bayesian Weight of Evidence for REACH ITS generation

EU FP6 project OSIRIS, Pillar 4
E. Rorije, B. Hakkert, T. Aldenberg,
J. Jaworska
D. Kroese, H. Buist
E. Pauné

RIVM
P&G
TNO
Simpple

Berlin, March 1-2, 2010
Overall GOAL of OSIRIS:

• (generation of) Integrated Testing Strategies (ITS)

- Transparent → document all choices
- Reproducible → algorithm, choices
- Objective → quantitative?
- Flexible → allowing extension of methodology
- Practicle → Implementation in a Webtool
  - Allowing user data entry
  - Allowing user test/model entry
ITS:
Most efficient strategy of fulfilling information requirement

Chemical Safety Assessment
REACH context (Annex VI)
4 steps are described:

1. Gather and share ALL available information
2. Consider information requirements cf tonnage-bands (Annexes VII-X)
3. Identify information gaps
4. Generate new testing data / propose testing strategy

In REACH (in vivo) testing is the last resort (art.13, 25 en Annex XI)
ITS:
Most efficient strategy of fulfilling information requirement

**Step 1:** Gather all available (Testing and Non-Testing) information

**Step 2+3:** If not sufficient for C&L and RA cf REACH

**Step 4:** Generate more information (non-Testing)

**Step 2+3:** If not possible and information not sufficient

**Step 4:** Is Exposure-Based Waiving an option?

**Step 2+3:** Still not sufficient for C&L and RA cf REACH

**Step 4:** Perform / Propose (in vivo) Testing as last resort!!
Step 1: Available sources of information (Annex VII-X and Annex XI)

- Human data
- In vitro
- In vivo guideline test
- Non guideline test
- Grouping & read across
- Exposure (-based waiving)

“Weight of Evidence”
Step 2 and 3: Identify data gaps

Available information + REACH information requirements (for C&L, RC) = sufficient ?

Non-guideline

= sufficient ???
Step 2 and 3: Identify data gaps

- Human data
- In vitro
- In vivo guideline test
- (Q)SAR
- non guideline tests
- Grouping & read across

Test Proposal

Data Gap?

use “Weight of Evidence”

Exposure (-based waiving)

Rorije - OSIRIS Stakeholder Workshop, Berlin

March 1-2, 2010
Step 2 and 3: Identify data gaps

By applying “Weight of Evidence”

- Express all information sources (“Evidence”) in one unit (“Weight”)
  - from tests
  - from models
  - from categories
  - etc.

- Decide if total Weight of Evidence is sufficient

- If not → identify most efficient way of filling information gap
How to apply Weight of Evidence? Compare Apples and Pears

→ OSIRIS: apply Bayesian statistics:
  • calculate a *probability* that a result is true, given a specific test outcome
  • add these probabilities (using Bayesian belief network)
  • compare the combined probability to a threshold
1. define (REACH) information requirement for every combination of purpose and endpoint

2. list all possible data sources

Regulatory framework, e.g. REACH

Toxicological endpoint, e.g. sensitization

Battery of available tests/methods

- In vivo guideline study
- Non guideline study
- In vitro test
- QSAR SAR Category

Test or Method fulfilling all requirements, e.g. LLNA
3. Determine Threshold, from:
   - a Gold Standard,
   - intra-test variability, or
   - expert judgment
But: not all Apples are equal

→ the need for **Quality Factors (QF)**
  
e.g. Klimisch(-like) codes for data quality
4. **Method Performance** compared to REACH Information Requirement (statistics)

5. **Quality Factor** of a *specific* result
   Modifying the (overall) Performance
6. Assign Cost to Optimize a Test proposal

- Regulatory Framework, e.g. REACH
  - Toxicological endpoint, e.g. sensitization
    - Information Requirement
      - Test or Method fulfilling all requirements, e.g. LLNA
        - Sensitivity/specificity towards Gold Standard OR expert estimates
          - QF = 1
            - Threshold Posterior Probability, e.g. P = 0.85
  - Prior Probability
    - Battery of available tests/methods
      - Available Information
        - In vitro test
          - Cost
            - QF = 0.8
              - Posterior Probability, e.g. P = 0.78
          - Cost
            - QF = 0.7
              - Posterior Probability, e.g. P = 0.75
          - Cost
            - QF = 0.6
              - Posterior Probability, e.g. P = 0.72
          - Cost
            - QF = 0.5
              - Posterior Probability, e.g. P = 0.68
      - QSAR SAR Category
        - Cost
          - QF = 0.8
          - Cost
            - QF = 0.7
            - Cost
              - QF = 0.6
              - Cost
                - QF = 0.5

- Information Requirement fulfilled!
Optimization function for Test Proposal

• If the cost function is “correct” (politically) then “Non-testing” is exhausted before (in vivo) Testing

• If Non-Testing is exhausted, and a REACH required test is performed, this effectively ignores all previously gathered WoE information.

• Do we need options to adapt testing protocols?
Case (proof-of-principle) of qWoE ITS: Skin Sensitization

- Excel implementation
- Hugin® implementation
- OSIRIS webtool implementation

- hydroxycitronellal 107-75-5
- dimethyl carbonate 616-38-6
- pentachlorophenol 87-86-5
<table>
<thead>
<tr>
<th>Gold Standard</th>
<th>OECD 429 LLNA</th>
<th>Endpoint:</th>
<th>Sensitization</th>
<th>Regulatory choice</th>
<th>Calculated</th>
<th>Sensitivity:</th>
<th>0.959</th>
<th>Specificity:</th>
<th>0.993</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test type:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity:</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specificity:</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality Factors:</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adapted sensitivity</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adapted specificity</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pr( C=1 )</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pr( C=0 )</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior Probabilities</td>
<td>Pr( C=1</td>
<td>T=1 )</td>
<td>99%</td>
<td>Pr( C=1</td>
<td>T )</td>
<td>50.00%</td>
<td>50.00%</td>
<td>50.00%</td>
<td>50.00%</td>
</tr>
<tr>
<td>Pr( C=0</td>
<td>T=0 )</td>
<td>96%</td>
<td>Pr( C=0</td>
<td>T )</td>
<td>50.00%</td>
<td>50.00%</td>
<td>50.00%</td>
<td>50.00%</td>
<td>50.00%</td>
</tr>
<tr>
<td>Threshold Probabilities (reproducability):</td>
<td>for a positive conclusion</td>
<td>79%</td>
<td>for a positive conclusion</td>
<td>50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for a negative conclusion</td>
<td>77%</td>
<td>for a negative conclusion</td>
<td>50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information Gap:</td>
<td>for a positive conclusion</td>
<td>29%</td>
<td>TEST PROPOSAL needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for a negative conclusion</td>
<td>27%</td>
<td>TEST PROPOSAL needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Case (proof-of-principle) of qWoE ITS: Skin Sensitization

- Excel implementation

- hydroxycitronellal 107-75-5

- DEREK alert found positive
- SMARTs alert found positive
- TOPKAT probability>0.7 positive
- TIMES-SS alert found positive
- MultiCASE alert found positive

- LLNA test (2 results) positive
Endpoint: Sensitization

Framework: OECD 429 LLNA - REACH A.VII

Gold Standard: OECD 429 LLNA

Information Requirement: OECD 429 LLNA

Prior Probability Threshold level: Positive 50% Pr(C=1) 80% Negative 50% Pr(C=0)

Battery of Available Tests
testnr. 1 2 3 4 5

type 429 LLNA - - - -

Test Result:

Quality Factors: 1.00 0.00 0.00 0.00 0.00
adapted sensitivity 0.959 0.500 0.500 0.500 0.500
adapted specificity 0.993 0.500 0.500 0.500 0.500
Pr(C=1) 50% 99% 99% 99% 99%
Pr(C=0) 50% 1% 1% 1% 1%

Posterior Probabilities

Pr(C=1 | T=1) 99%
Pr(C=0 | T=0) 96%

Posterior Probability of the qWoE conclusion
for a positive conclusion 99%
for a negative conclusion 1%

Threshold Probabilities (reproducability):
for a positive conclusion 79%
for a negative conclusion 77%

Information Gap:
for a positive conclusion -20%
for a negative conclusion 76%

WoE satisfied TEST PROPOSAL needed
### Quality Factors

<table>
<thead>
<tr>
<th>test nr.</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>test type</td>
<td>OECD 429 LLNA</td>
</tr>
</tbody>
</table>

#### General Questions

- Klimisch code of test quality (1,2,3 or 4)
  - 1 Performed under GLP (Y/N): Y
  - 2 Documentation OK (Y/N): Y
  - 3 Within Domain of Applicability (Y/N): Y
  - 4 etc…: Y

#### Test specific Questions

- 1 Coverage of all Mechanisms of Action (0..1): 1
- 2 Experimental issues (vehicle, test duration, etc): 1
- 3 etc…: 1

### OVERALL QUALITY FACTOR

1
Quality Factors

General Questions
Klimisch code of test quality (1, 2, 3 or 4)

1. Performed under GLP (Y/N)  Y
2. Documentation OK (Y/N)  Y
3. Within Domain of Applicability (Y/N)  Y
4. etc…

Test specific Questions
1. Coverage of all Mechanisms of Action (0..1)  1
2. Experimental issues (vehicle, test duration, etc)  1
3. etc…

OVERALL QUALITY FACTOR 0.8
Quality Factors

General Questions

Klimisch code of test quality (1,2,3 or 4)

1 Performed under GLP (Y/N)  
2 Documentation OK (Y/N)  
3 Within Domain of Applicability (Y/N)  
4 etc…

Test specific Questions

1 Coverage of all Mechanisms of Action (0..1)  
2 Experimental issues (vehicle, test duration, etc)  
3 etc…

OVERALL QUALITY FACTOR 0.64
Endpoint: Sensitization
Framework: OECD 429 LLNA

Gold Standard: OECD 429 LLNA

Battery of Available Tests

<table>
<thead>
<tr>
<th>testnr.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>test type</td>
<td>429 LLNA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Prior Probability

| Positive | Pr( C=1 ) | 50% |
| Negative | Pr( C=0 ) | 50% |

Threshold level:

- Positive: Pr( C=1 ) = 80%
- Negative: Pr( C=0 ) = 80%

Sensitivity: 0.959
Specificity: 0.993

Weight Factors:

- 1

Quality Factors:

- 0.64
- 0.00
- 0.00
- 0.00
- 0.00

Test Result: 1

Adapted sensitivity: 0.794
Adapted specificity: 0.816

Pr( C=1 ):

- 50%
- 81%
- 81%
- 81%
- 81%

Pr( C=0 ):

- 50%
- 19%
- 19%
- 19%
- 19%

Posterior Probabilities

| Pr( C=1 | T=1 ) | 99% |
| Pr( C=0 | T=0 ) | 96% |

Threshold Probabilities (reproducability):

- for a positive conclusion: 79%
- for a negative conclusion: 77%

Posterior Probability of the qWoE conclusion

- for a positive conclusion: 81%
- for a negative conclusion: 19%

Information Gap:

- for a positive conclusion: -2%
- for a negative conclusion: 58%

WoE satisfied

TEST PROPOSAL needed
### Endpoint: Sensitization

#### Prior Probability
- **Positive**: $Pr(C=1) = 50\%$
- **Negative**: $Pr(C=0) = 50\%$

#### Threshold level:
- **Positive**: $Pr(C=1) = 80\%$
- **Negative**: $Pr(C=0) = 80\%$

#### Gold Standard: OECD 429 LLNA

#### Framework: REACH A.VII

#### Information Requirement: OECD 429 LLNA

#### Battery of Available Tests
<table>
<thead>
<tr>
<th>testnr.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>test type</td>
<td>DEREK</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Test Result:

<table>
<thead>
<tr>
<th>Quality Factors</th>
<th>1.00</th>
<th>0.00</th>
<th>0.00</th>
<th>0.00</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>adapted sensitivity</td>
<td>0.788</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
</tr>
<tr>
<td>adapted specificity</td>
<td>0.506</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
</tr>
<tr>
<td>$Pr(C=1)$</td>
<td>50%</td>
<td>61%</td>
<td>61%</td>
<td>61%</td>
<td>61%</td>
</tr>
<tr>
<td>$Pr(C=0)$</td>
<td>50%</td>
<td>39%</td>
<td>39%</td>
<td>39%</td>
<td>39%</td>
</tr>
</tbody>
</table>

#### Posterior Probabilities
- **Positive**
  - $Pr(C=1 | T=1) = 99\%$
  - $Pr(C=0 | T=0) = 96\%$
- **Negative**
  - $Pr(C=1 | T) = 61.44\%$
  - $Pr(C=0 | T) = 38.56\%$

#### Threshold Probabilities (reproducability):
- **Positive conclusion**: 79% for a positive conclusion, 77% for a negative conclusion
- **Negative conclusion**: 77% for a positive conclusion, 39% for a negative conclusion

#### Information Gap:
- **Positive**
  - 18% for a positive conclusion, 18% for a negative conclusion
- **Negative**
  - 38% for a positive conclusion, 38% for a negative conclusion

**TEST PROPOSAL needed for a positive conclusion**

**TEST PROPOSAL needed for a negative conclusion**
Endpoint: Sensitization

Gold Standard OECD 429 LLNA

Framework: REACH A.VII

<table>
<thead>
<tr>
<th>Information Requirement</th>
<th>OECD 429 LLNA</th>
</tr>
</thead>
</table>

Prior Probability
- Positive: 50% Pr( C=1 )
- Negative: 50% Pr( C=0 )

Threshold level: 80%

Battery of Available Tests

<table>
<thead>
<tr>
<th>testnr.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>test type</td>
<td>DEREK</td>
<td>SMARTs</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Quality Factors:
- 1.00 1.00 0.00 0.00 0.00

Sensitivity: 0.959
Specificity: 0.993

Weight Factors: 1

Posterior Probabilities
- Pr( C=1 | T=1 ) 99%
- Pr( C=0 | T=0 ) 96%

Threshold Probabilities (reproducability):
- for a positive conclusion 79%
- for a negative conclusion 77%

Information Gap:
- for a positive conclusion 7%
- for a negative conclusion 49%

Posterior Probability of the qWoE conclusion
- for a positive conclusion 72%
- for a negative conclusion 28%

TEST PROPOSAL needed for a negative conclusion

Endpoint: Sensitization

Gold Standard OECD 429 LLNA

Framework: REACH A.VII

<table>
<thead>
<tr>
<th>Information Requirement</th>
<th>OECD 429 LLNA</th>
</tr>
</thead>
</table>

Prior Probability
- Positive: 50% Pr( C=1 )
- Negative: 50% Pr( C=0 )

Threshold level: 80%

Battery of Available Tests

<table>
<thead>
<tr>
<th>testnr.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>test type</td>
<td>DEREK</td>
<td>SMARTs</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Quality Factors:
- 1.00 1.00 0.00 0.00 0.00

Sensitivity: 0.788 0.776 0.500 0.500 0.500
Specificity: 0.506 0.522 0.500 0.500 0.500

Weight Factors: 1

Posterior Probabilities
- Pr( C=1 | T=1 ) 61.44%
- Pr( C=0 | T=0 ) 38.56%

Threshold Probabilities (reproducability):
- for a positive conclusion 79%
- for a negative conclusion 77%

Information Gap:
- for a positive conclusion 7%
- for a negative conclusion 49%

Posterior Probability of the qWoE conclusion
- for a positive conclusion 72%
- for a negative conclusion 28%

TEST PROPOSAL needed for a negative conclusion
### Endpoint:
Sensitization

### Prior Probability
- Positive: 50% \(Pr(\ C=1 )\)
- Negative: 50% \(Pr(\ C=0 )\)

### Threshold level:
- Positive: 80%

### Gold Standard
OECD 429 LLNA

### Framework:
REACH A.VII

### Battery of Available Tests
- testnr. 1
- test type: TIMES- SS
- Specificity: 0.993
- Test Result: 1

### Information Requirement
OECD 429 LLNA

### Test Type:
- DEREK SMARTs
- TIMES- SS
- TOPK Multi CASE

### Sensitivity:
- 0.788
- 0.776
- 0.562
- 0.670
- 0.877

### Specificity:
- 0.506
- 0.522
- 0.761
- 0.444
- 0.106

### Weight Factors:
- 1

### Posterior Probabilities
- \(Pr(\ C=1 \mid T=1 )\): 99%
- \(Pr(\ C=0 \mid T=0 )\): 96%

### Threshold Probabilities (reproducibility):
- for a positive conclusion: 79%
- for a negative conclusion: 77%

### Posterior Probability of the qWoE conclusion
- for a positive conclusion: 88%
- for a negative conclusion: 12%

### Information Gap:
- for a positive conclusion: -8%
- for a negative conclusion: 65%

### WoE satisfied
TEST PROPOSAL needed
Endpoint: Sensitization

Gold Standard: OECD 429 LLNA

Information Requirement: OECD 429 LLNA

Sensitivity: 0.959
Specificity: 0.993

Weight Factors: 1

Prior Probability Threshold level:

Positive 50% Pr( C=1 )
Negative 50% Pr( C=0 )

Threshold level: 80%

Battery of Available Tests

testnr. 1 2 3 4 5

type DEREK SMARTs TIMES- SS TOPK Multi CASE

Test Result: 1 1 1 1 1

Quality Factors:

0.80 0.80 0.80 0.80 0.80

adapted sensitivity
0.730 0.721 0.549 0.636 0.802
adapted specificity
0.504 0.518 0.709 0.456 0.184
Pr( C=1 )
50% 60% 69% 81% 83%
Pr( C=0 )
50% 40% 31% 19% 17%

Posterior Probabilities

Pr( C=1 | T=1 ) 99%
Pr( C=0 | T=0 ) 96%

Threshold Probabilities (reproducability):

for a positive conclusion 79%
for a negative conclusion 77%

Information Gap:

for a positive conclusion -3%
for a negative conclusion 60%

WoE satisfied
TEST PROPOSAL needed

Posterior Probability of the qWoE conclusion

for a positive conclusion 83%
for a negative conclusion 17%
Case (proof-of-principle) of qWoE ITS: Skin Sensitization

- Excel implementation
  
  - dimethyl carbonate 616-38-6
  
  - DEREK no alert found negative
  - SMARTs no alert found negative
  - TOPKAT probability<0.7 negative
  - TIMES-SS no alert found negative
  - MultiCASE no alert found negative

- LLNA test (1 results) negative
Endpoint: Sensitization

Prior Probability
Positive: 50% Pr(C=1)
Negative: 50% Pr(C=0)

Gold Standard: OECD 429 LLNA
Framework: REACH A.VII

Battery of Available Tests

<table>
<thead>
<tr>
<th>testnr.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>test type</td>
<td>DEREK</td>
<td>SMARTs</td>
<td>TIMES-SS</td>
<td>TOPK</td>
<td>Multi CASE</td>
</tr>
<tr>
<td>Sensitivity:</td>
<td>0.788</td>
<td>0.776</td>
<td>0.562</td>
<td>0.670</td>
<td>0.877</td>
</tr>
<tr>
<td>Specificity:</td>
<td>0.506</td>
<td>0.522</td>
<td>0.761</td>
<td>0.444</td>
<td>0.106</td>
</tr>
<tr>
<td>Test Result:</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Quality Factors:</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>adapted sensitivity</td>
<td>0.788</td>
<td>0.776</td>
<td>0.562</td>
<td>0.670</td>
<td>0.877</td>
</tr>
<tr>
<td>adapted specificity</td>
<td>0.506</td>
<td>0.522</td>
<td>0.761</td>
<td>0.444</td>
<td>0.106</td>
</tr>
<tr>
<td>Pr( C=1 )</td>
<td>50%</td>
<td>30%</td>
<td>15%</td>
<td>9%</td>
<td>7%</td>
</tr>
<tr>
<td>Pr( C=0 )</td>
<td>50%</td>
<td>70%</td>
<td>85%</td>
<td>91%</td>
<td>93%</td>
</tr>
</tbody>
</table>

Posterior Probabilities

| Pr( C=1 | T=1 ) | 99% |
| Pr( C=0 | T=0 ) | 96% |

Threshold Probabilities (reproducability):
for a positive conclusion: 79%
for a negative conclusion: 77%

Posterior Probability of the qWoE conclusion

for a positive conclusion: 8%
for a negative conclusion: 92%

Information Gap:

| for a positive conclusion | 71% |
| for a negative conclusion | -15% |

TEST PROPOSAL needed
WoE satisfied

H₃C⁻O⁻C=O⁻CH₃

Rorije - OSIRIS Stakeholder Workshop, Berlin

March 1-2, 2010
Case (proof-of-principle) of qWoE ITS: Skin Sensitization

- Excel implementation
  - Pentachlorophenol 87-86-5
  - DEREK: no alert found negative
  - SMARTs: alert found positive
  - TIMES-SS: no alert found negative
  - TOPKAT: probability<0.7 negative
  - MultiCASE: alert found positive
  - LLNA test (2 results) ?
March 1-2, 2010  Rorije - OSIRIS Stakeholder Workshop, Berlin

National Institute for Public Health and the Environment

Prior Probability Threshold level: Positive 50% Pr(C=1) 80% Negative 50% Pr(C=0) Gold Standard OECD 429 LLNA Battery of Available Tests

testnr. 1 2 3 4 5

test type DEREK SMARTs TIMES- SS TOPK AT Multi CASE

Information Requirement OECD 429 LLNA

Sensitivity: 0.959
Specificity: 0.993

Weight Factors: 1

Test Result:
0 1 0 0 1

Quality Factors:
1.00 1.00 1.00 1.00 1.00

adapted sensitivity 0.788 0.776 0.562 0.670 0.877
adapted specificity 0.506 0.522 0.761 0.444 0.106
Pr(C=1)
50% 30% 41% 28% 22.6%
Pr(C=0)
50% 70% 59% 72% 77%

Posterior Probabilities
Pr(C=1 | T=1) 99%
Pr(C=0 | T=0) 96%

Threshold Probabilities (reproducability):
for a positive conclusion 79%
for a negative conclusion 77%

Posterior Probability of the qWoE conclusion
for a positive conclusion 22%
for a negative conclusion 78%

Information Gap:
for a positive conclusion 57%
for a negative conclusion -1%

TEST PROPOSAL needed
WoE satisfied
Sensitization

Gold Standard OECD 429 LLNA

Information Requirement OECD 429 LLNA

Sensitivity: 0.959
Specificity: 0.993

Weight Factors: 1

Posterior Probabilities

Pr( C=1 | T=1 ) 99%
Pr( C=0 | T=0 ) 96%

Threshold Probabilities (reproducability):

for a positive conclusion 79%
for a negative conclusion 77%

Information Gap:

for a positive conclusion 53%
for a negative conclusion 4%

Prior Probability

Positive 50% Pr( C=1 )
Negative 50% Pr( C=0 )

Threshold level:

80%

Battery of Available Tests
testnr. 1 2 3 4 5

test type DEREK SMARTs TIMES- TOPK Multi CASE

Sensitivity: 0.788 0.776 0.562 0.670 0.877
Specificity: 0.506 0.522 0.761 0.444 0.106

Test Result: 0 1 0 0 1

Quality Factors: 0.80 1.00 1.00 1.00 1.00

adapted sensitivity 0.730 0.776 0.562 0.670 0.877
adapted specificity 0.504 0.522 0.761 0.444 0.106
Pr( C=1 ) 50% 35% 46% 33% 27%
Pr( C=0 ) 50% 65% 54% 67% 73%

Posterior Probabilities

Pr( C=1 | T=1 ) 99%
Pr( C=1 | T ) 34.8% 46.5% 33.3% 27.1% 26.7%
Pr( C=0 | T=0 ) 96%
Pr( C=0 | T ) 65.2% 53.5% 66.7% 72.9% 73.3%

Posterior Probability of the qWoE conclusion

for a positive conclusion 27%
for a negative conclusion 73%

Endpoint: Sensitization

Sensitization Framework: REACH A.VII

OECD 429 LLNA

OECD 429 LLNA

TEST PROPOSAL needed for a negative conclusion
TEST PROPOSAL needed for a negative conclusion
Case (proof-of-principle) of qWoE ITS: Skin Sensitization

• Excel implementation
  - Pentachlorophenol 87-86-5
    - DEREK no alert found negative
    - SMARTs alert found positive
    - TIMES-SS no alert found negative
    - TOPKAT probability<0.7 negative
    - MultiCASE alert found positive
  - LLNA test (2 results) positive
OSIRIS Intelligent testing strategy tool

OSIRIS web tool

This is the fourth version of the web tool implementation of the Integrated Testing Strategies (ITS) to be developed within OSIRIS.

New features in this version are:
• **Integration with Chemprop database**: test data contained in Chemprop datasets can be imported into a substance.
• **Current version includes two new endpoints, Aquatic Toxicity**, using a preliminary ITS, and **Sensitisation**, using an ITS that includes qWoE.
• **Advanced probabilistic-based decision methods** are supported by the tool. They are included in the ITS for Sensitisation.
You are here: Substances > hydroxycitronellal > New Assessment

New Assessment

Name: In silico Skin Sensitizer?

Endpoint: Sensitisation

Requirement: Annex VII: 1 tonne or more

Save or Cancel
Assessment execution

Name: In silico Skin Sensitizer?
Endpoint: Sensitisation
State: CONCLUDED

Execution result
A conclusion has been reached.
Substance is Sensitising, using a target probability of 80%.
The quantitative Weight of Evidence approach, using the tests in Used information, has obtained this result:
Sensitising: 85.9%
Not sensitising : 14.1%
using a Bayesian Network implementation.

Used information
In silico data items used
SMARTs  Skin sensitisation sensitising Reliability 1
DEREKfW  Skin sensitisation sensitising Reliability 1
TIMES-SS  Skin sensitisation sensitising Reliability 1
You are here: Substances > hydroxycitronellal > In silico Skin Sensitizer? > Assessment execution

Assessment execution

Name: In silico Skin Sensitizer?
Endpoint: Sensitisation
State: CONCLUDED

Execution result
A conclusion has been reached.
Substance is Sensitising, using a target probability of 30%.
The quantitative Weight of Evidence approach, using the tests in Used information, has obtained this result:
- Sensitising: 81.9%
- Not sensitising: 18.1%
using a Bayesian Network implementation.

Used information

In silico data items used
- SMARTs  Skin sensitisation sensitising Reliability 1
- DEREKfW Skin sensitisation sensitising Reliability 2
- TIMES-SS Skin sensitisation sensitising Reliability 2

Steps followed
Assessment execution

Name: Assessment 1
Endpoint: Sensitisation
State: EXECUTED

Execution result

Please, provide data of in-silico methods from the list below. There is insufficient information: the target probability 95% has not been reached.

Test types

TopKat Skin Sensitization

Create and assign in silico data
Assign in silico data

The quantitative Weight of Evidence approach, using the tests in Used information, has obtained this result:

Sensitising: 9.4%
Not sensitising: 90.6%

using a Bayesian Network implementation.

Used information

In silico data items used

- DEREK7W Skin sensitisation not sensitising Reliability 1
- SMARTs Skin sensitisation not sensitising Reliability 1
- TIMES-SS Skin sensitisation not sensitising Reliability 1
Conclusion:

Bayesian network tool to Weight of Evidence seems transparent, reproducible, flexible, objective, practical

Discussion:

specific choices / information required to make this work...

1. REACH Information Requirement → Threshold Probability
2. List of all possible data sources → statistical performance
3. Quality Factors – endpoint / method specific checklists?
4. Costs (animals, money, time) with an optimization algorithm

Acceptance?