Computational drug repositioning based on side-effects mined from social media

Timothy Nugent, Vassilis Plachouras, Jochen L Leidner

Drug repositioning methods attempt to identify novel therapeutic indications for marketed drugs. Strategies include the use of side-effects to assign new disease indications, based on the premise that both therapeutic effects and side-effects are measurable physiological changes resulting from drug intervention. Drugs with similar side-effects might share a common mechanism of action linking side-effects with disease treatment, or may serve as a treatment by "rescuing" a disease phenotype on the basis of their side-effects; therefore it may be possible to infer new indications based on the similarity of side-effect profiles. While existing methods leverage side-effect data from clinical studies and drug labels, evidence suggests this information is often incomplete due to under-reporting. Here, we describe a novel computational method that uses side-effect data mined from social media to generate a sparse undirected graphical model using inverse covariance estimation with l_1-norm regularization. Results show that known indications are well recovered while current trial indications can also be identified, suggesting that sparse graphical models generated using side-effect data mined from social media may be useful for computational drug repositioning.

Computational drug repositioning based on side-effects mined from social media

Timothy Nugent, Vassilis Plachouras, Jochen L. Leidner

Thomson Reuters, Corporate Research and Development, 1 Mark Square, London, EC2A 4EG, United Kingdom

* tim.nugent@thomsonreuters.com

Abstract

Drug repositioning methods attempt to identify novel therapeutic indications for marketed drugs. Strategies include the use of side-effects to assign new disease indications, based on the premise that both therapeutic effects and side-effects are measurable physiological changes resulting from drug intervention. Drugs with similar side-effects might share a common mechanism of action linking side-effects with disease treatment, or may serve as a treatment by "rescuing" a disease phenotype on the basis of their side-effects; therefore it may be possible to infer new indications based on the similarity of side-effect profiles. While existing methods leverage side-effect data from clinical studies and drug labels, evidence suggests this information is often incomplete due to under-reporting. Here, we describe a novel computational method that uses side-effect data mined from social media to generate a sparse undirected graphical model using inverse covariance estimation with ℓ_1-norm regularization. Results show that known indications are well recovered while current trial indications can also be identified, suggesting that sparse graphical models generated using side-effect data mined from social media may be useful for computational drug repositioning.

Introduction

Drug repositioning is the process of identifying novel therapeutic indications for marketed drugs. Compared to traditional drug development, repositioned drugs have the advantage of decreased development time and costs given that significant pharmacokinetic, toxicology and safety data will have already been accumulated, drastically reducing the risk of attrition during clinical trials. In addition to marketed drugs, it is estimated that drug libraries may contain upwards of 2000 failed drugs that have the potential to be repositioned, with this number increasing at a rate of 150-200 compounds per year [1]. Repositioning of marketed or failed drugs has opened up new sources of revenue for pharmaceutical companies with estimates suggesting the market could generate multi-billion dollar annual sales in coming years [2,3]. While many of the current successes of drug repositioning have come about through serendipitous clinical observations, systematic data-driven approaches are now showing increasing promise given their ability to generate repositioning hypotheses for multiple drugs and diseases simultaneously using a wide range of data sources, while also incorporating prioritisation information to further accelerate development time [4]. Existing computational repositioning strategies generally use similar approaches but attempt to link different concepts. They include

the use of transcriptomics methods which compare drug response gene-expression with disease gene-expression signatures [5–9], genetics-based methods which connect a known drug target with a genetically associated phenotype [10–14], network-based methods which link drugs or diseases in a network based on shared features [15–19], and methods that use side-effect similarity to infer novel indications [20–27].

Drug side-effects can be attributed to a number of molecular interactions including on or off-target binding, drug-drug interactions [28, 29], dose-dependent pharmacokinetics, metabolic activities, downstream pathway perturbations, aggregation effects, and irreversible target binding [20, 30]. While side-effects are considered the unintended consequence of drug intervention, they can provide valuable insight into the physiological changes caused by the drug that are difficult to predict using pre-clinical or animal models. This relationship between drugs and side-effects has been exploited and used to identify shared target proteins between chemically dissimilar drugs, allowing new indications to be inferred based on the similarity of side-effect profiles [20]. One rationale behind this and related approaches is that drugs sharing a significant number of side-effects might share a common mechanism of action linking side-effects with disease treatment - side-effects essentially become a phenotypic biomarker for a particular disease [21, 25]. Repositioned drugs can also be said to "rescue" a disease phenotype, on the basis of their side-effects; for example, drugs which cause hair growth as a side-effect can potentially be repositioned for the treatment of hair loss, while drugs which cause hypotension as a side-effect can be used to treat hypertension [21]. Examples of drugs successfully repositioned based on phenotypic rescue that have made it to market include exenatide, which was shown to cause significant weight loss as a side-effect of type 2 diabetes treatment, leading to a trial of its therapeutic effect in non-diabetic obese subjects [31, 32], minoxidil which was originally developed for hypertension but found to cause hair growth as a side-effect, leading to its repositioning for the treatment of hair loss and androgenetic alopecia [33, 34], and, perhaps most famously, sildenafil citrate which was repositioned while being studied for the primary indication of angina to the treatment of erectile dysfunction [35].

Existing repositioning methods based on side-effects, such as the work of Campillos *et al.* [20] and Yang and Agarwal [21], have used data from the SIDER database [36], which contains side-effect data extracted from drug labels, largely collected from clinical trials during the pre-marketing phase of drug development. Other resources include *Meyler's Side Effects of Drugs* [37], which is updated annually in the *Side Effects of Drugs Annual* [38], and the Drugs@FDA database [39], while pharmacovigilance authorities attempt to detect, assess and monitor reported drug side-effects post-market. Despite regular updates to these resources and voluntary reporting systems, there is evidence to suggest that side-effects are substantially under-reported, with some estimates indicating that up to 86% of adverse drug reactions go unreported for reasons that include lack of incentives, indifference, complacency, workload and lack of training among healthcare professionals [40–43]. Side-effects reported from clinical trials also have limitations due to constraints on scale and time, as well as pharmacogenomic effects [44]. A number of cancer drug studies have also observed that women are often significantly under-represented in clinical trials, making it difficult to study the efficacy, dosing and side-effects of treatments which can work differently in women and men; similar problems of under-representation also affect paediatrics, as many drugs are only ever tested on adults [45].

Recently, efforts to mine user-generated content and social media for public-health issues and side-effects have shown promising performance, demonstrating correlations

between the frequency of side-effects extracted from unlabelled data and the frequency of documented adverse drug reactions [46]. Despite this success, a number of significant natural language processing challenges remain. These include dealing with idiomatic expressions, linguistic variability of expression and creativity, ambiguous terminology, spelling errors, word shortenings, and distinguishing between the symptoms that a drug is treating and the side-effects it causes. Some of the solutions proposed to deal with these issues include the use of specialist lexicons, appropriate use of semantic analysis, and improvements to approximate string matching, modeling of spelling errors, and contextual analysis surrounding the mentions of side-effects [46, 47], while maintaining a list of symptoms for which a drug is prescribed can help to eliminate them from the list of side-effects identified [48]. Although much of the focus has explored the use of online forums where users discuss their experience with pharmaceutical drugs and report side-effects [49], the growing popularity of Twitter [50], which at the time of writing has over 300 million active monthly users, provides a novel resource upon which to perform large-scale mining of reported drug side-effects in near real-time from the 500 millions tweets posted daily [51]. While only a small fraction of these daily tweets are related to health issues, the sheer volume of data available presents an opportunity to bridge the gap left by conventional side-effects reporting strategies. Over time, the accumulation of side-effect data from social media may become comparable or even exceed the volume of traditional resources, and at the very least should be sufficient to augment existing databases. Additionally, the cost of running such a system continuously is relatively cheap compared to existing pharmacovigilance monitoring, presenting a compelling economic argument supporting the use of social media for such purposes. Furthermore, the issues related to under-representation described above may be addressed.

Freifeld *et al.* [52] presented a comparison study between drug side-effects found on Twitter and adverse events reported in the FDA Adverse Event Reporting System (FAERS). Starting with 6.9 million tweets, they used a set of 23 drug names and a list of symptoms to reduce that data to a subset of 60,000 tweets. After manual examination, there were 4,401 tweets identified as mentioning a side-effect, with a Spearman rank correlation found to be 0.75. Nikfarjam *et al.* [53] introduce a method based on Conditional Random Fields (CRF) to tag mentions of drug side-effects in social media posts from Twitter or the online health community DailyStrength. They use features based on the context of tokens, a lexicon of adverse drug reactions, Part-Of-Speech (POS) tags and a feature indicating whether a token is negated or not. They also used embedding clusters learned with Word2Vec [54]. They reported an F1 score of 82.1% for data from DailyStrength and 72.1% for Twitter data. Sarker and Gonzalez [55] developed classifiers to detect side-effects using training data from multiple sources, including tweets [56], DailyStrength, and a corpus of adverse drug events obtained from medical case reports. They reported an F1 score of 59.7% when training a Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel on all three datasets. Recently, Karimi *et al.* [57] presented a survey of the field of surveillance for adverse drug events with automatic text and data mining techniques.

In this study, we describe a drug repositioning methodology that uses side-effect data mined from social media to infer novel indications for marketed drugs. We use data from a pharmacovigilance system for mining Twitter for drug side-effects [58]. The system uses a set of cascading filters to eliminate large quantities of irrelevant messages and identify the most relevant data for further processing, before applying a SVM classifier to identify tweets that mention suspected adverse drug reactions. Using this data we apply sparse inverse covariance estimation to construct an undirected graphical model, which offers a way to describe the relationship between all drug pairs [59–61]. This is

achieved by solving a maximum likelihood problem using ℓ_1-norm regularization to make the resulting graph as sparse as possible, in order to generate the simplest graphical model which fully explains the data. Results from testing the method on known and proposed trial indication recovery suggest that side-effect data mined from social media in combination with a regularized sparse graphical model can be used for systematic drug repositioning.

Methods

Mining Twitter for drug side-effects

We used the SoMeDoSEs pharmacovigilance system [58] to extract reports of drug side-effects from Twitter over a 6 month period between January and June 2014. SoMeDoSEs works by first applying topic and volume filters to identify Tweets that contain keywords relating to drugs, are written in English, are not re-tweets and do not contain a hyperlink to a web page, since these posts are typically commercial offerings. Side-effect were then mapped to an entry in the FDA Adverse Event Reporting System. Tweets that pass these filters are then classified by a linear SVM to distinguish those that mention a drug side-effect from those that do not. The SVM classifier uses a number of natural language features including unigrams and bigrams, part-of-speech tags, sentiment scores, text surface features, and matches to gazetteers related to human body parts, side-effect synonyms, side-effect symptoms, causality indicators, clinical trials, medical professional roles, side effect-triggers and drugs.

For each gazetteer, three features were created: a binary feature, which is set to 1 if a tweet contains at least one sequence of tokens matching an entry from the gazetteer, the number of tokens matching entries from the gazetteer, and the fraction of characters in tokens matching entries from the gazetteer. For side-effect synonyms we used the Consumer Health Vocabulary (CHV) [62], which maps phrases to Unified Medical Language System concept universal identifiers (CUI) and partially addresses the issue of misspellings and informal language used to discuss medical issues in tweets. The matched CUIs were also used as additional features.

To develop the system, 10,000 tweets which passed the topic and volume filters were manually annotated as mentioning a side-effect or not. Using a split of 8,000 tweets for training, 1,000 for development, and 1,000 for testing, the SVM classifier that used all the features achieved a precision of 55.0%, recall of 66.9%, and F1 score of 60.4% when evaluated using the 1,000 test tweets. This is statistically significantly higher than the results achieved by a linear SVM classifier using only unigrams and bigrams as features (precision of 56.0%, recall of 54.0% and F1 score of 54.9%). One of the sources of false negatives was the use of colloquial and indirect expressions by Twitter users to express that they have experienced a side-effect. We also observed that a number of false positives discuss the efficacy of drugs rather than side-effects.

Twitter data

Over the 6 month period, SoMeDoSEs typically identified $\sim$700 tweets per day that mentioned a drug side-effect, resulting in a data set of 620 unique drugs and 2196 unique side-effects from 108,009 tweets, once drug synonyms had been resolved to a common name using World Drug Index [63] and drugs with only a single side-effect were excluded. We were also careful to remove indications that were falsely identified as side-effects

using drug indications from Cortellis Clinical Trials Intelligence [64]. We used this data to construct a 2196 row by 620 column matrix of binary variables X, where $x \in \{0,1\}$, indicating whether each drug was reported to cause each side-effect in the Twitter data set.

Calculating the sample covariance matrix

Using this data, we are able to form the sample covariance matrix S for binary variables as follows [65], such that element $S_{i,j}$ gives the covariance of drug i with drug j:

$$
\begin{aligned}
S_{i,j} &= \frac{1}{n-1} \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \\
&= \frac{1}{n-1} \sum_{k=1}^{n} x_{ki} x_{kj} - \bar{x}_i \bar{x}_j
\end{aligned}
\tag{1}
$$

where $\bar{x}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ki}$ and x_{ki} is the k-th observation (side-effect) of variable (drug) X_i. It can be shown than the average product of two binary variables is equal to their observed joint probabilities such that:

$$
\frac{1}{n-1} \sum_{k=1}^{n} x_{ki} x_{kj} = P(X_j = 1 | X_i = 1)
\tag{2}
$$

where $P(X_j = 1 | X_i = 1)$ refers to the conditional probability that variable X_j equals one given that variable X_i equals one. Similarly, the product of the means of two binary variables is equal to the expected probability that both variables are equal to one, under the assumption of statistical independence:

$$
\bar{x}_i \bar{x}_j = P(X_i = 1) P(X_j = 1)
\tag{3}
$$

Consequently, the covariance of two binary variables is equal to the difference between the observed joint probability and the expected joint probability:

$$
S_{i,j} = P(X_j = 1 | X_i = 1) - P(X_i = 1) P(X_j = 1)
\tag{4}
$$

Our objective is to find the precision or concentration matrix θ by inverting the sample covariance matrix S. Using θ, we can obtain the matrix of partial correlation coefficients ρ for all pairs of variables as follows:

$$
\rho_{i,j} = -\frac{\theta_{i,j}}{\sqrt{\theta_{i,i}\theta_{j,j}}}
\tag{5}
$$

The partial correlation between two variables X and Y given a third, Z, can be defined as the correlation between the residuals R_x and R_y after performing least-squares regression of X with Z and Y with Z, respectively. This value, denoted $\rho_{x,y|z}$, provides a measure

5/22

of the correlation between two variables when conditioned on the third, with a value of zero implying conditional independence if the input data distribution is multivariate Gaussian. The partial correlation matrix ρ, however, gives the correlations between all pairs of variables conditioning on all other variables. Off-diagonal elements in ρ that are significantly different from zero will therefore be indicative of pairs of drugs that show unique covariance between their side-effect profiles when taking into account (*i.e.* removing) the variance of side-effects profiles amongst all the other drugs.

Shrinkage estimation

For the sample covariance matrix to be easily invertible, two desirable characteristics are that it should be positive definite, *i.e.* all eigenvalues should be distinct from zero, and well-conditioned, *i.e.* the ratio of its maximum and minimum singular value should not be too large. This can be particularly problematic when the sample size is small and the number of variables is large (n < p) and estimates of the covariance matrix become singular. To ensure these characteristics, and speed up convergence of the inversion, we condition the sample covariance matrix by shrinking towards an improved covariance estimator T, a process which tends to pull the most extreme coefficients towards more central values thereby systematically reducing estimation error [66], using a linear shrinkage approach to combine the estimator and sample matrix in a weighted average:

$$S' = \alpha T + (1 - \alpha)S \qquad (6)$$

where $\alpha \in \{0, 1\}$ denotes the analytically determined shrinkage intensity. We apply the approach of Schäfer and Strimmer, which uses a distribution-free, diagonal, unequal variance model which shrinks off-diagonal elements to zero but leaves diagonal entries intact, *i.e.* it does not shrink the variances [67]. Shrinkage is actually applied to the correlations rather than the covariances, which has two distinct advantages: the off-diagonal elements determining the shrinkage intensity are all on the same scale, while the partial correlations derived from the resulting covariance estimator are independent of scale.

Graphical lasso for sparse inverse covariance estimation

A useful output from the covariance matrix inversion is a sparse ρ matrix containing many zero elements, since, intuitively, we know that relatively few drug pairs will share a common mechanism of action, so removing any spurious correlations is desirable and results in a more parsimonious relationship model, while the non-zero elements will typically reflect the correct positive correlations in the true inverse covariance matrix more accurately [68]. However, elements of ρ are unlikely to be zero unless many elements of the sample covariance matrix are zero. The graphical lasso [60, 61, 69] provides a way to induce zero partial correlations in ρ by penalizing the maximum likelihood estimate of the inverse covariance matrix using an ℓ_1-norm penalty function. The estimate can be found by maximizing the following log-likelihood using the block coordinate descent approach described by Friedman *et al.* [60]:

$$\log \det \theta - \operatorname{tr}(S'\theta) - \lambda\|\theta\|_1 \qquad (7)$$

Here, the first term is the Gaussian log-likelihood of the data, tr denotes the trace

6/22

operator and $\|\theta\|_1$ is the ℓ_1-norm - the sum of the absolute values of the elements of θ, weighted by the non-negative tuning paramater λ. The specific use of the ℓ_1-norm penalty has the desirable effect of setting elements in θ to zero, resulting in a sparse matrix, while the parameter λ effectively controls the sparsity of the solution. This contrasts with the use of an ℓ_2-norm penalty which will shrink elements but will never reduce them to zero. While this graphical lasso formulation is based on the assumption that the input data distribution is multivariate Gaussian, Banerjee *et al.* showed that the dual optimization solution also applies to binary data, as is the case in our application [61].

It has been noted that the graphical lasso produces an approximation of θ that is not symmetric, so we update it as follows [70]:

$$\theta \leftarrow \frac{(\theta + \theta^T)}{2} \tag{8}$$

The matrix ρ is then calculated according to Equation 5, before repositioning predictions for drug i are determined by ranking all other drugs according to their absolute values in ρ_i and assigning their indications to drug i.

Results and Discussion

Recovering known indications

To evaluate our method we have attempted to predict repositioning targets for indications that are already known. If, by exploiting hindsight, we can recover these, then our method should provide a viable strategy with which to augment existing approaches that adopt an integrated approach to drug repositioning [19]. Figure 1a shows the performance of the method at identifying co-indicated drugs at a range of λ values, resulting in different sparsity levels in the resulting ρ matrix. We measured the percentage at which a co-indicated drug was ranked amongst the top 5, 10, 15, 20 and 25 predictions for the target drug, respectively. Of the 620 drugs in our data set, 595 had a primary indication listed in Cortellis Clinical Trials Intelligence, with the majority of the remainder being made up of dietary supplements (*e.g.* methylsulfonylmethane) or plant extracts (*e.g.* *Agaricus brasiliensis* extract) which have no approved therapeutic effect. Rather than removing these from the data set, they were left in as they may contribute to the partial correlation between pairs of drugs that do have approved indications.

Results indiciate that the method achieves its best performance with a λ value of 10^{-9} where 42.41% (243/595) of targets have a co-indicated drug returned amongst the top 5 ranked predictions (Figure 1a). This value compares favourably with both a strategy in which drug ranking is randomized (13.54%, standard error ± 0.65), and another in which drugs are ranked according to the Jaccard index (28.75%). In Ye *et al.* [27], a related approach is used to construct a repositioning network based on side-effects extracted from the SIDER database, *Meyler's Side Effects of Drugs, Side Effects of Drugs Annual*, and the Drugs@FDA database [36–39], also using the Jaccard index as the measure of drug-drug similarity. Here, they report an equivilent value of 32.77% of drugs having their indication correctly predicted amongst the top 5 results. While data sets and underlying statistical models clearly differ, these results taken together suggest that the use of side-effect data mined from social media can certainly offer comparable performance to methods using side-effect data extracted from more conventional resources, while the use of a global statistical model such as the graphical
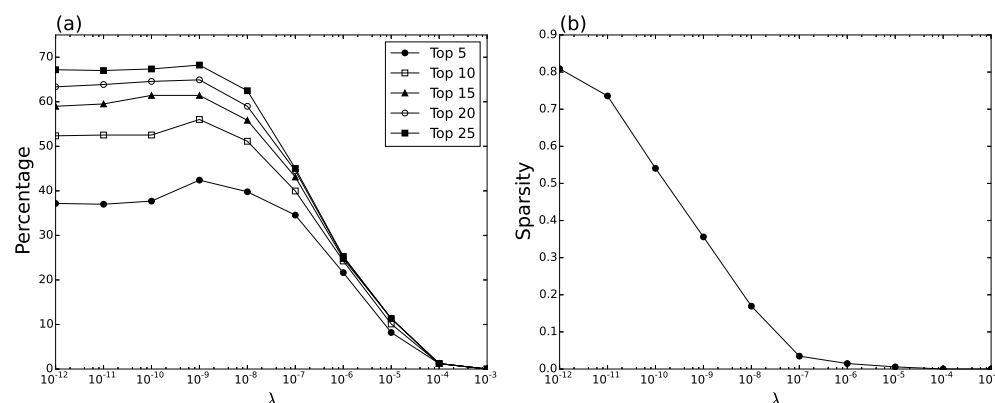
Figure 1. Recovery of known indications (a) Percentage at which a co-indicated drug is returned amongst the top 5, 10, 15, 20 and 25 ranked predictions for a given target, at different λ values - the parameter that weights the ℓ_1-norm penalty in the graphical lasso (Equation 7) (b) Sparsity of ρ matrix at different λ values, *i.e.* the number of non-zero elements in the upper triangle divided by $(n^2 - n)/2$.

lasso does result in improved performance compared to a pairwise similarity coefficient such as the Jaccard index.

To further investigate the influence of the provenance of the data, we mapped our data set of drugs to ChEMBL identifiers [71,72] which we then used to query SIDER for side-effects extracted from drug labels. This resulted in a reduced data set of 229 drugs, in part due to the absence of many combination drugs from SIDER (e.g. the antidepressant Symbyax which contains olanzapine and fluoxetine). Using the same protocol described above, best performance of 53.67% (117/229) was achieved with a slightly higher λ value of 10^{-6}. Best performance on the same data set using side-effects derived from Twitter was 38.43% (88/229), again using a λ value of 10^{-9}, while the randomized strategy achieved 12.05% (standard error ± 1.14), indicating that the use of higher quality side-effect data from SIDER allows the model to achieve better performance than is possible using Twitter data. Perhaps more interestingly, combining the correct predictions between the two datasources reveals that 30 are unique to the Twitter model, 59 are unique to the SIDER model, with 58 shared, supporting the use side-effect data mined from social media to augment conventional resources.

We also investigated whether our results were biased by the over-representation of particular drug classes within our data set. Using Using Cortellis Clinical Trials Intelligence, we were able to identify the broad class for 479 of the drugs (77.26%) in our data set. The five largest classes were benzodiazepine receptor agonists (3/14 drugs returned amongst the top 5 ranked predictions), analgesics (6/12), H_1-antihistamines (8/11), cyclooxygenase inhibitors (9/11), and anti-cancer (2/11). This indicates that the over-representation of H_1-antihistamines and cyclooxygenase inhibitors did result in a bias, and to a lesser extent analgesics, but that the overall effect of these five classes was more subtle (28/59 returned amongst the top 5 ranked predictions, 47.46%).

The best performance of our approach at the top 5 level is achieved when the resulting ρ matrix has a sparsity of 35.59% (Figure 1b and Figure 2) which both justifies the use of the ℓ_1-norm penalized graphical lasso, and generates a graphical model with approximately a third of the parameters of a fully dense matrix, while the comparable performance at λ values between 10^{-12} and 10^{-7} also indicates a degree of robustness to

8/22
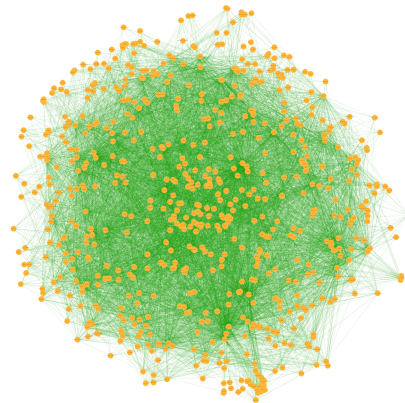
Figure 2. The overall layout of the side-effect network. Drugs are yellow, connecting edges are green. The layout is performed using a relative entropy optimization-based method [73]. In total, there are 616 connected nodes, with each having an average of 267 neighbours. Painkillers such as paracetamol and ibuprofen have the highest number of connections (587 and 585, respectively), which corresponds to them having the largest number of unique side-effects (256 and 224) reported on Twitter. The strongest connection is between chondroitin and glucosamine (partial correlation coefficient (PCC) 0.628), both of which are dietary supplements used to treat osteoarthritis, closely followed by the antidepressant and anxiolytic agents phenelzine and tranylcypromine (PCC 0.614).

the choice of this parameter. Beyond the top 5 ranked predictions, results are encouraging as the majority of targets (56.02%) will have a co-indicated drug identified by considering only the top 10 predictions, suggesting the method is a feasible strategy for prioritisation of repositioning candidates.

Predicting proposed indications of compounds currently in clinical trials

While the previous section demonstrated our approach can effectively recover known indications, predictions after the fact are - while useful - best supported by more forward-looking evidence. In this section, we use clinical trial data to support our predictions where the ultimate success of our target drug is still unknown. Using Cortellis Clinical Trials Intelligence, we extracted drugs present in our Twitter data set that were currently undergoing clinical trials (ending after 2014) for a novel indication (*i.e.* for which they were not already indicated), resulting in a subset of 277 drugs currently in trials for 397 indications. Figure 3 shows the percentage at which a co-indicated drug was ranked amongst the top 5, 10, 15, 20 and 25 predictions for the target. Similar to the recovery of known indications, best performance when considering the top 5 ranked predictions was achieved with a λ value of 10^{-9}, resulting in 16.25% (45/277) of targets having a co-indicated drug, which again compares well to a randomized strategy (5.42%, standard error ±0.32) or a strategy using the Jaccard index (10.07%).

Recovery of proposed clinical trial indications is clearly more challenging than known indications, possibly reflecting the fact that a large proportion of drugs will fail during trials and therfore many of the 397 proposed indications analysed here will in time prove false, although the general trend in performance as the sparsity parameter λ is adjusted tends to mirror the recovery of known indications. Despite this, a number of interesting predictions with a diverse range of novel indications are made that are supported by
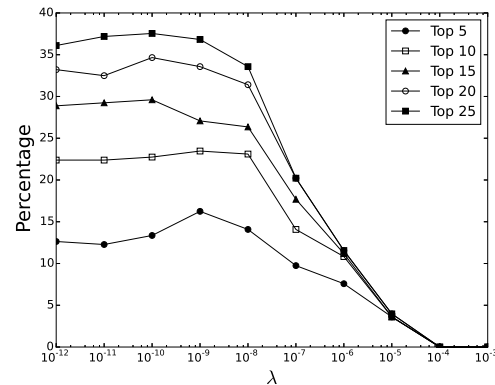
Figure 3. Recovery of proposed clinical trial indications. Percentage at which a co-indicated drug is returned amongst the top 5, 10, 15, 20 and 25 ranked predictions for a given target, at different λ values.

experimental and clinical evidence; a selection of 10 of the 45 drugs where the trial indication was correctly predicted are presented in Table 1. We further investigated three repositioning candidates with interesting pharmacology to understand their predicted results.

| Drug	Current indication	New indication	Evidence	PCC	Rank	ID	Title
ramelteon	insomnia	bipolar I disorder	ziprasidone	0.197	2	6991	Ramelteon for the Treatment of Insomnia and Mood Stability in Patients With Euthymic Bipolar Disorder
denosumab	osteoporosis	breast cancer	capecitabine	0.133	3	85503	Pilot Study to Evaluate the Impact of Denosumab on Disseminated Tumor Cells (DTC) in Patients With Early Stage Breast Cancer
meloxicam	inflammation	non-Hodgkin lymphoma	rituximab	0.131	1	176379	A Phase II Trial Using Meloxicam Plus Filgrastim in Patients With Multiple Myeloma and Non-Hodgkins Lymphoma
sulfasalazine	rheumatoid arthritis	diarrhea	loperamide	0.106	5	155516	Sulfasalazine in Preventing Acute Diarrhea in Patients With Cancer Who Are Undergoing Pelvic Radiation Therapy
pyridostigmine	myasthenia gravis	cardiac failure	digitoxin	0.100	4	190789	Safety Study of Pyridostigmine in Heart Failure
alprazolam	anxiety disorder	epilepsy	clonazepam	0.097	4	220920	Staccato Alprazolam and EEG Photoparoxysmal Response
oxytocin	Prader-Willi syndrome	schizophrenia	chlorpromazine	0.096	3	163871	Antipsychotic Effects of Oxytocin
interferon alfa	leukemia	thrombocythemia	hydroxyurea	0.094	3	73064	Pegylated Interferon Alfa-2a Salvage Therapy in High-risk Polycythemia Vera (PV) or Essential Thrombocythemia (ET)
etomidate	general anesthesia	depression	trazodone	0.091	5	157982	Comparison of Effects of Propofol and Etomidate on Rate Pressure Product and Oxygen Saturation in Patients Undergoing Electroconvulsive Therapy
guaifenesin	respiratory tract infections	rhinitis	ipratropium	0.090	5	110111	The Effect of Oral Guaifenesin on Pediatric Chronic Rhinitis: a Pilot Study

Table 1. Predicted indications for drugs currently in clinical trials. A selection of drugs which are currently in clinical trials for a new indication, and have a co-indicated drug ("Evidence") ranked amongst the top 5 predictions. "PCC" is the absolute partial correlation coefficient, "ID" is the Cortellis Clinical Trials Intelligence identifier. Average PCC scores for co-indicated drugs ranked amongst the top 5, 10, 15, 20 and 25 positions were 0.162, 0.0804, 0.0620, 0.0515, and 0.0468, respectively.
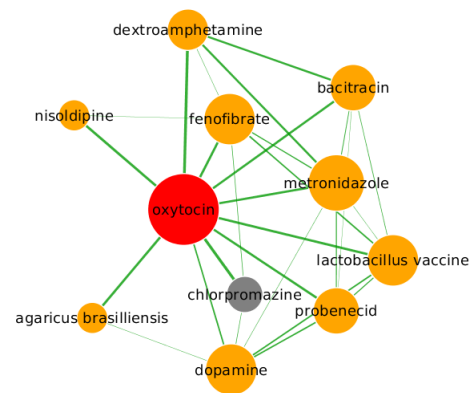
Figure 4. Predicted repositioning of oxytocin (red) for the treatment of schizophrenia based on its proximity to the schizophrenia drug chlorpromazine (grey). Drugs in the graph are sized according to their degree (number of edges), while the thickness of a connecting edge is proportional to the partial correlation coefficient between the two drugs. The graph layout is performed by Cytoscape [81] which applies a force-directed approach based on the partial correlation coefficient. Nodes are arranged so that edges are of more or less equal length and there are as few edge crossings as possible. For clarity, only the top ten drugs ranked by partial correlation coefficient are shown.

Oxytocin

Oxytocin is a nonapeptide hormone that acts primarily as a neuromodulator in the brain via the specific, high-affinity oxytocin receptor - a class I (Rhodopsin-like) G-protein-coupled receptor (GPCR) [74]. Currently, oxytocin is used for labor induction and the treatment of Prader-Willi syndrome, but there is compelling pre-clinical evidence to suggest that it may play a crucial role in the regulation of brain-mediated processes that are highly relevant to many neuropsychiatric disorders [75]. A number of animal studies have revealed that oxytocin has a positive effect as an antipsychotic [76,77], while human trials have revealed that intranasal oxytocin administered to highly symptomatic schizophrenia patients as an adjunct to their antipsychotic drugs improves positive and negative symptoms significantly more than placebo [78,79]. These therapeutic findings are supported by growing evidence of oxytocin's role in the manifestation of schizophrenia symptoms such as a recent study linking higher plasma oxytocin levels with increased pro-social behavior in schizophrenia patients and with less severe psychopathology in female patients [80]. The mechanisms underlying oxytocin's therapeutic effects on schizophrenia symptoms are poorly understood, but its ability to regulate mesolimbic dopamine pathways are thought to be responsible [75]. Here, our method predicts schizophrenia as a novel indication for oxytocin based on its proximity to chlorpromazine, which is currently used to treat schizophrenia (Figure 4). Chlorpromazine also modulates the dopamine pathway by acting as an antagonist of the dopamine receptor, another class I GPCR. Interestingly, the subgraph indicates that dopamine also has a high partial correlation coefficient with oxytocin, adding further support to the hypothesis that oxytocin, chlorpromazine and dopamine all act on the same pathway and therefore have similar side-effect profiles. Side-effects shared by oxytocin and chlorpromazine include hallucinations, excessive salivation and anxiety, while shivering, weight gain, abdominal pain, nausea, and constipation are common side-effects also shared by other drugs within the subgraph. Currently, larger scale clinical trials of intranasal oxytocin in schizophrenia are underway. If the early positive results hold up, it may signal the beginning of an new era in the treatment of schizophrenia, a field which has seen little progress in the development of novel efficacious treatments over recent years.
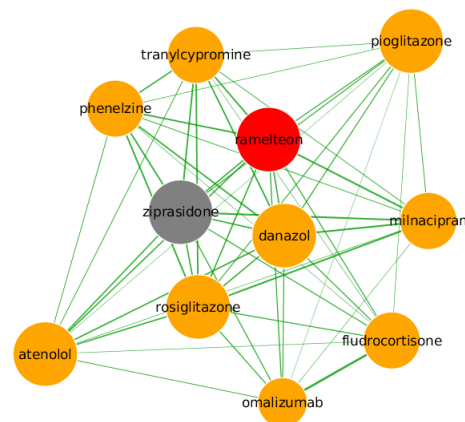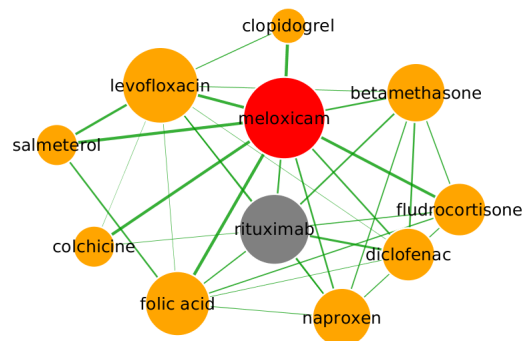
Figure 5. Predicted repositioning of ramelteon (red) for the treatment of bipolar I disorder based on its proximity to ziprasidone (grey). Along with ziprasidone, phenelzine, milnacipran and tranylcypromine are all used to treat mood disorders.

Ramelteon   392

Ramelteon, currently indicated for the treatment of insomnia, is predicted to be useful   393
for the treatment of bipolar depression (Figure 5). Ramelteon is the first in a new   394
class of sleep agents that selectively binds the MT_1 and MT_2 melatonin receptors   395
in the suprachiasmatic nucleus, with high affinity over the MT_3 receptor [82]. It is   396
believed that the activity of ramelteon at MT_1 and MT_2 receptors contributes to its   397
sleep-promoting properties, since these receptors are thought to play a crucial role in   398
the maintenance of the circadian rhythm underlying the normal sleep-wake cycle upon   399
binding of endogenous melatonin. Abnormalities in circadian rhythms are prominent   400
features of bipolar I disorder, with evidence suggesting that disrupted sleep-wake circadian   401
rhythms are associated with an increased risk of relapse in bipolar disorder [83]. As   402
bipolar patients tend to exhibit shorter and more variable circadian activity, it has been   403
proposed that normalisation of the circadian rhythm pattern may improve sleep and   404
consequently lead to a reduction in mood exacerbations. Melatonin receptor agonists such   405
as ramelteon may have a potential therapeutic effect in depression due to their ability   406
to resynchronize the suprachiasmatic nucleus [84]. In Figure 5, evidence supporting the   407
repositioning of ramelteon comes from ziprasidone, an atypical antipsychotic used to   408
treat bipolar I disorder and schizophrenia [85]. Ziprasidone is the second-ranked drug   409
by partial correlation coefficient; a number of other drugs used to treat mood disorders   410
can also be located in the immediate vicinity including phenelzine, a non-selective   411
and irreversible monoamine oxidase inhibitor (MAOI) used as an antidepressant and   412
anxiolytic, milnacipran, a serotonin–norepinephrine reuptake inhibitor used to treat   413
major depressive disorder, and tranylcypromine, another MAOI used as an antidepressant   414
and anxiolytic agent. The co-location of these drugs in the same region of the graph   415
suggests a degree of overlap in their respective mechanistic pathways, resulting in a   416
high degree of similarity between their side-effect profiles. Nodes in this subgraph also   417
have a relatively large degree indicating a tighter association than for other predictions,   418
with common shared side-effects including dry mouth, sexual dysfunction, migraine, and   419
orthostatic hypotension, while weight gain is shared between ramelteon and ziprasidone.   420

Meloxicam   421

Meloxicam, a nonsteroidal anti-inflammatory drug (NSAID) used to treat arthritis, is   422
predicted to be a repositioning candidate for the treatment of non-Hodgkin lymphoma,   423

Figure 6. Predicted repositioning of meloxicam (red) for the treatment of non-Hodgkin lymphoma based on its proximity to rituximab (grey).

via the mobilisation of autologous peripheral blood stem cells from bone marrow. By inhibiting cyclooxygenase 2, meloxicam is understood to inhibit generation of prostaglandin E_2, which is known to stimulate osteoblasts to release osteopontin, a protein which encourages bone resorption by osteoclasts [86,87]. By inhibiting prostaglandin E_2 and disrupting the production of osteopontin, meloxicam may encourage the departure of stem cells, which otherwise would be anchored to the bone marrow by osteopontin [88]. In Figure 6, rituximab, a B-cell depleting monoclonal antibody that is currently indicated for treatment of non-Hodgkin lymphoma, is the top ranked drug by partial correlation, which provides evidence for repositioning to this indication. Interestingly, depletion of B-cells by rituximab has recently been demonstrated to result in decreased bone resorption in patients with rheumatoid arthritis, possibly via a direct effect on both osteoblasts and osteoclasts [89,90], suggesting a common mechanism of action between meloxicam and rituximab. Further evidence is provided by the fifth-ranked drug clopidogrel, an antiplatelet agent used to inhibit blood clots in coronary artery disease, peripheral vascular disease, cerebrovascular disease, and to prevent myocardial infarction. Clopidogrel works by irreversibly inhibiting the adenosine diphosphate receptor P2Y12, which is known to increase osteoclast activity [91]. Similarly to the ramelteon subgraph, many drugs in the vicinity of meloxicam are used to treat inflammation including diclofenac, naproxen (both NSAIDs) and betamethasone, resulting in close association between these drugs, with shared side-effects in the subgraph including pain, cramping, flushing and fever, while swelling, indigestion, inflammation and skin rash are shared by meloxicam and rituximab.

While the side-effects shared within the subgraphs of our three examples are commonly associated with a large number of drugs, some of the side-effects shared by the three drug pairs such as hallucinations, excessive salivation and anxiety are somewhat less common. To investigate this relationship for the data set as a whole, we calculated log frequencies for all side-effects and compared these values against the normalized average rank of pairs where the side-effect was shared by both the query and target drug. If we assume that a higher ranking in our model indicates a higher likelihood of drugs sharing a protein target, this relationship demonstrates similar properties to the observations of Campillos *et al.* [20] in that there is a negative correlation between the rank and frequency of a side-effect. The correlation coefficient has a value of -0.045 which is significantly different from zero at the 0.001 level, although the linear relationship appears to break down where the frequency of the side-effect is lower than about 0.025.

Conclusions

In this study, we have used side-effect data mined from social media to generate a sparse graphical model, with nodes in the resulting graph representing drugs, and edges between them representing the similarity of their side-effect profiles. We demonstrated that known indications can be inferred based on the indications of neighbouring drugs in the network, with 42.41% of targets having their known indication identified amongst the top 5 ranked predictions, while 16.25% of drugs that are currently in a clinical trial have their proposed trial indication correctly identified. These results indicate that the volume and diversity of drug side-effects reported using social media is sufficient to be of use in side-effect-based drug repositioning, and this influence is likely to increase as the audience of platforms such as Twitter continues to see rapid growth. It may also help to address the problem of side-effect under-reporting. We also demonstrate that global statistical models such as the graphical lasso are well-suited to the analysis of large multivariate systems such as drug-drug networks. They offer significant advantages over conventional pairwise similarity methods by incorporating indirect relationships between all variables, while the use of the lasso penalty allows a sparse, parsimonious model to be generated with fewer spurious connections resulting in a simpler theory of relationships.

While our method shows encouraging results, it is more likely to play a role in drug repositioning as a component in an integrated approach. Whether this is achieved by combining reported side-effects with those mined from resources such as SIDER, or by using predictions as the inputs to a supervised learning algorithm, a consensus approach is likely to achieve higher performance by incorporating a range of different data sources in addition to drug side-effects, while also compensating for the weaknesses of any single method [19]. Limitations of our method largely stem from the underlying Twitter data [58]. Only a small fraction of daily tweets contain reports of drug side-effects, therefore restricting the number of drugs we are able to analyse. However, given that systems such as SoMeDoSEs are capable of continuously monitoring Twitter, the numbers of drugs and reported side-effects should steadily accumulate over time.

To address this, in the future it may be possible to extend monitoring of social media to include additional platforms. For example, Weibo is a Chinese microblogging site akin to Twitter, with over 600 million users as of 2013. Clearly, tools will have to be adapted to deal with multilingual data processing or translation issues, while differences in cultural attitudes to sharing medical information may present further challenges. Extensions to the statistical approach may also result in improved performance. Methods such as the joint graphical lasso allow the generation of a graphical model using data with observations belonging to distinct classes [92]. For example, two covariances matrices generated using data from Twitter and SIDER could be combined in this way, resulting in a single model that best represents both sources. An extension to the graphical lasso also allows the decomposition of the sample covariance graph into smaller connected components via a thresholding approach [93]. This leads not only to large performance gains, but significantly increases the scalability of the graphical lasso approach.

Another caveat to consider, common to many other repositioning strategies based on side-effect similarity, is that there is no evidence to suggest whether a repositioning candidate will be a better therapeutic than the drug from which the novel indication was inferred. While side-effects can provide useful information for inferring novel indications, they are in general undesirable and need to be balanced against any therapeutic benefits. Our model does not attempt to quantify efficacy or side-effect severity, but it might be possible to modify the natural language processing step during Twitter mining in order to capture comparative mentions of side-effects, since tweets discussing both the

15/22

therapeutic and side-effects of two related drugs are not uncommon. Incorporating this $_{510}$ information into our model may allow a more quantitative assessment of the trade-off $_{511}$ between therapeutic and side-effects to be made as part of the prediction. $_{512}$

Competing interests $_{513}$

The authors have declared that no competing interests exist. $_{514}$

References

1. Jarvis L. Teaching an old drug new tricks. Chemical and Engineering News. 2006 Feb;84(7):54–55.

2. Thomson Reuters. Knowledge-Based Drug Repositioning to Drive R&D Productivity. New York, NY, USA: The Thomson Reuters Corporation; 2012. Available from: `http://thomsonreuters.com/products/ip-science/04_038/drug-repositioning-cwp-en.pdf`.

3. Tobinick EL. The value of drug repositioning in the current pharmaceutical market. Drug News Perspect. 2009 Mar;22(2):119–125.

4. Hurle MR, Yang L, Xie Q, Rajpal DK, Sanseau P, Agarwal P. Computational drug repositioning: from data to therapeutics. Clin Pharmacol Ther. 2013 Apr;93(4):335–341.

5. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006 Sep;313(5795):1929–1935.

6. Hu G, Agarwal P. Human disease-drug network based on genomic expression profiles. PLoS ONE. 2009;4(8):e6536.

7. Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaokar P, Ferriero R, et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci USA. 2010 Aug;107(33):14621–14626.

8. Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A, et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med. 2011 Aug;3(96):96ra77.

9. Dudley JT, Sirota M, Shenoy M, Pai RK, Roedder S, Chiang AP, et al. Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. Sci Transl Med. 2011 Aug;3(96):96ra76.

10. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010 Dec;42(12):1118–1125.

11. Zhang J, Jiang K, Lv L, Wang H, Shen Z, Gao Z, et al. Use of genome-wide association studies for cancer research and drug repositioning. PLoS ONE. 2015;10(3):e0116477.

12. Wang ZY, Zhang HY. Rational drug repositioning by medical genetics. Nat Biotechnol. 2013 Dec;31(12):1080–1082.

13. Sanseau P, Agarwal P, Barnes MR, Pastinen T, Richards JB, Cardon LR, et al. Use of genome-wide association studies for drug repositioning. Nat Biotechnol. 2012 Apr;30(4):317–320.

14. Wang H, Gu Q, Wei J, Cao Z, Liu Q. Mining drug-disease relationships as a complement to medical genetics-based drug repositioning: Where a recommendation system meets genome-wide association studies. Clin Pharmacol Ther. 2015 May;97(5):451–454.

15. Krauthammer M, Kaufmann CA, Gilliam TC, Rzhetsky A. Molecular triangulation: bridging linkage and molecular-network information for identifying candidate genes in Alzheimer's disease. Proc Natl Acad Sci USA. 2004 Oct;101(42):15148–15153.

16. Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011 Jan;12(1):56–68.

17. Kohler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. Am J Hum Genet. 2008 Apr;82(4):949–958.

18. Vanunu O, Magger O, Ruppin E, Shlomi T, Sharan R. Associating genes and protein complexes with disease via network propagation. PLoS Comput Biol. 2010 Jan;6(1):e1000641.

19. Emig D, Ivliev A, Pustovalova O, Lancashire L, Bureeva S, Nikolsky Y, et al. Drug target prediction and repositioning using an integrated network-based approach. PLoS ONE. 2013;8(4):e60618.

20. Campillos M, Kuhn M, Gavin AC, Jensen LJ, Bork P. Drug target identification using side-effect similarity. Science. 2008 Jul;321(5886):263–266.

21. Yang L, Agarwal P. Systematic drug repositioning based on clinical side-effects. PLoS ONE. 2011;6(12):e28025.

22. Zhang P, Wang F, Hu J, Sorrentino R. Exploring the relationship between drug side-effects and therapeutic indications. AMIA Annu Symp Proc. 2013;2013:1568–1577.

23. Cheng F, Li W, Wu Z, Wang X, Zhang C, Li J, et al. Prediction of polypharmacological profiles of drugs by the integration of chemical, side effect, and therapeutic space. J Chem Inf Model. 2013 Apr;53(4):753–762.

24. Bisgin H, Liu Z, Kelly R, Fang H, Xu X, Tong W. Investigating drug repositioning opportunities in FDA drug labels through topic modeling. BMC Bioinformatics. 2012;13 Suppl 15:S6.

25. Duran-Frigola M, Aloy P. Recycling side-effects into clinical markers for drug repositioning. Genome Med. 2012;4(1):3.

26. Wang F, Zhang P, Cao N, Hu J, Sorrentino R. Exploring the associations between drug side-effects and therapeutic indications. J Biomed Inform. 2014 Oct;51:15–23.

27. Ye H, Liu Q, Wei J. Construction of drug network based on side effects and its application for drug repositioning. PLoS ONE. 2014;9(2):e87864.

28. Vilar S, Uriarte E, Santana L, Lorberbaum T, Hripcsak G, Friedman C, et al. Similarity-based modeling in large-scale prediction of drug-drug interactions. Nat Protoc. 2014 Sep;9(9):2147–2163.

29. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med. 2012 Mar;4(125):125ra31.

30. Xie L, Kinnings S, Xie L, Bourne P. Predicting the Polypharmacology of Drugs: Identifying New Uses through Chemoinformatics, Structural Informatics, and Molecular Modeling-Based Approaches. In: Barratt M, Frai D, editors. Drug Repositioning: Bringing New Life to Shelved Assets and Existing Drugs. Hoboken, NJ: John Wiley & Sons; 2012. .

31. Buse JB, Henry RR, Han J, Kim DD, Fineman MS, Baron AD. Effects of exenatide (exendin-4) on glycemic control over 30 weeks in sulfonylurea-treated patients with type 2 diabetes. Diabetes Care. 2004 Nov;27(11):2628–2635.

32. Ladenheim EE. Liraglutide and obesity: a review of the data so far. Drug Des Devel Ther. 2015;9:1867–1875.

33. Shorter K, Farjo NP, Picksley SM, Randall VA. Human hair follicles contain two forms of ATP-sensitive potassium channels, only one of which is sensitive to minoxidil. FASEB J. 2008 Jun;22(6):1725–1736.

34. Li M, Marubayashi A, Nakaya Y, Fukui K, Arase S. Minoxidil-induced hair growth is mediated by adenosine in cultured dermal papilla cells: possible involvement of sulfonylurea receptor 2B as a target of minoxidil. J Invest Dermatol. 2001 Dec;117(6):1594–1600.

35. Ghofrani HA, Osterloh IH, Grimminger F. Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. Nat Rev Drug Discov. 2006 Aug;5(8):689–702.

36. Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol. 2010;6:343.

37. Aronson JK. Meyer's Side Effects of Drug 6 Volume Set. Meyler's Side Effects of Drugs Series. Elsevier; 2015.

38. Ray SD. Side Effects of Drugs Annual: A worldwide yearly survey of new data in adverse drug reactions. v. 36. Elsevier Science; 2014.

39. Drugs@FDA: FDA Approved Drug Products;. Accessed: 2015-04-30. `http://www.accessdata.fda.gov/scripts/cder/drugsatfda/`.

40. Backstrom M, Mjorndal T, Dahlqvist R. Under-reporting of serious adverse drug reactions in Sweden. Pharmacoepidemiol Drug Saf. 2004 Jul;13(7):483–487.

41. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. Drug Saf. 2009;32(1):19–31.

42. Hazell L, Shakir SA. Under-reporting of adverse drug reactions: a systematic review. Drug Saf. 2006;29(5):385–396.

43. Tandon VR, Mahajan V, Khajuria V, Gillani Z. Under-reporting of adverse drug reactions: A challenge for pharmacovigilance in India. Indian J Pharmacol. 2015;47(1):65–71.

44. Evans WE, McLeod HL. Pharmacogenomics–drug disposition, drug targets, and side effects. N Engl J Med. 2003 Feb;348(6):538–549.

45. Jones N. Too few women in clinical trials? Nature. 2009 Jun;Available from: `http://dx.doi.org/10.1038/news.2009.549`.

46. Leaman R, Wojtulewicz L, Sullivan R, Skariah A, Yang J, Gonzalez G. Towards Internet-age Pharmacovigilance: Extracting Adverse Drug Reactions from User Posts to Health-related Social Networks. In: Proceedings of the 2010 Workshop on Biomedical Natural Language Processing. BioNLP '10. Stroudsburg, PA, USA: Association for Computational Linguistics; 2010. p. 117–125. Available from: http://dl.acm.org/citation.cfm?id=1869961.1869976.

47. Segura-Bedmar I, Martinez P, Revert R, Moreno-Schneider J. Exploring Spanish health social media for detecting drug effects. BMC Med Inform Decis Mak. 2015;15 Suppl 2:S6.

48. Sampathkumar H, Chen XW, Luo B. Mining adverse drug reactions from online healthcare forums using hidden Markov model. BMC Med Inform Decis Mak. 2014;14:91.

49. Chee BW, Berlin R, Schatz B. Predicting adverse drug events from personal health messages. AMIA Annu Symp Proc. 2011;2011:217–226.

50. Twitter;. Accessed: 2015-04-30. http://twitter.com/.

51. Twitter Usage Statistics;. Accessed: 2015-04-30. http://www.internetlivestats.com/twitter-statistics/.

52. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital Drug Safety Surveillance: Monitoring Pharmaceutical Products in Twitter. Drug Safety. 2014;37(5):343–350. Available from: http://dx.doi.org/10.1007/s40264-014-0155-x.

53. Nikfarjam A, Sarker A, O'Connor K, Ginn R, Gonzalez G. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. Journal of the American Medical Informatics Association. 2015;.

54. Mikolov T, Chen K, Corrado G, Dean J. Efficient Estimation of Word Representations in Vector Space; 2013.

55. Sarker A, Gonzalez G. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. J Biomed Inform. 2015 Feb;53:196–207.

56. Ginn R, Pimpalkhute P, Nikfarjam A, Pakti A, O'Connor K, Sarker A, et al. Mining Twitter for Adverse Drug Reaction Mentions: A Corpus and Classification Benchmark. In: Proceedings of the 4th Workshop on Building and Evaluating Resources for Health and Biomedical Text Processing; 2014. p. 1–8.

57. Karimi S, Wang C, Metke-Jimenez A, Gaire R, Paris C. Text and Data Mining Techniques in Adverse Drug Reaction Detection. ACM Comput Surv. 2015 May;47(4):56:1–56:39. Available from: http://doi.acm.org/10.1145/2719920.

58. Plachouras V, Leidner J, Garrow A. Large-Scale Mining of Social Media for Medical Drug Side Effects. Unpublished data;.

59. Meinshausen N, Bühlmann P. High-dimensional graphs and variable selection with the lasso. The Annals of Statistics. 2006;p. 1436–1462.

60. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostatistics. 2008 Jul;9(3):432–441.

61. Banerjee O, El Ghaoui L, d'Aspremont A. Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. The Journal of Machine Learning Research. 2008;9:485–516.

62. Zeng QT, Crowell J, Divita G, Roth L, Browne AC. Identifying Consumer-Friendly Display (CFD) Names for Health Concepts. In: AMIA Annual Symposium Proceedings; 2005. p. 859–863. Available from: `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1560732/`.

63. Thomson Reuters World Drug Index;. Accessed: 2015-04-30. `http://thomsonreuters.com/en/products-services/pharma-life-sciences/life-science-research/world-drug-index.html/`.

64. Thomson Reuters Cortellis Clinical Trials Intelligence;. Accessed: 2015-04-30. `http://go.thomsonreuters.com/cti/`.

65. Allen MP. Covariance and linear independence. Understanding Regression Analysis. 1997;p. 31–35.

66. Ledoit O, Wolf M. Honey, I shrunk the sample covariance matrix. UPF Economics and Business Working Paper. 2003;(691).

67. Schäfer J, Strimmer K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Statistical applications in genetics and molecular biology. 2005;4(1).

68. Jones DT, Buchan DW, Cozzetto D, Pontil M. PSICOV: precise structural contact prediction using sparse inverse covariance estimation on large multiple sequence alignments. Bioinformatics. 2012;28(2):184–190.

69. Hastie T, Tibshirani R, Wainwright M. Boca Raton, FL, USA: CRC Press; 2015.

70. Sustik MA, Calderhead B. GLASSOFAST: An efficient GLASSO implementation; 2012.

71. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012 Jan;40(Database issue):D1100–1107.

72. Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, et al. The ChEMBL bioactivity database: an update. Nucleic Acids Res. 2014 Jan;42(Database issue):D1083–1090.

73. Kovács IA, Mizsei R, Csermely P. A unified data representation theory for network visualization, ordering and coarse-graining. ArXiv e-prints. 2014 Sep;.

74. Gimpl G, Wiegand V, Burger K, Fahrenholz F. Cholesterol and steroid hormones: modulators of oxytocin receptor function. Progress in brain research. 2002;139:43–55.

75. Feifel D. Oxytocin as a potential therapeutic target for schizophrenia and other neuropsychiatric conditions. Neuropsychopharmacology. 2012 Jan;37(1):304–305.

76. Feifel D, Reza T. Oxytocin modulates psychotomimetic-induced deficits in sensorimotor gating. Psychopharmacology (Berl). 1999 Jan;141(1):93–98.

77. Lee PR, Brady DL, Shapiro RA, Dorsa DM, Koenig JI. Social interaction deficits caused by chronic phencyclidine administration are reversed by oxytocin. Neuropsychopharmacology. 2005 Oct;30(10):1883–1894.

78. Feifel D, Macdonald K, Nguyen A, Cobb P, Warlan H, Galangue B, et al. Adjunctive intranasal oxytocin reduces symptoms in schizophrenia patients. Biol Psychiatry. 2010 Oct;68(7):678–680.

79. Pedersen CA, Gibson CM, Rau SW, Salimi K, Smedley KL, Casey RL, et al. Intranasal oxytocin reduces psychotic symptoms and improves Theory of Mind and social perception in schizophrenia. Schizophr Res. 2011 Oct;132(1):50–53.

80. Rubin LH, Carter CS, Drogos L, Pournajafi-Nazarloo H, Sweeney JA, Maki PM. Peripheral oxytocin is associated with reduced symptom severity in schizophrenia. Schizophr Res. 2010 Dec;124(1-3):13–21.

81. Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD. Cytoscape Web: an interactive web-based network browser. Bioinformatics. 2010 Sep;26(18):2347–2348.

82. Owen RT. Ramelteon: profile of a new sleep-promoting medication. Drugs Today. 2006 Apr;42(4):255–263.

83. Jung SH, Park JM, Moon E, Chung YI, Lee BD, Lee YM, et al. Delay in the recovery of normal sleep-wake cycle after disruption of the light-dark cycle in mice: a bipolar disorder-prone animal model? Psychiatry Investig. 2014 Oct;11(4):487–491.

84. Wu YH, Ursinus J, Zhou JN, Scheer FA, Ai-Min B, Jockers R, et al. Alterations of melatonin receptors MT1 and MT2 in the hypothalamic suprachiasmatic nucleus during depression. J Affect Disord. 2013 Jun;148(2-3):357–367.

85. Nicolson SE, Nemeroff CB. Ziprasidone in the treatment of mania in bipolar disorder. Neuropsychiatr Dis Treat. 2007 Dec;3(6):823–834.

86. Rainsford KD, Ying C, Smith FC. Effects of meloxicam, compared with other NSAIDs, on cartilage proteoglycan metabolism, synovial prostaglandin E2, and production of interleukins 1, 6 and 8, in human and porcine explants in organ culture. J Pharm Pharmacol. 1997 Oct;49(10):991–998.

87. Ogino K, Hatanaka K, Kawamura M, Ohno T, Harada Y. Meloxicam inhibits prostaglandin E(2) generation via cyclooxygenase 2 in the inflammatory site but not that via cyclooxygenase 1 in the stomach. Pharmacology. 2000 Nov;61(4):244–250.

88. Reinholt FP, Hultenby K, Oldberg A, Heinegard D. Osteopontin–a possible anchor of osteoclasts to bone. Proc Natl Acad Sci USA. 1990 Jun;87(12):4473–4475.

89. Wheater G, Hogan VE, Teng YK, Tekstra J, Lafeber FP, Huizinga TW, et al. Suppression of bone turnover by B-cell depletion in patients with rheumatoid arthritis. Osteoporos Int. 2011 Dec;22(12):3067–3072.

90. Boumans MJ, Thurlings RM, Yeo L, Scheel-Toellner D, Vos K, Gerlag DM, et al. Rituximab abrogates joint destruction in rheumatoid arthritis by inhibiting osteoclastogenesis. Ann Rheum Dis. 2012 Jan;71(1):108–113.

91. authors listed N. P2RY12 and its role in osteoclast activity and bone remodeling. Bonekey Rep. 2012;1:232.

92. Danaher P, Wang P, Witten DM. The joint graphical lasso for inverse covariance estimation across multiple classes. J R Stat Soc Series B Stat Methodol. 2014 Mar;76(2):373–397.

93. Mazumder R, Hastie T. Exact Covariance Thresholding into Connected Components for Large-Scale Graphical Lasso. J Mach Learn Res. 2012 Mar;13:781–794.