
Supplementary information

A deep learning framework for drug repurposing via emulating clinical trials on real-world patient data

In the format provided by the authors and unedited

Supplemental Table 1. The definition of coronary artery disease (CAD) from observational health data.

| | |
|------------------|--|
| PMID | 16159046, 26524702, 28008010 |
| Criteria | <ul style="list-style-type: none"> • A history of coronary revascularization in the EHR • Or, history of acute coronary syndrome, ischemic heart disease, or exertional angina |
| Diagnostic codes | <p>ICD-9 codes: 410* to 414*</p> <p>ICD-10 codes: The best approximation are the following codes: I20* Angina pectoris I21* Acute myocardial infarction I22* Subsequent ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction I23* Certain current complications following ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction (within the 28 day period) I24* Other acute ischemic heart diseases I25* Chronic ischemic heart disease</p> |

Supplemental Table 2. The definition of heart failure from observational health data.

| | |
|------------------|--|
| PMID | 26524702, 26687987, 21156884, 15606986 |
| Criteria | <p>Any one of the following:</p> <ol style="list-style-type: none"> 1. (ICD-9) billing code 2. (ICPC-2-R) diagnosis code 3. "CHF" on the patient's problem list (free text or ICD-9) |
| Diagnostic codes | <p>ICD-9 codes: 402.01, 402.11, 402.91, 428.xx</p> <p>OR: 402.01 Hypertensive heart disease, malignant with CHF 402.11 Hypertensive heart disease, benign with CHF 402.91 Hypertensive heart disease, NOS with CHF 404.01 Hypertensive heart/renal disease, malignant with CHF 404.03 Hypertensive heart/renal disease, malignant with CHF + renal failure 404.11 Hypertensive heart/renal disease, benign with CHF 0 (0) 404.13 Hypertensive heart/renal disease, benign with CHF + renal failure 404.91 Hypertensive heart/renal disease, NOS with CHF 404.93 Hypertensive heart/renal disease, NOS with CHF + renal failure 425.xx Cardiomyopathy 428.xx Heart failure</p> <p>ICD-10 codes: I11 I13 I50 I42</p> <p>ICPC-2-R code: K77</p> |

Supplemental Table 3. The definition of stroke from observational health data.

| | |
|------------------|---|
| PMID | 29202795 |
| Diagnostic codes | <p>ICD-9 codes: V12.54, 438.0–438.9</p> <p>ICD 10 codes: Z86.73 I60-I69 subarachnoid hemorrhage (I60); intracerebral hemorrhage (I61); cerebral infarction (I63); and other transient cerebral ischemic attacks and related syndromes and transient cerebral ischemic attack (unspecified) (G458 and G459),</p> |

Supplemental Table 4. Main results for all 55 repurposing drugs.

| Drug name | # User | # Non-user | Pre.unbalanced covariates | Post.unbalanced covariates | # Covariates | Post.unbalanced ratio | Pre. ATE | Post. ATE |
|---------------------|--------|------------|---------------------------|----------------------------|--------------|-----------------------|----------|-----------|
| atorvastatin | 13099 | 39297 | 16.560 | 26.200 | 1300 | 0.020 | -0.029 | -0.050 |
| metoprolol | 9730 | 29190 | 38.308 | 23.231 | 1270 | 0.018 | -0.023 | -0.043 |
| fenofibrate | 1352 | 4056 | 39.340 | 13.200 | 1038 | 0.013 | -0.051 | -0.038 |
| rosuvastatin | 2420 | 7260 | 24.020 | 9.620 | 1097 | 0.009 | -0.063 | -0.030 |
| hydrochlorothiazide | 2001 | 6003 | 32.500 | 15.320 | 1076 | 0.014 | -0.055 | -0.029 |
| amlodipine | 4613 | 13839 | 21.340 | 8.300 | 1180 | 0.007 | -0.050 | -0.026 |
| pravastatin | 2007 | 6021 | 11.260 | 9.640 | 1085 | 0.009 | -0.016 | -0.022 |
| simvastatin | 1605 | 4815 | 10.060 | 13.240 | 1044 | 0.013 | -0.032 | -0.020 |
| lisinopril | 5876 | 17628 | 17.960 | 25.000 | 1200 | 0.021 | -0.002 | -0.020 |
| valsartan | 1316 | 3948 | 24.940 | 13.740 | 1026 | 0.013 | 0.010 | -0.015 |
| diltiazem | 1044 | 3132 | 28.360 | 13.080 | 1007 | 0.013 | -0.010 | -0.013 |
| omeprazole | 1916 | 5748 | 31.080 | 15.220 | 1084 | 0.014 | -0.052 | -0.011 |
| losartan | 4822 | 14466 | 22.680 | 7.720 | 1187 | 0.006 | -0.015 | -0.007 |
| fluoxetine | 505 | 1515 | 104.500 | 46.240 | 932 | 0.050 | -0.064 | -0.005 |
| atenolol | 845 | 2535 | 42.460 | 22.460 | 974 | 0.023 | -0.082 | -0.005 |
| metformin | 3258 | 9774 | 29.700 | 15.300 | 1131 | 0.014 | -0.052 | -0.004 |
| nebivolol | 713 | 2139 | 49.960 | 28.500 | 958 | 0.030 | -0.083 | -0.003 |
| clopidogrel | 6488 | 19464 | 27.700 | 7.340 | 1212 | 0.006 | -0.014 | 0.013 |
| levothyroxine | 2637 | 7911 | 39.520 | 9.380 | 1131 | 0.008 | -0.034 | 0.014 |
| escitalopram | 1123 | 3369 | 56.040 | 15.460 | 1025 | 0.015 | -0.036 | 0.016 |
| gabapentin | 1117 | 3351 | 74.800 | 23.220 | 1041 | 0.022 | 0.002 | 0.016 |
| pantoprazole | 2508 | 7524 | 21.100 | 9.780 | 1114 | 0.009 | 0.005 | 0.019 |
| sertraline | 932 | 2796 | 60.980 | 24.140 | 1013 | 0.024 | -0.036 | 0.021 |

| | | | | | | | | |
|--------------------|------|-------|---------|--------|------|-------|--------|-------|
| benazepril | 566 | 1698 | 55.120 | 44.620 | 907 | 0.049 | -0.068 | 0.025 |
| bupropion | 779 | 2337 | 77.920 | 29.900 | 979 | 0.031 | -0.050 | 0.026 |
| aspirin | 709 | 2127 | 35.260 | 31.600 | 952 | 0.033 | -0.010 | 0.030 |
| isosorbide | 1482 | 4446 | 33.320 | 9.560 | 1039 | 0.009 | 0.045 | 0.034 |
| prasugrel | 1316 | 3948 | 41.500 | 18.340 | 1019 | 0.018 | -0.043 | 0.036 |
| trazodone | 527 | 1581 | 128.580 | 53.440 | 947 | 0.057 | -0.006 | 0.039 |
| ramipril | 887 | 2661 | 25.340 | 14.840 | 973 | 0.015 | 0.020 | 0.043 |
| olmesartan | 571 | 1713 | 73.260 | 45.400 | 933 | 0.049 | -0.075 | 0.047 |
| citalopram | 672 | 2016 | 56.420 | 30.440 | 960 | 0.032 | -0.041 | 0.060 |
| duloxetine | 932 | 2796 | 116.300 | 20.900 | 1011 | 0.021 | -0.043 | 0.068 |
| canagliflozin | 960 | 2880 | 98.900 | 50.040 | 993 | 0.050 | -0.053 | 0.073 |
| potassium chloride | 1110 | 3330 | 43.460 | 20.240 | 1016 | 0.020 | 0.169 | 0.090 |
| ezetimibe | 938 | 2814 | 67.900 | 21.220 | 992 | 0.021 | -0.049 | 0.090 |
| glipizide | 675 | 2025 | 63.000 | 45.240 | 945 | 0.048 | 0.003 | 0.095 |
| zolpidem | 550 | 1650 | 88.840 | 41.940 | 927 | 0.045 | -0.015 | 0.106 |
| esomeprazole | 446 | 1338 | 101.660 | 57.560 | 903 | 0.064 | -0.072 | 0.108 |
| glimepiride | 789 | 2367 | 70.820 | 38.380 | 979 | 0.039 | -0.034 | 0.112 |
| venlafaxine | 606 | 1818 | 113.980 | 58.320 | 953 | 0.061 | -0.055 | 0.116 |
| carvedilol | 3959 | 11877 | 38.280 | 8.140 | 1154 | 0.007 | 0.198 | 0.124 |
| ranolazine | 587 | 1761 | 54.780 | 42.040 | 927 | 0.045 | 0.036 | 0.134 |
| sitagliptin | 1104 | 3312 | 55.400 | 25.940 | 1013 | 0.026 | -0.044 | 0.155 |
| ticagrelor | 905 | 2715 | 45.360 | 29.160 | 979 | 0.030 | -0.002 | 0.162 |
| furosemide | 1545 | 4635 | 50.880 | 17.080 | 1064 | 0.016 | 0.301 | 0.179 |
| montelukast | 908 | 2724 | 82.480 | 27.400 | 996 | 0.027 | -0.022 | 0.181 |
| spironolactone | 1292 | 3876 | 70.620 | 12.920 | 1034 | 0.013 | 0.393 | 0.190 |
| allopurinol | 865 | 2595 | 84.520 | 26.580 | 976 | 0.027 | 0.025 | 0.197 |
| alprazolam | 492 | 1476 | 110.960 | 49.180 | 907 | 0.054 | 0.006 | 0.204 |
| oxycodone | 575 | 1725 | 127.480 | 50.980 | 947 | 0.054 | -0.001 | 0.289 |
| tamsulosin | 1137 | 3411 | 66.140 | 27.060 | 1026 | 0.026 | 0.006 | 0.311 |
| apixaban | 710 | 2130 | 81.040 | 41.380 | 963 | 0.043 | 0.168 | 0.332 |
| rivaroxaban | 945 | 2835 | 79.080 | 29.400 | 1002 | 0.029 | 0.102 | 0.392 |
| warfarin | 685 | 2055 | 95.760 | 34.720 | 952 | 0.036 | 0.234 | 0.540 |

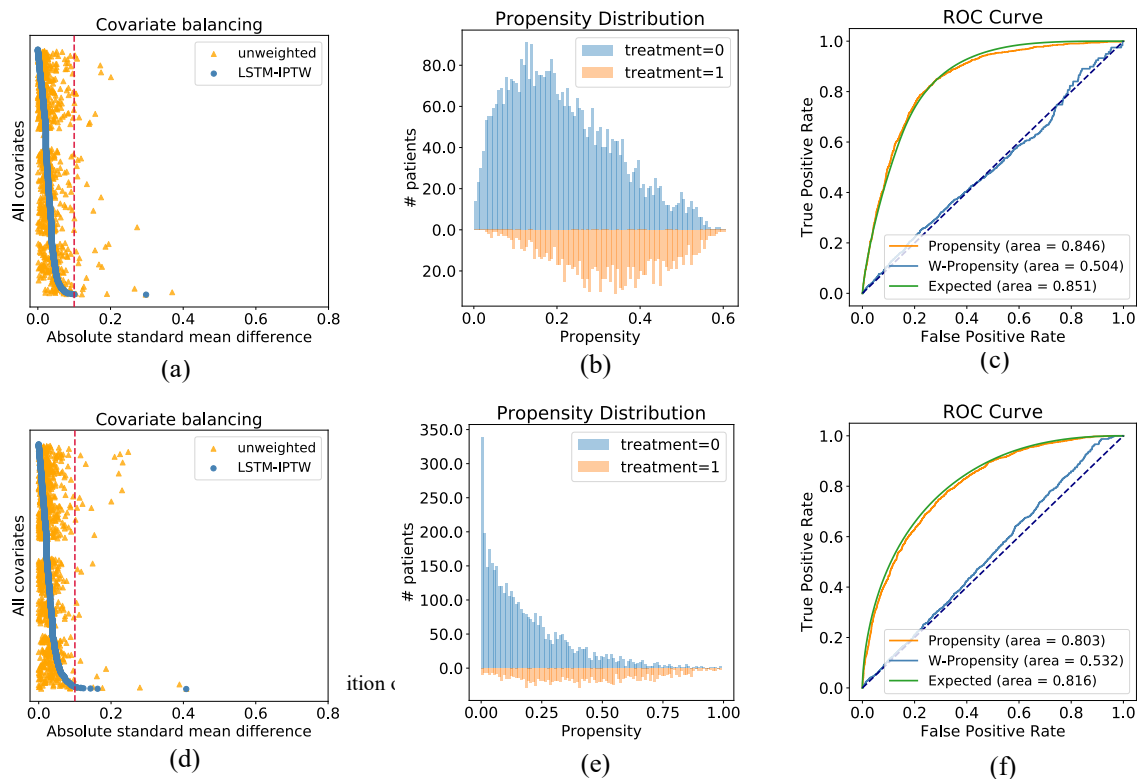
Supplemental Table 5. Main results for all 38 repurposing drug classes.

| Drug name | # User | # Non-user | Pre.unbalanced covariates | Post.unbalanced covariates | # Covariates | Post.unbalanced ratio | Pre. ATE | Post. ATE |
|-----------|--------|------------|---------------------------|----------------------------|--------------|-----------------------|----------|-----------|
| A02BA | 655 | 1965 | 49.450 | 10.500 | 557 | 0.019 | -0.025 | -0.028 |
| A02BC | 3812 | 10775 | 19.500 | 5.300 | 611 | 0.009 | -0.033 | -0.040 |

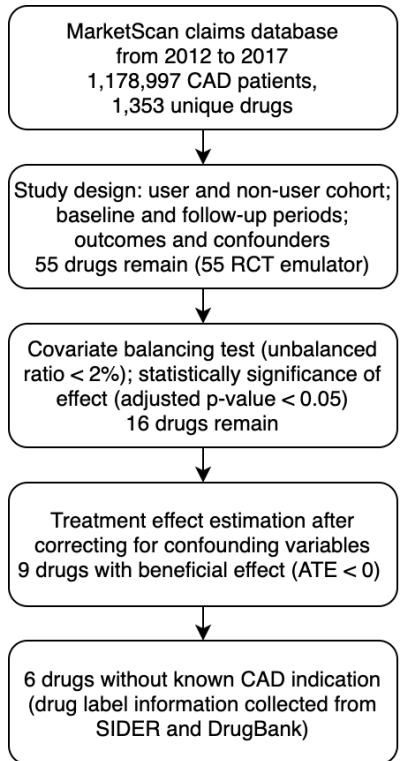
| | | | | | | | | |
|-------|-------|-------|--------|--------|-----|-------|--------|--------|
| A10AE | 597 | 1791 | 82.150 | 22.700 | 545 | 0.042 | 0.024 | -0.024 |
| A10BA | 3252 | 9756 | 32.000 | 14.650 | 604 | 0.024 | -0.065 | -0.086 |
| A10BB | 1373 | 4119 | 47.700 | 17.550 | 575 | 0.030 | -0.016 | -0.056 |
| A10BH | 1358 | 4074 | 57.550 | 29.850 | 573 | 0.052 | -0.039 | -0.091 |
| A10BJ | 735 | 2205 | 78.050 | 24.200 | 552 | 0.044 | -0.030 | -0.051 |
| A10BK | 1712 | 5136 | 77.700 | 30.400 | 584 | 0.052 | -0.039 | -0.075 |
| A11CC | 543 | 1629 | 79.600 | 13.050 | 556 | 0.024 | 0.030 | -0.030 |
| A12BA | 1150 | 3450 | 40.700 | 6.500 | 578 | 0.011 | 0.153 | 0.061 |
| B01AA | 677 | 1596 | 80.400 | 22.450 | 551 | 0.041 | 0.197 | 0.083 |
| B01AC | 8451 | 24429 | 19.100 | 4.900 | 632 | 0.008 | -0.017 | -0.030 |
| B01AF | 1619 | 4857 | 61.200 | 24.650 | 586 | 0.042 | 0.115 | -0.021 |
| C01DA | 1535 | 4605 | 37.600 | 4.900 | 581 | 0.008 | 0.053 | 0.011 |
| C01EB | 537 | 1611 | 61.600 | 12.150 | 532 | 0.023 | 0.019 | -0.008 |
| C03AA | 1989 | 5967 | 37.250 | 9.250 | 591 | 0.016 | -0.078 | -0.075 |
| C03CA | 1601 | 4803 | 48.200 | 6.350 | 588 | 0.011 | 0.293 | 0.128 |
| C03DA | 1395 | 4185 | 61.150 | 7.350 | 577 | 0.013 | 0.384 | 0.151 |
| C05AE | 1040 | 3120 | 44.150 | 8.300 | 569 | 0.015 | -0.011 | -0.023 |
| C07AA | 657 | 1971 | 63.800 | 11.850 | 549 | 0.022 | 0.053 | -0.010 |
| C07AB | 10359 | 28354 | 16.368 | 7.000 | 636 | 0.011 | -0.039 | -0.053 |
| C07AG | 4040 | 12120 | 23.850 | 4.250 | 611 | 0.007 | 0.187 | 0.056 |
| C08CA | 4801 | 14403 | 19.300 | 5.450 | 624 | 0.009 | -0.041 | -0.063 |
| C09AA | 7016 | 21048 | 15.650 | 10.650 | 629 | 0.017 | -0.007 | -0.047 |
| C09CA | 5895 | 17685 | 14.050 | 5.200 | 628 | 0.008 | -0.016 | -0.036 |
| C10AA | 11730 | 30838 | 29.600 | 15.150 | 641 | 0.024 | -0.026 | -0.064 |
| C10AB | 1412 | 4236 | 38.350 | 8.950 | 572 | 0.016 | -0.047 | -0.059 |
| C10AX | 979 | 2937 | 61.100 | 12.250 | 566 | 0.022 | -0.043 | -0.049 |
| G04CA | 1326 | 3978 | 56.800 | 25.100 | 580 | 0.043 | -0.019 | -0.062 |
| H03AA | 2641 | 7923 | 42.550 | 13.200 | 605 | 0.022 | -0.043 | -0.055 |
| M04AA | 949 | 2847 | 69.550 | 14.550 | 558 | 0.026 | 0.016 | -0.029 |
| N02AA | 607 | 1821 | 99.800 | 24.250 | 553 | 0.044 | -0.040 | -0.061 |
| N03AX | 1719 | 5157 | 69.750 | 12.450 | 593 | 0.021 | 0.014 | -0.032 |
| N05BA | 621 | 1863 | 86.350 | 13.400 | 549 | 0.024 | -0.025 | -0.013 |
| N05CF | 598 | 1794 | 70.500 | 10.300 | 550 | 0.019 | -0.026 | -0.036 |
| N06AB | 2793 | 8379 | 41.700 | 8.300 | 612 | 0.014 | -0.045 | -0.053 |
| N06AX | 2279 | 6837 | 65.450 | 11.350 | 601 | 0.019 | -0.054 | -0.066 |
| R03DC | 899 | 2697 | 65.250 | 11.300 | 563 | 0.020 | -0.035 | -0.034 |

Supplemental Table 6. Main results for all 7 repurposing drug combinations.

| Drug name | # User | # Non-user | Post unbalanced ratio | Pre.ATE | Post.ATE | Adjusted P-value |
|----------------------------|--------|------------|-----------------------|---------|----------|------------------|
| Metoprolol + Clopidogrel | 1237 | 3711 | 0.010 | -0.034 | -0.028 | < 0.05 |
| Metoprolol + Atorvastatin | 2158 | 6474 | 0.014 | -0.045 | -0.024 | < 0.05 |
| Lisinopril + Atorvastatin | 1145 | 3435 | 0.015 | -0.002 | -0.018 | < 0.05 |
| Lisinopril + Clopidogrel | 630 | 1890 | 0.013 | -0.018 | -0.012 | > 0.1 |
| Metoprolol + Lisinopril | 962 | 2886 | 0.011 | -0.028 | -0.012 | > 0.1 |
| Clopidogrel + Atorvastatin | 1477 | 4431 | 0.007 | -0.019 | 0.008 | > 0.1 |
| Carvedilol + Atorvastatin | 860 | 2580 | 0.011 | 0.124 | 0.112 | < 0.05 |



Supplemental Figure 1. Performance comparison of LSTM-IPTW and LR-IPTW on case drug: fenofibrate (*without* known CAD indication). The three figures on the top are results obtained from LSTM-IPTW, and the figures on the bottom are from LR-IPTW. Figure (a) and Figure (d) show the absolute SMD of each covariate in the original data (orange triangles) and in the weighted data (blue circles). Figure (b) and Figure (e) show the distribution of estimated propensity scores over user (orange area) and non-user (blue area) cohorts. Figure (c) and Figure (f) show the ROC curves for the propensity model (orange), expected value (green) and weighted propensity (blue).



Supplemental Figure 2. Flowchart of data collection and study process of identifying repurposed drug candidates