

eText

1	Amendments to the study protocol.	2
2	Covariates in the multivariable logistic regression analysis.	2
3	Comparison of bias scatter plots and Davies' existing bias plots.	3
4	Fullfillment of IV study reporting guidelines.	3
5	Conversion of relative risk estimates to additional cases per 1000 treated.	4

eFigures

1	Study design diagram.	5
2	Decision tree for outcome definitions.	5
3	Attrition of patients in the analysis cohort.	6
4	Instrument strength.	7
5	Multivariable logistic regression results.	8
6	Bias scatter plot for covariates in the any dementia analysis.	8

1 Amendments to the study protocol.

We made two amendments to the original study protocol (Walker et al, 2016):

- Exposure: ‘Centrally acting antihypertensives’ are primarily used for acute events, while ‘Loop diuretics’ are primarily used for heart failure, and so have been excluded from the analysis. We have also combined ‘Potassium-sparing diuretics and aldosterone antagonists’ and ‘Thiazides and related diuretics’ into a single category titled ‘Diuretics’ as prescriptions for the former in the data extract were rare.
- Control: Instead of using beta-adrenoceptor blocking drugs as our reference drug class for all analyses, we have used each drug class as the reference drug class in turn and presented all of the results in a matrix.

2 Covariates in the multivariable logistic regression analysis.

Previous history of coronary heart disease, coronary-bypass surgery or cerebrovascular disease.

Presence of one or more relevant Read codes on record.

Chronic illness, including cancer and arthritis.

Charlson index implemented using Read code lists. (Charlson et al, 1987; Khan et al, 2010) Code lists based on those by Taylor et al. (Taylor et al, 2017)

Socioeconomic position.

2010 English Index of Multiple Deprivation (IMD) at the ‘twentile’ level, where 1 represents the least deprived and 20 the most deprived.

Consultation rate.

Calculated by dividing the total number of clinic visits by the length of the patient record prior to the index date to give an average annual rate.

Alcohol status.

Recorded value (current, former or never).

Smoking status.

Most recent of recorded value (current, former or never) or Read code indicating a recorded value. Code lists based on those by Wright et al. (Wright et al, 2017)

Bodt mass index.

Recorded value if available, or a calculated value using the last recorded height and weight measurements. Measurements taken before the age of 25 were excluded to ensure adult measurements were used.

3 Comparison of bias scatter plots and Davies' existing bias plots.

Bias scatter plots contain the same information as those proposed by Davies et al but differ in their presentation. Instead of using a discrete y-axis to indicate each covariate, bias scatter plots facet by covariate. This allows you to plot the multivariable linear regression analysis estimate (absolute value) on the x-axis and the instrumental variable analysis estimate (absolute value) on the y-axis instead of as jittered points using a shared axis. The motivation behind altering the presentation is so that multiple analyses related to an outcome can be presented simultaneously. This is particularly beneficial when it is infeasible to present bias plots for each analysis, such as in this paper. Considering the facets together provides insight into the overall bias in the IV analysis in the same way that looking across the series of points presented on Davies' original bias plot does. A further advantage of the bias scatter plot presentation is that the free scales mean that both continuous and binary covariates can be presented on a single plot.

4 Fullfillment of IV study reporting guidelines.

State which population target parameter the study aims to estimate (eg, local average treatment effect, effect of treatment on the treated) and the assumptions on which it depends (eg, monotonicity or no effect modification).

The following statements are made in the paper: “the results were interpreted as the effect among patients whose prescription was affected by their physicians’ preference (known as the local average treatment effect)” and “Instrumental variable analysis requires that the instrument: (i) be associated with the exposure of interest; (ii) affect the outcome only through its effect on the exposure of interest; and (iii) have no common causes with the outcome. To obtain a point estimate for this analysis, we also make a fourth assumption of monotonicity.”

Report the association of instruments and exposure using a partial F-statistic.

The Cragg-Donald F statistic has been presented alongside the results for each of our analyses.

Report and test the association of observed potential confounding factors with both the exposure and the instrument.

See section ‘Assessment of bias’ in the paper.

With multiple instruments report the test for overidentifying restrictions, ie, the Sargan or the Hansen test.

See section ‘Sensitivity analyses’ in the paper.

For binary outcomes, exposures, and instruments, report a tabulation of the frequencies of each combination of instrument, exposure, and outcome, so readers can reconstruct basic results.

See eTable 5.

When using generalized linear models with binary outcomes, always use robust (sandwich estimators) or bootstrapped standard errors and take clustering of study participants into account where necessary.

The analysis used Stata’s ivreg2 command with ‘robust’ specified and clustering according to the physicians’ staff ID. The analysis was conducted in Stata version 15MP (StataCorp; College Station, TX).

5 Conversion of relative risk estimates to additional cases per 1000 treated.

The Larsson et al meta-analysis reports a number of relative risk estimates that can be compared to the results from this study. However, our study presents the additional cases per 1000 treated so we must convert the relative risk estimates to be on the same scale before comparing them. The following derivation was based on the details provided at <https://bestpractice.bmj.com/info/toolkit/learn-ebm/how-to-calculate-risk/>.

To calculate the additional cases per 1000 treated, we first calculate the number needed to treat (NNT). Consider the following formula that uses absolute risk reduction (ARR):

$$NNT = \frac{1}{ARR}$$

Another way to express ARR uses the absolute risk of events in the control group (ARC) and the absolute risk of events in the treatment group (ART):

$$ARR = ARC - ART$$

Therefore:

$$NNT = \frac{1}{ARC - ART}$$

Relative risk can also be expressed in terms of ARC and ART by combining the following formulas concerning relative risk reduction (RRR):

$$RRR = 1 - RR$$

$$RRR = \frac{ARC - ART}{ARC}$$

Consequently:

$$\begin{aligned} 1 - RR &= \frac{ARC - ART}{ARC} \\ (1 - RR) ARC &= ARC - ART \end{aligned}$$

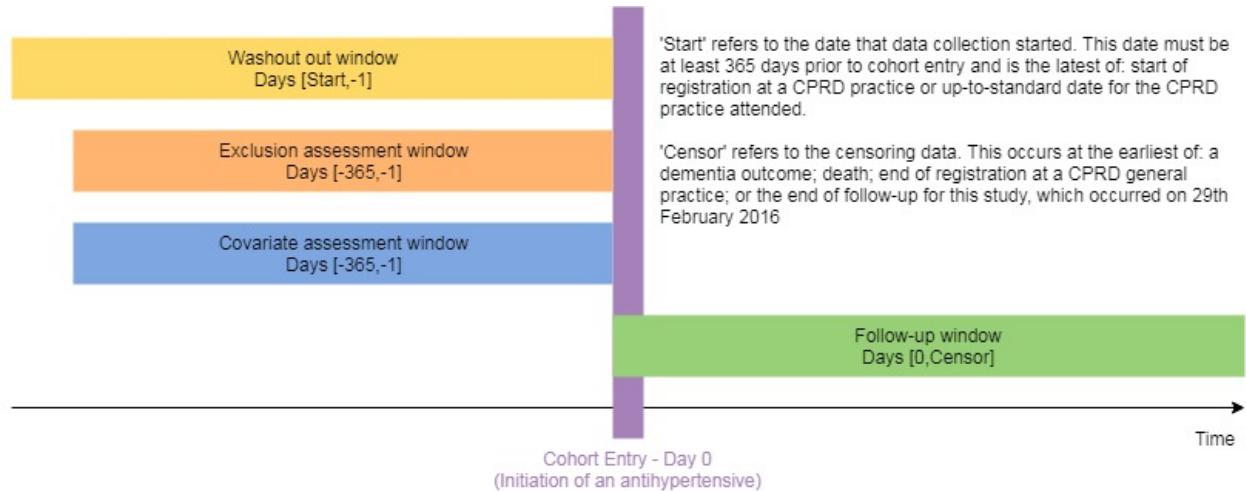
Ultimately, this means NNT can be expressed by RR as follows:

$$NNT = \frac{1}{(1 - RR) ARC}$$

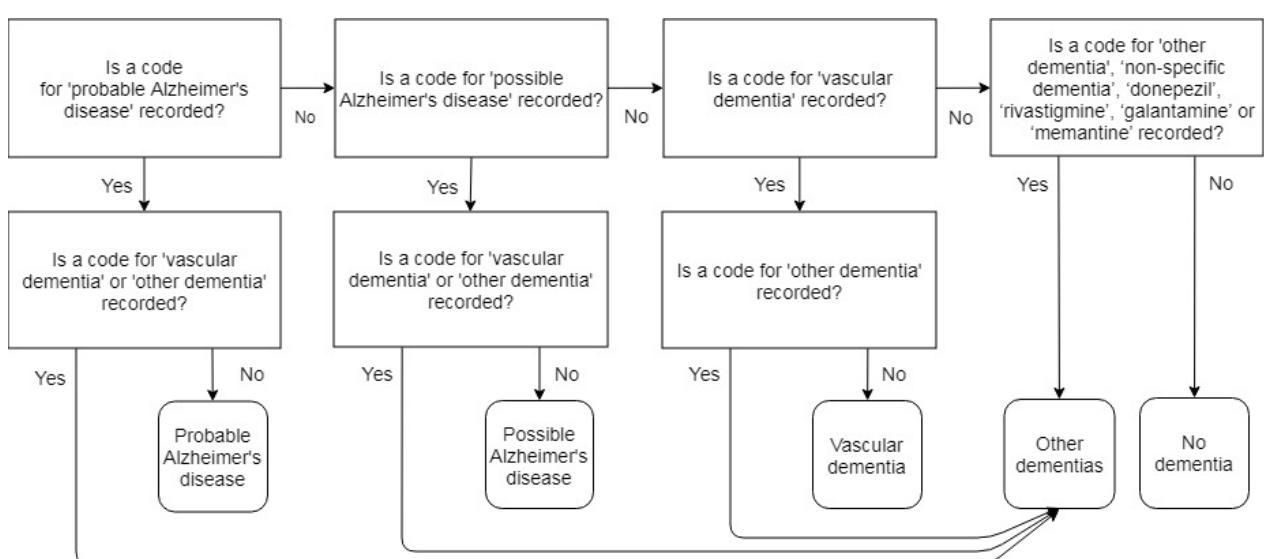
Finally, once we have obtained NNT, we can then calculate the additional cases per 1000 treated according to the following equation:

$$\text{Additional cases per 1000 treated} = -1 \left(\frac{1000}{NNT} \right)$$

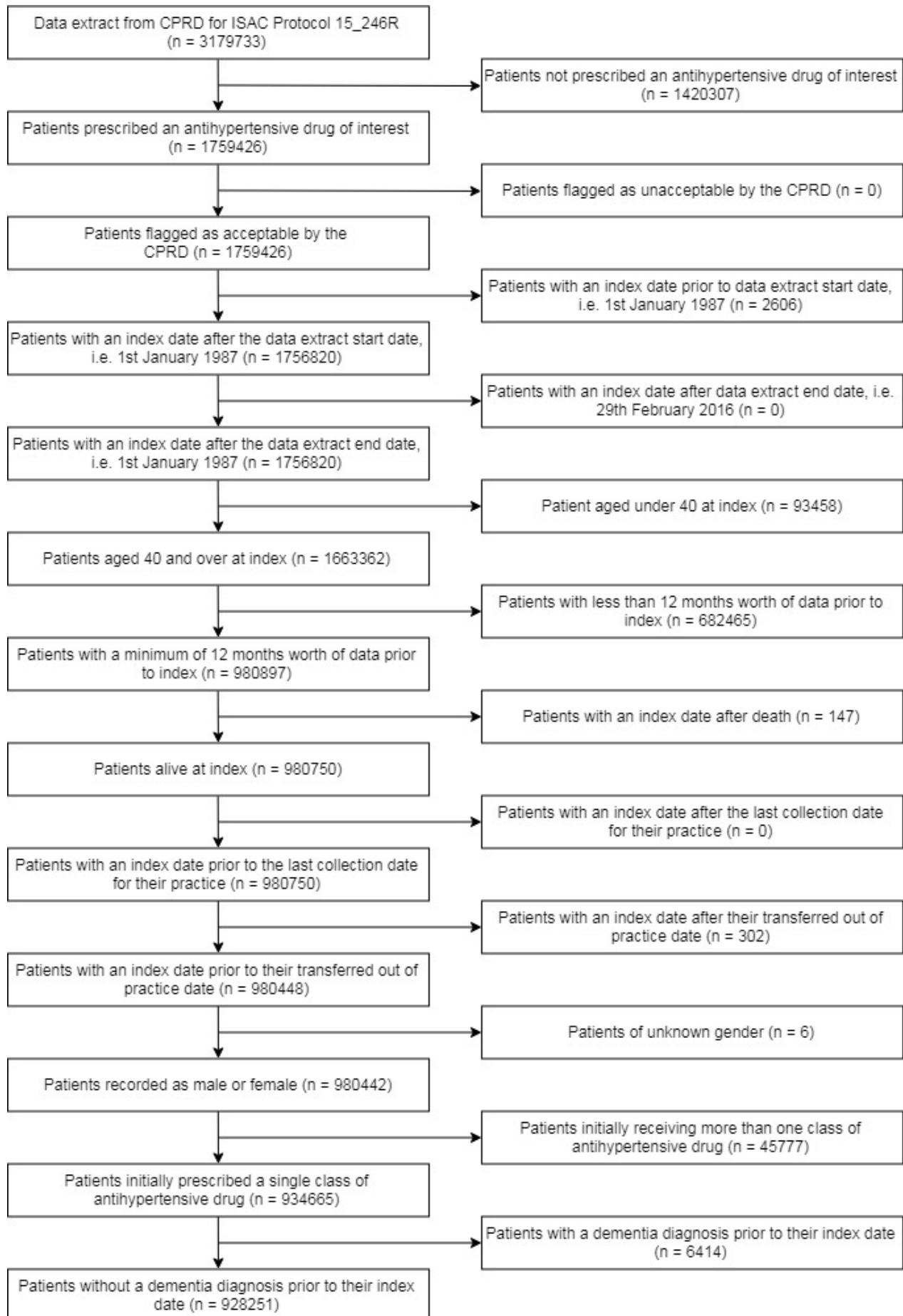
eFigure 1: Study design diagram.



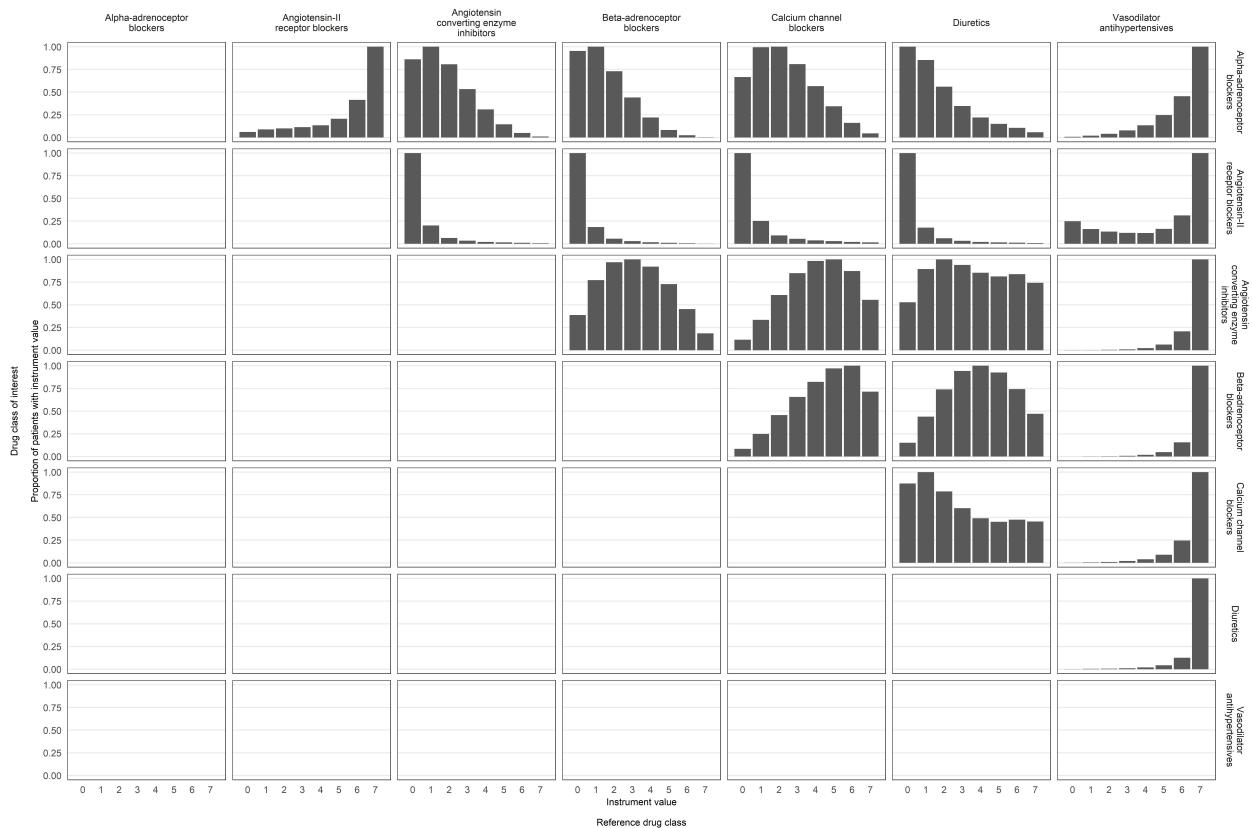
eFigure 2: Decision tree for outcome definitions.



eFigure 3: Attrition of patients in the analysis cohort.

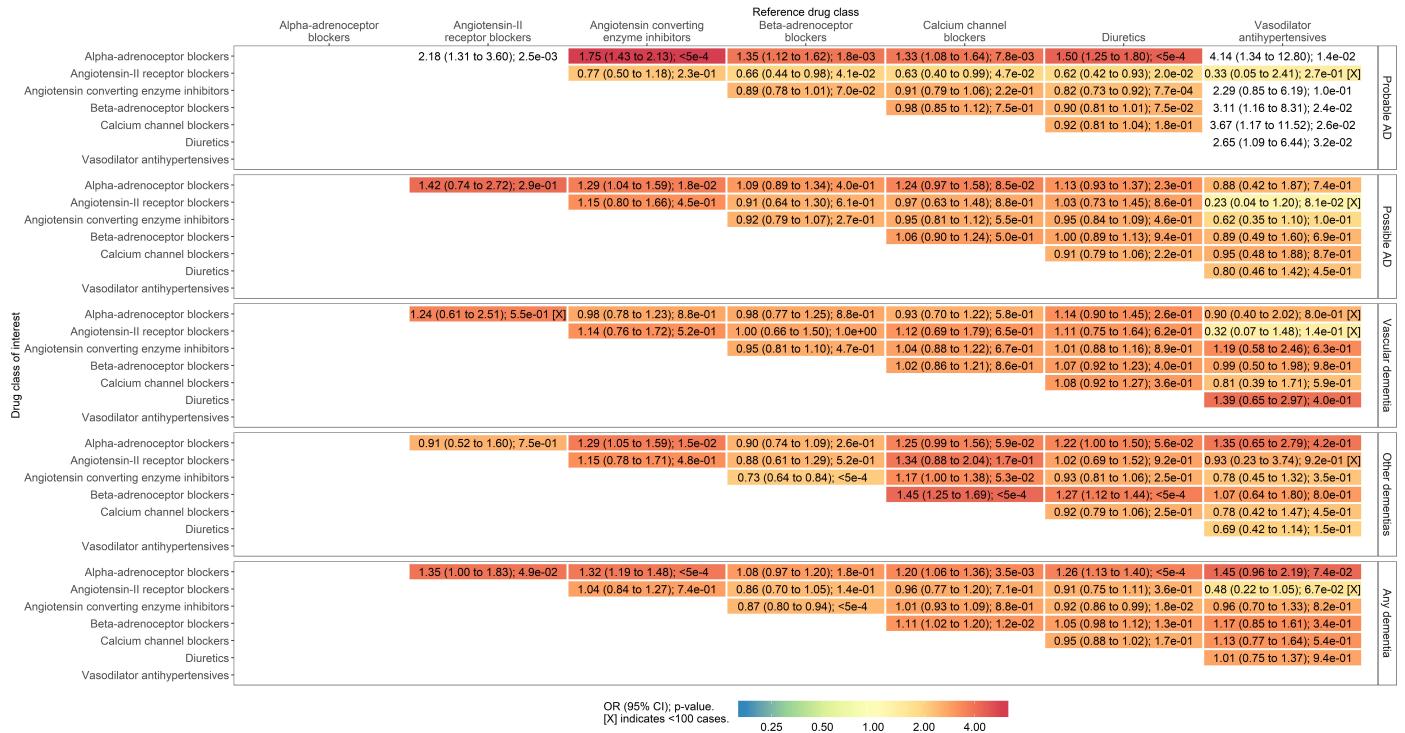


eFigure 4: Instrument strength.



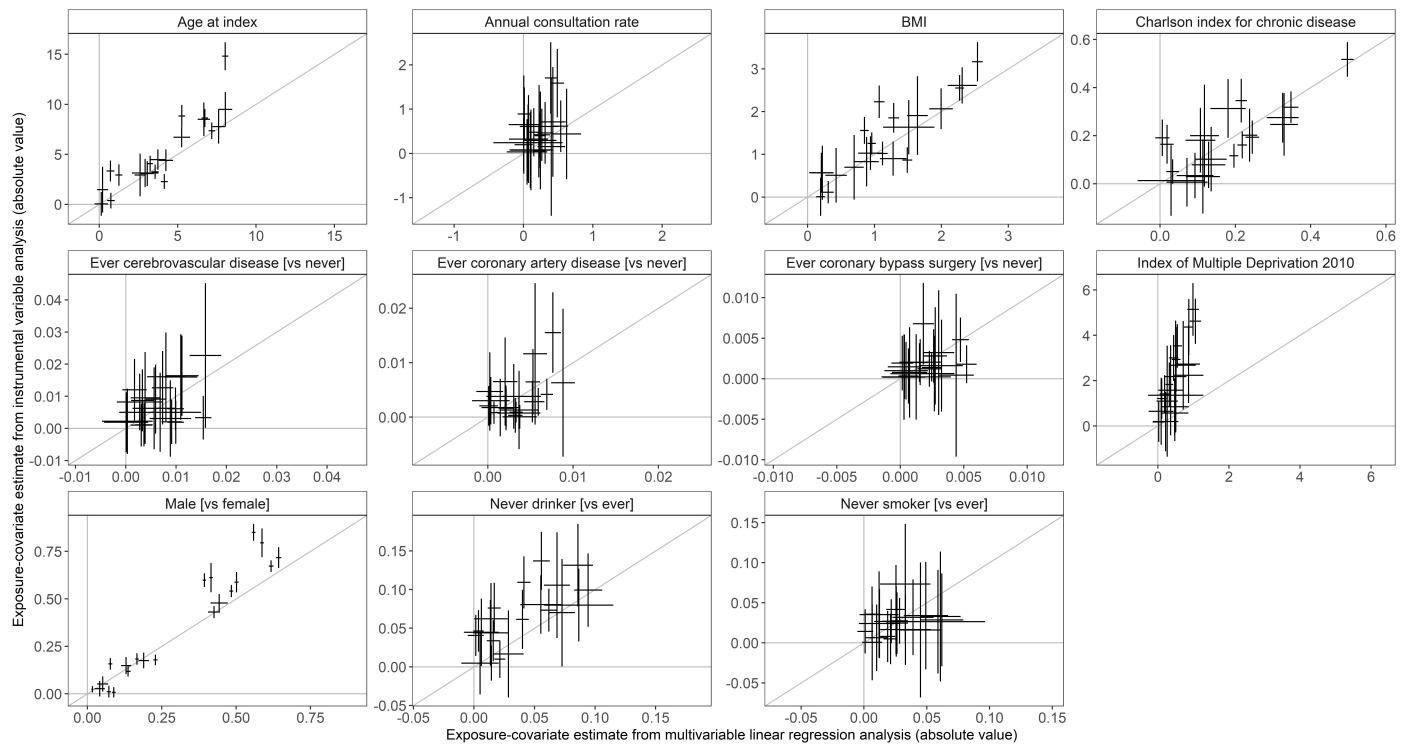
This plot is for the outcome 'any dementia', which is representative of the results obtained for all outcomes. The instrument is the number of times the treatment of interest has been prescribed to the previous seven patients, who were seen by the same physician and received either the treatment of interest or the reference drug.

eFigure 5: Multivariable logistic regression results.



For a drug class of interest, i.e. a ‘row’, estimates below zero indicate a protective effect of that drug class. For a reference drug class, i.e. a ‘column’, estimates above zero indicate a protective effect of that drug class. This matrix is antisymmetric, so the values not presented are the negation of those that are presented.

eFigure 6: Bias scatter plot for covariates in the any dementia analysis.



Each point on a scatter plot represents an individual analysis with the outcome ‘any dementia’. This plot is representative of the results obtained for all outcomes.