# The case-control study design


MANGER
JOHNSNOW

"On proceeding to the spot, I found that nearly all the deaths had taken place within a short distance of the [Broad Street] pump. There were only ten deaths in houses situated decidedly nearer to another street-pump."

John Snow
(the one that actually knew something…)

# Smoking and lung cancer?

3 of 86 male cancer patients were non-smokers

14 of 86 of healthy men were non-smokers

Müller FH, *Z. Krebsforsch* (1939); 49:57

# Smoking and lung cancer?

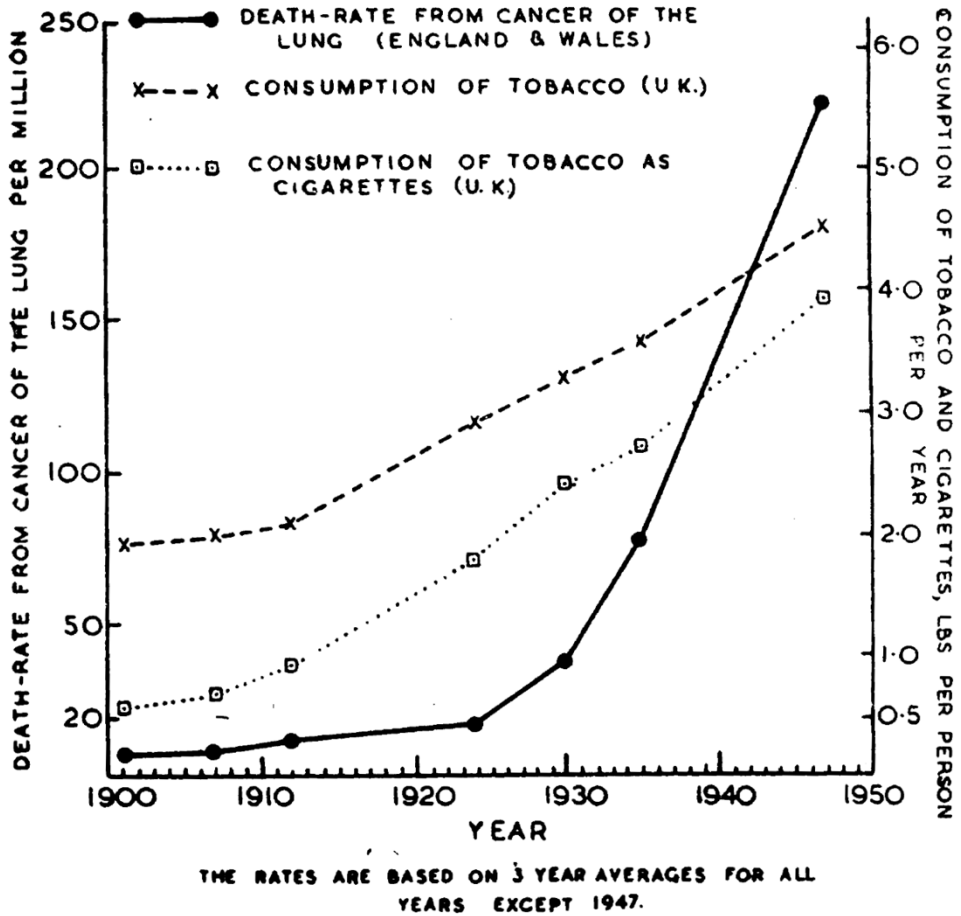
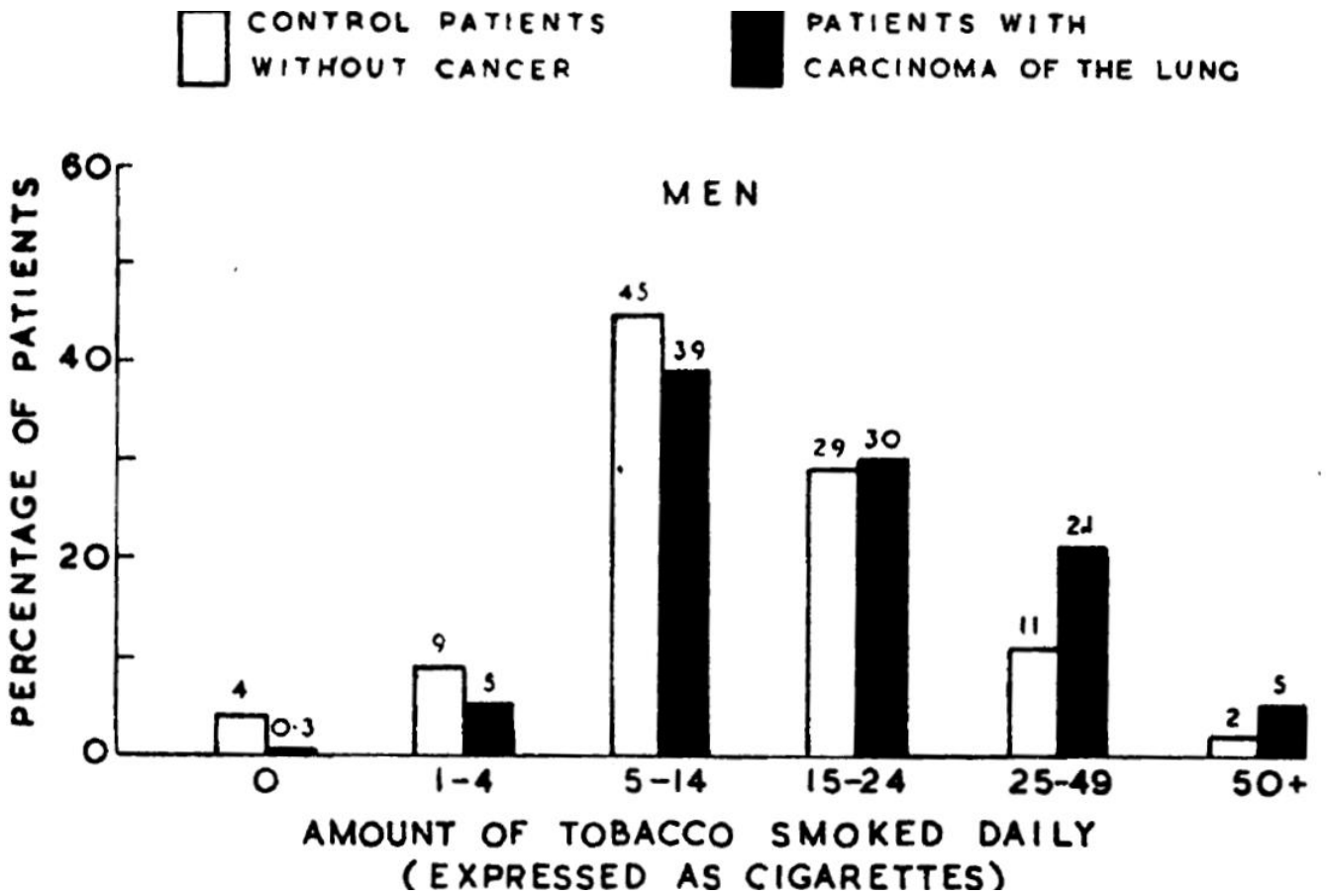



FIG. 2.—Death rate from cancer of the lung and rate of consumption of tobacco and cigarettes.

Doll & Hill. Br Med J 1950;2:739-48

# Smoking and lung cancer?




Doll & Hill. Br Med J 1950;2:739-48

**Cohort study**

A group of subjects using the drug under scrutiny and a group of non-users are followed over time with respect to the development of a certain outcome.

**Case-control study**

Subjects with a certain outcome (cases) and subjects without this outcome (controls) are mapped according to use of the drug under scrutiny.

# Cohort study

10,000 girls aged 20-25 years using 'the pill' are followed for three years.

Among these girls, 200 incident cases of deep vein thrombosis are recorded.

Among 20,000 girls NOT using 'the pill' (but same age and follow-up), 100 incident cases of deep vein thrombosis are recorded.

# Case-control study

300 girls aged 20-25 with incident deep vein thrombosis are identified. Among these girls, 80% had used ‘the pill’

Another 300 girls of the same age that have no record of deep vein thrombosis are identified. Among these girls, 50% have used ‘the pill’.

# Odds ratio

| | DVT Y | DVT N |
|---|---|---|
| **The pill Y** | 240 | 150 |
| **The pill N** | 60 | 150 |

$$OR = \frac{\left({}^{240}/_{60}\right)}{\left({}^{150}/_{150}\right)} = 4$$

# … but why!?

Use of appetite-suppressant drugs causes primary pulmonary hypertension

Relative risk ≈ 20

Baseline IR: 2 / 1 000000 person-years

If ALL Danes (≈6 mill) used these drugs, how many cases would I expect per year?

What if there was ”only” 100 000 users?

Abenhaim et al. NEJM 1996

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

# Identification of Associations Between Prescribed Medications and Cancer: A Nationwide Screening Study

Anton Pottegård [a,*], Søren Friis [b], René dePont Christensen [a], Laurel A. Habel [c], Joshua J. Gagne [d], Jesper Hallas [a]

[a] Clinical Pharmacology, Department of Public Health, University of Southern Denmark, Odense, Denmark
[b] Danish Cancer Society Research Center, Danish Cancer Society, Copenhagen Ø, Denmark
[c] Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA
[d] Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

ARTICLE INFO

*Article history:*
Received 12 January 2016
Received in revised form 11 March 2016
Accepted 11 March 2016
Available online xxxx

*Keywords:*
Cancer
Carcinogenicity
Chemoprevention
Drug evaluation
Pharmacology
Screening
Pharmacoepidemiology
Denmark

ABSTRACT

*Purpose:* We present a systematic screening for identifying associations between prescribed drugs and cancer risk using the high quality Danish nationwide health registries.
*Methods:* We identified all patients (cases) with incident cancer in Denmark during 2000–2012 (n = 278,485) and matched each case to 10 controls. Complete prescription histories since 1995 were extracted. Applying a two-phased case–control approach, we first identified drug classes or single drugs associated with an increased or decreased risk of 99 different cancer types, and further evaluated potential associations by examining specificity and dose–response patterns.
*Findings:* 22,125 drug–cancer pairs underwent evaluation in the first phase. Of 4561 initial signals (i.e., drug–cancer associations), 3541 (78%) failed to meet requirements for dose–response patterns and specificity, leaving 1020 eligible signals. Of these, 510 signals involved the use of single drugs, and 33% (166 signals) and 67% (344 signals) suggested a reduced or an increased cancer risk, respectively. While a large proportion of the signals were attributable to the underlying conditions being treated, our algorithm successfully identified well-established associations, as well as several new signals that deserve further investigation.
*Conclusion:* Our results provide the basis for future targeted studies of single associations to capture novel carcinogenic or chemopreventive effects of prescription drugs.


## 1. Introduction

cancer requires at least five years of regular use (Chan et al., 2012; Cuzick et al., 2015). Traditional approaches in pharmacovigilance

| 1 | Cancer | ATC | Drugname | Cases | Controls | OR | ORAll | p |
|---|---|---|---|---|---|---|---|---|
| 233 | Vulva and vagina (Squamos cell carcinoma) | D07AC01 | Betamethasone | 21 / 715 | 106 / 7,510 | 1.84 (1.13-3.00) | 1.07 | 0.01 |
| 234 | Vulva and vagina (Other) | G03CA03 | Estradiol | 50 / 157 | 255 / 1,856 | 2.39 (1.67-3.42) | 1.03 | 0.03 |
| 235 | Cervix uteri (Squamos cell carcinoma) | C09CA03 | Valsartan | 10 / 3,197 | 58 / 31,971 | 1.71 (0.87-3.35) | 1.02 | 0.09 |
| 236 | Cervix uteri (Squamos cell carcinoma) | G02BB01 | Vaginal ring with progestoge | 11 / 3,188 | 56 / 31,911 | 2.03 (1.05-3.90) | 0.96 | 0.01 |
| 237 | Cervix uteri (Squamos cell carcinoma) | L04AX01 | Azathioprine | 16 / 3,188 | 57 / 31,973 | 2.75 (1.57-4.81) | 1.34 | 0.08 |
| 238 | Cervix uteri (Adenocarcinoma) | N06AB03 | Fluoxetine | 11 / 709 | 58 / 7,123 | 1.88 (0.97-3.64) | 1.07 | 0.05 |
| 239 | Cervix uteri (Other) | C08CA02 | Felodipine | 10 / 381 | 32 / 3,858 | 3.58 (1.68-7.61) | 1.03 | 0.05 |
| 240 | Cervix uteri (Other) | R03AC02 | Salbutamol | 11 / 376 | 66 / 3,782 | 1.91 (0.98-3.72) | 1.12 | 0.01 |
| 241 | Corpus uteri (Adenocarcinoma, endometrioid) | A10BB12 | Glimepiride | 104 / 4,977 | 593 / 50,460 | 1.87 (1.51-2.33) | 0.95 | 0.09 |
| 242 | Corpus uteri (Adenocarcinoma, endometrioid) | B03BB01 | Folic acid | 26 / 5,070 | 154 / 50,639 | 1.72 (1.13-2.61) | 1.11 | <0.01 |
| 243 | Corpus uteri (Adenocarcinoma, endometrioid) | C02CA01 | Prazosin | 16 / 5,110 | 69 / 51,203 | 2.32 (1.35-4.01) | 0.98 | 0.07 |
| 244 | Corpus uteri (Adenocarcinoma, endometrioid) | C03AB01 | Bendroflumethiazide and pc | 886 / 3,738 | 6,961 / 39,620 | 1.38 (1.28-1.50) | 1.03 | 0.09 |
| 245 | Corpus uteri (Adenocarcinoma, endometrioid) | C03DA01 | Spironolactone | 79 / 5,004 | 520 / 50,341 | 1.57 (1.23-2.00) | 1.08 | 0.07 |
| 246 | Corpus uteri (Adenocarcinoma, endometrioid) | C03DB01 | Amiloride | 12 / 5,116 | 29 / 51,252 | 4.19 (2.14-8.22) | 1.09 | 0.07 |
| 247 | Corpus uteri (Adenocarcinoma, endometrioid) | C03EB01 | Furosemide and potassium | 11 / 5,117 | 45 / 51,213 | 2.52 (1.30-4.87) | 0.95 | 0.07 |
| 248 | Corpus uteri (Adenocarcinoma, endometrioid) | C09CA02 | Eprosartan | 10 / 5,119 | 53 / 51,216 | 1.92 (0.97-3.78) | 1.13 | 0.08 |
| 249 | Corpus uteri (Adenocarcinoma, endometrioid) | C09CA04 | Irbesartan | 46 / 5,072 | 259 / 50,881 | 1.82 (1.32-2.49) | 1.07 | <0.01 |
| 250 | Corpus uteri (Adenocarcinoma, endometrioid) | D07XC01 | Betamethasone | 13 / 5,066 | 73 / 50,777 | 1.79 (0.99-3.23) | 0.97 | 0.04 |
| 251 | Corpus uteri (Adenocarcinoma, endometrioid) | G03CX01 | Tibolone | 160 / 4,935 | 459 / 50,613 | 3.64 (3.03-4.38) | 1.28 | 0.02 |
| 252 | Corpus uteri (Adenocarcinoma, endometrioid) | G03DC02 | Norethisterone | 37 / 4,974 | 207 / 50,454 | 1.77 (1.24-2.51) | 1.30 | 0.07 |
| 253 | Corpus uteri (Adenocarcinoma, endometrioid) | G03FB01 | Norgestrel and estrogen | 65 / 5,040 | 312 / 50,785 | 2.09 (1.60-2.74) | 1.26 | 0.02 |
| 254 | Corpus uteri (Adenocarcinoma, endometrioid) | M04AA01 | Allopurinol | 73 / 5,027 | 340 / 50,715 | 2.19 (1.69-2.83) | 1.10 | 0.02 |
| 255 | Corpus uteri (Adenocarcinoma, endometrioid) | N02CC06 | Eletriptan | 15 / 5,111 | 65 / 51,162 | 2.33 (1.33-4.08) | 0.96 | 0.06 |
| 256 | Corpus uteri (Adenocarcinoma, endometrioid) | S01GX09 | Olopatadine | 12 / 5,095 | 73 / 50,951 | 1.65 (0.89-3.03) | 0.84 | 0.02 |
| 257 | Corpus uteri (Adenocarcinoma, other) | J01EB02 | Sulfamethizole | 23 / 783 | 142 / 8,020 | 1.65 (1.04-2.60) | 1.00 | 0.06 |
| 258 | Corpus uteri (Adenocarcinoma, other) | M01AB08 | Etodolac | 16 / 899 | 101 / 8,989 | 1.62 (0.94-2.79) | 1.02 | <0.01 |
| 259 | Corpus uteri (Sarcomas) | A10BB12 | Glimepiride | 14 / 558 | 73 / 5,646 | 2.56 (1.38-4.74) | 0.95 | 0.09 |
| 260 | Corpus uteri (Sarcomas) | S01EE01 | Latanoprost | 11 / 562 | 47 / 5,679 | 2.44 (1.25-4.79) | 0.94 | 0.07 |
| 261 | Corpus uteri (Adenocarcinoma, serous) | A10BA02 | Metformin | 22 / 396 | 155 / 4,005 | 1.52 (0.93-2.48) | 0.95 | 0.10 |
| 262 | Corpus uteri (Adenocarcinoma, serous) | C01AA05 | Digoxin | 14 / 401 | 75 / 4,089 | 2.01 (1.11-3.65) | 1.07 | 0.04 |
| 263 | Corpus uteri (Other) | D07AC01 | Betamethasone | 11 / 489 | 51 / 4,975 | 2.09 (1.08-4.04) | 1.07 | 0.08 |

Pottegård et al. EBioMedicine 2016 May; 7:73-9

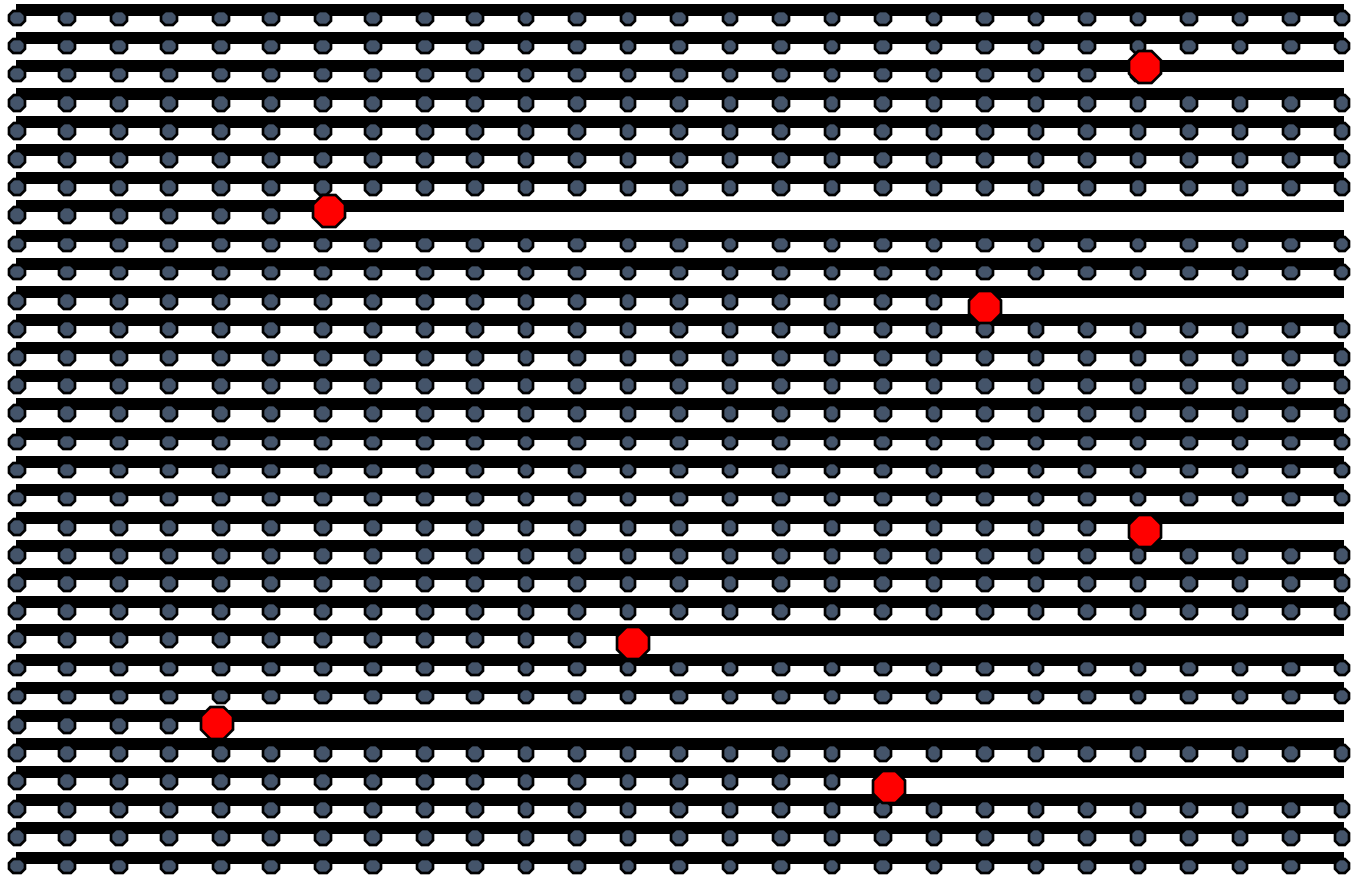
# The difficult part…

## Source population

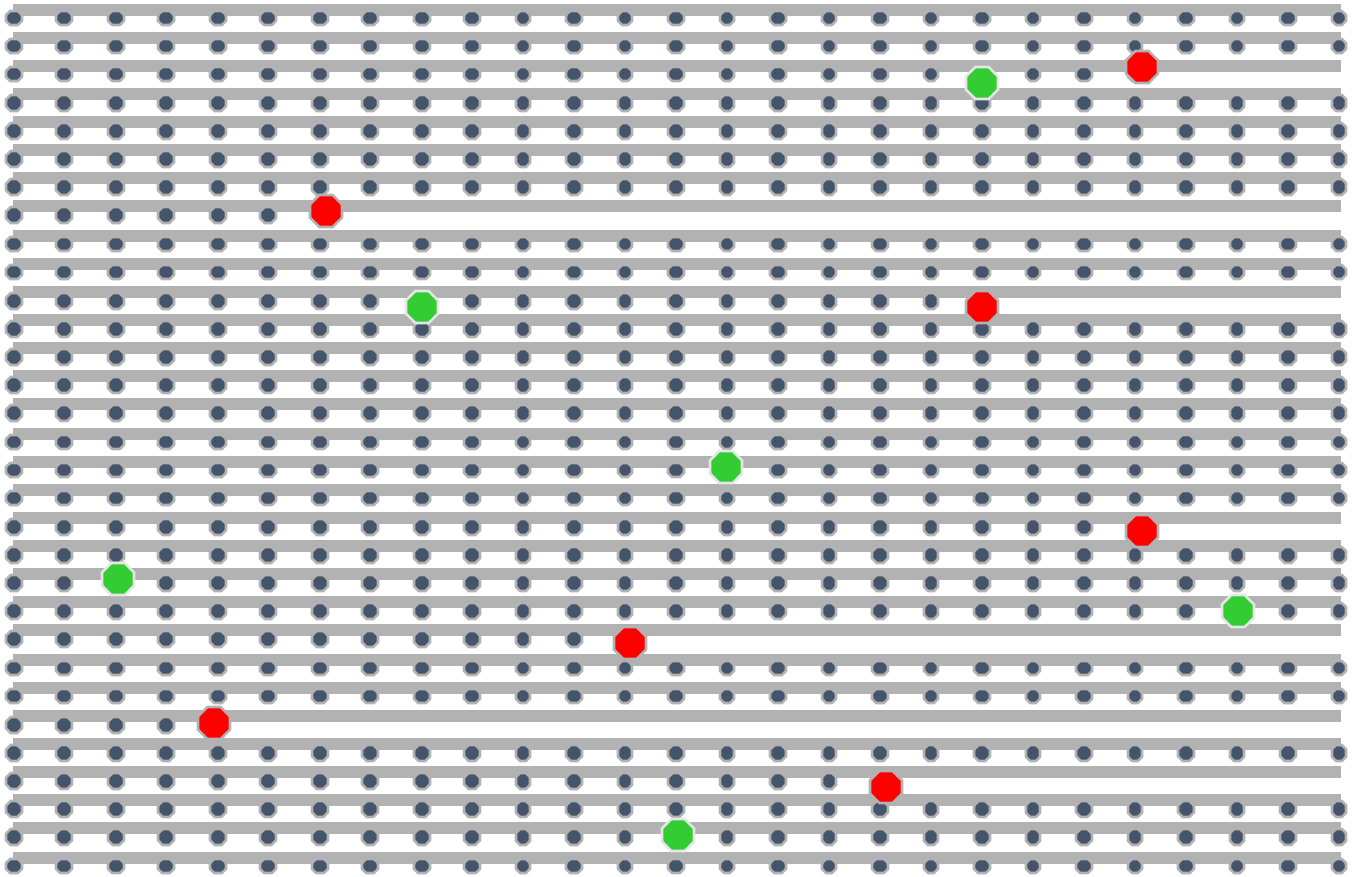
cohort

The ~~population~~ from which cases and controls are drawn (sampled).

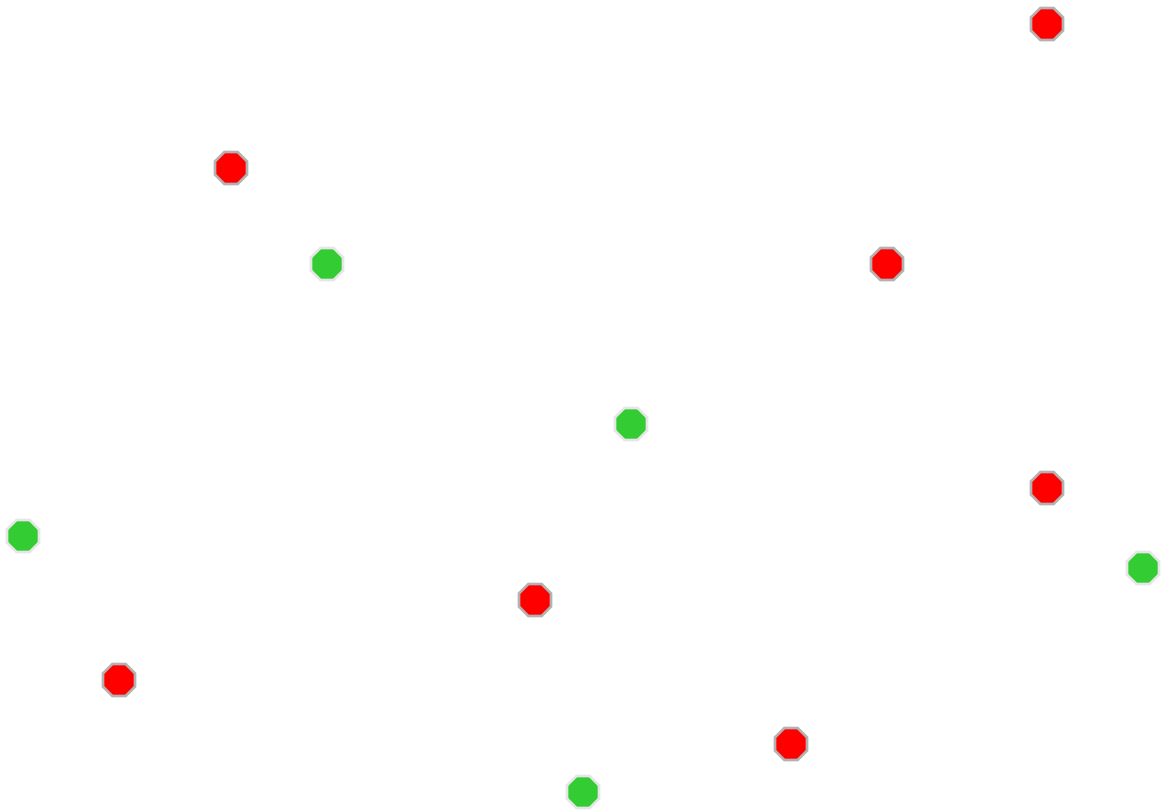
# Cohort design



# Case-control design



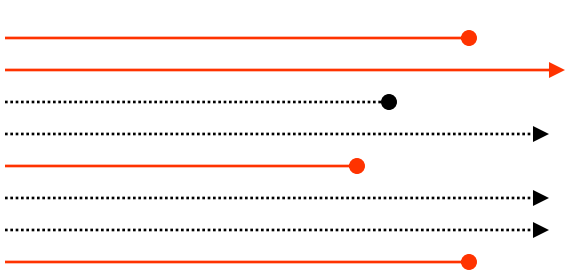
# Case-control design



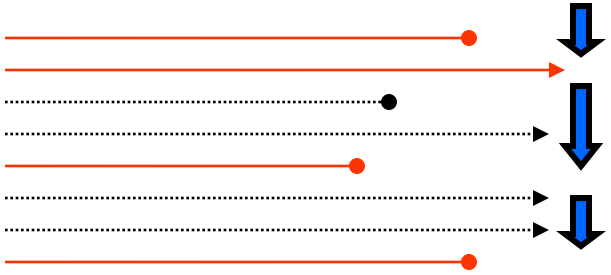
Has it always been like this?

**NO!**

# “Traditional” or Case-Non-Case or Cumulative Case-Control Studies



Persons

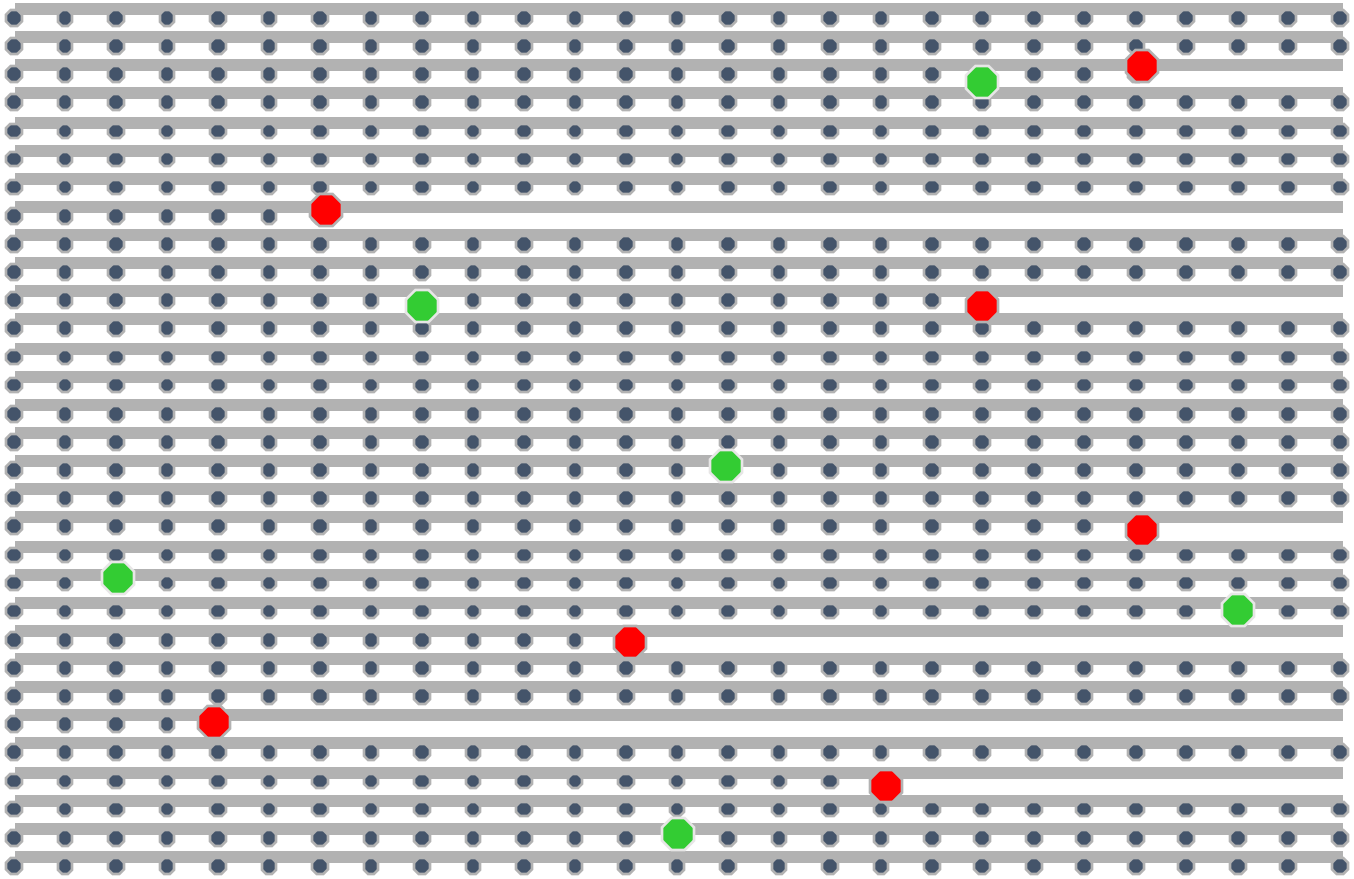


Use cases and a random sample of non-cases (controls)

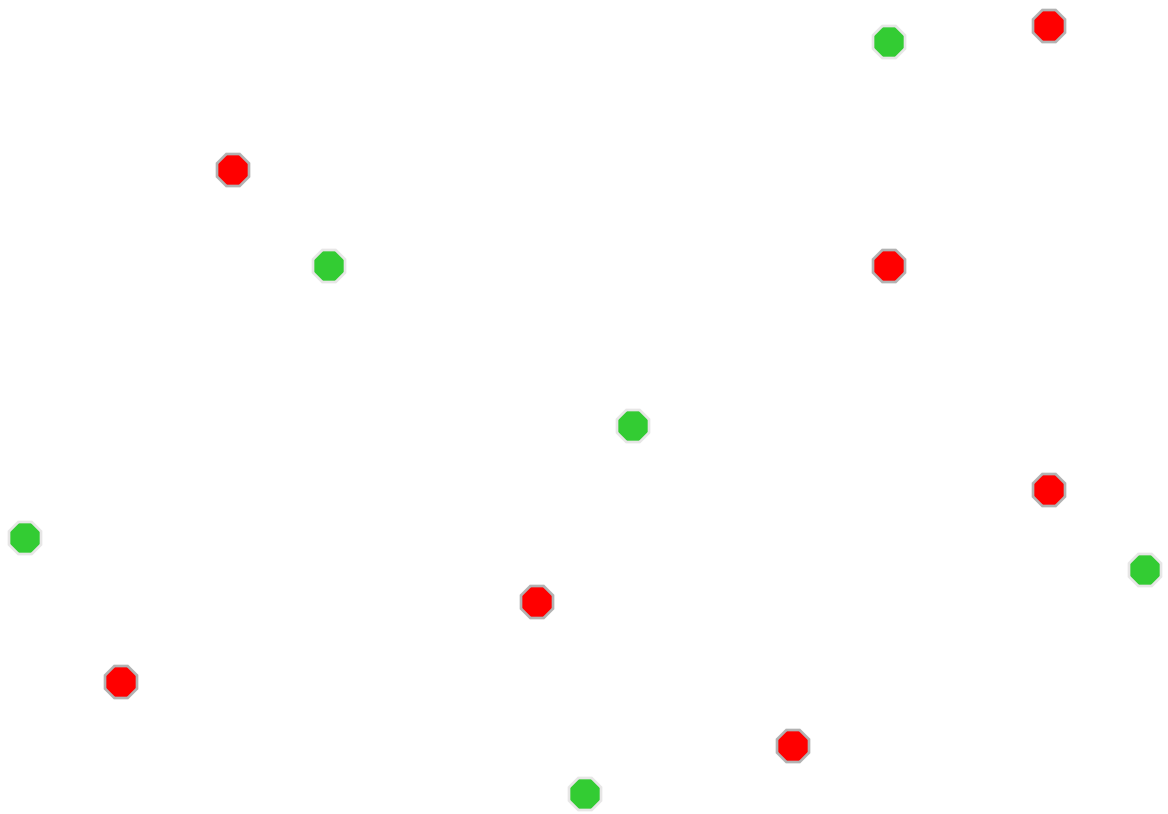
# A “case-control” study…

This study aimed to investigate the association between X use and the risk of Y in a case-control study. We analysed XXX database from 2002 to 2013. We defined “cases” as who underwent Y surgery between 2010 and 2013. “Controls” were patients with no history of Y between 2002 and 2013.

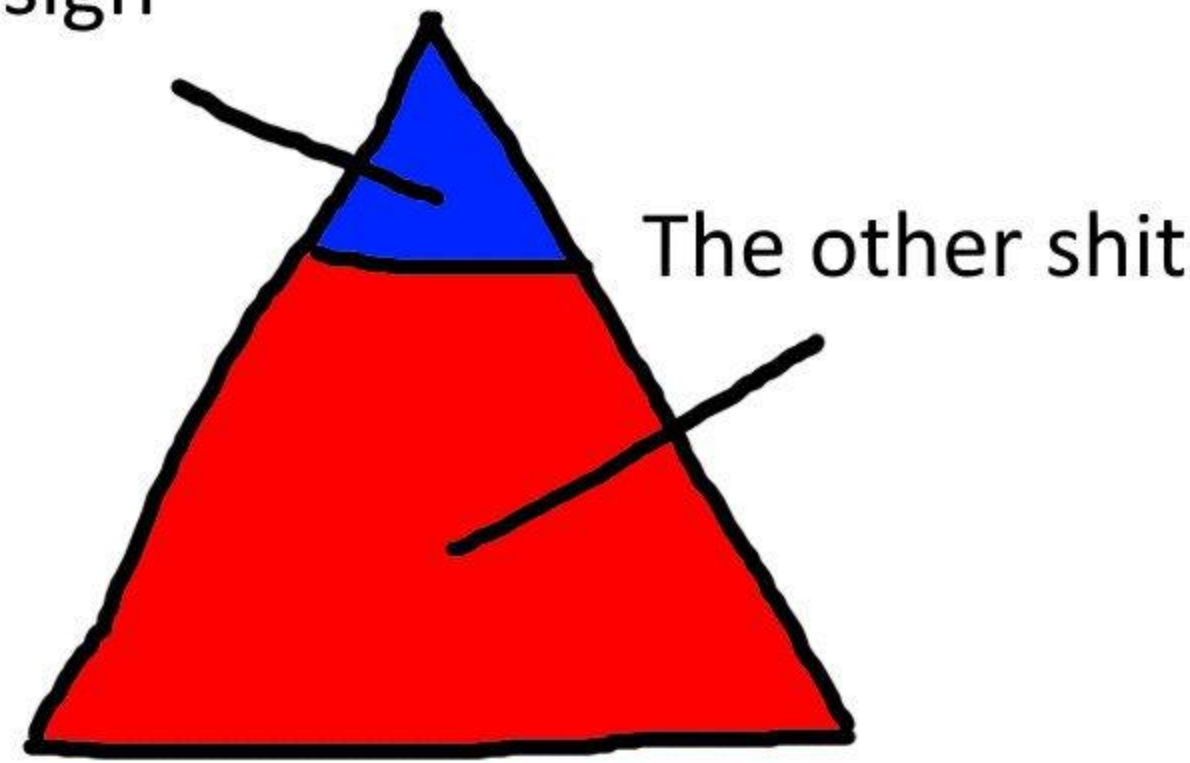
# Case-control design



# Case-control design



Thoughtful, well-conducted studies of any design



# Example

SHORT COMMUNICATION

BJC

British Journal of Cancer (2016), 1–5 | doi: 10.1038/bjc.2016.10

Keywords: lithium; colorectal cancer; adenocarcinoma; case–control; pharmacoepidemiology; Denmark

# Long-term use of lithium and risk of colorectal adenocarcinoma: a nationwide case–control study

Anton Pottegård[*,1], Zandra Nymand Ennis[2], Jesper Hallas[1,2], Boye L Jensen[3], Kirsten Madsen[3,4] and Søren Friis[5]

*[1]Clinical Pharmacology, Department of Public Health, University of Southern Denmark, DK-5000 Odense, Denmark; [2]Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, DK-5000 Odense, Denmark; [3]Department of Cardiovascular and Renal Research, University of Southern Denmark, DK-5000 Odense, Denmark; [4]Department of Pathology, Odense University Hospital, DK-5000 Odense, Denmark and [5]Danish Cancer Society Research Center, Danish Cancer Society, DK-2100 Copenhagen, Denmark*

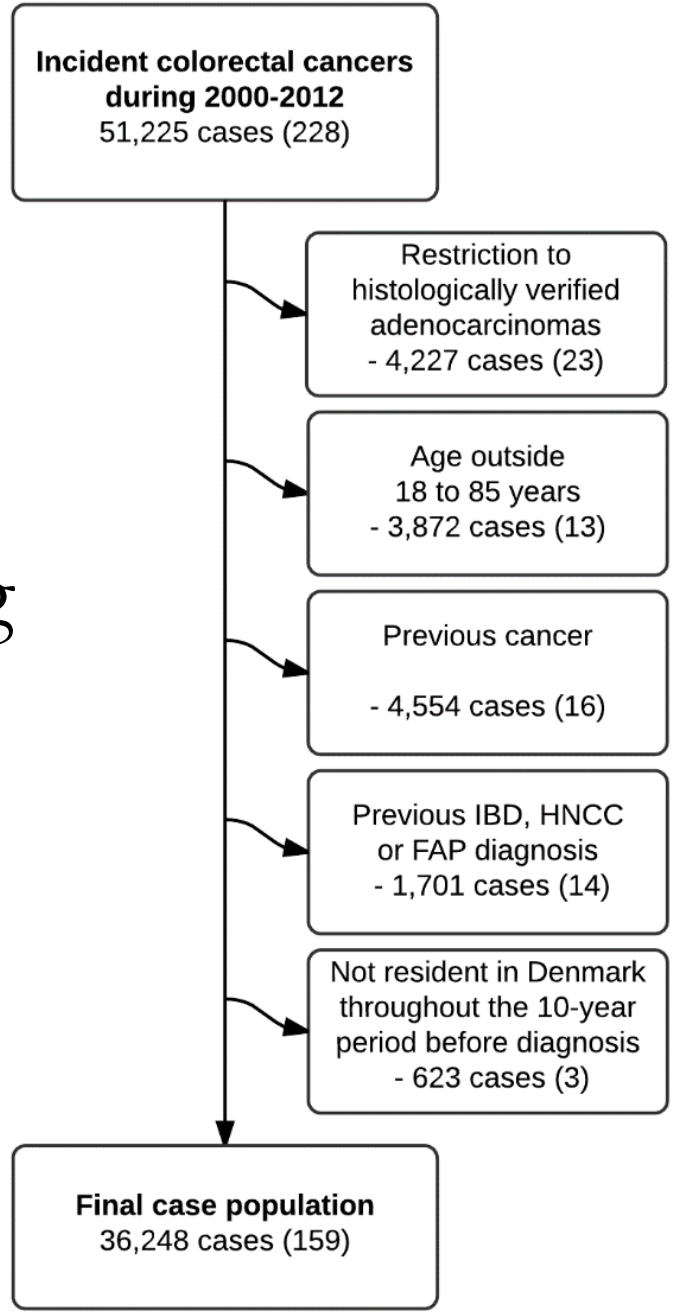
**Background:** Lithium accumulates in the colon and inhibits the enzyme GSK-3$\beta$ that possesses anti-carcinogenic effects. We therefore examined the association between lithium use and colorectal cancer risk in a nationwide study.

**Methods:** We used the Danish Cancer Registry to identify all patients diagnosed with incident colorectal adenocarcinoma during 2000–2012 ($n = 36\,248$). Using a matched case–control approach, we estimated the association between long-term use ($\geqslant$5 years) of lithium and risk of colorectal adenocarcinoma using conditional logistic regression.

**Results:** Long-term use of lithium was similar among cases (0.22%) and controls (0.20%), yielding an odds ratio of 1.13 (95% confidence interval (CI), 0.89–1.43) for colorectal adenocarcinoma. Dose–response, subgroup and other subanalyses returned neutral associations. However, ORs differed for colorectal subsites (proximal colon: 1.01 (95% CI, 0.66–1.55; distal colon: 1.52 (95%

$$Li^+$$

# Underlying cohort?



| Exposure group | Cases | Controls | Crude OR [1] | Adjusted OR [2] |
|---|---|---|---|---|
| Non-use | 36,089 | 360,909 | 1.00 (ref.) | 1.00 (ref.) |
| Ever use | 159 | 1,571 | 1.01 (0.86-1.19) | 1.08 (0.92-1.28) |
| Long-term use (≥ 5 years) | 78 | 734 | 1.06 (0.84-1.34) | 1.13 (0.89-1.43) |
| Duration of use | | | | |
| < 1 year | 21 | 277 | 0.76 (0.49-1.18) | 0.82 (0.53-1.28) |
| 1-4.99 years | 60 | 560 | 1.07 (0.82-1.40) | 1.15 (0.88-1.50) |
| 5-9.99 years | 50 | 506 | 0.99 (0.74-1.33) | 1.06 (0.79-1.41) |
| ≥ 10 years | 28 | 228 | 1.22 (0.83-1.81) | 1.29 (0.87-1.91) |

# Pros and cons?

# Pros

Statistically efficient

~~Less ressource demanding~~

Can (easily) look at multiple exposures at the same time

# Cons

Only provides relative estimates (in principal)

Not suited for multiple (different) outcomes

Less efficient with rare exposures

Control selection might ”go wrong”

Design often misunderstood

# … often misunderstood?!

Decision: rejection

Detailed comments from the meeting:
The committee felt this is a topical subject. This study is not the first of its kind, but it is a very big study and this is a strength.

**However the committee felt that the case-control methodology is intrinsically weak.**

**Table 1.** Characteristics of cases and controls at the index date.

| | Cases (*n* = 3571) | Controls (*n* = 35,582) |
|---|---|---|
| | | All |
| *Age* | | |
| Median (IQR) | 75 (64–83) | 75 (64–83) |
| *Sex* | | |
| Men | 1811 (50.7%) | 18,029 (50.7%) |
| *Current drug use* | | |
| VKA | 183 (5.1%) | 823 (2.3%) |
| Low-dose aspirin | 696 (19.5%) | 3436 (9.7%) |
| Other antiplatelet drugs | 197 (5.5%) | 782 (2.2%) |
| NSAID | 1220 (34.2%) | 4005 (11.3%) |
| SSRI | 429 (12.0%) | 2038 (5.7%) |
| Systemic corticosteroids | 384 (10.8%) | 1638 (4.6%) |

**When to consider?**

When you want to use MANY different exposure definitions

When outcome is rare

When computer power might be a limitation

**When best to avoid?**

If studying multiple outcomes

If exposure is rare

When absolute risks are central

When active comparators are considered