Comparison of National Cancer Registration Dataset and Rapid Cancer Registration Dataset

NPCA: Methodological Update 2023

Date of release: 07.09.2023

Cohort for data comparison: Patients diagnosed with prostate cancer between 1st January 2018 and 31st March 2021

Contents

Executive summary

In this NPCA methodological update we compare the National Cancer Registration Dataset (NCRD) and the Rapid Cancer Registration Dataset (RCRD) for four NPCA performance indicators. The RCRD raises the potential for more timely and more frequent reporting to providers. Alongside this, data quality also needs to be considered. Data from the NCRD and the RCRD for patients diagnosed between 1st January 2018 and 31st March 2021 were used. The RCRD and NCRD are English datasets and as such this methodological update isn't applicable to Welsh patients.

There was very good agreement between the NCRD and the RCRD concerning the proportion of patients who had an emergency readmission within 90 days of radical prostate cancer surgery. Good agreement was also found for the proportion of patients experiencing a genitourinary (GU) complication requiring a procedural/surgical intervention within 2 years of radical prostatectomy, and the proportion of patients receiving a procedure of the large bowel and a diagnosis indicating radiation toxicity (gastrointestinal (GI) complication) up to 2 years following radical prostate radiotherapy. However, there was relatively poor agreement for the proportion of men diagnosed with metastatic disease.

This report shows that there is scope for more timely reporting of prostate cancer services using the RCRD, and that informative performance indicators can be obtained. However, data quality issues, including poor completeness of disease staging data, limit the reliability of indicators and prevent an outlier process being undertaken. Clinical teams should be encouraged to improve their timely submission of staging data through the Cancer Outcomes and Services Dataset (COSD). Further, the absence of histology data (particularly Gleason score) from the RCRD means risk classifying is not possible and indicators of undertreatment and overtreatment cannot be produced. Gleason score is more likely to be included in the RCRD if all pathology reports are completed by pathology teams in both machine readable and free text formats.

Key messages

- Using data for patients diagnosed between 1st January 2018 and 31st March 2021, there was good agreement between the NCRD and the RCRD for three out of the four NPCA performance indicators investigated, showing that informative indicators can be obtained from the RCRD. Of particular note:
    - There was very good agreement for the proportion of patients who had an emergency readmission within 90 days of radical prostate cancer surgery.
    - There was relatively poor agreement for the proportion of men diagnosed with metastatic disease.
- The RCRD is available much sooner than the NCRD, enabling analysis of incidence and stage at diagnosis much closer to real time than was previously possible.

- The incompleteness of key data on stage (specifically the metastatic status within the TNM stage 4 category) and lack of Gleason grade precludes the ability to consider important details relating to treatment, toxicities relating to treatment and the ultimate outcome of those treatments in different patient groups. This shortfall limits the current value of the RCRD.
- Clinical teams should be encouraged to improve their timely submission of staging data (TNM) through the Cancer Outcomes and Services Dataset (COSD) and pathology teams should be encouraged to provide their reports in both machine readable and free text formats.

Introduction

The National Prostate Cancer Audit (NPCA) presents performance indicators on NHS providers in each annual report, accompanied by an outlier process to investigate performance which is significantly below the expected range. For England, indicators are calculated from data provided by National Cancer Registration and Analysis Service (NCRAS), however disruption during the COVID-19 pandemic meant that the standard cancer registration dataset, National Cancer Registration Dataset (NCRD) [1], was unavailable in time for the 2021 and 2022 NPCA annual reports which were published in January 2022 and 2023 respectively. Data were made available instead from the Rapid Cancer Registration Dataset (RCRD) system, which are released in much shorter time frames than the NCRD but undergo less stringent quality control. The speed of production of the RCRD and the pandemic has meant that several of the NCRD data items are unavailable or too incomplete for use and others should be interpreted with caution. In particular, the RCRD doesn't currently include Gleason score (cancer grade) which is used for risk stratification and is essential to calculating some of the indicators the NPCA reports. It is also integral to the audit's ability to identify outlier groups for a number of key parameters relating to "overtreatment", "undertreatment" and treatment related toxicity.

When preparing the 2021 NPCA annual report, the RCRD was available for patients diagnosed between 1 January 2018 and 31 December 2020. The NCRD was available for patients diagnosed up until 31 March 2019. It was therefore possible to compare the RCRD and the NCRD for the 15-month period from 1 January 2018 to 31 March 2019. This indicated that approximately 90% of diagnoses were captured by the RCRD system, and importantly this figure was consistent across providers. It was therefore decided to publish performance indicators using the RCRD for men diagnosed in England in the NPCA 2021 and 2022 annual reports, but with some uncertainty about data quality the formal outlier process was not undertaken.

NCRD vs RCRD

There is widespread interest in understanding how comparable the RCRD is to the NCRD as the RCRD raises the potential for more timely (within 6 months) and more frequent (at least quarterly) reporting to providers. This agile reporting could be more informative in some aspects for NHS providers and other stakeholders.

In March 2023, the NCRD became available for patients in England diagnosed to 31 March 2021. This made it possible to reproduce the 2021 and 2022 annual report indicators with the NCRD. This report compares the published indicators that were calculated for England using the RCRD, and the same indicators produced from the NCRD.

The aim of the report is to investigate the reliability of key NPCA performance indicators calculated using the RCRD by comparing with the same indicators produced from the NCRD.

Methods

Indicators and data sources

Both the RCRD and the NCRD were available for patients diagnosed between 1st January 2018 to 31st March 2021. The four NPCA annual report performance indicators that were compared were:

- Indicator 1: Proportion of men who had metastatic disease when they were first diagnosed (specialist multidisciplinary team (SMDT) level)
- Indicator 2: Proportion of patients who had an emergency readmission within 90 days of radical prostate cancer surgery (surgery centre level)
- Indicator 3: Proportion of patients experiencing at least one genitourinary (GU) complication requiring a procedural/surgical intervention within 2 years of radical prostatectomy (surgery centre level)
- Indicator 4: Proportion of patients receiving a procedure of the large bowel and a diagnosis indicating radiation toxicity (gastrointestinal (GI) complication) up to 2 years following radical prostate radiotherapy (radiotherapy centre level)

The NPCA has previously reported on the undertreatment and overtreatment of prostate cancer patients based on tumour stage and grade. However, these indicators can't be calculated using the RCRD as it is missing Gleason score, a key factor in risk classification.

Indicator 1 included patients diagnosed with prostate cancer between 1 April 2019 and 31 March 2021. NCRD Cancer Registry data and the equivalent RCRD data was used. For both data sources, the prevalence of metastatic disease at diagnosis was calculated for each SMDT by dividing the number of metastatic disease patients by the number of prostate cancer patients. Metastatic status was missing for ~40% of patients in the RCRD Cancer Registry data but for some of these patients overall stage was completed. Where metastatic status was missing but overall stage was assigned, metastatic stage was imputed, with stage 1, 2, and 3 cancer classified as non-metastatic and stage 4 classified as metastatic. Please note that for prostate cancer 'stage 4' is not synonymous with 'metastatic cancer' as 'stage 4' can include cancer that has spread to nearby lymph nodes but hasn't spread elsewhere so it is not metastatic disease [2]. This limitation was considered but it was still deemed beneficial to conduct the imputation. In the NCRD, 14,761 patients missing metastatic status were excluded from this analysis, and in the RCRD 19,709 patients with missing metastatic status after imputation were also excluded.

Indicator 2 included prostate cancer patients who had a radical prostate cancer surgery between 1 April 2019 and 31 March 2021. NCRD Hospital Episode Statistics Admitted Patient Care (HES APC) data and the equivalent RCRD data was used to identify patients who had radical prostate cancer surgery and emergency readmissions.

Indicator 3 included prostate cancer patients who had radical local treatment by prostatectomy between 1 January 2018 and 31 December 2019. Indicator 3 used NCRD HES

APC data and the equivalent RCRD data. Patients with any record of bladder cancer were excluded. National Radiotherapy Dataset (RTDS) data was used to identify patients who had received radiotherapy and exclude them from the cohort.

Indicator 4 included prostate cancer patients who had radical local treatment by radiotherapy between 1 January 2018 and 31 December 2019. Indicator 4 used the RTDS to exclude patients who had any record of brachytherapy treatment. Only patients who had radiotherapy to the prostate for non-metastatic curative treatment were included. HES APC data was used to identify GI complications. Patients with any record of bladder cancer were excluded. Patients who had a prostatectomy were excluded.

Indicator 1 was analysed at the SMDT level whereas indicators 2, 3 and 4 were analysed at the provider trust level. Only NHS SMDTs and provider trusts were included in this analysis. Indicators 2, 3 and 4 were adjusted for the age distribution, co-morbidity burden (Charlson score), cancer risk group distribution and deprivation (IMD deprivation score) of the patients in each provider trust. Cancer risk group and deprivation were available from Cancer Registry data.

All of the data was provided by the National Disease Registration Service (NDRS). This report uses data that has been provided by patients and collected by the NHS as part of their care and support. The data are collated, maintained and quality assured by the National Disease Registration Service, which is part of NHS England [3].

Analysis

To assess agreement visually between the NCRD and the RCRD, scatter plots were plotted for each indicator. For indicators with poor agreement, further investigation was carried out to ascertain whether this was associated with SMDT/provider size by producing scatter plots where the difference found between the NCRD and the RCRD was plotted against the number of prostate cancer patients in each SMDT/provider. To further investigate indicators with relatively poor or good correlation, agreement in the numerator and the denominator was plotted separately.

Results

From the NCRD, a total of 78,793 men were diagnosed with prostate cancer in England in the two years from 1 April 2019 to 31 March 2021. Of these men, 91% (n=71,954) were also included in the RCRD, with the remaining 6,939 in the NCRD only. The completeness of the RCRD, by comparison with the NCRD, varied between SMDTs from 84% to 98%. There were also 2,134 diagnoses recorded in the RCRD and not in the NCRD, evenly distributed across SMDTs.

Figure 1 shows four plots, one for each indicator, with the result from the RCRD on the vertical axis and the result from the NCRD on the horizontal axis. Each plotted point

represents a SMDT in Figure 1a and a provider trust in Figures 1b-1d. Providers or SMDTs, where similar results were obtained from the RCRD and the NCRD, appear close to the red line. The level of agreement varied across indicators, with very good agreement between the datasets for indicator 2 (Emergency readmissions, Figure 1b) but relatively poor agreement for indicator 1 (Metastatic at diagnosis, Figure 1a). For indicator 3 (GU complications, Figure 1c) and indicator 4 (GI complications, Figure 1d) agreement was not as good as for indicator 2. For indicator 1 (Figure 1a) there is also evidence of bias since most points are below the line of agreement, indicating higher proportions of metastatic patients found when using the NCRD.

To visually investigate whether the poor agreement found for indicator 1 (Metastatic at diagnosis, Figure 1a) is related to the size of the patient populations the SMDTs serve, Figure 2 plots the difference found between the NCRD and the RCRD by the SMDT size. Figure 2 shows no clear association between agreement and SMDT size.

To further investigate the relatively poor agreement of indicator 1 (Metastatic at diagnosis), the numerator and denominator were analysed separately (Figure 3). Comparing the denominator and numerator of indicator 1 (Figure 3A and 3B) agreement appears to be worse in the numerator. This indicates that “missingness” in the RCRD is more common among metastatic patients than in those who are non-metastatic since the numerator represents metastatic patients only.

For indicator 2 (Emergency readmission), agreement was also plotted separately for the numerator and denominator. In this case the overall agreement was better than that seen in either the numerator or denominator separately, with the dots gathered predominantly under the lines for both in Figure 4.

Discussion

This report compares the National Cancer Registration Dataset (NCRD) with the Rapid Cancer Registration Dataset (RCRD) across four NPCA performance indicators. The reliability of results using the RCRD as a proxy for the NCRD, is variable. The best agreement was observed in indicator 2 (emergency readmission) and the poorest in indicator 1 (metastatic at diagnosis). For indicator 1, as well as poor agreement there was evidence of bias with higher proportions of metastatic patients found by using the NCRD. These findings will help with interpretation of future NPCA outputs that use the RCRD.

Missing data may contribute to the poor agreement of indicator 1, with metastatic status missing for ~40% of patients in the RCRD compared with ~20% in the NCRD. “Missingness” was reduced by imputing some missing M stage from overall stage in the RCRD but it is unclear how reliable this imputation is; it is likely that some stage 4 non-metastatic patients were incorrectly classified as metastatic. This imputation also underestimates the difference between the datasets. Improvement in the completeness of staging data, particularly metastatic status, would be valuable for both the RCRD and the NCRD. This could be

achieved by clinical teams improving timely submission of staging data through the Cancer Outcomes and Services Dataset (COSD).

To investigate whether excluding those with missing metastatic status contributed to higher proportions of metastatic patients found when using the NCRD, we added back in the patients with missing metastatic status, however, this made little difference.

For indicator 1, agreement was worse in the numerator, which is the number of metastatic patients. This suggests that either a metastatic patient is more likely to be absent from the RCRD than one who is non-metastatic, or that a metastatic patient is more likely to be missing metastatic status, resulting in the bias towards lower rates of metastatic patients. Why this happens is unclear but it may arise from more complex SMDT processes for men with primary presenting metastatic disease.

Very good agreement was found in emergency readmission rates following radical prostate cancer surgery. The agreement of indicator 3 (GU complications) and indicator 4 (GI complications) was not as good but this may result from smaller event numbers of specific complications compared with all emergency readmissions, as in indicator 2.

For indicator 2 (Emergency readmission), overall agreement was better than that seen in either the numerator or denominator separately. This seems likely to result from the indicator using HES data for both its numerator and denominator; if a patient is missing from the RCRD their HES data will therefore be missing from both the numerator and denominator and the error will cancel out. This may also explain the relatively good agreement of indicators 2, 3 and 4 compared with indicator 1, as data from associated datasets provides both the numerators and denominators for indicators 2, 3 and 4. Associated datasets (e.g. HES APC, RTDS) have a better completeness than the main dataset (Cancer Registry dataset) as patients who undergo intensive cancer treatments are better captured in data than patients who don't undergo these treatments.

In this report we compare the RCRD and the NCRD, taking the NCRD to be the gold standard; this was deemed appropriate as the NCRD is more rigorously checked than the RCRD. Comparative analysis by NCRAS found that some patients in the RCRD do not appear in the NCRD [4]. This could arise for several reasons, e.g., a tumour initially thought to be prostate cancer may be found on further investigation to originate elsewhere. While the NCRD is the gold standard for this comparison it is worth noting that data completeness remains an issue, with metastatic status missing for ~20% of patients.

The RCRD has shown potential to permit more timely and frequent reporting by the NPCA. However, a key limitation is that it doesn't contain the patient's Gleason score. If Gleason score could be included in the RCRD, this would allow the NPCA to report on the undertreatment and overtreatment of prostate cancer patients, a critical element of the audit's outlier process. This is of particular importance to NPCA stakeholders and patients. Gleason score is more likely to be included in the RCRD if all pathology reports are completed by pathology teams in both machine readable and free text formats.

Conclusion

The reliability of performance indicators calculated using the RCRD as a proxy for the NCRD depends on the performance indicator being considered. Very good agreement was found in emergency readmission following radical prostate cancer surgery but agreement was not as good in the proportion of patients found to be metastatic at diagnosis. Indicators which used the same source of data for both numerator and denominator appeared to be more reliable. There is scope for more timely reporting around prostate cancer services using the RCRD and these results show that informative performance indicators can be obtained. However, the ability to carry out an assessment of "outliers" through established NPCA processes using the RCRD would be susceptible to data quality issues, while the lack of Gleason grading limits the range of indicators that can be used.

Patient summary

In its annual reports, the National Prostate Cancer Audit (NPCA) produces indicators that report on the performance of NHS providers (e.g., specialist multidisciplinary teams, surgery centres, radiotherapy centres). For the English data in the annual reports, the National Cancer Registration and Analysis Service (NCRAS) usually provides its standard, routinely collected dataset, the National Cancer Registration Dataset (NCRD). However, due to COVID-19 pandemic-related disruption, NCRAS instead provided the Rapid Cancer Registration Dataset (RCRD) for the 2021 and 2022 annual reports. The RCRD is available more quickly but there are concerns about its data quality as it undergoes less stringent quality control. There is widespread interest in understanding how comparable the RCRD is to the NCRD as the RCRD raises the potential for more timely and more frequent reporting.

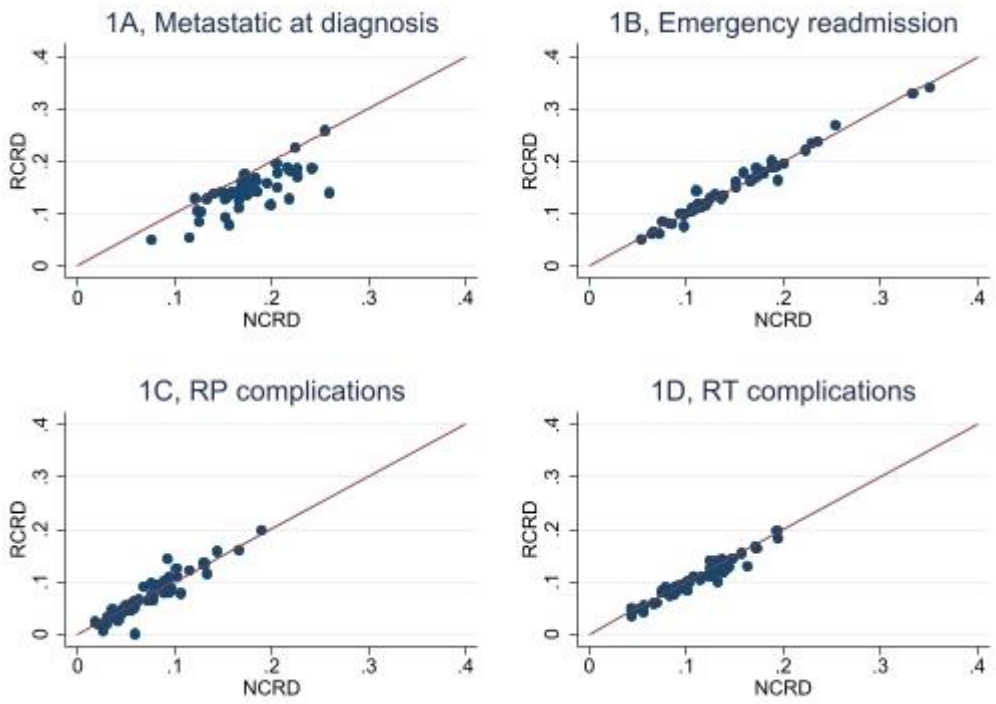
The NCRD and the RCRD were compared for the period 1st January 2018 to 31st March 2021. Overall, there was good agreement between the datasets for three out of the four NPCA performance indicators investigated. There was very good agreement in the proportion of patients who had an emergency readmission within 90 days of radical prostate cancer surgery. Additionally good agreement was found for the proportion of patients experiencing a genitourinary (GU) complication requiring a procedural/surgical intervention within 2 years of radical prostatectomy. It was also good for the proportion of patients undergoing a bowel-related procedure (such as a colonoscopy), which might indicate radiation-related toxicity on the lower bowel, particularly the rectum (gastrointestinal (GI) complication). This occurs more commonly in the 2-year period following radical prostate radiotherapy. When considering the accuracy of men presenting with metastatic disease at first diagnosis of their prostate cancer, there was relatively poor agreement between the rapid and standard reporting methods. Another consideration is that the RCRD doesn't include information on the grade of the cancer (how much the cancer cells look like normal cells) meaning risk

classifying is not possible and useful indicators about the undertreatment and overtreatment of patients cannot be produced.

Within this report, we demonstrated potential for more timely reporting around prostate cancer services with the RCRD, and these results show that informative performance indicators can be obtained, although there are data quality concerns for some of the performance indicators.

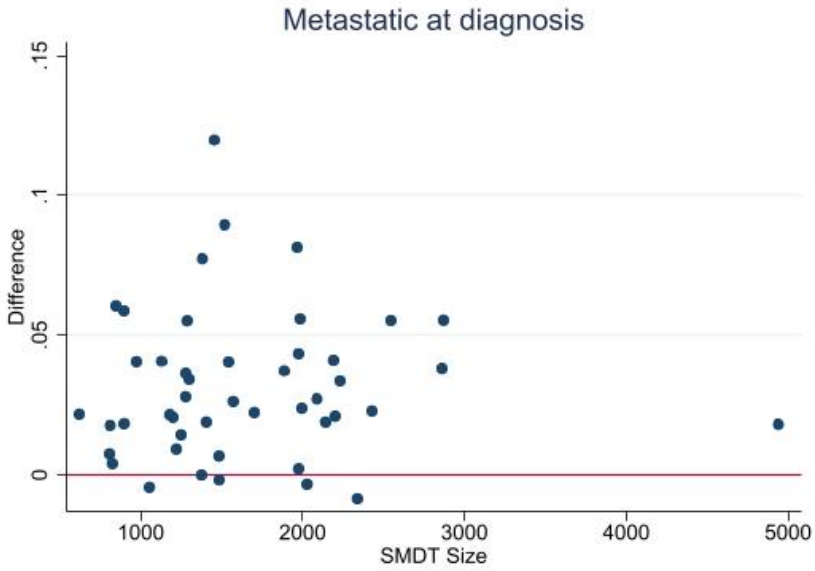
Tables and Figures

Figure 1. NPCA performance indicators using the NCRD and the RCRD



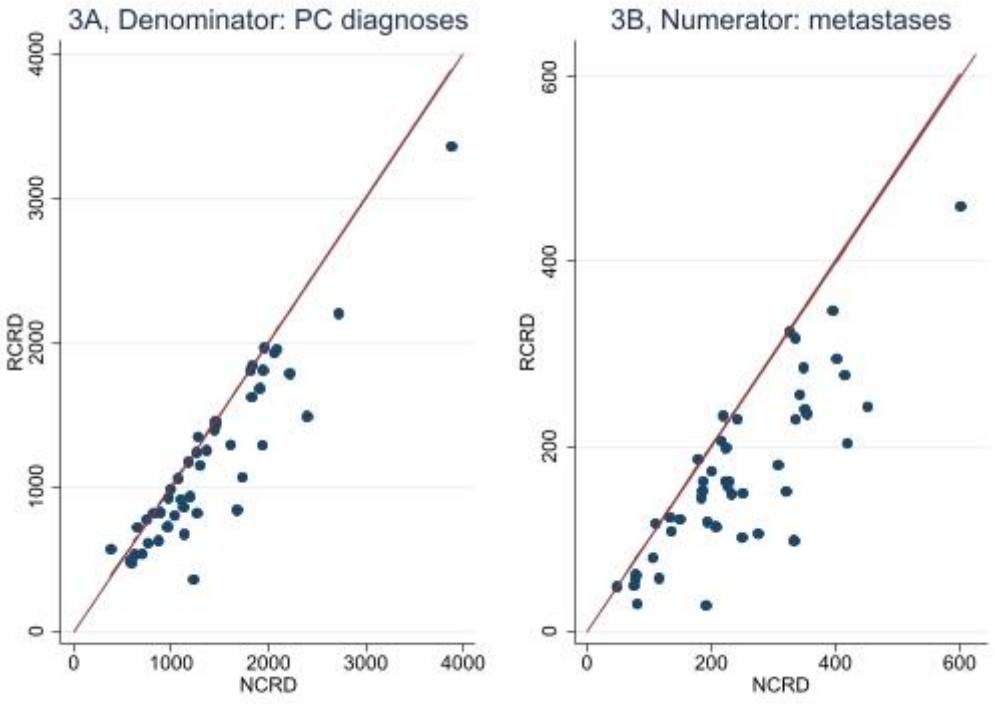
Footnote: Each plotted point represents a specialist multidisciplinary team (SMDT) in Figure 1a, and a provider trust in Figures 1b-1d. Providers, or SMDTs, where similar results were obtained from the Rapid Cancer Registration Dataset (RCRD) and the National Cancer Registration Dataset (NCRD) appear close to the red line which shows 1:1 equivalence. NPCA, National Prostate Cancer Audit; RP, radical prostatectomy; RT, radical prostate radiotherapy.

Figure 2. Metastatic at diagnosis performance indicator: difference between using the NCRD and the RCRD plotted against SMDT size



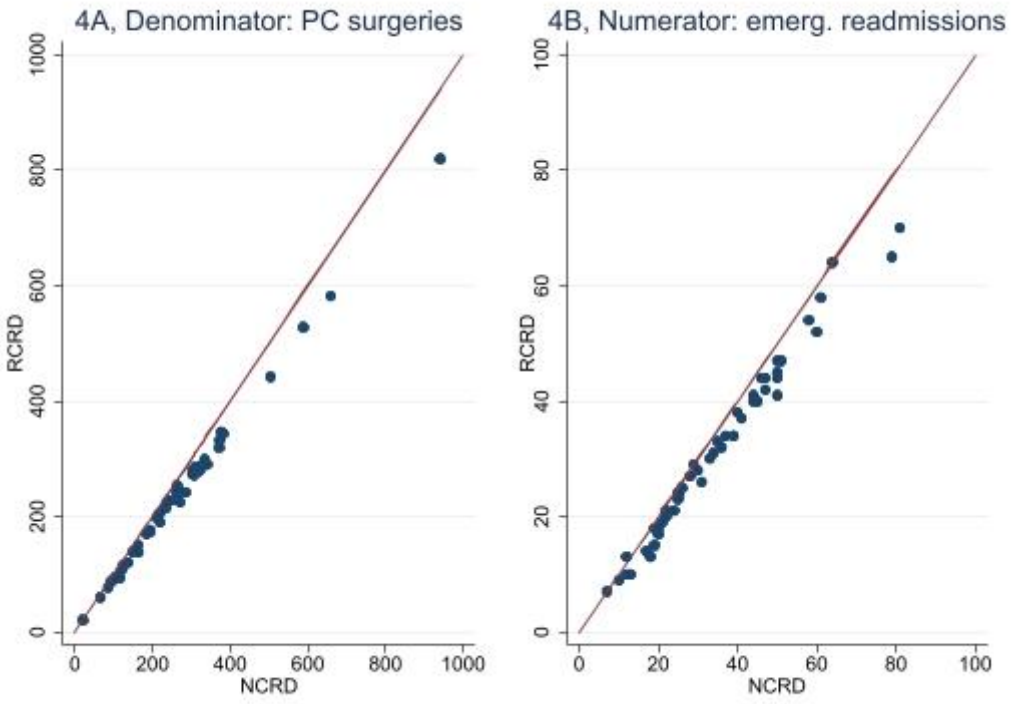
Footnote: Each plotted point represents a specialist multidisciplinary team (SMDT). NCRD, National Cancer Registration Dataset; RCRD, Rapid Cancer Registration Dataset.

Figure 3. Agreement in denominator of indicator 1 (number of prostate cancer diagnoses) and numerator of indicator 1 (number with metastases)



Footnote: Each plotted point represents a specialist multidisciplinary team (SMDT). SMDTs where similar results were obtained from the Rapid Cancer Registration Dataset (RCRD) and the National Cancer Registration Dataset (NCRD) appear close to the red line which shows 1:1 equivalence. Please note that the scales of these graphs differ. PC, prostate cancer.

Figure 4. Agreement in denominator of indicator 2 (number of radical prostate cancer surgeries) and numerator of indicator 2 (number with emergency readmission <90 days)



Footnote: Each plotted point represents a provider trust. Provider trusts where similar results were obtained from the Rapid Cancer Registration Dataset (RCRD) and the National Cancer Registration Dataset (NCRD) appear close to the red line which shows 1:1 equivalence. Please note that the scales of these graphs differ. PC, prostate cancer; emerg, emergency.

Acknowledgements

Data for this report is based on patient-level information collected by the NHS, as part of the care and support of cancer patients. The data is collated, maintained and quality assured by the National Disease Registration Service, which is part of NHS England.

Glossary

British Association of Urological Surgeons (BAUS)

Professional association for urological surgeons. Registered charity no: 1127044.

British Uro-oncology Group (BUG)

Professional association for clinical and medical oncologists specialising in the field of urology. Registered charity no: 1116828.

Charlson Co-morbidity Score

A scoring system used commonly to quantify the co-existence of other medical conditions (medical co-morbidities: see below). Many patients may have other medical conditions in addition to their prostate cancer. The score is calculated based on the absence and presence of specific medical problems in the Hospital Episode Statistics (HES) database.

Co-morbidity

Medical condition(s) or disease process(es) that are additional to the disease under investigation (in this case, prostate cancer).

Gleason Score

The Gleason score is a measure assigned by a pathologist to determine how aggressive an individual's prostate cancer is when the prostate cancer tissue is examined using a microscope. It is made up of two separate scores between 3 and 5 which are then added together to make a final score graded between 6 and ten. Along with PSA and TNM, the Gleason score can be used to predict how a prostate cancer might behave in the future. This process is used for risk stratification, i.e., to help to predict how a specific cancer might progress and/or respond to treatment.

Hospital Episode Statistics (HES)

A database that contains data on all patients treated within NHS trusts in England. This includes details of admissions, diagnoses and treatments.

Index of Multiple Deprivation (IMD)

The English Indices of Deprivation (ID) are a useful tool for targeting services to help tackle deprivation. They provide a means of identifying the most and least deprived areas in England and to compare whether one area is more deprived than another.

Metastatic Disease

When cancer has spread from its initial site of development in the prostate (the primary site) to distant sites of the body (the metastatic site(s)). These sites are mainly in the bones and lymph nodes in the first instance.

National Cancer Registration and Analysis Service (NCRAS)

A national body which collects, analyses and reports on cancer data for the NHS population in England.

Radical Prostatectomy

The surgical removal of all the prostate gland and the associated seminal vesicles. The latter are structures integrally associated with the prostate. Their function is to produce and store fluid which sustains the viability of sperm when it leaves the prostate.

Radiotherapy

The use of radiation to destroy cancer cells. There are different types of radiotherapy, including external beam radiotherapy (radiotherapy delivered from a radiation source outside the body) and brachytherapy (radiotherapy delivered directly by implanting a radiation source within the tumour itself).

Royal College of Surgeons of England (RCS)

An independent professional body committed to enabling surgeons to achieve and maintain the highest standards of surgical practice and patient care. As part of this it supports audit and the evaluation of clinical effectiveness of surgery.

Specialist Multidisciplinary Team (SMDT)

A team of specialists who coordinates the specialist treatment of men with prostate cancer. The SMDT enables local cancer units to access specialist prostate cancer services which may not be locally available. Specialist services include prostatectomy and radiotherapy.

Staging/stage

The anatomical extent of a cancer. This determines whether a cancer is confined within its primary site (localised disease) or whether it has spread to other areas of the body (metastatic spread). It is usually denoted by the TNM staging process where “T” represents the local stage, “N” the presence of lymph node involvement and “M” represents the presence of metastatic disease. T1 means the cancer is too small to be seen on a scan, T2 means the cancer is completely inside the prostate gland, T3 means the cancer has broken through the capsule (covering) of the prostate gland and T4 means the cancer has spread into other body organs nearby, such as the back passage, bladder, or the pelvic wall. N0 means that the nearby lymph nodes do not contain cancer cells and N1 means there are cancer cells in lymph nodes near the prostate. M0 means the cancer has not spread to other parts of the body and M1 means the cancer has spread to other parts of the body outside the pelvis.

References

[1] Henson, K. E., Elliss-Brookes, L., Coupland, V. H., Payne, E., Vernon, S., Rous, B., & Rashbass, J. (2020). Data Resource Profile: National Cancer Registration Dataset in England. *International Journal of Epidemiology*, 49(1), 16-16H. https://doi.org/10.1093/ije/dyz076

[2] Prostate cancer stages: Staging of Prostate cancer. Staging of Prostate Cancer | American Cancer Society. Available from: https://www.cancer.org/cancer/types/prostate-cancer/detection-diagnosis-staging/staging.html

[3] National Disease Registration Service. Available from: https://digital.nhs.uk/ndrs

[4] National Cancer Registration and Analysis Service. *Rapid Cancer Registration Dataset: data at 4th March (CAS2303).* Available from: http://www.ncin.org.uk/collecting_and_using_data/rcrd