|  |
| --- |
|  |
| **Janssen Research & Development\*** |
|  |
| **Study Protocol for Retrospective Observational Studies Using Secondary Data** |
|  |
| **RWE Safety Pilot: Characterizing the risks of cardiovascular disease among patients being treated for castration-resistant prostate cancer** |
|  |
| **Protocol [insert protocol number]** |

\* Janssen Research & Development (Janssen R&D) is a global organization that operates through different legal entities in various countries. Therefore, the legal entity acting as the sponsor for studies of Janssen R&D may vary. The term "sponsor" is used throughout the protocol to represent these various legal entities.

**Status:** [Draft][Pending Review][Approved]   
[Make sure to update the Date field when changing the Status!]

**Date:** [Day Month Year]

**Prepared by:** [Insert the legal entity name of the company preparing this document (eg, Janssen Research & Development, LLC; Janssen Research & Development, a division of Janssen Pharmaceutica NV]

**EDMS number:** [insert EDMS no, and version] [eg EDMS-ERI-1234567, 1.0]

**Confidentiality Statement**

The information in this document contains trade secrets and commercial information that are privileged or confidential and may not be disclosed unless such disclosure is required by applicable law or regulations. In any event, persons to whom the information is disclosed must be informed that the information is *privileged* or *confidential* and may not be further disclosed by them. These restrictions on disclosure will apply equally to *all* future information supplied to you that is indicated as *privileged or confidential*.

TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS 3

2. RESPONSIBLE PARTIES 4

2.1. Investigator(s) and Authors 4

2.2. Sponsor 5

3. AMENDMENTS AND UPDATES 5

4. RATIONALE AND BACKGROUND 5

5. STUDY Objectives 7

5.1. Primary Objectives 7

5.2. Secondary Objectives 8

6. Research METHODs 8

6.1. Overview 8

6.2. Study Design and Setting 8

6.2.1. IBM MarketScan® Commercial Claims and Encounters Database (CCAE) v1061 9

6.2.2. IBM MarketScan® Medicare Supplemental Database (MDCR) v1062 9

6.2.3. Optum® De-Identified Clinformatics® Data Mart Database - Date of Death (DOD) v1064 10

6.3. Study Population(s) 10

6.4. Outcomes, Follow-Up, and Time-at-Risk 14

7. Negative and Positive Control Outcomes 15

8. SAMPLE SIZE AND STUDY POWER 17

9. DATA ANALYSIS PLAN 18

9.1. Patient Characteristics Summary 18

9.2. Model Specification 19

9.3. Evidence Evaluation 20

10. STRengths and limitations of the research methods 22

10.1. Strengths 22

10.2. Limitations 23

11. PROTECTION OF HUMAN SUBJECTS 25

12. SAFETY DATA COLLECTION AND REPORTING 25

13. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS 26

14. list of tables & figures 26

15. ANNEX (LIST OF STAND-ALONE DOCUMENTS) 26

16. references 27

# LIST OF ABBREVIATIONS

ADT – Androgen deprivation therapy

AMI – Acute myocardial infarction

ASO – Administrative services only

AUC – Area under the (receiver operating) curve

CCAE – IBM Commercial Claims and Encounters database

CDM – Common Data Model

CPT – Current Procedural Terminalogy

CRPC - Castration-resistant prostate cancer

CVD – Cardiovascular disease

DMF – Death master file

DOD – Date of death  
ESRD – End-stage renal disease

ETL - Extract, transform, and load

FAERS - FDA Adverse Event Reporting System

GnRH – Gonadotropin-releasing hormones

HCPCS - Healthcare Common Procedure Coding System

HF – Heart failure

ICD – International Classification of Diseases

IQR – Interquartile range

IRB – Institutional Review Board

ITT – Intent-to-treat

MDCR – IBM Medicare Supplemental database

MDRR – Minimally detectable relative risk

MedDRA – Medical Dictionary for Regulatory Activities

OMOP – Observational Medical Outcomes Partnership

OHDA – Observational Health Data Analytics (department within Janssen Research & Development)

OR – Odds ratio

PCa – Prostate cancer  
PSADT – Prostate-specific antigen doubling time

ROC – Receiver operating curve

RWE – Real-world evidence

RWD – Real-world data

SD – Standard deviation

SNOMED – Systematized Nomenclature of Medicine

LOINC - Logical Observation Identifiers Names and Codes

# RESPONSIBLE PARTIES

## Investigator(s) and Authors

Dina Gifkins, PhD, MPH, Senior Director, Epidemiology

Bo Fan, MD, Medical Safety Officer, Global Medical Safety

Chun-yuan Guo, MD, Medical Safety Officer, Global Medical Safety

Gowtham Rao, MD, PhD, Senior Director, Epidemiology Analytics

Peter De Porre, MD, Executive Medical Director, Clinical Oncology

Gerhard Leitz, MD, Director, Clinical Research

Suneel Mundle, MD, Global Medical Affairs Leader

Mitchell M. Conover, PhD, Manager, Observational Health Data Analytics

James Weaver, Associate Director, Observational Health Data Analytics

*(others please add)*

## Sponsor

Paul Chang, MD, Chief Safety Officer, Pharmaceuticals, Janssen Research and Development

Paul Stang, PhD, Vice President, Global Epidemiology, Pharmaceuticals, Janssen Research and Development

# AMENDMENTS AND UPDATES

| **Number** | **Date** | **Section of Study Protocol** | **Amendment or Update** | **Reason** |
| --- | --- | --- | --- | --- |
| 1 |  |  |  |  |
| 2 |  |  |  |  |

# RATIONALE AND BACKGROUND

Prostate cancer (PCa) is second most common cancer among men and is the most frequently diagnosed cancer in 105 countries, with an estimated 1,276,106 new cases diagnosed in 2018 (approximately 7% of all incident cancer cases) (Bray et al. 2018), for an age standardized incidence risk of 29.3/100,000 population. For 2010-2014 the age–adjusted incidence rate of prostate cancer in the US was 126.7/100,000 (Howlader, 2017). Prostate cancer occurs primarily in older men, with the highest incidence rates in the US being in the 70-74 age group at 673.1 /100,000, and 65-69 years at 651.7 /100,000, There were almost no occurrences below the age of 40, and the incidence rate for ages 45-49 in 2011-2015was 36.0/100,000 population (Noone AM et al. 2018). Among a US study of men diagnosed with localized PCa between 2000 and 2014, 14.9% eventually developed metastatic hormone-sensitive PCa (Pascale et al. 2017).

Prostate cancer has a variable naturalhistory. A large proportion of all histopathological lesions currently diagnosed as prostate cancers progress either very slowly, or not at all, and some may even regress. For men with initially untreated PCa there is a progression rate to lethal disease of 0.5–1.5 % per year (Holmberg et al. 2014). Men newly diagnosed with PCa are likely to have low-grade disease which can remainindolent for more than a decade following diagnosiswith less than 6% progressing during the first 10 years of follow-up. The time between diagnosis and the appearance of clinical symptoms is estimated to be as much as 10 years for men in their 50s and 5 years for men in their 70s. However***,*** high-grade cancers frequentlyprogress and lead to prostate cancer mortality withina decade.(Albertsen 2015). In the US it was estimated that among men diagnosed with prostate cancer, 80% presented with localized disease, 12% had regional disease, and 4% had distant disease (Punnen et al. 2013). The majority of patients with localized PCa have a five-year survival near 100%; but once the tumor progresses developing distant metastasis, the disease often become incurable. The most common metastatic sites are bone and lymph nodes, but visceral metastases may also be present and may be associated with a more severe clinical course (Pascale et al. 2017)

Globally, there were an estimated 358,989 deaths due to PCa in 2018, for an age-standardised risk of 7.6/100,000 population, and PCa is the fifth leading cause of death from cancer in men (Bray, Ferlay et al. 2018). The prognosis for patients with CRPC is poor compared to those with castration sensitive PCa (Kirby 2011). It has been reported that 33% - 46% of NM-CRPC patients develop bone metastasis within 2 years of diagnosis, (Kirby, 2011) (Smith, 2011). Median prostate-specific antigen doubling time (PSADT) after CRPC diagnosis was 6.5 months in 1 trial of 160 patients, of whom 26.9% were metastatic at diagnosis (Shulman 2004) indicating that those with a rapidly rising PSA are at risk of developing metastatic disease. An examination of the placebo arms of clinical trials has shown that NM-CRPC has a median bone-metastasis free survival of approximately 2 years, (Hong, 2014).

Comorbidities for PCa patients on ADT include cardiovascular conditions, depression, diabetes , gastric acid disorders, hyperlipidaemia, and osteoporosis (Ng et al. 2018) Comorbidities specific to CRPC patients include hypertension, dyspnoea, cardiac disease (including ischaemic heart disease/angina, renal failure/impairment, diabetes mellitus, peripheral oedema, hypotension, urinary disorders, anaemia, digestive disorders, and respiratory infections (Hirst et al. 2012).

Older men with prostate cancer receiving long-term ADT have been shown to exhibit significant functional and physical impairment, putting them at increased risk for cardiovascular disease, falls and fracture, and cognitive impairment (Bylow 2008; Sun, 2017).

Published reports have suggested that there may be an association between ADT with GnRH analogues (with or without an anti-androgen) or bilateral orchiectomy and the incidence of cardiac disease, including myocardial infarction and mortality ([Levine et al, 2010](#_ultralink105)). Although no cardiac findings were reported in the nonclinical studies with abiraterone acetate, cardiac disorders, in theory, could develop considering the primary pharmacology of abiraterone (selective inhibition of CYP17 that results in an inhibition of androgen biosynthesis). Based on this current knowledge and the risks of hypokalaemia and fluid retention/oedema events related to mineralocorticoid excess in subjects receiving abiraterone acetate, cardiac disorders are considered an Important Identified Risk for abiraterone acetate.

Cardiovascular disorders are the most prevalent comorbidity among patients with prostate cancer, regardless of treatment. Prevalence has been estimated as follows (whites vs blacks): coronary heart disease, 36.8% vs 35.4%; congestive heart failure, 12.7% vs 19.7%; benign hypertension, 46.4% vs 64.5%; valvular disease, 6.5% vs 5.0%; cardiac conduction disorders, 12.9% vs 9.0%; peripheral vascular disease, 12.8% vs 17.4% (Fleming et al, 2006). A recent analysis of the Swedish National Prostate Cancer Registry reported that prostate cancer patients treated with endocrine therapy are at a 24% increased risk of having a non-fatal myocardial infarction, a 19% increased risk of cardiac arrhythmias, a 31% increased risk of ischaemic heart disease, and a 26% increased risk of heart failure compared to the general population (Laino, 2009). Risk factors for cardiac disease include high blood cholesterol and other lipids, tobacco use, physical inactivity, overweight and obesity, diabetes mellitus, ESRD and chronic kidney disease, fasting plasma glucose ≥ 100 mg/dL or drug treatment for elevated glucose, waist circumference ≥ 102 cm in men, BP ≥ 130 mm Hg systolic or > 85 mm Hg diastolic or undergoing any treatment for hypertension, and nutritional factors (Lloyd-Jones et al, 2010). Additionally, trials report that patients undergoing ADT have a 20% increased risk of any cardiac morbidity compared with men who do not undergo ADT, with increasing duration of treatment significantly associated with increasing risks (Fitzpatrick, 2008).

Representatives from Global Medical Safety, Clinical, and Epidemiology are collaborating to address open questions in the prostate cancer patient population and identified cardiovascular outcomes as a disease area that would be useful to further evaluate using real world data. Specifically, the team seeks to understand whether patients exposed to Zytiga are at increased risk of cardiovascular morbidity compared to other therapies among patients with metastatic castrate resistant prostate cancer. At the most recent American Association for Cancer Research (AACR) annual conference, an abstract was presented in which cardiovascular outcomes among Zytiga users was presented (Lu-Yao 2019). The study identified patients who initiated Zytiga after being diagnosed with prostate cancer and presents unadjusted/crude incidence rate ratio estimates, comparing rates of hospitalization for various cardiovascular outcome in the period after vs. before Zytiga initiation. The study has multiple methodological shortcomings that bias findings (e.g. failing to use an active comparator, immortal time bias). The team therefore seeks to fully evaluate whether Zytiga poses additional cardiovascular risk among prostate cancer patients relative to other treatments in the metastatic castrate resistant patient population, after controlling for the biases that were present in the previous analysis.

# STUDY Objectives

## Primary Objectives

The primary objectives of the study are:

1. Among patients with castration-resistant prostate cancer (CRPC), compare new users of abiraterone acetate plus predniso(lo)ne to new users of enzalutamide with respect to hospitalization for the following cardiovascular morbidity outcomes
   * Cardiovascular morbidity (ischemic stroke, hemorrhagic stroke, heart failure, acute myocardial infarction or sudden cardiac death)
   * Ischemic stroke
   * Hemorrhagic stroke
   * Hospitalization because of heart failure (HF)
   * Acute myocardial infarction (AMI)
   * Sudden cardiac death

## Secondary Objectives

The secondary objectives are to evaluate all primary objectives listed above among patients of the clinically relevant subgroups listed below, if sample size allows:

1. Patients with no cardiovascular disease (CVD) related diagnoses or procedures in the 180 days before index
2. Patients with any cardiovascular disease (i.e. at least one CVD-related diagnoses or procedure in the 180 days before the index date)
3. Patients with less severe cardiovascular disease (i.e. at least one but less than three cardiovascular diagnoses or cardiovascular procedures in the 180 days before the index date)
4. Patients with more severe cardiovascular disease (i.e. at least three CVD diagnoses or at least three CVD procedures in the 180 days before the index date)

# Research METHODs

## Overview

The proposed study will include clinical characterization and population-level effect estimation (observational causal inference) using administrative claims databases which include nation-wide samples of patients insured in the United States (US). We plan to conduct a new user comparative cohort design to compare incidence of safety events among men with CRPC who are initiating treatment with either abiraterone or enzalutamide. Eligibility criteria for patients in the study target and comparator cohorts are defined in Section 6.3. The primary purpose of this study is to evaluate safety events among patients being treated for CRPC. The study period began after August 31, 2012 since before that date one of the study drugs (enzalutamide) was not FDA-approved for the treatment of CRPC.

The proposed study has already been fully specified, although results remain blinded and have not been inspected by any study investigators. Fully specifying the analysis in the protocol stage allows us to use empirical evaluation to assess the feasibility of the study design *a priori,* for its ability to generate valid comparative risk estimates and inform which results will be unblinded. This process of empirically inspecting study design diagnostics and determining which analyses will be unblinded is described in greater detail in Section 9.3. In that section, we describe the process we used to determine the three effect estimates/outcomes (among all effects we proposed to estimate) that will be unblinded and assessed in this study.

## Study Design and Setting

Internal administrative claims databases that will be utilized include: IBM MarketScan® Commercial Claims and Encounters (CCAE), IBM MarketScan® Medicare Supplemental (MDCR), Optum® De-Identified Clinformatics® Data Mart Database – Date of Death (DOD). All 3 databases have been standardized into the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), which includes a standard representation of health care experiences (such as information related to drug utilization and condition occurrence) as well as common vocabularies for coding clinical concepts and enables consistent application of analyses across multiple disparate data (OHDSI, 2020; Voss, 2015). Complete specifications for the extract, transform, and load (ETL) process for each database is available at: <https://github.com/OHDSI/ETL-CDMBuilder>. All analyses will be performed independently within each of these three databases to produce three source-specific results for each analysis. No patient-level data will be pooled across the databases for any analysis, in part to preserve internal validity of the comparative analyses within each database and avoid the potential risk of ‘double-counting’ cases for duplicate patients. Instead, meta-analysis estimates will be computed based on per-database aggregate statistics.

### IBM MarketScan® Commercial Claims and Encounters Database (CCAE) v1061

IBM MarketScan® Commercial Database (CCAE) represent data from individuals enrolled in United States employer-sponsored insurance health plans. The data includes adjudicated health insurance claims (e.g. inpatient, outpatient, and outpatient pharmacy) as well as enrollment data from large employers and health plans who provide private healthcare coverage to employees, their spouses, and dependents. Additionally, it captures laboratory tests for a subset of the covered lives. This administrative claims database includes a variety of fee-for-service, preferred provider organizations, and capitated health plans.

The major data elements contained within this database are outpatient pharmacy dispensing claims (coded with National Drug Codes (NDC), inpatient and outpatient medical claims which provide procedure codes (coded in CPT-4, HCPCs, ICD-9-CM or ICD-10-PCS) and diagnosis codes (coded in ICD-9-CM or ICD-10-CM). The data also contain selected laboratory test results (those sent to a contracted thirds-party laboratory service provider) for a non-random sample of the population (coded with LOINC codes). This database captures mortality that occurs during the course of an inpatient hospitalization. However, mortality outside of the inpatient setting is not noted and appears as disenrollment from the database. Additional information and some descriptive characteristics about the IBM MarketScan® CCAE database is available at the following link: <https://catalog.rwe.jnj.com/index#jnjsearches?dataSetUri=%2Fdataset%2Ff2f27aca-60f6-491d-8ed8-fff1858d178e.xml>

### IBM MarketScan® Medicare Supplemental Database (MDCR) v1062

IBM MarketScan® Medicare Supplemental Database (MDCR) represents health services of retirees in the United States with primary or Medicare supplemental coverage through privately insured fee-for-service, point-of-service, or capitated health plans. These data include adjudicated health insurance claims (e.g. inpatient, outpatient, and outpatient pharmacy). Additionally, it captures laboratory tests for a subset of the covered lives.

The major data elements contained within this database are outpatient pharmacy dispensing claims (coded with National Drug Codes (NDC), inpatient and outpatient medical claims which provide procedure codes (coded in CPT-4, HCPCs, ICD-9-CM or ICD-10-PCS) and diagnosis codes (coded in ICD-9-CM or ICD-10-CM). The data also contain selected laboratory test results (those sent to a contracted thirds-party laboratory service provider) for a non-random sample of the population (coded with LOINC codes). This database captures mortality that occurs during the course of an inpatient hospitalization. However, mortality outside of the inpatient setting is not noted and appears as disenrollment from the database. Additional information and some descriptive characteristics about the IBM MarketScan® MDCR database is available at the following link: <https://catalog.rwe.jnj.com/index#jnjsearches?dataSetUri=%2Fdataset%2F4d653f44-36f7-4ccb-a81d-a1dce3be1c62.xml>

### Optum® De-Identified Clinformatics® Data Mart Database - Date of Death (DOD) v1064

Optum® De-Identified Clinformatics® Data Mart Database (OptumInsight, Eden Prairie, MN) is an adjudicated US administrative health claims database for members of private health insurance, who are fully insured in commercial plans or in administrative services only (ASOs), Legacy Medicare Choice Lives (prior to January 2006), and Medicare Advantage (Medicare Advantage Prescription Drug coverage starting January 2006).  The population is primarily representative of commercial claims patients (0-65 years old) with some Medicare (65+ years old) however ages are capped at 90 years.  It includes data captured from administrative claims processed from inpatient and outpatient medical services and prescriptions as dispensed, as well as results for outpatient lab tests processed by large national lab vendors who participate in data exchange with Optum.  This dataset also provides date of death (month and year only) for members with both medical and pharmacy coverage from the Social Security Death Master File and location information for patients is at the US state level. It is important to note that after 2011, reporting frequency for death changed due to changes in national reporting requirements; however, patients were not included in this analysis until 2012, so we do not expect meaningful changes in classification of death outcomes over the course of the study. Family identifiers are provided and utilized to infer mother-child linkages.

The major data elements contained within this database are outpatient pharmacy dispensing claims (coded with National Drug Codes (NDC), inpatient and outpatient medical claims which provide procedure codes (coded in CPT-4, HCPCs, ICD-9-CM or ICD-10-PCS) and diagnosis codes (coded in ICD-9-CM or ICD-10-CM).  The data also contain selected laboratory test results (those sent to a contracted thirds-party laboratory service provider) for a non-random sample of the population (coded with LOINC codes). Additional information and some descriptive characteristics about the Optum® De-Identified Clinformatics® Data Mart Database is available at the following link: <https://catalog.rwe.jnj.com/index#jnjsearches?dataSetUri=%2Fdataset%2F5115ce8f-54d7-46f8-bef5-7fba67943b75.xml>

## Study Population(s)

In the primary analyses, we select patients from the aforementioned databases at the point of their observed drug exposure to either abiraterone (the target exposure group) or enzalutamide (the comparator exposure group). We include patients whose index date is preceded by at least one condition occurrence of primary malignant neoplasm of the prostate at any point and who have observed use of ADT therapy 30 days before the index date or earlier. Since abiraterone is indicated to be prescribed concomitantly with oral prednisone or prednisolone, we also require that patients in the abiraterone cohort had a fillof oral predniso(lo)ne within 30 days (either before or after) the index date. Patients were only included in the analysis after August 31, 2012, which is the date that enzalutamide was approved for treatment of CRPC (Pazdur, 2012). Thus, this study only includes time where the two treatments being studied (abiraterone and enzalutamide) were indicated for patients with similar clinical conditions and disease stage.

In additional sub-analyses, we further stratify these cohorts according to the following criteria if sample size allows:

* No cardiovascular disease in prior six months: we require that in the 180 days before the index date there be no occurrences of conditions or procedures indicating cardiovascular disease.
* Any cardiovascular disease in prior six months: we require that in the 180 days before the index date there be at least one occurrence of a condition or procedure indicating cardiovascular disease.
* Less severe cardiovascular disease in prior six months: we require that in the 180 days before the index date there be 1) at least one occurrence of a condition or procedure indicating cardiovascular disease, 2) no more than two conditions indicating cardiovascular disease and 3) no more than two procedures indicating cardiovascular disease.
* More severe cardiovascular disease: we require that in the 180 days before the index date there be at least three occurrences of condition or at least three occurrences of procedures indicating cardiovascular disease in the 180 days before the index date

We use the web-based Atlas tool (a point-and-click interface for specifying study design in analyses of large observational databases) to develop target cohorts (initiators of abiraterone) and comparator cohorts (initiators of enzalutamide) for each of the above-described analyses. In Table 1 below, we provide a general overview of the logic used to define the target (abiraterone) and comparator (enzalutamide) cohorts. In Table 2, all cohorts (including outcome cohorts) are listed with hyperlinks that lead to their full specification in the Atlas web-interface. Details on the concept sets used to define these cohorts are available in Atlas.

**Table 1.** General overview of logic used to define target (abiraterone) and comparator (enzalutamide) cohorts.

|  |  |
| --- | --- |
| **Cohort Component** | **Description of Specification** |
| **Concept Sets** | * [RWE CTF] Abiraterone (Zytiga)   + <https://epi.jnj.com/atlas/#/conceptset/3462/conceptset-expression> * Enzalutamide   + <https://epi.jnj.com/atlas/#/conceptset/2806/conceptset-expression> * [EPI\_693] Oral predniso(lo)ne   + <https://epi.jnj.com/atlas/#/conceptset/9617/conceptset-expression> * Primary malignant neoplasm of prostate   + <https://epi.jnj.com/atlas/#/conceptset/9071/conceptset-expression> * CPRC Definition Androgen Deprivation Therapy   + <https://epi.jnj.com/atlas/#/conceptset/10186/conceptset-expression> * [EPI\_720] Conditions indicating established CVD (strata)   + <https://epi.jnj.com/atlas/#/conceptset/9762/conceptset-expression> * [EPI\_720] Procedures indicating established CVD (strata)   + <https://epi.jnj.com/atlas/#/conceptset/9764/conceptset-expression> |
| **Initial Event Criteria** | People with a drug era of *[RWE CTF] Abiraterone (Zytiga) or Enzalutamide* for the first time in the person's history who have   * continuous observation of at least 365 days prior and 0 days after event index date * limit initial events to: **earliest event per person**. |
| **Initial Event Inclusion Criteria or Additional Qualifying Inclusion Criteria** | For people matching the Primary Events, include patients who meet all of the criteria below:   * at least 1 occurrences of a condition occurrence of *Primary malignant neoplasm of prostate* where event starts between all days Before and 0 days Before index start date * Having any of the following criteria:   + at least 1 occurrences of a drug exposure of *CPRC Definition Androgen Deprivation Therapy* where event starts between all days Before and 0 days Before index start date   + at least 1 occurrences of a procedure of *CPRC Definition Androgen Deprivation Therapy* where event starts between all days Before and 0 days Before index start date * at least 1 occurrences of a drug exposure of *[EPI\_693] Oral predniso(lo)ne* where event starts between 30 days Before and 30 days After index start date   + This requirement only applies to the abiraterone cohorts and is not applied for the enzalutamide cohorts   Limit initial events to: **earliest event per person** |
| **Inclusion Rules** | Cohort: general  In the un-stratified cohort, no additional inclusion rules are applied.  Cohort: no CVD in prior 6-mo  Having ALL of the following criteria:   * exactly 0 occurrences of a condition occurrence of *[EPI\_720] Conditions indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date * exactly 0 occurrences of a procedure of *[EPI\_720] Procedures indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date   Limit qualifying cohort to: **earliest event per person.**  Cohort: any CVD in prior 6-mo  Having ANY of the following criteria:   * at least 1 occurrence of a condition occurrence of *[EPI\_720] Conditions indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date * at least 1 occurrence of a procedure of *[EPI\_720] Procedures indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date   Limit qualifying cohort to: **earliest event per person.**  Cohort: less severe CVD in prior 6-mo  Having ANY of the following criteria:   * at least 1 and at most 2 occurrences of a condition occurrence of *[EPI\_720] Conditions indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date * at least 1 and at most 2 occurrences of a procedure of *[EPI\_720] Procedures indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date   Limit qualifying cohort to: **earliest event per person.**  Cohort: more severe CVD in prior 6-mo  Having ANY of the following criteria:   * at least 3 occurrences of a condition occurrence of *[EPI\_720] Conditions indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date * at least 3 occurrences of a procedure of *[EPI\_720] Procedures indicating established CVD (strata)* where event starts between 180 days Before and 0 days Before index start date   Limit qualifying cohort to: **earliest event per person.** |
| **Event Persistence Criteria** | Custom Drug Era Exit Criteria   * This strategy creates a drug era from the codes found in the specified concept set. If the index event is found within an era, the cohort end date will use the era's end date. Otherwise, it will use the observation period end date that contains the index event. * Use the era end date of *[RWE CTF] Abiraterone (Zytiga) or Enzalutamide*   + allowing 30 days between exposures   + adding 0 days after exposure end   + using days supply and exposure end date for exposure duration. |
| **Censoring Events** | * No censoring events |
| **Cohort Eras** | * Collapse cohort by era with a gap size of 0 days. |

**Table 2**. 17 Atlas cohorts (six target cohorts, seven outcome cohorts)

|  |  |
| --- | --- |
| **Cohort ID** | **Cohort Title** |
| *Target cohorts* |  |
| [13186](https://epi.jnj.com/atlas/#/cohortdefinition/13186) | [EPI\_720 - T01] New Users of abiraterone + prednisone with CRPC |
| [13248](https://epi.jnj.com/atlas/#/cohortdefinition/13248) | [EPI\_720 - T01] New Users of abiraterone + prednisone with CRPC (no CVD in prior 6-mo) |
| [13185](https://epi.jnj.com/atlas/#/cohortdefinition/13185) | [EPI\_720 - T01] New Users of abiraterone + prednisone with CRPC (any CVD in prior 6-mo) |
| [13187](https://epi.jnj.com/atlas/#/cohortdefinition/13187) | [EPI\_720 - T01] New Users of abiraterone + prednisone with CRPC (less severe CVD in prior 6-mo) |
| [13191](https://epi.jnj.com/atlas/#/cohortdefinition/13191) | [EPI\_720 - T01] New Users of abiraterone + prednisone with CRPC (more severe CVD in prior 6-mo) |
| *Comparator cohorts* |  |
| [13245](https://epi.jnj.com/atlas/#/cohortdefinition/13245) | [EPI\_720 - C01] New users of enzalutamide with CRPC |
| [13247](https://epi.jnj.com/atlas/#/cohortdefinition/13247) | [EPI\_720 - C01] New users of enzalutamide with CRPC (no CVD in prior 6-mo) |
| [13246](https://epi.jnj.com/atlas/#/cohortdefinition/13246) | [EPI\_720 - C01] New users of enzalutamide with CRPC (any CVD in prior 6-mo) |
| [13316](https://epi.jnj.com/atlas/#/cohortdefinition/13316) | [EPI\_720 - C01] New users of enzalutamide with CRPC (less severe CVD in prior 6-mo) |
| [13319](https://epi.jnj.com/atlas/#/cohortdefinition/13319) | [EPI\_720 - C01] New users of enzalutamide with CRPC (more severe CVD in prior 6-mo) |
| *Outcome cohorts* |  |
| [13320](https://epi.jnj.com/atlas/#/cohortdefinition/13320) | [EPI\_720 - O01] Cardiovascular morbidity (hospitalization for ischemic stroke, hemorrhagic stroke, heart failure, acute myocardial infarction or sudden cardiac death) |
| [13321](https://epi.jnj.com/atlas/#/cohortdefinition/13321) | [EPI\_720 - O02] Cardiovascular morbidity - ischemic stroke hospitalization |
| [13323](https://epi.jnj.com/atlas/#/cohortdefinition/13323) | [EPI\_720 - O03] Cardiovascular morbidity - hemorrhagic stroke hospitalization |
| [13324](https://epi.jnj.com/atlas/#/cohortdefinition/13324) | [EPI\_720 - O04] Cardiovascular morbidity - heart failure hospitalization |
| [13325](https://epi.jnj.com/atlas/#/cohortdefinition/13325) | [EPI\_720 - O05] Cardiovascular morbidity - acute myocardial infarction (AMI) hospitalization |
| [13326](https://epi.jnj.com/atlas/#/cohortdefinition/13326) | [EPI\_720 - O06] Cardiovascular mortality - sudden cardiac death (in-hospital) |

\* Text descriptions of each cohort and code lists for each concept set can be found using the provided links then selecting the “export” tab and the “Text View” sub-menu. These text descriptions are also available in Annex Document 3 (see Section 15 of this document).

## Outcomes, Follow-Up, and Time-at-Risk

Following completion of a feasibility analysis in which we will assess diagnostics to see whether comparisons between Zytiga and Xtandi are possible, we will compare incidence of various outcomes including ischemic stroke, hemorrhagic stroke, heart failure, acute myocardial infarction, sudden cardiac death, and a composite of all of the aforementioned cardiovascular outcomes. The codes used to define these outcomes were developed during a prior research effort which assessed safety outcomes across various comparisons of first-line antihypertensive medications (Suchard 2019). Follow-up for outcomes begins 30 days after the index date. This is intended to reduce the possibility of immortal time bias since, to enter the cohort, abiraterone users were required to have observed predniso(lo)ne use within 30 days of their index date and no comparable requirement is applied to the enzalutamide group. Additionally, we require that patients have at least one day at-risk during follow-up; in other words, we exclude any patients who exited the cohort before day 31. This restriction is intended to prevent the analysis from balancing potentially confounding variables in a population that includes patients who are never eligible to experience the outcome events being compared. We are aware that this design feature limits our ability to study events that occur within 30 days of treatment initiation. However, we believe that it is necessary to reduce the likelihood of immortal time bias and, furthermore, the short-term safety of these treatments have already been characterized by randomized clinical trials.

In the intent-to-treat analysis, follow-up time is only censored if the patient experiences the outcome, dies, or disenrolls from the database. For the as-treated analysis, we also censor follow-up time when patients discontinue use of the study drugs (i.e. either abiraterone or enzalutamide), both of which are daily oral medications. To do so, we use the available prescribing data that accompanies pharmacy claims recorded in administrative claims data (pill count, days supply, number of refills). We define patients as having discontinued use when greater than 30 days is observed where the patient would be expected to have no supply of the study drug. For example, a patient whose last observed abiraterone prescription has a days supply of 20 would be censored on day 50. It is important to note that while abiraterone is indicated for concomitant use with predniso(lo)ne, we do not censor patients if we failed to observe continuous use of predniso(lo)ne. Since cheap generic forms of predniso(lo)ne are available, it is likely that patients could continue to use it even if we do not observe it in our data (e.g. if they paid for their predniso(lo)ne prescription in cash).

Failure to censor follow-up in the case of treatment switching can bias results towards the null which would be particularly problematic in this study since bias towards the null has the potential to limit our ability to recognize important safety events. In all analyses, patients are not allowed to have overlapping follow-up time in the two cohorts being compared. We conduct three analyses that vary analytic choices about how to treat patients who appear in both exposure cohorts. To be clear, the analyses varying these analytic choices are distinct from the analyses that vary censoring of follow-up based on as-treated vs. intent-to-treat designs. In the first analysis, we keep only the first observed exposure such that patients who entered one exposure cohort are never eligible to enter the other exposure cohort. In a separate analysis, we allow patients to appear in both exposure cohorts, but censor follow-up time if at any point the patient enters the other exposure cohort. This approach is only applied for as-treated analyses since, in the intent-to-treat analysis, we follow-up until the end of available database time and allowing patients to appear in both exposure groups allows patients to have overlapping follow-up time in the two exposure cohorts. In another analysis, patients who appear in both exposure cohorts are removed from the analysis entirely. It is important to emphasize that with this design, cohort entry may be conditioned on events that take place in the future (i.e. whether or not they later went on to initiate the alternate exposure). Conditioning cohort entry on events that occur during follow-up can produce bias. As an illustrative example, it could be the case that such an exclusion increases observed mortality rate. This is because patients who do not die are more likely to be included because they must live long enough to initiate the alternate exposure.

## Negative and Positive Control Outcomes

Negative control outcomes, i.e., outcomes known not to be causally associated with any of the exposure cohorts [2], will be used for empirical calibration [3, 4]. Conditions must meet the following requirements to be considered as negative controls [5]: (1) have no known drug-condition association in published literature with any study exposure (2) not listed in the “Adverse Drug Reactions” or “Post-marketing” section in the US product label, (3) not considered a FAERS signal [4], (4) have no indication or contraindication listed in the OMOP Vocabulary for the drug-condition pair, (5) are not general concepts (e.g. finding of trunk), (6) are not considered a drug induced concept, or (7) not considered a pregnancy related concept.

The same analysis that is performed for each pairwise comparison to assess the risk of SUB (see Sections 9.3, 9.4, 9.5) is also performed to assess the risk of each negative control outcome. Because the negative control qualifying criteria support the *a priori* assertion of no effect, we assume the true relative risk (RR) for each negative control outcome is 1, and the difference between RR=1 and the observed effect estimate is considered error, encompassing both random and potentially systematic.

In addition to negative control outcomes, we also generate and include synthetic positive control outcomes. Positive control outcomes are developed on the basis the observed negative controls, but where the true effect size is artificially increased to a desired effect size by injection of additional, simulated outcomes [3]. To preserve confounding, these additional outcomes are sampled from predicted probabilities generated using a predictive model fit to identify patients similar to those with observed negative control outcomes during the study time-at-risk. Synthetic positive control outcomes are differentially inserted among target patients to artificially create positive control outcomes with true relative risks of 1.5, 2, and 4. Using both negative and positive controls, we fit a systematic error model and with which we calibrated the hazard ratio, confidence interval, and p-value to account for residual error not addressed by other methods for confounding control (i.e. propensity score adjustment) [3]. The 124 negative control outcomes used in this study are listed in the Table 3 below.

**Table 3.** 124 conditions used as negative control outcomes for the empirical calibration

|  |  |  |
| --- | --- | --- |
| Abnormal posture | Generalized hyperhidrosis | Onychomycosis due to dermatophyte |
| Acquired renal cystic disease | Hemangioma of skin and subcutaneous tissue | Open-angle glaucoma - borderline |
| Acquired trigger finger | Hemorrhage of rectum and anus | Osteoarthritis of hip |
| Actinic keratosis | Herpes zoster without complication | Osteoarthritis of knee |
| Acute bronchitis | Hypercalcemia | Paresthesia |
| Acute cystitis | Hyperparathyroidism due to renal insufficiency | Parkinson's disease |
| Acute pharyngitis | Hyperpigmentation of skin | Plantar fascial fibromatosis |
| Acute respiratory failure | Hypertrophy of breast | Polyneuropathy due to drug |
| Acute sinusitis | Impacted cerumen | Polyp of colon |
| Acute stress disorder | Inflamed seborrheic keratosis | Post-inflammatory pulmonary fibrosis |
| Acute upper respiratory infection | Ingrowing nail | Presbyopia |
| Adjustment disorder with depressed mood | Intestinovesical fistula | Pressure ulcer of buttock |
| Adjustment disorder with mixed emotional features | Kidney stone | Primary localized osteoarthrosis of pelvic region |
| Alcohol abuse | Leukocytosis | Primary open angle glaucoma |
| Allergic rhinitis | Leukopenia | Pulmonary emphysema |
| Altered sensation of skin | Localized enlarged lymph nodes | Pyelonephritis |
| Anesthesia of skin | Localized osteoarthrosis uncertain if primary OR secondary | Recurrent major depressive episodes, moderate |
| Ascites | Localized swelling, mass and lump, lower limb | Renal disorder due to type 2 diabetes mellitus |
| Atelectasis | Localized, primary osteoarthritis of the shoulder region | Retention of urine |
| Bacteremia | Loss of consciousness | Scar conditions and fibrosis of skin |
| Benign hypertensive renal disease with renal failure | Lumbar spondylosis | Sciatica |
| Candidiasis of mouth | Lumbar sprain | Segmental and somatic dysfunction |
| Cervical disc disorder with radiculopathy | Lumbosacral radiculopathy | Senile hyperkeratosis |
| Cervical radiculopathy | Lumbosacral spondylosis without myelopathy | Sleep apnea |
| Cervical spondylosis without myelopathy | Lumbosacral stenosis | Somatic dysfunction of sacral region |
| Chronic kidney disease due to type 2 diabetes mellitus | Malnutrition | Somatic dysfunction of thoracic region |
| Chronic nonalcoholic liver disease | Melanocytic nevus of trunk | Spasm of back muscles |
| Chronic obstructive pulmonary disease with acute lower respiratory infection | Merkel cell carcinoma | Spinal stenosis in cervical region |
| Chronic pain syndrome | Microscopic hematuria | Spinal stenosis of lumbar region |
| Chronic sinusitis | Migraine | Tear film insufficiency |
| Combined form of senile cataract | Mixed hyperlipidemia | Thoracic spondylosis without myelopathy |
| Dependence on supplemental oxygen | Moderate major depression, single episode | Thyrotoxicosis without goiter OR other cause |
| Diabetic polyneuropathy | Moderate recurrent major depression | Transient cerebral ischemia |
| Diaphragmatic hernia | Mononeuritis | Type 2 diabetes mellitus |
| Displacement of cervical intervertebral disc | Morbid obesity | Ulcer of foot |
| Diverticulosis of large intestine without diverticulitis | Multiple sclerosis | Uncomplicated asthma |
| Epiretinal membrane | Myopia | Uncomplicated inguinal hernia |
| Gallstone | Neck sprain | Ureteric stone |
| Gastroesophageal reflux disease without esophagitis | Nicotine dependence | Verruca vulgaris |
| Gastroparesis syndrome | Non-toxic multinodular goiter | Vitreous opacities |
| Generalized anxiety disorder | Nuclear senile cataract |  |
| Generalized enlarged lymph nodes | Ocular hypertension |  |

# SAMPLE SIZE AND STUDY POWER

Since the sample size is fixed in retrospective studies (the data has already been collected), and the true effect size is unknown, we do not apply the widely-used approach of calculating sample sizes necessary to observe effects of various hypothetical strengths. Instead, we specify the analysis a priori and then calculate, for each separate analysis, the minimum detectable relative risk (MDRR) given the known sample size for that analysis. The calculation includes a targeted type I error rate (alpha) of 0.05 (2-sided) and a type II error rate (beta) of 0.20 (power=80%) and reports the minimum hazard ratio detectable given the final target and comparator patient count, outcome event count, and time-at-risk (Schoenfeld, 1983).

The MDRR provides a range of effect estimates that we can reasonably expect to produce a significant finding, given how much sample we have available in our study. For example, a minimally detectable risk ratio of 1.50 would indicate that we are powered to observe a statistically significant finding (based on alpha=0.05) when the true risk ratio is 1.50 or greater. On the other hand, a minimally detectable risk ratio of 4.00 would indicate that we are only powered to observe a statistically significant finding (based on alpha=-0.05) if the true risk ratio is 4.0 or greater. Risk ratios of this magnitude are rare, especially for safety outcomes like the ones being considered; therefore such a study would be considered to be underpowered.

Due to the low incidence of CRPC among younger people, the study-eligible sample in the IBM Commercial Claims and Encounters (CCAE) database is too low to proceed with analyses, and so this database will not be used in the analysis. In Table 4 and Table 5, we present the counts of patients in each exposure group that are included in the propensity-score matched (1:1 and 1:100 variable ratio, respectively) sample for various analyses in the Optum Extended DOD database and the IBM Medicare Supplemental database. These samples correspond to the analyses of the composite cardiovascular morbidity outcome (ischemic stroke, hemorrhagic stroke, heart failure, acute myocardial infarction, or sudden cardiac death), where the study population excludes anyone with prior history of the outcome. The corresponding matched samples included in analyses of other outcomes are larger than the ones shown below.

**Table 4.** Count of patients in 1:1 propensity-score matched samples in the analysis of the composite cardiovascular morbidity outcome

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Optum Extended DOD | | IBM Medicare Supplemental | |
|  | Abiraterone + predniso(lo)ne | Enzalutamide | Abiraterone + predniso(lo)ne | Enzalutamide |
| As-treated analyses |  |  |  |  |
| Keep first | 1,619 | 1,619 | 1,002 | 1,002 |
| Remove all | 1,242 | 1,242 | 734 | 734 |
| Keep all | 2,179 | 2,179 | 1,389 | 1,389 |
| Intent-to-treat analyses |  |  |  |  |
| Keep first | 1,813 | 1,813 | 1,117 | 1,117 |
| Remove all | 1,413 | 1,413 | 864 | 864 |

**Table 5.** Count of patients in 1:100 variable-ratio propensity-score matched samples in the analysis of the composite cardiovascular morbidity outcome

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Optum Extended DOD | | IBM Medicare Supplemental | |
|  | Abiraterone + predniso(lo)ne | Enzalutamide | Abiraterone + predniso(lo)ne | Enzalutamide |
| As-treated analyses |  |  |  |  |
| Keep first | 1,619 | 1,849 | 1,002 | 1,343 |
| Remove all | 1,242 | 1,482 | 734 | 1,110 |
| Keep all | 2,179 | 2,668 | 1,389 | 1,974 |
| Intent-to-treat analyses |  |  |  |  |
| Keep first | 1,813 | 2,141 | 1,117 | 1,543 |
| Remove all | 1,413 | 1,731 | 864 | 1,299 |

# DATA ANALYSIS PLAN

## Patient Characteristics Summary

***Descriptive characterizations***

The demographics of patients within each treatment cohort (including stratified cohorts) will be described as well as other conditions, drugs, and procedures used in the 30 days, 180 days, 360 days, and all time preceding the index date. The full set of characteristics that will be reported can be found in the population characteristics tab of the Shiny app (link available in Annex section - Document 5). Continuous variables will be presented as mean (SD) and median (IQR), while categorical data will be presented as n (%).

***Incidence rates***

For calculation of incidence rates, the number of persons with each event, the incidence proportion, and the incidence rate adjusted for person-time according to each at-risk time window for each study population and each of the outcomes of interest will be reported:

* Cardiovascular morbidity and mortality
  + Stratified by treatment cohorts
  + Stratified by baseline history of CVD present or absent if sample size allows
    - Among those with history of CVD, further stratified by less evere or more severe CVD

We will estimate incidence rates in the crude/unmatched study population as well as the matched study population. Patients will be excluded from the analysis if they have a history of the outcome at any point in the database before the index date. So, for the analysis assessing the outcome of ischemic stroke hospitalization, any patients who were hospitalized for ischemic stroke before their index date were excluded from the analysis. For the calculation of incidence rates for all time-at-risk definitions, patients will be followed from 30 days past their index date through death, end of enrollment, or occurrence of the outcome of interest, whichever comes first. For incidence rates within the exposure-specific time window, person-time will be captured as cumulative duration of exposure based on drug eras that have been previously defined via the OMOP Common Data Model (CDM).

## Model Specification

For the purpose of contextualizing the event rates and quantifying relative risk while controlling for additional confounding factors, which could potentially influence unadjusted incidence rates, a new user cohort design will be used to conduct comparative analyses if the exposed (abiraterone) population can be appropriately matched to the selected comparator population (enzalutamide) based on a defined set of patient and clinical characteristics, using propensity score matching. Propensity scores will be used as an analytic strategy to reduce potential confounding due to imbalance between the target and comparator cohorts in baseline covariates. The propensity score is the probability of a patient being classified in the target cohort vs. the comparator cohort, given a set of observed covariates. The propensity score will be estimated for each patient, using the predicted probability from a regularized logistic regression model, fit with a Laplace prior (LASSO) and the regularization hyperparameter selected by optimizing the likelihood in a 10-fold cross validation, using a starting variance of 0.01 and a tolerance of 2e-7. Covariates that occur in fewer than 0.1% of the combined target and comparator cohort in a pairwise comparison will be excluded prior to model fitting. Patients in the target cohort will be matched to patients in the comparator cohort, using a ratio of either 1:1 matching or 1:100 variable ratio matching, using a greedy matching algorithm with a caliper of 0.25 of the standard deviation of the propensity score distribution. Comparative analyses will only be run in the case that feasibility assessments show appropriate comparable clinical cohorts can be identified based on diagnostics (see Evidence Evaluation below). Using this approach to matching, the hazard ratios we produce will estimate the average effect in the population treated with abiraterone (as opposed to the average effect in the entire study population which may differ if there is effect heterogeneity by treatment status).

Cox proportional hazards will be used to estimate the hazards of each outcome in the target cohort, relative to the comparator cohort. The final outcome model will be summarized by providing the hazards ratio and associated 95% confidence interval. The number of persons, amount of time-at-risk, and number of outcomes in each cohort will also be reported. Additionally, a Kaplan-Meier plot will be generated to characterize the contour of risk over time for the outcome of interest.

## Evidence Evaluation

Covariate balance will be summarized in tabular form by showing the mean value for all baseline covariates in the target and comparator cohort, with the associated standardized mean difference computed for each covariate. Once the propensity score model is fit, we will plot the propensity score distribution of the target and comparator cohorts to evaluate the comparability of the two cohorts. The plot will be scaled to the preference score, normalizing for any imbalance in cohort size. The area under the Receiver Operating Characteristic (ROC) curve (AUC) will be reported. An attrition diagram will be provided to detail the loss of patients from the original target cohort, T, and comparator cohort C to the subpopulations that remain after all design considerations have been applied.

As described in Section 7, the distribution of estimates on negative control outcomes (i.e. the empirical null distribution) describes the residual error of a study specification after confounding control has been implemented. Negative control outcomes will be used to evaluate the potential impact of residual systematic error in the study design, and to facilitate empirical calibration of the p-value for the outcomes of interest. Negative control outcomes are concepts known not to be associated with either the target or comparator group, such that we can assume the true relative risk should equal 1. For each negative control outcome, the study design described above will be implemented and the effect estimate will be recorded. We will also generate and include synthetic positive control outcomes. Positive control outcomes developed on the basis the observed negative controls, but where the true effect size is artificially increased to a desired effect size by injection of additional, simulated outcomes [3]. Using both negative and positive controls, we fit a systematic error model and with which we calibrated the hazard ratio, confidence interval, and p-value to account for residual error not addressed by other methods for confounding control (i.e. propensity score adjustment) [3].The empirical null distribution will then be applied to the target outcome of interest to calibrate the p-value. Empirical calibration serves as an important diagnostic tool to evaluate if the residual systematic error is sufficient to cast doubt on the accuracy of the unknown effect estimate. The calibration effect plot and calibration probability plots will be generated for review. We will report the traditional p-value and empirically calibrated p-value for each negative control, as well as each outcome of interest.

The estimation study described in the protocol above has already been fully specified using the Atlas web-interface. The following link leads to that specification, which can be reviewed in detail: <https://epi.jnj.com/atlas/#/estimation/cca/91>

We have executed and reviewed the study diagnostics (described above) in detail and used them to make decisions on which analyses should be unblinded once the protocol is approved. The diagnostics for each analysis are presented in an interactive, web-based Shiny app which can be accessed using the following link: <https://sharedshiny.jnj.com/user/mconove1/EPI720_3/>

In collaboration with colleagues within the Observational Health Data Analytics (OHDA) department, we specified a rule for determining when to not move forward with an analysis on the basis of covariate imbalance between the two groups we would like to compare. This data-driven rule is applied a priori (before the adjusted effect estimates are unblinded to study investigators) in order to determine which analyses we expect to be sufficiently unbiased to provide an informative effect estimate. We only plan to proceed with analyses where all covariates that satisfy one of the following conditions: [1] an absolute standardized mean difference < 0.10 (after adjustment), [2] an absolute difference in (un-standardized) prevalence < 0.05 (after adjustment). Rule [1] is widely used in observational research conducted using these data (Austin, 2009). However, we added the second rule since the covariate is unlikely to confound the effect as long as the actual difference in prevalence isn’t extreme (i.e. greater than 0.05). As an additional protection against unblinding analyses where results may be biased, we also had five study team member manually review all covariates that failed criteria [1] but passed criteria [2] and flagged all covariates that at least one reviewer considered to be a plausible risk-factor for any of the cardiovascular outcomes being studied. Doing so, we reviewed 473 covariates and found that 419 may conservatively be considered to be confounders. Using this list, we also choose to not unblind any analysis where one of these 419 covariates was imbalanced with an absolute standardized mean difference ≥ 0.10, even if it’s absolute difference in prevalence was < 0.05.

Initially, based on the objectives of this protocol across 3 databases with multiple subgroup analyses, there were 600 analyses included as part of this feasibility assessment. This includes the combinations of 5 target/comparator combinations, 6 outcomes, two databases, and 10 analysis variations. Of these 600 analyses, 9 analyses met all of the criteria for unblinding, all of which were analyses in the Optum Extended DOD database comparing the primary/unstratified target and comparator cohorts (new users of abiraterone + predniso(lo)ne versus new users of enzalutamide among patients with CRPC). Furthermore, the 9 analyses recommended for unblinding included only the as-treated analyses, specifically those where all patient observations were kept, regardless of whether an individual patient appears in both the target and comparator group (due to switching). The specific outcomes and propensity-score matching settings for the 9 analyses recommended for unblinding are displayed in Table 6 below. The minimally-detectable risk ratios for these analyses ranged between 1.50 and infinite. It is likely implausible that the real effect exceeds a risk ratio of 2. If this is indeed the case, all analyses with minimally-detectable risk ratios >2 will produce a null finding, even if a real effect with risk ratio < 2 exists (type II error). Thus, we applied an additional rule blinding any analysis where the minimally detectable risk ratio was > 2. Finally, four analyses are recommended for unblinding. In other words, we are confident that we are able to control for the bias that exists in observational data for these four outcomes: myocardial infarction (both with 1:1 matching and 1:100 variable matching), heart failure, and ischemic stroke. Given we do not expect the matching ratio to influence the extent or direction of the effect estimate, for ease of interpretation we will move forward with one analysis for each outcome that met our criteria using only 1:100 variable matching.

**Table 6.** Outcomes, propensity-score matching settings, and minimally-detectable risk ratios (MDRRs) for the 9 analyses that met the covariate-based criteria for unblinding

|  |  |  |  |
| --- | --- | --- | --- |
| Final blinding recommendation | Outcome | Propensity-score matching | MDRR |
| Unblinded | Acute myocardial infarction (AMI) | 1:100 variable ratio | 1.78 |
| Acute myocardial infarction (AMI) | 1:100 variable ratio | 1.99 |
| Heart failure | 1:100 variable ratio | 1.50 |
| Blinded | Ischemic stroke | 1:1 matching\* | 1.82 |
| Ischemic stroke | 1:1 matching | 2.03 |
| Hemorrhagic stroke | 1:100 variable ratio | ∞ |
| Hemorrhagic stroke | 1:1 matching | ∞ |
| Sudden cardiac death | 1:100 variable ratio | 2.41 |
| Sudden cardiac death | 1:1 matching | 2.48 |

\* Only one analysis that employed a 1:1 propensity-score matching ratio was recommended for unblinding and its corresponding 1:100 variable ratio matching analysis was also recommended for unblinding. Thus we chose not to unblind the 1:1 matched analysis for ischemic stroke.

# STRengths and limitations of the research methods

## Strengths

* Use of large observational datasets that include a nation-wide sample of insured US adults and provide complementary perspectives about treatment utilization and effects within the US population. The outcomes we propose to study are rare and it is unlikely that experimental or randomized research would have sufficient statistical power to meaningfully study them.
* Cohort studies allow direct estimation of incidence rates following exposure of interest, and the new-user design can capture early events following initiation of treatment. By aligning the start of follow-up at the point of treatment initiation for both the target exposure and the comparator exposure, the new user design prevents bias that can result from comparing populations that differ with respect to the time they’ve been on the treatment. Anchoring the index date on treatment initiation also clarifies the temporal distinction between potentially confounding baseline variables (e.g. comorbid conditions) which precede treatment and variables that are the result of treatment (which may be mediators of the causal effects being studied and should not be included in adjustment.
* Propensity score adjustment allows balancing on many baseline potential confounders. In this study, the outcomes of interest are rare and controlling for confounding by modeling the relationship between covariates and outcomes would result in a substantial loss in statistical power. Propensity score methods are a better analytical choice since they allow for confounding to be controlled by modeling the relationships between covariates and exposures.
* The use of a set of negative control outcomes allows for estimation of residual bias inherent to the study design and data. By using the distribution of error observed in the estimates produced by analyses of the 192 negative control analyses, we seek to correct (or recalibrate) effect estimates and confidence intervals, taking into account random and non-random sources of bias.

## Limitations

* The study will be subject to the limitation that some confounders may be unmeasured or inadequately represented in claims data, including weight, smoking status, many clinical measurements, frailty, and lifestyle behaviors, such as diet and exercise. Adjustment by propensity score may not completely remove bias due to unmeasured or mis-specified confounders. However, this study compares two treatments (abiraterone and enzalutamide) which are prescribed to patients with similar clinical conditions and disease stage. Using an active comparator reduces the likelihood that these potential confounders are imbalanced and bias estimates (Schneeweiss et al., 2007).
* Informative right-censoring over time may mean differences in outcome observed later in follow-up are actually the result of imbalance between the patients in each exposure group who remain in the analysis. As implemented in this new-user study design, propensity score matching produces a study population which is balanced at baseline on various confounders that preceded the index date. However, informative censoring of follow-up time is possible, meaning confounders may become imbalanced over the course of the study due to patients being lost to follow-up. While analytic strategies are available to address this problem (e.g. inverse-probability of censoring weights) we have not employed them here.
* In any observational study of secondary-health databases, misclassification of the outcome is an important concern. Non-differential outcome misclassification (i.e. misclassification that does not differ between the two exposure groups being compared) will bias results towards the null while differential outcome misclassification could bias estimates in any direction. In this instance, we study serious acute medical events which are likely to be recorded in administrative claims data and electronic medical record. However, sensitivity and specificity of outcome ascertainment in database analyses is rarely perfect. In these analyses, we exclude any patients who had the outcome of interest before exposure initiation. However, it is possible that outcomes observed during follow-up (e.g. heart failure) were actually present before exposure was initiated and just not observed in the available pre-exposure data.
* We know a priori that we substantially under-ascertain the outcome of sudden cardiac death. Both the IBM databases (CCAE and MDCR) only capture mortality that occurs during the course of an inpatient hospitalization. Mortality outside of the inpatient setting is not noted and simply results in disenrollment from the database. In the Optum DOD database, death outside of the inpatient setting is ascertained from the Social Security Administration’s Death Master File (DMF). It is important to note that there were important changes in the accuracy of the DMF pre-post 2011. One study found that the DMF ascertained in-hospital mortality with a sensitivity of 71.4% before 2011 and 28.9% post-2011 (Levin, 2019). Even if a meaningful effect of these exposures on the outcome sudden cardiac death does exist, under-ascertainment of the outcome has the potential to bias the result to the null. Thus, we plan to interpret any null findings from that analysis with appropriate caution.
* Causality between drug exposure and any given event cannot be drawn for individual cases. These analyses are intended to estimate the average effect of exposure (abiraterone vs. enzalutamide) among those who were treated with abiraterone (i.e. the average population-level effect in the treated).
* In this analysis we impose no upper limit on the length of follow-up, meaning that if a person does not experience one of the censoring events, their follow-up time will extend until the end of the database. There are several implications to this design choice.
  + First, we may be concerned if we observe meaningfully different amounts of available follow-up in the database for the two drugs being compared (abiraterone and enzalutamide). However, the initial diagnostics presented in the Shiny app do not seem to indicate a meaningful difference in the distribution of available follow-up time / time-at-risk.
  + Second, we present a single estimate of the hazard ratio, which is averaged over the duration of follow-up. For the primary comparison (i.e. not stratified by CVD risk factors), the maximum follow-up available was 2,410 days (6.6 years) in the intent-to-treat analysis and 2,006 days (5.5 years) in the as-treated analysis. However, in these analyses the median follow-up was 355 days (1.0 years) and 148 days (0.41 years), respectively, indicating that half the cohort had less than a year of follow-up available in these analyses. As pointed out by Miguel Hernan, “The magnitude of the average HR depends on the length of follow-up because the average HR ignores the distribution of events during the follow-up.” (Hernán, 2010). In order to address this short-coming, we also plan to consider Kaplan-Meier survival curves which, when fit with the appropriate confidence intervals, will allow us to explore whether the survival functions differ significantly from one another at any given point in follow-up. Sudden and/or dramatic changes in the separation of the survival curves over the course of follow-up will indicate that that the follow-up averaged hazard ratio estimate may be masking important heterogeneity in the effect over time. However, Kaplan-Meier curves suffer from their own limitations. Specifically, a comparison of the survival probabilities at a given point in the Kaplan-Meier curve does not account for “the distribution of events between baseline and *t*.” However, we consider these two approaches to be complimentary to one another and, together, capable of informing the evidence.
* Using these data, we are limited in our ability to meaningfully differentiate patients with less severe vs. more severe cardiovascular disease. In this study, we define less severe cardiovascular disease as having one or two CVD-related diagnoses or procedures in the 180 days before index and we define more severe cardiovascular as having three or more CVD-related diagnoses or procedures in the 180 days before index. While we understand that this definition is somewhat crude, we believe that using it to stratify estimates will provide valuable information about whether effects are heterogeneous with respect to baseline cardiovascular condition.
* Our definition of abiraterone initiation requires a concomitant predniso(lo)ne prescription within 30 days of first observed treatment. However, predniso(lo)ne use is likely predniso(lo)ne in these data, which means there may be patients who do initiate abiraterone and predniso(lo)ne at the same time but, due to data misclassification, are not included in this analysis. We do not expect this limitation to meaningfully bias the results we produce in a study population comprised of patients who do have predniso(lo)ne use recorded in the data.
* It was our intention to examine multiple sub-groups of patients with varying degrees of baseline cardiovascular risk. However, for the vast majority of our proposed analyses, the available sample in our data sources was not sufficient to produce reliable estimates within these sub-groups. While we have elected not to unblind many of these analyses (on the basis of empirical diagnostics), we believe the process of specifying them and evaluating their feasibility has already been valuable. This pilot project is partially intended to foster greater collaboration between teams within Janssen and inform a wider understanding of study feasibility as it relates to proposed future analyses. We believe the process we developed as part of this study will be highly valuable in future oncology safety studies where we are likely to require similar assessments and decision-making around unblinding potentially limited results.

# PROTECTION OF HUMAN SUBJECTS

The New England Institutional Review Board (IRB) has determined that studies conducted in IBM MarketScan CCAE, MDCR, MDCD, and Optum Extended DOD are exempt from study-specific IRB review, as these studies do not qualify as human subjects research. All patient data are deidentified in this study. All study reports will contain aggregate data only and will not identify individual patients or physicians. At no time during the study will the sponsor receive patient identifying information except when it is required by regulations in case of reporting adverse events.

# SAFETY DATA COLLECTION AND REPORTING

This study uses coded data that already exist in a electronic database. In this type of database, the minimum criteria for reporting an adverse event (i.e., identifiable patient, identifiable reporter, a suspect product, and event) are not available, and adverse events are not reportable as individual case safety reports. The study results will be assessed for medically important results and regardless of findings, will be shared with the safety management team (who is actively collaborating on this project) in order to inform next steps.

# PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Once finalized, the protocol will be registered at either [www.clinicaltrials.gov](http://www.clinicaltrials.gov) or <http://www.encepp.eu/> (European Network of Centres for Pharmacoepidemiology and Pharmacovigilance). Results will be reported to the registration location within 12 months of completion. Additionally, results will be submitted for peer-reviewed publication. All results will be shared in the form of presentation at scientific conferences and/or as peer-reviewed publications.

# list of tables & figures

|  |
| --- |
| Include table shells to indicate anticipated format of output. Other tables and figures to support the protocol can also be included here. Note that summary statistics will be provided in an automated process if study is executed by Epi Analytics. |

# ANNEX (LIST OF STAND-ALONE DOCUMENTS)

|  |
| --- |
| Documents listed in the Annex can be maintained separately from the study protocol (e.g., variable definitions, sample case report form, detailed Statistical Analysis Plan, STROBE or ENCePP checklist, etc.). All annexes should be clearly identifiable and provided on request. Write “None” if there is no document or list documents in a table as indicated below. |

|  |  |
| --- | --- |
| **Document Number** | **Title** |
| 1 | Concept set expressions, included SNOMED concepts, and included source codes for concepts used to define target, comparator, and outcome cohorts.   * *File:* conceptSetExpressions - Exposures and Outcomes.xlsx |
| 2 | Concept set expressions, included SNOMED concepts, and included source codes for concepts used to define cardiovascular-disease stratification.   * *File:* conceptSetExpressions - CVD (stratification).xlsx |
| 3 | Fully specified human-readable descriptions of study cohorts:   * *File:* Human-Readable Cohort Descriptions.docx |
| 4 | A descriptive comparison of baseline characteristics for initiators of abiraterone versus enzalutamide, among patients with castration-resistant prostate cancer   * *Link:* <https://epi.jnj.com/atlas/#/cc/characterizations/671> |
| 5 | Web-based R Shiny application displaying study diagnostics (and table shells with blinded results)   * *Link*: <https://sharedshiny.jnj.com/user/mconove1/EPI720_3/> |

# references

Albertsen, P. C. (2015). "Observational studies and the natural history of screen-detected prostate cancer." Curr Opin Urol **25**(3): 232-237.

Austin P. "Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples." Stat Med. 2009 Nov 10;28(25):3083-107. https://doi.org/10.1002/sim.3697

Banefelt, J., A. Liede, J. Mesterton, J. Stalhammar, R. K. Hernandez, P. Sobocki and B. E. Persson (2014). "Survival and clinical metastases among prostate cancer patients treated with androgen deprivation therapy in Sweden." Cancer Epidemiol **38**(4): 442-447.

Bernard, B., V. Muralidhar, Y. H. Chen, S. S. Sridhar, E. P. Mitchell, C. A. Pettaway, M. A. Carducci, P. L. Nguyen and C. J. Sweeney (2017). "Impact of ethnicity on the outcome of men with metastatic, hormone-sensitive prostate cancer." Cancer **123**(9): 1536-1544.

Bernard, B. and C. J. Sweeney (2015). "Management of metastatic hormone-sensitive prostate cancer." Curr Urol Rep **16**(3): 14.

Brawley, O. W. (2012). "Prostate cancer epidemiology in the United States." World J Urol **30**(2): 195-200.

Bray, F., J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal (2018). "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." CA Cancer J Clin **68**(6): 394-424.

Dalela, D., M. Sun, M. Diaz, P. Karabon, T. Seisen, Q.-D. Trinh, M. Menon and F. Abdollah (2019). "Contemporary Trends in the Incidence of Metastatic Prostate Cancer Among US Men: Results from Nationwide Analyses." European Urology Focus **5**(1): 77-80.

Engholm, G. F., J; Christensen, N; Hansen, HL; et al. (2018, June 2018). "Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 8.1 " Retrieved Nov. 26, 2018, from <http://www.ancr.nu>.

Hernán MA. "The hazards of hazard ratios." Epidemiology. 2010 Jan;21(1):13-5. DOI: 10.1097/EDE.0b013e3181c1ea43. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653612/

Hirst, C. J., C. Cabrera and M. Kirby (2012). "Epidemiology of castration resistant prostate cancer: a longitudinal analysis using a UK primary care database." Cancer Epidemiol **36**(6): e349-353.

Holmberg, L. and M. Van Hemelrijck (2014). "The biology and natural history of prostate cancer: a short introduction." Recent Results Cancer Res **202**: 1-7.

International Agency for Research on Cancer. (2018). "Cancer Today." Retrieved Nov. 29, 2018, from <http://gco.iarc.fr/today/home>.

Kantar Health. (2018). "Western Europe Patient Metrics: Sources and Methodology." Retrieved Dec. 6, 2018, from <http://cancermpact.khapps.com/#id/3458>.

Levin, M. A., et al. (2019). "Alive or dead: Validity of the Social Security Administration Death Master File after 2011." Health Serv Res 54(1): 24-33.

Liede, A., O. Günther, B. Bennett and S. Wong (2012). Prevalence of non-metastatic castration-resistant prostate cancer in Europe. European Society for Medical Oncology. Vienna, Austria.

Lu-Yao, G., et al. (2019). Clinical outcomes among patients treated with abiraterone acetate for advanced prostate cancer with pre-existing cardiovascular conditions, AACR.

Mottet, N., M. De Santis, E. Briers, L. Bourke, S. Gillessen, J. P. Grummet, T. B. Lam, H. G. van der Poel, O. Rouvière, R. C. N. van den Bergh and P. Cornford (2018). "Updated Guidelines for Metastatic Hormone-sensitive Prostate Cancer: Abiraterone Acetate Combined with Castration Is Another Standard." European Urology **73**(3): 316-321.

Ng, H. S., B. Koczwara, D. Roder and A. Vitry (2018). "Development of comorbidities in men with prostate cancer treated with androgen deprivation therapy: an Australian population-based cohort study." Prostate Cancer Prostatic Dis **21**(3): 403-410.

Nguyen-Nielsen, M., A. Liede, M. L. Maegbaek, M. Borre, N. Harving, R. K. Hernandez, H. Toft Sorensen and V. Ehrenstein (2015). "Survival and PSA-markers for mortality and metastasis in nonmetastatic prostate cancer treated with androgen deprivation therapy." Cancer Epidemiol **39**(4): 623-632.

Noone AM, Howlader N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ and C. K. (eds). (2018). "SEER Cancer Statistics Review, 1975-2015." Retrieved Nov. 27, 2018, from <https://seer.cancer.gov/csr/1975_2015/> based on November 2017 SEER data submission, posted to the SEER web site, April 2018.

OHDSI. “Definition and DDLs for the OMOP Common Data Model (CDM)” CommonDataModel, https://github.com/OHDSI/CommonDataModel. Accessed 2 February 2020.

Pascale, M., C. N. Azinwi, B. Marongiu, G. Pesce, F. Stoffel and E. Roggero (2017). "The outcome of prostate cancer patients treated with curative intent strongly depends on survival after metastatic progression." BMC cancer **17**(1): 651-651.

Pazdur, R. U.S. Food & Drug Administration: Center for Drug Evaluation and Research. (2012). "NDA Approval Letter for Xtandi®" (NDA 203415). Silver Spring, MD. URL: https://www.accessdata.fda.gov/drugsatfda\_docs/appletter/2012/203415Orig1s000ltr.pdf

Punnen, S. and M. R. Cooperberg (2013). "The epidemiology of high-risk prostate cancer." Curr Opin Urol **23**(4): 331-336.

Scher, H. I., K. Solo, J. Valant, M. B. Todd and M. Mehra (2015). "Prevalence of Prostate Cancer Clinical States and Mortality in the United States: Estimates Using a Dynamic Progression Model." PLoS ONE **10**(10): e0139440.

Schneeweiss S, Patrick AR, Stürmer T, Brookhart MA, Avorn J, Maclure M, Rothman KJ, Glynn RJ. "Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results." Med Care. 2007 Oct;45(10 Supl 2):S131-42. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2905666/

SEER. (2018). "Cancer Stat Facts: Prostate Cancer." Retrieved Feb. 13, 2019, from <https://seer.cancer.gov/statfacts/html/prost.html>.

Schoenfeld, D.A. "Sample-size formula for the proportional-hazards regression model." Biometrics, 1983. 39(2): p. 499-503.

Sternberg, C. N., E. S. Baskin-Bey, M. Watson, A. Worsfold, A. Rider and B. Tombal (2013). "Treatment patterns and characteristics of European patients with castration-resistant prostate cancer." BMC Urology **13**: 58-58.

Suchard, M. A., et al. (2019). "Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis." Lancet 394(10211): 1816-1826.

Voss, E.A., et al., "Feasibility and utility of applications of the common data model to multiple, disparate observational health databases." J Am Med Inform Assoc, 2015. 22(3): p. 553-64.