BCPT Basic & Clinical Pharmacology & Toxicology

Basic & Clinical Pharmacology & Toxicology Doi: 10.1111/bcpt.12716

Performance of the High-dimensional Propensity Score in a Nordic Healthcare Model

Jesper Hallas and Anton Pottegård

Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Odense, Denmark

(Received 24 October 2016; Accepted 22 November 2016)

Abstract: The high-dimensional propensity score (hdPS) is increasingly used as a tool to adjust for confounding in observational studies of drug effects. It was developed within very rich data sources, for example the American claims databases. Thus, it is unknown whether it can be applied in settings that provide little more than primary care prescriptions and diagnoses from hospital contacts, as in the Nordic data sources. Our objective was to evaluate the performance of hdPS under such circumstances. As our case, we chose the association between use of selective cyclooxygenase-2 inhibitors (coxibs) and traditional NSAIDs (tNSAIDs) and the risk of upper GI bleeding. Using Danish health registries, we identified 110,285 incident users of coxibs and 575,980 incident users of tNSAIDs and followed them for 90 days with respect to the occurrence of serious upper GI bleeding. Data were analysed using Cox regression, estimating the coxib/tNSAID hazard ratio (HR). Values below 1.00 indicate a lower estimated hazard with coxibs. We build hdPS models with inclusion of up to 500 diagnosis and 500 prescription drug covariates. The crude HR was 1.76 (95% confidence interval: 1.57–1.97), decreasing to 1.12 (1.00–1.26) and 0.99 (0.88–1.12) after adjustment for age and sex and 11 pre-selected confounders, respectively. A hdPS with inclusion of 500 most prevalent diagnoses and 500 most prevalent prescription drugs resulted in a HR of 0.89 (0.77–1.02). These estimates were consistently lower when the analysis was restricted to non-users of low-dose aspirin. The estimate based on 500 diagnoses alone was higher than an estimate based on 500 prescription drugs alone (0.99 *versus* 0.91). We conclude that hdPS does work within a Nordic setting that prescription data are more effective than diagnosis data in achieving confounder adjustment and that hdPS seems more effective than simple confounder adjustment by variables selected on the basis of clinical reasoning.

The propensity score technique has become immensely popular as a tool to adjust for confounding in pharmacoepidemiological studies of drug effects. Its popularity is founded on its ability to achieve and demonstrate covariate balance in cohort studies, while at the same time analysing factors that drive choice of one treatment alternative over the other and effectively addressing problems like potential overfitting [1]. In 2009, Schneeweiss proposed that propensity scores could be automatically modelled on the basis of hundreds of covariates, the so-called high-dimensional propensity score (hdPS) [2]. It was demonstrated that this approach achieved more plausible effect estimates than conventional PS modelling based on clinically selected variables or simple multivariable modelling. The rationale for hdPS is that by selecting a myriad of variables, solely on the basis of their frequency, the method effectively exploits the data set's inherent capacity for covariate adjustment.

The method has been developed and used mainly on data sets generated from a North American healthcare model. Such data are very rich, and because they serve primarily a billing purpose, they typically include all services, diagnostic procedures, diagnoses and treatments offered to the patient, both from general practice, outpatient specialist care and hospitalizations. The Nordic healthcare model offers full tax-funded health care to all citizens, and because billing for the individual patient is less of an issue, data sources are not as rich and diverse as the North American. In- and outpatient contacts in hospital settings are recorded, but in principle, a person can be hospitalized for several weeks having little more recorded than a few discharge diagnoses and the admission and discharge dates. Prescriptions from primary care are available in all Nordic countries [3], but diagnoses and procedures from primary care are generally not available.

We undertook this study to evaluate how hdPS performs in a Nordic setting, using only outpatient prescription records and diagnoses from in- or outpatient hospital contacts. As our subject matter, we chose the association between choice of selective COX-2-inhibitors (coxibs) over traditional non-steroidal anti-inflammatory drugs (tNSAIDs) and the risk of upper gastrointestinal bleeding (UGB). This association is notoriously difficult to analyse by conventional multivariable modelling [2,4], as coxibs are preferentially channelled to persons at higher risk of UGB [5]. The individuals' risk profile is not necessarily fully captured in the data, thus resulting in a bias against the coxibs [4]. Incidentally, the coxib-NSAID case was one of the motivating examples in the original presentation of hdPS [2].

Methods

We used nationwide Danish data on prescription fillings and hospital diagnoses and compared the 90-day risk of UGB among new users of coxibs and tNSAIDs. From clinical reasoning, we selected a number of covariates that would be potential confounders, and we used up to the 500 most prevalent drugs and 500 most prevalent diagnoses to build hdPS models.

Author for correspondence: Jesper Hallas, Clinical Pharmacology and Pharmacy, University of Southern Denmark, J.B. Winsløws Vej 19, 2, 5000 Odense C, Denmark (e-mail jhallas@health.sdu.dk).

Data sources. The data sources have been described in previous publications [6–8]. In brief, we used the Danish National Prescription Register [6], which has data on all prescriptions from 1995 onwards for all Danish residents. The recorded information includes the prescription holder, the date of filling, the dispensed product, including substance, quantity, brand name, unit strength and number of units. Substances are encoded according to the ATC system developed by WHO, and quantities are expressed by the defined daily doses methodology, also developed by WHO. The Danish National Patient Register [7] has records on all hospital contacts, including hospitalizations from 1977 and outpatient and emergency room contacts since 1994. Diagnoses were encoded according to the ICD8 from 1977 and the ICD10 from 1994. ICD9 was never used in Denmark. Finally, we used the Danish civil registration system [8], which accounts for all births, death and migrations, thereby allowing us to perform censoring as appropriate.

These data sources were linked using the central person register (CPR) number, a 10-digit unique identifier assigned to all Danish residents, and used by all health registers [9]. The linkage and data hosting were performed by Statistics Denmark, a governmental institution which collects data for a variety of statistical and scientific purposes [9].

Population. We included individuals who fulfilled the following criteria: (i) they were born in 1950 or earlier and (ii) they filled a prescription on either a coxib (only celecoxib, rofecoxib and etoricoxib were available on the market during the study period) or a tNSAID during the period of 1 January 2000 to 31 December 2004 after not having filled any such prescription for at least 2 years. No other in- or exclusion criteria were applied.

Analysis. The analysis conformed to a conventional cohort study. Individuals were classified as users of either coxibs or tNSAIDs on the basis of their index prescription, and they were followed to either hospitalization with a diagnosis of UGB (ICD10: K250, K252, K254, K256, K260, K262, K264, K266, K270, K272, K274, K276, K280, K282, K284, K286 or K290), death, emigration or the end of the study period, whichever was earliest. The choice of 90 days of follow-up was based on the observation that the risk of UGB is highest within the first few months of NSAID therapy [10] and that NSAIDs are often prescribed for fairly short treatment periods [11].

Individuals were characterized at baseline (day of filling a NSAID/coxib prescription) according to criteria that were preselected on the basis of clinical reasoning. These were age, sex, a history of uncomplicated peptic ulcer, bleeding peptic ulcer or alcohol abuse and use of proton pump inhibitors, SSRIs, low-dose aspirin, clopidogrel, statins, systemic corticosteroids or vitamin K antagonists.

In addition, we identified the 500 most prevalent diagnoses for the study population. In doing so, we used four-digit ICD10 codes (e.g. I20.9: Angina pectoris, unspecified) and applied a look-back period from 1994 onwards, corresponding to the introduction of the ICD10 system in Denmark. Similarly, we identified the 500 most used drugs at baseline. We used the full ATC code (e.g. A02BC01: omeprazole) and applied a look-back period of 6 months; that is, if there was a prescription for omeprazole within 6 months before the index date, the individual was considered as exposed to omeprazole. Registrations with incomplete codes, for example using preliminary ATC with only five digits or ICD10 with only three digits, were disregarded. Finally, we constructed a PS model, estimating the probability that an individual was using a coxib rather than a tNSAID, conditional on baseline characteristics.

Diagnoses and drugs were ranked according to their prevalence. When analysing the level of confounder adjustment achieved by for example 60 diagnosis covariates, we only included the 60 most prevalent diagnoses. In addition, we performed an analysis based on prioritized covariates, using the approach described by Schneeweiss *et al.* [2]. In brief, covariates were ranked according to their potential for exerting confounding effect, on the basis of their own association with the outcome and their prevalence among users of coxibs and tNSAIDs, and those with the highest potential for exerting confounding were included first. We included all of the 500 most prevalent diagnoses and 500 most prevalent prescriptions as candidates for prioritizing.

The association between use of coxib over tNSAID and ensuing UGB was carried out using conventional Cox regression, estimating the hazard ratio for coxibs *versus* tNSAIDs. We did not take into account whether the coxib of tNSAID dispensing covered the full 90-day observation period or whether any switching occurred or other treatment was initiated. Thereby, our analyses conform to an intention-to-treat principle. The following exploratory analyses were conducted as follows: (i) a crude analysis with no adjustment for covariates; (ii) adjustment for age (as a continuous variable) and sex; (iii) adjustment for age, sex and clinically selected covariates; (iv) adjustment for a PS based on age, sex and the top 500 diagnoses and drugs; (v) an explorative analysis using a PS based on age, sex and a variable number of diagnosis or prescription drug covariates in the PS model, to establish the dose–response effect between number of covariates included and the level of confounder adjustment achieved; and finally (vi) a similar analysis based on prioritized covariates. In all our main analyses, we applied asymmetrical trimming [12], that is, excluded all individuals (users of coxibs or tNSAIDs) with PS below the 2.5 percentile of the PS among the coxib users or above the 97.5 percentile for the tNSAID users. All of these analyses were repeated by restricting to non-users of low-dose aspirin, as it has been shown that use of low-dose aspirin is a strong effect modifier of bleeding risk and that the benefit of choosing coxibs over tNSAIDs in terms of upper GI bleeding risk is nearly absent in users of low-dose aspirin [13,14].

To investigate how the hdPS approach behaved in our study population, we also charted the number of covariates that each individual was positive for; for example, when including 100 diagnoses and 100 drugs in the hdPS model, how many drug or diagnosis markers were positive for the individual patient. This distribution of 'number of positive markers per patient' was calculated as a function of the number of PS covariates included in the analysis.

Finally, we charted the degree of area overlap between PS distributions for coxib and tNSAID users as a function of the number of included covariates in the hdPS model. The underlying rationale is that with an increasingly rich PS model, the two distributions will be increasingly separated [15]. It has been shown that PS models that predict choice of treatment very well, that is, have very little area overlap between distributions of PS for the two treatments, are statistically inefficient in terms of estimating treatment effect and may amplify bias by unmeasured confounders [16].

Other. All calculations were performed using STATA 14.1 (StataCorp, College Station, TX, USA). According to Danish law, review by an ethics committee is not required for pure register studies [9].

Results

We identified 110,285 users of coxibs and 575,980 users of tNSAIDs. Their characteristics are detailed in table 1. Compared to users of tNSAIDs, users of coxibs were older, more often women, had a higher prevalence of peptic ulcer antecedents and used low-dose aspirin, statins, systemic corticosteroids, vitamin K antagonists and proton pump inhibitors. tNSAID users were

followed for 137,874 person-years and had 1182 outcomes, for a crude incidence rate of 8.6 per 1000 person-years. The corresponding figures for coxib users were 26,560 person-years, 400 outcomes and a rate of 15.1 per 1000 person-years.

The results of the main analysis are shown in table 2. The crude and the simple analysis adjusted for age and sex showed a higher hazard of UGB for users of coxib. After inclusion of clinically selected covariates, the HR was 0.99 (95% confidence interval [CI]: 0.88–1.12), that is indicating a similar risk of bleeding among users of coxibs and tNSAIDs. The inclusion of the top 40 covariates (20 diagnoses and 20 prescription drugs) resulted in a HR of 0.86 (CI: 0.75–0.99), while further inclusion of covariates did not result in further lowering of the estimate. There was consistently lower HR when choosing coxibs over tNSAIDs among non-users of low-dose aspirin in all models. We found that inclusion of prescription drug covariates alone was more effective than diagnosis covariates alone with respect to confounder adjustment (HR = 0.91 [CI: 0.80–1.04] with 500 prescription drugs *versus* HR = 0.99 [0.87–1.13] with 500 diagnoses). The effect of trimming was a modest lowering in the observed HRs, without any appreciable loss of precision (data not shown).

The numbers of covariates that tested positive in study participants are shown in fig. 1, as a function of the number of covariates included in the PS model. With inclusion of the top 500 diagnoses and 500 drugs, 643,598 (93.7%) were positive for at least one covariate, but individuals with more than 50 positive covariates were extremely rare (<0.06%). The corresponding distribution of positive covariates for diagnoses and prescription drugs viewed separately is shown in fig. S1.

The degree of overlap between PS distributions is shown in fig. 2. The overlap decreased from 71%, achieved by 10 diagnoses and 10 drugs, to 64% achieved with the full range of diagnoses and drugs. The distribution of PS for coxibs and tNSAIDs with inclusion of 500 diagnoses and 500 prescription drugs is shown in fig. 3.

Table 1.
Characteristics of incident users of selective COX-2 inhibitors and traditional NSAIDs in Denmark during 2000–2004.

	Traditional NSAIDs (N = 575,980)	Selective COX-2 inhibitors (N = 110,285)
Age, median (IQR)	62.9 (56.6–72.5)	71.4 (61.3–80.2)
Male sex	269,090 (46.7%)	37,414 (33.9%)
History of		
Uncomplicated ulcer	11,604 (2.0%)	6637 (6.0%)
Bleeding peptic ulcer	3508 (0.6%)	2446 (2.2%)
Alcohol abuse	9806 (1.7%)	1979 (1.8%)
Current use of		
Low-dose aspirin	79,011 (13.7%)	23,578 (21.4%)
Clopidogrel	1812 (0.3%)	670 (0.6%)
Vitamin K antagonists	7782 (1.4%)	3850 (3.5%)
SSRIs	32,070 (5.6%)	11,520 (10.4%)
Proton pump inhibitors	30,559 (5.3%)	18,130 (16.4%)
Statins	36,204 (6.3%)	8025 (7.3%)
Systemic corticosteroids	28,246 (4.9%)	11,037 (10.0%)

SSRI, Selective serotonin reuptake inhibitors; NSAIDs, Non-steroidal anti-inflammatory drugs; COX, Cyclooxygenase; IQR, Interquartile range.

Discussion

We have shown that the hdPS approach does work within a Nordic setting, limited to the use of prescription data and diagnoses from in- or outpatient hospital contacts. However, the incremental adjustment from what was achieved by the clinically selected covariates was modest, even with inclusion of large numbers of covariates. Particularly inclusion of large numbers of diagnosis variables alone was ineffective.

A meta-analysis of randomized trials has found a 60% lower rate of peptic ulcer bleeding when using coxibs, compared to tNSAIDs [17]. We found a somewhat lower benefit, even with inclusion of large numbers of covariates. There are several potential explanations. Firstly, most users of tNSAIDs consume fairly small quantities [11]. As we followed the individuals for 90 days, some of them were likely to have stopped treatment before the end of follow-up. Thereby, our analysis is in effect an intention-to-treat analysis. Secondly, the preferred tNSAID in our population was ibuprofen (50% of tNSAID users) and the preferred coxib was rofecoxib (52% of users). In observational studies, ibuprofen has been established as one of the least GI toxic tNSAIDs [18,19]. In a recent meta-analysis of observational studies, celecoxib had an OR of 1.5, ibuprofen 1.8 and rofecoxib 2.3 for peptic ulcer complications [18]. Thereby, we would expect a low contrast of GI toxicity between coxibs and tNSAIDs in our study. Thirdly, we had a high proportion of users of low-dose aspirin among the coxib users, 21.4%. Trials have established that the benefit of coxibs in terms of reduced UGB risk is virtually absent in users of low-dose aspirin [13,14]. Accordingly, we also found substantially lower hazard ratios when we restricted the data to non-users of low-dose aspirin. Finally, there may be residual confounding. Coxibs are known to produce fewer ulcer complications than tNSAIDs and they are therefore channelled to users who are at particular risk [5]. This is a typical example of confounding by indication, which is notoriously difficult to remove completely in observational research. Importantly, our estimates are in line with what has been found in other recent observational studies [2,20,21].

The prioritized covariate inclusion did not appear to perform better than the model that included covariates based on frequency. There was one possible advantage of prioritizing; the confidence intervals were marginally narrower, as demonstrated by a lower ratio between upper and lower limits. This is possibly explained by the fact that the prioritizing algorithm is unlikely to pick covariates that are instruments, that is, related to exposure without having any direct effect on the outcome. Inclusion of such variables is likely to decrease precision [15].

One concern would be that inclusion of a very high number of covariates would lead to efficient separation of the PS distributions of the coxib and tNSAID cohorts. However, as judged from fig. 2, it is apparent that this effect is very

Table 2.
Performance of different levels of confounder adjustment. Cohort study of 110,285 users of selective COX-2 inhibitors and 575,980 users of traditional NSAIDs using upper gastrointestinal bleeding within 90 days as outcome. The analyses with inclusion of propensity scores were subjected to asymmetrical trimming. See text for explanation.

Analytic approach	Number of diagnosis covariates included	Number of prescription covariates included	Total number of covariates	Hazard ratio (95% CI)	
				Entire Cohort	Non-users of low-dose aspirin
Crude analysis	NA	NA	NA	1.76 (1.57–1.97)	1.40 (1.19–1.64)
Adjusted by age and sex	NA	NA	NA	1.12 (1.00–1.26)	0.90 (0.76–1.05)
Adjusted by age, sex and clinically selected covariates	3	7	10	0.99 (0.88–1.12)	0.75 (0.64–0.89)
Adjusted by PS model	10	10	20	0.89 (0.77–1.02)	0.76 (0.63–0.92)
	20	20	40	0.86 (0.75–0.99)	0.77 (0.63–0.93)
	50	50	100	0.86 (0.76–0.99)	0.71 (0.59–0.86)
	100	100	200	0.89 (0.77–1.01)	0.67 (0.55–0.82)
	200	200	400	0.88 (0.77–1.01)	0.66 (0.54–0.81)
	500	500	1000	0.90 (0.79–1.03)	0.68 (0.56–0.83)
	0	10	10	0.91 (0.79–1.04)	0.76 (0.63–0.92)
	0	20	20	0.89 (0.78–1.02)	0.81 (0.67–0.98)
	0	50	50	0.89 (0.77–1.01)	0.75 (0.62–0.90)
	0	100	100	0.89 (0.78–1.01)	0.70 (0.57–0.85)
	0	200	200	0.90 (0.79–1.03)	0.69 (0.57–0.84)
	0	500	500	0.91 (0.80–1.04)	0.72 (0.60–0.87)
	10	0	10	1.07 (0.94–1.21)	0.91 (0.76–1.09)
	20	0	20	1.03 (0.91–1.17)	0.88 (0.73–1.05)
	50	0	50	1.02 (0.89–1.16)	0.86 (0.71–1.03)
	100	0	100	1.03 (0.91–1.18)	0.86 (0.71–1.03)
	200	0	200	1.00 (0.88–1.14)	0.80 (0.66–0.96)
	500	0	500	0.99 (0.87–1.13)	0.79 (0.65–0.95)
Adjusted by PS model based on prioritized covariates	1	9	10	1.14 (1.02–1.29)	0.91 (0.77–1.07)
	1	19	20	1.15 (1.02–1.29)	0.91 (0.78–1.07)
	7	43	50	0.99 (0.87–1.11)	0.78 (0.66–0.92)
	31	69	100	0.92 (0.81–1.04)	0.70 (0.59–0.83)
	97	103	200	0.93 (0.82–1.05)	0.69 (0.58–0.82)
	260	240	500	0.96 (0.85–1.08)	0.71 (0.60–0.84)

NA, Not applicable; PS, Propensity score; NSAIDs, Non-steroidal anti-inflammatory drugs; COX, Cyclooxygenase.

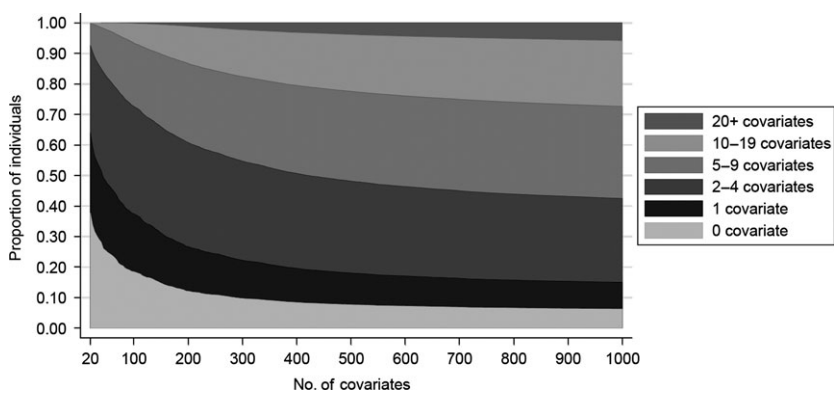

Fig. 1. Proportion of patients who have 0, 1, 2-4, 5-9, 10-19 or 20 + positive variables as a function of the number of covariates included. Both diagnosis and prescription variables were included.

limited. This is likely to be explained by the limited gain in the number of positive covariates for the single individual that is achieved by including 500 covariates compared to 50–100 covariates (fig. 1).

The primary strength of the analysis is the use of nationwide health register data, capturing all Danish users of coxibs and tNSAIDs obtained via prescription. Furthermore, the large data material allowed us to exemplify the application of hdPS with little interference from stochastic variation.

Our study also has limitations. Ibuprofen, a tNSAID, is available over the counter in Denmark. However, two-thirds of ibuprofen was obtained by prescription during the study period [22], and elderly individuals such as those in our study population will have a financial incentive to obtain their

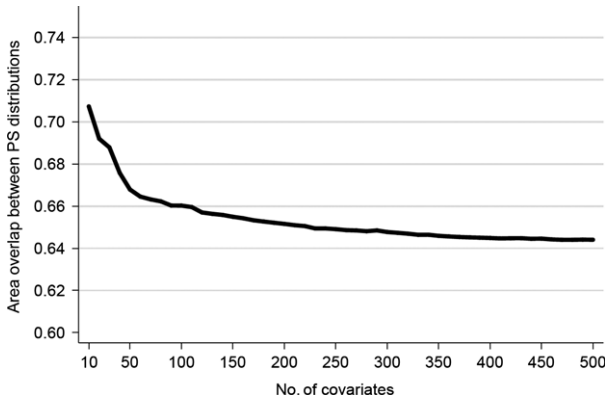

Fig. 2. Degree of area overlap between propensity score distributions of selective COX-2 inhibitors and traditional NSAIDs as a function of the number of diagnosis and prescription covariates included in the propensity score model. NSAIDs, Non-steroidal anti-inflammatory drugs; COX, Cyclooxygenase.

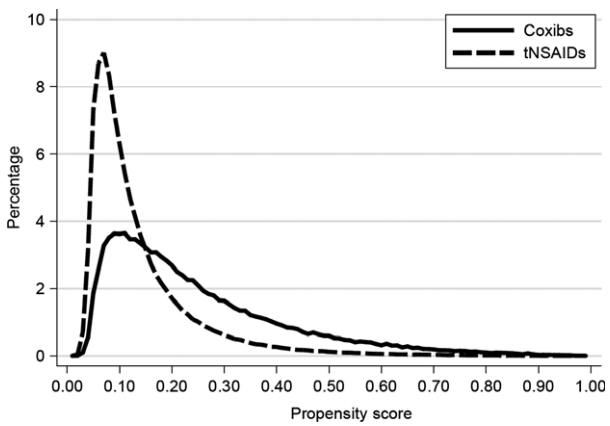

Fig. 3. Distribution of high-dimensional propensity scores in 686,265 users of tNSAIDs and coxibs. In all, 500 diagnosis and 500 prescription covariates were included. The area of the distributions has each been standardized to have an area of 1. The *y*-axis indicates the percentage of the total area found within a 0.01-wide band on the *x*-axis. tNSAIDs, Traditional non-steroidal anti-inflammatory drugs; coxibs, Selective cyclooxygenase-2 inhibitors.

NSAIDs via prescription, in order to receive reimbursement. Even so, the use of ibuprofen obtained over the counter is unlikely to be an important bias in our study [23]. Another limitation of our study is that it only outlines the performance of hdPS for a single drug-outcome association. Further methodological work is warranted to document application of hdPS in other clinical scenarios. In addition to diagnosis and prescription data, this might also include additional data, for example, from clinical databases [24].

In our example, the added benefit of hdPS compared to using clinically selected variables was modest. However, in contrast to clinically selected variables, hdPS can be automated. It may thus be useful when there are few candidates for clinically selected variables or when they are unavailable. Further, hdPS may be useful for screening for multiple unknown associations in large data sets. It would be impossible to select confounders manually for each tested association, but the hdPS can be applied without use of clinical reasoning [25].

We conclude that hdPS is effective as a principle for confounder adjustment in a setting, such as the Nordic one, that only provides data on prescriptions and diagnoses from hospital contacts and that it does provide some modest incremental confounder adjustment compared to a traditional approach. However, the ultimate position of hdPS in the pharmacoepidemiological armamentarium in such settings is currently unknown, and further work is needed to fully appraise the value of the hdPS approach under these circumstances.

References

1 Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic Clin Pharmacol Toxicol 2006;**98**:253–9.
2 Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiol Camb Mass 2009;**20**:512–22.
3 Wettermark B, Zoëga H, Furu K, Korhonen M, Hallas J, Nørgaard M *et al.* The Nordic prescription databases as a resource for pharmacoepidemiological research–a literature review. Pharmacoepidemiol Drug Saf 2013;**22**:691–9.
4 Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiol Camb Mass 2006;**17**:268–75.
5 MacDonald TM, Morant SV, Goldstein JL, Burke TA, Pettitt D. Channelling bias and the incidence of gastrointestinal haemorrhage in users of meloxicam, coxibs, and older, non-specific non-steroidal anti-inflammatory drugs. Gut 2003;**52**:1265–70.
6 Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health 2011;**39**(7 Suppl):38–41.
7 Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health 2011;**39**(7 Suppl):30–3.
8 Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011;**39**(7 Suppl):22–5.
9 Thygesen LC, Daasnes C, Thaulow I, Brønnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. Scand J Public Health 2011;**39**(7 Suppl):12–6.
10 Lanas A, García-Rodríguez LA, Arroyo MT, Gomollón F, Feu F, González-Pérez A *et al.* Risk of upper gastrointestinal ulcer bleeding associated with selective cyclo-oxygenase-2 inhibitors, traditional non-aspirin non-steroidal anti-inflammatory drugs, aspirin and combinations. Gut 2006;**55**:1731–8.
11 Fosbøl EL, Gislason GH, Jacobsen S, Abildstrom SZ, Hansen ML, Schramm TK *et al.* The pattern of use of non-steroidal anti-inflammatory drugs (NSAIDs) from 1997 to 2005: a nationwide study on 4.6 million people. Pharmacoepidemiol Drug Saf 2008;**17**:822–33.
12 Stürmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution–a simulation study. Am J Epidemiol 2010;**172**:843–54.
13 Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A *et al.* Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. Celecoxib Long-term Arthritis Safety Study. JAMA 2000;**284**:1247–55.
14 Schnitzer TJ, Burmester GR, Mysler E, Hochberg MC, Doherty M, Ehrsam E *et al.* Comparison of lumiracoxib with naproxen and ibuprofen in the Therapeutic Arthritis Research and

Gastrointestinal Event Trial (TARGET), reduction in ulcer complications: randomised controlled trial. Lancet Lond Engl 2004;**364**:665–74.
15 Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. Am J Epidemiol 2006;**163**:1149–56.
16 Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ *et al.* Effects of Adjusting for Instrumental Variables on Bias and Precision of Effect Estimates. Am J Epidemiol 2011;**174**:1213–22.
17 Rostom A, Muir K, Dubé C, Jolicoeur E, Boucher M, Joyce J *et al.* Gastrointestinal safety of cyclooxygenase-2 inhibitors: a Cochrane Collaboration systematic review. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2007;**5**:818–28, 828.e1–5; quiz 768.
18 Castellsague J, Riera-Guardia N, Calingaert B, Varas-Lorenzo C, Fourrier-Reglat A, Nicotra F *et al.* Individual NSAIDs and upper gastrointestinal complications: a systematic review and meta-analysis of observational studies (the SOS project). Drug Saf 2012;**35**:1127–46.
19 Henry D, Lim LL, Garcia Rodriguez LA, Perez Gutthann S, Carson JL, Griffin M *et al.* Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis. BMJ 1996;**312**:1563–6.
20 Garbe E, Kloss S, Suling M, Pigeot I, Schneeweiss S. High-dimensional versus conventional propensity scores in a comparative effectiveness study of coxibs and reduced upper gastrointestinal complications. Eur J Clin Pharmacol 2013;**69**:549–57.
21 Bakhriansyah M, Souverein PC, de Boer A, Klungel OH. Gastrointestinal toxicity among patients taking selective Cox-2 inhibitors or conventional NSAIDs alone or combined with proton pump inhibitors: a case control study. Pharmacoepidemiol Drug Saf 2016;**25** (Suppl 3):p444.
22 http://medstat.dk/da/viewDataTables/medicineAndMedicalGroups/%7B%22year%22:[%222004%22,%222003%22,%222002%22],%22region%22:[%220%22],%22gender%22:[%22A%22],%22ageGroup%22:[%22A%22],%22searchVariable%22:[%22people_count%22],%22errorMessages%22:[],%22atcCode%22:[%22M01AE01%22],%22sector%22:[%220%22]%7D.
23 Schmidt M, Hallas J, Friis S. Potential of prescription registries to capture individual-level use of aspirin and other nonsteroidal anti-inflammatory drugs in Denmark: trends in utilization 1999–2012. Clin Epidemiol 2014;**6**:155–68.
24 Mainz J, Kristensen S, Bartels P. Quality improvement and accountability in the Danish health care system. Int J Qual Health Care J Int Soc Qual Health Care 2015;**27**:523–7.
25 Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S *et al.* The US Food and Drug Administration's Mini-Sentinel program: status and direction. Pharmacoepidemiol Drug Saf 2012;**21**(Suppl 1):1–8.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Proportion of patients who have 0,1,2-4,5-9,10-19 or 20 + positive variables as a function of the number of covariates included.