Journal of Clinical Epidemiology 182 (2025) 111767

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Emulating an existing trial of treatments for prostate cancer using real-world data: challenges and lessons learned

Caroline Chesang[a,*], Linda D. Sharples[a], Christen M. Gray[a,b], Julie Nossiter[c,d], Jan van der Meulen[c,d], Thomas E. Cowling[c,d], Ruth H. Keogh[a]

[a]*Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT, London, UK*
[b]*AstraZeneca, BioPharma Medical Evidence, Real World Data Science, 1 Francis Crick Avenue, Cambridge, CB2 0AA, UK*
[c]*Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK*
[d]*Clinical Effectiveness Unit, Royal College of Surgeons of England, London, WC2A 3PE, UK*

Accepted 19 March 2025; Published online 25 March 2025

Abstract

Objectives: If randomized controlled trials can be successfully emulated using real-world data (RWD), confidence in the validity of RWD for estimating treatment effects for questions that have not been assessed in trials increases. We used routinely collected administrative and clinical national linked datasets from England to emulate the PR07 trial for high-risk prostate cancer patients, which compared the effects of radiotherapy added to hormone therapy (RT+HT) within 8 weeks of randomization (the ''grace period'') and hormone therapy (HT) only on all-cause mortality. We highlight methodological choices required and challenges encountered in emulating this trial.

Study Design and Setting: Patients diagnosed with prostate cancer from 2014 to 2020 were identified from the routine national linked datasets. Diagnosis was taken as the time zero. As few patients initiated radiotherapy within 8 weeks of diagnosis, we considered target trials with grace periods of 4-6 months. Estimands of interest were hazard ratios (HRs) and survival probabilities over 7 years. The cloning-censoring-and-weighting (CCW) approach was used to control for measured confounding and to allow for the grace period. We also used an extension (the ''landmark-CCW'' approach), in which we consider several time-origins post-diagnosis, enabling us to use a grace period of 8 weeks as in PR07.

Results: A total of 2,690 patients were eligible for inclusion in the emulated trial. The CCW analysis using a grace period of 6 months gave an estimated HR of 0.48 (95% confidence interval [CI]: 0.34−0.60) and 7-year survival estimates of 80.7% (95% CI: 74.3−87.0) for the RT+HT strategy and 65.6% (95% CI: 62.8−68.1) for HT only strategy, and corresponding risk difference of 15.1% (95% CI: 11.5−18.9). The corresponding HR from the landmark-CCW approach was 0.58 (95% CI: 0.51−0.65) and with survival estimates of 80.7% (95% CI: 77.7−83.8) for RT+HT strategy and 69.8% (95% CI: 68.2−71.4) for the HT only strategy, and a risk difference of 10.9% (95% CI: 6.3−15.9).

Conclusion: Our findings from the emulated trial using RWD are broadly consistent with those from PR07, with RT+HT estimated to result in better survival compared to HT only. However, the findings were not replicated exactly, with PR07 reporting an HR of 0.77 (95% CI: 0.61−0.98) over 7 years of follow-up. Differences may be in part due to challenges in defining time zero and allowing for a treatment grace period of the same duration as in PR07. Our study considered ways in which these challenges can be addressed, and our findings affirm the utility of RWD for estimating treatment effects.

Keywords: Trial emulation; Time-varying treatment; All-cause mortality; Prostate cancer; Grace period; Clone-censor-weight; Landmarking

Funding: Caroline Chesang was supported by a PhD studentship co-funded by the Economic and Social Research Council and AstraZeneca (ES/P000592/1). Ruth H. Keogh was funded by UK Research and Innovation (Future Leaders Fellowship MR/S017968/1, MR/X015017/1). This study was undertaken alongside the National Prostate Cancer Audit. The audit is commissioned by the Healthcare Quality Improvement Partnership as part of the National Clinical Audit and Patient Outcomes Programme, funded by NHS England and the Welsh Government (https://www.hqip.org.uk/national-programmes). Neither HQIP nor NHS England or the Welsh Government had any involvement in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

* Corresponding author. Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.

E-mail address: Caroline.Chesang1@student.lshtm.ac.uk (C. Chesang).

https://doi.org/10.1016/j.jclinepi.2025.111767

Plain Language Summary

Clinical trials are the best way to test whether treatments work, but they are expensive, take years to complete, and focus on narrow research questions. If we can use real-world data (RWD) such as patient health records to mimic these trials, we may be able to answer additional medical questions relevant to patients who are prescribed these medications. This study aimed to see if we could recreate the results of a past clinical trial (PR07) using national health data from England. The PR07 trial looked at two treatments for high-risk prostate cancer: hormone therapy (HT) alone and radiotherapy added to hormone therapy (RT+HT) within 8 weeks of starting the study. This trial was chosen for several reasons. It included high-risk patients who were studied for up to 7 years, meaning that there was enough information to study survival outcomes. Being a major UK-based trial, it was useful for comparing with the data recorded in UK national health data. The trial also allowed some flexibility in when treatment started, which was an interesting factor to examine. Its main goal was to see if adding radiotherapy to hormone therapy provided extra benefits to patients. We wanted to see whether the trial results were replicated using the UK health data. If results were replicated, it would provide confidence in further exploration of questions not covered by the trial. In a trial, randomization time is the point where doctors allocate patients to one of the treatments being assessed, usually 1 new treatment and 1 control treatment. Patients in the study are then followed up from that point. However, defining a starting point in a non-trial setting, using UK national health data, is not as straightforward. In the PR07 trial, patients started RT+HT within 8 weeks of randomization. In the UK datasets, we use diagnosis as a starting point, because it is available for both treatment groups. However, very few patients started RT+HT within 8 weeks of diagnosis. To address this, we allowed for longer periods of 4−6 months from diagnosis for RT+HT initiation. Recent developments provided statistical methods for estimating the cause-effect relationship between treatment received and survival patterns. In this study, we show how these approaches can be used to estimate survival patterns if all patients had RT+HT within 4−6 months from diagnosis compared with if no patients had any RT within 4−6 months. We account for the different treatment initiation times and the main differences between the RT+HT and HT only patients. Using these methods, we estimate that RT+HT has an 11%−15% higher survival rate at 7 years compared to HT only. We discuss similarities and differences between these findings and those in the original PR07 trial.

1. Introduction

Prostate cancer (PCa) is the second most common cancer in men and the third most common cancer overall, after breast and lung cancers [1]. PCa follows a variable course in different patients as a result of its highly heterogeneous nature. Treatment depends on risk stratification and preferences of clinicians and patients [2,3]. Although trials have been conducted in PCa, there remain unanswered questions about treatments, not all of which can be answered using randomized controlled trials (RCTs), so that we may consider using observational data for this task.

Target trial emulation provides a framework for using real-world data (RWD) to estimate causal treatment effects, drawing on the design principles of RCTs [4]. The idea is to explicitly specify a protocol for the RCT that one would like to perform if it were feasible (the ''target trial''). We apply this framework, with causal inference analysis methods, to emulate an existing, published RCT. Benchmarking emulated trials against existing trials could provide evidence of the ability of RWD to accurately estimate treatment contrasts, hence providing greater confidence in using RWD for emulation of new target trials which have not been conducted. Examples of benchmarking have been conducted in different disease areas [5−10].

In this study, we aimed to use routine English national linked data to emulate an existing trial (PR07) of treatment for men with high-risk PCa. The PR07 trial found that early addition of radiotherapy to hormone therapy (RT+HT) increased overall survival compared to hormone therapy (HT) alone [11−13].

This article discusses the challenges of emulating an existing trial and how we addressed them. Key challenges in emulating the PR07 trial were definition of ''time zero'' (corresponding to the time of randomization) and the fact that radiotherapy (RT) was initiated within a grace period of 8 weeks.

We use the cloning-censoring-and-weighting (CCW) method to address confounding and to allow for a treatment strategy that has a grace period [14]. We extend this by combining the CCW approach with landmarking, to consider different time zeros in the emulated trial [15−17]. We obtain estimates of hazard ratios (HRs), the main measure of treatment effect in the PR07 trial. We also obtain survival probabilities under the 2 treatment strategies, which do not suffer from the inbuilt selection bias that affects HRs [18−20].

2. Materials and methods

The STrengthening the Reporting of OBservational studies in Epidemiology checklist guided reporting.

2.1. Data sources

We used routine national linked datasets for PCa patients diagnosed from April 1, 2014 to March 31, 2020 in

What is new?

Key findings

- Estimated treatment effects from an emulated trial of prostate cancer treatments using routinely collected national datasets were consistent with those from the PR07 trial.
- Changes in the patient selection and treatment delivery between the PR07 trial era and the emulated trial explained some differences in results.

What this adds to what is known?

- Defining time zero in an emulated trial is not straightforward when randomization is not aligned with a documented event such as diagnosis and requires careful attention.
- Defining a grace period within the cloning-censoring-and-weighting (± landmarking) approach can successfully address bias due to delays in treatment and confounding.

What is the implication and what should change now?

- Assessing the sensitivity of results from the randomized controlled trial and the trial emulations using real-world data to different modeling and analysis assumptions can provide insight into differences in findings and reassurance about validity of findings.
- Our study provides some confidence that real-world data could be used to compare the benefits of different treatment strategies in prostate cancer patient groups for which trial evidence is lacking but also highlights some of the considerations needed.
- Future work could assess alternative definitions of treatment timing, bias in selection, particularly of the control group definition, and different approaches to handling missing data.

England to emulate the PR07 trial. Datasets included the National Cancer Registration Dataset, Hospital Episode Statistics Admitted Patient Care data, Office for National Statistics death records, and National Radiotherapy Dataset (Table 1) [21−23]. Follow-up ranged from 37 to 2,849 days (7.8 years), ending on January 18, 2022.

The National Cancer Registration and Analysis Service, run by National Health Service Digital, collects and links annual cancer data from healthcare providers for all patients diagnosed in England. Under a data-sharing agreement, National Cancer Registration and Analysis Service provides anonymized PCa data to the National Prostate Cancer Audit team, which identifies variations in clinical care and supports initiatives to improve PCa research.

To assess whether datasets could emulate the PR07 trial, we examined completeness and definitions of eligibility criteria, treatment alignment and outcomes, and availability of key variables for treatment classification and confounding control. We verified the use of standardized measurements, consistent follow-up periods, and temporal alignment.

2.2. The target trial and challenges of trial emulation

Table 2 shows protocol components for PR07 and the emulated trial. Below we summarize key challenges of emulating the PR07 trial using available datasets, which inform interpretation of results from the emulated trial and comparison with the RCT.

2.2.1. Defining time zero and the treatment grace period

In the PR07 trial, randomization was allowed within 12 weeks after starting first-line HT. In the RWD, individuals are followed from diagnosis, but the time of initiating HT is not recorded, making choice of "time zero" that aligns with timing of randomization challenging. We consider two approaches to defining time zero in the emulated trial. First, we used the date of diagnosis as time zero. Second, we considered time zeros at a range of landmark times (0, 4,…, 24 weeks) from diagnosis, which may align better with randomization in the RCT. We selected a maximum of 24 weeks for the landmark times, as in PR07 pre-randomization investigations such as hemoglobin tests took a maximum of 6 months to complete.

Choice of time zero in the emulated trial also has implications for our ability to estimate the effects of an RT+HT strategy, because RT may not be initiated immediately. In clinical practice, timing of treatment initiation depends on clinical tests, administrative procedures, and the patient's availability. In the PR07 trial, RT could be initiated within 8 weeks of randomization. Using diagnosis as time zero in the emulated trial, we found that only 1 patient initiated RT within 8 weeks, while 51.2% of patients who initiate RT did so within 6 months of diagnosis. We therefore considered grace periods of longer duration (4, 5, and 6 months) when using diagnosis as time zero. In the landmark approach, where time zero is defined in a series of landmark times, a grace period of 8 weeks for RT initiation was taken, matching the grace period used in PR07.

The correct handling of a grace period in the analysis predominantly addresses immortal time that would arise if the misalignment of meeting eligibility and treatment initiation were ignored [4,24−26].

2.2.2. Defining a similar population to the PR07 trial

Inclusion and exclusion criteria in RCTs can be difficult to fully emulate using available datasets. The RWD captures data on most variables required to define the eligibility criteria, but some criteria were not recorded

(Table 2). There was substantial missingness in some variables needed to define eligibility, including performance status (50%), prostate-specific antigen (PSA) values (60%), and Gleason scores (82%). We restricted to patients with all eligibility variables recorded. PR07 excluded patients with malignant or nonmalignant diseases resulting in life expectancy of <5 years and with contraindications to pelvic RT, neither of which are explicitly recorded in RWD. For inclusion, hemoglobin, platelets, and other biomarkers had to be recorded within 4 weeks before randomization, and PSA tests within 4 weeks before HT. However, except for PSA values, these biomarkers, test timing, and HT duration were missing in our RWD. This could be one source of unobserved difference in the overall health of the emulated trial population compared to the trial population, and the absence of these variables in the RWD demonstrates the difficulty in replicating trial eligibility.

2.2.3. Confounding

In the analysis, we controlled for potential confounders: age, number of comorbidities, Gleason score, PSA levels, income deprivation quintile, performance status, T-staging, number of tumors, ethnicity, and year of diagnosis [27,28]. These variables were measured at baseline only. RT initiation is time-varying, and its initiation could be influenced by time-varying features. While the baseline measurements capture the most important confounders in PCa, recording of comorbidities (such as myocardial infarction and stroke) is incomplete in the Hospital Episode Statistics Admitted Patient Care dataset. If present, these comorbidities might influence clinicians' decisions around prescribing RT [29,30].

There was some missingness in confounding variables, but this was low among individuals meeting eligibility criteria and so we used a complete case analysis under an assumption of data being missing completely at random.

2.3. Statistical analysis

2.3.1. Time zero at diagnosis: cloning-censoring-and-weighting approach

We used the CCW approach to enable estimation of the causal estimand, as this method allows for treatment strategies with a grace period as well as controlling for measured confounders [14,31]. CCW involves first creating two copies of the data, one for each treatment strategy (RT+HT and HT) (cloning step). This ensures that the baseline characteristics of patients are balanced in both groups and eliminate confounding at baseline. The dataset corresponding to a given treatment strategy is then modified by artificially censoring individuals at the time they deviate from their strategy (censoring step). To mimic the HT strategy, individuals who receive RT within the grace period are censored at RT initiation. However, we emphasize that the HT strategy allows patients to initiate RT after the grace period. To mimic the RT+HT strategy, individuals are censored at the end of the grace period if they have not initiated RT by that time (Supplementary Fig 1). Censoring is assumed to be dependent on covariates and was addressed using inverse probability of censoring weights (IPCW), where the models for the weights included the baseline confounders (weighting step) [32]. We used Cox regression to estimate the IPCW for HT strategy dataset, as censoring could occur at any time before the end of the grace period. For the RT+HT strategy dataset, the IPCW were estimated using logistic regression as censoring occurred only at end of the grace period. We estimated marginal survival probabilities under each treatment strategy using IPCW-weighted Kaplan-Meier analysis applied to the HT and RT+HT strategy datasets and obtained marginal risk differences up to 7 years [33]. To estimate an HR comparing the two treatment strategies, we combined the two datasets and applied an IPCW-weighted Cox proportional hazards regression with the only covariate being the binary indicator for treatment strategy [34].

2.3.2. A landmark-CCW approach

After setting time zero to a series of landmark times (0, 4,…, 24 weeks post-diagnosis), we used a landmark extension of the CCW approach (''landmark-CCW''). A dataset was created for each landmark time. The dataset for each landmark time was restricted to individuals who remained alive and under observation, and who had not initiated RT before then (Supplementary Fig 2). The cloning and

Table 1. Summary of the routinely collected administrative and clinical national linked datasets for prostate cancer patients in England

Dataset	Contents	Number of patients
National Cancer Registration Dataset (NCRD) including Cancer Outcomes and Services Dataset (COSD)	Includes baseline diagnostic and demographic characteristics of people diagnosed with cancer in England, including age, date of diagnosis, prostate-specific antigen (PSA) level, Gleason score, and Tumor-Node-Metastases (TNM) staging [21].	261,919
Hospital Episode Statistics Admitted Patient Care (HES APC) dataset	Includes records of patient, administrative and geographical information, including hospital admissions, dates of discharge, diagnosis codes, and operations received [22].	255,822
National Radiotherapy Dataset (RTDS)	Includes information on radiotherapy treatments, the primary disease being treated, primary procedures performed, radiotherapy intent, and doses [23].	88,448
Office for National Statistics (ONS) death records	Includes causes and dates of deaths, and demographic details of deceased patients.	261,919

Table 2. Summary of the target trial (PR07) and the emulated trial protocol components to study treatments in high-risk prostate cancer patients using the cloning-censoring-and-weighting (CCW) approach and the landmark-CCW approach

Protocol component	Target trial: PR07 trial	Emulation using routinely collected administrative and clinical national linked datasets from England
Eligibility criteria	Histological diagnosis of prostatic adenocarcinoma; high-risk patient as defined by clinical stage (T3 and T4, N0 or NX, M0 and included T2), Gleason scores 8–10, and PSA $\geq$ 20 ng/mL; age $<$80 y; no history of cancer and absence of specified comorbidities; no cytotoxic anticancer therapy prior to randomization; no nonprotocol anticancer therapy until documented progression of the prostate cancer; laboratory tests completed within 4 wks before randomization; no radical prostatectomy; baseline PSA taken within 12 wks prior to any hormone therapy; bone scans confirming no bony metastases; and histologically negative lymph node dissections. Exclusions: no history of previous malignancies or nonmalignancies resulting in life expectancy of $<$5 y; no small-cell or transitional-cell carcinoma in biopsy specimen; no prior treatment for carcinoma of the prostate; and no contraindications to pelvic radiotherapy.	As in the target trial, using patients diagnosed from 2014 to 2020. Differences in criteria include that these datasets had no information on laboratory tests; when PSA levels were recorded; duration of life expectancy when diagnosed with certain malignancies or nonmalignancies; contraindications to pelvic radiotherapy; indicator of whether a bone scan was conducted and when that was done; biopsy specimen indicating no small-cell or transitional cell carcinoma; treatments received before randomization; and lymph node dissections recorded within 12 wks of randomization.
Treatment strategies	Strategy 1: Initiate ablative long-term hormone therapy (choose either luteinizing hormone-releasing hormone agonist or bilateral orchiectomy) within 4 wks of randomization. Use of radiotherapy was only allowed if disease progression occurred and was at the discretion of the physician. Strategy 2: As in strategy 1, plus the addition of radiation therapy within 8 wks of randomization. The pelvic radiation dose was 45 Gy in 25 fractions, followed by 20–24 Gy to the prostate in 10–12 fractions. Patients unsuitable for whole pelvis RT received 60–65 Gy to the prostate in 35–9 fractions.	Strategy 1: Initiate hormone therapy soon after diagnosis. Radiotherapy is allowed after the grace period. Strategy 2: Use hormone therapy and initiate radiotherapy within the grace period after diagnosis. We considered grace periods of 8 wks (as in the target trial) and 4, 5, and 6 mo. Radiation dose and fractions are as specified in the target trial.
Assignment procedures	Randomization with minimization stratified by the following factors: institution, PSA levels at diagnosis, Gleason scores, choice of hormonal therapy, prior hormonal therapy, and method of lymph node staging.	Randomization is emulated via adjustment for baseline covariates, including demographics, measures of disease status, and comorbidities. The covariates include age, ethnicity, year of diagnosis, Gleason score, T-staging, PSA levels, number of comorbidities, performance status, number of tumors, and income deprivation quintile.
Follow-up period	Starts at randomization and ends at the earliest of death, loss to follow-up, or end of study (7 y), whichever occurs first.	Starts at diagnosis and ends at the earliest of death, loss to follow-up, or end of study (maximum of 7 y), whichever occurs first.
Outcome of interest	Overall survival defined as the time from randomization to the time of death from any cause or to the date of last follow-up.	Overall survival defined as the time from diagnosis to the time of death from any cause or to the date of last follow-up.
Causal contrasts of interest	Intention-to-treat effect measured using a hazard ratio and differences in risk up to 7 y.	Per-protocol effect measured using a hazard ratio and differences in risk up to 7 y.
Analysis plans	Overall survival probabilities estimated using with the Kaplan-Meier method. Log-rank test stratified by the minimizing factors at randomization. HRs estimated with the Cox model.	(1) Cloning-censoring-and-weighting (CCW) analysis for strategies 1 and 2 using 4–6 months grace period. (2) Landmark-CCW analysis for strategies 1 and 2 using 8 wk grace period.

CCW, cloning-censoring-and-weighting; HRs, hazard ratios; PSA, prostate-specific antigen; RT, radiotherapy.

censoring steps of the CCW approach were then implemented for each landmark dataset, using a grace period of 8 weeks for RT initiation. The models for the IPCW were estimated as in the CCW approach for each landmark dataset. The landmark datasets were then combined for the final analysis, with a weighted Kaplan-Meier analysis fitted for

survival probability estimates and corresponding risk differences, and a weighted Cox model for HRs. Marginal survival probabilities and risk differences refer to the combined eligible population across the landmark times, which differ from the population in the CCW approach. Therefore, in additional analysis, we used standardization to target the same estimand in the landmark-CCW and CCW analyses. A conditional Cox model was fitted to the combined landmark-CCW datasets, including a binary indicator for treatment strategy and potential confounders (the same set as in the weights models). Conditional survival probabilities were obtained using Breslow's estimator under each treatment strategy for each individual in the original CCW analysis, that is, for each individual meeting the emulated trial eligibility criteria at diagnosis. Survival probabilities for each treatment strategy were averaged over individuals to give marginal estimates and corresponding risk differences.

2.3.3. Assumptions and uncertainty

To interpret results from the analyses as estimates of causal estimands of interest, we rely on key assumptions of no unmeasured confounding (all confounders have been measured and controlled for), positivity (each PCa patient

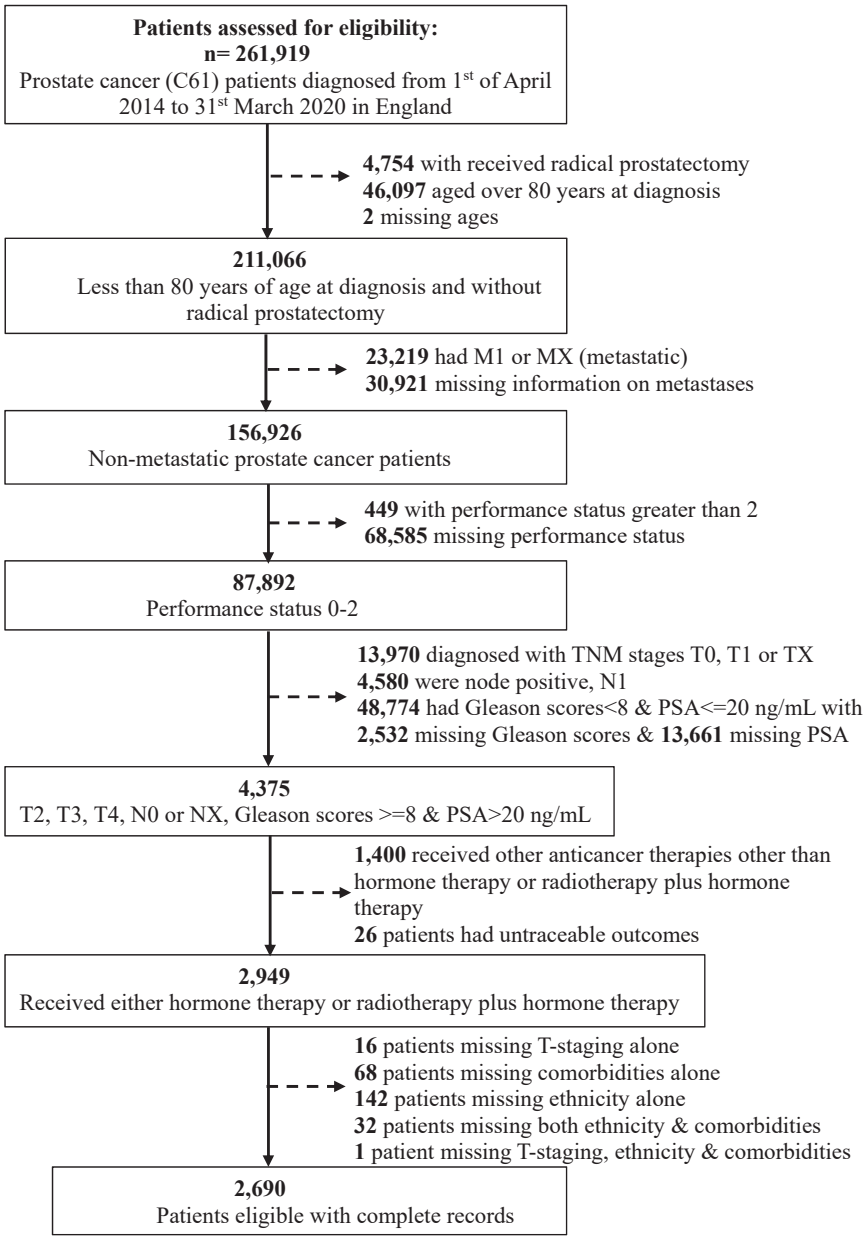

Figure 1. Emulated trial profile using the routinely collected administrative and clinical national datasets for patients diagnosed between 2014 and 2020 with prostate cancer. The corresponding trial profile from the PR07 trial can be found in [11].

Table 3. Summary of the overall baseline characteristics of the emulated trial using complete cases from the routinely collected administrative and clinical national linked datasets (RWD) and the PR07 trial [11]

Variable	PR07 overall characteristics, $n = 1,205$ (%)	RWD overall characteristics, $n = 2,690$ (%)
Total	1,205 (100.0)	2,690 (100.0)
Age		
Mean (SD)	69.7 (6.1)	70.3 (5.9)
Gleason scores		
8	978 (81.2)	997 (37.1)
9–10	218 (18.1)	1,693 (62.9)
Performance status[a]		
0	943 (78.3)	1,920 (71.4)
1	245 (20.3)	635 (23.6)
2	17 (1.4)	135 (5.0)
T-staging		
T2	146 (12.1)	689 (25.6)
T3	1,000 (83.0)	1,836 (68.3)
T4	57 (4.7)	165 (6.1)
Number of tumors		
1	-	2,290 (85.1)
2	-	335 (12.5)
3 or more	-	65 (2.4)
Prostate-specific antigen (PSA) levels		
20–50 ng/mL	900 (74.7)	1,689 (62.8)
>50 ng/mL	305 (25.3)	1,001 (37.2)
Number of comorbidities		
0	-	2,196 (81.6)
1	-	424 (15.8)
2 or more	-	70 (2.6)
Income deprivation quintile		
1: least deprived	-	554 (20.6)
2	-	628 (23.3)
3	-	578 (21.5)
4	-	506 (18.8)
5: most deprived	-	424 (15.8)
Year of diagnosis		
2014–2016	-	979 (36.4)
2017–2019	-	1,594 (59.3)
2020	-	117 (4.3)
Ethnicity		
White	-	2,555 (95.0)
Mixed	-	11 (0.4)
Asian	-	33 (1.2)
Black	-	57 (2.1)
Other	-	34 (1.3)

RWD, real-world data; SD, standard deviation.

Results are presented as ''number (%)'' or ''mean (SD)'' where indicated, and ''-'' indicates that the information was not reported.

[a] Performance status: 0 = able to perform duties normally without restriction; 1 = restricted in physical strenuous activity but able to walk and do light work; 2 = able to walk and capable of all self-care, but unable to carry out any work. Up and about more than 50% of waking hours.

has a non-zero probability of being assigned to each treatment strategy), and consistency (the individual's observed outcome under their actual treatment strategy is the same as their potential outcome). It is difficult to assess unmeasured confounding in practice. The consistency assumption was believed to be justified as RT treatment in England is well standardized. For the positivity assumption, we assessed the distribution of weights for the two treatment strategies, as extremely large weights can indicate that certain patients have a very low probability of receiving RT+HT or HT given their covariates. All approaches rely on correct model specification for the weights.

Bootstrapping with 1,000 replicates was used to obtain 95% confidence intervals (CIs) based on a normal distribution for the estimates.

3. Results

The data included 261,919 individuals diagnosed with PCa between April 2014 and March 2020. More than 55% of patients (143,243 of 261,919 in Fig 1) were not eligible for the PR07 trial and excluded. Of the remainder, 97.5% had 1 or more missing eligibility criteria. One or more covariates were missing for 259 of 2,949 eligible patients (8.7%) (Fig 1, Supplementary Table 1). We restricted analysis to the remaining 2,690 patients (Table 3). Of these, 2,201 (82%) received RT+HT and 489 (18%) received HT alone. Median time of initiating RT was 161 days (range: 49 to 1,863 days) after diagnosis (Supplementary Fig 4). Only 1 (0.04%) patient initiated RT within 8 weeks following diagnosis. For the CCW approach, 314 (11.7%) individuals initiated RT within 4 months, 946 (35.2%) within 5 months, and 1,377 (51.2%) within 6 months of diagnosis. For landmark times, $0, 4, 8, 12, 16, 20,$ and 24 weeks after diagnosis, the number of patients initiating RT within a grace period of 8 weeks were as follows, where the percentages are relative to the number meeting eligibility criteria at the landmark time: 1 (0.04%), 7 (0.3%), 143 (5.3%), 687 (25.6%), 1082 (42.6%), 814 (410%), and 509 (35.0%).

Baseline characteristics for participants in the PR07 and emulated trials are summarized in Table 3. Mean ages for both trials were similar. PR07 patients had worse stage, but better Gleason scores, performance status, and PSA values compared to the emulated trial population (Table 3).

3.1. Hazard ratios and risk estimates

Results are presented in Table 4 and Figure 2. The PR07 trial reported HR of 0.77 (95% CI: 0.61–0.98). Estimated survival probabilities at 7 years were 74.0% (95% CI: 70.0–77.0) for RT+HT and 66.0% (60.0–68.0) for HT only. The CCW analysis using 4, 5, and 6 months grace periods resulted in HRs of 0.64 (95% CI: 0.28–0.78), 0.54 (95% CI: 0.30–0.73), and 0.48 (95% CI: 0.34–0.60).

Table 4. Estimated HR and survival probabilities at 7 years from the PR07 trial and the emulated trial using the cloning-censoring-and-weighting (CCW) approach with 4-6 months grace period (time within which radiotherapy is initiated in the radiotherapy treatment strategy) and the landmark-CCW approach

Trial/analysis approach	HR (95% CI)	Survival probability (95% CI): RT+HT	Survival probability (95% CI): HT only	Risk difference (95% CI)
PR07 trial	0.77 (0.61, 0.98)	74.0% (70.0, 77.0)	66.0% (60.0, 68.0)	-
Emulated trial				
CCW: grace period of 4 mo	0.64 (0.28, 0.78)	85.5% (66.0, 100.0)	71.0% (63.6, 74.8)	14.2% (10.7, 19.8)
CCW: grace period of 5 mo	0.54 (0.30, 0.73)	81.2% (70.3, 86.3)	68.4% (64.0, 69.4)	12.8% (6.2, 21.2)
CCW: grace period of 6 mo	0.48 (0.34, 0.60)	80.7% (74.3, 87.0)	65.6% (62.8, 68.1)	15.1% (11.5, 18.9)
Landmark-CCW: grace period of 8 wks	0.58 (0.51, 0.65)	80.7% (77.7, 83.8)	69.8% (68.2, 71.4)	10.9% (6.3, 15.9)

CCW, cloning-censoring-and-weighting; CI, confidence interval; HR, hazard ratio; HT, hormone therapy; RT, radiotherapy.

Estimated survival probabilities at 7 years using 4−6 months grace periods were 85.5% (95% CI: 66.0−100.0), 81.2% (95% CI: 70.3−86.3), and 80.7% (95% CI: 74.3−87.0) for the RT+HT strategy vs 71.0% (95% CI: 63.6−74.8), 68.4% (95% CI: 64.0−69.4), and 65.6 (95% CI: 62.8−68.1) for the HT only strategy. Corresponding estimated risk differences for the 4−6 months grace periods were 14.2% (95% CI: 10.7−19.8), 12.8% (95% CI: 6.2−21.2), and 15.1% (95% CI: 11.5−18.9), respectively.

The HR from the landmark-CCW approach was 0.58 (95% CI: 0.51−0.65) with survival probability estimates of 80.7% (95% CI: 77.7−83.8) for RT+HT and 69.8% (95% CI: 68.2−71.4) for HT only, with risk difference of 10.9% (95% CI: 6.3−15.9). These results show that the estimated survival under RT+HT and HT strategies for the emulated trial is better than in PR07, with the difference being greater for the RT+HT strategy.

3.2. Assessing identifiability assumptions

For the positivity assumption, there was overlap in the distribution of weights in both treatment strategies (Supplementary Fig 5). The standardized mean differences at the end of the 6-month grace period showed good covariate balance after the weighting step in the CCW approach (Supplementary Fig 3) [35,36].

4. Discussion

We emulated the PR07 trial, or a closely related trial, using routine national linked datasets in England. Results indicate improved overall survival from RT+HT compared to HT only. Despite the same overall conclusion as PR07 trial, numerical results differed, with our emulated trial analysis giving HRs further from one, and larger risk differences, estimating a greater effect of RT+HT. The treatment effect in our emulated trial suggests that results from both trials are in the same direction. Despite RWD including frailer patients (poor performance status, Gleason scores, and PSA values), survival probabilities were lower under both treatment strategies in the PR07 trial. Differences in results may stem from residual confounding and variation in patient characteristics. Since PR07 trial (1995−2005), there have been significant advances in RT treatment for PCa [37]. Treatment guidelines now favor RT+HT for high-risk patients [38−40]. Additionally, some men now receive both RT and surgery and excluding those treated with surgery could introduce selection bias.

Aligning time zero and accounting for time-varying treatment status were key challenges in emulating the PR07 trial. We addressed these by considering several grace periods for RT initiation. HRs from the CCW moved slightly further from one with longer grace periods. Consistent with the PR07 trial, we allowed those following the HT only strategy to receive RT after the grace period. This could in part explain why shorter grace periods gave a weaker treatment effect. Information on the numbers of patients in the HT only arm of PR07 who later received RT was not available. The landmark-CCW approach aimed to address the challenge of aligning time zero with that in the PR07 trial, and to allow for a grace period of the same duration. We obtained similar estimates from this approach, although with narrower CIs as we would expect. Future work could further explore the impacts of the timing of RT initiation relative to diagnosis and consider control strategies in which RT is delayed for a longer period.

A key strength of our study was the use of routine national data from reliable sources, which enabled us to use variables needed to define eligibility and to control for confounding. An important limitation was substantial missing data on variables that defined eligibility for the emulated trial, which could alter the study sample and limit generalizability of our findings. Missing data in the emulated trial cohort made it difficult to expand our results to a wider population, although our focus was on emulating the RCT rather than on extending to different or wider populations. However, there were relatively few missing adjustment variables among eligible patients (8.7%). Information on the timing and duration of HT was also unavailable.

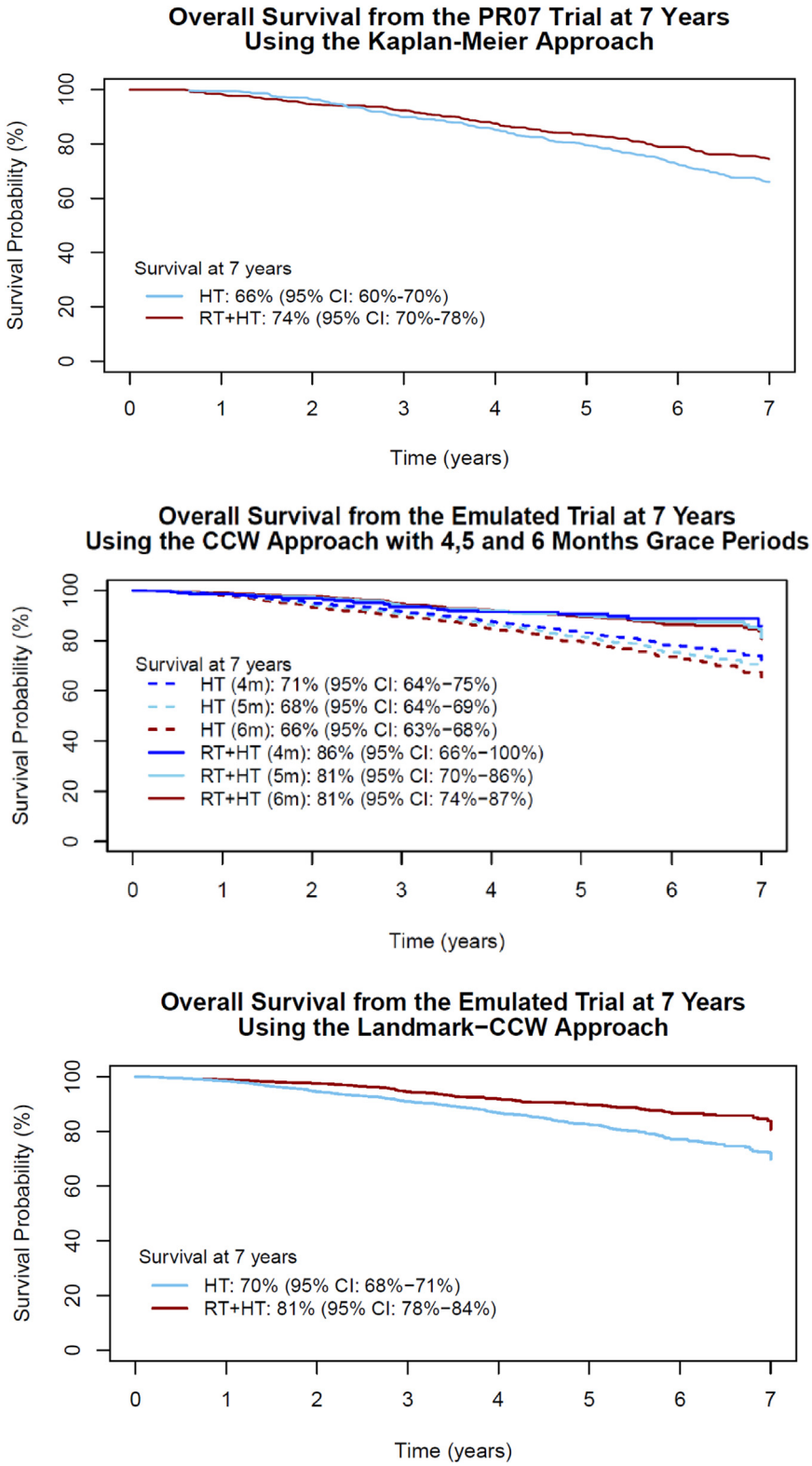

Figure 2. Estimated survival curves from the PR07 trial and the emulated trial using the cloning-censoring-and-weighting (CCW) approach with 4–6 months grace period (time within which radiotherapy is initiated in the radiotherapy treatment strategy) and the landmark-CCW approach.

5. Conclusion

The quality of RWD varies and combining data from multiple sources can help capture a wider range of patient characteristics for analysis. Trial emulation using high-quality national RWD can address issues such as confounder adjustment and delayed treatment. However,

difficulties in defining ''time zero'', residual confounding and heterogeneous populations may affect the accuracy of estimates. Assessing sensitivity of results to different modeling and analysis assumptions and understanding the differences between RCT results and trial emulations using RWD can provide reassurance. Our study provides some confidence that RWD could be used to compare different treatment strategies in PCa patient groups for which trial evidence is lacking but also highlights some of the considerations needed. Future work could assess alternative definitions of treatment timing, potential selection bias in the HT only strategy, and different approaches to handling missing data.

Ethics approval

This study received ethical approval by the London School of Hygiene and Tropical Medicine Research Ethics Committee in the United Kingdom (reference: 27294/RR/29250).

CRediT authorship contribution statement

Caroline Chesang: Writing − original draft, Writing − review & editing, Methodology, Formal analysis, Data curation. **Linda D. Sharples:** Writing − review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Christen M. Gray:** Writing − review & editing, Supervision, Funding acquisition. **Julie Nossiter:** Writing − review & editing, Data curation. **Jan van der Meulen:** Writing − review & editing, Data curation. **Thomas E. Cowling:** Writing − review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Ruth H. Keogh:** Writing − review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

C. G. reports a relationship with AstraZeneca that includes employment and equity or stocks. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work uses data that have been provided by patients and collected by the NHS as part of their care and support, including data that are collated, maintained, and quality assured by the National Disease Registration Service, which is part of NHS England. Access to the data was facilitated by the NHS England Data Access Request Service.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2025.111767.

Data availability

The authors do not have permission to share data.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209−49. https://doi.org/10.3322/caac.21660.

[2] NICE Guidance - Prostate cancer: diagnosis and management. Prostate cancer: diagnosis and management. (2019). BJU Int 2019;124(1):9−26. https://doi.org/10.1111/bju.14809.

[3] Gnanapragasam VJ, Bratt O, Muir K, Lee LS, Huang HH, Stattin P, et al. The Cambridge Prognostic Groups for improved prediction of disease mortality at diagnosis in primary non-metastatic prostate cancer: a validation study. BMC Med 2018;16:1−10. https://doi.org/10.1186/s12916-018-1019-5.

[4] Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183(8):758−64. https://doi.org/10.1093/aje/kwv254.

[5] Schneeweiss S. Real-world evidence of treatment effects: the useful and the misleading. Clin Pharmacol Ther 2019;106(1):43−4. https://doi.org/10.1002/cpt.1405.

[6] Lodi S, Phillips A, Lundgren J, Logan R, Sharma S, Cole SR, et al. Effect estimates in randomized trials and observational studies: comparing apples with apples. Am J Epidemiol 2019;188(8):1569−77. https://doi.org/10.1093/aje/kwz100.

[7] Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation differences vs. Biases when calibrating real-world evidence findings against randomized controlled trials. Clin Pharmacol Ther 2020;107(4):735−7. https://doi.org/10.1002/cpt.1793.

[8] Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. Circulation 2021;143(10):1002−13. https://doi.org/10.1161/CIRCULATIONAHA.120.051718.

[9] Wang SV, Schneeweiss S, Franklin JM, Desai RJ, Feldman W, Garry EM, et al. Emulation of randomized clinical trials with nonrandomized database analyses: results of 32 clinical trials. JAMA 2023;329(16):1376−85. https://doi.org/10.1001/jama.2023.4221.

[10] Baptiste PJ, Wong AY, Schultze A, Clase CM, Leyrat C, Williamson E, et al. Cardiorenal effects of Angiotensin-converting enzyme inhibitors and Angiotensin receptor blockers in people underrepresented in trials: analysis of routinely collected data with emulation of a reference trial (ONTARGET). Am J Epidemiol 2024;193(12):1785−95. https://doi.org/10.1093/aje/kwae137.

[11] Warde P, Mason M, Ding K, Kirkbride P, Brundage M, Cowan R, et al. Combined androgen deprivation therapy and radiation therapy for locally advanced prostate cancer: a randomised, phase 3 trial. Lancet 2011;378(9809):2104−11. https://doi.org/10.1016/s0140-6736(11)61095-7.

[12] Brundage M, Sydes MR, Parulekar WR, Warde P, Cowan R, Bezjak A, et al. Impact of radiotherapy when added to androgen-

deprivation therapy for locally advanced prostate cancer: long-term quality-of-life outcomes from the NCIC CTG PR3/MRC PR07 randomized trial. J Clin Oncol 2015;33(19):2151−7. https://doi.org/10.1200/JCO.2014.57.8724.

[13] Mason MD, Parulekar WR, Sydes MR, Brundage M, Kirkbride P, Gospodarowicz M, et al. Final report of the intergroup randomized study of combined androgen-deprivation therapy plus radiotherapy versus androgen-deprivation therapy alone in locally advanced prostate cancer. J Clin Oncol 2015;33(19):2143−50. https://doi.org/10.1200/JCO.2014.57.7510.

[14] Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. BMJ 2018;360:k182. https://doi.org/10.1136/bmj.k182.

[15] Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol 1983;1(11):710−9. https://doi.org/10.1200/JCO.1983.1.11.71.

[16] Van Houwelingen HC. Dynamic prediction by landmarking in event history analysis. Scand J Stat 2007;34(1):70−85.

[17] Dafni U. Landmark analysis at the 25-year landmark point. Circ Cardiovasc Qual Outcomes 2011;4(3):363−71. https://doi.org/10.1161/CIRCOUTCOMES.110.957951.

[18] Aalen OO, Cook RJ, Røysland K. Does Cox analysis of a randomized survival study yield a causal treatment effect? Lifetime Data Anal 2015;21:579−93. https://doi.org/10.1007/s10985-015-9335-y.

[19] Hernán MA. The hazards of hazard ratios. Epidemiology 2010;21(1):13−5. https://doi.org/10.1097/EDE.0b013e3181c1ea43.

[20] Stensrud MJ, Aalen JM, Aalen OO, Valberg M. Limitations of hazard ratios in clinical trials. Eur Heart J 2019;40(17):1378−83. https://doi.org/10.1093/eurheartj/ehy770.

[21] Henson KE, Elliss-Brookes L, Coupland VH, Payne E, Vernon S, Rous B, et al. Data resource profile: national cancer registration dataset in England. Int J Epidemiol 2020;49(1):16. https://doi.org/10.1093/ije/dyz076.

[22] Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. BMJ 2013;346:f2350. https://doi.org/10.1136/bmj.f2350.

[23] Sandhu S, Sharpe M, Findlay Ú, Roe C, Broggio J, Spencer K, et al. Cohort profile: radiotherapy dataset (RTDS) in England. BMJ Open 2023;13(6):e070699. https://doi.org/10.1136/bmjopen-2022-070699.

[24] Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol 2008;167(4):492−9. https://doi.org/10.1093/aje/kwm324.

[25] Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernán MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. Int J Biostat 2010;6(2):18. https://doi.org/10.2202/1557-4679.1212.

[26] Huitfeldt A, Kalager M, Robins JM, Hoff G, Hernán MA. Methods to estimate the comparative effectiveness of clinical strategies that administer the same intervention at different times. Curr Epidemiol Rep 2015;2(3):149−61. https://doi.org/10.1007/s40471-015-0045-5.

[27] Zhou EH, Ellis RJ, Cherullo E, Colussi V, Xu F, Chen WD, et al. Radiotherapy and survival in prostate cancer patients: a population-based study. Int J Radiat Oncol Biol Phys 2009;73(1):15−23. https://doi.org/10.1016/j.ijrobp.2008.04.001.

[28] Aas K, Berge V, Myklebust TA, Fossa SD. Comparative survival outcomes of high-risk prostate cancer treated with radical prostatectomy or definitive radiotherapy regimens. Eur Urol open Sci 2021;26:55−63. https://doi.org/10.1016/j.euros.2021.01.011.

[29] Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data resource profile: hospital episode statistics admitted patient care (HES APC). Int J Epidemiol 2017;46(4):1093. https://doi.org/10.1093/ije/dyx015.

[30] Morgan A, Sinnott SJ, Smeeth L, Minassian C, Quint J. Concordance in the recording of stroke across UK primary and secondary care datasets: a population-based cohort study. BJGP Open 2021;5(2):BJGPO.2020.0117. https://doi.org/10.3399/BJGPO.2020.0117.

[31] Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. Int J Epidemiol 2020;49(5):1719−29. https://doi.org/10.1093/ije/dyaa057.

[32] Willems S, Schat A, van Noorden MS, Fiocco M. Correcting for dependent censoring in routine outcome monitoring data by applying the inverse probability censoring weighted estimator. Stat Methods Med Res 2018;27(2):323−35. https://doi.org/10.1177/0962280216628900.

[33] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53(282):457−81.

[34] Cox DR. Regression models and life-tables. J R Stat Soc Ser B (Methodological) 1972;34(2):187−202.

[35] Ahmed A, Husain A, Love TE, Gambassi G, Dell'Italia LJ, Francis GS, et al. Heart failure, chronic diuretic use, and increase in mortality and hospitalization: an observational study using propensity score methods. Eur Heart J 2006;27(12):1431−9. https://doi.org/10.1093/eurheartj/ehi890.

[36] Greifer N. Cobalt: covariate balance tables and plots. R Package Version 4.5.5 [Software]. 2024. Available at: https://cran.r-project.org/web/packages/cobalt/vignettes/cobalt.html. Accessed May 3, 2024.

[37] Parry MG, Nossiter J, Cowling TE, Sujenthiran A, Berry B, Cathcart P, et al. Patient-reported functional outcomes following external beam radiation therapy for prostate cancer with and without a high-dose rate brachytherapy boost: a national population-based study. Radiother Oncol 2021;155:48−55. https://doi.org/10.1016/j.radonc.2020.10.019.

[38] Widmark A, Klepp O, Solberg A, Damber JE, Angelsen A, Fransson P, et al. Endocrine treatment, with or without radiotherapy, in locally advanced prostate cancer (SPCG-7/SFUO-3): an open randomised phase III trial. Lancet 2009;373(9660):301−8. https://doi.org/10.1016/s0140-6736(08)61815-2.

[39] Bolla M, Van Tienhoven G, Warde P, Dubois JB, Mirimanoff RO, Storme G, et al. External irradiation with or without long-term androgen suppression for prostate cancer with high metastatic risk: 10-year results of an EORTC randomised study. Lancet Oncol 2010;11(11):1066−73. https://doi.org/10.1016/s1470-2045(10)70223-0.

[40] Jones CU, Hunt D, McGowan DG, Amin MB, Chetner MP, Bruner DW, et al. Radiotherapy and short-term androgen deprivation for localized prostate cancer. N Engl J Med 2011;365(2):107−18. https://doi.org/10.1056/NEJMoa1012348.