The Potential of High-Dimensional Propensity Scores in Health Services Research: An Exemplary Study on the Quality of Care for Elective Percutaneous Coronary Interventions

Dirk Enders, Christoph Ohlmeier, and Edeltraut Garbe

Objective. Evaluating the potential of the high-dimensional propensity score (HDPS) to control for residual confounding in studies analyzing quality of care based on administrative health insurance data.

Data Source. Secondary data from 2004 to 2009 from three German statutory health insurance providers.

Study Design. We conducted a retrospective cohort study in patients with elective percutaneous coronary interventions (PCIs) and compared the mortality risk between the in- and outpatient setting using Cox regression. Adjustment for predefined confounders was performed using conventional propensity score (PS) techniques. Further, an HDPS was calculated based on predefined and empirically selected confounders from the database.

Principal Findings. Conventional PS methods showed a decreased mortality risk for outpatient compared to inpatient PCIs, while trimming of patients with nonoverlap in the HDPS distribution and weighting resulted in a comparable risk. Most comorbidities were less prevalent in the HDPS-trimmed population compared to the original one.

Conclusion. The HDPS methodology may reduce residual confounding by rendering the studied cohort more comparable through restriction. However, results cannot be generalized for the entire study population. To provide unbiased results, full assessment of all unmeasured confounders from proxy information in the database would be necessary.

Key Words. Residual confounding, unmeasured confounding, administrative data

Administrative data from health insurance providers reflect routine care and are therefore a valuable data source for health services research. However, one of the most prominent limitations of observational studies based on those

data sources is the lack of important confounder information. This is of special importance in studies comparing treatment outcomes related to procedures conducted in the inpatient and outpatient setting where disease severity probably constitutes an important potential confounder.

Due to the increasing specialization of the outpatient setting, percutaneous coronary intervention (PCI), a minimally invasive procedure used to open blocked or narrowed (stenotic) coronary arteries, is increasingly shifted from the in- to the outpatient setting (Robra, Swart, and Vogt 2010). The decision regarding the treatment setting depends on the severity of the underlying disease and the patient's risk factor profile, but also on other factors. In a recent consensus article of the German Cardiac Society and the Medical Advisory Board of the Statutory Health Insurances, the authors stated a list of criteria for which inpatient treatment is recommended (Reinecke et al. 2015). Nevertheless, in practice, many patients treated in the inpatient setting turned out to be eligible for an outpatient treatment as well. Thus, evidence on the comparison of the quality of care in either treatment setting is still necessary. However, this evidence is largely missing, mainly due to differences in quality assurance measures (Boy, Hahn, and Kociemba 2008; Albrecht et al. 2009; AQUA-Institut 2010; Jeschke et al. 2013). Therefore, in a recent study conducted by Ohlmeier et al. (2014) based on German statutory health insurance data, patients with an elective PCI in the inpatient and outpatient setting were compared with regard to potential subsequent complications such as myocardial infarction, ischemic stroke, further coronary intervention, or death. To adjust for confounding, the authors used a multivariable Cox regression analysis, a propensity score (PS)-matched analysis, and an inverse-probability-of-treatment (IPT)-weighted analysis with predefined covariables. They observed that patients in both settings had comparable risks with respect to myocardial infarction, stroke, or coronary reintervention. However, a decrease of the all-cause mortality risk by 42 percent was found for patients with PCIs in the outpatient setting. This observation seemed implausible, as the authors expected similar mortality risks in the inpatient and outpatient setting. At most, a slightly increased mortality risk was expected in the outpatient setting due to a high degree of standardization and close surveillance in the inpatient setting as well as fast access

Address correspondence to Dirk Enders, Dipl.-Math., Leibniz Institute for Prevention Research and Epidemiology – BIPS, Achterstraße 30, 28359 Bremen, Germany; e-mail: enders@leibniz-bips.de. Christoph Ohlmeier, Dr. PH, and Edeltraut Garbe, Dr. Med., are with the Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany. Edeltraut Garbe, Dr. Med., is also with the Core Scientific Area “Health Sciences” at the University of Bremen, Bremen, Germany. Christoph Ohlmeier, Dr. PH, is also with the IGES Institut GmbH, Berlin, Germany.

to the intensive care unit or heart surgery in the case of complications, all of which could lead to lower mortality rates in the inpatient setting (Srinivas et al. 2009). Furthermore, the observed extent of the risk reduction seemed implausible, too, as rare fatal PCI complications should only contribute a small percentage to the all-cause mortality. The authors concluded that, despite a comprehensive confounder adjustment, the observed difference in the mortality risk might be attributable to residual or unmeasured confounding; in particular, the lack of information on the severity of the underlying disease and associated risk factors might have influenced the decision on the PCI setting. Patients with more severe coronary heart disease or serious risk factors might have preferentially been treated in the inpatient setting. However, this could not be verified with the information available in the database.

To control for unmeasured or residual confounding, high-dimensional propensity scores (HDPSs) have been proven useful in past studies based on health insurance databases (Schneeweiss et al. 2009; Garbe et al. 2013). In an HDPS analysis, the underlying database of the study is automatically searched for hundreds of empirical confounders, which are used in the subsequent analyses. Therefore, an HDPS analysis essentially adjusts for a large number of covariables that are confounders in the study at hand. However, under the assumption that these confounders collectively serve as proxies for all unmeasured or residual confounders, an HDPS analysis can adjust for the unmeasured or residual confounders as well (Sauer et al. 2013). However, the assumption that the automatically searched covariates of the HDPS serve as proxies for all unmeasured confounders is debatable and often not verifiable. Most HDPS analyses have been applied in studies concerning the comparative effectiveness of medications. To our knowledge, few studies in health services research applied the HDPS methodology almost exclusively analyzing health care costs (Goudie et al. 2015; Vekeman et al. 2015) and not the quality of care.

In this study, we evaluated the potential of the HDPS methodology to control for residual and unmeasured confounding in studies analyzing the quality of care by reanalyzing the original study by Ohlmeier et al. on the mortality risk of patients with an elective PCI in the in- and outpatient setting.

METHODS

The original study as well as this study share the same data source and study design. However, the statistical analyses differed between both studies as described below.

Data Source

Both studies are based on pseudonymized claims data from three German health insurance providers included in the German Pharmacoepidemiological Research Database (GePaRD) using data from 2004 to 2009. GePaRD contains information on demographic variables, hospitalizations, outpatient physician visits, and outpatient prescriptions. The database has been described in detail elsewhere (Behr, Andersohn, and Garbe 2010).

Design of the Study

We conducted a retrospective cohort study in patients with an elective PCI. Only patients with valid information on sex, year of birth, and region of residence were included. To ensure that only elective PCIs were considered in this study, patients with an inpatient PCI and an admission diagnosis of acute coronary syndrome or recurrent myocardial infarction were excluded. The admission diagnosis was chosen to identify elective PCIs, because the main discharge diagnosis would not have allowed a clear distinction between myocardial infarction leading to hospital admission and myocardial infarction occurring during the hospital stay, for example, as a complication of an elective PCI. Cohort entry was defined as the date of the PCI. Cohort exit was defined as the date of death, the end of the individual insurance period, or the end of the study period (31.12.2009), whichever came first. Patients were considered dead if the reason for the end of an insurance period was coded with death. If cohort exit was not due to death, the respective patient was censored. Covariable information was obtained in the 365 days before cohort entry and comprised age, sex, region of residence, year of cohort entry, 30 diseases of the Elixhauser comorbidity measure (ECM), and other cardiovascular comorbidities (past stroke, recurrent myocardial infarction, past acute coronary artery syndrome, other coronary artery disease, dyslipidemia). The comorbidities were obtained from main and secondary inpatient diagnoses as well as secured ambulatory diagnoses. Coding of PCIs and of all comorbidities is provided in Tables A1 and A2 in the Appendix SA2.

Analyses in the Original Study

In the primary analysis of the original study, the mortality risk was compared between patients with in- and outpatient PCIs using a multivariable Cox regression with the following covariables: year of cohort entry, age at cohort

entry (continuous), sex, number of diseases of the ECM, past acute coronary syndrome, coronary artery disease, dyslipidemia, past stroke, and past acute bypass surgery.

In the PS analyses of the original study, the PS was defined as the probability of an elective PCI in the outpatient setting. It was estimated using a logistic model with the PCI setting as the dependent variable and age, sex, region of residence, comorbidities of the ECM, as well as other comorbidities and interventions as explanatory variables (Table A2 in the Appendix SA2). In the PS-matched analyses, patients in the outpatient setting were matched to patients in the inpatient setting at a ratio of 1 : 1 using greedy 6 - > 1 digit matching (Parsons 2001). In the IPT-weighted analysis, the contribution of each patient to the log-likelihood of the Cox regression was weighted by the inverse of the probability of receiving a PCI in the respective setting. The IPT-weighted analysis was also conducted using stabilized weights, where the weights for patients with a PCI in the outpatient and inpatient setting were multiplied by the proportion of patients in the respective setting. Furthermore, to evaluate the impact of extremely low and high weights, three analyses were also performed using truncated stabilized weights, where the lowest and highest 1, 5, and 10 percent of the stabilized weights were truncated to the 1, 5, 10 percent and 99, 95, and 90 percent quantile of the stabilized weight distribution, respectively. All analyses of the original study revealed a decreased mortality risk for the outpatient setting of at least 40 percent (Figure 2; Ohlmeier et al. 2014).

HDPS Analysis

In this study, a HDPS was calculated using the variables of the PS model above plus automatically selected covariables from the underlying database. The HDPS algorithm aims at identifying the empirically most influential confounders and thus selects all variables strongly associated with the exposure and outcome but also prevalent enough for the confounding to come into effect (Schneeweiss et al. 2009). In the first step of the algorithm, the m different data dimensions in the database and the granularity of the codes in the respective dimension must be specified. “Granularity” is thereby defined as the number of digits of a diagnosis code that were used, and “dimensions” is a term referring to the number of different claims sources used to search diagnosis codes. In our study, we obtained all codes in the 365 days before cohort entry in the following $m = 6$ data dimensions (code type and granularity of the codes in parentheses): inpatient diagnoses (International Classification of

Figure 1: Distribution of the High-Dimensional Propensity Score in the Cohort, Matched Cohort, and Trimmed Cohort

Diseases German Modification [ICD-10 GM] codes, 3-digit), ambulatory diagnoses without exclusion diagnoses (ICD-10 GM codes, 3-digit), inpatient procedures (Operationen- und Prozedurenschlüssel [OPS] codes, 4-digit), outpatient procedures (OPS codes, 4-digit), ambulatory treatments (Einheitlicher Bewertungsmaßstab [EBM] codes, 5-digit), and ambulatory prescribed medications (Anatomical–therapeutic–chemical [ATC] codes, 7-digit). In each data dimension, the prevalence of each code is then determined and the top n codes are selected. We selected the $n = 200$ most prevalent codes for our study. From each of the n codes, three binary variables are constructed representing the recurrence of the documentation of that code, that is, how frequently the code occurred (≥ 1 times, $\geq$ median number of times or ≥ 75 th percentile of the number of times). The $m \times n \times 3$ binary variables are then sorted by the potential to control for confounding based on the bias ranking method (Bross 1966; Schneeweiss et al. 2009). After the sorting, the top k variables are selected for the calculation of the HDPS. In our study, we selected the top $k = 500$ variables from the $m \times n \times 3 = 3,600$ variables to be included in the model to calculate the HDPS.

In studies using the PS (or HDPS), it is important that there is sufficient overlap in the PS (or HDPS) distribution between patients in the two groups of comparison to assure comparability (Stürmer et al. 2014). This is more important in studies using the HDPS, as the inclusion of more than 500 covariables in the calculation of the HDPS likely leads to a good discrimination of patients in both groups. Trimming of patients above and below two cutoff values of the HDPS is generally recommended, although cutoff values are often arbitrary, for example, trimming at some percentile of the HDPS distribution (Stürmer et al. 2010) or trimming below a HDPS value of 0.1 or above 0.9 (Crump et al. 2009). The data-dependent approach by Crump et al. (2009) chooses the cutoff value α in a way that leads to the most precise effect estimate among all datasets that are trimmed below α and above $1-\alpha$ without introducing bias.

In our main HDPS analysis, we trimmed the cohort using the approach by Crump et al. and performed a weighted Cox regression in the trimmed cohort using stabilized IPT weights based on the HDPS. To examine the balance of covariables after trimming and weighting, the standardized difference in the original cohort and the weighted standardized difference in the trimmed cohort were calculated (Austin and Stuart 2015). The validity of the positivity assumption was examined by calculating the range of the stabilized weights in the trimmed cohort.

Analogously to the PS analyses of the original study, we conducted several sensitivity analyses, that is, a HDPS-matched analysis and weighted analyses based on stabilized and truncated IPT weights. Furthermore, to evaluate the influence of the number of covariables used for the calculation of the HDPS, the main and sensitivity analyses were repeated with $k = 100, 200$, and 1,000 selected covariables. If covariables are included in the HDPS that are only weakly associated with the outcome but strongly with the exposure, the precision in the effect estimates might be decreased (Brookhart et al. 2006). Therefore, we also repeated the analyses in which the ranking of the variables was solely based on the bivariate association between the variable and the outcome (outcome ranking method) and calculated the HDPS with the top $k = 500$ variables based on this ranking method.

RESULTS

The study cohort comprised 30,313 patients with an elective PCI, of which 4,269 (14.1 percent) were conducted in the outpatient setting. A flowchart of

the study cohort is displayed in Figure A1 of the Appendix SA2. A detailed description of the characteristics of the cohort can be found in Ohlmeier et al. (2014). The mortality rate in the study cohort was 31.35 (95%-CI: 30.02–32.72) per 1,000 person-years.

The distribution of the HDPS is displayed in Figure 1. The range of the HDPS was 2.8×10^{-10} to 0.980 in the inpatient and 0.001–0.996 in the outpatient setting. Notably, most of the patients in the inpatient setting received a low HDPS smaller than 0.1, while the HDPS in the outpatient setting was uniformly distributed. Trimming excluded patients with a HDPS lower than 0.057 and larger than 0.943 from the cohort. The trimmed population consisted of 7,668 patients treated in hospital and 3,964 treated in the outpatient setting, and the mortality rate reduced to 18.62 (95%-CI: 17.02–20.32) per 1,000 person-years. In Table A3 and A4 of the Appendix SA2, the distribution of the variables included in the PS as well as the top 20 variables selected by the HDPS algorithm are displayed for the excluded and included patients of the trimmed cohort. We observed that, regarding the variables included in the PS model, the prevalence of almost all comorbidities was higher in the excluded patients. Regarding the top 20 selected variables of the HDPS, we observed that besides variables similar to the ones in the PS, unexpected comorbidities such as congenital malformations of cardiac septa or presence of cardiac and vascular implants and grafts and intensive complex care were more prevalent in the excluded patients. On the other hand, presence of cardiologic treatments (code: additional flat rate cardiology II) was reduced after trimming. As seen in Figure 1, the overlap of the HDPS distributions improved after trimming. The stabilized IPT weights in the trimmed population ranged from 0.15 to 14.83 (mean = 1.02, SD = 0.92). After trimming and weighting of the patients, we observed a nonsignificant increase of 20 percent in the mortality risk between the inpatient and the outpatient setting (Figure 2). The trimming and weighting resulted in a standardized difference of less than 0.1 for all variables included in the PS as well as for the top 20 covariables selected by the HDPS algorithm (Figures A2 and A3 in the Appendix SA2). As can be seen by Table A4 and Figure A3 in the Appendix SA2, most of the top 20 empirically selected confounders are closely related to the theory-based confounders (e.g., heart failure, hypertension, renal failure, diabetes mellitus type 2). Thus, as many of the theory-based confounders turn up in the empirically selected covariables as well, we concluded that the HDPS method performs well in selecting important confounders.

The results of the sensitivity analyses are displayed in Figure 2. The HDPS-matched analysis was based on 3,034 (71.1 percent) patients treated in

the outpatient setting and 3,034 (11.6 percent) patients treated in hospital and resulted in an increased mortality risk of 26 percent, although still not significant, between the in- and outpatient setting. Additional truncation of the stabilized weights did not substantially change the results of the IPT-weighted analyses. The results of the HDPS analyses changed only slightly, when 100, 200, or 1,000 covariables instead of 500 were selected for the calculation of the HDPS (Figures A4–A6 in the Appendix SA2). Similarly, only little changes were observed when using the outcome ranking method instead of the bias ranking method (Figure A7 in the Appendix SA2).

DISCUSSION

In this study, we applied HDPS methods to control for residual confounding due to disease severity, which was presumed to have affected a recently conducted study on the comparison of the mortality risk after elective PCI between the in- and outpatient setting, in which a lower mortality risk was observed in the outpatient setting compared to the inpatient setting. Trimming patients with nonoverlap in the HDPS distribution and conducting a stabilized IPT-weighted analysis in this trimmed cohort resulted in a comparable mortality risk between patients treated in the in- and outpatient setting; that is, the presumed implausible direction and magnitude of the risk reduction, which was observed in the original study, were not observed in the HDPS

Figure 2: Hazard Ratio of Death in the Outpatient versus Inpatient Setting in the Adjusted, PS, and HDPS Analyses

Note. HR, Hazard Ratio; LCL, UCL, 95%-Confidence Interval

analysis. The results were consistent across alternative implementations of the HDPS, that is, in an HDPS-matched analysis and in IPT-weighted analyses using truncated weights, and they are robust with respect to the number of covariables and the sorting algorithm used in the HDPS algorithm. However, the results of the HDPS analyses must be interpreted with caution mainly due to potentially limited generalizability and unbiasedness.

Generalizability

First, trimming of the population leads to a restricted population, and thus, the results often cannot be generalized for the entire population. In our example, the original population of 30,313 patients was reduced to 11,632 patients after trimming, which is a reduction of more than 60 percent in the sample size. The same applies for an analysis based on an HDPS-matched cohort. In our study, trimming led to the exclusion of patients with a low HDPS in the inpatient setting. Likewise, the HDPS-matched analysis excluded mostly patients in the inpatient setting with a low HDPS that missed a match with a comparable HDPS in the outpatient setting (Figure 1). Patients with a low HDPS exhibited a high mortality risk (Table 1). These patients were characterized by higher prevalences in most of the manually selected as well as automatically selected comorbidities (Tables A3 and A4 in the Appendix SA2). Thus, the observed comparable mortality risk between the in- and outpatient setting was observed in a population, which was generally healthier than the original population, although the difference between the trimmed and original cohort is not clearly identifiable due to the large number of covariables included in the HDPS.

Unbiasedness

Trimming of patients guarantees that the same range of the HDPS and thus a comparable profile in the covariables included in the calculation of the HDPS is present in patients with inpatient as well as outpatient PCIs. The trimmed population using the automatically selected covariables of the HDPS algorithm is therefore more comparable on observable covariables than the original population or a trimmed population solely based on the manually selected covariables of the PS. Weighting of patients in the trimmed cohort generally leads to a balance in the covariables of the HDPS algorithm. If all unmeasured confounders were fully assessed by proxy information in the HDPS, the weighted and trimmed cohort would be balanced with respect to all measured

Table 1: Mortality Rates in the Overall Cohort and by PCI Setting

	Overall				Inpatient Setting				Outpatient Setting			
	Number of Patients	Cases	Person-time (in 1,000 PY)	Mortality Rate (per 1,000 PY)	Number of Patients	Cases	Person-time (in 1,000 PY)	Mortality Rate (per 1,000 PY)	Number of Patients	Cases	Person-time (in 1,000 PY)	Mortality Rate (per 1,000 PY)
Before matching	30,313	2,087	66.62	31.50	26,044	1,924	57.02	33.74	4,269	163	9.599	16.98
In quintiles												
of the HDPS												
1st quintile	6,062	703	11.30	62.21	6,057	703	11.29	62.28	5	0	0.014	0
(2.8×10^{-10})												
<HDPS ≤ 0.002)												
2nd quintile	6,063	426	13.54	31.46	6,039	425	13.48	31.52	24	1	0.054	18.52
(0.002)												
< HDPS ≤ 0.012)												
3rd quintile	6,063	433	13.84	31.29	5,941	429	13.55	31.61	122	4	0.27	14.93
(0.012)												
< HDPS ≤ 0.052)												
4th quintile	6,062	319	14.24	22.40	5,283	273	12.40	22.00	779	46	1.83	25.07
(0.052)												
< HDPS ≤ 0.246)												
5th quintile	6,063	206	13.70	15.03	2,724	94	6.27	14.98	3339	112	7.43	15.08
(0.246)												
< HDPS < 0.996)												
After matching	6,068	236	14.09	16.75	3,034	109	7.06	15.43	3,034	127	7.03	18.06

and unmeasured confounders and the estimate of the hazard ratio in the trimmed population would be unbiased for the true hazard ratio in that trimmed population (conditional on the prespecified covariables included in the Cox regression model).

However, although the HDPS adjusts for hundreds of covariables and may adjust for proxies of residual confounders, further unmeasured confounders completely unrelated to the information in the underlying database can never be ruled out. Furthermore, an accurate representation of the unmeasured confounders by proxy information depends on the type of unmeasured confounder and is questionable in general. Thus, the effect estimate in the trimmed/matched cohort might still be biased. Furthermore, the bias of the effect estimate might also increase in the HDPS analysis, because imbalance of unmeasured confounders between treated and untreated patients might increase if measured confounders are balanced in the course of the HDPS analysis (Brooks and Ohsfeldt 2013). In our study, the top 20 selected covariables of the HDPS included, for example, presence of cardiac and vascular implants and grafts, intensive complex care, and cardiologic treatments (code: additional flat rate cardiology II) as important empirical confounders, which could be considered indicators of complex cases of coronary artery diseases. However, as other proxies of disease severity likely exist, the HDPS algorithm accounts at most partially for unmeasured confounding of disease severity.

The implementation of the HDPS in an IPT-weighted analysis might also lead to biased results, if individuals with high weights dominate the analysis. This is known as a practical violation of the positivity assumption (Cole and Hernán 2008). As more than 500 covariables were included in the model to calculate the HDPS, some data constellations might lead to extraordinarily high or low weights. The use of stabilized weights is therefore typically recommended to reduce the chance of extremely high or low weights (Robins, Hernan, and Brumback 2000). In case of high stabilized weights, a truncation of the weights can be performed to examine whether the results are affected by extremely high weights (Cole and Hernán 2008). A rule of thumb is to truncate the weights so that the mean of the weights is close to 1 and the maximum weight is less than 10. As the stabilized weights were already in a moderate range after trimming, a truncation of the weights in addition to trimming affected the results only slightly.

Besides issues regarding generalizability and unbiasedness of the results, the additional benefit of the HDPS compared to conventional methods is further not consistently reported in the literature. For example, in case of a study population with a small sample size, one study showed that the HDPS

algorithm only performed better than the conventional PS technique provided that a further correction method was applied (Rassen et al. 2011). Two further studies also found no additional benefit of the HDPS method compared to conventional PS methods (Toh, Garcia Rodriguez, and Hernan 2011; Li et al. 2014), which suggests that the conventional methods already captured all relevant confounders, but also that the HDPS algorithm might still miss important confounding variables.

In this study, we did not perform analyses implementing the HDPS in other ways, that is, stratification with respect to, for example, quintiles of the HDPS or including the HDPS as a covariable into the Cox model. The reason was that we observed heterogeneous effect estimates of the PCI setting on different levels of the HDPS (Table 1). Therefore, a stratified analysis would not be valid (Stürmer et al. 2014). Further, a Cox regression that uses the HDPS as a covariable would be complicated as (1) the correct function of the HDPS in the regression model needs to be specified and (2) interactions with the PCI setting and HDPS need to be included in the model.

STRENGTHS AND LIMITATIONS

We performed comprehensive HDPS analyses in a study based on a large administrative dataset from health insurance providers. As the results of the HDPS might be sensitive to the selected covariables (Rassen and Schneeweiss 2012), we performed several sensitivity analyses with different numbers of selected covariables and with a different sorting algorithm. The results were consistent across the sensitivity analyses.

With the exception of the ICD codes, the maximum detail of the codes was used. In the definition of the granularity of the ICD codes, a four-digit granularity could have been specified to better identify the underlying diseases. However, we decided to consider only on a three-digit granularity to obtain reasonable prevalences of the codes.

The choice of the empirical confounders in HDPS analyses is generally based on bivariate associations of the confounder and the exposure and the bivariate association of the confounder and the outcome. Thus, some of the selected variables might not have a confounding influence in the presence of the other variables. Also, some selected variables might be colliders of some unmeasured variables and thus may introduce bias. However, the impact of bias due to the inclusion of colliders is usually negligible (Myers et al. 2011; Liu et al. 2012).

The exclusion of emergency PCIs in the inpatient setting might not have been successful in all patients. However, the review of patient profiles did not indicate an inclusion of PCIs, for example, due to an acute coronary syndrome to a relevant extent. Furthermore, those patients would have likely been excluded from the matched or trimmed population.

In general, the HDPS method can be applied for any health care database which provides information on a patient-level basis. Ideally, the database is divided into different data dimensions, containing data of different sectors of the underlying health care system. However, the algorithm is implemented in a *SAS* software package, which is only available to evaluate the effect of a binary exposure on a binary outcome at the moment (Rassen et al. 2016). Thus, studies involving more complex definitions of exposure or outcome variables cannot be analyzed easily using the HDPS algorithm.

CONCLUSION

We observed that the large number of empirical confounders identified during the application of the HDPS method allowed for a better discrimination of patients with differing characteristics compared to the PS method. Therefore, patients without adequate controls, who were included in the original PS analyses, are excluded from the HDPS analysis. This is particularly important in studies comparing the quality of care between patient populations with, for example, a differing disease severity, where the indication of treatment options under study might not be available for all patients in the study. Thus, it can be assumed that the HDPS method is useful to reduce residual confounding by restricting the population to a comparable subgroup. However, the HDPS method critically relies on the often not verifiable assumption that proxy information for important unmeasured confounders can be obtained from the underlying database and that all unmeasured confounders are accurately represented by the proxy information. In case unmeasured confounders remain, the HDPS method might even increase the bias. Further, the results are not generalizable for the original study population and thus need to be interpreted with caution.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The authors are grateful to all statutory health insurance providers that provided data for this study, namely the

AOK Bremen/Bremerhaven, the Techniker Krankenkasse (TK), and the Handelskrankenkasse (hkk).

Disclosures: This study was funded by the Central Research Institute of Ambulatory Health Care. The authors had complete autonomy for the process of establishing the protocol, carrying out the analyses, and interpreting the results. This also includes the full right to publish the results without limitation. The authors are working/worked for an institute that occasionally performs studies for pharmaceutical industries. These companies include Bayer, Celgene, GlaxoSmithKline, Mundipharma, Novartis, Sanofi-Aventis, Sanofi Pasteur MDS, and STADA. Edeltraut Garbe has been a consultant to Bayer-Schering, Nycomed, Teva, GlaxoSmithKline, Schwabe, and Novartis. Christoph Ohlmeier left BIPS and joined the IGES Institut GmbH after conduct of the study. Except for finalization of the manuscript, the entire analysis and interpretation of the results was performed under BIPS affiliation.

Disclaimer: None.

REFERENCES

- Albrecht, A., B. Levenson, S. Gohring, W. Haerer, N. Reifart, G. Ringwald, and B. Troger. 2009. "The QuIK-Registry of the German Society of Cardiologists in Private Practice: Countrywide and Benchmarking Quality Assurance in Invasive Cardiology." *Deutsche Medizinische Wochenschrift* 134 (suppl 6): S211–3.
- AQUA-Institut. 2010. *Qualitätsreport 2009*. Göttingen: AQUA-Institut.
- Austin, P. C., and E. A. Stuart. 2015. "Moving Towards Best Practice When Using Inverse Probability of Treatment Weighting (IPTW) Using the Propensity Score to Estimate Causal Treatment Effects in Observational Studies." *Statistics in Medicine* 34 (28): 3661–79.
- Behr, S., F. Andersohn, and E. Garbe. 2010. "Risk of Intracerebral Hemorrhage Associated With Phenprocoumon Exposure: A Nested Case-Control Study in a Large Population-Based German Database." *Pharmacoepidemiology and Drug Safety* 19 (7): 722–30.
- Boy, O., S. Hahn, E. Kociemba, and BQS-Fachgruppe Kardiologie. 2008. *Koronarangiographie und Perkutane Koronarinterventionen (PCI)*. Düsseldorf: BQS Institut.
- Brookhart, M. A., S. Schneeweiss, K. J. Rothman, R. J. Glynn, J. Avorn, and T. Sturmer. 2006. "Variable Selection for Propensity Score Models." *American Journal of Epidemiology* 163 (12): 1149–56.
- Brooks, J. M., and R. L. Ohsfeldt. 2013. "Squeezing the Balloon: Propensity Scores and Unmeasured Covariate Balance." *Health Services Research* 48 (4): 1487–507.
- Bross, I. D. J. 1966. "Spurious Effects from an Extraneous Variable." *Journal of Chronic Diseases* 19 (6): 637–47.
- Cole, S. R., and M. A. Hernán. 2008. "Constructing Inverse Probability Weights for Marginal Structural Models." *American Journal of Epidemiology* 168 (6): 656–64.

- Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik. 2009. "Dealing with Limited Overlap in Estimation of Average Treatment Effects." *Biometrika* 96 (1): 187–99.
- Garbe, E., S. Kloss, M. Suling, I. Pigeot, and S. Schneeweiss. 2013. "High-Dimensional Versus Conventional Propensity Scores in a Comparative Effectiveness Study of Coxibs and Reduced Upper Gastrointestinal Complications." *European Journal of Clinical Pharmacology* 69 (3): 549–57.
- Goudie, A., L. Dynan, P. W. Brady, E. Fieldston, R. J. Brilli, and K. E. Walsh. 2015. "Costs of Venous Thromboembolism, Catheter-Associated Urinary Tract Infection, and Pressure Ulcer." *Pediatrics* 136 (3): 432–9.
- Jeschke, E., H. T. Baberg, P. Dirschedl, K. Heyde, B. Levenson, J. Malzahn, T. Mansky, M. Mockel, and C. Gunster. 2013. "Complication Rates and Secondary Interventions After Coronary Procedures in Clinical Routine: 1-Year Follow-up Based on Routine Data of a German Health Insurance Company." *Deutsche Medizinische Wochenschrift* 138 (12): 570–5.
- Li, L., W. M. Vollmer, M. G. Butler, P. Wu, E. O. Kharbanda, and A. C. Wu. 2014. "A Comparison of Confounding Adjustment Methods for Assessment of Asthma Controller Medication Effectiveness." *American Journal of Epidemiology* 179 (5): 648–59.
- Liu, W., M. A. Brookhart, S. Schneeweiss, X. Mi, and S. Setoguchi. 2012. "Implications of M Bias in Epidemiologic Studies: A Simulation Study." *American Journal of Epidemiology* 176 (10): 938–48.
- Myers, J. A., J. A. Rassen, J. J. Gagne, K. F. Huybrechts, S. Schneeweiss, K. J. Rothman, M. M. Joffe, and R. J. Glynn. 2011. "Effects of Adjusting for Instrumental Variables on Bias and Precision of Effect Estimates." *American Journal of Epidemiology* 174 (11): 1213–22.
- Ohlmeier, C., R. Linder, D. Enders, R. Mikolajczyk, W. Haverkamp, D. Horenkamp-Sonntag, and E. Garbe. 2014. "Evaluating Methods for Intersectoral Comparison of Quality of Care. A Routine Data Analysis of Elective Percutaneous Coronary Interventions." *Methods of Information in Medicine* 53 (4): 269–77.
- Parsons, L. S. 2001. "Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques. (Paper 214-26)." In *Proceedings of the 26th Annual SAS Users Group International Conference, Long Beach, California, April 22–25, 2001*. Cary, NC: SAS Institute, Inc.
- Rassen, J. A., and S. Schneeweiss. 2012. "Using High-Dimensional Propensity Scores to Automate Confounding Control in a Distributed Medical Product Safety Surveillance System." *Pharmacoepidemiology and Drug Safety* 21 (suppl 1): 41–9.
- Rassen, J. A., R. J. Glynn, M. A. Brookhart, and S. Schneeweiss. 2011. "Covariate Selection in High-Dimensional Propensity Score Analyses of Treatment Effects in Small Samples." *American Journal of Epidemiology* 173 (12): 1404–13.
- Rassen, J. A., M. Doherty, W. Huang, and S. Schneeweiss. 2016. "Pharmacoepidemiology Toolbox" [accessed on May 20, 2016]. Available at <http://www.hdpharmacoepepi.org>
- Reinecke, H., M. Braun, L. Frankenstein, G. Gorge, A. Kerlin, S. Knoblich, Y. von Kodolitsch, B. Lengsfelder, B. Levenson, D. Pfeiffer, B. Reichle, G. Steinbeck,

- J. Reinöhl, and P. Dirschedl. 2015. "Criteria for the Necessity and Duration of Hospital Treatment for Coronary Angiographies and Coronary Interventions." *Der Kardiologe* 9 (4): 295–302.
- Robins, J. M., M. A. Hernan, and B. Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11 (5): 550–60.
- Robra, P., E. Swart, and T. Vogt. 2010. *Veränderungen des Umfangs der Vertragsärztlichen Leistungen Durch Leistungsverlagerungen Zwischen dem Stationären und dem Ambulanten Sektor*. Magdeburg: ISMHE.
- Sauer, B., M. A. Brookhart, J. A. Roy, and T. J. Vanderweele. 2013. "Covariate Selection." In *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099*, edited by P. Valentgas, N. A. Dreyer, P. Nourjah, S. R. Smith, and M. M. Torchia, pp. 93–108. Rockville, MD: Agency for Healthcare Research and Quality.
- Schneeweiss, S., J. A. Rassen, R. J. Glynn, J. Avorn, H. Mogun, and M. A. Brookhart. 2009. "High-Dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data." *Epidemiology* 20 (4): 512–22.
- Srinivas, V. S., S. M. Hailpern, E. Koss, E. S. Monrad, and M. H. Alderman. 2009. "Effect of Physician Volume on the Relationship Between Hospital Volume and Mortality during Primary Angioplasty." *Journal of the American College of Cardiology* 53 (7): 574–9.
- Stürmer, T., R. Wyss, R. J. Glynn, and M. A. Brookhart. 2014. "Propensity Scores for Confounder Adjustment When Assessing the Effects of Medical Interventions Using Nonexperimental Study Designs." *Journal of Internal Medicine* 275 (6): 570–80.
- Stürmer, T., K. J. Rothman, J. Avorn, and R. J. Glynn. 2010. "Treatment Effects in the Presence of Unmeasured Confounding: Dealing with Observations in the Tails of the Propensity Score Distribution—A Simulation Study." *American Journal of Epidemiology* 172 (7): 843–54.
- Toh, S., L. A. Garcia Rodriguez, and M. A. Hernan. 2011. "Confounding Adjustment via a Semi-Automated High-Dimensional Propensity Score Algorithm: An Application to Electronic Medical Records." *Pharmacoepidemiology and Drug Safety* 20 (8): 849–57.
- Vekeman, F., M. Gauthier-Loiselle, E. Faust, P. Lefebvre, R. Lahoz, M. S. Duh, and P. Sacco. 2015. "Patient and Caregiver Burden Associated with Fragile x Syndrome in the United States." *Am J Intellect Dev Disabil* 120 (5): 444–59.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Appendix SA2: Operationalization of Variables and Additional Results.