Method Title

Luis Pedro Coelho

Instituto de Medicina Molecular

*Suggested Page Limit: 3 pages.*

**Summary Sentence**: include 1 short sentence description of your method that can be used in a main text table.

**Background/Introduction**

Please try to address the following points:

* What is the motivation for your approach? This will include any previous work and observations that you have made about the NCI-DREAM data to suggest your approach is a good one. Provide the reader with an intuition of how you approached the problem
* What is the underlying methodology used (e.g., SVM or regression)?
* Where there any novel approaches taken in regards to feature selection, data imputation, ranking, etc?

**Methods**

The main idea behind the final solution was the following

Two drugs will act in a synergistic fashion if they disturb different genes with related functionality.

For the implementation of this idea, I relied, again, on the GO terms (molecular function and biological process only).

First, the media and DMSO data was used to determine a baseline variation and genes were considered disturbed if they moved more than 1.5 standard deviations away from the mean at any time point. This is a noisy measurement, but it will only be used very indirectly. As above, genes were also grouped together by GO terms.

Each drug was thus characterized by two signatures:

1. A set of disturbed genes
2. A set of disturbed GO terms

For each pair of drugs, I computed the correlation between its disturbed genes and its disturbed GO terms.

There is a roughly linear correlation between these two values, but the distance to the regression line is our measure of synergy. Drugs pairs for which the GO terms are more correlated than predicted by the correlation at the gene level were predicted to be synergistic.

The methods should cover a full description of your methods so a reader can reproduce them. Please cover how you processed the data, if any data was imputed or manipulated in any way (e.g., you mapped data onto pathways or combined different datasets), the underlying algorithm, any modifications to the underlying method of importance, the incorporation of outside data (e.g., genesets or GO terms), and the approach to ranking.

**Conclusion/Discussion**

This section should include a short summary and any insights gained during the algorithm. For example, which dataset was most informative? You can include future directions. You may also add some discussion on the general performance of your methodology (if you wish) and if there were pitfalls, what are they?

**References**

**Authors Statement**

LPC developed the methodology, implemented it, and wrote the report.