New methods to detect combinations of drugs associated with cancer

Rachel D. Melamed
Department of Biological Sciences
University of Massachusetts, Lowell
198 Riverside St, Lowell, MA 01854
Rachel_Melamed@uml.edu

Abstract

Combinations of common drugs may, when taken together, have unexpected effects on diseases like cancer. It is not feasible to test for all combination drug effects in clinical trials, but in the real world, drugs are frequently taken in combination. Then, there may be undiscovered effects protecting users of drug combinations from cancer–or increasing their risk. By analyzing health claims data containing numerous people exposed to drug combinations, we have an opportunity to evaluate the association of drug combinations with cancer risk. Discovering these effects can not only contribute to prevention of cancer, but also suggest new clinical uses for drug combinations to prevent or treat cancer. Here we describe new methods for discovery of drug combination effects on cancer. Our approach emulates a randomized trial where one arm would have been assigned to take a drug alone, while the other arm takes the drug together with a second drug. Because discovery of associations from observational data is prone to spurious results due to confounding, we develop a number of strategies to distinguish confounding from biomedically relevant findings.

Introduction

Late onset chronic diseases are responsible for a bulk of deaths, with cancer ranking as the second leading cause of death. Despite intensive research, we have not identified effective ways of preventing most cases, and new treatments are urgently needed. To address this problem, one approach is to investigate what common exposures could be associated with disease risk. Among people over 40 years old, one third use two or more drugs, and 20% use more than five medications[1]. We propose that combinations of common medications could influence cancer risk. These effects are unlikely to have been uncovered in randomized trials, but examples exist. Whie estrogen replacement as a treatment for menopause is not linked to adverse cancer outcomes, combining estrogen with progesterone appeared to increase risk of breast cancer over a three year follow-up period[2]. Statins, prescribed for high cholesterol, have been linked to reduced incidence of multiple cancer types in epidemiological studies, and they are in trials for cancer therapy. Trials have investigated anti-cancer effects of statins combined with celecoxib (anti-inflammatory) and metformin (a diabetic drug)[3]. Therefore, common drugs may both positively and negatively impact disease.

As tens of thousands of common drug combinations could be tested for any of a dozen types of common cancer, these effects are unlikely to be discovered through experimental trials. Outside of the setting of clinical trials, one approach to discovering such drug effects is analysis of observational data. Using such health data, we can follow health outcomes in people exposed to the drug combination. Although reusing existing data is fast, cheap, and allows systematic discovery, new methods are needed to find drug combination effects. Unlike experimental data, observational data is subject to confounding: people are prescribed drugs based on their health, and some health conditions incur increased risk of cancer. For instance, smoking increases lung cancer risk, meaning that people who take smoking cessation drugs are likely at increased risk of lung cancer. Without accounting for such variation in health, the drugs and drug combinations most associated with cancer would be dominated by such spurious associations.

The typical approach to discovering a biomedically relevant association from observational data is to attempt to emulate a randomized trial that could have investigated the same question. Two study cohorts, consisting of users or non-users of a drug are "enrolled", and then the analysis compares their outcomes, adjusting for confounders of the association. While we do not know of other drug combination-wide studies, we have reviewed previous drug-wide association studies of cancer (single drug, not combinations)[4–7]. Because of their systematic nature, these studies make some simplifying assumptions. Particularly, they adjust for a limited set of confounders such as age, sex and a summary comorbidity index. It is not clear if these measures are adequate. One tactic, called "empirical calibration" compares the effect of a drug on outcomes of interest against its association with selected "negative control" outcomes[8,9]. For instance, we expect most common drugs do not affect risk of ingrown nails. This systematic approach allows us to summarize the distribution of effect estimates on a range of outcomes. The authors suggest that we can estimate, and thus remove, the bias by assuming all effect estimates are centered around the overall bias. One weakness of this assumption is that bias is often specific to the particular pair of exposure and outcome. For instance, people taking smoking cessation drugs are at risk of lung cancer, but not breast cancer. As well, randomly selected negative controls may not suffer the same bias; we expect no confounding effect associating smoking cessation agents with ingrown nail.

Therefore, here we develop new methods for *drug combination-wide association studies.* Our aim is a systematic discovery of drug combinations that, like other drug-wide methods, allows unpredjudiced investigation of any observable drug combination. We accept that due to the nature of observational data, our results must be replicated and verified by future experimental studies. But, we make extensive efforts to identify the effect of confounding, and we use a number of sensitivity analysis to pinpoint associations that are most robustly supported. We expect both our methods and the identified drug combination associations can provide advances toward clinically relevant insights from observational data.

Results

Observational data to emulate a randomized trial of drug combinations

In order to set up our approach, we first design a randomized trial that would test the association of any arbitrary combination of drugs with a common cancer outcome. Our design is inspired by the case control study. We make a design to estimate the effect of a combination of the arbitrary pair of drugs A and B. Our design enrolls new users of some drug A who have never before taken either drug A or B, and with no history of cancer. Then, our hypothetical trial would randomly assign some people to additionally take drug B (Fig 1A). Finally, we would follow these two groups to observe incidence of cancer. Analyses of the results could follow all people randomized, called an *intention-to-treat* analysis. An alternative analysis could censor follow-up time for those who discontinued use of drug A, known as a *per-protocol* analysis.

To emulate such a trial, we use a set of health claims data from Truven MarketScan (now IBM), containing 150 million drug users. Our study uses data with coded diagnoses, prescriptions, and procedures, alongside basic demographic characteristics, over a twelve year span from 2003 to 2015. We observe 9,502 drug combinations with a median of 4,979 combination users per drug pair (Fig 1B), and with a median of 20 exposed people who eventually develop cancer across all cancer types (Fig 1C). To summarize, we perform this emulation of a randomized trial for 9,502 pairs of a drug A and B. For the rest of this manuscript, we refer to the first drug as drug A, and the second drug as drug B. In theory, we might expect the same true effect if the order of the drugs were reversed, but in practice sample size limitations mean we rarely observe both orders.

Approach to emulating the randomized trial

We divide our main design into three steps, analogous to steps of a randomized trial: 1) enrollment (creation of our cohorts); 2) randomization (adjusting for time-varying confounding); and 3) analysis of outcomes and estimation of effects.

In the enrollment step, as mentioned, we compile all new users of drug A, who have never taken drug B or had the outcomes of interest. While in the randomized setting, some subset would be immediately assigned to take the combination of drugs A and B, in the observational setting, some people will start drug B immediately, while others might start drug B later. While we could only compare those who start drug B immediately after drug A against those who

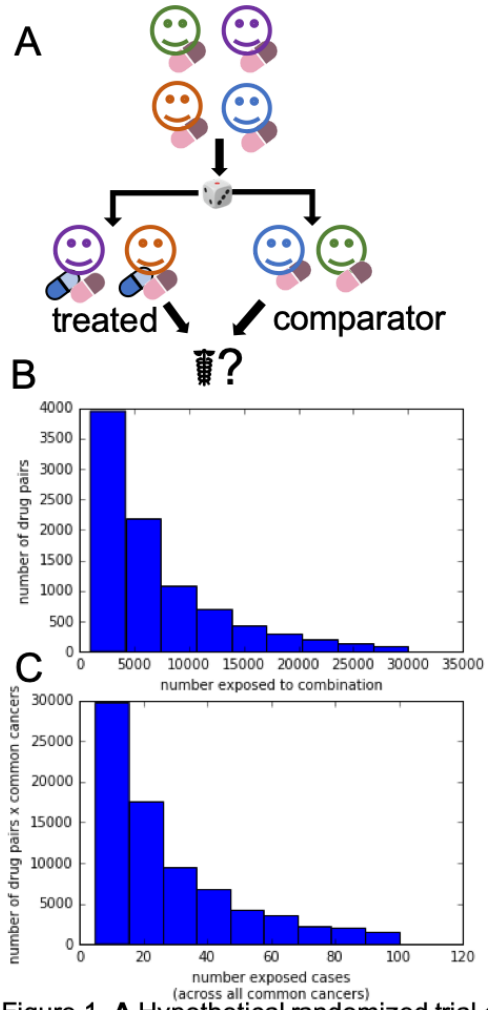

Figure 1. **A** Hypothetical randomized trial of drug combinations. **B.** Total numbers of people exposed to a drug pair, across 9,502 drug pairs. **C.** Number of cancer cases per drug pair, aggregated across all common cancers.

take drug A alone, this choice would drastically reduce our data size. In order to capture all person-time under joint exposure of the drug combination, we follow the approach of emulating a *sequence of* randomized trials[10,11]. This tactic divides follow-up time after initiation of drug A for each person into arbitrary but reasonable windows (we choose 3 or 6 month windows). Each time window defines a trial in the sequence of trials: a person who has not initiated drug B yet at follow-up time window *w* remains eligible to "enroll" in a trial that starts at that time window (Figure 2). Therefore, we are able to include all joint users of drugs A and B in one of these trials. For instance, a person who started the drug combination in window 3 is in the treated group for trial 3, but is in the comparator group for windows 1 and 2. This person is not eligible to join trials starting after trial 3. We can easily account for this replicated data by indicating the repeated observations in our statistical analysis (see Methods).

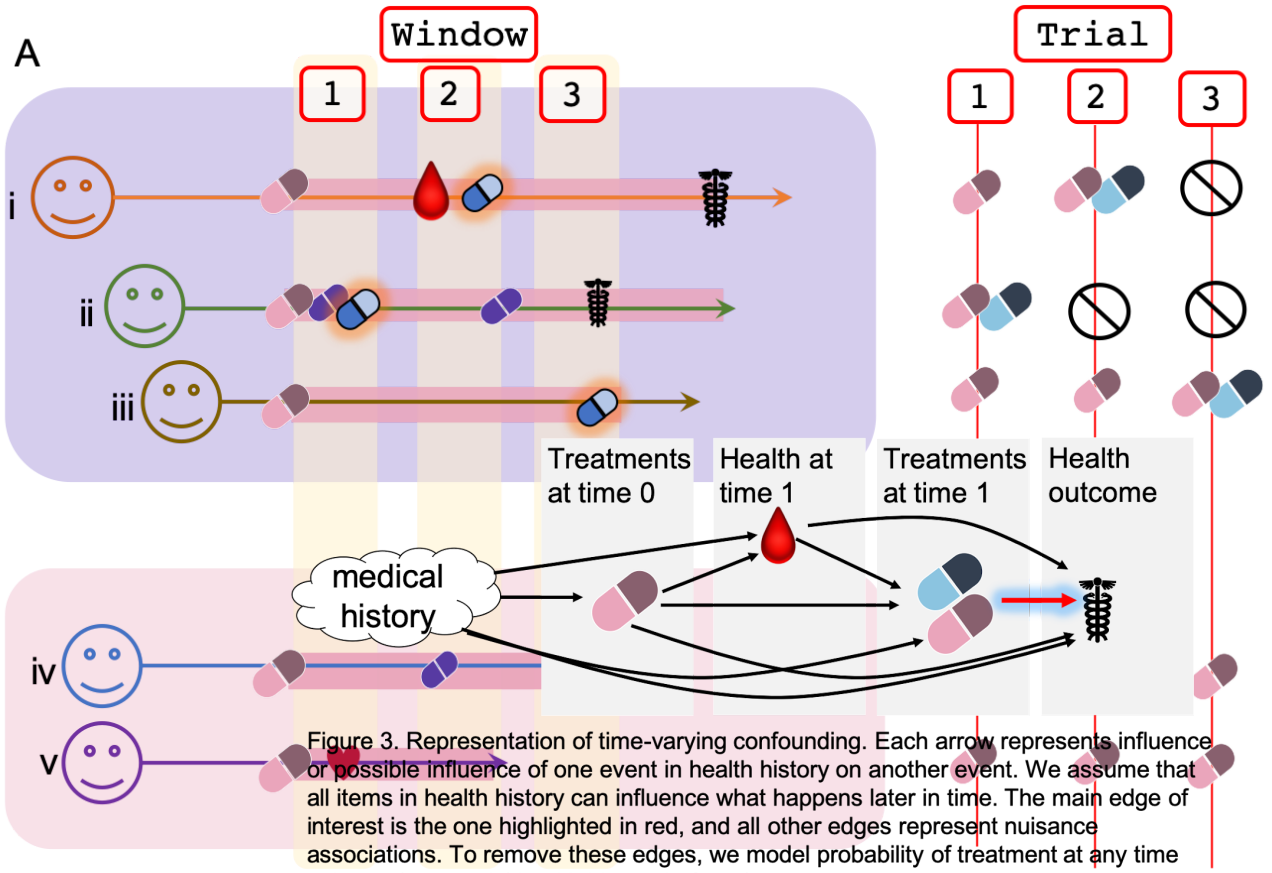

Figure 3. Representation of time-varying confounding. Each arrow represents influence or possible influence of one event in health history on another event. We assume that all items in health history can influence what happens later in time. The main edge of interest is the one highlighted in red, and all other edges represent nuisance associations. To remove these edges, we model probability of treatment at any time point w (exposed, health history before w).

Figure 2. Enrollment for trial 1. **A.** Medical history for five people exposed to drug A (pink pill) is represented with time along the horizontal axis. Follow-up time is divided into 3 windows. The top three people are eventually exposed to the combination of drugs A and B (blue pill), while the bottom two people remain in the comparator group (drug A alone). **B.** Representation of the expansion of person-time for these five people into a sequence of randomized trials. Person 1 is unexposed at time window 1 (trial 1), and thus is in the comparator group, but still eligible to participate in trial 2. At trial 2 they receive the combination, and thus are assigned to the joint-exposure group in trial 2, and they are not eligible to participate in trial 3.

In the randomization step, we emulate randomization by adjusting for time varying confounding. It is crucial to not only adjust for disparities between the treated and comparator groups at baseline initiation of drug A. This is because changes in health over time can

influence which people are prescribed drug B (Figure 3). For instance, onset of menopause ("Health at time 1" in Figure 3) might prompt prescription of drugs to treat the symptoms, and menopause also increases risk of breast cancer. Therefore, if we do not adjust for time-varying confounding we would find a spurious association between menopause-related drugs and breast cancer. Previous drug-wide methods that analyze use of single drugs only have not required time-varying analysis, and these methods have adjusted only for selected confounders. Instead, we consider that all medical history before time window *w* could include possible confounders. Because of the time-specific nature of medical data, colliders[12,12] are unlikely: this would require both the treatment and the exposure to affect some variable that occurs *before* either of them. While errors in the medical record induce some cases of reverse causality, we describe measures taken to guard against this issue later.

To model confounding, we calculate probability of initiating drug B, based on both time window-specific and overall medical history. Confounding also can influence whether someone discontinues drug A, an issue that also arises in analysis of the data from a randomized trial. In a per-protocol analysis of randomiszed trials, person-time is censored after discontinuation of drug A. In order to best capture the effect of exposure to drug combinations, we also censor person-time after discontinuation of drug A. To account for confounding impacting censoring, we also model probability of discontinuation in the same way as we model initiation of drug B. Finally, we obtain time to cancer outcome using a survival analysis. The probabilities calculated in the randomization step are used as weights to remove the confounding-related association between drug and outcome in a marginal structural model[13]. In this manner, we can estimate the association between each drug combination and each common cancer, for those where enough cases are observed.

Drug-wide studies may be influenced by a wide range of confounders

Our analysis yields estimates for 9,502 drug combinations, for a median of 22 cancers where more than 100 people in the cohort study were diagnosed with that type of cancer. As mentioned, previous drug-wide studies typically account for only a handful of confounders. We investigate the possible effects of considering a narrow range of confounders. In Figure 4, we compare the effect estimates that consider no confounders (unweighted); those derived from cohort studies that only adjust for age, sex, year, and Charlson comorbidity index (minimal confounding control); and our method of including a wide range of possible confounders (weighted estimates). We find that globally, the effect estimates are greatly reduced. Because this is a general effect, across many drug combinations and cancer outcomes, we conclude that this is likely to be due to a number of assorted confounding effects, rather than due to widespread effects of drug combinations of cancer. This suggests that other drug-wide

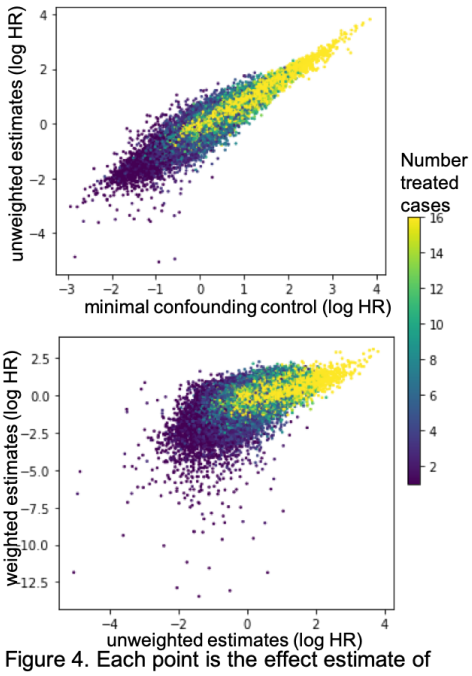

Figure 4. Each point is the effect estimate of one drug combination on one outcome, across 200,000 cohort studies. Top: the unweighted and minimally adjusted estimates are largely similar to each other. Bottom: a global reduction in effect estimates with weighting, due to control for confounding. By highlighting the number of treated cases, we can see that extreme “protective” (negative log hazard ratio) effects are associated with a low number of exposed cases, and likely largely spurious.

studies would benefit from considering a wide range of confounders. However, the results also suggest that there is likely substantial remaining confounding. The ability to support such conclusions is a key advantage of taking a systematic drug (combination)-wide approach to discovery.

Initial results of the drug-combination wide scan and remaining confounding

Across around 200,000 resulting effect estimates, we investigate selected examples to assess the nature of the remaining confounding. Among drugs most associated increased risk of cancer, we find two obvious patterns. First, drugs that are used in the treatment of cancer appear associated with rates of multiple types of cancer. For example, fentanyl, morphine, and other pain treatments, as well as anti-nausea medications such as prochlorzerpine or ondansetron appear co-prescribed in cancer cases. This is likely due to incompleteness of the medical record: these people are overwhelmingly likely to be under treatment for cancer. Then, this is a case of *reverse causality,* where the medications are a result of the diagnosis, rather than the other way around. Reverse causality is usually associated with many common cancers, as in Figure 5A, as many cancers involve surgery, chemotherapy, and drugs to manage the symptoms and compications of these procedures. We later present a simple heuristic to sort these cases.

The second pattern includes likely cases of confounding, which is usually specific to one type of cancer. For instance, kidney disease impacts glaucoma, and we find that people taking a combination of two glaucoma treatments (latanaprost and travaprost), as compared to those who take only latanoprost, are at higher risk of kidney cancer (Figure 5B). This is overwhelmingly likely to be due to the confounding effect of kidney disease on risk of both kidney cancer, and glaucoma (and thus these drugs). Note that this combination is specifically associated with kidney cancer, which is logical due to the specific confounding effect of kidney disease. Such examples indicate remaining confounding, a separate issue from reverse causality. Again, our systematic exposure-wide assessment enables identification of this issue.

Because the presence of confounding in observational studies is well known, approaches like the Hill criteria seek to distinguish epidemiologic signals with strongest evidence of biomedically relevant effects[14]. We first focus on two criteria: consistency and specificity.

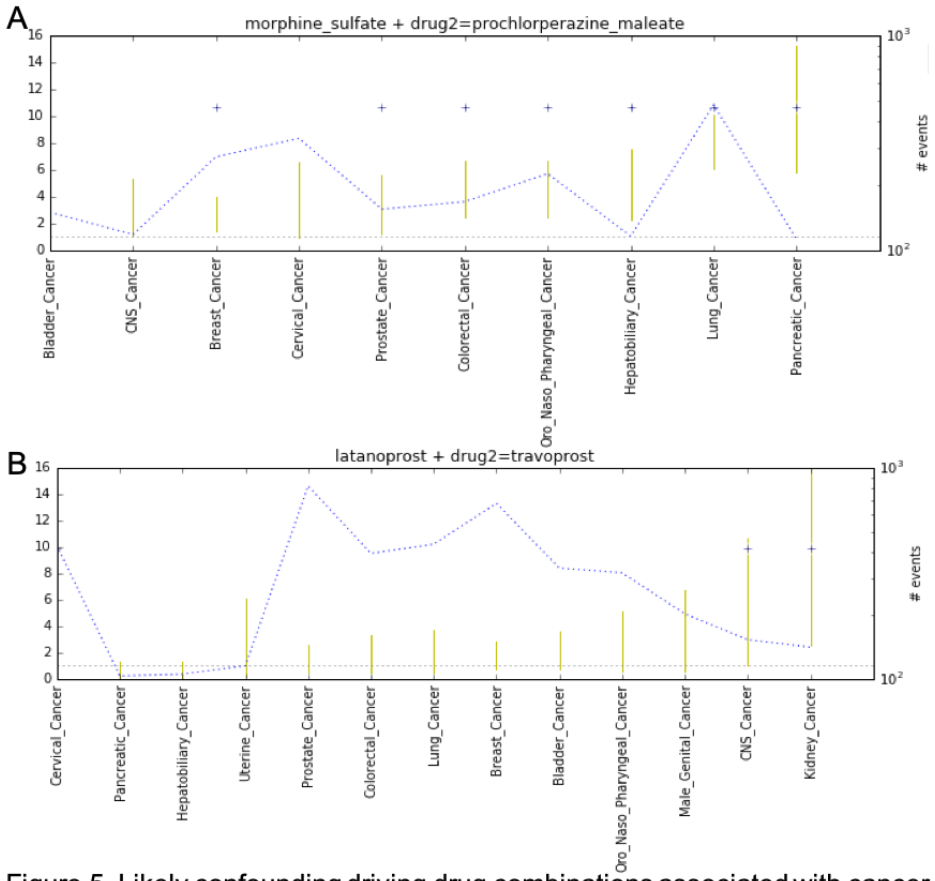

Figure 5. Likely confounding driving drug combinations associated with cancer. Each yellow line represents the confidence interval of effect of one drug combination on cancer (hazard ratio scale, + symbols denote significance). The dashed blue line indicates number of incident cases (events) of each cancer type.

Consistency: First, we examine the consistency among results that we would expect to be consistent with each other. Some previous findings of drugs impacting cancer, like statins, appear to impact multiple cancer types. Our examples of confounding (rather than reverse causality), on the other hand, are specific to a single cancer. Therefore, our current analysis seeks drug effects that are consistent across multiple common cancer types. As well, we consider consistency among drugs sharing an active ingredient as another form of evidence.

Taking inspiration from the consistency heuristic, we implement an analytical approach to integrate the results of multiple cohort studies and distinguish the strongest signals from confounding. Because of the size of the data and the even larger size of the data when expanded into a sequence of windows of person-time, we split up any large cohort studies into *replicate* cohort studies containing disjoint study populations. We integrate these multiple effect estimates in a new Bayesian hierarchical model, to estimate the per-combination effects on any cancer (Figure 6). The first level of the hierarchy assumes that the replicate cohort study effects of the drug combination on each common cancer are centered around the true per-cancer effect of that combination. Then, at the next level, our model specifies that effects of the combination on a particular cancer type are centered around the overall pan-cancer effect of the drug comb

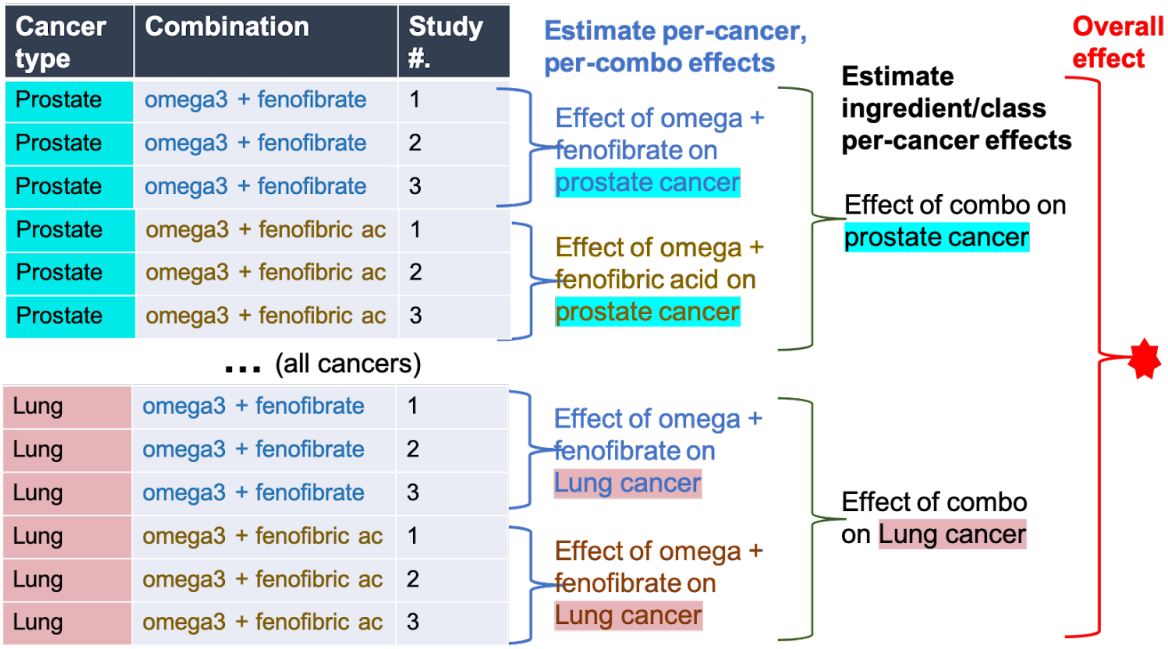

Figure 6. A hierarchical model for identifying a combination of drug A (here omega 3 acids (fish oil)) combined with any drug B containing fenofibric acid as an active ingredient.

ination. In a slightly different model, we estimate the effects for a combination of drug A and all members of the set of drugs B that share some particular active ingredient. This model only differs from the previous model in that it has one extra layer in the hierarchy: we assume that the per-cancer effects for the drug are drawn from the overall effect of that ingredient class on that type of cancer. The effect estimates yielded by the Cox regression comprise the data input to the model. After fitting the model to the effect estimates, we obtain the posterior distribution of the overall effect of the combination on cancer. We classify protective drugs as those where the upper 1% posterior interval is less than zero, and predisposing drugs as those where the bottom 1% is greater than zero.

Specificity: Second, another criterion examines whether the associations are specific to the outcome of interest. We could imagine that people in poorer health simply take more medications; many indicators of poor health, such as sedentary lifestyle, increase overall cancer risk. But, we may expect in that case that these people may have increased risk of other non-cancer diseases. Therefore, the association would not be specific to cancer but would correspond to general sickliness. We create a set of negative control outcomes, matching each common cancer with a non-cancer outcome with a roughly similar incidence rate in our study population (Figure 7A). These outcomes are chosen to span a wide range of body systems. We apply the same hierarchical model to our negative control outcomes, which allows us to identify drug combinations that are generally associated with illness.

Detecting reverse causality: We found many examples of reverse causality, similar to the one illustrated in Figure 5A. As mentioned, we expect that these are due to erroneous missing information about a cancer diagnosis, where eventually the diagnosis is recorded in the coded data, but only after some amount of treatment. It is reasonable to assume that when cancer in fact causes the treatment, the diagnosis of cancer will occur close in time to the treatment. We develop two tests to detect these cases. First, we test among cancer cases,

whether the duration of time between the trial start date and date of cancer diagnosis is shorter among those treated with the combination (Fig 7B). Second, we test whether this temporal relationship is stronger for cancer types more strongly associated with the drug combination. In this manner, we can exclude the obvious cases of reverse causality

Sensitivity analyses to strengthen findings

After the previous series of tests, we are left with only a few drug-combination wide associatons. We further subject these signals to a set of sensitivity analysis. These test whether the results are sensitive to some of the design decisions we made., We re-run the same drug wide association study for the candidate signals, twice, each with two different sets of assumptions. For each, we perform the consistency analysis to assess whether the drug pair impacts all-cancer risk.

First, we change the window width used for defining trial start periods. While our primary design used 6 month windows, we create a secondary analysis that uses 3 month windows to define trial start periods. Note that this creates more follow-up periods, and results in a larger person-time matrix.Our second sensitivity analysis tests the effect of varying the definition of drug combination. While our main analysis censored person-time when the person stopped taking drug A, this results in a very limited follow-up time, where most people are followed for no more than a year. We expect this can only allow us to detect smaller drug effects. Removing this requirement, we can follow time after exposure for an average of 3.25 years, which is as long as people are followed for cancer outcomes in some randomized trials[2]. Removing the censoring greatly increases the number of cancers we can observe in follow up, but it also increases the size of our expanded person-time matrix, necessitating smaller replicate cohort studies.

Our final sensitivity analysis assesses the association of amount of drug B dispensed with cancer incidence. We find

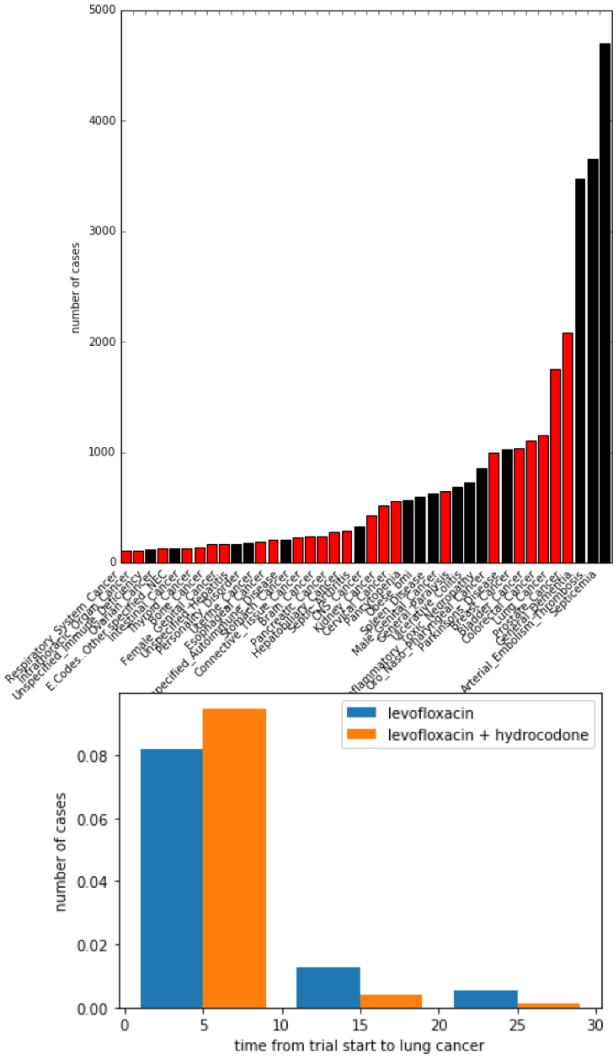

Figure 7. **A**. Overview of frequency of incidence of cancers (red) and of negative controls (black) (among one drug combination cohort study. **B.** Reverse causal effect of lung cancer on hydrocodone can be detected due to shorter time span from combination to cancer diagnosis ($p = 8x10^{-6}$, rank sum test)

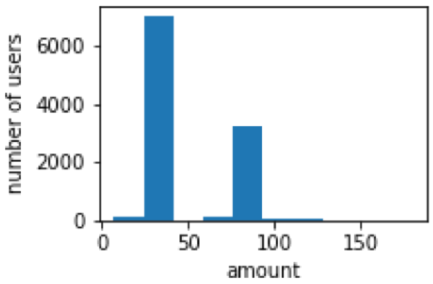

Figure 8: Frequency of prescription for various amounts of fenofibrate

empirically that most drugs have a few common amounts dispensed, and we simply categorize users as in the "low" amount group or the "high" group. For instance, for fenofibrate users two main amounts dominate. Because the number of people exposed to a drug combination and later getting cancer is rather low, we do not further categorize drug amounts. We consider the signal has passed this test if 1) both categories are significant and 2) the effect size is stronger for the high group than for the low group.

Combining the sensitivity analyses with the reverse causality test, we narrow down our signals to the final associations. Our final results are presented in Table 1 and Table 2 (see end). We find 50 single drug pairs are associated with either an elevated or a reduced rate of cancer. Among these, many can be further excluded from a cancer effect by the negative control test. Further, by combining signals across drugs in ingredient categories, we can prioritize drug combinations with a consistent effect in the ingredient class. Among our protective findings, only the combination of omega-3 fatty acids plus fenofibrate was able to be tested as an ingredient class (where fenofibrate is an ingredient of both fenofibric acid and fenofibrate). The ingredient class also was associated with reduced risk of cancer. As well, the sensitivity analysis for amount of medication dispensed showed stronger protective effects for the high amount versus the lower amount (average log hazard ratio of -2.9 for the high amount, versus -1.6 for the low amount).

Discussion

We have described a systematic method to perform a drug combination-wide association study across all cancer types. While further studies are needed, one interesting finding is that a combination of Omega 3 acids (fish oil) with fenofibrate may be protective of cancer. Our negative control analysis indicates that this is not merely an effect of increased health consciousness but is specific to cancers. Our analysis of amount dispensed and other sensitivity analysis further supports the results over a 3 year follow up period. This pair of drugs both impact circulating lipid levels. Circulating lipids are thought to impact cancer development via the Warburg effect. Therefore a biomedically relevant effect is plausible.

Our analysis has a number of limitations. Foremost, our results display evidence of remaining confounding. We previously showed that the propensity score is not able to model many important types of confounding[15]. This is because the medical record is incomplete. Many unpredictable factors can influence drug prescription, ranging from various comorbidities in medical history, to the patient's gender, ethnicity, and geographical location[16]. Therefore, future work must improve upon the propensity score and enable better modeling of confounding. Our results can provide a starting point for these efforts by providing a large set of cases of real-world confounding. A second limitation of our work is the limited follow-up time available for observing drug combination effects. In our analysis with no censoring, we have 3.25 years of follow-up time available. This is short, but it is in line with some previously discovered combination drug effects[2].

Our unique approach has a number of strengths. First, this work represents, to our knowledge, the first application of marginal structural models for discovery of drug combinations impacting health outcomes. While marginal structural models have been used to describe effects of time-varying treatment, our unique design allows us to expand on these methods to

describe the time-varying effect of adding a second drug on top of treatment with an initial drug. Another unique method is our hierarchical model for combining drug effect estimates on related diseases to create an overall summary of the effect on a disease category. While some other studies have combined drug effect estimates using a hierarchical model[17], this approach is the first to use hierarchical models to identify an estimate across multiple diseases that is more robust to many types of confounding.

A final and crucial strength of our design is its systematic nature. By performing a systematic analysis across all drug combinations, we are able to compare methods and discover their weaknesses. For instance, we showed that methods that only account for a few general confounders are quite vulnerable to bias. Our thorough consideration of sensitivity to our study design, and our attempts to quantify remaining confounding, increase confidence in our results. We expect that this approach an provide a basis for future work seeking to discover drug combination effects from health data.

Methods

Expansion of medical record into person time

We divide up follow-up time into windows of time since initiation of drug A. Our main analysis uses windows of 24 weeks, and a subsequent sensitivity analysis uses windows of 12 weeks. For each window, we obtain patients' time-specific data, including if they started drug B, and other medical history of that time, including incidence of cancer. As outlined in Figure 2, we expand person-time into repeated trials. Specifically, people are eligible for a trial as long as 1) they are still observed in the data set (and not censored, depending on the design) and 2) they they have not yet taken drug B, and they have not yet had the outcome of interest (here, a specfic cancer). In trial 1, the person's full time observed is the amount of time available for follow-up. For subsequent trials, follow-up time begins only after trial initiation. In Table 3, we illustrate the expansion info person-time for person i and person v from Figure 2. The shorter the windows, the more rows will be in this table. As well, the longer the follow-up time (which increases when no censoring is performed), the more rows will be in this table.

Table 3: Expansion into trials

Person	Trial	Follow-up	Treatment	Outcome	comment
i	1	1	A	no	person i is not eligible for trials after trial 2 because they took the combination at trial 2
i	1	2	A	no	
i	1	3	A	no	
i	1	4	A	yes	
i	2	1	A+B	no	
i	2	2	A+B	no	

i	2	3	A+B	yes	
v	1	1	A	no	person v is only followed for 2 windows, they have no drug B or outcome in either.
v	1	2	A	no	
v	2	1	A	no	
v	2	2	A	no	

Modeling confounding

Taking the marginal structural model approach, we must model both time-varying confounding and baseline confounding. We implement this with a sparse high-dimensional regression. $p(treatment_t \mid health_t, baseline\ info)$. We include as variables in ***health***$_t$ and ***baseline*** all treatments and diagnoses. Baseline also includes sex and cubic spline variables to model the effect of age, calendar time, and the number of total diagnoses and treatments the person has as a measure of overall medical burden. Typically, our data has over 10,000 features, but we filter the features observed in fewer than 100 people. We fit the model using elastic net logistic regression using the scikit learn package, performing a grid search with cross-validation to choose values for the regularization parameter and the l1-ratio. For each model, we perform the hyperparameter tuning separately. After regularization, typically only a few hundred features have a non-zero coefficient in the logistic regression model. Finally, we obtain the predicted probabilities from the model.

Marginal structural models and survival analysis

Using the predicted probabilities of treatment, we use marginal structural models to estimate the association with a cancer outcome. To this end, we use the predicted probabilities from our models to create *treatment weights* that estimate the probabilitiy of each person receiving the treatment sequence they in fact received, given their medical history. So, for person *i,* this treatment sequence entails first being treated with drug A alone in time 1: $p(drug\ B_0 = 0 \mid health_0, baseline\ info)$, then being treated with the drug combination in time 2: $p(drug\ B_1 = 1 \mid health_1, baseline\ info)$. We obtain the probabilities at each time point from our model, and we multiply these to obtain the overall probability of the sequence of treatments up through person-time observation. Finally, we use the inverse of these probabilities to weight each time interval in follow-up time in the Cox regression using the R *survival* package[13]. This part of our analysis is similar to what has been implemented in other studies modeling a sequence of randomized trials with observational data[10,11]. As those previous publications proved that using the *stabilized weights* reduces extreme weights without adding bias, we use stabilized weights.

For designs where we censor person-time when the users stop taking drug A, we use the same procedure to adjust for time varying confounding of discontinuation for censoring. This is because time-varying confounding could cause censoring and could thus bias our results[13].

Censoring

In our primary analysis, we censor people when they discontinue drug A. This is a design choice that allows us to estimate the effect specifically of simultaneous exposure to drugs A and B, rather than possibly sequential exposure. As mentioned immediately above, when censoring those who discontinue drug A, we must account for time-varying influences on censoring. We obtain the probabilities of censoring exactly in the same manner as obtaining the probabilities of initiation of drug B, simply modeling $p(discontinue\ drug\ A_0 = 0 \mid health_0, baseline\ info)$ using a high-dimensional sparse logistic regression model. Finally, we obtain the weights in the same manner described above. To weight each observation adjusting for confounding of both discontinuation of drug A and of treatment with drug B, we multiply the two weights together. While this makes the assumption that the two treatment decisions are independent, by performing the sensitivity analysis without censoring, we are able to assess whether that assumption impacts the results.

Hierarchical model

The hierarchical model shown in Figure 6 is presented in Appendix Model. The slightly simpler version without the ingredient layer is very similar. The model uses as input the log hazard ratios of the treatment effects estimated from the Cox regression model. These are assumed to have a normal distribution in the Wald test. Our model suggests that these estimates are samples derived from a normal distribution centered around a true effect estimate. We fit the model using Stan, and obtain the posterior distribution of the all-cancer effect from this model. We use the same model to obtain the effect estimate across all negative control outcomes, replacing the cancer estimates with the estimates of the effect of each drug combination on incidence of each negative control health condition.

Sensitivity analyses and other tests

We repeat our drug-wide analysis, but only testing drugs that have effects tested in our main analysis. For the first sensitivity analysis, we alter the size of the windows, and for the second sensitivity analysis, we do not censor when drug A is discontinued. Because, as mentioned above, both of these alterations can result in large increases in the size of the expanded person-time data set, we must reduce the number of drug A users included in some analyses.

To perform the reverse causality test, we gather data on time to cancer for the cancer most strongly associated with the drug combination. Considering only the people who eventually get that type of cancer, we use the rank-sum test to evaluate whether the time to cancer is significantly shorter for those taking the combination versus those in the single-drug group. These values are presented in Table 1 and 2, "reverse causality test" column.

To categorize combination users into those with "low" and "high" amount of drug, we obtain the two most common amounts prescribed. We set the cut point to be half way between those two amounts.

Table 1: Drug combinations associated with reduced risk of cancer. The negative control test is passed for those in yellow (no reduced risk of non-cancers)

1 Drug(A)	Drug 2 (B)	1% posterior	99% posterior	1% posterior, negative controls	99% posterior, negative controls	50% posterior--unweighted	reverse causality test	1% posterior, 3 month windows	99% posterior, 3 month windows	reverse causality test, 3 month windows	1% posterior, no censoring	99% posterior, no censoring	reverse causality test, no censoring
tiotropium_bromide	diclofenac_sodium	-0.826	-0.325	-0.201	0.192	-0.185	0.089	-1.412	-0.466	0.056	-0.377	-0.095	0.84
metformin_hydrochloride	oseltamivir_phosphate	-0.916	-0.282	-0.447	0.049	-0.029	0.58				-0.679	-0.17	0.405
spironolactone	naproxen	-0.974	-0.186	-0.674	-0.086	-0.023	0.991	-1.081	-0.032	0.973	-0.516	-0.017	0.346
fluticasone_propionate_salmeterol_xinafoate	mometasone_furoate	-0.628	-0.173	-0.58	-0.011	-0.192	0.554	-0.762	-0.169	0.299	-0.363	-0.03	0.02
niacin	penicillin_v_potassium	-0.879	-0.148	-0.683	-0.015	-0.154	0.684	-1.269	-0.355	0.845	-0.338	-0.028	0.297
escitalopram_oxalate	oseltamivir_phosphate	-0.757	-0.12	-0.72	-0.172	-0.047	0.929	-1.711	-0.499	0.991	-0.609	-0.06	0.416
pantoprazole_sodium	fluticasone_propionate	-0.497	-0.112	-0.33	-0.026	-0.064	0.679	-0.711	-0.074		-0.299	-0.051	
donepezil_hydrochloride	olanzapine	-0.856	-0.11	-0.352	0.417	-0.044	0.293	-1.422	-0.172	0.116	-0.641	-0.069	0.016
insulin_aspart__recombinant	glucagon_hydrochloride	-0.816	-0.104	-0.52	-0.065	-0.134	0.918	-0.999	-0.196	0.648	-0.675	-0.143	0.634
amlodipine_besylate	pravastatin_sodium	-0.484	-0.094	-0.121	0.164	-0.023	0.935	-0.612	-0.158	0.41	-0.349	-0.083	0.283
omega-3-acid_ethyl_esters	fenofibrate	-0.844	-0.092	-0.394	0.225	-0.254	0.466	-1.098	-0.129	0.935	-0.514	-0.016	0.282
fluticasone_propionate_salmeterol_xinafoate	cefdinir	-0.948	-0.085	-0.613	-0.218	-0.107	0.408	-1.077	-0.157	0.025	-0.512	-0.176	0.063
nifedipine	codeine_phosphate_guaifenesin	-0.723	-0.076	-0.477	0.103	-0.038	0.494	-1.121	-0.245	0.418	-0.439	-0.022	0.631
ezetimibe_simvastatin	fenofibrate	-0.609	-0.063	-0.494	0.04	-0.033	0.527	-0.833	-0.092	0.29	-0.259	-0.072	0.818

glyburide	penicillin_v_potassium	-0.802	-0.059	-0.698	-0.151	-0.047	0.198	-0.902	-0.038	0.184	-0.489	-0.096	0.984
fexofenadine_hydrochloride	cefdinir	-0.934	-0.049	-0.634	0.205	-0.038	0.555	-0.771	-0.025	0.406	-0.365	-0.084	0.46
budesonide_formoterol_fumarate	oseltamivir_phosphate	-1.062	-0.046	-0.761	0	-0.065	0.039				-0.702	-0.084	0.082
niacin	oseltamivir_phosphate	-0.838	-0.026	-0.296	0.25	-0.049	0.879				-0.623	-0.177	0.242
valsartan	penicillin_v_potassium	-0.513	-0.017	-0.436	0.065	-0.016	0.406	-0.763	-0.076	0.469	-0.289	-0.002	0.279
omeprazole	fluticasone_propionate	-0.541	-0.012	-0.305	-0.055	-0.048	0.831	-0.783	-0.18	0.212	-0.311	-0.006	0.312

Table 2: Drugs combinations associated with increased risk of cancer. The negative control test passes for those in yellow (no increased risk of non-cancer). But, most of these are cancer therapies and likely due to reverse causality.

Drug 1 (A)	Drug 2 (B)	1% posterior	99% posterior	1% posterior, negative controls	99% posterior, negative controls	50% posterior--unweighted	reverse causality test	1% posterior, 3 month windows	99% posterior, 3 month windows	reverse causality test, 3 month windows	1% posterior, no censoring	99% posterior, no censoring	reverse causality test, no censoring
polyethylene_glycol_3350	acetaminophen_oxycodone_hydrochloride	0.18	0.723	-0.241	0.331	1.072	0.045	0.233	0.937				
tramadol_hydrochloride	acetaminophen_propoxyphene_napsylate	0.078	0.581	-0.17	0.26	0.592	0.528	0.059	0.823				
carvedilol	acetaminophen_propoxyphene_napsylate	0.008	0.471	-0.213	0.34	0.519	0.142	0.09	0.534	0.066	0.224	0.457	0.646
polyethylene_glycol_3350	acetaminophen_propoxyphene_napsy	0.272	0.838	-0.015	0.417	0.937	0.003	0.331	1.076				

	late												
divalproex_sodium	acetaminophen_propoxyphene_napsylate	0.2	0.878	-0.064	0.399	0.795	0.684				0.137	0.649	0.322
tamsulosin_hydrochloride	acetaminophen_propoxyphene_napsylate	0.163	0.45	-0.044	0.28	0.643	0.987	0.111	0.392		0.366	0.534	
fluoxetine_hydrochloride	acetaminophen_propoxyphene_napsylate	0.07	0.579	-0.23	0.257	0.742	0.015	0.086	0.758	0.005	0.253	0.551	0.199
risperidone	acetaminophen_propoxyphene_napsylate	0.301	1.183	-0.019	0.709	0.951	0.391				0.295	0.847	0.624
acetaminophen_propoxyphene_napsylate	alprazolam	0.02	0.56	-0.09	0.304	0.751	0.004	0.15	0.803				
insulin_glargine__recombinant	ciprofloxacin_hydrochloride	0.024	0.328	-0.038	0.281	0.72	0.029	0.015	0.418				
amphetamine_salt_combination	ciprofloxacin_hydrochloride	0.086	0.742	-0.019	0.397	0.577	0.509	0.275	1.132				
colchicine	furosemide	0.131	0.945	0.092	0.763	0.722	0.007	0.047	0.89				
glucose_meter_test_control_strips	furosemide	0.227	0.821	0.463	1.05	0.745	0.924	0.06	0.863		0.258	0.631	
lansoprazole	furosemide	0.103	0.758	0.094	0.869	0.553	0.001	0.124	0.925				
clonidine_hydrochloride	hydralazine_hydrochloride	0.137	0.622	0.19	0.712	0.624	0.026	0.228	0.699		0.154	0.525	
omeprazole	lansoprazole	0.058	0.681	0.011	0.406	0.906	0.053						
glyburide	levofloxacin	0.106	0.505	0.205	0.564	0.814	0.018	0.004	0.471				
insulin_glargine__recombinant	lorazepam	0.041	0.462	0.073	0.475	1.109	0.076	0.036	0.549				
sucralfate	lorazepam	0.034	0.854	0.095	0.762	1.014	0.017	0.415	1.237				

risperidone	moxifloxacin_hydrochloride	0.133	1.012	-0.133	0.604	0.809	0.273			0.332	0.881	0.072
ramipril	nitrofurantoin_monohydrate_nitrofurantoin__m	0.085	0.859	0.121	0.825	0.707	0.009					
morphine_sulfate	ondansetron	0.001	0.848	-0.158	0.533	1.041	0.601					
morphine_sulfate	ondansetron_hydrochloride	0.386	0.923	-0.072	0.439	1.281	0.013	0.245	0.944			
sucralfate	ondansetron_hydrochloride	0.31	1.064	0.05	0.747	1.265	0.022	0.445	1.256			
insulin_glargine__recombinant	potassium_chloride	0.085	0.518	0.326	0.589	0.528	0.014	0.195	0.618			
clonidine_hydrochloride	potassium_chloride	0.071	0.662	0.39	0.724	0.617	0.002	0.106	0.651			
glyburide	potassium_chloride	0.03	0.678	0.596	1.057	0.68	0.118	0.001	0.575			
pioglitazone_hydrochloride	potassium_chloride	0.02	0.559	0.369	0.821	0.836	0.005	0.015	0.53	0.048	0.363	
morphine_sulfate	warfarin_sodium	0.134	0.92	0.338	1.014	0.879	0.702			0.017	0.518	0.166

Supplementary Model:

```
data{
  ///
      int <lower=1> N; // tot
      int <lower=1> Noutcome;
      real cancer_effect_sd; // prior variance for mean of eff
      real percancer_devsd; // prior variance for mean of eff
      real across_cancers_sd; // prior variance for mean of eff
      vector[N] expvals;
      vector[N] se;
      //int<lower=0, upper=1> outcome_is_cancer[Noutcome];
      int N_class;
      int<lower=1,upper=N_class> drug2[N];

      //int<lower=0, upper=1> cancer_indicator[N];
      int<lower=1, upper=Noutcome> outcome_id[N];


}
parameters{
      vector[N_class] eff_per_drug_cancer_tilde[Noutcome];
      real class_cancer_effect;
      real<lower=0> class_allcancer_sd;

      vector<lower=0>[Noutcome] outcome_sd_across_drugs; // for each outcome, the spread across
drugs might be different
  vector[Noutcome] class_per_cancer_effect_tilde;
}
transformed parameters{
      real Xeff_rep[N];
      vector[N_class] eff_percancer[Noutcome];
```

```
  vector[Noutcome] class_per_cancer_effect = class_per_cancer_effect_tilde * class_allcancer_sd +
class_cancer_effect;
      for(i in 1:Noutcome){
        eff_percancer[i] = eff_per_drug_cancer_tilde[i] * outcome_sd_across_drugs[i] +
class_per_cancer_effect[i];

      }
      for (i in 1:N){
           Xeff_rep[i] = eff_percancer[outcome_id[i]][drug2[i]] ;
        }
}

model {
  // these 2 are the effect & var of effect overall of cancer of drugs in that group
  class_cancer_effect ~ normal(0, cancer_effect_sd);
  class_allcancer_sd ~ normal(0,across_cancers_sd); // overall, how wide is the spread across
cancers

  // then these are the variations across cancers of that
  class_per_cancer_effect_tilde ~ normal(0, 1);

  // then we get to the per-drug variation from that for each cancer: the random deviations, and
what we scale these by per cancer
  outcome_sd_across_drugs ~ normal(0, percancer_devsd);
  for(i in 1:Noutcome){
      eff_per_drug_cancer_tilde[i] ~ normal(0,1); // each of these being random spread around the
per-cancer effect, across drugs in that class
  }
       if(run_estimation==1){
         expvals ~ normal(Xeff_rep, se);     }
 }
```

1. Hales, cm, Servais, j, Martin, cb & Kohen, d. Prescription drug use among adults aged 40–79 in the United States and Canada. in *NCHS Data Brief, no 347* (National Center for Health Statistics, 2019).
2. Investigators, W. G. for the W. H. I. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative Randomized Controlled Trial. *JAMA* **288**, 321–333 (2002).
3. Longo, J., Leeuwen, J. E. van, Elbaz, M., Branchard, E. & Penn, L. Z. Statins as Anticancer Agents in the Era of Precision Medicine. *Clin Cancer Res* **26**, 5791–5800 (2020).
4. Patel, C. J., Ji, J., Sundquist, J., Ioannidis, J. P. A. & Sundquist, K. Systematic assessment of pharmaceutical prescriptions in association with cancer risk: a method to conduct a population-wide medication-wide longitudinal study. *Sci Rep* **6**, 31308 (2016).
5. Pottegård, A. *et al.* Identification of Associations Between Prescribed Medications and Cancer: A Nationwide Screening Study. *EBioMedicine* **7**, 73–79 (2016).
6. Støer, N. C. *et al.* Drug Use and Cancer Risk: A Drug-Wide Association Study (DWAS) in Norway. *Cancer Epidemiol Biomarkers Prev* **30**, 682–689 (2021).
7. McDowell, R. D., Hughes, C., Murchie, P. & Cardwell, C. A systematic assessment of the association between frequently prescribed medicines and the risk of common cancers: a series of nested case-control studies. *BMC Med* **19**, 22 (2021).
8. Schuemie, M. J., Hripcsak, G., Ryan, P. B., Madigan, D. & Suchard, M. A. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *PNAS* 201708282 (2018) doi:10.1073/pnas.1708282114.
9. Shaddox, T. R., Ryan, P. B., Schuemie, M. J., Madigan, D. & Suchard, M. A. Hierarchical models for multiple, rare outcomes

using massive observational healthcare databases. *Statistical Analysis and Data Mining: The ASA Data Science Journal* **9**, 260–268 (2016).

10. Danaei, G., García Rodríguez, L. A., Cantero, O. F., Logan, R. W. & Hernán, M. A. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *Journal of Clinical Epidemiology* **96**, 12–22 (2018).
11. Danaei, G., Rodríguez, L. A. G., Cantero, O. F., Logan, R. & Hernán, M. A. Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res* **22**, 70–96 (2013).
12. VanderWeele, T. J. On the Relative Nature of Overadjustment and Unnecessary Adjustment: *Epidemiology* **20**, 496–499 (2009).
13. Hernán, M. A. & Robins, J. M. *Causal Inference*.
14. Schünemann, H., Hill, S., Guyatt, G., Akl, E. & Ahmed, F. The GRADE approach and Bradford Hill's criteria for causation. *Journal of epidemiology and community health* **65**, 392–5 (2011).
15. Melamed, R. D. Using indication embeddings to represent patient health for drug safety studies. *JAMIA Open* **3**, 422–430 (2020).
16. Melamed, R. D. & Rzhetsky, A. Patchwork of contrasting medication cultures across the USA. *Nature Communications* **9**, (2018).
17. Shahn, Z. *et al.* G-Computation and Hierarchical Models for Estimating Multiple Causal Effects From Observational Disease Registries With Irregular Visits. *AMIA Jt Summits Transl Sci Proc* **2019**, 789–798 (2019).