Subject: Importing compounds with assay data from Pubchem into Datawarrior
Posted by Jo W on Fri, 23 Jul 2021 11:12:44 GMT
View Forum Message <> Reply to Message

Many thanks Thomas and colleagues
Data warrior is a "life saver" for non coding people like me - thank you so much for you and your team for developing it.

Its also one of (I think it is THE) easiest software to use for a variety of informatics / prediction modeling etc.

Importing / uploading to Datwarrior is usually really straightforward compared to other programs.

For example, it's possible to download compounds from Pubchem and via a text / csv file import into Datawarrior with no editing. This works really well and is easy to do.

However when trying to download, say 30,000 compounds that are also P450 inhibitors an error comes up, maybe because the request has too many data points or maybe it cant be done via a web page (Pubchem talks about Restful / ftp but for non coding people this seems very difficult to do - 2 hours trying to figure this out did not prove fruitful!

My questions/requests/comments are:

1. A request - Can Datawarrior create a built in search for Pubchem along similar lines to the Chembl facility? Pubchem seems to have so many more compounds with accompanying biodata in it compared to Chembl.

2. In the meantime - how can you import compounds with specific biodata from Pubchem into datawarrior?

Can you use Datawarriors URL import? I tried this for example with a list of P450 enzyme inhibition active compounds from Pubchem (by copying and pasting the url, but this did not work:
https://pubchem.ncbi.nlm.nih.gov/protein/P10633#section=Chem icals-and-Bioactivities

Any ideas regarding the URL or any other "non coding" ways to get data form Pubchem into Datawarrior? As stated above, you cannot simply download a csv file from Pubchem (at least for obtaining biodata) from the above link as an error occurs, even though Pubchem provides a downlaod "button" for it.

Many thanks in advance
Jon