Novelty in Drug Innovation

NBER Value of Medical Research White Paper

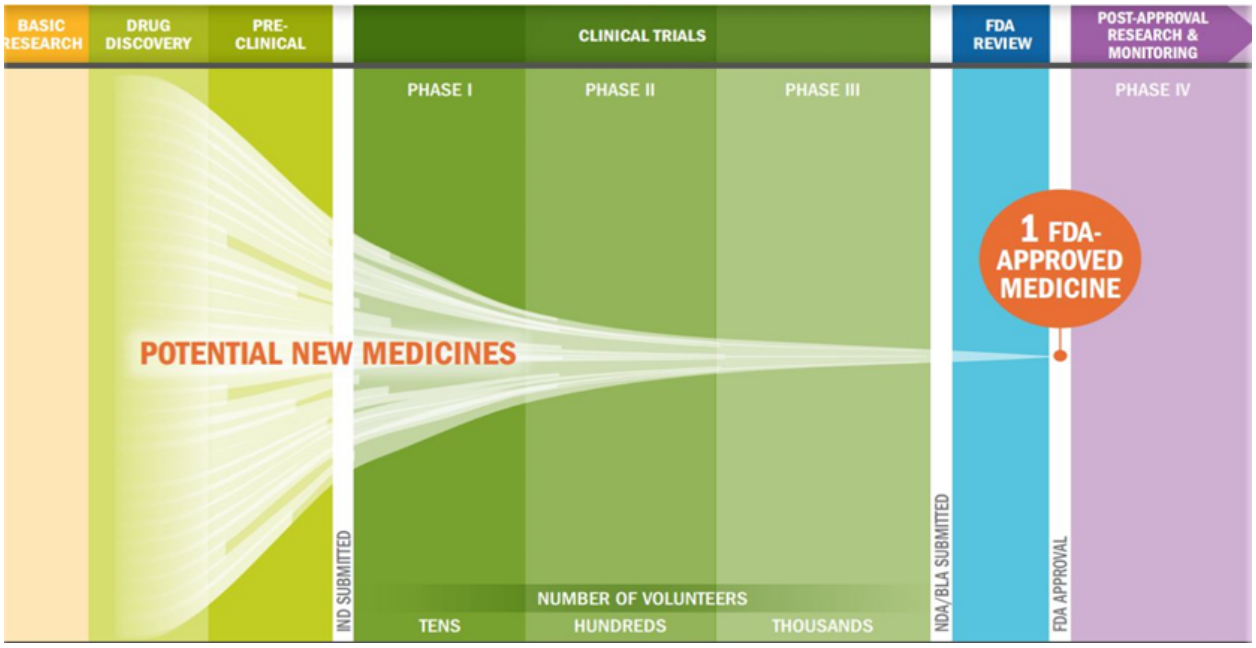
Joshua Krieger (w/Danielle Li and Dimitris Papanikolaou)

September 1, 2016

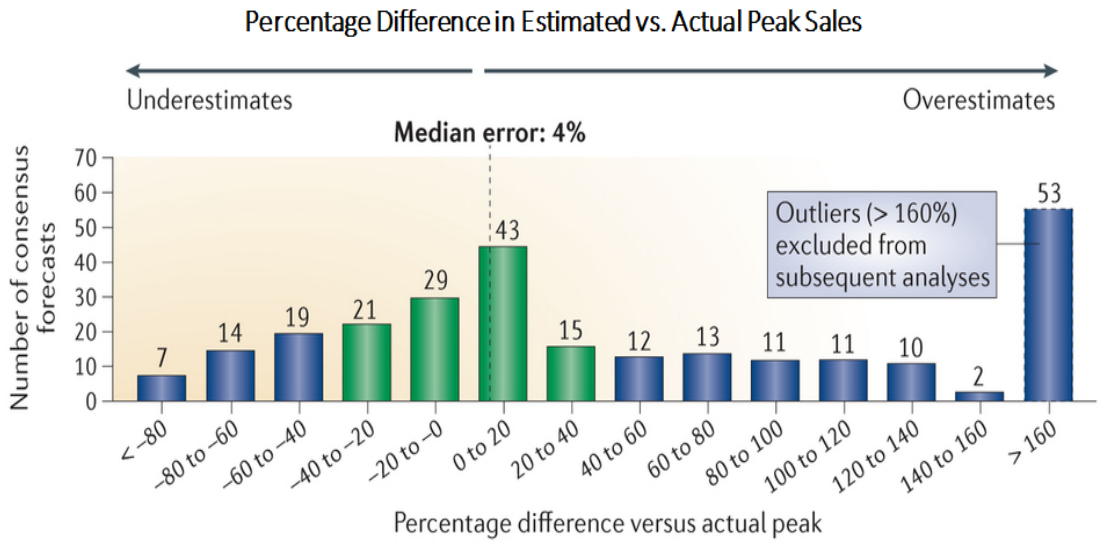
Why Measuring Drug Novelty Matters

Consumers want innovations that will improve their health and extend their lives. Yet a steady stream of new medical products alone may not translate into improved health outcomes because many new products are very similar to existing ones. While such incremental innovations may improve efficacy or access on the margin, sustainable improvements in health and welfare depend on the development of products that treat medical conditions in novel ways.

One of the most important areas of health innovation is the development of new drugs. Unfortunately, the process of drug discovery is difficult and outcomes are uncertain, so drug development projects fail at different stages and for a variety of reasons. Bringing a drug through the development and regulatory process to market is a highly uncertain process (in terms of both likelihood of success, and the profitability for drugs that do succeed) and can take many years.



Source: http://www.phrma.org/sites/default/files/pdf/rd_brochure_022307.pdf



Source: Cha et al., Nature Drug Discovery (2013)

Despite this uncertainty, breakthrough drugs can have incredible value in improving health and welfare. Consider the Hepatitis C drug Sovaldi, which in spite of a $84,000 a year price tag is still considered "cost-effective" due to a greater than 90% cure rate. Another example is the drug Ventoclax, which in the FDA gave an accelerated approval status to in 2016. This new drug led to complete or partial remission for chronic lymphocytic leukemia patients. In both cases, these new drugs provided enormous value over previous therapies.

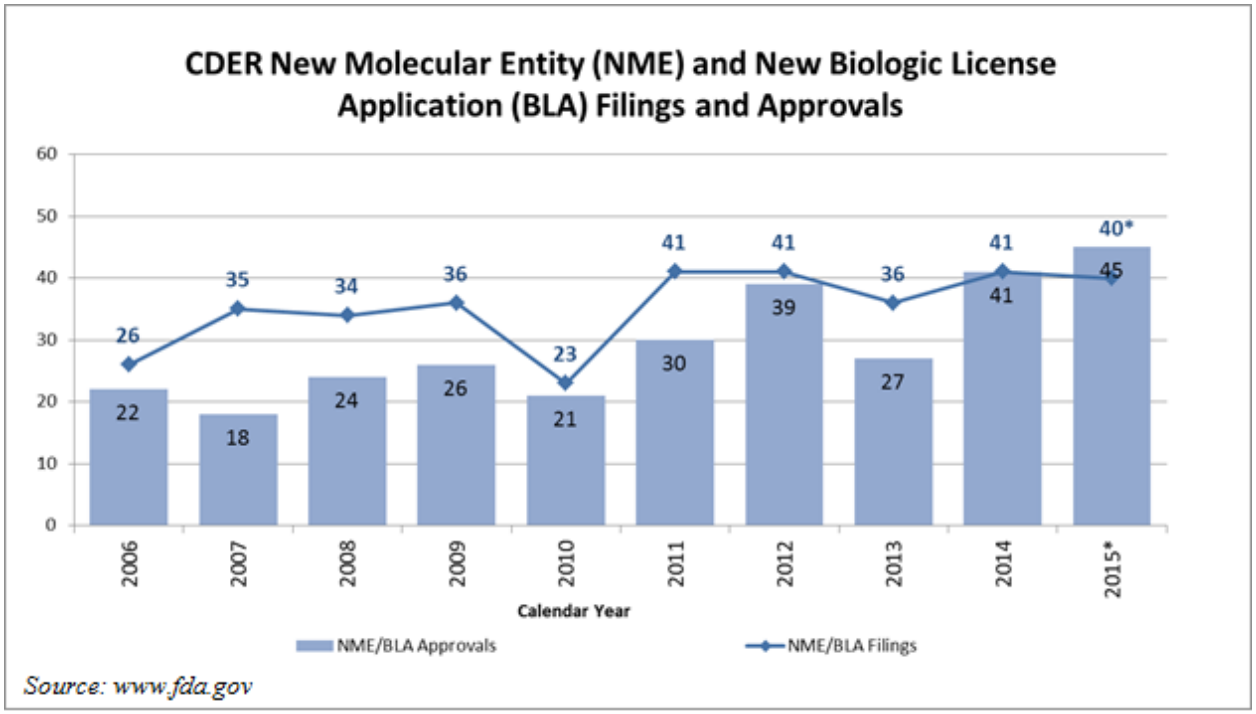
However, not all drugs have such clear value. For example, some drugs fall into the category of "copycat" or "me-too" drugs, where the new drug is merely a slight variation of an older drug, with no meaningful new clinical benefits. One such example of "me-too" drug is AstraZeneca's Nexium, a new treatment developed to replace Prilosec, the company's existing acid reflux blockbuster. Nexium conferred essentially no treatment advantage over Prilosec -- in fact the two drugs were nearly identical molecules -- but had the advantage of extending AstraZeneca's portfolio of patented heartburn drugs by nearly a decade. This example highlights a common popular narrative in which pharmaceutical firms invest in lower-risk me-too drugs, rather than in truly innovative therapies. If firms can leverage their marketing abilities to extract profits from these (less-risky) drugs that add less social value, then the pharmaceutical firms' incentives may be diverging from the general public's benefit.

Developing a measure of the medical novelty of drugs is important for policymakers, R&D executives, and investors who want to judge the quality, risk and overall health of drug development pipelines. If these parties can separate the potentially breakthrough drugs from the "me-too" drugs, then they can more effectively evaluate their own initiatives and allocate resources strategically. This white paper describes the approaches that researchers have taken in measuring novelty in drug innovation, and then describes a new method involving chemical similarity measures.

Previous approaches to measuring drug novelty

Research on pharmaceutical productivity has measured drug innovation using a variety of methods. This section highlights some of the methods used in the the literature, but is not a comprehensive list of papers that use some measure of drug development novelty. Generally, these approaches suffer from some common weaknesses: 1) failing to separate the volume of new drugs from the novelty of new drugs, 2) inability to measure the relative novelty of drugs within a particular group or "class", or 3) relying on the (ex-post) outcomes of development projects to determine the (ex-ante) novelty of the drug.

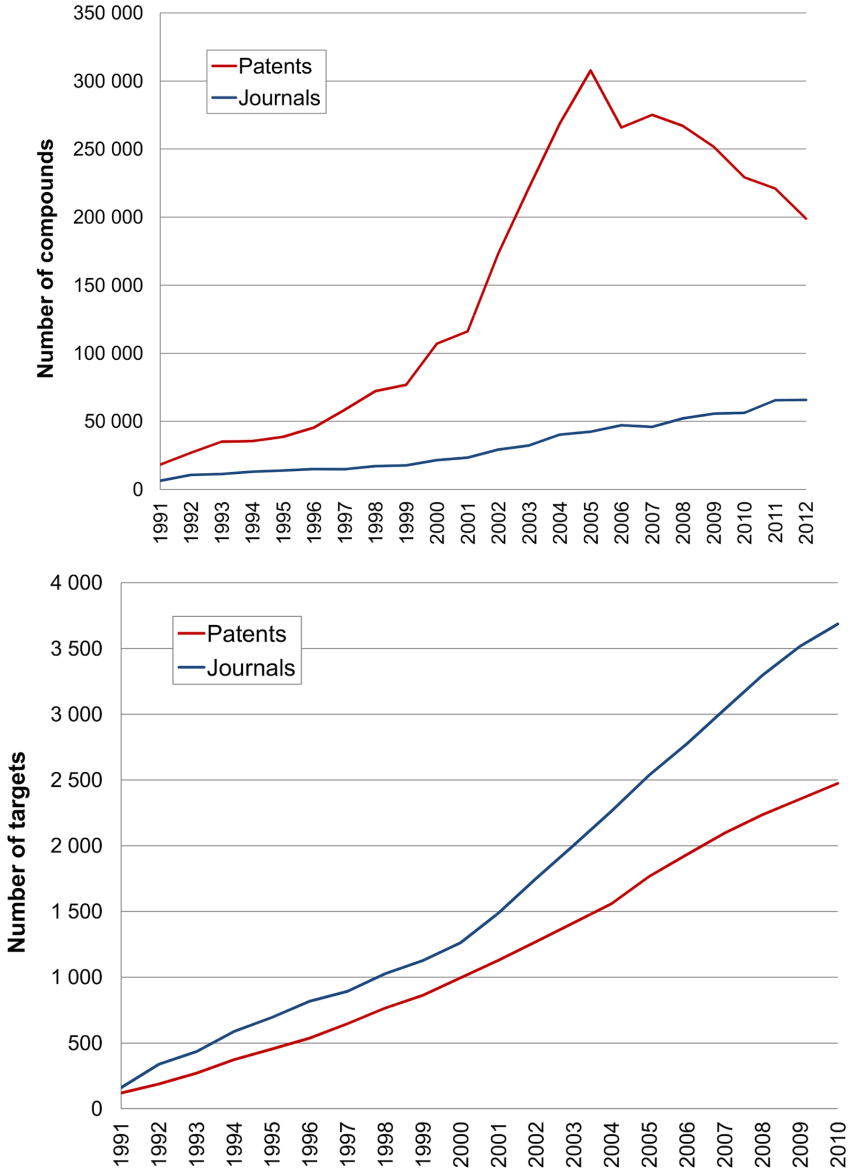
Counting new drugs



The most common way that the media reports on productivity in the pharmaceutical industry is by counting the number of new FDA drug approvals. While this rate is a reasonable measure of the industry's output, it does not necessarily reflect the medical value of this output, or the efforts that led to the output. Firms may, for instance, simply be developing "me-too" versions of previously approved drugs. Further, it is possible that new approvals and medical innovation may even be negatively related: if firms take large risks and focus on developing new types of compounds for unmet need, then one may even expect approval rates to decrease. By counting only approved drugs, we see only survivors without seeing the number and quality of development attempts.

An alternative to counting new approvals, is to count new drug compounds reported at earlier stages of the development process. Blume-Kohout & Sood (2013) used the number of new drugs entering

preclinical research and number entering clinical trials as tracked in Citeline's Pharmaprojects pipeline data as their measure of drug development productivity. Southan et al. (2013) tracked new drug compounds disclosed in both patent filings and journal articles over time. While these rates still do not capture the nature of these new compounds (e.g. the direction or novelty of innovation), they provide a better picture of output in the earlier stages of the drug development pipeline. Additionally, these authors also counted the number of new drug targets (e.g. biological pathways, channels, proteins, enzymes and receptors through which a drug acts) reported in patents and publications. This count gives a better sense of how much the drug R&D industry is exploring over time, as it captures how many new approaches are being studied. However, counting new targets does not reveal the distribution of drug projects across new and old targets, and it does not parse the novel approaches to hitting a drug target from the recycled versions.



Source: Southan et al. (2013)

Order of entry

Another notion of drug novelty that has appeared in the literature is the order of entry into a market or "class."[1] These type of measures are helpful because they are easy to observe and their meaning is intuitive (e.g. easy to understand why the first entrant is more novel than the fifth entrant). Most commonly, references to a drug's order of entry are limited to order of drug approval (DiMasi & Paquette, 2004; DiMasi & Faden, 2011; Lanthier et al., 2013). One can also analyze the time between first and follow on entrants as a proxy for the intensity of competition, or to evaluate pricing behavior.

One of the limitations of the basic order of entry approach is that it fails to distinguish between "me-too" drug additions, and new drugs that bring novel properties to an existing market or class. Lanthier et al. (2013) accounts for this issue by distinguishing between first in class, addition to class, and advance in class. However, this approach requires subjective judgements about a set of drugs' characteristics and advantages over one another.

Another issue with the studies categorizing the order or type of entry is that they are usually limited to drugs that have survived the clinical trial gauntlet and reached the market. The first drug approved for a therapeutic indication or drug class is not necessarily the first drug developed for that group, especially if drug development in that area required cumulative innovation and learning from prior failures. One approach to handle this limitation is to open up the order of entry characterizations to earlier development dates. Guedj & Scharfstein (2004) ranks drugs by their order of entry, in terms of early development dates, into each pharmacological groups in the Pharmaprojects database (which tracks drug development pipelines through time). While this approach covers a larger set of drugs (both the inputs and outputs of the drug development process), the downside of this approach is still the issue of separating copycat entrants from real follow-on innovation, and to do so in an objective manner.

Ex-post designations

A different indicator of drug novelty is whether a drug qualifies for official FDA designations such as the Breakthrough Therapies, Priority Review status, and the orphan drug designation. These designations have the advantage of encapsulating the type of novelty that consumers care about. For example, the FDA defines assigns the breakthrough status based when "clinical evidence indicates that the drug may demonstrate substantial improvement over existing therapies on one or more clinically significant endpoints, such as substantial treatment effects observed early in clinical development."

Dranove et al. (2014) provides an example of how these designations can be used in the systematic study of policy on drug novelty. The authors combine all three of the FDA designations mentioned above to characterize the "innovativeness" of a given drug. These approach is useful when examining

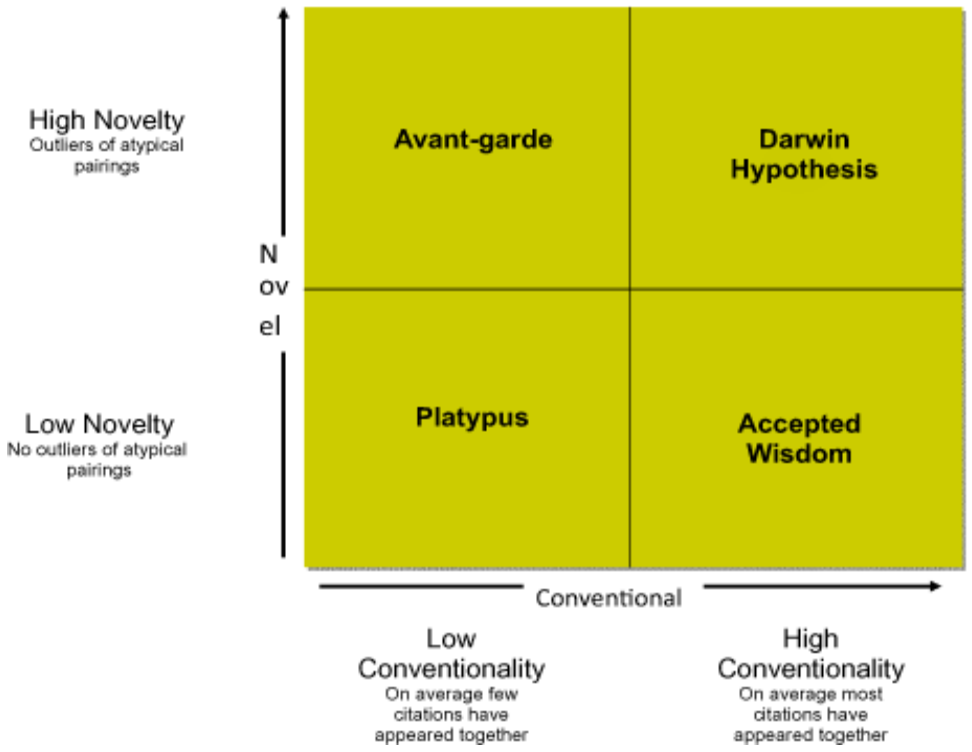
[1] A drug class is a frequently used but somewhat vaguely defined term. Usually a class is defined by the drug's therapeutic market (e.g. small cell lung cancer), biological target and actions (e.g. nucleoside reverse transcriptase inhibitor), or a combination of the two.

late-stage development projects, but cannot be applied to early-stage drugs who have not yet had the opportunity to gain one of these designations. Of all the novelty measures in the literature, these types of categories most closely align with the notion of innovativeness that leads to increases in patient welfare, but they have a blindspot when it comes to early-stage drug candidates. As a result, these measures tend to label drugs as novel only if they have been successful. This approach would miss novel projects that ultimately fail. Because bringing novel compounds to market may be riskier, this type of survivorship bias may leads us to conclude that firms that face many failures because they take more risks are in fact less innovative than firms that develop more established classes of drugs.

Content based measures of novelty and similarity

Outside of the medical sphere, researchers have also defined novelty as the recombination of existing ideas for useful applications (Fleming, 2001). One application of this approach is to define innovations by their representative documents (e.g. patents, academic publications), and use the combination of text, keywords, and citations in these documents to estimate the relative novelty of those documents.

Uzzi et al. (2013) estimate a measure of novelty using the distribution of unusual combinations of citations in academic journal articles. The authors look at all the pairs of citations in an article and evaluate how frequently each pair appears together in the entire universe of literature relative to each citation's overall appearances. From these calculations, the authors can describe the average novelty of co-citations in a given paper, and whether or not the paper has any atypical combinations. In a follow-up publication by the same author group, (Mukherjee et al., 2015), the authors describe different types of research outputs based on these measures:



A related approach is to use topic modeling to develop similarity measures between documents, and use the distribution of those similarity measures to calculate novelty. Nanda et al. (2013) use the text found in patent applications and calculate the similarity of that text to all the previous patent applications in same technology areas. Lin & Wilbur (2007) describe the PubMed related articles algorithm (PMRA) that they developed to measure the similarity of articles in the National Library of Medicine's database based on their titles, text and keywords. While the original purpose of PMRA algorithm was to facilitate search (and it has become a permanent feature of the PubMed website), one can also limit similarity comparisons to articles published prior to a focal article, and use their PMRA scores to assign novelty categories.

These content- and components-based approaches to measuring novelty are immediately applicable at the time of the idea or invention's first disclosure. Furthermore, they are attractive to researchers because they are quantifiable measures with both ordinal and cardinal values, such that they provide values that are easily comparable to one another. The downside to these approaches is that they are abstractions that are removed from tangible consumer welfare improvements, so they must be thoroughly validated as proxies for true novelty. The construction of these measures also requires documents or objects that are structured consistently, and contain sub-components whose presence and order is clearly observable to both human and algorithm. But the principles of these methods might still be applied to pharmaceutical innovations, where drug candidates and their performance are often represented in patents and publications.

Measuring Drug Development Novelty Through Compound Similarity

Ideally, a measure of drug novelty would have many of the features of the approaches in the literature, but without the limitations involving ex-post analysis and within group comparisons. The quantifiable measure would distinguish between different combinations of a drug's component parts and be applied objectively and immediately at the time of the innovation's first disclosure. Furthermore, this ideal measure's validity as a proxy for true medical product novelty would be verifiable across a number of dimensions, including prior measures of novelty and performance in clinical development and in the market.

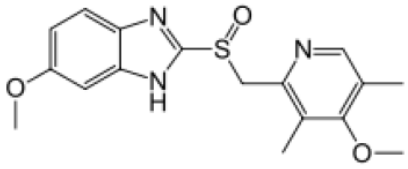
We propose a new measure of drug novelty that uses chemical informatics compound similarity measures to provide such an objective and quantifiable estimate of a drug's novelty. The approach compares the chemical structure of a new investigational drug compound to previously introduced drug candidates, and uses the distribution of their chemical similarities to determine the drug's overall novelty.

Before describing these calculations, it is informative to consider what information is encoded in a drug compound. At the time an organization begins developing a small-molecule drug compound,[2] the

[2] This approach is currently limited to small molecule drugs. Biologic drugs and vaccines are not included in this discussion.

developers know the compound's chemical structure (i.e. the sequence of bonded atoms and stable spatial arrangements), which can be represented graphically, through a molecular formula (e.g. $C_{17}H_{14}O_4S$) or through standardized chemical informatics identification systems such as the simplified molecular-input line-entry system (SMILES) or the International chemical Identifier (InChI) system. The developers usually also know what disease they plan to treat with the drug, and they (sometimes) know what biological target that the drug acts on in the body.

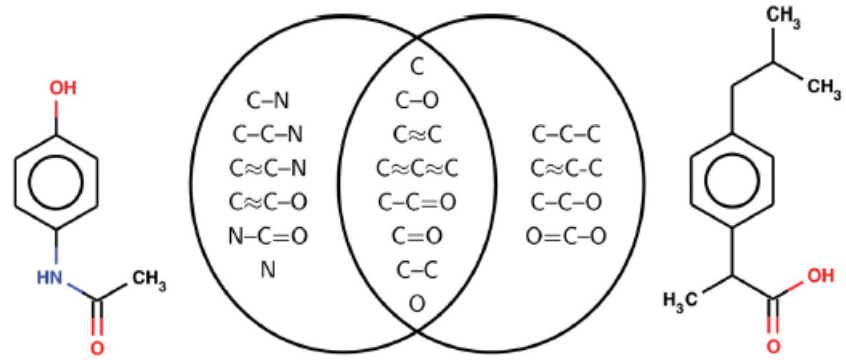
Example: Prilosec (omeprazole)



- SMILES code:
  CC1=CN=C(C(=C1OC)C)CS(=O)C2=NC3=C(N2)C=C(C=C3)OC
  - Standardized representation of molecular structure
- Target-Action: *Potassium-transporting ATPase inhibitor*
  - Describes drug impact on biological function: what does it act on, how does it act?
- ICD-9: *Potassium-transporting ATPase inhibitor*
  - Describes drug indication: what condition is it intended to treat?

In chemical informatics research the "similar property principle" means that one can expect that similar compounds have similar properties and biological activities (Johnson and Maggiora, 1990). As a result drug similarity measures are useful for building chemical libraries for drug screening (Wawer et al., 2014), quantifying the quality and "drug-like" properties of a compound (Bickerton et al., 2012), and expanding medicinal chemistry techniques (Maggiora et al., 2013). The most prevalent chemical similarity measure is the pairwise tanimoto score. The tanimoto calculation takes a pair of chemical structures as inputs, breaks them down to their component fragments (known as the chemical "fingerprint"), and calculates the Jaccard coefficient (a.k.a. tanimoto distance) between two sets of fragments. To illustrate how this calculation is performed, an example comparing the structures of Tylenol (paracetamol) and Advil (ibuprofen) is displayed below.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$



Tylenol (paracetamol)

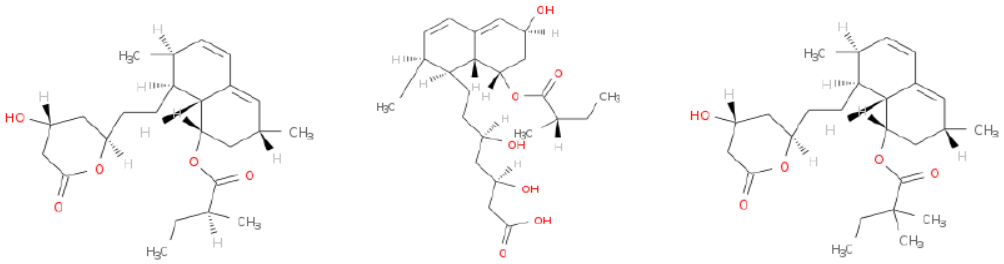
$\frac{8 \text{ in common}}{18 \text{ total}} = \mathbf{0.44}$

Advil (ibuprofen)

To turn pairwise drug similarity scores into novelty measures, we use the pairwise similarity scores between a focal drug and all drugs developed prior to that focal drug. Once we have the distribution of those scores, we can create a variety of novelty measures. For example,

1) maximum similarity of all prior drug candidates
2) maximum similarity of all prior candidates in the same indication or target group
3) maximum similarity of all prior candidates developed within the same firm
4) number of prior candidates with a similarity score x or greater

Tanimoto Examples: Cholesterol Reducing Drugs



Mevacor (Similarity Score=0.25)

Pravachol (Similarity Score =0.61)

Zocor (Similarity Score =0.82)

1. Mevacor (Lovostatin) is 1st FDA approved statin (Sep 1987): similarity to prior molecules is 0.25
2. Pravachol (Pravastatin) is 2nd (Oct 1991): pair-wise similarity to Mevacor is 0.61. Overall similarity to prior molecules is 0.61
3. Zocor (Simvastatin) is 3rd (Dec 1991): pair-wise similarity to Mevacor is 0.82. Pairwise to Pravachol is 0.52. Overall similarity to prior molecules is 0.82

Limitations

While attractive for many reasons, these compound similarity-based measures of drug novelty are not perfect. First, not all molecularly similar compounds function the same in the body. Sometimes a small tweak in the chemical structure can lead to drastic changes in the the body. For example, Thalidomide is a drug comprised of two mirror image molecules: one is a sedative, the other causes birth defects. Second, sometimes chemically dissimilar compounds have quite similar biological effects. Crestor and Lipitor have different structural profiles, but are often prescribed interchangeably by doctors.

While imperfect, chemical informatics studies have shown that Tanimoto similarity measures are useful for identifying drug qualities and novelty (O'Hagan et al., 2015; Baldi & Nasr, 2010; Bickerton et al., 2012). Furthermore, these measures can be validated by comparing to prior measures of drug novelty and using data from clinical studies and performance in the product market.

References:

Baldi, Pierre, and Ramzi Nasr. 2010. "When is Chemical Similarity Significant? The Statistical Distribution of Chemical Similarity Scores and Its Extreme Values." *Journal of Chemical Information and Modeling* 50(7): 1205-1222.

Bickerton, G. Richard, Gaia V. Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L. Hopkins. 2012. "Quantifying the chemical beauty of drugs." *Nat Chem* 4(2): 90-98.

Cha, Myoung, Bassel Rifai, and Pasha Sarraf. 2013. "Pharmaceutical forecasting: throwing darts?" *Nat Rev Drug Disco*v 12(10): 737-738.

DiMasi, Joseph A., and Laura B. Faden. 2011. "Competitiveness in follow-on drug R&D: a race or imitation?" *Nat Rev Drug Discov* 10(1): 23-27.

DiMasi, Joseph A., and Cherie Paquette. 2004. "The economics of follow-on drug research and development." *PharmacoEconomics* 22(2): 1-14.

Dranove, David, Craig Garthwaite, and Manuel Hermosilla. 2015. "Pharmaceutical Profits and the Social Value of Innovation." NBER Working Paper.

Fleming, Lee. 2001. "Recombinant Uncertainty in Technological Search." *Management Science* 47(1): 117-132.

Fleming, Lee, Santiago Mingo, and David Chen. 2007. "Collaborative Brokerage, Generative Creativity, and Creative Success." *Administrative Science Quarterly* 52(3): 443-475.

Gerald, Maggiora, Vogt Martin, Stumpfe Dagmar, and Bajorath Jürgen. 2014. "Molecular Similarity in Medicinal Chemistry." *Journal of Medicinal Chemistry* 57(8): 3186-3204.

Guedj, Ilan, and David Scharfstein. 2004. "Organizational Scope and Investment: Evidence from the Drug Development Strategies and Performance of Biopharmaceutical Firms." NBER Working Paper.

Lanthier, Michael, Kathleen L. Miller, Clark Nardinelli, and Janet Woodcock. 2013. "An Improved Approach To Measuring Drug Innovation Finds Steady Rates Of First-In-Class Pharmaceuticals, 1987–2011." *Health Affair*s 32(8): 1433-1439.

Lin, Jimmy, and W. John Wilbur. 2007. "PubMed related articles: a probabilistic topic-based model for content similarity." *BMC Bioinformatics* 8(1): 1-14.

Mukherjee, Satyam, Brian Uzzi, Ben Jones, and Michael Stringer. 2016. "A New Method for Identifying Recombinations of Existing Knowledge Associated with High-Impact Innovation." *Journal of Product Innovation Management* 33(2): 224-236.

Nanda, Ramana, Ken Younge, and Lee Fleming. 2013. "Innovation and Entrepreneurship in Renewable Energy." In Adam Jaffe, and Benjamin Jones (Eds.), *The Changing Frontier: Rethinking Science and Innovation Policy*, University of Chicago Press.

O'Hagan, Steve, Neil Swainston, Julia Handl, and Douglas B. Kell. 2015. "A 'rule of 0.5' for the metabolite-likeness of approved pharmaceutical drugs." Metabolomics 11(2): 323-339.

Southan, Christopher, Peter Varkonyi, Kiran Boppana, Sarma A. R. P. Jagarlapudi, and Sorel Muresan. 2013. "Tracking 20 Years of Compound-to-Target Output from Literature and Patents." *PLoS ONE* 8(10): e77142.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. "Atypical Combinations and Scientific Impact." *Science* 342(6157): 468-472.

Wawer, Mathias J., Kejie Li, Sigrun M. Gustafsdottir, Vebjorn Ljosa, Nicole E. Bodycombe, Melissa A. Marton, Katherine L. Sokolnicki, Mark-Anthony Bray, Melissa M. Kemp, Ellen Winchester, Bradley Taylor, George B. Grant, C. Suk-Yee Hon, Jeremy R. Duvall, J. Anthony Wilson, Joshua A. Bittker, Vlado Dančík, Rajiv Narayan, Aravind Subramanian, Wendy Winckler, Todd R. Golub, Anne E. Carpenter, Alykhan F. Shamji, Stuart L. Schreiber, and Paul A. Clemons. 2014. "Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling." *Proceedings of the National Academy of Sciences* 111(30): 10911-10916.