Advantages in Utilizing an Integrated *In Silico* Solution for ICH M7 Expert Review

SOT March 2016, New Orleans

Chris Barber
Director of Science
chris.barber@lhasalimited.org
A total solution to ICH M7
..developed through collaboration

ICH Harmonised Tripartite Guideline

Assessment and Control of DNA Reactive (Mutagenic)
Impurities in Pharmaceuticals to Limit Potential
Carcinogenic Risk

ICH M7

- Impurity identification
- Database searching
- Expert prediction
- Statistical prediction
- Expert assessment
- Purge mitigation
- Classification
- Reporting
- Control
- Test
Identifying potential mutagenic impurities

ICH HARMONISED TRIPARTITE GUIDELINE

ASSESSMENT AND CONTROL OF DNA REACTIVE (MUTAGENIC) IMPURITIES IN PHARMACEUTICALS TO LIMIT POTENTIAL CARCINOGENIC RISK

M7

Impurity identification

ICH M7

Observed impurities

Expected impurities

Synthesis-derived

Degradation-derived

• Starting material
• Reagent
• Catalyst / ligand
• Intermediates
• By-products

• Decomposition
• Reaction with
  - Packaging
  - Excipients
Identifying impurities from degradation

**Degradation Product**: A molecule resulting from a chemical change in the drug molecule brought about over time and/or by the action of light, temperature, pH, water, or by reaction with an excipient and/or the immediate container/closure system.

ICH M7

…observed during manufacture or stability studies

ICH 3QB

“potential degradation pathways” including from interaction with excipients and/or packaging

‘relevant stress conditions’
- Light
- Heat
- Humidity
- Acid/base hydrolysis
- Oxidation
Identifying impurities from degradation

'relevant stress conditions’
• Light
• Heat
• Humidity
• Acid/base hydrolysis
• Oxidation

ICH M7
ICH 3QB

Expert review

Photochemical (37)
Rearrangement (31)
Elimination (49)
Addition (77)
Hydrolysis (80)
Oxidation (126)

Zeneth includes predictions of reaction of the API under different conditions in the presence of excipients, solvents and degradants
Publications on Zeneth for degradation prediction

• An expert system to predict the forced degradation of organic molecules
  • Parenty.. *Mol. Pharmaceutics*, 2013, **10**, 2962

• Forced degradation studies to assess the stability of drugs and products
  • Singh.. *Trends in Analytical Chem*, 2013, **49**, 71

• Mutagenic Impurities: Precompetitive/Competitive Collaborative and Data Sharing Initiatives
  • Elder.. *Org. Process Res. Dev*, 2015, **19**, 1486

• Tools and workflow for structure elucidation of drug degradation products
  • Foti.. *Trends in Analytical Chem*, 2015, **49**, 89

• Strategies To Address Mutagenic Impurities Derived from Degradation in Drug Substances and Drug Products
  • Kleinman.. *Org Process R & D*, 2015, **19**, 1447
    • Kleinman.. *Mol. Pharmaceutics*, 2014, **11**, 4179
A total solution to ICH M7
...developed through collaboration

ICH HARMONISED TRIPARTITE GUIDELINE

ASSESSMENT AND CONTROL OF DNA REACTIVE (MUTAGENIC) IMPURITIES IN PHARMACEUTICALS TO LIMIT POTENTIAL CARCINOGENIC RISK

ICH M7

Impurity identification

+ synthetic route
+ expert review...

ICH M7

Database searching

Expert in silico prediction

Statistical in silico prediction

Expert assessment

Purge mitigation

Classification

Reporting

Control

Test
Undertaking a database search

Database searching

Public data

Proprietary data

Substructure / similarity search

Further supporting information for expert review

Mutagen Carcinogen ✓ ✓

Class 1

Mutagen Carcinogen ✓ ?

Class 2

Non-Mutagen ✓

Class 5

M7 (Muller) classification
Undertaking a database search
Vitic Nexus – an authoritative source of data

Sources include:
• FDA
• NTP
• ISSSTY
• Kirkland
• Hansen, Bursi
• MPDB
• Literature

In vitro genetox
• 164,001 | 10,246

In vivo genetox
• 10,595 | 2,723

Overall-call genetox
• 17,322 | 8,932

Carcinogenicity
• 16,419 | 3,865

Aromatic amines
• 3,639 | 424

Intermediates
• 16,971 | 1,088

Excipients
• 2,435 | 975

Vitic contains raw, summary, and Lhasa overall call data + literature references

Currently 400K records, 20K compounds across a wide range of tox endpoints
A total solution to ICH M7
developed through collaboration

ICH Harmonised Tripartite Guideline

Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk

ICH M7

Database searching

Expert in silico prediction

Statistical in silico prediction

Expert assessment

Test

Classification

Purge mitigation

Reporting

Control

Impurity identification
ICH M7 allows the use of *in silico* predictions

- “The absence of structural alerts from two complementary (Q)SAR methodologies (expert rule-based and statistical) is sufficient to conclude that the impurity is of no mutagenic concern”

- “If warranted, the outcome of any computer system-based analysis can be reviewed with the use of expert knowledge in order to provide additional supportive evidence on relevance of any positive, negative, conflicting or inconclusive prediction and provide a rationale to support the final conclusion”

In silico predictions should comply with OECD principles

• “To facilitate the consideration of a (Q)SAR model for regulatory purposes, it should be associated with the following information:

  1. a defined endpoint;
  2. an unambiguous algorithm;
  3. a defined domain of applicability;
  4. appropriate measures of goodness-of-fit, robustness and predictivity;
  5. a mechanistic interpretation, if possible.”
*In silico* predictions should also...

- Make predictions and have good accuracy
  - Within your chemical space
- Provide some meaningful measure of confidence
  - Transparent, demonstrably relevant
- Be regularly updated with new knowledge
- Ideally be well understood by regulatory authorities
- Be transparent and highlight areas of uncertainty
  - To support expert review
- Make predictions that you understand and can support or overturn
How independent are statistical and expert systems?

• ICH M7 doesn’t prescribe how to build the 2 systems but the expectation is that they are independent…

Expert Model

Statistical Model

- Expert pre-define fragments used by statistical system to learn weightings
- Fragmentation rules enable statistical system to independently discover both fragments and their significance
- Statistically found fragments imported into expert system
- Proprietary datasets not seen by statistical system

Range of techniques used to discover relevant fragments. Expert only adds them when supported by a clear scientific rationale, mechanism, literature...
• Derek Nexus - an expert in silico prediction system

• Key endpoints of relevance for M7

- Carcinogenicity (ALL)
  - Carcinogenicity
  - Photocarcinogenicity
- Genotoxicity (ALL)
  - Chromosome damage (ALL)
  - Mutagenicity (ALL)
    - Mutagenicity
      - in vitro
      - in vivo
    - Photomutagenicity
  - Non-specific genotoxicity (ALL)
- Irritation (ALL)
Reviewing a prediction for mutagenicity from Derek

Is it alerting?

Y

Active

Accuracy of predictions

<table>
<thead>
<tr>
<th></th>
<th>Certain</th>
<th>Probable</th>
<th>Plausible</th>
<th>Equivocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>016: Quinolino</td>
<td>100</td>
<td>98</td>
<td>80</td>
<td>49</td>
</tr>
</tbody>
</table>

Assessing confidence in predictions made by knowledge-based systems.
Judson… Toxicol. Res., 2013, 2, 70
Reviewing a prediction for mutagenicity from Derek

Is it alerting?

Contains features in known false negatives?

- Yes
  - Inactive with misclassified features

- No
  - Contains unknown features?
    - Yes
      - Inactive with unclassified features
    - No
      - Inactive

It’s difficult, but important, to make negative predictions
Williams… *Reg. Tox. and Pharmacol.* 2016, 76, 79
Establishing best practise in the application of expert review of mutagenicity under ICH M7
Barber… *Reg. Tox. and Pharmacol.* 2015, 73, 367
Reviewing an inactive with unclassified

Inactive with unclassified

- Is a negative prediction

Highlights fragments in contexts not seen in public data
- May be present in confidential data
- Could be innocuous

Expert review recommended

It’s difficult, but important, to make negative predictions
Williams... *Reg. Tox. and Pharmacol.* 2016, 76, 79
Establishing best practise in the application of expert review of mutagenicity under ICH M7
Barber... *Reg. Tox. and Pharmacol.* 2015, 73, 367
Reviewing an inactive with misclassified

Inactive with misclassified

Is a negative prediction

Highlights fragments seen in false positives

Many possible reasons for activity...
- ‘Innocent spectator’
- Inconsistent data
- Missing alert

Expert review recommended

It's difficult, but important, to make negative predictions
Williams... Reg. Tox. and Pharmacol. 2016, 76, 79
Establishing best practise in the application of expert review of mutagenicity under ICH M7
Barber... Reg. Tox. and Pharmacol. 2015, 73, 367
Expert review – reconsider the source data...

<table>
<thead>
<tr>
<th>Report</th>
<th>Structure</th>
<th>Name</th>
<th>CAS Number</th>
<th>Formula</th>
<th>Weight</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image.png" alt="Structure" /></td>
<td>(Z)-2,3-Dimethoxypro...</td>
<td>106686-63-9</td>
<td>C11H14O2</td>
<td>178.23</td>
<td>1.000000</td>
</tr>
</tbody>
</table>

**Results**

<table>
<thead>
<tr>
<th>CAS Number</th>
<th>Result</th>
<th>Result type</th>
<th>Assay</th>
</tr>
</thead>
<tbody>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Negative</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
<tr>
<td>106686-63-9</td>
<td>Positive</td>
<td>Single experim...</td>
<td>Bacterial revers...</td>
</tr>
</tbody>
</table>

**Details**

Features of the query structure (highlighted in red) and the mutagenicity structural alert are misclassified. The query structure does not match any structural alerts in the mutagenicity test.

**Nearest neighbours**

Displaying most similar compounds (from Lhasa).
Reviewing negative predictions from Derek

- Performance: how often does each category occur?

<table>
<thead>
<tr>
<th>Private data 1 (n=325)</th>
<th>Inactive</th>
<th>Inactive with misclassified features</th>
<th>Inactive with unclassified features</th>
<th>Inactive with unclassified &amp; misclassified features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>86%</td>
<td>9%</td>
<td>5%</td>
<td>-</td>
</tr>
<tr>
<td>Private data 2 (n=416)</td>
<td>89%</td>
<td>7%</td>
<td>3%</td>
<td>-</td>
</tr>
<tr>
<td>Private data 3 (n=1669)</td>
<td>90%</td>
<td>7%</td>
<td>3%</td>
<td>-</td>
</tr>
</tbody>
</table>
Reviewing negative predictions from Derek

- Performance: how right are compounds in each category?

<table>
<thead>
<tr>
<th></th>
<th>Private data 1 (n=325)</th>
<th>Private data 2 (n=416)</th>
<th>Private data 3 (n=1669)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inactive</td>
<td>94%</td>
<td>87%</td>
<td>91%</td>
</tr>
<tr>
<td>Inactive with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>misclassified features</td>
<td>86%</td>
<td>86%</td>
<td>93%</td>
</tr>
<tr>
<td>Inactive with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unclassified features</td>
<td>86%*</td>
<td>94%*</td>
<td>95%*</td>
</tr>
<tr>
<td>Inactive with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unclassified &amp;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>misclassified features</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* low n
A total solution to ICH M7
..developed through collaboration

ICH Harmonised Tripartite Guideline

Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk

M7

ICH M7

Database searching

Expert in silico prediction

Statistical prediction

Expert assessment

Purge mitigation

Classification

Reporting

Control

Test

Impurity identification
Sarah Nexus – a statistical model for mutagenicity

- Supplied with model built with Lhasa-curated public data
- Optimised to learn mutagenicity...
  - Fragmentation designed for reactivity-driven endpoints
  - Self-organising Hierarchical Network to maximise information gain
  - Decision-tree to reduce the chance of coincidence
- Explicit applicability domain
- Confidence score is provided for each prediction
- Predictions are transparent and therefore interpretable
A self-organising hypothesis network produces transparent predictions.
A self-organising hypothesis network produces transparent predictions

- Allows the most specific hypothesis to be applied
- Links back to training set

Examples
Larger fragments
Smaller fragments
Root hypothesis

More specific hypothesis
More general hypothesis
Reviewing a prediction from Sarah Nexus

For the 'Mutagenicity in vitro' endpoint the prediction is:

**NEGATIVE**

with 31% confidence

The compound is predicted to be negative with 31% confidence for the 'Mutagenicity in vitro' endpoint in the model 'Sarah Model - 1.1.19'. Supporting hypotheses containing similar examples from the training set have been found.
Possible outcomes from Sarah Nexus

For the 'Mutagenicity in vitro' endpoint the prediction is:

NEGATIVE
with 42% confidence

For the 'Mutagenicity in vitro' endpoint the prediction is:

EQUIVOCAL

For the 'Mutagenicity in vitro' endpoint the prediction is:

POSITIVE
with 100% confidence

OUTSIDE DOMAIN

Displaying 'outside domain features', click above to view the original structure
Building your own statistical model using Sarah Nexus

- Model building allows you to import your own data – appending or replacing the default data set
### Internal Cross-Validation using Sarah Nexus

#### Sarah Model Building - Results

<table>
<thead>
<tr>
<th>Cross-Validation - Fold 1</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Accuracy</th>
<th>B-Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>123</td>
<td>30</td>
<td>147</td>
<td>37</td>
<td>80%</td>
<td>80%</td>
<td>77%</td>
<td>83%</td>
<td>80%</td>
</tr>
<tr>
<td>Cross-Validation - Fold 2</td>
<td>107</td>
<td>31</td>
<td>132</td>
<td>35</td>
<td>80%</td>
<td>79%</td>
<td>75%</td>
<td>83%</td>
<td>78%</td>
</tr>
<tr>
<td>Cross-Validation - Fold 3</td>
<td>128</td>
<td>21</td>
<td>134</td>
<td>37</td>
<td>82%</td>
<td>82%</td>
<td>78%</td>
<td>86%</td>
<td>86%</td>
</tr>
<tr>
<td>Cross-Validation - Fold 4</td>
<td>120</td>
<td>38</td>
<td>140</td>
<td>31</td>
<td>79%</td>
<td>79%</td>
<td>79%</td>
<td>79%</td>
<td>76%</td>
</tr>
<tr>
<td>Cross-Validation - Fold 5</td>
<td>122</td>
<td>22</td>
<td>147</td>
<td>49</td>
<td>79%</td>
<td>79%</td>
<td>71%</td>
<td>87%</td>
<td>85%</td>
</tr>
<tr>
<td>Total Validation</td>
<td>600</td>
<td>142</td>
<td>720</td>
<td>189</td>
<td>80%</td>
<td>80%</td>
<td>76%</td>
<td>84%</td>
<td>81%</td>
</tr>
</tbody>
</table>

---

**Performance**

- **ROC**: 30 (8.4%)
- **Accuracy**: 147 (40.5%)
- **Structures**: 123 (34.3%)
External Validation using Sarah Nexus
A total solution to ICH M7
..developed through collaboration

ICH HARMONISED TRIPARTITE GUIDELINE

ASSESSMENT AND CONTROL OF DNA REACTIVE (MUTAGENIC) IMPURITIES IN PHARMACEUTICALS TO LIMIT POTENTIAL CARCINOGENIC RISK

ICH M7

Database searching
Expert in silico prediction
Statistical in silico prediction
Expert assessment
Purge mitigation
Classification
Reporting
Control
Test
Expert review – the M7 Decision Matrix

O.O.D. = out of domain

<table>
<thead>
<tr>
<th>O.O.D.</th>
<th>-</th>
<th>+</th>
<th>Equiv</th>
<th>-</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>Negative</td>
<td>Positive</td>
<td>Unclassified</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

confident
Performance of Sarah and Derek...

### Frequency of Outcomes (%)

<table>
<thead>
<tr>
<th>O.O.D</th>
<th>2</th>
<th>0</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>10</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Equiv</td>
<td>8</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>-</td>
<td>38</td>
<td>2</td>
<td>12</td>
</tr>
</tbody>
</table>

### Probability of being positive (%)

<table>
<thead>
<tr>
<th>O.O.D</th>
<th>29</th>
<th>-</th>
<th>62</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>30</td>
<td>-</td>
<td>70</td>
</tr>
<tr>
<td>Equiv</td>
<td>23</td>
<td>-</td>
<td>51</td>
</tr>
<tr>
<td>-</td>
<td>8</td>
<td>15</td>
<td>34</td>
</tr>
</tbody>
</table>

- Lhasa data sharing consortium group (n=777; 32% positive)
Expert review of *in silico* predictions

- M7 guidelines
  - “If warranted, the outcome of any computer system-based analysis can be reviewed with the use of expert knowledge..”

- We recommend that you always do some review….
- Be guided by the software and your knowledge…
  - Use confidence measures if they are proven to indicate accuracy
  - Use software that is transparent and highlights areas of concern
  - Use the expert commentary, mechanism and references
  - Look at relevant close examples from the models & databases

- Depth of your analysis and the detail you report will vary
  - …from a cursory analysis to a well-supported argument
Likely to conclude positive
Very strong evidence would be needed to overturn both predictions

Likely to conclude positive
Lack of a second prediction suggests insufficient evidence to draw any other conclusion

Uncertain
Likely to conclude positive without strong evidence to overturn a positive prediction

In silico prediction 1
Positive  Positive  Positive  Negative  Negative

In silico prediction 2
Positive  O.O.D. or equivocal  Negative  O.O.D. or equivocal  Negative

O.O.D. = out of domain

Uncertain
Conservatively could assign as positive. May conclude negative with strong evidence showing feature driving a ‘no prediction’ is present in the same context in known negative examples (without deactivating features)

Likely to conclude negative
Expert review should support this conclusion – e.g. by assessing any concerning features (misclassified, unclassified, potentially reactive...)

Establishing best practise in the application of expert review of mutagenicity under ICH M7 Barber... Reg. Tox. and Pharmacol. 2015, 73, 367
Expert review of *in silico* predictions

- There are publications describing expert analysis for M7
  - *In silico* methods combined with expert knowledge rule out mutagenic potential of pharmaceutical impurities: an industry survey
  - (Q)SAR assessments of potentially mutagenic impurities: A regulatory perspective on the utility of expert knowledge and data submission
    - Powley.. Reg. Tox. and Pharmacol. 2015, 71, 295
  - Establishing best practise in the application of expert review of mutagenicity under ICH M7
    - Barber… Reg. Tox. and Pharmacol. 2015, 73, 367
  - Principles and procedures for implementation of ICH M7 recommended (Q)SAR analyses
    - Amberg… Reg. Tox. and Pharmacol. 2016, asap
- Training course at the GTA meeting (May 4th, Delaware)
A total solution to ICH M7
..developed through collaboration

ICH Harmonised Tripartite Guideline

Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk

ICH M7

Database searching

Expert in silico prediction

Statistical in silico prediction

Expert assessment

Purge mitigation

Classification

Reporting

Control

Test
M7 Classification


• M7 classification helps define how to control impurities…
Batch process against M7 settings

- User can add additional data
- Searches for carc’ and mut’ data from Lhasa and custom database
ICH M7 class generated and report produced

Each impurity is classified according to whether there is Ames or Carcinogenicity information in addition to the Derek Prediction

Users can also input experimental results for mutagenicity or carcinogenicity which updates the ICH M7 Class
A total solution to ICH M7
..developed through collaboration

ICH HARMONISED TRIPARTITE GUIDELINE

ASSESSMENT AND CONTROL OF DNA REACTIVE (MUTAGENIC) IMPURITIES IN PHARMACEUTICALS TO LIMIT POTENTIAL CARCINOGENIC RISK

ICH M7

Database searching

Expert in silico prediction

Statistical in silico prediction

Expert assessment

Purge mitigation

Classification

Reporting

Control

Test
Mirabilis – supporting expert assessment of purge

• Concept is part of the M7 guidelines

- Risk of potential mutagenic impurity
  - Ames or 2 in silico models...
  - Treat as a non-mutagen
  - Control & monitor in final API

- Present a purge argument for absence...
  - Risk mitigated

- If negative, treat as a non-mutagen
- If positive, control & monitor in final API


Introducing Mirabilis

• 11 companies sponsoring development of software and science

• Software will support estimating & expert-review of purge values
  
  • Web-based to support collaborative (internal) use
  
  • Support a consistent industry-standardised approach
  
  • Provide supporting commentary and data for expert review
  
  • Generate reports suitable for regulatory submission
Scope of science to be covered by Mirabilis

- Chromatography
- Scavenger resins
- Physical processes
  - Liquid-liquid extractions
  - Solid-liquid extractions
- Reactivity
- Solubility
  - Precipitation
- Evaporation
- Recrystallisation
## Physicochemical Parameters

<table>
<thead>
<tr>
<th>Reactivity</th>
<th>Purge Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly Reactive</td>
<td>100</td>
</tr>
<tr>
<td>Moderately reactive</td>
<td>10</td>
</tr>
<tr>
<td>Low Reactivity / un-reactive</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solubility</th>
<th>Purge Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freely Soluble</td>
<td>10</td>
</tr>
<tr>
<td>Moderately soluble</td>
<td>3</td>
</tr>
<tr>
<td>Sparingly Soluble</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Volatility</th>
<th>Purge Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boiling point &gt;20°C below that of the reaction/ process solvent</td>
<td>10</td>
</tr>
<tr>
<td>Boiling point +/- 10°C that of the reaction/ process solvent</td>
<td>3</td>
</tr>
<tr>
<td>Boiling point &gt;20°C above that of the reaction/ process solvent</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ionisability</th>
<th>Purge Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ionisation potential of GI significantly different to that of the desired product</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical Processes – chromatography</th>
<th>Purge Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromatography – GI elutes prior to desired product</td>
<td>100</td>
</tr>
<tr>
<td>Chromatography – GI elutes after desired product</td>
<td>10</td>
</tr>
<tr>
<td>Others evaluated on an individual basis</td>
<td></td>
</tr>
</tbody>
</table>

### A theoretical approach to estimating purge factors

**Organic Process Research & Development**

2013, 17, 221

**Risk Assessment of Genotoxic Impurities in New Chemical Entities: Strategies To Demonstrate Control**

Andrew Teasdale, *, † David Elder, ‡ Sou-Jen Chang, § Sophie Wang, ‖ Richard Thompson, ‡ Nancy Benz, ‡ and Ignacio H. Sanchez Flores "

AstraZeneca, Charter Way, Silk Road Business Park, Macclesfield, Cheshire SK10 2NX, United Kingdom

GlaxoSmithKline, Park Road, Ware, Hertfordshire SG12 0DP, United Kingdom

Abbott Laboratories, 200 Abbott Park Road, PA71, Building AP-50, Abbott Park, Illinois 60064-6220, United States

Amgen Inc., 1 Amgen Center Drive, Thousand Oaks, California 91320, United States

---

**Theoretical Approach**

- **Reactivity**
  - Highly Reactive: 100
  - Moderately Reactive: 10
  - Low Reactivity / Un-reactive: 1

- **Solubility**
  - Freely Soluble: 10
  - Moderately Soluble: 3
  - Sparingly Soluble: 1

- **Volatility**
  - Boiling point >20°C below that of the reaction/ process solvent: 10
  - Boiling point +/- 10°C that of the reaction/ process solvent: 3
  - Boiling point >20°C above that of the reaction/ process solvent: 1

- **Ionisability**
  - Ionisation potential of GI significantly different to that of the desired product: 100

- **Physical Processes – Chromatography**
  - Chromatography – GI elutes prior to desired product: 100
  - Chromatography – GI elutes after desired product: 10
  - Others evaluated on an individual basis
Using kinetic studies to support a purge value

- Measure rates of loss of ‘impurity’ whilst changing temp
- Can derive rate constant \((k)\) and activation energy \((E_a)\)
Mirabilis – supporting expert assessment of purge

• 11 companies are currently sponsoring development

  • **Software**
    • To apply methodology agreed by industry and regulators
    • To facilitate efficient internal decision-making and reporting

  • **Science**
    • Support capture of expert-driven knowledge
    • Undertake analytical studies to
      • fill in knowledge gaps
      • test performance

• Product release scheduled for end of 2016
A total solution to ICH M7
..developed through collaboration

ICH Harmonised Tripartite Guideline

Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk

M7

ICH M7

Database searching

Expert in silico prediction

Statistical in silico prediction

Expert assessment

Purge mitigation

Classification

Reporting

Control

Test
Summary

• Lhasa has produced a suite of integrated tools to address M7 with the support of members and regulators
  • Developing scientifically robust, transparent and trusted tools to meet our charitable purpose
    • “advance scientific knowledge and understanding through the use of computer-aided reasoning in chemistry and the life sciences”

• Acknowledgements
  • All the members involved in pre-competitive Lhasa collaborations
  • Dedicated staff at Lhasa

• Come to stand 518 for further details and demonstrations.
Thank you and Questions

Lhasa Limited is a not-for-profit organisation and educational charity that facilitates collaborative data sharing projects in the pharmaceutical, cosmetics and chemistry-related industries.
What type of data was used to build the model?

<table>
<thead>
<tr>
<th>Source</th>
<th>+ve</th>
<th>-ve</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feng</td>
<td>880</td>
<td>961</td>
<td>1841</td>
</tr>
<tr>
<td>NTP</td>
<td>591</td>
<td>1234</td>
<td>1825</td>
</tr>
<tr>
<td>Helma</td>
<td>341</td>
<td>343</td>
<td>684</td>
</tr>
<tr>
<td>ISSSTY</td>
<td>3389</td>
<td>2473</td>
<td>5862</td>
</tr>
<tr>
<td>FDA CFSAN</td>
<td>4300</td>
<td>4115</td>
<td>8415</td>
</tr>
<tr>
<td>Kirkland</td>
<td>358</td>
<td>334</td>
<td>692</td>
</tr>
<tr>
<td>Benchmark</td>
<td>3503</td>
<td>3009</td>
<td>6512</td>
</tr>
<tr>
<td>MPDB</td>
<td>40</td>
<td>491</td>
<td>531</td>
</tr>
<tr>
<td>Bursi</td>
<td>2401</td>
<td>1936</td>
<td>2177</td>
</tr>
<tr>
<td>Vitic Summary Call</td>
<td>2079</td>
<td>2838</td>
<td>4917</td>
</tr>
<tr>
<td>DX Mut Examples</td>
<td>300</td>
<td>10</td>
<td>310</td>
</tr>
<tr>
<td>Acid Halide Data</td>
<td>3</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td>Member Data</td>
<td>9</td>
<td>15</td>
<td>24</td>
</tr>
</tbody>
</table>
Sarah’s confidence score correlates with accuracy

<table>
<thead>
<tr>
<th>Confidence Score</th>
<th>0-20%</th>
<th>20-40%</th>
<th>40-60%</th>
<th>60-80%</th>
<th>80-100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPV</td>
<td>58%</td>
<td>74%</td>
<td>85%</td>
<td>93%</td>
<td>92%</td>
</tr>
<tr>
<td>NPV</td>
<td>62%</td>
<td>80%</td>
<td>95%</td>
<td>96%</td>
<td>97%</td>
</tr>
</tbody>
</table>

- Less confident predictions deserve greater scrutiny
- Default settings optimum for regulatory submissions

ESTABLISHING BEST PRACTICE FOR THE APPLICATION OF A NOVEL STATISTICAL-BASED MODEL TO EVALUATE POTENTIAL MUTAGENIC IMPURITIES UNDER ICH M7

Sarah vs. external dataset of confidential data from Lhasa members
It’s difficult, but important, to make negative predictions. Williams… *Reg. Tox. and Pharmacol*. 2016, 76, 79

<table>
<thead>
<tr>
<th>Class</th>
<th>Definition</th>
<th>Proposed action for control (details in Section 7 and 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Known mutagenic carcinogens</td>
<td>Control at or below compound-specific acceptable limit</td>
</tr>
<tr>
<td>2</td>
<td>Known mutagens with unknown carcinogenic potential (bacterial mutagenicity positive*, no rodent carcinogenicity data)</td>
<td>Control at or below acceptable limits (appropriate TTC)</td>
</tr>
<tr>
<td>3</td>
<td>Alerting structure, unrelated to the structure of the drug substance; no mutagenicity data</td>
<td>Control at or below acceptable limits (appropriate TTC) or conduct bacterial mutagenicity assay; If non-mutagenic = Class 5 If mutagenic = Class 2</td>
</tr>
<tr>
<td>4</td>
<td>Alerting structure, same alert in drug substance or compounds related to the drug substance (e.g., process intermediates) which have been tested and are non-mutagenic</td>
<td>Treat as non-mutagenic impurity</td>
</tr>
<tr>
<td>5</td>
<td>No structural alerts, or alerting structure with sufficient data to demonstrate lack of mutagenicity or carcinogenicity</td>
<td>Treat as non-mutagenic impurity</td>
</tr>
</tbody>
</table>
This document is intended to outline our general product direction and is for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon. The development, release, and timing of any features or functionality described for Lhasa Limited’s products remains at the sole discretion of Lhasa Limited.