Using big data for pharmacovigillance I

The problems in reporting-based and in big-data based pharmacovigilance

Jesper Hallas, professor, DrMedSc
Clinical Pharmacology, pharmacy and environmental medicine
Odense Universitetshospital/SDU

1

PV and PE

PV signal generated by spontaneous reporting → Signal confirmed by PE

2

PV and PE

PV signal generated by spontaneous reporting → **Signal generated by PE**

3

Spontaneous reporting, limitations

- Extensive and variable(!) underreporting
- Still dependent on the individual clinicians' suspicions
- Purpose is widely misunderstood
- Cannot reject a suspicion
- Sensitive to publicity
- Not effective for cumulative, long-term effects
- .
- .
- .

4

Big data analyses, limitations

- Extensive and variable(!) underreporting
- Still dependent on the individual clinicians' suspicions
- Purpose is widely misunderstood
- Cannot reject a suspicion
- Sensitive to publicity
- Not effective for cumulative, long-term effects

5

Big data analyses, limitations

- Vulnerability to confounding
- Signals generated by clinical context

6

Confounding

- RR may go up or down, because of confounding
- False positive signals:
  - ->full-scale epidemiological analysis
- False negative signals
  - No follow-up

7

Ideal epidemiological design for screening

- Automatic, decision-free
- Robust towards confounding
- (simple)

8

Epidemiological design choices

- Cohort
- Case-control
- Self-controlled designs

9

Epidemiological design choices

- Cohort
- Case-control
- **Self-controlled designs**
  - **Case-crossover**
  - **Symmetry analysis**
  - **…**

10

Rationale for symmetry analysis

A woman aged 75, mildly hypochondriac and nursing home resident, started a thiazide and an antidepressant for the first time in her life in 2007.

Thiazides cause depression.

Which drug had the highest probability of being prescribed first?

Hallas, Epidemiology 1996

11

PHARMACO-EPIDEMIOLOGY

Hypothesis-free screening of large administrative databases for unsuspected drug-outcome associations

Jesper Hallas[1,2] · Shirley V. Wang[1] · Joshua J. Gagne[1] · Sebastian Schneeweiss[1] · Nicole Pratt[3] · Anton Pottegård[2]

Received: 12 January 2018 / Accepted: 21 March 2018

Abstract
Active surveillance for unknown or unsuspected adverse drug effects may be carried out by applying epidemiological techniques to large administrative databases. Self-controlled designs, like the symmetry design, have the advantage over conventional design of adjusting for confounders that are stable over time. The aim of this paper was to describe the output of a comprehensive open-ended symmetry analysis of a large dataset. All drug dispensings and all secondary care contacts in Denmark during the period 1995–2012 for persons born before 1950 were analyzed by a symmetry design. We analyzed all drug–drug sequences and all drug–disease sequences occurring during the study period. The identified associations were ranked according to the number of outcomes that potentially could be attributed to the exposure. In the main analysis, 29,891,212 incident drug therapies, and 21,300,000 incident diagnoses were included. Out of 186,758 associations tested in the main analysis, 43,575 (23.3%) showed meaningful effect size. For the top 200 drug–drug associations, 47% represented unknown associations, 24% represented known adverse drug reactions, 30% were explained by mutual indication or reverse causation. For the top 200 drug–disease associations the proportions were 31, 15, and 55%, respectively. Screening by symmetry analysis can be a useful starting point for systematic pharmacovigilance activities if coupled with a systematic post-hoc review of signals.

Keywords Pharmacovigillance · Pharmcoepidemiology · Self-controlled design · Databases · Screening

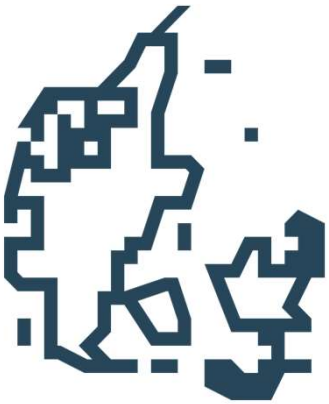
Background

About 1–3% of all newly marketed drugs are withdrawn because of adverse effects that are not known at the time of authorization [1, 2], thus necessitating a systematic surveillance of marketed drugs. For decades, the primary tool in generating signals about adverse drug effects after marketing has been spontaneous reporting [3]. This approach has several well-known limitations. First, there is massive and, most importantly, highly variable underreporting [4, 5], which is often the underlying cause of signals in spontaneous reporting schemes [6]. Second, individual case reports require an individual patient or clinician to connect the drug and the adverse event as potentially being causally related. It is therefore likely that many inconspicuous adverse drug reactions (e.g., those

12

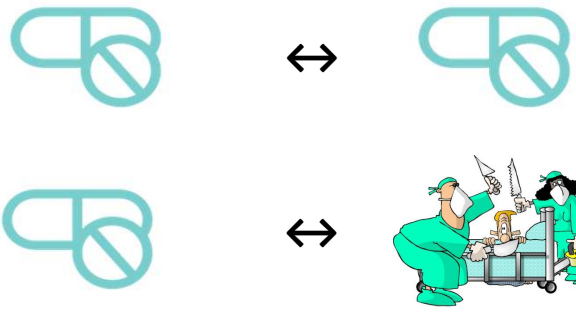
All-by-all symmetry analysis, data

- All Danes born before 1950, data from 1995-2012
- All prescriptions, n = 479,420,576
- All diagnoses, n = 80,865,480



13

All-by-all symmetry analysis



14

Hypotesis free symmetry analyses, output

Number of digits in ATC codes	Number of digits in ICD10 code	Analysis	Number of individual-level drug-drug and drug- disease pairs analyzed	Number of different associations analyzed	Number of different associations showing statistical significance (%)
7	4	Most extensive	232,303,361,934	3,099,493	305,965 (9.9%)
4	3	Main analysis	38,154,194,500	186,758	43,575 (23.3%)

15

16

Top association drug-drug pairs

Rank	Exposure drug (ATC)	Outcome drug (ATC)	Exposure drug first/last	Sequence ratio (95% confidence interval)	Log10(p)	Interpretation
1	NSAIDs (M01A)	Opioids (N02A)	73793 / 34542	2.14 (2.11 - 2.16)	<-307	**MIC**
2	Opioids (N02A)	Drugs for constipation (A06A)	58893 / 25127	2.34 (2.31 - 2.38)	<-307	**ADR**
3	Antithrombotic agents (B01A)	Cholesterol lowering drugs (C10A)	58859 / 29759	1.98 (1.95 – 2.01)	<-307	MIC
4	High-ceiling diuretics (C03C)	Potassium (A12B)	37271 / 11268	3.31 (3.24 - 3.38)	<-307	ADR
5	Thiazides (C03A)	Potassium (A12B)	35047 / 10128	3.46 (3.39 - 3.54)	<-307	ADR
6	Opioids (N02A)	Propulsives (A03F)	42322 / 19817	2.14 (2.10 - 2.17)	<-307	ADR
7	Other analgesics (N02B)	Drugs for constipation (A06A)	48385 / 28207	1.72 (1.69 - 1.74)	<-307	TDC
8	NSAIDs (M01A)	Other analgesics (N02B)	45941 / 27478	1.67 (1.65 - 1.70)	<-307	MIC
9	NSAIDs (M01A)	Corticosteroids for systemic use, plain (H02A)	36346 / 19479	1.87 (1.83 - 1.90)	<-307	MIC
10	Opioids (N02A)	Drugs affecting bone structure and mineralization (M05B)	20776 / 5484	3.79 (3.68 - 3.90)	<-307	MIC
11	Potassium (A12B)	Potassium-sparing agents (C03D)	21464 / 6472	3.32 (3.23 - 3.41)	<-307	MIC
12	Antithrombotic agents (B01A)	Drugs for constipation (A06A)	26205 / 12348	2.12 (2.08 - 2.17)	<-307	**Unknown**
13	NSAIDs (M01A)	Anti-ulcer drugs (A02B)	31295 / 18351	1.71 (1.67 - 1.74)	<-307	ADR
14	Antithrombotic agents (B01A)	Other analgesics (N02B)	40143 / 27300	1.47 (1.45 - 1.49)	<-307	Unknown
15	NSAIDs (M01A)	Drugs for constipation (A06A)	20841 / 9888	2.11 (2.06 - 2.16)	<-307	**TDC**
16	Opioids (N02A)	Antacids (A02A)	21060 / 10818	1.95 (1.90 - 1.99)	<-307	TDC
17	Antithrombotic agents (B01A)	Antidepressants (N06A)	27913 / 17906	1.56 (1.53 - 1.59)	<-307	Unknown

UNIVERSITY OF SOUTHERN DENMARK.DK

17

Top association drug-drug pairs

Rank	Exposure drug (ATC)	Outcome drug (ATC)	Exposure drug first/last	Sequence ratio (95% confidence interval)	Log10(p)	Interpretation
1	NSAIDs (M01A)	Opioids (N02A)	73793 / 34542	2.14 (2.11 - 2.16)	<-307	**MIC**
2	Opioids (N02A)	Drugs for constipation (A06A)	58893 / 25127	2.34 (2.31 - 2.38)	<-307	**ADR**
3	…	…	…	…	…	…
4	…	…	…	…	…	…
5	…	…	…	…	…	…
6	…	…	…	…	…	…
7	…	…	…	…	…	…
8	…	…	…	…	…	…
9	…	…	…	…	…	…
10	…	…	…	…	…	…
11	…	…	…	…	…	…
12	Antithrombotic agents (B01A)	Drugs for constipation (A06A)	26205 / 12348	2.12 (2.08 - 2.17)	<-307	**Unknown**
13	…	…	…	…	…	…
14	…	…	…	…	…	…
15	…	…	…	…	…	…
16	…	…	…	…	…	…
17	…	…	…	…	…	…

**P<0.00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001**

UNIVERSITY OF SOUTHERN DENMARK.DK

18

Top association drug-disease pairs

Rank	Exposure drug (ATC)	Outcome (ICD10)	Exposure drug first/last	Sequence ratio (95% confidence interval)	Log10(p)	Interpretation
1	Anti-ulcer drugs (A02B)	Dyspepsia (K30)	17487 / 7115	2.46 (2.39 - 2.53)	<-307	RC
2	Opioids (N02A)	Other intervertebral disc disorders (M51)	14985 / 5552	2.70 (2.62 - 2.78)	<-307	RC
3	Opioids (N02A)	Dorsalgia (M54)	16670 / 9858	1.69 (1.65 - 1.73)	<-307	RC
4	Opioids (N02A)	Osteoporosis without pathological fracture (M81)	11804 / 5563	2.12 (2.06 - 2.19)	<-307	RC
5	Corticosteroids for systemic use, plain (H02A)	Shoulder lesions (M75)	8084 / 2297	3.52 (3.36 - 3.69)	<-307	RC
6	Macrolides, lincosamides and streptogramins (J01F)	Pneumonia, organism unspecified (J18)	14603 / 9984	1.46 (1.43 - 1.50)	-190.3	RC
7	Macrolides, lincosamides and streptogramins (J01F)	Other chronic obstructive pulmonary disease (J44)	10277 / 6067	1.69 (1.64 - 1.75)	-237.4	RC
8	Topical treatment of hemorrhoids and anal fissures (C05A)	Other diseases of anus and rectum (K62)	6787 / 2695	2.52 (2.41 - 2.63)	<-307	RC
9	Anti-ulcer drugs (A02B)	Cholelithiasis (K80)	9390 / 5554	1.69 (1.64 - 1.75)	-215.7	RC
10	Beta-lactam antibacterials, penicillins (J01C)	Pneumonia, organism unspecified (J18)	15473 / 11721	1.32 (1.29 - 1.35)	-114.2	RC
11	Opioids (N02A)	Other spondylopathies (M48)	7639 / 4203	1.82 (1.75 - 1.89)	-218.4	RC
12	Opioids (N02A)	Other functional intestinal disorders (K59)	9301 / 5982	1.55 (1.51 - 1.61)	-158.3	RC
13	Corticosteroids for systemic use, plain (H02A)	Osteoporosis without pathological fracture (M81)	6026 / 2742	2.20 (2.10 - 2.30)	-269.0	ADR
14	Opioids (N02A)	Sequelae of injuries of lower limb (T93)	5994 / 2774	2.16 (2.07 - 2.26)	-258.7	RC
15	NSAIDs (M01A)	Pneumonia, organism unspecified (J18)	9443 / 6225	1.52 (1.47 - 1.57)	-145.3	Unknown
16	Macrolides, lincosamides and streptogramins (J01F)	Heart failure (I50)	8354 / 5152	1.62 (1.57 - 1.68)	-166.7	Unknown
17	Macrolides, lincosamides and streptogramins (J01F)	Malignant neoplasm of bronchus and lung (C34)	4604 / 1440	3.20 (3.02 - 3.39)	<-307	RC
18	Antithrombotic agents (B01A)	Other anaemias (D64)	8124 / 5072	1.60 (1.55 - 1.66)	-155.1	ADR

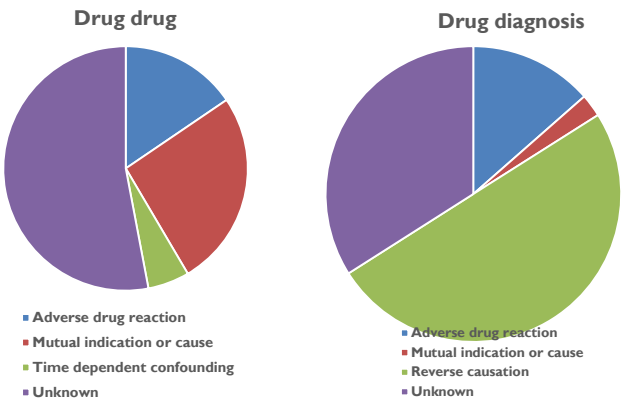
19

Top association drug-disease pairs

Rank	Exposure drug (ATC)	Outcome (ICD10)	Exposure drug first/last	Sequence ratio (95% confidence interval)	Log10(p)	Interpretation
1	Anti-ulcer drugs (A02B)	Dyspepsia (K30)	17487 / 7115	2.46 (2.39 - 2.53)	<-307	RC
2	…	…	…	…	…	RC
3	…	…	…	…	…	RC
4	…	…	…	…	…	RC
5	…	…	…	…	…	RC
6	…	…	…	…	…	RC
7	…	…	…	…	…	RC
8	…	…	…	…	…	RC
9	…	…	…	…	…	RC
10	…	…	…	…	…	RC
11	…	…	…	…	…	RC
12	…	…	…	…	…	RC
13	…	…	…	…	…	…
14	…	…	…	…	…	…
15	…	…	…	…	…	…
16	…	…	…	…	…	…
17	…	…	…	…	…	…
18	…	…	…	…	…	…

20

Interpretation of top 200 drug-drug and drug-diagnosis associations



21

Screening of big data, observations and reflections

- All prioritised outcomes represent true associations, all "needles"
  - "Big results" not big data
- Less than half represent potential unknown adverse drug reactions
- Some associations are generated by clinical context.
  - only to a limited extent removable by design

22

Follow-up of findings

- Known? Interpretable ? Obvious confounders ?
- Specific to any sub-exposure or any sub-outcome ?
- Supported by other types of evidence ?
- Specificity of association ?
  - Hierachy of ATC and ICD10
- ACNU HdPS ?
- Dose-response association ?
- Orthogonal predictions ?
- Validation in other data sources ?