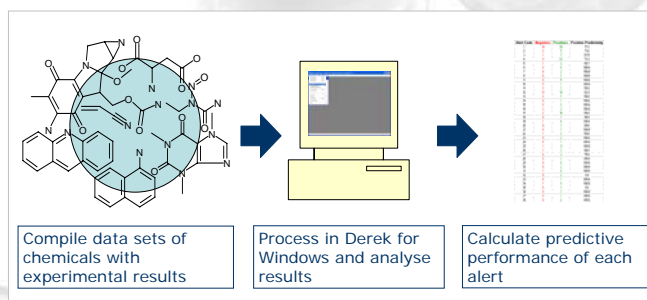


Figure 2 – Methodology for alert validation

Chromosome Damage Results

The Derek for Windows knowledge base currently contains fifty-eight alerts for the prediction of *in vitro* chromosome damage. Figure 5 summarises the coverage of chromosome damage alerts by chemicals in the data set, and shows that approximately 75% of the alerts are activated by at least one chemical. Figure 6 summarises the positive predictivity for the alerts activated.



Figure 5 – Chromosome damage alerts activated by chemicals in the data set

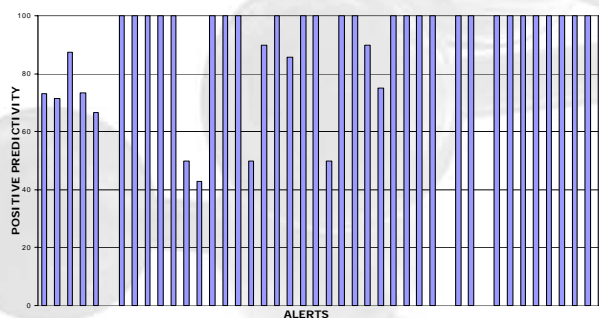


Figure 6 – Positive predictivity for each chromosome damage alert

Twenty-eight alerts were associated with a predictivity of 100%, although some of these are new alerts (right-hand side of Figure 6) which were developed using the data set.

Alerts associated with both Ames test mutagenicity and chromosome damage activity often exhibited comparable performance for both endpoints as would be expected for chemicals operating by a common mechanism of action. Table 1 shows results for selected alerts, well-covered by the data sets, acting through DNA adduct formation.

Alert	Mutagenicity	Chromosome damage
N-Nitro or nitroso compound	91.4%	73.1%
Epoxide	60.5%	71.4%
Alkylating agent	68.4%	73.3%
Aromatic nitro compound	83.3%	90.0%

Table 1 – Positive predictivities of selected alerts for both Ames test mutagenicity and chromosome damage