How to model temporal changes in epidemiological data

Treatment trajectories in men with prostate cancer

EUGENIO VENTIMIGLIA

Owing to the improvements in detection, diagnostics, and treatment, many men currently diagnosed with prostate cancer (PCa) have a low risk of PCa death. For many men PCa has become a chronic disease with a small risk of progression that remains even decades after date of diagnosis. In this thesis PCa was used as an example of a chronic disease, since it holds all the main characteristics required by the WHO definition of a chronic disease. Against this background, the aim of this PhD thesis was to create models that can be used to quantify the probability of different treatment trajectories and to assess the duration of certain disease states, while accounting for patient characteristics, disease severity, and primary treatment.

In Paper I a state transition model was developed for prediction of disease trajectories. The developed state transition model showed good consistency with a follow-up spanning up to 30 years.

In Paper II a state transition model was developed and validated using age, Charlson comorbidity index, and a drug comorbidity index (DCI) based on filled drug prescriptions collected at a population-based level to estimate life expectancy.

In Paper III the state transition model proposed in Paper I was used to assess 30-year PCa trajectories in men managed with active surveillance, in order to identify the ideal candidates for this management strategy.

In Paper IV the state transition model from Paper I was updated including the castration resistant PCa (CRPC) state as an additional state and estimating the duration of the CRPC state as well as its outcomes.

In Paper V, the updated state transition model from Paper IV was used to model long term outcomes for men with PCa managed with watchful waiting (WW). Since WW is currently recommended for men with PCa and life expectancy less than 10 years, the state transition model from Paper II was used to estimate life expectancy.

Keywords: prostate cancer, state transition model, life expectancy, active surveillance, watchful waiting

Eugenio Vendimiglia, Department of Surgical Sciences, Urology, Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden.

© Eugenio Vendimiglia 2023

ISSN 1651-6206
URN urn:nbn:se:uu:diva-499770 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-499770)
To Marta and Anna
## Overview of papers included in this thesis

<table>
<thead>
<tr>
<th>Paper</th>
<th>Design</th>
<th>Study population</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Method paper</td>
<td>Men diagnosed with PCa and registered in NPCR</td>
<td>To model disease trajectories for men with PCa</td>
</tr>
<tr>
<td>II</td>
<td>Method paper</td>
<td>PCa-free controls in PCBaSe</td>
<td>To model comorbidity trajectories for men and to predict their life expectancy</td>
</tr>
<tr>
<td>III</td>
<td>Cohort study</td>
<td>Men diagnosed with low and intermediate-risk PCa, registered in NPCR, and managed with deferred treatment</td>
<td>To model very long-term outcomes (30 years) for men with PCa and managed with active surveillance (AS) in order to identify the best candidates for AS</td>
</tr>
<tr>
<td>IV</td>
<td>Cohort study</td>
<td>Men with CRPC selected among 40% of the Swedish male adult population with available PSA data</td>
<td>To implement the state transition model from Paper I including CRPC, with the aim to estimate the duration of the CRPC state as well as outcomes for men in the CRPC state</td>
</tr>
<tr>
<td>V</td>
<td>Cohort study</td>
<td>Men in NPCR managed with watchful waiting</td>
<td>To model very long-term outcomes for men with PCa and managed with watchful waiting based on the updated model from paper IV and estimating life expectancy using the methodology from Paper II</td>
</tr>
</tbody>
</table>

CRPC: castration resistant prostate cancer  
NPCR: National Prostate Cancer Register of Sweden  
PCa: prostate cancer  
PCBaSe: Prostate Cancer data Base Sweden  
PSA: prostate specific antigen
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the publisher.
Contents

Introduction ................................................................. 11
Chronic diseases .......................................................... 11
Prostate cancer as a paradigm for chronic disease ............... 11
Prostate cancer staging, grading and risk categories .......... 12
Prostate cancer management strategies ............................ 14
Castration resistant prostate cancer (CRPC) ....................... 15
Comorbidity assessed by use of the Charlson comorbidity index (CCI) and drug comorbidity index (DCI) ............ 15

Aims of the studies .......................................................... 19
Overall aim ..................................................................... 19
Paper I ........................................................................ 19
Paper II ........................................................................ 19
Paper III ....................................................................... 19
Paper IV ....................................................................... 20
Paper V ........................................................................ 20

Materials and Methods .................................................. 21
The National Prostate Cancer Register .............................. 21
Prostate Cancer data Base Sweden .................................. 21
Paper I ........................................................................ 22
  Study population ........................................................ 22
  State transition model ................................................ 23
  Non-registered transitions .......................................... 24
Paper II ........................................................................ 24
  Study population ........................................................ 24
  State transition model ................................................ 24
  Validation .................................................................. 25
Paper III ....................................................................... 25
  Study population ........................................................ 25
  State transition model ................................................ 25
  Estimates .................................................................. 26
Paper IV .......................................................................................................................... 26
  Study population ......................................................................................................... 26
  Definition of castration sensitive PCa (CSPC) and castration resistant PCa (CRPC) ... 27
  State transition model ............................................................................................... 27
  Estimates ....................................................................................................................... 28
Paper V .......................................................................................................................... 28
  Study population ......................................................................................................... 28
  Life expectancy ............................................................................................................ 29
  State transitions .......................................................................................................... 29
  Microsimulation after censoring date ......................................................................... 29
  Estimates ....................................................................................................................... 30
Results ............................................................................................................................. 31
  Paper I .......................................................................................................................... 31
  Paper II ......................................................................................................................... 33
  Paper III ......................................................................................................................... 35
  Paper IV ......................................................................................................................... 36
  Paper V .......................................................................................................................... 37
Discussion ....................................................................................................................... 41
  Paper I .......................................................................................................................... 41
  Paper II ......................................................................................................................... 42
  Paper III ......................................................................................................................... 43
  Paper IV ......................................................................................................................... 43
  Paper V .......................................................................................................................... 44
Ethical considerations .................................................................................................. 46
  Data collection and registration in NPCR .................................................................... 46
  Data enrichment in PCBaSe ......................................................................................... 46
  Risk assessment ........................................................................................................... 47
  Ethical approvals ......................................................................................................... 47
Conclusions ..................................................................................................................... 48
Acknowledgements ........................................................................................................ 49
References ...................................................................................................................... 51
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADT</td>
<td>Androgen Deprivation Therapy</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen Receptor</td>
</tr>
<tr>
<td>AS</td>
<td>Active Surveillance</td>
</tr>
<tr>
<td>CCI</td>
<td>Charlson Comorbidity Index</td>
</tr>
<tr>
<td>CRPC</td>
<td>Castration Resistant Prostate Cancer</td>
</tr>
<tr>
<td>CSPC</td>
<td>Castration Sensitive Prostate Cancer</td>
</tr>
<tr>
<td>DCI</td>
<td>Drug Comorbidity Index</td>
</tr>
<tr>
<td>DT</td>
<td>Doubling time</td>
</tr>
<tr>
<td>EAU</td>
<td>European Association of Urology</td>
</tr>
<tr>
<td>ERL</td>
<td>Estimated Remaining Lifetime</td>
</tr>
<tr>
<td>GGGG</td>
<td>Gleason Grading Group</td>
</tr>
<tr>
<td>ICD</td>
<td>International statistical Classification of Diseases and related health problems</td>
</tr>
<tr>
<td>INCA</td>
<td>Information Network for Cancer care</td>
</tr>
<tr>
<td>ISUP</td>
<td>International Society of Urological Pathology</td>
</tr>
<tr>
<td>NCCN</td>
<td>National Comprehensive Cancer Network</td>
</tr>
<tr>
<td>NPCR</td>
<td>National Prostate Cancer Register</td>
</tr>
<tr>
<td>PCa</td>
<td>Prostate Cancer</td>
</tr>
<tr>
<td>PCBaSe</td>
<td>Prostate Cancer dataBase of Sweden</td>
</tr>
<tr>
<td>PSA</td>
<td>Prostate Specific Antigen</td>
</tr>
<tr>
<td>RP</td>
<td>Radical Prostatectomy</td>
</tr>
<tr>
<td>RT</td>
<td>Radiotherapy</td>
</tr>
<tr>
<td>RT$_{adj/salv}$</td>
<td>Adjuvant/salvage Radiotherapy</td>
</tr>
<tr>
<td>STHLM-0</td>
<td>Stockholm PSA and Biopsy Register</td>
</tr>
<tr>
<td>UPSAC</td>
<td>Uppsala-Örebro Prostate-specific antigen Cohort</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WW</td>
<td>Watchful Waiting</td>
</tr>
<tr>
<td>YLD</td>
<td>Years Lived with Disability</td>
</tr>
</tbody>
</table>
Introduction

Chronic diseases

The World Health Organization (WHO) defines chronic disease as a disease that lasts a long time, progresses slowly, and is not spread from person to person [1]. The Global Burden of Disease study 2017 reported an increase (20%) in the years lived with disability (YLD) from 1990 to 2017 [2]. This was attributable to non-communicable diseases, as infectious diseases were not among the top 20 primary causes of YLDs worldwide [3]. Chronic disease multi-morbidity is common in industrialized countries, and its prevalence rises with age: it is estimated that at least 50% of individuals above age 65 is affected by more two or more chronic comorbid conditions [3]. During the last decades, several previously fatal conditions have become chronic diseases due to advances in diagnosis and treatment [4]. These disease pathways (referred to as disease trajectories below) may span over several decades and include a variety of treatments and outcomes. Besides, both changes in age and comorbidities will impact on treatment decision. Therefore, it becomes essential to predict long term disease trajectories to inform patients and healthcare professionals on the outcomes of these conditions and provide information on the need of future health care [5].

Prostate cancer as a paradigm for chronic disease

According to the GLOBOCAN 2020 estimates, prostate cancer (PCa) is the second most frequently diagnosed cancer in men worldwide [6] accounting for 14% of all new cancer diagnosis. Despite a high incidence and prevalence, especially in older men, PCa only represents the 5th cause of cancer-related death worldwide [6]. However, there are large regional variations of this estimate, for instance in Sweden PCa is the leading cause of cancer mortality in men [7]. These differences are explained by geographic variation in socioeconomic status, life expectancy, and PCa diagnostic intensity [6].

Indeed, there is a wide variation in the clinical course of prostate cancer (PCa) [8]. Due to advancements in detection, diagnostics, and treatment, many men currently diagnosed with PCa have a low risk of PCa death during the first 15 years after diagnosis. For many men PCa has become a chronic disease, with a limited risk of progression that persists even decades after
diagnosis [9]. This has at least two main implications. First, low-risk PCa can be managed conservatively, deferring active treatment until/if signs of progression occur without increasing the cancer-specific risk of death [8]. Second, many treated and untreated men with PCa die of other causes, making PCa a chronic condition which affects quality of life, but does not reduce life expectancy [10]. For these reasons, in this thesis PCa will be treated as a proof of principle for a chronic disease, since it holds all the main characteristics required by the WHO definition.

Prostate cancer staging, grading and risk categories

As hinted before, PCa is a heterogeneous disease which spans from very low-risk, indolent PCa up to metastatic lethal disease [8]. PCa histological grading system, i.e. the Gleason grading system, is historically considered the main prognostic factor of PCa [11,12]. This grading depicts the architectural pattern of prostate cancer tissue, ranging from well differentiated, Gleason grade 1, to very poorly differentiated, Gleason grade 5. Since prostate adenocarcinomas are typically histologically heterogeneous, grading is difficult. The Gleason Score, which ranges from 2 to 10, is calculated as the sum of the dominant and second most dominant patterns [13]. The International Society of Urological Pathology (ISUP) advised in 2005 that the Gleason Score be based on the highest-grade pattern rather than the second most prominent pattern [14]. Gleason grade 4 was also assigned to typical cribriform glands, which were previously classified as grade 3. Because of this grade migration, the Gleason Score has inflated, making it difficult to compare survival in men with prostate cancer based on Gleason Score between time periods before and after 2005 [14]. In 2014, the Gleason Score was divided into five prognostic categories, Gleason Grade Groups (GGG) 1–5, with Gleason 7 being divided into two groups, GGG 2: 3+4 and GGG 3: 4+3 [15]. It has been questioned recently whether the use of GGG really improves clinical management of PCa [16]. The 2017 version of TNM staging is the one currently in use for PCa [17]. In detail, T1a and T1b represent incidentally detected PCa during transurethral resection of the prostate (TURP) and T1c is PCa detected in a work-up of an elevated PSA with no palpable tumor at digital rectal examination (DRE). T2 tumors are confined within the prostate and are palpable at DRE. T3-T4 are tumors that extend through the prostatic capsule (T3a), invade the seminal vesicles (T3b) or other structures of the pelvic region (T4). Lymph node metastasis in the pelvic region is classified as stage N1 and metastasis to lymph nodes outside the pelvis is described by stage M1a. M1b and M1c represent metastasis to bone and visceral organs.

The stratification of prostate cancer by risk of progression is the main factor that determines the management strategy. By use of TNM stage, Gleason score, and PSA levels it is possible to categorize the of risk of progression and
relapse with reasonable accuracy. In the five papers of this thesis, we defined risk categories according to a modification of the National Comprehensive Cancer Network (NCC) categorization: very low-risk, low-risk (not very low), intermediate-risk, high-risk or locally advanced, regionally metastatic, and distant metastases. Table 1 details the clinical features of these risk categories.

Table 1. Prostate cancer risk categories

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Clinical features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low-risk</td>
<td>T1c</td>
</tr>
<tr>
<td></td>
<td>Positive cores ≤ 33%</td>
</tr>
<tr>
<td></td>
<td>Total cancer extent in systematic biopsies ≤ 8 mm</td>
</tr>
<tr>
<td></td>
<td>Gleason score 6</td>
</tr>
<tr>
<td></td>
<td>PSA&lt;10 ng/ml,</td>
</tr>
<tr>
<td></td>
<td>Prostate volume &lt; 90cc</td>
</tr>
<tr>
<td></td>
<td>&gt;5 core biopsies performed</td>
</tr>
<tr>
<td></td>
<td>PSA density &lt; 0.15 ng/ml/cc</td>
</tr>
<tr>
<td>Low-risk, not very low</td>
<td>Gleason score 6</td>
</tr>
<tr>
<td></td>
<td>PSA&lt;10 ng/ml,</td>
</tr>
<tr>
<td></td>
<td>T1 or T2 with at least one of the following characteristics:</td>
</tr>
<tr>
<td></td>
<td>• prostate volume ≥ 90cc</td>
</tr>
<tr>
<td></td>
<td>• PSA density ≥ 0.15 ng/ml/cc</td>
</tr>
<tr>
<td></td>
<td>• ≤5 core biopsies performed</td>
</tr>
<tr>
<td></td>
<td>• positive cores &gt; 33%</td>
</tr>
<tr>
<td></td>
<td>• total cancer extent in systematic biopsies &gt; 8 mm</td>
</tr>
<tr>
<td></td>
<td>• or Gleason score 6, T1 or T2 with</td>
</tr>
<tr>
<td></td>
<td>10 ng/ml ≤ PSA&lt;15 ng/ml</td>
</tr>
<tr>
<td>Intermediate-risk</td>
<td>Gleason score 7 (3+4)</td>
</tr>
<tr>
<td></td>
<td>PSA&lt;10 ng/ml</td>
</tr>
<tr>
<td></td>
<td>T1 or T2</td>
</tr>
<tr>
<td>High-risk or locally ad-</td>
<td>T4 or N1 or PSA 50-99.9 ng/ml</td>
</tr>
<tr>
<td>vanced</td>
<td></td>
</tr>
<tr>
<td>Distant metastases</td>
<td>M1 or PSA ≥ 100 ng/ml</td>
</tr>
</tbody>
</table>
Prostate cancer management strategies

According to the European Association of Urology (EAU) guidelines [8], it is possible to manage PCa with the following strategies:

- Deferred treatment, which consists of either active surveillance (AS) and watchful waiting (WW). AS aims to avoid unnecessary treatment in men with clinically localized PCa who do not require immediate treatment, but at the same time achieve the correct timing for curative treatment in those who eventually do. Watchful waiting refers to conservative management for patients deemed unsuitable for curative treatment from the outset, and patients are clinically ‘watched’ for the development of local or systemic progression with (imminent) disease-related complaints, at which stage they are then treated palliatively according to their symptoms in order to maintain quality of life.

- Curative treatment, consisting of either radical prostatectomy (RP) and radiotherapy (RT). RP is the surgical curative treatment for PCa which eradicates cancer while, whenever possible, preserving pelvic organ function. This procedure involves removing the entire prostate with its capsule intact and the seminal vesicles, followed by vesico-urethral anastomosis. Lymph node dissection is performed in patients with a pre-operative risk of lymph node invasion ≥ 5%. RP has excellent oncological outcomes in men with organ confined disease (up to 99% cancer specific survival, [10]), however it is burdened by postoperative urinary incontinence and sexual dysfunctions. RT has equivalent oncological outcomes as compared to RP in organ confined PCa, however with a different profile in terms of side effects [18].

- Adjuvant and salvage radiotherapy: RT may serve also as a secondary treatment following RP either in case of disease persistence/recurrence or as an adjuvant treatment. PSA becomes undetectable following RP, therefore a steady and consistent post-operative PSA rise is considered as a sign of biochemical recurrence, especially when above 0.2 ng/ml [8]. To date, there is not a precise consensus regarding timing and PSA cutoff for delivering salvage RT, as well as the optimal candidates for adjuvant RT.

- Androgen deprivation therapy (ADT): hormonal therapy which aims to reduce androgenic stimulation of prostate cancer cells. This can be achieved by either suppressing the production of testicular androgens or inhibiting the action of circulating androgens at the level of the androgen receptor.
Castration resistant prostate cancer (CRPC)

In men on ADT, the development of castration resistance (CRPC) is defined by castrate serum testosterone < 50 ng/dL or 1.7 nmol/L plus either:

1) three consecutive rises in PSA at least one week apart resulting in two 50% increases over the nadir, and a PSA > 2 ng/mL or
2) the appearance of new lesions: either two or more new bone lesions on bone scan or a soft tissue lesion.

Two major overlapping mechanisms, androgen-receptor (AR)-independent and AR-dependent mechanisms, are thought to mediate the onset of CRPC. Once men reach the CRPC state, their median survival is 3.5 years [19]. For this reason, during the last ten years CRPC has become one of the main research focus in PCa, with novel drugs being developed and approved in this setting.

Comorbidity assessed by use of the Charlson comorbidity index (CCI) and drug comorbidity index (DCI)

The presence of comorbidities affects life expectancy, both in cancer and non-cancer patients [20]. Therefore, it is important to include changes in comorbidity that occur over time in analysis of treatment trajectories in men with chronic diseases, since comorbidities affect treatment choice and disease progression [21]. A major challenge in epidemiological studies is how to comprehensively assess comorbidities, as well as how to describe change in comorbidities over time [21]. In 1987, Charlson et al introduced the concept of the Charlson comorbidity index (CCI) [22], which considers 17 different comorbid conditions (Table 2) in order to provide a score associated with the risk of death at 1 year. The CCI gained wide diffusion and is nowadays one of the most employed tools in clinical and epidemiological studies [23]. However, it has several limitations. First, it was developed in hospitalized patients, although it is regularly used in the outpatient setting. Second, the scores attributed to the different comorbidities are now obsolete, for example any solid cancer contributes with 2 points irrespectively of cancer stage, whereas AIDS gives a score of 6 [22]. Despite these limitations, CCI remains widely accepted and used: in 2020, 1,498 studies mentioned CCI, an all-time high-score since its introduction 1987 [24]. CCI was recently shown to be an effective tool for the assessment of comorbidities in register-based studies [25].
Table 2. Diagnoses and their weights in the Charlson comorbidity index (CCI)

<table>
<thead>
<tr>
<th>Condition (ICD-9-CM codes)</th>
<th>CCI Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Myocardial infarction</strong></td>
<td>1</td>
</tr>
<tr>
<td>Old myocardial infarction (412)</td>
<td></td>
</tr>
<tr>
<td><strong>Congestive heart failure</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Peripheral vascular disease</strong></td>
<td>1</td>
</tr>
<tr>
<td>Aortic aneurysm (441)</td>
<td></td>
</tr>
<tr>
<td>Peripheral vascular disease NOS (443.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Cerebrovascular disease</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Connective tissue disorder</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Peptic ulcer</strong></td>
<td>1</td>
</tr>
<tr>
<td>Acute peptic ulcer NOS (533)</td>
<td></td>
</tr>
<tr>
<td><strong>Pulmonary disease</strong></td>
<td>1</td>
</tr>
<tr>
<td>COPD (496)</td>
<td></td>
</tr>
<tr>
<td><strong>Liver disease</strong></td>
<td>1</td>
</tr>
<tr>
<td>Alcoholic cirrhosis of liver (571.2)</td>
<td></td>
</tr>
<tr>
<td>Chronic hepatitis (571.4)</td>
<td></td>
</tr>
<tr>
<td>Cirrhosis of liver NOS (571.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td>1</td>
</tr>
<tr>
<td>Diabetes mellitus (250)</td>
<td></td>
</tr>
<tr>
<td><strong>Diabetes complication</strong></td>
<td>2</td>
</tr>
<tr>
<td><strong>Renal disease</strong></td>
<td>2</td>
</tr>
<tr>
<td>Chronic glomerulonephritis (582)</td>
<td></td>
</tr>
<tr>
<td>Nephritis and nephropathy (583)</td>
<td></td>
</tr>
<tr>
<td><strong>Cancer (malignant neoplasm)</strong></td>
<td>2</td>
</tr>
<tr>
<td>of the prostate (185)</td>
<td></td>
</tr>
<tr>
<td>of the testis (186)</td>
<td></td>
</tr>
<tr>
<td>of the kidney (189.0)</td>
<td></td>
</tr>
<tr>
<td>of the descending colon (153.2)</td>
<td></td>
</tr>
<tr>
<td>of the main bronchus (162.2)</td>
<td></td>
</tr>
<tr>
<td>of the thyroid (193)</td>
<td></td>
</tr>
<tr>
<td>melanoma (skin) NOS (172)</td>
<td></td>
</tr>
<tr>
<td>of the bladder (188)</td>
<td></td>
</tr>
<tr>
<td>Lymphoma NOS (202.8)</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td></td>
</tr>
<tr>
<td><strong>Metastatic cancer</strong></td>
<td>6</td>
</tr>
<tr>
<td>testis cancer with positive retroperitoneal lymph nodes (202.8)</td>
<td></td>
</tr>
<tr>
<td><strong>HIV disease</strong></td>
<td>6</td>
</tr>
</tbody>
</table>

A proposal for the management of comorbidities change over time was proposed by Lindhagen and Garmo [26]. Changes in comorbidity, as measured by the Charlson Comorbidity index (CCI), occur multiple times and with
different sizes (1, 2, 3, or 6), which makes CCI difficult to describe when using conventional time to event analysis. Lindhagen and Garmo proposed a method using a multi-state approach to model temporal changes in comorbidity for cancer patients using data from PCBaSe between 1992 and 2012. This method assumes that a change in CCI is an irreversible process, i.e., CCI accumulates over time and cannot decrease, and it is based on a state transition model approach with states and state transitions (Figure 1). In this approach, follow-up time is discretized in time steps of four weeks. At each time step an individual can experience death, a change in CCI, or remain in the previous CCI state. A three-step modelling-simulation approach yields the estimates of CCI increase and survival:

1. The vital status of each individual is estimated at the end of every time step
2. If the individual survived, the model determines whether a CCI occurred
3. When a CCI change occurred, this step defines the entity of CCI change

With this approach it is possible to model and monitor comorbidity trajectories over time, since age and CCI are updated at each time step.

Figure 1. State transition model for change in Charlson comorbidity index (CCI) over time

Possible Charlson comorbidity index (CCI) states (blue) and final death-state (black) in the state transition model. The arrows reflect the possible changes in CCI and the possibility of death.

As previously discussed, CCI has important limitations. Gedeborg and Garmo recently proposed a novel drug comorbidity index (DCI) based on filled drug prescriptions and compared it to CCI in order to predict survival [26] in the PCa-free control population of PCBaSe as well as in breast-cancer free controls of the Swedish breast cancer register. They found that DCI was a better predictor when compared to CCI in different age strata and may serve as an
estimator of remaining survival in health-care registers containing data on filled drug prescriptions. Moreover, it can be modelled as previously shown for CCI.
Aims of the studies

Overall aim
The aim of this PhD thesis was to create models that can be used to quantify the probability of different treatment trajectories and to assess the duration of certain disease states, while accounting for patient characteristics, disease severity and primary treatment. A statistical model that accounts for these temporal changes will inform future medical decision making which will be useful given that also other previously lethal diseases will become chronic, and the concept of healthy ageing and its health-economic impact will rely on these temporal effects.

Paper I
The aim of the first paper was to develop a state transition model that predicts disease trajectories until 25 years after diagnosis using PCa as a paradigm for a chronic disease. More specifically, we used data on PCa from PCBaSe to predict disease trajectories in men with different disease severity, comorbidity profile, and management strategy using state transition models.

Paper II
The aim of the second paper was to develop a model using age, Charlson comorbidity index, and a drug comorbidity index (DCI) based on filled drug prescriptions collected at a population-based level to estimate life expectancy. We developed life tables for each level of age, CCI and DCI, and allowed this to change during follow-up by use of a state transition model.

Paper III
The aim of the third paper was to apply the state transition model proposed in Paper I in order to assess PCa trajectories in men with low or intermediate-risk PCa managed with active surveillance. In this context, we wanted to
identify men who would benefit the most from active surveillance in terms of treatment-free years and low risk of PCa death.

Paper IV
The aim of the fourth paper was to implement the state transition model proposed in Paper I with the inclusion of a new state, castration resistant prostate cancer (CRPC). We then analyzed the duration of the CRPC state as well as the survival outcomes for men reaching this state according to different baseline PCa risk group.

Paper V
The aim of the fifth paper was to use the updated model from paper IV in order to estimate long term PCa trajectories in men initially managed with watchful waiting (WW). We used the model from Paper II in order to estimate baseline estimated remaining lifetime (ERL) for included men, since WW is a management strategy indicated in men with ERL < 10 years. We then identified those men who are the best candidates to be managed with WW.
Materials and Methods

The National Prostate Cancer Register

All papers in this thesis are based on data in the National Prostate Cancer Register (NPCR) of Sweden. Data on incident cases of prostate cancer has been registered in NPCR since 1998 and has a capture rate of 98% as compared with the Swedish Cancer Registry, to which reporting is mandated by law [27]. Completeness and validation of the data is checked at each of the six Regional Cancer Centre (RCC) throughout the country before trans-mission to the Swedish Cancer Registry and NPCR. Corrections and updates of previous years are made continuously. Detailed data on prostate cancer diagnosis, work-up and treatment is provided to NPCR through four separate forms: diagnostic data, primary treatment and separate forms for radical prostatectomy and radical radiotherapy. Since 2007 each form is reported to NPCR at an online platform, the Information Network for Cancer care (INCA). Variables registered include date and health care unit of diagnosis, TNM classification, Gleason Score, serum levels of PSA at diagnosis, and primary treatment delivered within six months after date of diagnosis [28,29]. In order to ensure that data in NPCR is correct there are multiple logical controls in the online registration, and there are manuals to guide staff that performs the registration, as well as annual meeting for staff who perform registration (www.npcr.se). A formal audit of data has been performed [30]. Despite the high capture rate and the high-fidelity of registered data [31], there are some issues with NPCR when trying to use it for the assessment of treatment and disease trajectories. First, and most important, there are virtually no data on follow-up after primary treatment, this information is only available for a very limited set of men. Second, there is no information regarding comorbidities, meaning that it is not possible to estimate treatment trajectories also in this setting. Third, both active surveillance and watchful waiting were registered as ‘conservative therapy’ until 2007 and only after that date are these two strategies separately registered.

Prostate Cancer data Base Sweden

Since 2008, by use of the Swedish personal identity number, NPCR has been linked with several other health care registers and demographic databases,
including The National Patient Registry, The Prescribed Drug Registry, The Cause of Death Registry, The Multi-Generation Register, and The Longitudinal integration database for health insurance and labor market studies (LISA), with the aim of constructing Prostate Cancer data Base Sweden (PCBaSe), a database for registry-based research (Figure 2). These linkages have been performed multiple times since 2008 in order to include more cases and a longer follow-up. In the studies included in this thesis, we used data in PCBaSeTraject which is a refined version of PCBaSe 4 with information on cause of death, comorbidities, drug use, socio-economic status, and treatment changes [28]. In the following paragraphs, it will be discussed more in detail how both PCAs and comorbidities trajectories were obtained and estimated.

Figure 2. Linkages between NPCR and other health care registers and demographic databases in Prostate Cancer data Base Sweden

Paper I

Study population

The study reported in Paper I was based on 118,743 men diagnosed with PCAs during 1992-2014. Included men were classified into five groups based on the primary management strategy: active surveillance (AS), radical prostatectomy (RP), radiotherapy (RT) performed as either external beam radio therapy or brachy therapy, watchful waiting (WW), and androgen deprivation therapy (ADT). Specific risk categories were defined for each management strategy. For AS risk categories, we included men with very low-risk, low-risk (but not very low), and favorable intermediate risk (i.e. Gleason score 7 (3+4), PSA<10 ng/ml, T1 or T2); the least favorable intermediate-risk category was not considered as suitable for AS. Curative treatment consisted of RP or RT, which could be either primary treatment or initiated after a period of AS, i.e. deferred RP or RT. Further states were watchful waiting, either as a primary strategy or following active surveillance, and androgen deprivation therapy (ADT). Treatment states were divided into specific risk categories based on TNM stage, pathological Gleason score and PSA levels.
State transition model

A specifically developed state transition model (Figure 3) was applied to our study population. After PCa diagnosis, included men entered their primary state according to their primary management strategy (AS, RP, RT, WW, ADT) and their treatment specific risk category (AS₁-AS₅, RP₁-RP₉, RT₁-RT₉, WW₁-WW₆, and ADT₁-ADT₉). Transition probabilities to another state were based on age, comorbidity, history of previous treatments, and treatment-specific risk category. All transitions were considered irreversible and state transitions were allowed until a final absorbing state, i.e. PCa-death or death from other causes, was reached. Transitions between treatment-specific risk categories were not considered, i.e. each man stayed in his designated treatment risk category until a new treatment was introduced.

![State transition model diagram](image)

**Figure 3. States and state transitions in the state transition model in Paper I**

State transition model of transitions (arrows) between states (circles) for men diagnosed with prostate cancer. The states are active surveillance (AS), watchful waiting (WW), curative treatment; radical prostatectomy (RP) or radiotherapy (RT), adjuvant or salvage radiotherapy following RP (RT-adj/salv), androgen deprivation therapy (ADT), death from other causes, and prostate cancer death. Dashed lines represent the choice of primary treatment following diagnosis (not part of the model), solid lines are transitions included in the models. Multi-colored circles represent transient states with colors indicating the proportion of men with increasing disease risk categories defined by data at date of diagnosis. Orange circles represent absorbing states. Dashed circles represent additional information gathered to facilitate estimates of transition probabilities, i.e. biopsy and Charlson Comorbidity Index (CCI).
Our dataset was randomly split in two equally large subsets, a training set that was used to estimate transition probabilities, and a second one that was used to assess internal validity of the simulation process. Follow-up time was discretized into 4-week intervals. At the end of each time step a man either remained in his current state or transited to a new state. Age, CCI, and cancer characteristics were updated in each time step. Next, we estimated state transition probabilities, which were determined from both registered and non-registered data regarding date of treatment change. We considered a change in CCI as a state transition. Eventually, we run a microsimulation, i.e. an individual-level simulation based on the state-transition model [32].

Non-registered transitions

Briefly, for all treatment state transitions, except the transition from AS → WW, the date of treatment change was retrieved from PCBaSeTraject (training set of data). This lack of data was considered as a missing data problem that was solved by use of information on dates of prostate biopsy as registered in the National Patient Registry in combination with recordings of WW and AS in NPCR. Briefly, the continued and timely use of prostate biopsies was used as an indication of adherence to AS; details on this approach have been described previously [33].

Paper II

Study population

This study is based on the PCa-free control men from PCaBaSe 4.0, who were matched to men with PCa in PCBaSe according to year of birth and county of residence. All men aged 65-90 years at entrance to the cohort in 1 Jan 2007 to 31 December 2013 were included. Follow-up ended at date of death, date of emigration, or 21 December 2017, whichever occurred first. Both the Charlson comorbidity index (CCI) and drug comorbidity index (DCI) were calculated at date of entry to the cohort and in each consecutive year until end of follow-up. For CCI, we used a cumulative CCI including all events dating back to 10 years prior to entry to the cohort. For the DCI, we used prescriptions filled in the previous year.

State transition model

This method predicts life expectancy at a population-level based on current age, CCI and DCI, while taking into account future changes in comorbidity as a dynamic process, using population-based data on Swedish men. We run a microsimulation of changes in vital status, CCI, and DCI for individual study
subjects. The proportion of deaths in each time step was calculated and used to create a survival curve. For each combination of age (65, 66, …, 90), CCI (0, 1, 2, …, 10), and DCI (-0.75, -0.5, …, 13.5), we ran the microsimulation using 10,000 identical men. The life expectancy was calculated as the area under the survival curve emanating from the simulation.

Validation
To evaluate the performance of our model predictions of life expectancy, we compared simulated and observed survival for men based on age, CCI and DCI. In addition, we compared the observed change in mean CCI and DCI over time following cohort entry and the corresponding change in simulated mean CCI. In addition, we estimated the probability of death within the next year, corresponding to the applied life tables. Finally, we assessed the validity of our method in a cross-validation by splitting our comparison cohort of PCBaSe 4.0 based on the six health care regions in Sweden. For each health care region, data from men in the other five regions were used to estimate transition probabilities.

Paper III
Study population
We identified men aged 40–75 years at date of diagnosis in Prostate Cancer data Base Sweden Traject (PCBaSeTraject) diagnosed in 1992-2014 with very-low, low-risk, and favorable intermediate-risk PCa (i.e. Gleason score 7 (3+4), PSA<10 ng/ml, T1 or T2).

State transition model
We report estimates from the previously described simulated state transition model (Paper I), to describe the full PCa trajectory or men managed with AS up to 30 years after diagnosis (Figure 4). The key assumption of our simulation was that after a transition from AS to a new treatment, the outcome is similar to that for men who received this treatment as primary treatment, according to the updated disease characteristics at date of transition.

In Paper III, microsimulation was performed for 100,000 men for all combinations of age at diagnosis, PCa risk categories, PSA levels, and Charlson Comorbidity Index until age 85 or follow-up of 30 years, whichever event occurred first.
Figure 4. State transition model of transitions (arrows) between states (circles) for men diagnosed with prostate cancer and managed with active surveillance

The states are active surveillance (AS), watchful waiting (WW), radical prostatectomy (RP) or radiotherapy (RT), adjuvant or salvage radiotherapy following RP (RT-adj/salv), androgen deprivation therapy (ADT), death from other causes, and prostate cancer death.

Estimates

Time-specific prevalence estimates for each combination of age, comorbidity, and PCa risk category were provided. In order to assess outcome of AS for each of these combinations, we described the relationship between PCa death and the proportion of lifetime after diagnosis without active PCa treatment under the assumption that lower risk of PCa death and longer lifetime after diagnosis without active PCa treatment represented the highest benefit of AS.

Paper IV

Study population

We selected men from the Uppsala-Örebro Prostate-specific antigen (PSA) Cohort (UPSAC) and the Stockholm PSA and Biopsy Register (STHLM-0), which include PSA data for 40% of the male Swedish population [34,35]. These cohorts were then linked to the National Prostate Cancer Register (NPCR) of Sweden and to several other national health care registers, including the Prescribed Drug Register, the Patient Register, and the Cause of Death Register. The following step consisted in the identification of men with a PCa diagnosis who on January 1, 2006 or thereafter had filled prescriptions for gonadotropin-releasing hormone (GnRH) agonists and had PSA levels recorded.
Definition of castration sensitive PCa (CSPC) and castration resistant PCa (CRPC)

A man was considered in the castration sensitive PCa (CSPC) state once he had undergone 3 mo of GnRH medication over a period of 6 mo, or on the day of bilateral orchidectomy. During follow-up, men were regarded as being in the CSPC condition based on the same standards. Men who had had androgen deprivation treatment (ADT) for three months or who had undergone surgical castration entered the castration resistant PCa (CRPC) state at the first date of doubling of their PSA nadir value with the last value being $>2$ ng/ml, or an absolute rise in PSA of $\geq 5$ ng/ml. Both CSPC and CRPC have pre-defined risk categories, which have been previously reported [36,37]. Briefly we calculated a “combined PSA kinetics risk” according to the equation:

$$\log(\text{PSA at CRPC}) - 1.4 \times \log(\text{PSA doubling time})$$

We then created eight CRPC risk categories that we used in our model.

State transition model

Transition probabilities for death from other causes were those from the original state transition model (Paper I). New transition probabilities were estimated for the following transitions: CSPC $\rightarrow$ CRPC, CSPC $\rightarrow$ PCa death, and CRPC $\rightarrow$ PCa death (Figure 5). The probabilities of events within a 28-d period were modeled using logistic regression. We then ran a microsimulation of every man in our set of data using a 28-d time interval until death or end of the simulated follow-up time (10 yr). To decrease random errors, we ran the simulation 100 times for each man.
Figure 5. State transition model of transitions (arrows) between states (circles) for men diagnosed with prostate cancer and who had started GnRH agonists

The states are gonadotropin releasing hormone agonist (GnRH), castration sensitive PCa (CSPC), castration resistant PCa (CRPC), death from other causes, and death from PCa.

Estimates

We then extracted and compared the observed and predicted cumulative incidence of transitions to the CRPC state and death from PCa. We then cross-validated the models by applying the model from the UPSAC cohort to the STHLM-0 cohort and vice versa. Eventually, we estimated the time spent in the CRPC state as well as the proportion of men who died of PCa during follow-up according to CRPC risk categories.

Paper V

Study population

We identified men in Prostate Cancer data Base Sweden (PCBaSe) who had been diagnosed with non-metastatic prostate cancer between the years 2007 and 2019, and who were registered with watchful waiting as treatment strategy in The National Prostate Cancer Register (NPCR) Sweden. European Association for Urology guidelines recommend watchful waiting for men with a life expectancy less than ten years. In order to apply this restriction in the selection of our study population, taking the uncertainty in estimation of life
expectancy, we restricted the study population to men with <12 years life expectancy.

Life expectancy
Life expectancy was calculated based on age, Charlson Comorbidity Index (CCI) and a Drug Comorbidity Index (DCI), by use of the state transition model described in Paper II.

State transitions
We considered transitions to androgen deprivation therapy (ADT), castration resistant prostate cancer, prostate cancer death, or death from other causes (Figure 6), as described in the updated state transition model from Paper IV. The watchful waiting-specific risk categories were based on a modified version of the National Comprehensive Cancer Network risk categorization.

![State transition model of transitions (arrows) between states (circles) for men diagnosed with prostate cancer and managed with watchful waiting](image)

Figure 6. State transition model of transitions (arrows) between states (circles) for men diagnosed with prostate cancer and managed with watchful waiting

The states are watchful waiting (WW), androgen deprivation therapy (ADT), castration sensitive PCa (CSPC), castration resistant PCa (CRPC), death from other causes, and death from PCa.

Microsimulation after censoring date
To increase follow-up time up to 20 years, state transitions before censoring were based on observed data and then combined with simulated data in the non-observed period. The microsimulation was based on information on age, state-specific prostate cancer risk category, history of changes in CCI, current
CCI, and the treatment history at date of censoring. To increase precision 100 replicates of each man were simulated.

Estimates
From the combined observed and simulated patient states over time, we visualized proportions of disease states after 5, 10, 15, 20 years of follow-up using Sankey diagrams stratified for prostate cancer risk group at diagnosis. Using prostate cancer risk category specific logistic regression models we predicted the proportions of men who died of prostate cancer, received ADT, or reached the CRPC state, and the proportion of life time on ADT in relation to life expectancy. We further made the assumption that lower risk of PC death and more life-years after diagnosis without PCa treatment represented the greatest benefit of watchful waiting. To illustrate this, we plotted the proportion of men with prostate cancer death vs the proportion of lifetime spent without prostate cancer treatment.
Results

Paper I

The internal validity of the simulation was assessed by comparing the cumulative incidence of the first transitions observed in PCBaSeTraject with the simulated first transitions for the same men up to 25 years after diagnosis. The simulation was run 100 times for each man with the goal of decreasing random error. There was an almost perfect overlap between observed and simulated first transitions for all groups (Figure 7). At ten years after RP, 21% of the observed PCBaSeTraject cohort had transitioned to adjuvant/salvage radiotherapy vs. 20% as predicted by the state transition model. Estimates at 20-years were still consistent across all the analysed primary management strategies (e.g., 5% of men primarily managed with RP received GnRH agonists in both observed and simulated data.)
Figure 7. Cumulative incidence of first observed and simulated transition according to primary management strategy in men with prostate cancer in PCBaSe

Cumulative incidence of first observed and simulated transition according to primary management strategy in men with prostate cancer in PCBaSe. Graphs show the cumulative incidence of first observed transition in PCBaSeTraject (continuous line) compared to the cumulative incidence of first simulated transition (dashed lines). AA: anti-androgens; GnRH: gonadotropin releasing hormone agonists; PC: prostate cancer; RP: radical prostatectomy; RT: radiotherapy; WW: watchful waiting.

A good consistency was found for second transitions as well as for transitions to final absorbing states.
Paper II

Figure 8 shows the life expectancy tables derived from the state transition model. For each age, there was a decrease in life expectancy for a fixed CCI, when increasing the DCI and for a fixed DCI, when increasing the CCI.

![Figure 8. Life expectancy based on age, Charlson comorbidity index and drug comorbidity index](image)

The comparison between simulated and observed survival for men based on age at entrance to the cohort, and CCI and DCI status showed consistent results (Figure 9).

As part of the validation process, we compared simulated and observed survival for men based on age, CCI and DCI (Figure 9). No evident systematic differences were found.
Figure 9. Simulated and observed survival for men based according to age at entrance to the cohort, Charlson comorbidity index (CCI), and drug comorbidity index (DCI)
Paper III

Figure 10 shows the time specific prevalence of each state during follow-up time. Transition to radical active treatment was much more common in young men with intermediate-risk PCa, e.g. 76% in men age 55 than in old men with less aggressive PCa, e.g. 24% in men age 70 with very low-risk PCa.

![Graphs showing transitions from various states to death from other causes, prostate cancer death, and active surveillance or watchful waiting for prostate cancer patients at different ages.](image)

Figure 10. Transitions until age 85 for men with Charlson comorbidity index 0 on active surveillance according to age at diagnosis, prostate cancer risk category, and secondary treatment strategy.

We estimated the remaining lifetime spent without active PCa treatment as the ratio between the proportion of the area under the curve for AS and WW in Figure 2 and the total remaining lifetime. We then compared the remaining lifetime spent without active PCa treatment with the risk of PCa death in Figure 11. Men diagnosed at age 60 or below with intermediate-risk PCa had little benefit of AS in terms of time without active PCa treatment as compared to men above age 65 with very low-risk PCa (29-33% vs. 62-77%) and a considerably higher risk of PCa death (12-15% vs 3-5%).
Figure 11. Relationship between prostate cancer mortality and life-time spent without active treatment in men 55 to 70 years old with very low-risk, low-risk, and intermediate-risk prostate

Paper IV

To build the state transition model, we identified two cohorts: (1) men from STHLM-0 or UPSAC who reached the CSPC state (n = 7263 men) and (2) men who transitioned to the CRPC state (n = 3899 men). Figure 12 shows the comparison between the cumulative incidence of observed and predicted transitions to CRPC, showing a good agreement. Transition to CRPC was much more common and rapid for men in the highest CSPC risk category (55% at 2 yr) compared to the lowest risk category (30% at 2 yr). Table 3 details time spent in the CRPC state and proportion of Pca death. During the first 10 yr following castration, time spent in CRPC state varied from 1.1 yr for the highest risk category to 3.9 yr for the lowest risk category. The proportion of men who died from PCa within 10 yr ranged from 93% for the highest risk category to 54% for the lowest risk category. These estimates are in good agreement with observed data, confirming that the internal validity of our model was good.
Figure 12. Cumulative incidence of transition to the castration resistant prostate cancer (CRPC) state for men in the castration sensitive prostate cancer state (CSPC) according to their risk category.

Table 3. Predicted and observed estimates of time spent in castration resistant prostate cancer (CRPC) state and proportion of men dying from prostate cancer (PCa) at the end of follow-up time according to different CRPC risk categories

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Time in CRPC (years)</th>
<th>Proportion of PCa death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed vs Predicted</td>
<td>Observed – Predicted</td>
</tr>
<tr>
<td>CRPC&lt;sub&gt;1-2&lt;/sub&gt;</td>
<td>3.84 vs 3.95</td>
<td>53% vs 54%</td>
</tr>
<tr>
<td>CRPC&lt;sub&gt;3-4&lt;/sub&gt;</td>
<td>3.10 vs 3.14</td>
<td>68% vs 71%</td>
</tr>
<tr>
<td>CRPC&lt;sub&gt;5-6&lt;/sub&gt;</td>
<td>2.07 vs 2.10</td>
<td>83% vs 84%</td>
</tr>
<tr>
<td>CRPC&lt;sub&gt;7-8&lt;/sub&gt;</td>
<td>1.04 vs 1.11</td>
<td>87% vs 93%</td>
</tr>
</tbody>
</table>

**Paper V**

There were 5234 (71%) men diagnosed with prostate cancer in NPCR between 2007-2019 who were registered with watchful waiting as primary treatment strategy and had a life expectancy less than 10 years. An additional 2124 men with a life expectancy between 10 and 12 years were used to enhance precision.
in modelling the relation between life expectancy and outcome. Figure 13 shows disease trajectories for men with life expectancy less than ten years. At five years after diagnosis, the majority (66%) of men with low-risk PCa were still on watchful waiting vs only 36% of those with high-risk PCa. At ten years after diagnosis few men were still on watchful waiting, whereas most patients had already died from other causes with similar proportions across risk groups (57% in low- vs 60% in high-risk). At 20-year follow-up time, transitions to CRPC and PCa death were more common in men with high-risk PCa (13% and 16% respectively) vs low-risk PCa (7% and 10%).

Figure 13. States and state transitions during follow-up for men with prostate cancer managed with watchful waiting
ADT: androgen deprivation therapy
CRPC: castration resistant prostate cancer
WW: watchful waiting

The proportion of men who started ADT, and the proportion of lifetime on ADT, increased non-linearly with increasing life expectancy (Figure 14), as well as the proportion of men who transitioned to CRPC and who died of prostate cancer. For example, in men with low-risk cancer, the proportion who died of prostate cancer death was twice as high for men with 10-year life expectancy compared to men with 5-year life expectancy (13% vs 7%).
Figure 14. Graphical association between oncological outcomes of men on watchful waiting and expected remaining lifetime at diagnosis

ADT: androgen deprivation therapy

CRPC: castration resistant prostate cancer

Life expectancy based on data at date of diagnosis and prostate cancer risk category were associated with earlier receipt of ADT and higher prostate cancer death (Figure 15). Men with low-risk cancer and 10-year life expectancy spent 82% of their remaining life without prostate cancer treatment and with a 13% risk of prostate cancer death within 10 years. In contrast, for a man with high-risk cancer corresponding numbers were 60% and 22%. 
Figure 15. Association between proportion of life-years without receiving any PCa treatment after diagnosis and risk of prostate cancer (PCa) death
Discussion

Paper I

In the first paper a novel state transition model, based on longitudinal data in health care registers including men with PCa, was developed to predict long term disease trajectories. Real-world treatment choices and forecasting of health-related consequences represent a difficult process. At the time of diagnosis, PCa disease trajectories might thus be challenging to forecast due to the long life expectancy, for example > 20 years for healthy men with low-risk PCa [38] and the availability of several therapy choices [9]. Despite the fact that therapeutic alternatives are advised by international guidelines based on illness status, there are a number of combinations of disease states and life expectancy for which there is little evidence to support treatment choices [38]. Furthermore, there are no long-term follow-up data for recently introduced management measures, such as active surveillance, making it impossible to anticipate outcomes. State-transition models have previously been attempted to be used in urology [39,40] and other medical fields [41], mostly from a health-economy perspective. These approaches, however, did not rely on a thorough and population-based data source like PCBaSeTraject. The findings from Paper I offer a proof-of-concept for state transition models that can be used to study a variety of different chronic conditions for which long-term outcome predictions are also required. We were able to create the simulation tool PCBaSeSim, which accurately models treatment changes up to 25 years following diagnosis, using comprehensive nationwide, population-based data of around 120,000 Swedish males with PCa. We compared real-world follow-up data from PCBaSeTraject with simulated data from PCBaSeSim in order to evaluate the reliability of our state transition model. Moreover, we reported consistency between the observed and simulated transitions, indicating that the PCBaSeSim estimations were reliable and accurate. The currently proposed state transition modelling approach importantly differs from conventional survival analyses and prediction tools [42], which focus on a single time-dependent outcome, with little information regarding the actual path that leads to a specific outcome and the changes in the disease trajectory. For this reason, conventional survival analyses inevitably result in a major loss of information and accuracy, especially when dealing with extended follow-up times as it usually happens in chronic diseases. Besides, with PCBaSeSim it is possible to describe disease trajectories during a very long follow-up time and to make
long-term predictions even in the absence of “observed” (i.e. registered) data. Eventually, the flexibility of the model makes it possible to include novel modelling parameters and states. The main limitation of our proposed model is the absence of an external validation. However, PCBaSeSim is ready to be tested in an such a setting, and collaborators are welcome to validate the simulation program. Another limitation is the use of administrative data for the definition of comorbidities, although it was previously shown that the accuracy of ICD codes for discharge diagnoses in the Patient Registry is high in the range of 85–95% [31].

Paper II

In the second paper, age and changes in comorbidity based on the Charlson Comorbidity Index (CCI) and a new Drug Comorbidity Index (DCI) were used to quantify life expectancy at a population level. The observed and simulated survival curves were similar up to 9 years of follow-up for men with higher age, CCI and DCI. The models accurately predicted changes in CCI and DCI. Measurement of comorbidity can be done in several ways [43]. The most often used indicator of comorbidity status is the CCI [22], which was developed with the goal of predicting 1-year mortality and was initially based on 17 medical disorders and their severity. Typically used for discharge diagnosis, CCI has been updated for use with administrative data such as ICD 9 and ICD 10 coding in health care registries. In a similar way, in order to supplement CCI, it was recently developed a prescription-based comorbidity index based on fillings in a Prescription Medication Registry [44,45]. The advantage of the proposed DCI is that it improves CCI's existing forecast capabilities. Since comorbidity measurements were developed to predict the risk of death usually from one to three years, their utility for long-term forecasts is limited. Therefore, a preferred alternative is based on estimated life expectancy. A limitation of our model is that it was developed for men. Although the estimates for females will be different, it is expected that the same basic methodology will be applicable. Another limitation of the current dataset is that prostate cancer was not a comorbidity at baseline, since all men were free of prostate cancer at time of cohort entry. Men diagnosed with prostate cancer during follow-up were, however, considered in the dynamic CCI. Of notice, this methodology can only be applied to geographical regions where similar registries are available as it is key to capture the dynamic process of changes in comorbidities. Future work to improve our proposed methods will involve calibrations based on proportion of deaths from injuries not related to CCI or DCI as well as inclusion of information about CCI and DCI changes prior to the current stage (i.e. allow for non-Markovian properties).
Paper III

In the third paper, the model from Paper I was applied to investigate long term outcomes in men diagnosed with PCa who were managed with AS. The risk of PCa death and remaining life-years without active PCa treatment represented the main study outcomes. Old men with low-risk PCa had the longest time without active treatment and the lowest risk of PCa death, whereas young men with intermediate-risk PCa showed little benefit from being managed with AS. In this context, the longest reported AS follow-up in the literature is approximately 15 years, mostly by institutions that pioneered the idea of AS, such as the Sunnybrook Active Surveillance program, the Gothenburg screening trial, and the Johns Hopkins University School of Medicine AS cohort [46,47]. These studies differ from the study in Paper III in a many significant ways, and share some limitations (i.e. the comparatively small numbers of men followed for longer than 10 years and the short follow-up periods). First, unlike the existing series with long-term data, our simulation was based on population-based data with a capture of 99% of all males diagnosed with PCa in the Swedish Cancer Registry, to which reporting is required by law, whereas the existing series with long-term data began as experimental programs at tertiary referral. This means that it is likely that a share of men included in Paper III would not have been selected for AS at a center of excellence. Besides, since in Paper III the follow-up schedule was much dependent on the treating physician, this introduces the issue of difficult standardization of follow-up schemes, which may eventually result in a lack of adherence. The strengths of our study include the high-quality, population-based data on PCa characteristics, treatments, updated information on comorbidity, and a virtually complete nationwide capture that allowed us to model a 30-year follow-up based on a population-based cohort. There are some other limitations in Paper III. We were unable to calculate PSA density for men diagnosed prior to 2008; moreover, registration in the NPCR did not differentiate between AS and watchful waiting before 2006. In addition, tumor grade inflation phenomenon occurred as a consequence of the International Society of Urological Pathology's modification of Gleason scoring in 2005 [14]. Our sensitivity analyses confirmed that these additional limitations had no impact on the model estimates.

Paper IV

In the fourth paper, the state transition model from Paper I was updated including castration resistant PCa (CRPC) as an additional state. Despite the interest in CRPC owing to the advent of novel drugs for this disease state, there is little information regarding the time spent in the CRCP state in clinical practice. To the best of our knowledge, this is the first report on disease
trajectories using a population-based cohort of men on ADT who transition to CRPC that includes the time spent in the CRPC state. The comprehensive follow-up and the high-quality data from the population-based registers used in the study. The survival observed in our study, ranging from 1 to 4 yr and varying according to the risk category, is in line with a previous report based on clinical practice data from The Netherlands [48]. Our study has some limitations. The introduction of the CSPC state, which could only be reached if PSA measurement occurred while the patient was exposed to GnRH, allowed us to solve the absence of data on testosterone levels. Moreover, there may be regional variations in the PSA measurement patterns between the two cohorts (STHML-0 and UPSAC). The lack of information on the occurrence of metastases is also a drawback, since men with nonmetastatic CRPC live longer than men with metastatic CRPC [49]. Another limitation is the lack of information regarding chemotherapy, which is not usually recorded in the Prescribed Drug Register. Regarding the performance of the model, we showed that it reliably predicts both transitions to CRPC and death. The use of a combination of PSA at transition to CRPC and PSA doubling time (PSA DT) to model the risk of death for men with CRPC is in accordance with pre-existing evidence associating both PSA at the date of CRPC and shorter PSA DT with shorter survival [50]. We argue that the combination of PSA at the date of CRPC and PSA DT results in a clear proxy for disease aggressiveness, as shown by the different outcomes for men according to risk category.

Paper V

In the fifth paper, the updated state transition model from Paper IV was applied in order to model long term outcomes for men with PCa and managed with watchful waiting (WW). Since WW is currently recommended by the EAU guidelines as a management strategy for men with PCa and life expectancy < 10 yr, we used the state transition model from Paper II in order to estimate life expectancy at diagnosis for all included men. In this nationwide, population-based cohort of men with prostate cancer, cancer characteristics at date of diagnosis and life expectancy were strongly associated with life-time spent without ADT, time spent on ADT, and risk of PCa death. We observed that the majority of men with high risk PCa and longer life expectancy received ADT before dying, and one out of four of those men eventually developed CRPC and died of PCa. As compared to men with low-risk PCa, men with high-risk PCa had a shorter remaining lifetime spent without ADT and a two-fold higher risk of prostate cancer death. Several trials and observational studies have reported the outcome for men with prostate cancer managed with watchful waiting [51]. In a frequently cited Swedish study of untreated early stage prostate cancer diagnosed in elderly men during the late seventies, most cancers had an indolent course during the first 10-15 years and most men died
of other causes [9]. However, when survival exceeded 15 years, prostate cancer mortality rose from 15/1000 person-years to 44/1000 person-years. Similarly, a cohort from the US [52] showed that in men with low-risk disease prostate cancer was not the most common risk of death. The potential loss in life years should be balanced against the risk of decreased quality of life brought by treatment initiation [53]. ADT is well known for its metabolic side effects, not only impairing quality of live, but also increasing comorbidity burden and possibly reducing life expectancy [54]. For instance, a man diagnosed with high-risk PCa and 10-year life expectancy, will spend half of his remaining lifetime on ADT. In this context, the key assumption is the proper estimation of life expectancy at diagnosis. This study is the first major nationwide, population-based study to investigate long-term outcomes of patients managed with watchful waiting in clinical practice. It was only possible to perform this analysis using the state transition methodology from Paper I and Paper IV for the estimation of prostate cancer disease trajectories, whereas the state transition methodology from Paper II was used for the remaining lifetime estimation. We were able to construct a 20-year follow-up using a population-based cohort thanks to the high-quality of the available data, including prostate cancer characteristics, treatments, and time-updated measures of comorbidity. The study has some limitations. The reason for the transition from watchful waiting to ADT is not recorded in the National Prostate Cancer Register or in any other register. This transition was therefore captured indirectly through a previously described algorithm [28]. There are, however, Swedish guidelines with detailed criteria for watchful waiting and instructions for follow-up. Adherence to these guidelines was recently shown to be good [55,56]. The definition of watchful waiting in the National Prostate Cancer Register is consistent with international guidelines [57]. Another limitation is the potential misclassification of the cause of death in older men [58]. It was recently shown in a patient record review that little evidence for prostate cancer progression could be found in ~30% of men >85 years who were coded as dead from prostate cancer in the National Cause of Death Register. This may overestimate the proportion of prostate cancer deaths and explain why a some of these deaths occur before a transition to castration resistant prostate cancer. The misclassification of cause of death could potentially also indirectly bias the algorithm used to identify state transition to castration resistant prostate cancer.
Ethical considerations

Data collection and registration in NPCR

Data on men diagnosed with prostate cancer have been collected and registered in the National Prostate Cancer Register (NPCR) of Sweden since 1998. Data in NPCR are handled according to the General Data Protection Regulation (GDPR). As in all Swedish clinical cancer registers, the “opt-out” approach was applied in NPCR, i.e. men with prostate cancer are informed that they are included in NPCR unless they actively express to be excluded. At the time of diagnosis and treatment for prostate cancer, all men receive information that their data will be included in NPCR and that non-participation is available but there is no collection of written informed consent. Information in Swedish is available via the NPCR web site on the rights of registered men and the legal requirements for registration and the responsible authorities (https://cancercentrum.se/samverkan/vara-uppdrag/kunskapsstyrning/kvalitetsregister/registratorades-rattigheter/).

Data enrichment in PCBaSe

Data from NPCR were then used in Prostate Cancer data Base Sweden (PCBaSe) version 4.0, a research project that was approved by the Research Ethics Board in Uppsala (Dnr 2016-239). Data in NPCR, including the personal identity number, were transferred to Statistics Sweden (SCB) and The National Board of Health and Welfare (SoS) and cross-linked with data from other health registers and demographic databases in order to create the database for PCBaSe 4.0. In addition to cases from NPCR, six men, free of prostate cancer, matched on birth and county of residence, were selected at Statistics Sweden. Each case and each comparator are assigned a unique code number that is kept as the sole identifier in files exported outside of SCB and SoS and a code key is kept at SoS for five years. In Paper 4 and 5, additional data were used and this amendment approved by the Research Ethics Authority, Stockholm (Dnr 2020-02773).
Risk assessment

The studies of this thesis include previously documented data only i.e. no additional diagnostic or laboratory examinations were performed, and no direct contact was made with the study men. The study men are under no risk of physical harm; however, personal integrity is inevitably at risk as in all clinical cancer research. To minimize the risk of violating personal integrity, information on personally identity is only kept by institution with high level of data security. The personal identity number has been replaced and only necessary variables are included in the study files of this thesis. Furthermore, the data is only presented on a group level, not individually. We therefore argue that the scientific value of our research outweighs the minimal risk of violating personal integrity.

Ethical approvals

The Regional Ethical Review Board of Uppsala University approved the studies in this thesis. Paper I, II, and III: Dnr 2016/239 PCBaSe 4.0 Regionala etikprövningsnämnden Uppsala, PI Pär Stattin. Paper IV and V: Dnr 2016/239 PCBaSe 4.0 Regionala etikprövningsnämnden Uppsala och Dnr 2020-02773 Etikprövningsmyndigheten Stockholm avd 3 medicin ”Inventering av PSA provsvar samt resultat av nålprovtagning från prostata”, PI Hans Garmo.
Conclusions

The goal of this thesis consisted in the development, application, and validation of novel state transition models, based on longitudinal data in health care registers including men with PCa and PCa-free controls, to predict long term disease trajectories in chronic diseases. State transition models differs in a significant way from conventional survival analyses and prediction tools, which usually focus on a specific outcome over time, with little information regarding the actual path that led to a specific outcome and the changes in the disease trajectory along the same path [42]. Conventional survival analyses inevitably result in a consistent loss of information and accuracy, especially when dealing with extended follow-up times, as it may happen in chronic diseases. When modeling disease trajectories, one must also consider changes in comorbidities and life expectancy. In this thesis, we propose an additional state-transition model for changes in comorbidities based on Charlson comorbidity index (CCI) and drug comorbidity index (DCI) changes. It is possible to model and accurately predict temporal changes in disease trajectories over a very long-term follow-up time. Aside from clinical implications, our models provide information applicable to healthcare resources allocation. Since it is possible to keep track of specific disease trajectories including the mean time spent in each state, estimates for costs can be obtained by use of our models.
Acknowledgements

**Pär Stattin** – Professor of Urology and main supervisor of this PhD project. Pär guided me since my arrival to Sweden and made it possible to start this amazing research experience. I am proud to say that during these years I grew up from the scientific standpoint thanks to your guidance, also learning from my mistakes that you were ready to point out in a very respectful and constructive way. Thank you for showing me the welcoming attitude of the Swedish people and for tolerating my attitude to deliver at the very last minute.

**Hans Garmo** – Senior statistician and co-supervisor. Always coming up with the most brilliant, insightful, and helpful suggestion whenever a problem was in sight. Always ready to help as a mentor, but especially as a friend. Your teachings I will always keep with me: R-coding and statistics, chantarelle mushroom picking, how (not) to ice skate, how to swing a golf club, and the secrets for a flawless trip to Dalarna.

**Francesco Montorsí** – Professor of Urology, co-supervisor and Head of Department at San Raffaele Hospital, Milan, Italy. For allowing me to work on this thesis and for always supporting me.

**Lars Lindhagen, Rolf Gedeborg, Marcus Westerberg** – Thank you for building together, one brick at a time, all the state transition models. It was a pleasure to collaborate with you.

**The members of the steering group of NPCR** – For their efforts in creating such a high-quality collection of data, and for allowing me to use it in my research.

**Kristina and Manfred** – Or, as I call them, my “Swedish parents”. Thank you for welcoming me in your house and for taking care of me.

**Cinzia and Giovanni** – For your tireless efforts and support.
**Lucia, Maria Vittoria, and Angelo** – My starting point, where I know I can always come back, no matter what.

**Marta and Anna** – Simply for being there. For tolerating all the nights and weekends I spent at work. For your love and unconditional support. For enjoying this Swedish experience all together, always as a family.
References


[38] European Association of Urology. EAU Guidelines - Prostate Cancer 2022.


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)