
Subject: thresholds and weights of toxicity risk
Posted by [nbehrnd](#) on Tue, 22 Aug 2023 15:45:36 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear Thomas,

the generation of evolutionary libraries can be restrained by fitness criterion «toxicity risk». Based on some preliminary tests, I have difficulties to comprehend the thresholds available. Departing from «nasty functions» as one of DW's calculated properties, one could assume DW counts functional groups. The suggested default is 2, an integer. Then, choosing a toxicity risk less or equal to zero would render the restraint a constraint.

Interestingly, however, DataWarrior equally permits to enter an upper threshold as a real/floating number (a negative one like `-1.23` does not hinder DW to work either).

Can you please share some insight how to engage well these thresholds? Do the weights act on a linear penalty scale to provide the fitness criterion greater influence on the overall fitness of the molecules suggested? (DW's internal documentation (via F1) includes the string «toxicity» only once.)

Best regards,

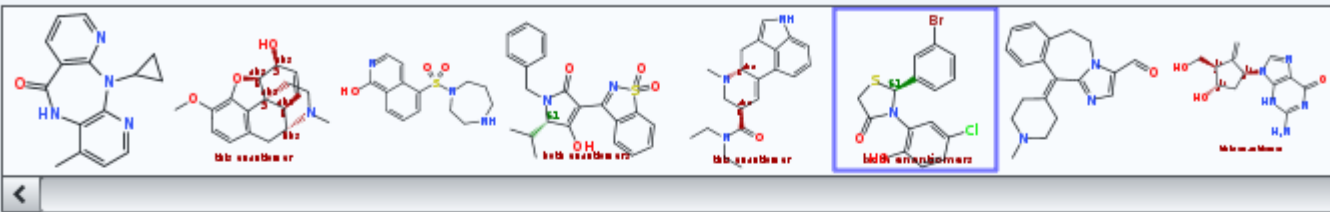
Norwid

File Attachments

1) [toxicity_risk.png](#), downloaded 618 times

Build Evolutionary Library

Root generation compounds: Build at task execution time



(Select sub-structures to protect them from being changed)

Cycle Create compounds like

Compounds per cycle

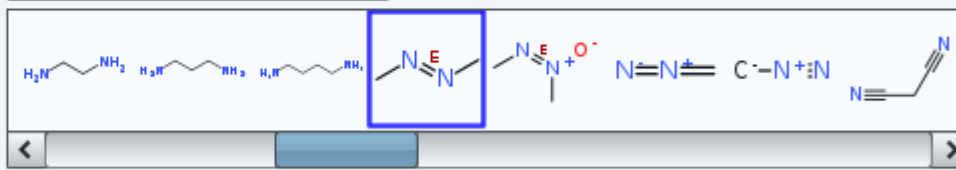
Compounds survive a cycle Total run count:

Fitness Criteria

Create molecules

Descriptor used:

Weight:



Prefer 'Toxicity Risk' \geq and \leq Weight:

Subject: Re: thresholds and weights of toxicity risk
 Posted by [thomas](#) on Thu, 24 Aug 2023 12:58:04 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear Norwid,

I admit that fitness criteria are not covered by the manual, because originally they were very simple.

The calculated combined toxicity risk of a compound is simply the sum of risk contributions from the 4 toxicity classes (mutagenic, tumorigenic, irritant, and reproductive effective. A low risk

contributes 1, a high risk contributes 2 to the combined score. Thus, a non-existing combined risk would be 0 and the very unlikely theoretical maximum is 8.

Threshold may be fractional numbers, because they describe the sigmoid curve that translate the risk (0..8) into a fitness (1..0)

The determination, whether a compound is assigned a no, low or high risk is based on a statistical approach based on whether and which fragments are present from a predefined fragment list, which are frequently found in toxic compound of the particular toxicity class.

Hope, this explains it somewhat...

Subject: Re: thresholds and weights of toxicity risk
Posted by [nbehrnd](#) on Fri, 25 Aug 2023 08:21:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear Thomas,

the explanation indeed provides the additional insight reading the file `properties.html` [1] in the GitHub repository did not. For the interested future reader: the relevant illustrated section starts on page 3 of the attached print to .pdf attached below.

Thank you,

Norwid

[1] <https://github.com/thsa/datawarrior/blob/master/src/html/properties/properties.html>

File Attachments

1) [Calculated Compound Properties.pdf](#), downloaded 883 times
