

(Dis)concordance of comorbidity data across ICD codes, medical charts, and self-reports

Abdullah Sherifdeen (✉ abdullah.sherifdeen@monash.edu)

Monash University Faculty of Medicine Nursing and Health Sciences <https://orcid.org/0000-0003-2028-4170>

Jeremy Millar

Monash University Faculty of Medicine Nursing and Health Sciences

Catherine Martin

Monash University Faculty of Medicine Nursing and Health Sciences

Melanie Evans

Monash University Faculty of Medicine Nursing and Health Sciences

Gabriella Tikellis

Monash University Faculty of Medicine Nursing and Health Sciences

Susan Evans

Monash University Faculty of Medicine Nursing and Health Sciences

Research article

Keywords: Prostate cancer, concordance, self-reports, comorbidities

Posted Date: September 12th, 2019

DOI: <https://doi.org/10.21203/rs.2.14363/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on September 11th, 2020. See the published version at <https://doi.org/10.1186/s12913-020-05713-5>.

Abstract

Background Benchmarking outcomes across settings commonly requires risk-adjustment for co-morbidities that must be derived from extant sources that were designed for other purposes. A question arises as to the extent to which differing available sources for health data will be concordant when inferring the type and severity of co-morbidities, how close these are to the “truth”. We studied the level of concordance for same-patient comorbidity data extracted from administrative data (coded from International Classification of Diseases, Australian modification, 10th edition [ICD-10AM]), from the medical chart audit, and data self-reported by men with prostate cancer who had undergone a radical prostatectomy.

Methods We included six hospitals (5 public and 1 private) contributing to the Prostate Cancer Outcomes Registry-Victoria (PCOR-Vic) in the study. We listed eligible patients from the PCOR-Vic who underwent a radical prostatectomy between January 2017 and April 2018 for the Health Information Manager in each hospital, who provided each patient’s associated ICD-10AM comorbidity codes. Medical charts were reviewed to extract comorbidities used to generate the Charlson Comorbidity Index. The self-reported comorbidity questionnaire (SCQ) was distributed through PCOR-Vic to eligible men.

Results The percentage agreement between the administrative data, medical charts and self-reports ranged from 92% to 99% in the 122 patients (from 217 eligible participants, 56%), who responded to the questionnaire. The prevalence-adjusted bias-adjusted kappa (PABAK) coefficient was from 0.83 to 0.98 for all conditions aside from cancer, reflecting a strong level of agreement for the absence of comorbidities. Conversely, the presence of comorbidities showed a poor level of agreement between data sources. There was concordance on 213/277 (77%) comorbidities when comparing medical charts and administrative data; 102/238 (30%) comorbidities when comparing medical chart and self-reports; and 34/150 (23%) comorbidities when comparing administrative data and self-reports.

Conclusion Relying on a single data source to generate comorbidity indices for risk-modelling purposes may fail to capture the reality of a patient’s disease profile. There does not appear to be a ‘gold-standard’ data source for the collection of data on comorbidities.

Background

Prostate cancer (CaP) is the most common non-skin cancer in men worldwide.(1) In Australia it represents the second leading cause of cancer-related mortality in males.(2) Optimal disease management prevents progression of the cancer and preserves quality of life through avoidance of unnecessary treatment. The need to monitor the health and wellbeing of men with CaP has been recognised by physicians and healthcare workers, leading to the establishment of clinical quality registries. A clinical registry “collect(s) uniform data to evaluate specific outcomes for a population defined by a specific disease, condition or exposure that serves one or more predefined scientific, clinical or policy purpose”.(3)

The Prostate Cancer Outcome Registry-Victoria (PCOR-Vic) was developed in 2009 as a clinical quality registry, to measure and report on quality of care, using benchmarking of performance at a clinician and hospital level. Benchmarking is one of the most effective strategies for quality improvement, as it provides useful information to medical professionals on where to improve clinical practice.(4) PCOR-Vic collects data on aspects relating to the diagnosis, treatment, quality of life and outcomes of men diagnosed with CaP.(5) Quality of care is examined using specific quality of care indicators (see supplementary *Table 1*); selected after a literature review and a prioritisation process (modified Delphi process (6)) which considered the importance and feasibility of each proposed quality indicator. Indicators were categorised according to whether they assessed structures, processes or outcomes of care.(7) Both clinical (e.g. positive surgical margins and 90-day prostate cancer specific mortality (PCSM) following treatment) and patient-reported outcomes (urinary, bowel and sexual quality of life) are reported.

Positive surgical margins and 90-day mortality are currently risk-adjusted using the National Comprehensive Cancer Network (NCCN) risk model, which considers the clinical stage prior to treatment, the prostate specific antigen (PSA) level and the Gleason score at biopsy, but not patient comorbidities. (8) Patient-reported outcomes are stratified according to the treatment modalities (surgery, radiotherapy, androgen deprivation therapy (ADT), active surveillance/watchful waiting). If comorbidities are found to be independently associated with outcomes such as positive surgical margins (PSM) and PCSM, it may be appropriate to include them in the risk models used to generate benchmark reports for hospitals and clinicians. Additionally, the inclusion of comorbidities in modelling may be used for pre-operative risk stratification, provision of information during the informed consent process with patients and their families, as well as in decision-making regarding the suitability for patients to undergo surgery.

However, prior to adjusting for comorbidities in the risk models, it is essential to understand the extent to which they are accurately captured in the data source from which they will be extracted. Comorbidity data may be derived from International Classification of Diseases (ICD) codes, manually abstracted from medical charts or collected directly from patients. It is unclear as to whether there would be concordance between each of these data sources and what would be considered the 'gold standard'. Clinical notes might be incomplete, the patient may not be informed of all comorbidities and coding system may not document all comorbidities.

The aim of this project was to examine the completeness and agreement of comorbidity data obtained from three data sources: ICD coded administrative data, medical charts and self-reports in men contributing to PCOR-Vic who have had a radical prostatectomy following a diagnosis of prostate cancer.

Materials And Methods

Study design

A retrospective cohort study design was employed for this project. Men contributing to PCOR-Vic who had undergone a radical prostatectomy between January 2017 and April 2018 at one of six convenient hospitals were eligible to participate in this study. The five public hospitals included four metropolitan and one regional hospital, while the private hospital was located in metropolitan Melbourne. These six major health care facilities were chosen to maximise the sample size at each site.

A patient information and consent form, along with the questionnaire, were sent to eligible men.

Comorbidity data sources

Three data sources were used to compare comorbidities:

- The Charlson Comorbidity Index (CCI) was calculated from the ICD–10AM codes, using work developed by Sundararajan et al.(9,10) The CCI is a universally employed index used to quantify comorbidity in ill patients, with weightings being assigned to each comorbidity depending on severity (see supplementary Table 2).(11) It captures 19 comorbidities. For the purposes of inter-data comparison, a modified weighted CCI (mCCI) was developed which contained only the eight variables present in both the CCI and the self-reported comorbidity questionnaire (SCQ). As the CCI weighted three comorbidities (liver disease, diabetes, cancer) based on their severity (moderate or severe), two mCCI scores were computed for each patient to factor in the possibility that a patient had a milder or more severe version of the particular comorbidity, as the severity of each of these three conditions was not captured in the self-reported comorbidity survey.
- The SCQ made available to men was to be completed either on paper, online or via telephone (see supplementary Table 3).(12) The questionnaire included eight comorbidities captured by the CCI and five additional comorbidities: hypertension, depression, osteoarthritis, back pain and dementia.
- A medical chart audit was undertaken to capture CCI comorbidities reported in both the CCI and the SCQ, identified during the hospital admission related to the radical prostatectomy. For patients treated in the private hospital, the medical audit was undertaken using medical charts held in consulting rooms.

Statistical analysis

Characteristics of the respondents are reported as medians and interquartile ranges due to the skewed distribution or ordinal nature of variables such as age, PSA levels, Gleason score. The Index of Relative Socio-economic Advantage and Disadvantage (IRSAD) was used to assess socio-economic status. An IRSAD score of 0 represented the greatest societal disadvantage and a score of ten represented the greatest societal advantage.(13) Statistical significance was defined as a p-value < 0.05.

Concordance was calculated using (1) the kappa statistic, (2) the PABAK statistic (kappa adjusted for prevalence bias); and (3) the Gwet's AC1 statistic which calculates the "conditional probability that two

randomly selected raters will agree, given that no agreement will occur by chance".(14) The Gwet's AC1 statistic is founded on the assumption that chance agreement should not exceed 0.5.(18) Previous studies have highlighted how this statistic is resistant to changes in prevalence, and remains close to the percentage of agreement.(15) The level of agreement for the three inter-rater statistics was defined according to Supplementary *Table 4*.(16) "Poor" agreement was defined as a score less than zero, "slight" agreement was defined as a score of 0 to 0.20, "fair" agreement was defined as a score of 0.21 to 0.40, "moderate" agreement was defined as a score of 0.41 to 0.60, "substantial" agreement was defined as a score of 0.61 to 0.80 and "almost perfect" agreement was defined as a score of 0.81 to 1.00.

Stata 15 (StataCorp, College Station, Texas) was used for data analysis.

The ethics application for this study was approved by the Monash University Human Research Ethics Committee (HREC/18/MonH/62). Governance approvals were obtained from each hospital.

Results

The recruitment frame is described in *Figure 1*. *Table 1* describes the characteristics of the cohort contributing to this analysis. The median age of the 217 male participants was 66 years (IQR, 9 years). The median PSA level and Gleason score at diagnosis was 6.91 ng/mL (IQR, 4.7 ng/mL) and 7 (IQR, 1.), respectively. The majority of patients had an IRSAD score ≥ 8 indicating they resided in a more advantageous socio-economic area. Most patients in this study were classified in the NCCN intermediate risk group.

The median CCI and mCCI score based on the administrative data and medical chart audit were each 2 (IQR, 0). The median mCCI score based on the patient self-reports' mild and severe assumptions were 0 (IQR, 0), due to the large proportion of patients not self-reporting any comorbidity. With the assumption that all self-reported patients had cancer, the median mCCI score based on the patient self-reports' mild and severe assumptions was 2 (IQR, 0) and 6 (IQR, 0), respectively. All further analysis report only on comparisons across the three data sources using the mCCI.

The median comorbidity count, for the administrative data, self reports and medical chart review were 1 (IQR, 0), 1 (IQR, 1), and 2 (IQR, 2) respectively (*Table 1*).

Figure 2A-D depicts the distribution of comorbidities for the 112 patients for whom data was available in all three sources, while *Figure 2E* and *Figure 2F* demonstrates distribution of comorbidities for the 201 patients who had data recorded in both the administrative datasets and medical charts. All distributions follow a discrete number scale due to the CCI being a non-continuous comorbidity index.

Based on administrative and medical chart distributions (*Figure 2A-B*), no patients had a CCI score of 0 or 1, as cancer was recorded for each patient, and this carries a weighting of 2 in the CCI. In 22.3% of patients, the administrative dataset had a mCCI score ≥ 3 , in contrast to only 5.7% of patients in the medical charts. The patient self-report CCI distributions indicate that 56.3% of patients had CCI of zero

(no self-reported comorbidities) (*Figure 2C-D*). In 0.9% of patients, the CCI score was ≥ 6 based on the mild comorbidity assumption (*Figure 2C*), with this figure increasing to 25% for the severe comorbidity assumption (*Figure 2D*).

The patient self-report distributions were notably more spread compared to the administrative dataset and medical chart-based distributions.

Medical charts vs administrative datasets

Table 2 reports on the level of statistical concordance between the medical chart and administrative datasets. In total 201 cases were compared; 272 comorbidities were identified through audit of medical chart and 218 through administrative data. There was a match on 213/277 (77%) unique comorbidities reported across both datasets. Four out of the thirteen analysed conditions were unable to have a reporting statistic computed due to the prevalence of the condition being either 0% or 100% in both of the compared datasets. Of the remaining comorbidities, the kappa statistic reported “poor” concordance between the medical chart and administrative data other than for diabetes, which had a “substantial agreement”. According to the PABAK and the Gwet AC1 statistic, nine out of thirteen conditions had “almost perfect” agreement between both data sources, reflecting the very low incidence of these comorbidities in the cohort.

Medical chart vs self-reports

Table 3 compares comorbidities captured by medical chart and SCQ data. Of the 121 cases present in both cohorts, medical charts detected 237 comorbidities, while patients reported 203 comorbidities. However, only 102 of the 338 unique comorbidities (30%) were matched between the two datasets. Based on the kappa score, concordance ranged from “poor” to “substantial”, with the best agreement being attributed to diabetes with a kappa score of 0.71. Agreement was considered as “substantial” for eight of thirteen conditions based on both the PABAK and Gwet’s AC1 statistic. Cancer was classified to have extremely “poor” concordance, with negative concordance scores being cited. Back pain and osteoarthritis were suggested to have “moderate” agreement based on the PABAK score, but “substantial” agreement based on the Gwet’s AC1 statistic.

Administrative data vs self-reports

Table 4 summarises the level of concordance between administrative data and patient self-report data using the SCQ. Only eight conditions were comparable. Of the 121 charts compared between these two datasets, administrative data identified 120 comorbidities while 64 were self-reported by patients. There was a match on 34 of the 150 comorbidities (23%) across the two groups. Due to the relatively small number of cases reported for each comorbidity in both datasets, the kappa score reported “poor” agreement between the administrative data and self-reports, with the exception of diabetes which had

“almost perfect” agreement. With the exception of cancer, the PABAK and Gwet AC1 statistic reported “almost perfect” agreement between both datasets for the remaining conditions.

Comparison across the three data sources

Table 5 provides a summary of the concordance of comorbidity data across the three data sources. Apart from diabetes, the Kappa statistic suggested a poor to slight agreement for seven of the eight comorbidities. In contrast, the PABAK and Gwet’s AC1 statistic suggested almost perfect agreement across the datasets, aside from cancer. Based on all three concordance statistics, cancer had the lowest level of agreement across the three datasets.

Discussion

This study examined the concordance of comorbidity data extracted from administrative data, medical charts and self-reported. Our findings showed that agreement between comorbidity data collected by the Victorian Admitted Episodic Dataset, medical charts, and self-reports by men with CaP who have undergone a radical prostatectomy varied across the analysed comorbidities.

In terms of the comparison of all three data sources, a “slight”, less than “moderate” agreement was observed for the majority of comorbidities when calculated using the kappa. However, the PABAK and Gwet’s AC1 statistic suggested an “almost perfect” agreement (asides from cancer) for all conditions. Interpretation of these findings is difficult when contradictory results are presented depending on the statistical test used. The higher agreement for the latter two statistics was due to a substantial number of patients not having certain diseases based on ≥ 1 data sources.

In light of this, the raw data for this study was analysed when investigating inter-rater concordance. No studies have previously compared these three data sources, so it is not possible to discuss our findings in context with others. However, studies have assessed concordance between two of these three groups and these are discussed below.

Concordance between self-reports and other data sources

Cancer was one of the most interesting self-reported conditions not reported by the majority of patients with only 26% of respondents reporting the condition. A negative statistic was reported with the kappa, PABAK and Gwet’s AC1 when this condition was analysed between the self-reports and other datasets, highlighting “poor” agreement. Additionally, when comparing medical charts and self-reports, 74% ($n = 121$) of patients had cancer in the medical charts, even though they did not self-report the condition themselves. A similar observation was noted with the administrative datasets, with 76% of patients compared with the self-reports failing to state that they had cancer even though this was recorded in the administrative dataset.

The finding that men did not report having cancer may be due to the framing of the SCQ question that asked: “Do you have cancer?”. Given that the selected cohort of patients had undergone surgery it may be that they believe having the prostate gland removed meant they no longer had cancer. However, while for many men this is the case, 27–53% of men will develop prostate-specific antigen (PSA) recurrence up to 10 years after radical prostatectomy and 16–35% of patients receive second-line treatment within five years of surgery.(17)

Other studies have noted that the framing of a question relating to conditions such as cancer can impact reporting. Katz et al analysed concordance of Charlson comorbidities between the medical chart and self-reports, and found that there was only “moderate” concordance (kappa of 0.45 (95% CI: 0.28 to 0.62)) between the medical chart and self-reports in regard to the presence of cancer.(18) The researchers had intentionally omitted asking patients whether their tumour (if present) had been treated within the last 5 years. More interestingly, the Spearman correlation (a statistical measure used to quantify the statistical dependence between the ranking of two variables) between the comorbidity questionnaire and the medical chart-based Charlson index increased from 0.63 (P = 0.0001) to 0.70 (P = 0.0001) after the tumour condition was excluded. This highlights how the explicit wording of questions relating to cancer can result in various levels of recall bias.

There was a three-fold difference in reporting of depression (6 vs 17 reports), a nine-fold difference in reporting of back pain (4 vs 38 reports) and a nearly two-fold difference in reporting of osteoarthritis between medical charts and self-reports. Studies have highlighted that these conditions, in particular depression, are under-diagnosed in hospitals, with physicians not actively investigating whether patients have the condition due to its non-acute nature.(19) This has been especially noted with back pain, a chronic but non-life threatening comorbidity.(20) Patients with osteoarthritis have been noted to view their condition as part of the ageing process (21), with 50% of patients with severe knee pain not reporting to their physician about it (22) and thus reducing their inclination to actively seek medical assistance for their condition. In a study of 2380 community-dwelling patients aged 55–85 years a comparison of medical charts and self-report data showed a kappa of 0.31 (95% CI: 0.27–0.35) for osteoarthritis, with 21.8% of patients stating that they were affected by this condition even though their medical chart did not support the claim.(23) Those with a mobility limitation were more likely to self-report their condition than those without this limitation (OR: 2.68, 95% CI 2.10–3.44).(23) Ultimately, this highlights that patient- and physician- specific views towards certain comorbidities can influence their tendency of being recorded in the medical charts or self-reports.

Concordance between administrative data versus other datasets

In this study, we found that there was strong concordance between the administrative comorbidity data and the other studied data sources based on the reporting statistics used. However, after a more nuanced and deeper investigation, agreement with regard to the reporting of comorbidities was modest, at best.

With conditions such chronic pulmonary disease and rheumatic disease, cases were often recorded in the medical chart but not in the administrative datasets.

Although the PABAK and Gwet's AC1 statistic suggested "almost perfect" agreement for the majority of conditions between the administrative datasets and the other data sources, seven of the thirteen comorbidities compared between the medical charts and administrative datasets were identified to have cases recorded in the former but not in the latter data source. This was visibly depicted in the CCI distributions of both datasets, with 26% of patients having a $CCI \geq 3$ in the medical charts in contrast to only 7% of patients in the administrative datasets. Of these seven conditions, chronic pulmonary disease was recorded 25 times in the medical chart but was not coded in the administrative dataset. A similar observation was seen with rheumatic disease and cerebrovascular disease, where it was recorded in the medical chart of 8 patients but was not coded for in the administrative dataset. There may be a few reasons for this observed discrepancy. Several studies have identified that hospital coders prioritise the coding of symptomatic comorbidities over asymptomatic ones, due to the higher level of hospital funding associated with the former (24) or due to Australian guidelines dictating that "additional diagnoses can only be assigned if they affect patient care during admission".(25, 26) The first point was somewhat proven in the case of chronic pulmonary disease in this study, as it is a condition that may only manifest depending on certain environmental and physiological stimuli.(27) The Australian guidelines surrounding coding comorbidities is interesting, given that in the Australian Modification of the ICD-10AM codes, coders are provided with ample space (fifty slots) to state any secondary diagnoses of a particular patient.(26) Other studies have highlighted how the level of experience of the hospital coder can impact the accuracy of the ICD-10AM codes.(28) Differences in coding practice between inexperienced and experienced coders have been shown to exist.(24, 29, 30, 31)

The "substantial" to "almost perfect" agreement reported by the kappa, PABAK and Gwet's AC1 statistic for diabetes may be attributed to the highly scrutinized nature of this condition in clinical settings.(32, 33) Clinical coders are required to document ICD codes for conditions which require health services resources.(24) The codes are used to assign a Diagnosis Related Group (DRG), which in turn translates to funding for the health service. For patients with Type I diabetes, blood sugar levels are usually required at least twice daily, and insulin must be administered, usually by nursing staff. Type II diabetes is a condition that is monitored strictly within hospitals.(34)

Strengths and limitations:

This study has several strengths. This is the first time that comorbidities have been compared across medical charts, administrative data and patient self-reports. Given the increasing focus on the use of patient reported data, this study enhances our knowledge on the reliability and accuracy of such data. This is particularly important as self-reported comorbidity data is a cost-effective way of collecting comorbidity data compared to the extraction of data from medical charts or administrative datasets.(35)

While the SCQ survey has been validated, with good test-retest reliability being reported (36, 37), this is the first time that it has been examined in a prostate cancer population. This is despite the fact that it has been recommended by ICHOM as the preferred method for collecting comorbidities in men with localised prostate cancer (38). Our finding that less than 30% of self-reported comorbidities reported by the patient appeared in the medical chart or coded administrative data, suggests that more research is required before we can use it to risk adjust health outcomes. While pre-operative administration of the SCQ will likely improve the likelihood that patients will self-report cancer, there remain other comorbidities such as heart disease, which were reported by patients but not documented in the other data sources.

However, this study has a number of limitations which will impact the interpretation of the findings. One relates to the potential discrepancy in interpretation of the definitions for each comorbidity. A list of ICD-10AM codes pertaining to certain conditions such as heart disease, lung disease and kidney disease were used to identify whether a patient was regarded to have the condition or not. These same criteria were not used uniformly in the other data sources. Indeed, the SCQ purposely diluted the complexity of the comorbidity labels to allow the conditions to be understood by patients "without any prior medical knowledge".(12) This likely impacted concordance. For example, in this analysis, ischaemic heart disease (IHD) was not classified as "heart disease" as the ICD-10AM codes pertaining to the two Charlson comorbidities of heart disease (myocardial infarction and congestive heart failure) did not consider IHD. However, patients with the condition may have self-reported it as heart disease.

Non-response bias may have influenced our results, given that we only had a response rate of 55.3%. It is not possible to know whether the non-responders would differ systematically from responders in terms of their comorbidities and whether they were recorded in the medical chart and administrative data. Another bias likely introduced into this study relates to recall bias as men may have had trouble recalling comorbidities.

The sample size for this study was relatively small (N = 217), preventing sub-group analysis, such as whether there were differences based on type of hospital (public/private), where documentation practices may differ. Also, more nuanced findings were unable to be revealed, as shown in the difference in CCI distributions for the administrative datasets for n = 112 (population whom had data across all three data sources) and n = 201 (sample who had data across administrative dataset and medical charts).

Conclusions

This study has shown that there are limitations in the recording of comorbidities for each dataset, that impacts the ability to generate comorbidities indices. Using the SCQ as a risk adjustment tool will likely over represent comorbidities when compared to the use of medical reports or administrative data. The Charlson score, initially developed to be calculated from data extracted from medical charts, is likely to under-represent the presence of comorbidities compared to using administrative codes. In terms of raw counts of comorbidities, there were 272 and 218 in the medical charts and administrative codes, respectively.

Given the relative discordance of comorbidity data, it appears that computing comorbidity indices from a chosen data source may not accurately capture the true health profile of the patient. This is a significant barrier that requires further investigations to determine the prognostic capacity of comorbidities on PSM or PCSM, outcomes that are currently reported by PCOR-Vic.

Further work is required to ensure that comorbidity-specific information is accurately collected for eventual risk-modelling purposes in hospital and clinician benchmarking reports. The SCQ can be tailored to include non-medical jargon that is familiar to patients, and possibly include more contemporary questions. Future studies should also ensure that more robust statistical measures are employed when reporting on the level of agreement.

Abbreviations

ICD-10AM: International Classification of Diseases-10th edition (Australian modified)

PCOR-Vic: Prostate Cancer Outcomes Registry-Victoria

SCQ: Self-reported comorbidity questionnaire

PABAK: Prevalence-adjusted bias-adjusted kappa

CaP: Prostate cancer

PCSM: Prostate cancer specific mortality

NCCN: National Comprehensive Cancer Network

PSA: Prostate specific antigen

ADT: Androgen deprivation therapy

PSM: Positive surgical margins

CCI: Charlson Comorbidity Index

mCCI: Modified weighted CCI

Declarations

Ethics approval and consent to participate

The ethics application for this study was approved by the Monash University Human Research Ethics Committee (HREC/18/MonH/62). Governance approvals were obtained from each hospital.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due to confidentiality but may be partially available from the corresponding author on reasonable request.

Competing interests

No relevant disclosures

Funding

No relevant disclosures

Authors' contributions

AS was the primary investigator for this research project. JM provided clinical expertise in the data analysis phase. CT offered guidance with statistical analysis. ME played a key role in the data extraction process from the Prostate Cancer Outcomes Registry (PCOR). GT was a contributor in the drafting of the project's results. SE overlooked and supervised the project.

Acknowledgements

Not applicable

References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *Cancer J Clin.* 2014;64(1):9–29.
2. Cancer Australia. Prostate cancer statistics: Australian Government 2018 [9/10/2018]. Available from: <https://prostate-cancer.canceraustralia.gov.au/statistics>.
3. Gliklich RE, Dreyer NA, Leavy MB. *AHRQ Methods for Effective Health Care. Registries for Evaluating Patient Outcomes: A User's Guide.* Rockville (MD): Agency for Healthcare Research and Quality (US); 2014.
4. McNeil J, Evans S, Brennan A, Read C., editor. *Registry Science Handbook.* Melbourne; Monash University; 2017.

5. Evans SM, Millar JL, Wood JM, Davis ID, Bolton D, Giles GG, et al. The Prostate Cancer Registry: monitoring patterns and quality of care for men diagnosed with prostate cancer. *BJU International*. 2013;111(4b):E158-E166.
6. Nag N, Millar J, Davis ID, Costello S, Duthie JB, Mark S, et al. Development of Indicators to Assess Quality of Care for Prostate Cancer. *EurUrol Focus*. 2018;4(1):57–63.
7. Reese AC, Cowan JE, Brajtbord JS, Harris CR, Carroll PR, Cooperberg MR. The quantitative Gleason score improves prostate cancer risk assessment. *Cancer*. 2012;118(24):6046–6054.
8. National Comprehensive Cancer Network. National Comprehensive Cancer Network: NCCN History 2017 [9/10/2018]. Available from: <https://www.nccn.org/about/history.aspx>.
9. Sundararajan V, Henderson T, Perry C, Muggivan A, Quan H, Ghali W. New ICD–10 version of the Charlson Comorbidity Index predicted in-hospital mortality. *J Clin Epi* 2004. 2004;57:1288–94.
10. Sundararajan V, Henderson T, Perry C, Muggivan A, Quan H, Ghali W. New ICD–10 version of the Charlson Comorbidity Index predicted in-hospital mortality. *J Clin Epi* 2004. 2004;57:1288–94.
11. Charlson M, Pompei P, Ales K, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–383.
12. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The self-administered comorbidity questionnaire: A new method to assess comorbidity for clinical and health services research. *Arthritis Care Res*. 2003;49(2):156–163.
13. Australian Bureau of Statistics. Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia. Canberra: Australian Bureau of Statistics; 2018.
14. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol*. 2013;13(1):61.
15. Gwet K. Inter-Rater Reliability: Dependency on Trait Prevalence and Marginal Homogeneity. *Stat Methods Inter-Rater Reliabil Assess Ser*. 2002;2:1–9.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
17. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. *JAMA*. 1999;281(17):1591–1597.
18. Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical chart review? *Med Care*. 1996;34(1):73–84.

19. Lee H-J, Choi EJ, Nahm FS, Yoon IY, Lee PB. Prevalence of unrecognized depression in patients with chronic pain without a history of psychiatric diseases. *Kor J Pain*. 2018;31(2):116–124.
20. D'Astolfo CJ, Humphreys BK. A chart review of reported musculoskeletal pain in an Ontario long term care facility. *BMC Geriatr*. 2006;6:5.
21. Lo TKT, Parkinson L, Cunich M, Byles J. Discordance between self-reported arthritis and musculoskeletal signs and symptoms in older women. *BMC Musculoskel Dis*. 2016;17:494.
22. Bedson J, Mottram S, Thomas E, Peat G. Knee pain and osteoarthritis in the general population: what influences patients to consult? *Fam Prac*. 2007;24(5):443–453.
23. Kriegsman DMW, Penninx BWJH, Van Eijk JTM, Boeke AJP, Deeg DJH. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly: A study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidemiol*. 1996;49(12):1407–1417.
24. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring Diagnoses: ICD Code Accuracy. *Health Serv Res*. 2005;40(5 Pt 2):1620–1639.
25. Romano PS, Mark, D. Bias in the coding of hospital discharge data and its implications for quality assessment. *JSTOR*. 1994;32(1):81–90.
26. Faculty of Health Sciences, editor. *The International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian modification (ICD–10-AM)*. Sydney: National Centre for Classification in Health; 2006.
27. Romano PS, Roos LL, Luft HS, Jollis JG, Doliszny K, The Ischemic Heart Disease Patient Outcomes Research T. A comparison of administrative versus clinical data: coronary artery bypass surgery as an example. *J Clin Epidemiol*. 1994;47(3):249–260.
28. Needham DM, Scales DC, Laupacis A, Pronovost PJ. A systematic review of the Charlson comorbidity index using Canadian administrative databases: a perspective on risk adjustment in critical care research. *J Crit Care*. 2005;20(1):12–19.
29. Leal JR, Laupland KB. Validity of ascertainment of co-morbid illness using administrative databases: a systematic review. *Clinical Microbiol Infect*. 2010;16(6):715–721.
30. Susser SR, McCusker J, Belzile E. Comorbidity information in older patients at an emergency visit: self-report vs. administrative data had poor agreement but similar predictive validity. *J Clin Epidemiol*. 2008;61(5):511–515.
31. Scott IA, Ward M. Public reporting of hospital outcomes based on administrative data: risks and opportunities. *Med J Aus*. 2006;184(11):571–575.

- 32.Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research Are clinical data derived from administrative databases a reliable alternative to chart review? *J Clin Epidemiol.* 2000;53(4):343–349.
- 33.Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical chart data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol.* 2004;57(10):1096–1103.
- 34.Dickerson LM, Ye X, Sack JL, Hueston WJ. Glycemic control in medical inpatients with type 2 diabetes mellitus receiving sliding scale insulin regimens versus routine diabetes medications: a multicenter randomized controlled trial. *Annals Fam Med.* 2003;1(1):29–35
- 35.Olomu AB, Corser WD, Stommel M, Xie Y, Holmes-Rovner M. Do self-report and medical chart comorbidity data predict longitudinal functional capacity and quality of life health outcomes similarly? *BMC Health Serv Res.* 2012;12:398.
- 36.Stolwijk C, van Tubergen A, Ramiro S, Essers I, Blaauw M, van der Heijde D, et al. Aspects of validity of the self-administered comorbidity questionnaire in patients with ankylosing spondylitis. *Rheumatol.* 2014;53(6):1054–1064.
- 37.Baron M, Schieir O, Hudson M, Steele R, Janelle-Montcalm A, Bernstein J, et al. Evaluation of the clinimetric properties of the Early Inflammatory Arthritis–self-administered comorbidity questionnaire. *Rheumatol.* 2009;48(4):390–394.
38. Kim AH, Roberts C, Feagan BG, Banerjee R, Bemelman W, Bodger K, et al. Developing a Standard Set of Patient-Centred Outcomes for Inflammatory Bowel Disease-an International, Cross-disciplinary Consensus.*J Crohns Colitis.* 2018;12(4):408–418

Tables

Table 1: Characteristics of study sample (N=217)

Characteristic (N=217)	n	Median* (IQR)
Age	217	66.0 (9.0)
PSA level at diagnosis	200	6.91 (4.7)
Gleason sum score at diagnosis	211	7 (1.)
NCCN		
Very Low	25	-
Low	22	-
Intermediate	120	-
High	40	-
Very High	2	-
Metastatic	3	-
IRSD		
0-1	18	-
2-3	20	-
4-5	46	-
6-7	48	-
≥8	85	-
CCI		
CCI-administrative data	203	2 (0)
CCI- medical charts	214	2 (0)
Modified CCI (mCCI)		
mCCI – administrative data	203	2 (0)
mCCI- medical charts	214	2 (0)
mCCI- self reported [mild]	122	0 (0)
mCCI- self reported [severe]	122	0 (0)
mCCI- self reported [mild, cancer assumed]	122	2 (0)
mCCI- self reported [severe, cancer assumed]	122	6 (0)
Comorbidity count		
administrative data	203	1.0 (0)
medical charts	214	2.0 (2.0)
self reported	122	1.0 (1.0)
*Medians (interquartile range –IQR) reported due to skewed distribution of variables.		

Table 2: Data table comparing medical chart vs administrative data (n=201)

Comorbidity	Medical chart data n, (%)	Administrative data n, (%)	Percentage agreement %, (95% CI)	κ statistic (95% CI)	Prevalence-adjusted κ statistic (95% CI)	Gwet's AC1 statistic (95% CI)
Myocardial infarction	6 (3.0)	0 (0)	97 (95-99)	0 (0-0)	0.94 (0.89-0.99)	0.97 (0.94-0.99)
Congestive heart failure						
Diabetes without chronic complications	21 (10.5)	16 (8.0)	94 (90-97)	0.61 (0.42-0.81)	0.87 (0.80-0.94)	0.92 (0.88-0.97)
Diabetes with chronic complications						
Mild liver disease	3 (1.5)	0 (0)	99 (97-100)	0 (0-0)	0.97 (0.94-1.00)	0.98 (0.97-1.00)
Moderate or severe liver disease						
Any malignancy*	201 (100)	201 (100)	NC	NC	NC	NC
Metastatic solid tumour						
Rheumatic disease**	4 (2.0)	0 (0)	98 (96-100)	0 (0-0)	0.96 (0.92-1.00)	0.98 (0.96-1.00)
Comorbidity	Medical chart data n, (%)	Administrative data n, (%)	Percentage agreement %, (95% CI)	κ statistic (95% CI)	Prevalence-adjusted κ statistic (95% CI)	Gwet's AC1 statistic (95% CI)
Ulcer or stomach disease	3 (1.5)	0 (0)	99 (97-100)	0 (0-0)	0.97 (0.94-1.00)	0.98 (0.97-1.00)
Renal disease	4 (2.0)	1 (0.5)	98 (95-100)	-0.01 (-0.02-0.00)	0.95 (0.91-0.99)	0.97 (0.95-1.00)
Chronic pulmonary disease	25 (12.4)	0 (0)	88 (83-92)	0 (0-0)	0.75 (0.66-0.84)	0.86 (0.80-0.92)
Hemiplegia	0 (0)	0 (0)	NC	NC	NC	NC
AIDS/HIV	0 (0)	0 (0)	NC	NC	NC	NC
Peripheral vascular disease	1 (0.5)	0 (0)	100 (99-100)	0 (0-0)	0.99 (0.97-1.00)	1.00 (0.99-1.00)
Dementia	0 (0)	0 (0)	NC	NC	NC	NC
Cerebrovascular disease	4 (2.0)	0 (0)	98 (96-100)	0 (0-0)	0.96 (0.92-1.00)	0.98 (0.96-1.00)
Total count	272	218				

NC: not calculable (prevalence was either 0% or 100%). *includes leukaemia and lymphoma **same

Table 3: Data table comparing medical chart vs SCQ data (n=121)

Comorbidity	Medical chart Present, n (%)	SCQ data Present, n (%)	Percentage agreement %, (95% CI)	K statistic (95% CI)	Prevalence-adjusted k statistic (95% CI)	Gwet's AC1 statistic (95% CI)
Heart disease	2 (1.6)	9 (7.4)	93 (88-97)	0.16 (-0.15-0.46)	0.85 (0.76-0.95)	0.92 (0.86-0.97)
Diabetes	9 (7.4)	10 (8.3)	96 (92-99)	0.71 (0.47-0.96)	0.92 (0.85-0.99)	0.95 (0.91-1.00)
Liver disease	1 (0.8)	1 (0.8)	98 (96-100)	-0.01 (-0.02-0.00)	0.97 (0.92-1.00)	0.98 (0.96-1.00)
Cancer	121 (100)	31 (25.6)	26 (18-34)	0 (0-0)	-0.49 (-0.65-0.33)	-0.40 (-0.60-0.19)
Rheumatic disease*	2 (1.7)	7 (5.8)	94 (90-98)	0.20 (-0.16-0.57)	0.88 (0.80-0.97)	0.94 (0.89-0.99)
Ulcer or stomach disease	1 (0.8)	4 (3.3)	96 (92-99)	-0.01 (-0.04-0.01)	0.92 (0.85-0.99)	0.96 (0.92-1.00)
Renal disease	3 (2.5)	1 (0.8)	97 (93-100)	-0.01 (-0.03-0.01)	0.93 (0.87-1.00)	0.97 (0.93-1.00)
Chronic pulmonary disease	12 (9.9)	5 (4.1)	91 (86-96)	0.31 (0.01-0.61)	0.82 (0.71-0.92)	0.90 (0.83-0.96)
High blood pressure	56 (46.3)	51 (42.2)	84 (78-91)	0.68 (0.55-0.81)	0.69 (0.55-0.82)	0.69 (0.56-0.82)
Anaemia or other blood disease	6 (5.0)	3 (2.5)	96 (92-99)	0.43 (0.01-0.84)	0.92 (0.85-0.99)	0.96 (0.92-1.00)
Depression	6 (5.0)	17 (14.1)	88 (82-94)	0.30 (0.04-0.55)	0.75 (0.63-0.87)	0.85 (0.77-0.93)
Back pain	4 (3.3)	38 (31.4)	72 (64-80)	0.14 (0.01-0.27)	0.44 (0.28-0.60)	0.61 (0.46-0.75)
Osteoarthritis/degenerative arthritis	14 (11.6)	26 (21.5)	75 (67-83)	0.12 (-0.08-0.31)	0.50 (0.35-0.66)	0.66 (0.53-0.79)
Total count	237	203				
*(same as connective tissue disorders)						

Table 4: Data table comparing administrative data vs SCQ data (n=112)

Comorbidity	Administrative data Present, n (%)	SCQ data Present, n (%)	Percentage agreement %, (95%CI)	K statistic (95% CI)	Prevalence-adjusted κ statistic (95% CI)	Gwet's AC1 statistic(95% CI)
Heart disease	0 (0)	9 (8.0)	92 (87-97)	0 (0-0)	0.84 (0.74-0.94)	0.91 (0.85-0.97)
Diabetes	7 (6.3)	10 (8.9)	97 (94-100)	0.81 (0.60-1.00)	0.95 (0.89-1.00)	0.97 (0.93-1.00)
Liver disease	0 (0)	1 (0.9)	99 (97-100)	0 (0-0)	0.98 (0.95-1.00)	0.99 (0.97-1.00)
Cancer	112 (100)	27 (24.1)	24 (16-32)	0 (0-0)	-0.52 (-0.68-0.36)	-0.43 (-0.64-0.23)
Rheumatic disease*	0 (0)	7 (6.3)	94 (89-98)	0 (0-0)	0.88 (0.78-0.97)	0.93 (0.88-0.99)
Ulcer or stomach disease	0 (0)	4 (3.6)	96 (93-100)	0 (0-0)	0.93 (0.86-1.00)	0.96 (0.93-1.00)
Renal disease	1 (0.9)	1 (0.9)	98 (96-100)	-0.01 (-0.02-0.00)	0.96 (0.91-1.00)	0.98 (0.96-1.00)
Chronic pulmonary disease	0 (0)	5 (4.5)	96 (92-99)	0 (0-0)	0.91 (0.83-0.99)	0.95 (0.91-1.00)
Total count	120	64				
*(same as connective tissue disorders)						

Table 5: Comparison of comorbidities recorded in administrative data vs SCQ data vs medical charts (n=112)

Comorbidity	Administrative data Present, n (%)	Medical chart Present, n (%)	SCQ data Present, n (%)	Percentage agreement % (95% CI)	K statistic (95% CI)	Prevalence-adjusted κ statistic (95% CI)	Gwet's AC1 statistic (95% CI)
Heart disease	0 (0)	2 (1.8)	9 (8.0)	94 (90-98)	0.08 (-0.07-0.23)	0.88 (0.81-0.95)	0.94 (0.90-98)
Diabetes	7 (6.3)	9 (8.0)	10 (8.9)	96 (93-99)	0.71 (0.51-0.90)	0.92 (0.86-0.98)	0.95 (0.91-0.99)
Liver disease	0 (0)	1 (0.9)	1 (0.9)	99 (97-100)	0 (-0.01-0.00)	0.98 (0.94-1.00)	0.99 (0.97-1.00)
Cancer	112 (100)	112 (100)	27 (24.1)	49 (44-55)	0 (0-0)	-0.01 (-0.12-0.10)	0.19 (0.07-0.31)
Rheumatic disease*	0 (0)	2 (1.8)	7 (6.3)	95 (92-98)	0.10 (-0.08-0.28)	0.90 (0.84-0.97)	0.95 (0.91-0.99)
Ulcer or stomach disease	0 (0)	1 (0.9)	4 (3.6)	97 (94-100)	-0.01 (-0.02-0.00)	0.94 (0.89-0.99)	0.97 (0.94-1.00)
Renal disease	1 (0.9)	3 (2.7)	1 (0.9)	97 (94-100)	-0.01 (-0.02-0.00)	0.94 (0.89-0.99)	0.97 (0.94-1.00)
Chronic pulmonary disease	0 (0)	12 (10.7)	5 (4.5)	92 (88-96)	0.15 (0.01-0.29)	0.83 (0.75-0.92)	0.91 (0.86-0.96)
Total count	120	142	64				
*(same as connective tissue disorders)							

Figures

Figure 1: Recruitment frame

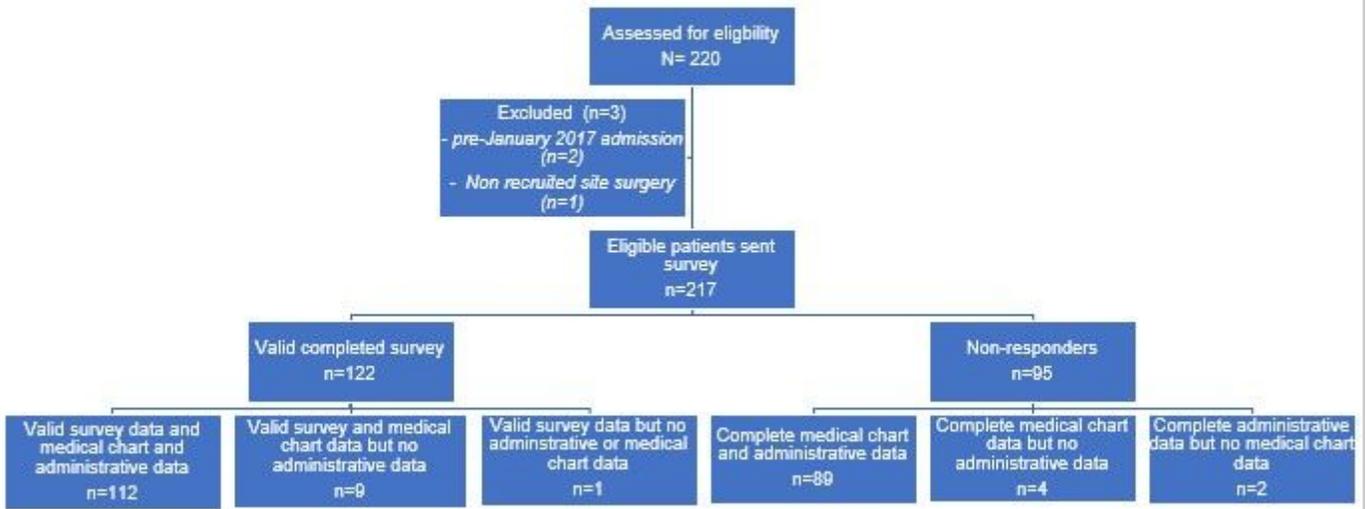


Figure 1

Recruitment frame

Figure 2: Modified ~~Charlson~~ distributions

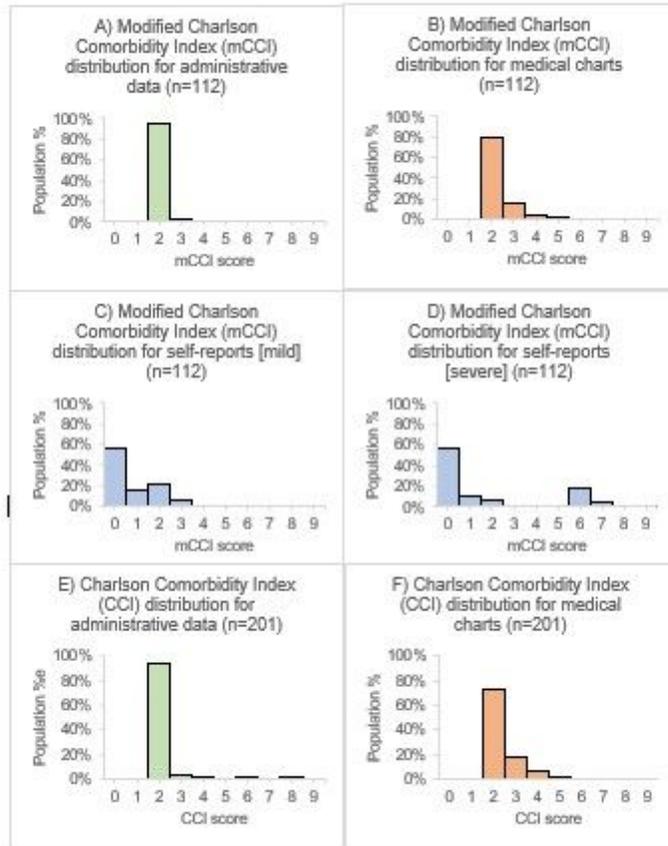


Figure 2

Modified Charlson distributions

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.docx](#)