Reviewer’s report

Title: GenDrux: A guided system for supporting prescription by gene expression-based sensitivity

Version: 4  Date: 19 November 2010

Reviewer: Adrien Coulet

Reviewer’s report:

The new scope of the paper make it much more consistent.

- Major Compulsory Revisions
  none

- Minor Essential Revisions

When introducing Lingpipe, authors must mention that Lingpipe includes LingMed that enable extraction of Medline info, as e-utils does.

In the state of the art, authors mention PharmGKB, lingpipe, Banner and Abner. That seems good. What are the interest and the drawback of these apporaches if you compare to GenDrux.

It appears crucial to clarify the section about unsupervised appraoches. It can be understood that PharmGKB is unsupervised and have to be downloaded. Resources and approaches to populate resources must be distinguish since it is two different aspects.

The query choosen is really restrictive because an abstract with breast+cancer+BRCA1+tamoxifen but without gene and drug terms will be out of the search of space.

Why not also generating all possible queries with breast+cancer+ all possible {gene name+drug name} pairs of some lexicon? If MeSH annotations resolve this issue, it must be explained.

Limitation of this query must be discussed.

The motivation of the query strategy with e-utils is obscure. Do you get all the pmids mentionning breast-cancer+gene+drug? If your algorithm "loop", I understand that you got them all for each period (and not only the first 500 fro each period)? If so why just querying once without restriction on time periods? The final set of pmids will be similar? This must be clarified

I strongly disagree with the sentence "Making predictions of the actual relationships between genes and drugs from a computational-driven assessment of the text is therefore, not possible."

"This archive file serves as the knowledgebase" is an abusive use of knowledge
base. It seems to be a database in this case.

- Discretionary Revisions
  none

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.