Recommending New Target Conditions for Drug Retesting Using Temporal Patterns in Clinical Trials: A Proof of Concept

Zhe He, Chunhua Weng
Department of Biomedical Informatics, Columbia University
Disclosure

• Both authors disclose that they have no financial relationships with commercial interests.
Learning Objective

• After attending this session, the learners will be able to:
  • Analyze the temporal pattern of drug retesting in retrospective clinical trials
  • Leverage the metadata in clinical trial summaries to narrow the search for new target conditions
Background

- De novo drug discovery
- Drug repurposing: discovery of novel indication of existing drugs

Duloxetine

- Depression
- Stress urinary incontinence

Successful drug repurposing cases were mostly identified by serendipity

Computational methods have been proposed

Approach

- ClinicalTrials.gov & its use for drug repurposing (Zhang et al. 2014)

1. **Eligibility**

   - Ages Eligible for Study: 40 Years to 60 Years
   - Genders Eligible for Study: Both
   - Accepts Healthy Volunteers: Yes

   **Criteria**
   - Inclusion Criteria:
     1. Healthy volunteers:
        - Males or females
        - On no medications except for the contraceptive pill and without medical illnesses in the last three months.
        - Non-smokers
        - 40 - 60 years of age.
     2. T2DM subjects:
        - Males or females
        - Diagnosis of T2DM
        - 40 - 60 years of age
        - HbA1C: 6.5 - 9.5%
        - Duration of diabetes 1 - 10 years
        - Diabetes treated with diet, or tablets only.

   **Exclusion Criteria:**
   - Healthy volunteers:
     - Pregnancy
     - Lack of contraception in women of child bearing age
     - Chronic medical conditions
   - Current smokers
   - Evidence of ischaemia on ECG
   - Drop attacks
   - Alcohol or drug abuse
   - Psychiatric illness

Drug retesting often occurred in conditions whose trials employed similar eligibility criteria. We explore the feasibility of using the data from CT.gov to narrow the search for drug repurposing targets.
Configuration for Drug Retesting

1. What drugs were often retested on different conditions?
2. How similar are the eligibility criteria of trials on A and B?
3. Can we leverage drug retesting patterns to recommend new target conditions for existing drugs?
Data Preparation (1)

Trial summaries from CT.gov

Extracting metadata of trials

Indexing trials by conditions

Extracting common eligibility features

Extracting n-grams from free-text EC

Partially match a UMLS concept?

Yes

Normalizing to a UMLS CUI

Retained CUIs appearing 3% of trials

Common Eligibility Feature:

e.g., Type 2 diabetes trials:

Metformin;

Contraceptive method;

.....

Data Preparation (2)

- 59,716 drug intervention trials between 2003 and 2013
- Included drugs used in >= 5 trials on the same condition in a year
- Formulated each retesting case as a quintuple:
  - Drug: *Duloxetine*
  - Initial condition: *Depression* first tested in 1995
  - Retested condition: *Stress urinary incontinence* first tested in 2004
- Excluded “placebo” from the dataset
- # of drugs: 550
- # of conditions: 451
- # of drug-condition pairs: 4,351
Network Visualization of Drug Retesting Patterns
# Pairwise Temporal Analysis of Drug Retesting Cases

# of retested drugs (# of pairs of target conditions)

<table>
<thead>
<tr>
<th>Yr 2 Yr 1</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>46(2982)</td>
<td>34(2278)</td>
<td>26(1212)</td>
<td>18(1035)</td>
<td>22(864)</td>
<td>13(560)</td>
<td>9(221)</td>
<td>9(284)</td>
<td>9(155)</td>
<td>11(251)</td>
</tr>
<tr>
<td>2004</td>
<td>--</td>
<td>39(1276)</td>
<td>31(787)</td>
<td>24(491)</td>
<td>21(516)</td>
<td>13(333)</td>
<td>9(236)</td>
<td>5(47)</td>
<td>3(20)</td>
<td>10(95)</td>
</tr>
<tr>
<td>2005</td>
<td>--</td>
<td>--</td>
<td>31(821)</td>
<td>30(554)</td>
<td>18(180)</td>
<td>15(471)</td>
<td>9(231)</td>
<td>8(95)</td>
<td>3(11)</td>
<td>8(67)</td>
</tr>
<tr>
<td>2006</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>24(454)</td>
<td>20(256)</td>
<td>15(435)</td>
<td>14(292)</td>
<td>11(108)</td>
<td>7(61)</td>
<td>7(57)</td>
</tr>
<tr>
<td>2007</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>19(333)</td>
<td>17(218)</td>
<td>14(179)</td>
<td>10(129)</td>
<td>4(82)</td>
<td>7(28)</td>
</tr>
<tr>
<td>2008</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>22(183)</td>
<td>16(152)</td>
<td>8(61)</td>
<td>3(17)</td>
<td>5(20)</td>
</tr>
<tr>
<td>2009</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>13(385)</td>
<td>13(91)</td>
<td>4(24)</td>
<td>5(33)</td>
</tr>
<tr>
<td>2010</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>13(144)</td>
<td>5(50)</td>
<td>4(11)</td>
</tr>
<tr>
<td>2011</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>13(143)</td>
<td>6(86)</td>
</tr>
<tr>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>7(80)</td>
</tr>
</tbody>
</table>
Analysis of Condition Relatedness

- **Hypothesis:**
  - Drug retesting often occurred between conditions whose trials used similar eligibility criteria

- Similarity: # of shared Common Eligibility Features (CEFs)

- Aggregated the retested drugs investigating the same pair of conditions

- Analyzed the distribution of # condition pairs over # of retested drugs
Shared CEFs of Conditions involving Drug Retesting

Avg # of CEFs shared by any two conditions: 52

Avg # of shared CEFs of condition pairs involving drug retesting is 139
Drug X will be recommended for Condition B if:

- Drug X
- Condition A
- # Shared CEFs > threshold
- Condition B
- Drug Y

Tested
Recommended
Drug Retesting Recommendation

- **No. of candidate drugs**
- **No. of different conditions**

<table>
<thead>
<tr>
<th>Threshold values (minimum No. of shared CEFs)</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>120</th>
<th>140</th>
<th>160</th>
<th>180</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>543</td>
<td>543</td>
<td>540</td>
<td>529</td>
<td>455</td>
<td>318</td>
<td>270</td>
<td>218</td>
<td>170</td>
<td>129</td>
</tr>
<tr>
<td>20</td>
<td>364</td>
<td>360</td>
<td>345</td>
<td>318</td>
<td>270</td>
<td>193</td>
<td>170</td>
<td>156</td>
<td>183</td>
<td>129</td>
</tr>
</tbody>
</table>
Limitations

- Do not work for new conditions and drugs
- Concept-level common eligibility features
  - “myocardial infarction within the last five years”
- Data quality issues in ClinicalTrials.gov
Future Work

• Drug retesting path linking multiple conditions over time
• Tuning the parameters, e.g., empirical threshold values
• Enriching the drug repurposing prediction method with SNOMED CT, DrugBank, OpenFDA
• Will formally evaluate the method with precision, recall, and f-measure.
Summary

• Drug retesting often occurred between conditions whose trials used similar eligibility criteria for participant selection.

• Leverage the design patterns in drug intervention trials to recommend potential new conditions for drug retesting.

• Provide very preliminary proof of concept.

• More sophisticated models should be developed to further test this idea.
Acknowledgements

Funding support:

• *National Library Medicine*
  
  R01 LM009886 (PI: Weng)

• *National Center for Advancing Translational Science*
  
  UL1 TR000040 (PI: Ginsberg)
Thank you!

Questions?

Contact:

Dr. Zhe He

zh2132@cumc.columbia.edu
Additional Slides
Analysis of Temporal Patterns

• For the 10-year time window, we constructed a 10 x 10 matrix.
  • Row $i$ and column $j$ being each year during the time window.
  • $d_{i,j}$ represents the number of distinct drugs that were first studied for one condition in year $i$ and later for a different condition in year $j$.
  • $c_{i,j}$ represents the number of distinct pairs of conditions in which a drug was tested for one condition in year $i$ and later for a different condition in year $j$. 

## Most Frequent Initial and Retested Conditions

<table>
<thead>
<tr>
<th>Top five frequent initial conditions</th>
<th># of condition pairs</th>
<th># of retested drugs</th>
<th>Top five frequent retested conditions</th>
<th># of condition pairs</th>
<th># of retested drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory tract diseases</td>
<td>173</td>
<td>35</td>
<td>Skin diseases</td>
<td>140</td>
<td>14</td>
</tr>
<tr>
<td>Carcinoma</td>
<td>167</td>
<td>46</td>
<td>Digestive system diseases</td>
<td>133</td>
<td>30</td>
</tr>
<tr>
<td>Vascular disease</td>
<td>167</td>
<td>30</td>
<td>Gastrointestinal diseases</td>
<td>133</td>
<td>30</td>
</tr>
<tr>
<td>Immunoproliferative disorders</td>
<td>164</td>
<td>39</td>
<td>Urologic diseases</td>
<td>124</td>
<td>10</td>
</tr>
<tr>
<td>Lymphoproliferative disorders</td>
<td>164</td>
<td>39</td>
<td>Neoplasm metastasis</td>
<td>117</td>
<td>19</td>
</tr>
</tbody>
</table>
Recommending Drug Retesting Candidate

A prediction was made if:

(1) a drug has been tested for the initial condition but has never been tested for the possible different condition

(2) there exists another drug that has been tested for both conditions

(3) the number of shared CEFs between two conditions is above a threshold.
Examples of Recommendations

For example, “Ranolazine” was predicted as a drug to be retested for myocardial infarction, because

(1) “Ranolazine” was tested for ischemia but has never been tested for myocardial infarction

(2) there exists another drug (i.e., “Ticagrelor”) that was tested for ischemia first and then retested for myocardial infarction,

(3) ischemia and myocardial infarction had 112 shared CEFs.

This prediction was confirmed by Hale et al.

Our prediction of “Everolimus” for treating rheumatic disease was confirmed by Yoon et al.
