From Department of Neurobiology, Care Sciences and Society
Karolinska Institutet, Stockholm, Sweden

REAL-WORLD OUTCOMES IN ONCOLOGY: ASSESSING THE VALUE OF REWARDABLE INNOVATION

Nahila Justo

Stockholm, 2022

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-617-1
Cover illustration: Shutterstock. Photo ID: 1660588972

Real-World Outcomes in Oncology
Assessing the Value of Rewardable Innovation

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Nahila Justo

The thesis will be defended in public at Stockholm, 19 August 2022, 9:00am.

Principal Supervisor:
Linus Jönsson
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Neurogeriatrics

Co-supervisor:
Nils Wilking
Karolinska Institutet
Department of Oncology-Pathology

Opponent:
Richard Sullivan
King's College London
King's Institute of Cancer Policy
Conflict and Health Research Group

Examination Board:
Niklas Zethraeus
Karolinska Institutet
Department of Learning, Informatics, Management and Ethics
Division for Medical Management Centre

Lars-Åke Levin
Linköping University
Department of Health, Medicine and Caring Sciences
Division of Society and Health

Johan Ahlgren
Örebro University
Faculty of Medicine and Health
Dept of Oncology

I would like to start by thanking Linus, my main supervisor. You have perfectly balanced the right amount of empowering when I could manage on my own, with the indispensable handholding (or even dragging) when I couldn't. Thanks also for being so generous with your knowledge and experience, since the first conversation about the crazy idea of going back to school, through today, finally crossing the finish line. To Nils, thank you also for sharing your knowledge, but even more for the encouragement. Your compliments of my progress were very timely, and highly appreciated.

I would also like to thank my manager, Anne Delaney. Without your support, I could not have made it. You kept your promise, even in difficult circumstances and that is a testimony of your character. Monique Martin and Jeanni Van Loon, you went the extra mile to make it possible a few years back and I am grateful.

To all my co-authors, thank you for your sharing your knowledge, your keen insights, and your critical thinking. Special thanks to Johan Dalen, Bernd Schweikert, Ana Baños Hernaez, Jonas Nilsson, and Prof. Michael Drummond, with whom I had the opportunity to share much more than just the work on these studies.

To my colleagues at Evidera, who inspire me to always want to learn more and do better. I learn from you every day. I will not miss my studies, because I have you.

To my family, Rickard, Leandro, Juan, and Marina. Thank you for believing in me and for understanding my choices. And thank you to the extended family Patrik, Ana, Peter, Caroline, Mariela, Viviana, Sonia, Annachiara, and Milena for always being there for me.

Y para ustedes mis amores, Astor y Bianca. Espero que esta travesura de mamá les contagie la curiosidad, las ganas de estudiar, y la liviandad de sentirse cómodos diciendo 'no sé' con la convicción de que no hay nada que no puedan aprender. Los amo.

POPULAR SCIENCE SUMMARY OF THE THESIS

New medicines are granted preferential terms and pricing if they demonstrate that they are innovative. Innovativeness is more than newness; it implies added therapeutic benefit beyond the existing treatment options. In some countries, the demands are higher, and they need to demonstrate additional benefits. These additional benefits of rewardable innovation may be delivered in several areas such as help in indications with high unmet needs or very severe, tend to underserved populations or at the end of life, be more convenient for patients or their caregivers, generate economic benefits, constitute a step-change in the management of the disease, or bring about more innovation. Our ever-growing understanding of cancer has translated into the development of many new therapies that claimed to be innovative. We investigated some of those innovations introduced in the treatment of lung, prostate, and breast cancer.

The first study evaluated the progression of survival of patients with non-small cell lung cancer (NSCLC) in the US between January 1973 and December 2012 in relation with the introduction of new drugs. Based on data from SEER-Medicare, we learned that the proportion of patients still alive one year after the diagnosis almost doubled but remained low (39%). We also found that 38 different therapies were used to treat these patients, but most of them had not been approved in this indication where, until 2012, innovation had been limited. We also studied patients with this disease in Sweden as they compare with the general population, based on data from national registries. We confirmed that, already at diagnosis, patients with NSCLC are burdened with many other concomitant diseases. In the year following NSCLC diagnosis, they also got significantly more and newer ailments, some a consequence of the cancer progression and/or the shared risk factor, and others related to the treatments. Patients with NSCLC were in poor health already before and got much worse after. Many more new treatments were approved for these patients after the end of our study period and soon it will be possible to evaluate their innovativeness.

In the second study we investigated new treatments for metastatic castration-resistant prostate cancer (mCRPC) used in Germany between January 2013 and December 2015. Until recently, patients with mCRPC had very limited treatment options if their disease had progressed despite surgical or chemical castration; only docetaxel was used. A similar chemotherapy, cabazitaxel was approved in 2011 and then came two innovative oral drugs, abiraterone and enzalutamide. We observed the disease at presentation and treatments used by 447 patients with mCRPC. Since most patients were still alive at the end of our study period, we could not assess their impact of these drugs on their prognosis; but we found that they were quickly incorporated into their routine care. While doctors tended to start younger patients with the old docetaxel, before resorting to the newer drugs; Abiraterone was the most frequently used front therapy and they lasted for significantly longer on treatment than with any of the other drugs. When looking at the sequence in which they were used, we found over 70 distinct treatment pathways, which suggests that doctors are tailoring the care of these patients according to their needs or preferences, rather than following a uniform recipe.

The last study evaluated the benefits delivered by trastuzumab to patients with early or metastatic HER2+ breast cancer, from January 2000 (when it was first approved) and December 2021, in Sweden. Based on a combination of estimates of clinical effects from clinical trials with real-world data from the national registries. We also used results of observational studies in Sweden to gather the necessary information on costs, quality of life (translated in utilities), productivity losses, and caregiver burden. With all these different pieces of information, we built two models (one in early and the other in metastatic disease) to estimate the value delivered over 20 years. More than 15,000 patients have been treated with trastuzumab, which meant that Sweden gained 25,844 life-years that would have been lost without this intervention. If we adjust these years of life gained by their quality of life, the gain was 13,437 quality-adjusted life-years. According to Swedish guidance in the evaluation of health technologies, the monetary value of these gains is equivalent to of 8.7 trillion SEK.

Overall, we can conclude that while some new drugs have delivered value, their degree of innovativeness varied, as did their uptake. The type of secondary data that we used was appropriate to evaluate most of the value-generating attributes mentioned but, for clinical effects, evidence emanating from clinical trials remains crucial.

ABSTRACT

Recent years have seen a remarkable expansion of therapeutic options available for many forms of cancer. Evidence for the efficacy of these treatments have mainly come from randomized controlled trials, however questions often remain regarding the actual use, effectiveness and value of innovative therapies when used under the circumstances of routine clinical practice. In this thesis, we aim to assess the value and contribution of new oncology treatments for common cancers (lung, prostate, and breast) in early and/or advanced stages, based on data generated under 'real world' conditions in routine care.

In paper I, we conducted a systematic review and mapping of the availability of real-world data (RWD) and use of evidence generated from such data (RWE), with focus on four South American countries. Findings were validated through workshops with regional experts. We identified 407 unique databases, and reported details included geographic scope, database type, population, and outcomes captured. The quality of RWD varied across countries, and we found that RWE was not consistently used to inform health care decision making. The main use of RWE was for pharmacovigilance studies, and to lesser extent for health technology assessment and for pricing decisions.

In Paper II, we investigated therapeutic innovation in the care of patients diagnosed with advanced or metastatic non-small cell lung cancer (NSCLC) in the US between 1991 and 2012. Based on data from SEER-Medicare, we examined the association between the degree of innovation (measured as an innovation index or mean medication vintage) and overall survival. Results indicated that therapeutic innovation was associated with only a slightly improved 1-year survival (odds ratio (OR): 1.05 [95% confidence interval (CI): 1.04–1.05]).

Paper III described the occurrence of comorbidities in patients with NSCLC based on national registry data from Sweden during 2006–2013. Comorbidities that may be associated with prognosis, disease progression or share risk factors with NSCLC were identified and assessed before and after the NSCLC diagnosis. 3,834 NSCLC patients were compared with 15,332 matched controls. The comorbidity prevalence at baseline was significantly higher in NSCLC patients with an OR of 2.44 (95% CI: 2.27–2.63), and the incidence rate ratio (IRR) of newly diagnosed comorbidities during the year after diagnosis was 32.5 (95% CI: 31.0–34.2).

In Paper IV, we described treatment patterns in patients with metastatic castration-resistant prostate cancer (mCRPC) based on health insurance data from Germany for the period January 2013 to December 2015. 447 patients were continuously enrolled for 12 months before being started on treatment with abiraterone, cabazitaxel, docetaxel, or enzalutamide. Over 70 distinct treatment pathways were identified. Abiraterone was the most commonly prescribed while cabazitaxel was the least commonly prescribed therapy. Abiraterone patients also had longest treatment duration.

Paper V aimed to estimate the life-cycle value of trastuzumab for early (EBC) and metastatic (MBC) breast cancer in Sweden. Aggregate data on trastuzumab-treated patients from national registries was combined with data from RCTs and economic studies in Markov models to estimate overall survival, lifetime costs, and quality-adjusted life years (QALYs). Over 15,000 patients have been treated with trastuzumab, generating 25,844 life-years and 13,437 QALYs gained, at a monetary value of 8.7 trillion SEK.

In conclusion, based on RWD, we found that innovative oncology therapies have delivered value in the care of patients with advanced or metastatic NSCLC, metastatic CRPC, and early or metastatic HER2+ BC, and other based on RWE over the past decades. However, this was not true for all new medicines introduced and the benefit derived from their use was not uniform. RWE can support value assessment of innovation, mainly in dimensions beyond therapeutic benefit.

LIST OF SCIENTIFIC PAPERS INCLUDED IN THESIS

I *Real-World Evidence in Healthcare Decision Making: Global Trends and Case Studies From Latin America*. **Justo N**, Espinoza MA, Ratto B, Nicholson M, Rosselli D, Ovcinnikova O, García Martí S, Ferraz MB, Langsam M, Drummond MF. Value Health. 2019 Jun;22(6):739-749.

II *Retrospective observational cohort study on innovation in oncology and progress in survival: How far have we gotten in the two decades of treating patients with advanced non-small cell lung cancer as a single population?*
Justo N, Nilsson J, Korytowsky B, Dalen J, Madison T, McGuire A. PLoS One. 2020 May 12;15(5):e0232669

III *Comorbidities and relevant outcomes, commonly associated with cancer, of patients newly diagnosed with advanced non-small-cell lung cancer in Sweden.* Linden S, Redig J, Banos Hernaez A, Nilsson J, Bartels DB, **Justo N**. Eur J Cancer Care (Engl). 2020 Jan;29(1):e13171

IV *Insights into treatment patterns in the routine care of patients diagnosed with advanced castration-resistant prostate cancer in Germany after the introduction of innovative new therapies*. **Justo N**, Schweikert B, Simon A, Waldeck AR, Meinhardt M, Samel YR, Goebell PJ. Clinical Oncology & Research. 2020 Sept; 3(9): 2-8. DOI: 10.31487/j.COR.2020.09.04

V *Determining the lifecycle value of trastuzumab based on registry data in Sweden.* **Justo N**, Wilking N, Jonsson L. Unsubmitted manuscript

CONTENTS

LIST OF ABBREVIATIONS

ACA	Affordable Health Care Act
AIFA	Italian Medicines Agency
AMNOG	Act on the Reform of the Market for Medical Products
aNSCLC	advanced or metastatic non-small cell lung cancer
ASMR	Amélioration du Service Médical Rendu
BC	breast cancer
BH3	Bcl-2 homology 3
CBA	cost–benefit analysis
CCI	Charlson Comorbidity Index
CEA	cost-effectiveness analysis
CEESP	Commission d'Évaluation Économique et de Santé Publique
CI	confidence interval
CMA	cost-minimisation analysis
CMS	Centers for Medicare and Medicaid Services
COPD	chronic obstructive pulmonary disease
CSDD	Center for the Study of Drug Development
CTLA-4	cytotoxic T-lymphocyte antigen 4
CUA	cost-utility analysis
DALY	disability-adjusted life year
DARWIN EU	Data Analysis and Real World Interrogation Network
DRG	diagnoses-related group
EBC	early breast cancer
EGFR	epidermal growth factor receptor
EHDS	European Health Data Space
EHR	electronic health record
EMA	European Medicines Agency
EMR	electronic medical record
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
FDA	Food and Drug Administration
GBA	Federal Joint Committee
GDPR	General Data Protection Regulation
GFL	Gesundheitsforen
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HAS	French High Authority for Health
HER2	human epidermal growth factor receptor
HGF	hepatocyte growth factor
HR	hazard ratio
HRQoL	health-related quality of life
HTA	health technology assessment

HYE	healthy-year equivalent
ICBP	International Cancer Benchmarking Partnership
ICER	Incremental cost-effectiveness ratio
ICD-O	International Classification of Diseases for Oncology
IDN	Integrated delivery networks
IF	Innovation Frontier
ILD	interstitial lung disease
INCA	Information Network for Cancer care
InnovInd	Innovation Index
IP	intellectual property
IPL	individual patient-level data
IQWiG	Institute for Quality and Efficiency in Health Care
IR	incidence rate
IRR	incidence rate ratio
ISPOR	International Society for Pharmacoepidemiology and Outcomes Research
IV	instrumental variable
KM	Kaplan-Meier
LVEF	left ventricular ejection fraction
mAb	monoclonal antibody
MBC	metastatic breast cancer
MCDA	multi-criteria decision analysis
mCRPC	metastatic castration-resistant prostate cancer
MHRA	Medicines and Healthcare products Regulatory Agency
MI	myocardial infarction
MMV	mean medication vintage
MR	mortality rate
NCI	National Cancer Institute
NCR	National Cancer Register
NICE	National Institute for Health and Care Excellence
NIS	non-interventional study
NKBC	Surveillance, Epidemiology, and End Results
NSCLC	non-small cell lung cancer
OR	odds ratio
OS	overall survival
PAES	post-authorisation efficacy study
PARP	poly adenosine diphosphate ribose polymerase
PASS	post-authorisation safety studies
PCORI	Patient Centered Outcomes Research Institute
PIN	personal identification number
PRO	patient-reported outcome

PROM	patient-reported outcome measure
QALY	quality-adjusted life year
QoL	quality of life
RCT	randomised controlled trial
RDP	regulatory data protection
RMST	restricted mean survival time
RWD	real-world data
RWE	real-world evidence
SD	standard deviation
SE	standard error
SEER	Surveillance, Epidemiology, and End Results
SEK	Swedish krona
SHCR	Skåne Health Care Register
SMR	Service Médical Rendu
SURVMARK-2	Cancer Survival in High-Income Countries
TLV	Dental and Pharmaceutical Benefits Agency
TNM	tumour, node, metastasis
US	United States
VEGF	vascular endothelial growth factor
WHO	World Health Organization
WTP	willingness-to-pay
ZIN	National Health Care Institute

1 INTRODUCTION AND LITERATURE REVIEW

1.1 THE INEVITABILITY OF CANCER AND ITS EVER-GROWING BURDEN

Robert Weinberg said in his seminal book *The Biology of Cancer* that "cancer is an inevitability; if we succeeded in avoiding the death traps set by all the other usual diseases, sooner or later most of us would become victims of cancer" (Weinberg 2014). This inevitability is a driver of the ever-growing magnitude of cancer burden as the world population grows and ages. The latest report published by the International Agency for Research on Cancer, GLOBOCAN 2020, estimates that 19.3 million people are diagnosed with cancer each year worldwide, and 10 million die as a consequence of the disease. Cancer represents the first- or second-leading cause of premature death in two-thirds of the countries worldwide. The increasing prevalence of risk factors in high-income countries, paired with the economic and epidemiologic transition in low-income regions, has resulted in cancer burden that is expected to increase by 50% by 2040 (Sung, Ferlay et al. 2021).

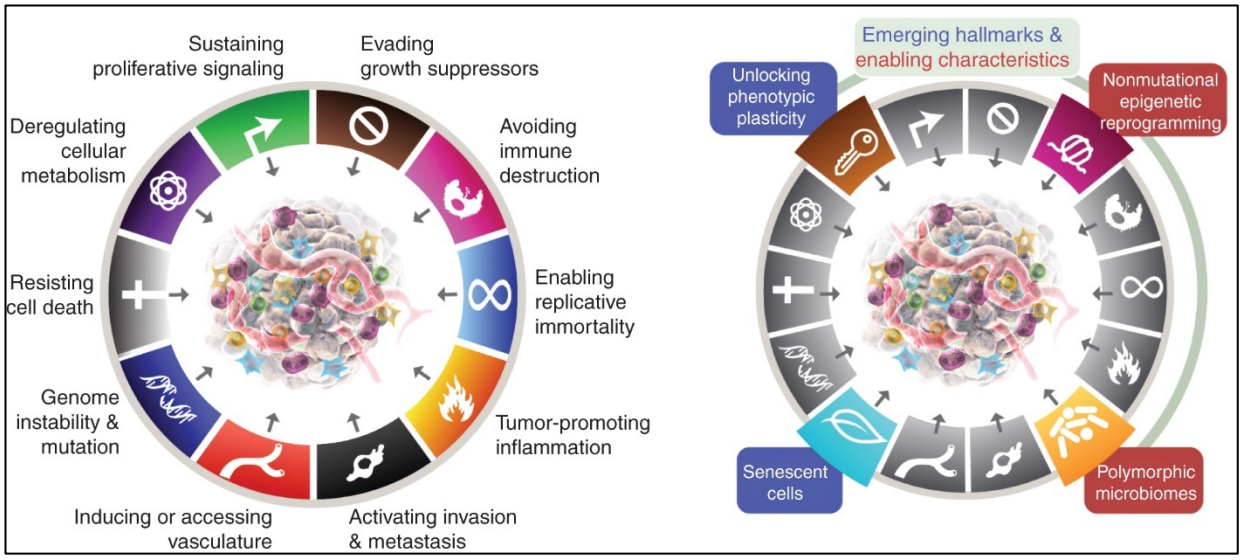
Results of the second phase of the Cancer Survival in High-Income Countries (SURVMARK-2) project conducted by the International Cancer Benchmarking Partnership (ICBP) suggest that cancer survival continues to increase in these countries, despite geographical disparities and across tumour types (Arnold, Rutherford et al. 2019). This longer survival compounds the effect of growing incidence rates on the prevalence of the disease, thus multiplying the cost of treatment for patients, and of maintenance and follow-up for survivors. The most recent European study estimated that the total cost of cancer in 2018 (including the 27 member states of the European Union, Iceland, Norway, Switzerland, and the United Kingdom) was €199 billion and healthcare expenditure accounted for about half (Hofmarcher, Lindgren et al. 2020). Similarly, the United States (US) National Cancer Institute projected total direct medical costs for cancer care to be $208.9 billion in 2020 and reported that the total economic burden borne by patients with cancer in 2019 was $21.1 billion (including out-of-pocket and time costs) (Mariotto, Enewold et al. 2020, Yabroff, Mariotto et al. 2021, National Cancer Institute July 2021). This heavy clinical, epidemiological, and economic burden further highlights the importance of the cancer conundrum.

1.2 CANCER AND ITS CURE, THE ELUSIVE ENIGMA

Cancer predates humans; paleontological research has traced cancer in pre-historic species, though the first record of cancer was in Edwin Smith Papyrus (3000 BC). Since then, all civilisations, have registered descriptions of the disease and their research, but it was only in the 19th century that the grounds to understand cancer started to emerge. Important milestones are the initial hypotheses' formulation on the hereditary nature of cancer and the inception of research on metastasis (1829), the introduction of cell theory (1838), the discovery of X-rays (1895), and the establishment and general spread of surgical pathology. The 20th century delivered robust foundations for experimental and clinical research, myriad diagnostic tools (from devices to screening campaigns, and from classification systems to

analytic capabilities), and the discovery and development of anticancer drugs with varying degrees of success across cancer types. In the words of cancer historian Steven Hajdu “the 25 years from 1970 and 1995 are the high-water mark in clinical oncology, and this is the period when oncology turned from art to science” (Hajdu 2011, Hajdu 2012, Hajdu 2012, Hajdu and Darvishian 2013, Hajdu and Vadmal 2013, Hajdu, Vadmal et al. 2015). Since the turn of the century, our knowledge of the disease has been fundamentally transformed by the molecular characterisation of numerous cancer genomes (Stratton, Campbell et al. 2009). Today, we know that cancer is a collection of diseases characterised by the persistent accumulation of genetic and epigenetic changes in cells replicating uncontrollably, and it can occur anywhere in the body. Whether triggered by inherited and/or acquired factors, oncogenesis is a process that involves several steps that enable those uncontrollably replicating cells to evade external growth-promoting or growth-inhibiting influences, avoid programmed cell death (apoptosis), recruit blood vessels (angiogenesis), and disseminate (metastasis) (Lyman 2009, Spira, Yurgelun et al. 2017).

Figure 1. The Hallmarks of Cancer



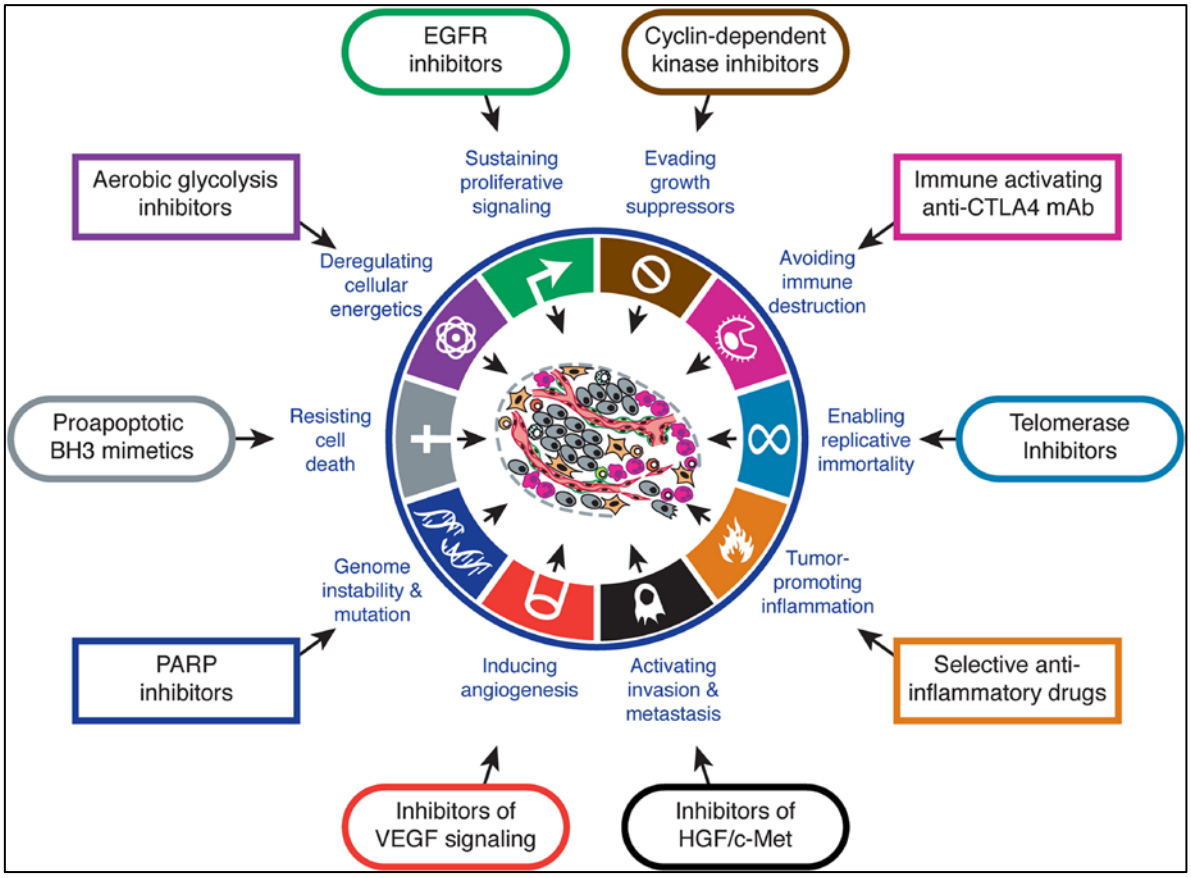
Reprinted from Cancer Discovery 2022;12(1):31-46, Douglas Hanahan, Hallmarks of Cancer: New Dimensions. Reproduced with permission from AACR [License Number 5346120319223]

Hanahan and Weinberg summarised the biological capabilities acquired by cancer cells in the development of tumours and metastasis in their pivotal enunciation of the Hallmarks of Cancer. The saga starts in the year 2000 with the six original hallmarks, i.e., proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis (Hanahan and Weinberg 2000). In 2011, the authors incorporated two enabling characteristics, genome instability and mutation and tumour-promoting inflammation, two emerging hallmarks reprograming energy metabolism and evading immune destruction (Hanahan and Weinberg 2011). More recently, after validating the emerging hallmarks as part of the core set, the authors incorporated the following new emerging hallmarks and enabling characteristics: unlocking phenotypic plasticity, non-mutational epigenetic reprogramming, polymorphic microbiomes, and

senescent cells (Hanahan 2022). Figure 1 presents an overview of the last two generations and serves as a clear illustration of the piecemeal nature of our understanding of the disease.

This evolution of our understanding of the disease has been driven mainly by the aspiration to find a cure, or at least, a treatment that reduces and/or delays the negative consequences of cancer. Every step forward taken by translational and clinical research have resulted in an innovative therapeutic opportunity. At first, surgical pathology, radiotherapy, and chemotherapy were promising innovations. Then, came biologics and new targeting agents. Most recently, cell and gene therapy hold promise for further improvement in outcomes. Hanahan and Weinberg described the therapeutic targeting enabled by discoveries in molecular biology and provided several examples in the graph reproduced as Figure 2 (Hanahan and Weinberg 2011). Yet, not all these innovations have delivered on their promise or, at least, some have delivered more benefits than others.

Figure 2. Therapeutic Targeting of the Hallmarks of Cancer



Reprinted from Hallmarks of Cancer: Next Generation. Douglas Hanahan and Robert Weinberg. Cell. 2011 Mar 4;144(5):646-74. Reproduced with permission from Elsevier [License Number 5340161054607]

Abbreviations: BH3 = Bcl-2 homology 3; CTLA-4 = cytotoxic T-lymphocyte antigen 4; EGFR = epidermal growth factor receptor; HGF = hepatocyte growth factor; mAb = monoclonal antibody; PARP = poly adenosine diphosphate ribose polymerase; VEGF = vascular endothelial growth factor

Several studies have linked progress in dropping mortality with the improvement in healthcare delivery attained thanks to the introduction of medical innovations. Particularly in oncology, where innovation has tended to come in waves, indication-specific studies have assessed survival and other outcomes before and after new treatments had been introduced (Lindskog, Wahlgren et al. 2017, MacEwan, Yin et al. 2017, Maiese, Evans et al. 2018, Justo, Nilsson et al. 2020, MacEwan, Majer et al. 2021, Ramagopalan, Leahy et al. 2021). Other studies that have also leveraged geographic variability apart from the longitudinal dimension, confirmed this connection (Lichtenberg 2014, Lichtenberg 2019, MacEwan, Dennen et al. 2020, Lichtenberg 2022). Moreover, the spill over effect of healthcare innovation in high-income countries, where innovation originates, is credited to have also driven improvements in longevity in poorer countries, in spite of the problems with affordability of new therapies (Khullar, Fisher et al. 2019). The variability in methods and specifications of the concept of innovation in these and other studies begs the question, what is innovation and how do we measure it?

1.3 DEFINING AND MEASURING INNOVATION IN ONCOLOGY AND BEYOND

1.3.1 More Than Just New

Thomas Edison said that "the value of an idea lies in the using of it," and this applied dimension of innovation is the one upon which most definitions used in healthcare agree. In general, the concept of innovation is associated with more than just "newness". As opposed to invention (that implies newly discovered molecular entities), innovation is typically considered in the context of its applications (Belloso 2020, Hofmann, Branner et al. 2021).

The World Health Organization (WHO) defines health innovation "as a new or improved solution with the transformative ability to accelerate positive health impact" (World Health Organization 2022). But, while added benefit is a common element of the definition of innovation across regulators, health technology assessment (HTA) agencies, and other endorsement bodies, there is no consensus on what constitutes added benefit (de Solà-Morales, Cunningham et al. 2018). For example, Prof. Di Masi, from the Tufts Center for the Study of Drug Development (CSDD), argued that new-use Efficacy Supplement approvals for secondary indications also constitute innovations (DiMasi 2013), but others disagree.

1.3.2 Rewardable Innovation

In 2007, a group of experts from different countries gathered in a Drug Innovation Workshop to develop the so-called Erice Statement: "An innovation in the field of medicinal products consists of a completely or partially new active substance or biological entity or combinations of such entities acting against a disease, relieving symptoms or preventing a disease through pharmacological or molecular mechanisms, and developed and made available as a medicinal product that can improve the quality of patient management and outcomes. The present definition of drug innovation may also include new indications, technological and manufacturing processes, new formulations (including combinations) and delivery systems of known drugs." The experts also highlighted the distinction between the meaning of

innovation and the value of a specific health intervention (Erice statement on drug innovation 2008). This distinction will be examined in the context of the discussion of policies and measures to incentivise and reward pharmaceutical innovation, as there is a risk for “double counting” benefits.

Soon after the Erice Statement, the National Institute for Health and Care Excellence (NICE) in England, commissioned a study on the value of innovation and other benefits that resulted in the so-called Kennedy report. Upholding the principles of cost-effectiveness, the Kennedy report acknowledges that, a higher price can be claimed when an innovation goes beyond than just newness and improvement over existing products; when it also “offers something more: a step-change in terms of outcomes for patients”. Furthermore, the report also advises that clear and measurable criteria should be set to define “step-change” for transparency, and to avoid double-counting benefits (Kennedy 2009).

This connection between the degree of innovativeness and premium pricing was adopted by NICE and other HTA bodies (see Section 1.4.2) and remained a topic for discussion in the literature until this day. Jeffrey Aronson and colleagues coined the term “rewardable innovation” and discuss a list of features such as newness, novelty, usefulness, cost-effectiveness, and the origin of the innovation (revolutionary vs. evolutionary (Aronson, Ferner et al. 2012).

In the report of the International Society for Pharmacoepidemiology and Outcomes Research (ISPOR) Task Force, the authors specify 12 such elements with the conceptual background theory and measurement approach for each of them. These are QALYs, severity of disease, net and opportunity costs, labour productivity, adherence-improving factors, risk and fear of contagion, reduction in uncertainty due to a new diagnostic, insurance value, value of hope, real option value, equity, and scientific spill overs (Lakdawalla, Doshi et al. 2018). Figure 1, reproduces the schematic version of the elements of value proposed by the Task Force.

Figure 3. Elements of Value according to the ISPOR Task Force Report

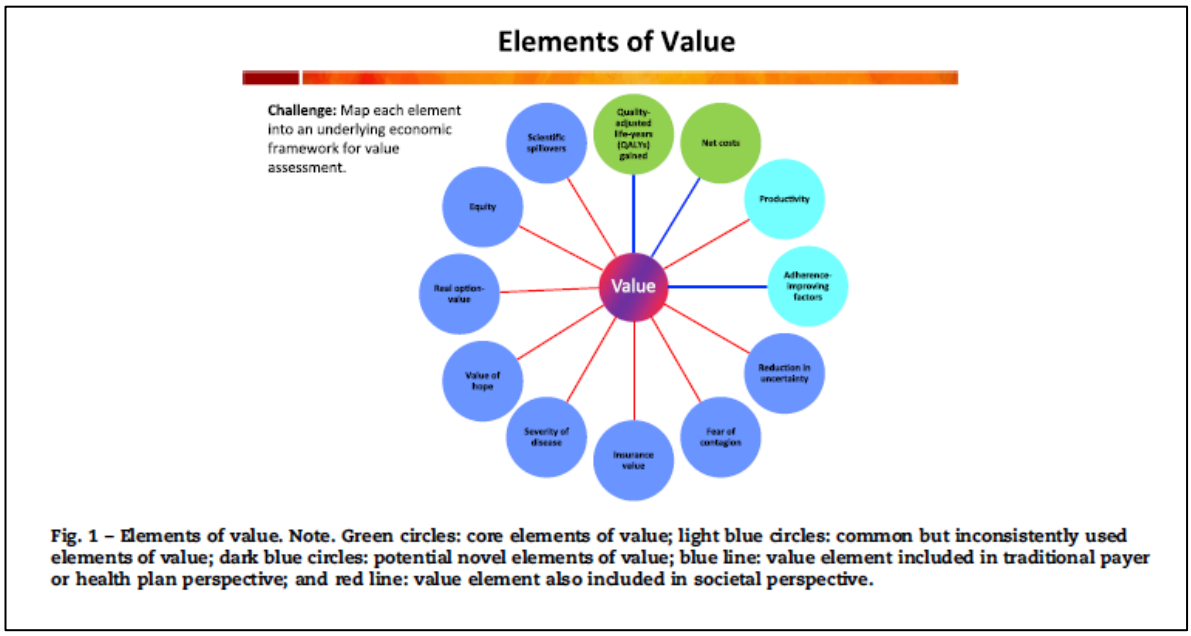


Fig. 1 – Elements of value. Note. Green circles: core elements of value; light blue circles: common but inconsistently used elements of value; dark blue circles: potential novel elements of value; blue line: value element included in traditional payer or health plan perspective; and red line: value element also included in societal perspective.

Reprinted from Value in Health, 2018 Feb;21(2):131-139., N. Lakdawalla et al. Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report 3. Reproduced with permission from Elsevier [License Number 5353000497365]

Other experts have also tried to more precisely define all the dimensions in which innovation can generate value. Sara Hofmann and colleagues and Rejon-Parrilla and colleagues and recently conducted two targeted literature reviews on this topic and developed inventories of items useful to examine and compare the approach to innovation assessment by different decision makers (Hofmann, Branner et al. 2021, Rejon-Parrilla, Espin et al. 2022).

In our work, the following dimensions in which oncology therapies can deliver value, are considered rewardable innovation:

- **Therapeutic benefit**: as evaluated against a relevant comparator, in terms of effectiveness, safety, and/or quality of life
- **Step-change novelty**: in disease management, typically related to the disruptiveness of the new intervention, its "breakthrough status"
- **Unmet need**: in the care of the underlying disease, due to ineffectiveness of the standard of care or the lack thereof
- **Severity** of underlying disease
- **End-of-life** care
- **Public health benefit**: in terms of size of the population affected (e.g., rare diseases), social needs, tending to underserved groups, reduction in health inequalities, etc.
- **Patient and/or carer convenience**: such as mode of administration, regimen complexity, pill burden, improved adherence, treatment travel time, user "friendliness", etc.

- **Economic benefit:** such as direct costs, budgetary impact, indirect costs, non-healthcare resource use, etc.
- **Dynamic effects**: enabling effects of current innovation on future developments, such as considerations pertaining to incremental innovation, spill-over effects, or real option value

1.4 CREATING INCENTIVES, ASSESSING THEIR VALUE, AND REWARDING INNOVATIONS

Since rewardable innovation can deliver benefits in one or more of the aforementioned dimensions, different stakeholders have adopted different ways to assess the value of innovative medical interventions.

1.4.1 Patent and Marketing Authorisation

For an innovative medicine to become available, the journey begins with the intellectual property (IP) system and regulatory agencies.

There are three principal criteria to be fulfilled for a drug to receive a patent. It has to be novel (not previously described or published), useful (industrial application and expected therapeutic benefit), and innovative (includes inventive step, not obvious for person with access to all previous knowledge). Even in this initial phase, it is the pharmacological use of a molecule, what is "patentable," and its therapeutic effect (Sampat and Williams 2019). This is why drug repurposing is also patentable, though proponents of more radical innovation definitions have referred to this practice as "evergreening" (Dutfield 2017).

If the clinical development program yields positive results, and the sponsor can substantiate evidence of quality, efficacy, and safety of the innovative medicine, the sponsor seeks marketing authorization from regulatory agencies such as the Food and Drug Administration (FDA), the European Medicines Agency (EMA), or the Medicines and Healthcare products Regulatory Agency (MHRA). Regulatory approval requires a positive balance of the risk-benefit assessment, though not requirements pertaining to novelty. Once marketing authorisation is granted, the effective patent protection starts, but its duration is still counted as of the date of the patent application, and it formally lasts 20 years in most developed countries. In special circumstances, the term can be extended. For example, in the European Union, a supplemental protection certificate can extend the data protection for up to five years because registration process of pharmaceuticals drugs before market approval is significantly longer than that for other patentable goods. Other special circumstances in which data protection can be extended are the coverage of paediatric development plans, or the more ample market exclusivity granted for 10 years to orphan medicinal products. Though, it should be clarified that the regulation is complex and not fully harmonized across countries (Garattini, Badinella Martini et al. 2022).

Patent protection in the IP system and data protection in the regulatory system are the mechanisms to incentivise innovation as they delay generic competition. Without the profit-

eroding effect of competition, sponsors of innovative health technologies charge premium prices for long yet varying periods of time that can be extended for more than 20 years. This principle holds to some extent even in single-payer settings, where the sponsor's monopoly power meets the health-care provider's monopsony power, because pricing and reimbursement systems have been designed to base pricing negotiations on the expected value of these innovations, i.e. far from the marginal cost of producing it (Garrison 2010).

An interesting note from Aronson and colleagues, who discussed ways to reward step-change innovation, and propose mechanisms other than pricing such as reduced tax on profits for the innovator (a patent box) or value-based patenting (patent extension) (Aronson, Ferner et al. 2012). These considerations are rare in the literature about rewardable innovation that mostly refers to pricing and reimbursement, which we will discuss next.

1.4.2 HTA Review for Pricing and Reimbursement

In many countries, once the regulatory process is successfully completed, marketing-authorisation holders proceed to seek HTA endorsement and, in most cases, the nature and degree of innovation is also considered in decisions pertaining to pricing and reimbursement. The definition of innovation and its assessment varies across agencies (Hofmann, Branner et al. 2021).

1.4.2.1 Economic Evaluation of Innovative Health Technologies

In 1991, Australia became the first jurisdiction to request economic evaluations as part of the process for assessing value for reimbursement decisions. This policy became mandatory in 1993 and was shortly followed by New Zealand and multiple Canadian provinces. Now, most countries in the European Union, several US payers and countries in Latin America and Asia have incorporated economic evaluations to their decision-making processes (Augustovski, Alcaraz et al. 2015, Drummond, Sculpher et al. 2015, Beletsi, Koutrafouri et al. 2018, Garattini and Padula 2019) to maximise efficiencies in the allocation of scarce resources. In this review, we have focused on the US and Europe.

HTA bodies that review and approve pricing applications, such as the National Institute for Health and Care Excellence (NICE) in England and Wales, the Italian Medicines Agency (Agenzia Italiana del Farmaco, AIFA), the National Health Care Institute (Zorginstituut Nederland, ZIN [formerly College voor zorgverzekeringen]) in the Netherlands, and the Dental and Pharmaceutical Benefits Agency (Tandvårds- och läkemedelsförmånsverket, TLV) in Sweden, have dealt with the need to maximise health benefits of the population they serve, given a budgetary constraint, by assessing the value of novel therapies compared with the best available alternative, in terms of incremental costs and incremental benefits against a willingness-to-pay (WTP) threshold, which is has not been explicitly stated in most countries. The following three methodological approaches have allowed for the comparison of alternative interventions (two of them, even across therapeutic areas).

If the alternative interventions all give the same health outcome, a cost-minimisation analysis (CMA), can be conducted in which only the total costs with each alternative are considered.

In cost-effectiveness analysis (CEA) outcomes are measured in natural or physical units such as life years gained, disease exacerbation crises averted, risk reduction of an adverse outcome, correctly diagnosed cases, or disease progression delay in months. The difference in cost between alternative interventions is divided by the difference in outcome. This ratio, referred to as the incremental cost-effectiveness ratio (ICER), expresses the cost-effectiveness of the intervention of interest compared to an alternative in terms of how much has to be paid for each additional unit of effectiveness/outcome. Lower ICER indicates better cost-effectiveness.

Cost-utility analysis (CUA) is similar to CEA, but health improvements are measured as gains in preference-based measures such as quality-adjusted life years (QALYs), disability-adjusted life years (DALYs), or healthy-year equivalents (HYEs); the order in which they are presented follows the frequency with which they are used. These measures incorporate the strength of preferences for different health states, and can therefore result in socially efficient resource allocations. Another advantage is that they allow comparisons of interventions across therapeutic areas.

Extensive literature can be found on these methods (Slothuus 2000, Kobelt and Office of Health Economics (London England) 2002, Drummond, Sculpher et al. 2015) and critiques to each also abound (Gyrd-Hansen 2005, Coast 2009, Edwards, Charles et al. 2013, Buchanan and Wordsworth 2015, Culyer and Chalkidou 2019). Some of these critiques propose alternative approaches, as described below.

1.4.2.2 Assessment of Comparative Clinical Benefit of New Health Technologies

German authorities and the US public sector question the use of economic evaluations for the assessment of the merits of new health technologies.

The US public sector remains reluctant to the use of cost-effectiveness assessments as its premises require the acceptance of the ethical basis of utilitarianism. Several initiatives have repeatedly been rejected on ethical, legal, and political grounds (Neumann 2004). Even the transformational *Patient Protection and Affordable Health Care Act (ACA) from 2010* states that “The Patient-Centered Outcomes Research Institute ... shall not develop or employ a dollars per quality adjusted life year (or similar measure that discounts the value of a life because of an individual's disability) as a threshold to establish what type of health care is cost effective or recommended. The Secretary shall not utilise such an adjusted life year (or such a similar measure) as a threshold to determine coverage, reimbursement, or incentive programs under title XVIII” (Neumann and Weinstein 2010). Instead, the government relies solely in the assessment of comparative clinical benefit of new therapies.

Similarly, in Germany, since 2011, and according to the *Act on the Reform of the Market for Medical Products* (Arzneimittelmarkt-Neuordnungsgesetz, AMNOG), the Federal Joint

Committee (Gemeinsamer Bundesausschuss, GBA) decides on the evidence-based magnitude of the additional benefit, which guides the price discount negotiations. Robust methods for the comparative clinical benefit assessment have been issued the Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, IQWiG). It is worth mentioning, that the GBA has created the German Health Care Innovation Fund endowed with €1.2 billion for the establishment phase (2016-2024) and later extended funding (additional €200 per year until 2024) to support the development of new integrated care models (Theidel and von der Schulenburg 2016, Wenzl and Paris 2018, Berghöfer, Göckler et al. 2020).

In France, assessment of new therapies has long followed a similar logic, though economic evaluations are gaining traction. Traditionally, the French High Authority for Health (HAS) and the Transparency Committee would evaluate the medical service provided (Service Médical Rendu, SMR) by a new health intervention and then classify them according to the expected Improvement of Therapeutic Benefit (Amélioration du Service Médical Rendu) or ASMR Levels I to V, depending on the degree of innovation and efficacy compared to other treatments in the same class. However, since 2013, for the products classified as ASMR I, II or III, additional economic criteria are applied, though no cost-effectiveness threshold has been set. This additional assessment is meant to inform price negotiations when the expected budgetary impact is significant (over €20 million) (Drummond, de Pouvourville et al. 2014, Toumi, Remuzat et al. 2015, Angelis, Lange et al. 2018).

In Italy, AIFA recently adopted a multidimensional approach for the assessment of the degree of innovation of new therapeutics. Since 2018, they can be designated fully innovative, conditionally innovative or non-innovative, depending on the assessment terms of the therapeutic need, the added therapeutic value, and the quality of clinical evidence, following the GRADE (Grading of Recommendations Assessment, Development and Evaluation) system. The sponsors of therapies deemed innovative, get commercial advantages for up to 36 months, including coverage in the innovative drug fund, exemption from compulsory discount mechanisms, and immediate incorporation into regional formularies (Fortinguerra, Perna et al. 2021).

1.4.2.3 Social Value Judgement Considered by HTAs

The assessment of additional therapeutic benefit (or lesser harm) is typically based on the scientific value judgement of the efficacy (or effectiveness in some cases) in terms of mortality, morbidity, and health-related quality of life (HRQoL), but also safety and cost-effectiveness (where applicable) of new health technologies. Yet, most HTAs recognise that these constitute only a fractional dimension of the overall value and consider that their contributions are multi-dimensional (Nicod and Kanavos 2016, Angelis, Lange et al. 2018, Hofmann, Branner et al. 2021).

Thus, some HTAs include in their deliberation processes social value factors such as:

- Disease burden and severity: NICE and TLV via WTP threshold, IQWiG via added benefit assessment, HAS/Commission for Economic and Public Health Evaluation (Commission d'Évaluation Économique et de Santé Publique – CEESP) via SMR, AIFA, and ZIN
- Unmet need and availability of therapeutic alternatives: NICE via clinical need criterion, TLV via WTP threshold, HAS/CEESP as part of the SMR, AIFA, and ZIN
- Prevalence of the disease and its rarity: NICE, TLV, IQWiG lowers demands on evidence, HAS/CEESP, AIFA via accelerated procedure, ZIN
- Product's position in the therapeutic strategy as preventive, curative or symptom-amelioration: NICE and HAS/CEESP
- Encouragement of innovation: NICE, TLV if captured by the incremental cost-effectiveness ratio (ICER), IQWiG only if added therapeutic benefit, HAS/CEESP as part of the ASMR, AIFA, and ZIN
- Target population's life expectancy if life-extending at end-of-life: NICE.

In some cases, these considerations have been formalised in specific weights for health gains or higher WTP thresholds, though it remains unclear the way in which they effectively impact decision making across all countries (Golan, Hansen et al. 2011, Claxton, Sculpher et al. 2015, Nicod, Berg Brigham et al. 2017, Sculpher, Claxton et al. 2017, Frutos Perez-Surio, Gimeno-Gracia et al. 2019, Hofmann, Branner et al. 2021).

1.4.2.4 Dealing with Trade-offs: Augmented CEA, Extended CEA, and Expanded Multiple-criteria Decision Analysis

To deal with the expectable trade-offs between and across the different dimensions in which rewardable innovation is assessed, experts in the field group most of the dimensions missing in the traditional approaches into "two types of aggregation issues" (Phelps, Lakdawalla et al. 2018).

To deal with the multidimensionality of value, Phelps et al. highlight the issue of aggregation into a single metric, of additional dimensions in which novel health technologies generate value, not all of which are accounted for in the traditional CEA (Phelps, Lakdawalla et al. 2018). In the aforementioned report published by Lakdawalla et al., the ISPOR Task Force incorporate traditional and novel elements into a so-called "Augmented Cost-Effectiveness Analysis" (Lakdawalla, Doshi et al. 2018).

The other type of aggregation issues pertains to distributional aspects, or the "aggregation of cost and benefit information across individuals to a population level" for decision-making. Thus, to deal with these equity concerns and financial risk protection, ISPOR highlights a new methodology developed by a group of Harvard professors called "Extended Cost-Effectiveness Analysis" (Verguet, Kim et al. 2016). This approach allows for the comparison of alternative policies and interventions by examining and quantifying four dimensions for each population subgroup (relative to the equity matter in consideration), i.e., health gains,

private expenditures averted, financial risk protection afforded, and the net costs of the policy.

Finally, in the absence of pure dominance, these approaches by themselves are not conductive to assert a ranking of alternative options so, we need to take an additional step to inform decision-making for the adoption of a new health technology. Among all possible ways to consolidate the information into a single measure of value, ISPOR proposes to use multi-criteria decision analysis (MCDA) (Marsh, M et al. 2016, Thokala, Devlin et al. 2016, Phelps, Lakdawalla et al. 2018).

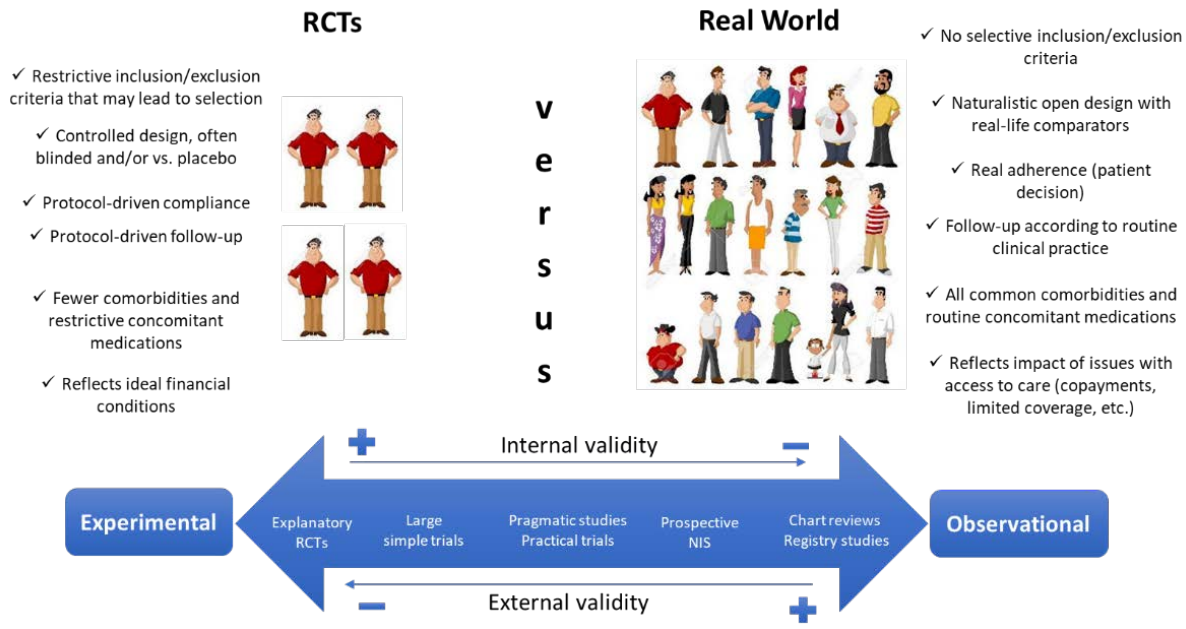
1.5 EVIDENCE BASE TO ASSESS REWARDABLE INNOVATION

For the assessment of rewardable innovation, regulators and HTA bodies have laid out guidance for the type of evidence they require. Randomised controlled trials (RCTs) have long been considered the golden standard when it comes to understanding benefits and risks of medical interventions. Yet, during the past decade, the use of data collected in routine clinical practice (or real-world data - RWD) has gained traction and is now being incorporated into formal guidance. The availability of vast and deep data repositories, paired with the unprecedented growth and acceleration in the development of analytic capacity (both computational and scientific), offer the opportunity for clinical and epidemiological research to complement and supplement evidence generated in a controlled environment. In the following sections we present basic concepts and circumstances of these supplementary use of RWD.

1.5.1 Experimental vs Observational Studies

As with innovation, we define real-world data (RWD) and real-world evidence (RWE) through their use and application. Throughout this work, we refer to RWD as the health data collected along routine clinical practice as opposed to in conventional controlled trials. With RWE, we refer to the evidence generated through the aggregation and analysis of record- and/or patient-level RWD, which may involve the processing of one or multiple data sources through linkage. Thus, RWE is more than just information, it is the information organised in such a way to serve as basis for conclusions or judgements (Garrison, Neumann et al. 2007, ISPOR Task Force 2013, Berger, Sox et al. 2017, Makady, de Boer et al. 2017).

Figure 4. Comparison of Evidence Emanating from RCTs vs. Routine Clinical Practice



Abbreviation: NIS = non-interventional study; RCT = randomised controlled trial

In Figure 4, we present and illustrate some key differences between these two types of evidence. The first contrast refers to the study designs that generate these types of evidence, with a focus on their use for the assessment of rewardable innovations. Randomised controlled trials (RCT) are experimental studies in which the intervention is tested in different phases to assess its quality, efficacy, and safety. The prospective design of an experimental study entails clearly pre-specified participants selection criteria (often restrictive), follow-up schedule and assessment methods, and well-defined endpoints. For the phases in which comparative assessments are conducted, these studies also entail randomisation, control groups, and blinding. These elements of the experimental design, aim to provide unbiased estimates of the effect of the intervention in the trial population, which confers the results high internal validity. In contrast, in observational studies, the researcher does not intervene during the study conduct but only observes and reports, thus minimising the risk for the observer effect or surveillance bias that can affect results. Furthermore, the advantages of observational studies are that they include more diverse and representative populations and allow to capture effects of the interventions in routine clinical practice, as opposed to the ideal circumstances in RCTs pertaining to treatment adherence, follow-up, and financial constraints for the overall care of these patients including but not limited to diagnostics and supportive care.

Traditionally, experimental studies and meta-analyses of experimental studies have been at the summit of the hierarchical ranking of evidence designed by prestigious institutions such as the Cochrane Collaboration (Woolf 2000, Evans 2003, Higgins and Green 2011) but there is a broader understanding of the circumstances in which RWE is a better option, or at least a necessity at par with RCTs. Typical examples are evidentiary needs on natural history or

burden of diseases or to assess health interventions where randomising patients to no treatment pose ethical issues (e.g., small indications with high unmet need).

1.5.2 Established Sources of RWD

RWE studies can leverage existing data collected along the healthcare continuum. Following, we present some of the typical types of secondary data.

1.5.2.1 Clinical Registries

These registries, created as open or closed cohorts defined by disease or treatment groups, are used for understanding natural history of the disease, monitor quality of care (healthcare provider performance) and/or results (effectiveness and safety), assess the impact of health interventions on specific clinical outcomes, benchmark across regions or countries, etc. Typically, they capture data in almost-real time and allow for the assessment of long-term clinical outcomes.

Cancer registers were among the first clinical registers to be developed and have evolved and spread worldwide (Wagner 1991, Ferlay, Colombet et al. 2021). While most of them offer limited information beyond diagnosis and death, there are a number of exceptions that also collect deep clinical information on disease presentation (such as stage and grade at diagnosis, tumour histology and certain biomarkers), necessary procedures (such as surgery or radiation therapies), systemic anti-cancer treatments even if administered while in-hospital. An even smaller number of these cancer registries they also collect certain outcomes not available in administrative registries (such as progression-free survival, treatment response, or patient-reported outcomes). They may lack in traceability through the different settings of healthcare provision, as they typically rely on reporting from secondary care facilities. This limitation is typically overcome through linkage, which is what we did in our studies.

1.5.2.2 Electronic Medical/Health Records (EMRs/EHRs)

Whether it is the digital version of the old paper charts or a more advanced interoperable and/or linkable version of it, EMRs/EHRs allow to extract data directly from their structured fields and sometimes even mine entries in unstructured format. The opportunities for research are immense but their strengths and limitations depend on the capabilities of the systems. In some cases, they can even count with linked files such as lab results or imaging. They offer the most recent data, and, in some cases, it is possible to reach patients through their healthcare providers to gather additional information or administer surveys.

In the US, withing integrated delivery networks (IDN), full traceability through the different healthcare settings within the network (healthcare provider) is possible, as is supplementing standard structured data with bespoke data collection from the patient EHRs.

1.5.2.3 Administrative and Insurance-Claims Databases

Data resources created and maintained for purposes other than research, that can be repurposed for retrospective longitudinal or cross-sectional analyses. These databases, are those resources primarily collected for the reimbursement of healthcare services provided, based on routine processes with a high level of automation and subject to regular audits and quality controls, thus resulting in a high-quality data capture. A key limitation of this type of data is the absence of important clinical information such as biomarkers, performance status, and surrogate outcomes like progression-free survival or treatment response. They are typically used for cost and burden of illness studies, to ascertain disease incidence or prevalence, and to investigate treatment or prescription patterns. In some cases, they can be used in comparative effectiveness or safety studies, but they rely only on standard coding systems (ATC for medications, ICD for diagnoses, NOMESCO for procedures, etc.) and these classifications not always allow to ascertain the clinical specifications of subpopulations (e.g., disease severity, biomarkers, etc.) or the adjudication of certain endpoints.

1.5.2.4 Primary Data Collection

In some cases, existing data resources are inadequate and primary data collection in real-world conditions is needed. These studies can be retrospective (chart reviews) or prospective. Typical examples are the post-authorization efficacy or safety studies. Whether voluntary or regulatory-mandated, these observational studies enrol and follow patients with a certain condition and treated with the therapy under study as well as comparators, without interfering in the routine care as decided by the physician (and sometimes with the patient). These studies can also be conducted in registries but, until recently, the majority of prospective studies would be site based.

1.5.3 Use of RWE to Evaluate Innovativeness

In oncology, comparative effectiveness research has been particularly prolific, which has triggered fruitful discussions and pushed the discipline forward (Ramsey, Veenstra et al. 2011, Deverka, Lavallee et al. 2012, Thariani, Veenstra et al. 2012, Esmail, Roth et al. 2013, Simonds, Khoury et al. 2013). While RCTs remain the golden standard, RWE has started being incorporated in formal assessments by regulators and HTAs.

In the US, the public sector continues to consider comparative clinical benefit as the only source of rewardable innovation and, in 2009, comparative effectiveness research become a more relevant policy instrument, as the *American Recovery and Reinvestment Act* (Section 804) earmarked $1.1 billion for the development and dissemination of this type of studies in the two years that followed and created the Federal Coordinating Council for Comparative Effectiveness Research. One year later, the ACA had instituted the Patient Centered Outcomes Research Institute (PCORI), which is aimed at evolving and improving the methodological guidance for this type of research (Vernon, Golec et al. 2010, Ali, Hanger et al. 2011, Manchikanti, Caraway et al. 2011). The approval of the 21st Century Cures Act in 2016, furthered this trend as it formalized the acceptability of the use of RWE in drug

marketing-approval programs. Subsequently, in 2018 the FDA published the "Framework for FDA's Real-World Evidence Program" and since then has been issuing guidance on data standards, on the assessment of EHRs and for registries to support submission, for external comparators, and for medical devices, as well as developed trainings and other resources for the industry and for researchers.

In Europe, there are four key initiatives that laid the ground for advances in the use of RWE in regulatory submissions. The first was the launch of the Adaptive Pathways Pilot in 2014, aimed at accelerating access to medicines in areas of high unmet need, whereby an initial conditional approval is granted in a restricted indication, subject to an iterative evaluation process based on RWE (Eichler, Oye et al. 2012, Eichler, Baird et al. 2015). The second, was the launch of the EMA Initiative for Patient Registries in 2015. In this framework, EMA developed guidance on the conduct of registry-based studies, created an inventory of registries (the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance [ENCePP]), and organized and hosted disease-specific stakeholder collaboration workshops. The third milestone was the establishment of the Data Analysis and Real World Interrogation Network (DARWIN EU®) by the EMA, a coordination center for the delivery of pan-European RWE research for regulatory purposes. The fourth was the European Commission initiative to create a European Health Data Space (EHDS), which aims at regulating the capture, access, and use of health data across Europe.

HTA agencies and payers have been using RWE for some time, but mostly limited to enable conditional reimbursement. These have taken several names ('coverage with evidence development' in the US and Germany, 'managed entry agreements' in Germany, 'only in research' in the UK, France, and Sweden, 'conditionally funded field evaluation' in Canada, 'monitored use' in Spain, etc.) but the concept is similar, reimbursement is conditional on the generation of additional evidence based on RWD (Carbonneil, Quentin et al. 2009, Baird, Banken et al. 2014).

1.5.4 Key Methodological Challenges in the Use of RWD for Causal Inference

The use of RWD to estimate treatment effects implies that the study design and the analytical approach ought to be fit-for-purpose to deal with the two main risks that cause erroneous or misleading results, that is confounding and bias. There is vast literature about the topics so we will not describe all the methodological challenges in comparative effectiveness studies (Delgado-Rodríguez and Llorca 2004, Rothman, Greenland et al. 2008, Lash, Fox et al. 2009, Rothman 2012, Hernan and Robins 2020), but will rather focus on introducing those relevant for our work. We have grouped them according to the way in which RWE departs from traditional Phase III RCTs: a) Enrolment b) Index, c) Randomization, and d) Outcome assessment.

1.5.4.1 Enrolment and Risk of Information Bias

When subjects are enrolled in clinical trials, a thorough assessment of their situation at baseline is conducted which allows to clearly distinguish pre-existing conditions and characteristics from those that occur during follow up and may be attributable to the treatment. In RWE studies, the consequences of departing from the full and uniformed assessment, is the risk of information bias. For example, if the registry has incomplete data on pre-existing comorbidities that are captured for the first-time during follow-up they could be misclassified as outcomes or mediators, instead of confounders. The use of RWE to evaluate treatment performance based on RWE requires careful consideration in the selection and specification of covariates. It is good practice to conduct only pre-specified analyses that are based on theoretical biological plausibility and properly qualify patients at baseline. Furthermore, counting only on secondary data may lead to misclassifying the exposure. For example, if prescription data is used to define exposure, all those subjects that did not fill their prescriptions or did not take the drugs would be misclassified. If nothing can be done to validate the data, methods exist to quantify the potential bias and conduct sensitivity analyses around the sources of uncertainty (Funk and Landi 2014).

1.5.4.2 Index Date and Risk of Time-Dependant Bias

The second departure from the RCT setup pertains to the choice of time zero, the time of origin. In RCTs, all subjects are indexed at comparable times, which are anchored in an objective milestone (such as date of enrolment and signing of Informed Consent Form, the date of randomization, date of occurrence of an event qualifying for inclusion in the trial, etc.). In RWE studies, particularly when assessing effectiveness or safety of a therapy as compared with untreated patients, the cohort identification and index assignment is often time-, event-, or exposure-based. Misspecification of index may lead to time-dependant bias, also known as survivor-selection or immortal-time bias, because it occurs when time-at-risk before index is not comparable between the cohorts or when there is window in which the outcome cannot occur (immortal time). There is rich literature on pharmacoepidemiology design and methods to guide the choice of index date for RWE studies, particularly when it comes to assess real-world effectiveness of new health interventions (Suissa 2008, Hernan and Robins 2020, Backenroth 2021, Lash, VanderWeele et al. 2021, Hatswell, Deighton et al. 2022). The consensus is that the new-user design is preferable, and methods guidelines recommend it (Ray 2003, Cox, Martin et al. 2009, Hernán, Sauer et al. 2016), together with advice on the establishment of appropriate wash-out period. However, there are several trade-offs that should be considered (Johnson, Bartman et al. 2013). For example, the sample depletion resulting from the exclusion of prevalent cases may jeopardize the powering of the analyses. Alternatives exist, such as the prevalent new-user design which relies on time-conditional propensity scores, but they are more complex and place higher demands on the data (Suissa, Moodie et al. 2017).

Apart from time-dependant exposures, time-varying covariates may also introduce bias in longitudinal studies. When the time trends in covariates is stochastic or varies independently

from exposure and outcome, it is sufficient if time-to-event analyses account for their varying impact on the baseline hazard for the outcome. However, in the presence of path dependence, it is important to determine whether the pattern of change over time is endogenous to the causal relation under study, as it may lead to simultaneity bias or even reverse causation (Goodliffe 2003, Wolkewitz, Allignol et al. 2012).

1.5.4.3 Randomization and Risk of Confounding

One of the main challenges to use RWE in causal inference is the risk of confounding. When a risk factor or exogenous determinant is associated with both the exposure and the outcome, it may "confound" the causal relation leading to spurious results. This is because, in the absence of randomisation, the assumption of exchangeability may not be met. Typical examples in oncology are age at diagnosis and certain pre-existing comorbidities that affect the treatment decision and the risk of the outcome. Thus, study design and analytical methods ought to secure comparability between groups to be assessed.

Confounders that are known and measured are included in causal-inference analyses to obtain estimates that are "net" of the confounding influence so as to represent the "true effect". Traditionally, this would be done by controlling for these co-variates in regression analyses. The exponential growth of available data and progress in computational capacity allowed for the development of advanced methods that further traditional approaches like propensity score matching, inverse-probability weighting, and multi-level regressions. For example, machine learning is being applied for prediction modelling to improve propensity score estimation when dealing with high dimensionality, as is the case for typical of advanced modelling in oncology, that may require processing of genomic, proteomic, metabolomic data (Canfield, Kemeter et al. 2018) (Linden and Yarnold 2017) (Wyss, Schneeweiss et al. 2018).

Since not all confounders are known or measured, additional methods to control for unmeasured confounding have been developed, such as the use of instrumental variables (IV). Valid instruments are those that affect the exposure but are fully independent from the outcome. The IV method relies on unstable assumptions, and they do not always exist. When unmeasured confounding cannot be addressed in the design or the analysis, it is good practice (and requested by certain regulatory and HTA bodies) to assess its impact on the results through quantitative bias analysis (Lash, Fox et al. 2009, Leahy, Kent et al. 2022).

1.5.4.4 Outcome Assessment and Risk of Other Types of Information Bias

The last departure from the RCT setting pertains to the assessment of outcomes as, RWD often lacks the specificity and granularity to ascertain outcomes in comparable time windows and with comparably robust methods or measurements which may affect the accuracy and precision of estimates. For example, insurance claims data tends to be more prone to censoring as switches in plans cause higher loss-to-follow up than typically recorded in RCTs. Furthermore, missingness is a problem in all study variables but, in the assessment of outcomes, it may affect the entire study population.

This is particularly the case in oncology, where clinical development programs of new therapies typically start in metastatic patients where few or no therapeutic options exist. Thus, given the high unmet need in many tumour types, they qualify for accelerated access schemes, often on the basis of results of single-arm trials, thus requiring additional evidence generated from RWD. One of the main challenges to constitute those external control arms is that certain outcomes, such as tumour response or progression-free survival, are measured with specific classifications used mainly in trial settings as is the case of Response Evaluation Criteria in Solid Tumours criteria. Since this standard is not always used in routine care, Flatiron Health developed a method to abstract real-world tumour response for a number of trials and arms with data abstracted from the physicians' unstructured clinical notes of the imaging obtained for the radiographic disease evaluation and their assessment of disease change (Griffith, Tucker et al. 2019). Other strategies for data expansion through linkage across types of RWD or augmentation through additional data collection can be effective to limit the impact of missingness (information bias) in RWE studies.

2 RESEARCH AIM AND OBJECTIVES

We aimed to assess the value and contribution of new oncology treatments in terms of extended survival, improved disease symptom control, reduced toxicity, and/or expanded alternatives to better tailor disease management, based on data generated in routine clinical practice in the care of patients diagnosed with the most incident cancers (lung, prostate, and breast) in early and/or advanced stages. This overall goal of the research program was broken down in a series of objectives in each of the constituting studies.

The first study was a literature review and expert consultation aiming to clarify basic RWE concepts, evaluate its use by diverse stakeholders and healthcare decision-makers, identify methodological challenges, and synthesise these concepts in several case studies. The notions and definitions have been presented and discussed in the previous section, and the article published in Value in Health (Paper I) is centred on the case studies in four Latin-American countries (Justo, Espinoza et al. 2019).

The other four constituting studies leveraged existing data collected along the cancer care continuum to understand the different dimensions in which biopharmaceutical innovation could create value, from extending life to enabling the tailoring of disease management by expanding the treatment armamentarium. As such, their specific objectives were:

- The first of these studies (Paper II) investigated therapeutic innovation in the care of patients diagnosed with advanced or metastatic non-small cell lung cancer (NSCLC) in the US. At the time of the study conception, the life expectancy of these patients was significantly shortened by the disease but the previously limited treatment options were being expanded with several new therapies, thus allowing the assessment of the impact of innovation in a population with high unmet clinical need. More specifically, the objective of this study was to inventory the level of therapeutic innovation introduced between 1991 and 2012 in the US to treat patients diagnosed with advanced or metastatic NSCLC and evaluate its impact on survival(Justo, Nilsson et al. 2020).
- The second study (Paper III), also focusing on advanced or metastatic NSCLC, looked into comorbid conditions that exacerbate the poor prognosis of patients and which are typically associated with the same underlying risk factors as lung cancer, with the progression of the disease, and/or with its treatments. Thus, these conditions may play different roles in the causal path from advanced or metastatic NSCLC to survival as, depending on the timing and clinical characterisation, they may be confounders, mediators, and/or effect modifiers and their influence may be time-varying. As such, any treatment effectiveness analysis ought to apply the right corrective strategy in each case. Thus, a thorough characterisation of these conditions was warranted, and the specific objective of this study was to describe the occurrence of several conditions associated with the disease and with its treatment, in the 12 months before (comorbidities) and 12 months after (adverse outcomes) the diagnosis of advanced or metastatic NSCLC in Sweden(Linden, Redig et al. 2020).

- In the third study (Paper IV), we studied patients with metastatic castration-resistant prostate cancer (mCRPC). Similarly, to the context of our study in lung cancer, at the time of conception of this study, new treatment alternatives were being introduced in mCRPC, an indication in which the only available options were continued hormonal therapies that could be supplemented with docetaxel, or watchful surveillance with symptom control. Thus, investigating the way in which these innovative treatments were being used was warranted and the primary objective of our study was to describe treatment patterns of patients diagnosed with mCRPC in Germany, in terms of sequencing and duration, and characterised by their demographic and clinical characteristics (Justo, Nilsson et al. 2020).
- The last study (Paper V) investigated one specific innovation in the treatment of patients diagnosed with breast cancer. With an over 20-year hindsight of the use of trastuzumab and its clinical and economic consequences, this ex-post evaluation of its benefits allowed us to construct a real world Weberian “ideal type” for innovation in oncology. This study aimed to estimate the life-cycle value created by trastuzumab based on RWD and characterise the split of the welfare surplus between the consumer (in this case, the Swedish Society) and the producer (in this case, the innovator).

3 MATERIALS AND METHODS

3.1 LITERATURE REVIEW

Paper I presents results of a literature review conducted at the outset of this doctoral education program. The search strategy and strings used for the identification of sources of RWD in Latin America can be found in the Supplemental Online Materials published with the article, together with a description of the expert consultation. Additionally, this systematic review was supplemented with a targeted literature review in each of the relevant topics discussed in Section 1, including grey literature and official sources (EMA, FDA, European Union, and HTA bodies). Finally, it is worth mentioning that the studies constituent of this thesis also drew from the scientific and the grey literature. The literature review in the first study was supplemented with an expert consultation process that guided the translation of the concepts under investigation into four case studies in Latin America, expounding the existence, use, and acceptability of RWE in healthcare decision-making. For the other papers, the literature review which informed the design and implementation of the three retrospective cohort studies and the health-economics model.

3.2 SUMMARY SPECIFICATIONS OF THE CONSTITUTING STUDIES

Table 1 presents key elements of the study design, setting, and analytical approach. So, no duplicate content was presented in the constituting articles, following this summary presentation of materials and methods, we discuss some of the key components of the study design and methodological challenges faced in the conduct of these studies.

Table 1. Overview of Design, Materials, and Methods in the Constituting Papers

	Paper II	Paper III	Paper IV	Paper V
Country	United States	Sweden	Germany	Sweden
Study Population	All patients enrolled in Medicare and diagnosed with advanced or metastatic NSCLC between 1/1/1991 and 30/06/2012	Advanced or metastatic NSCLC cohort: All adult patients, newly diagnosed with advanced or metastatic NSCLC between 1/1/2006 and 31/12/2013, registered as residents in the healthcare regions Skåne and Västra Götaland Reference cohort: adult subjects from the general population matched 1:4 on age, sex, and county of residence at the time of advanced or metastatic NSCLC diagnosis of the case to which they were matched	All male patients affiliated with one of the sickness funds contributing data to Vilua Research Database, diagnosed with prostate cancer treated with any of the four treatments indicated for mCRPC (docetaxel, abiraterone, cabazitaxel, or enzalutamide) between 1/1/2013 and 31/12/2015	In palliative setting: All female patients diagnosed with HER2+ MBC in stages IIIB or IV, between 1/1/2000 and 31/12/2021 In adjuvant/neoadjuvant setting: all female patients diagnosed with HER2+ EBC in stages I-IIIA, between 1/1/2005 and 31/12/2021
Exposure	Degree of innovation proxies: medication-vintage reflecting use of 37 oncology therapies	Diagnosis of advanced or metastatic NSCLC	Docetaxel and three (then) innovative therapies: abiraterone, cabazitaxel, and enzalutamide	Trastuzumab
Outcomes	OS	Adverse outcomes commonly associated with NSCLC and its treatments, including haematological, cardiovascular, respiratory, gastrointestinal, musculoskeletal, metabolic, and other diseases and conditions	Treatment patterns described by duration and sequence	For patients with MBC, outcomes were progression-free survival and overall survival. For patients with EBC, outcomes were recurrence free survival, BC-related mortality, and all-cause mortality
Data sources	Cohort identification, medication use, outcomes, and covariates: Linked database SEER-Medicare Innovation inventory: Clinical Outcome Labelling Claims Database (PROLabels), scientific literature identified and extracted through Medline	Cohort identification: Swedish Cancer Register for the advanced or metastatic NSCLC cohort and Total Population Register for the matched reference cohort (selected by Statistics Sweden) Co-variates and outcomes: Skåne Healthcare Register and Västra Götaland Healthcare Register linked with the National Cause of Death	Cohort identification, medication use, outcomes, and covariates: the Vilua Health Research Database (formerly known as Gesundheitsforen Leipzig)	Model inputs were derived from data provided by the Swedish National Quality Register for Breast Cancer and eHälsomyndigheten, and supplemented with estimates obtained from other Swedish observational studies and clinical trials

	Paper II	Paper III	Paper IV	Paper V
	and Embase, and grey literature, from online repositories (FDA, EMA, and Swedish Medical Products Agency)	Register		
Study Measures and Main Analyses	For OS: Royston-Parmar flexible model (Royston and Parmar 2002) using restricted cubic splines for time-varying covariates, to estimate HR and 95% CIs. For 1-year OS: Logistic regression to estimate OR and 95% CIs	NSCLC cohort compared with the general population on: - Prevalence of comorbidities at baseline: Logistic regression to estimate OR and 95% CIs - Incidence of adverse outcomes in the year following index: Poisson regression to estimate Incidence Rate Ratios and 95% CIs - Mortality: Poisson regression to estimate Mortality Rate Ratios and 95% CIs Additionally, median overall survival was estimated using Kaplan-Maier	Descriptive analyses for patient characteristics and treatment sequence Median (95% CI) treatment duration estimated using Kaplan-Meier method Mean (95% CI) treatment duration estimated using the restricted mean survival time(Therneau T 2015)	Lifecycle value was measured as i) total Life-years gained, ii) quality-adjusted life-years gained, iii) productivity gains, and iv) healthcare cost savings, and estimated by aggregating clinical and economic benefits synthesised through two 3-state Markov models comparing trastuzumab-containing regimens vs standard of care at time of introduction in each setting, and evaluated separately in ten 5-year age groups. MBC: Starting in pre-progression state, patients transitioned to progressive disease and/or to death, the model was run for 120 monthly cycles as of 2000. EBC: Starting in remission, patients transition to recurrence, and/or to death, the model was run for 50 yearly as of 2005

Abbreviations: CI = confidence interval; EBC = early breast cancer; EMA = European Medicines Agency; FDA = Food and Drug Administration; HER2 = human epidermal growth factor receptor; HR = hazard ratio; MBC = metastatic breast cancer; mCRPC = metastatic castration-resistant prostate cancer; NSCLC = non-small cell lung cancer; OR = odds ratio; OS = overall survival; SEER = Surveillance, Epidemiology, and End Results

3.3 DATABASES

In the four constituting studies of this thesis, we leveraged different types of data sources and we hereby described them grouped by type

3.3.1 Clinical Registries

Three clinical registries were used in the constituting studies, the Surveillance, Epidemiology and End Results (SEER) Program in the US for Paper II, the National Cancer Register (NCR) in Sweden for Paper III, and the National Breast Cancer Quality Registry (NKBC) in Sweden for Paper V. Following, we present a basic description of these sources and the data used in each case.

3.3.1.1 SEER

The US National Cancer Institute (NCI) created the SEER Program in 1971 to consolidate data from population-based cancer registries held by the participating states. Over the years, the register expanded its coverage in terms of variables captured and contributing states. SEER-Research is the database with clinical information on the cases detected and the treatments planned for first line, though follow-up is limited. Thus, linkage to Medicare health insurance data for treatment, and to the National Center for Health Statistics for date and cause of death, has been an indispensable research resource for decades facilitated by a collaborative effort of the Centers for Medicare and Medicaid Services (CMS) and the NCI. Deterministic record linkage is done based on a deterministic algorithm based on social security number, name, sex, birth date, and date of death (Potosky, Riley et al. 1993, Warren, Klabunde et al. 2002, Enewold, Parsons et al. 2020).

Medicare is the federally funded program that provides health insurance for the elderly (95% of persons aged 65 or older are eligible) and other smaller populations of eligible patients (e.g., persons with end-stage renal disease or disabled) who made 17% of all enrolees as of July 2012 (closest reporting time to the end of follow-up in our study) and dropped to 13% in 2020 (Statistics). All beneficiaries are eligible for Part A (hospital, skilled-nursing facility, hospice and some home-health care services), and about 96% of them choose to pay a monthly premium to enrol in Part B (physician and outpatient services). About one-third of these patients also opt for Medicare Part C (Medicare Advantage or managed care) and join a privately managed healthcare plan, typically with a health maintenance organisation. Our study used data on all healthcare encounters covered in Parts A, B, and C from 1991 until end of follow-up (31 December 2013). As of 2006, we also included data from Medicare Part D, which provides prescription drug coverage for beneficiaries who purchase the benefit (over 50% since its inception and increasing till over 70% today) (Statistics). The generalisability and validity of these data has been extensively tested and the government has developed an information portal with an extensive list of publications (US National Cancer Institute).

3.3.1.2 Swedish NCR

In 1958, surveillance data on all cancer cases started being collected in the NCR, and reporting is compulsory for all healthcare providers in the country. Since the mid-80s, registration, coding, and quality control routines are the responsibility of the oncological centres in each of the six medical regions in which Sweden is divided for administrative purposes. The registry contains information on the patient and the disease at presentation (demographics, site, histology, stage, basis and date of diagnosis, etc.) but limited follow-up data (date and cause of death, date of migration). Data available, while limited, is of high quality as found in a validation study conducted in 2008 and the major quality assurance and correction work performed continuously (Barlow, Westergren et al. 2009, National Board of Health and Welfare [Socialstyrelsen] 2022). Given the limited data on follow-up, supplementing it via linkage is typically necessary. In Sweden, this is possible thanks to the existence of a 12-digit national personal identification number (PIN) issued to all Swedish national and residents by the National Tax Authority and recorded in all the national and regional registries existing in the country enabling linkage, not only with other health databases but also statistics on social security, labour market, taxes, etc. (Ludvigsson, Otterblad-Olausson et al. 2009).

In Paper III, we used the NCR to identify patients with advanced or metastatic NSCLC and the Total Population Register (Ludvigsson, Almqvist et al. 2016) for Statistics Sweden to select a matched cohort from the general population. The PINs from both cohorts corresponding to people living in the Västra Götaland, Skåne, and Halland regions were then sent to the respective regional registries that extracted baseline and follow-up information (more information on these registries in the section for Electronic Medical Records). Additionally, the most recent mortality data was also linked and provided by the National Board of Health and Welfare from the National Cause of Death Registry (Brooke, Talbäck et al. 2017).

3.3.1.3 NKBC

The NKBC, like other Quality Registries in Sweden, was created in 2008 for quality assurance and benchmarking across regions, and for research. Data is collected through a platform developed by the Information Network for Cancer Care (INCA), on all primary breast cancer cases. The Registration Form, used at case intake to establish a baseline, captures information on the patient, the disease at presentation (including demographics, biometrics, relevant elements of their clinical history, location, stage, histology, and biomarkers such as human epidermal growth factor receptor 2 [HER2] and hormone receptors), the diagnosis, the intended front-line treatment, and the reporting hospital. Then, come the treatment forms: Operation (which also include additional clinical data such as Ki67-status, confirmation of tumour, node, metastasis [TNM] values, histopathology grading, etc.), Direct Reconstruction / Oncoplastic Surgery, Adjuvant Treatment (pre- and postoperative oncological therapies that were actually administered, including systemic anti-cancer therapies and radiotherapy), Changes in the planned treatment (though no reason for

regimen change), and the last form, for Follow-up, Remote Metastases / Postoperative Loco-regional Recurrence (though these data are only disclosed for the Stockholm-Gotland healthcare region due to underreporting in the rest). Furthermore, since 2020, patient-reported outcome measures (PROM) are also collected (Regional Cancer Centrum). Despite the underreporting of relapses and recurrent cases, data in the NKBC is of high quality. A recent validation study found almost full coverage (99.9% across all regions), low missingness (<5% or most variables reported), and high exact agreement (>90%) as compared with data extracted from the medical charts of 800 patients randomly selected (Löfgren, Eloranta et al. 2019).

3.3.2 Health Insurance Claims Data

As we already described in the previous section, Paper II used SEER-linked Medicare claims data. Similarly, Paper IV leveraged the German database Vilua Healthcare Research (formerly known as Gesundheitsforen [GFL] and Arvato Health Analytics), which consolidates claims data from several sickness funds covering approximately four million individuals out of 70 million in the statutory health insurance system (Gesetzliche Krankenversicherung) and is representative of the overall German population in terms of age, sex, and morbidity (Pöllinger, Schmidt et al. 2019). The sickness funds routinely collect all information relevant for the reimbursement of prescribed medications and healthcare providers in the two large sectors: inpatient services (which are paid for directly by the sickness funds) and ambulatory services (which are reimbursed based on lump sum payments to the physicians' regional associations) (Busse, Blümel et al. 2017). Thus, Vilua contained information on patient characteristics, diagnoses and dates of primary and secondary care outpatient visits, hospitalisation, prescriptions, and cost for almost four million insured persons and, for almost 3 million of them, more granular information was available including procedures, diagnoses-related groups (DRGs), or inability to work (the "Feingranularer Datensatz"). Since the insured can switch between sickness funds, typically <5% of the subjects in the database fluctuates, though the proportion of switchers is even lower for persons with chronic diseases and among the elderly (Werner, Reitmeir et al. 2005).

3.3.3 Electronic Medical/Health Records

In our work, this type of data was used for Paper III. In Sweden, the 21 healthcare regions are responsible for the provision of health services but, the National Board of Health and Welfare (Socialstyrelsen) systematically consolidates data on secondary care, prescribed medications redeemed in community pharmacies, and social services into national administrative registries. Yet, other data elements, particularly pertaining to primary care, remain geographically fragmented. So, certain regions have created more complete registries that are also made available for research. Such is the case of the Skåne Health Care Register (SHCR), an administrative registry that combines information extracted and transferred from EMRs, with administrative application sources on all consultations used for reimbursement purposes. The registry provides information on primary and secondary diagnoses, type of encounter (physical visit, telephone visit, emergency care, etc.), type of facility (primary, secondary

outpatient, tertiary inpatient), type of caregiver/health professional, surgical and medical procedures, DRG codes and reimbursed costs (Löfvendahl, Schelin et al. 2020). Furthermore, a recent validation study has demonstrated its reliability for cancer research (Shen, Schelin et al. 2021). Likewise, the VEGA Health Register contains similar information for residents of the Region Västra Götaland and has also been used for epidemiological research similar to ours (Agnafors, Norman Kjellström et al. 2019, Agnafors, Torgerson et al. 2020, Kruse and Thoreson 2021).

3.4 COHORT IDENTIFICATION AND CHARACTERISATION

The identification of eligible patients in these studies was based on codes from the International Classification of Diseases for Oncology (ICD-O) (World Health Organization) but they were not sufficiently specific, thus requiring additional eligibility criteria to accurately define the target patient population. The additional criteria were needed to supplement the International Classification of Diseases codes with clinical characteristics of the disease, mainly pertaining to staging, histology, and or specific biomarkers. In the following sections, we describe the key delimiters to the topographic codes for each cancer under study.

3.4.1 Advanced or Metastatic NSCLC (Papers II and III)

In both studies, during the cohort-identification period, histological type was available and also coded using ICD-O in the NCR in Sweden and SEER-Medicare in the US, but the operational definition of this patient population required special attention to the selection of histological codes to accurately and comprehensively capture patients with NSCLC. Usually, the most prevalent histological types (adenocarcinoma, squamous-cell carcinoma, and large-cell carcinoma) were used to define this patient population, but the other non-small cell types had to be included in both studies. Additionally, in SEER-Medicare, NSCLC only started being recorded in the 2000 coinciding with the adoption of ICD-O-3 and, before then, many of these cases were classified as “Not otherwise specified” (See Figure 3 in Paper II). Similarly, in the Swedish Cancer Register, ICD-O-2 was used until 2004 and ICD-O-3 since 2005. So, for the survival analyses conducted in Sweden for the period 1973 to 2014, to validate findings in the US (Paper II), the old histopathological diagnosis classification (Patologisk Anatomisk Diagnos – PAD WHO/HS/CANC/24.1) was used because it offered consistency for the ascertainment of histological types through the entire period. To minimise the risk of misclassification bias, the cohort identification started with all topographical codes for lung cancer (ICD-O-3 C34.x), and then excluded cases registered with morphological codes for small-cell lung cancer, and neuroendocrine tumours as they are treated differently.

3.4.2 mCRPC (Paper IV)

The third study used claims data from several sickness funds in Germany that lacked clinical specificity to accurately determine disease stage at index. Thus, the identification of patients with metastatic castration-resistant disease relied on adherence to current treatment label and disease management guidelines at the time of patient follow-up, which only recommended

the use of docetaxel, cabazitaxel, abiraterone, or enzalutamide for the treatment of metastatic disease following chemical or surgical castration (European Medicines Agency (EMA) 2011, European Medicines Agency (EMA) 2012, European Medicines Agency (EMA) 2013, Horwich, Hugosson et al. 2013, Parker, Gillessen et al. 2015). Since docetaxel was also indicated for other cancers, the cohort identification criterion was further refined by excluding patients receiving docetaxel before the first record of prostate cancer in the data. While these assumptions were validated with clinical experts, the risk of misclassification due to off-label used was not completely eliminated.

3.4.3 HER2+ Breast Cancer (Paper V)

In the fourth study, the identification of patients with HER2+ early breast cancer (EBC) was conducted by the NKBC with a high degree of reliability, though data on patients with metastatic breast cancer (MBC) was incomplete because the registry does not capture consistently disease recurrence or treatment for patients with metastatic disease. Thus, the estimation of the number of patients with MBC treated with trastuzumab was approximated using annual sales data and assuming the same age distribution as in the adjuvant setting. Patients diagnosed with or who progress to metastatic disease tend to be slightly older, but not necessarily those who receive treatment due to the age-capped stopping rule for the use of trastuzumab in several Swedish hospitals. This assumption was confirmed when we found that 48% of patients 70 years or older with HER2+ EBC did not receive trastuzumab treatment, as compared with 14% of those younger than 70 (see Figure 2 in Paper V). And this finding relates to another challenge in the cohort identification for this study, the potential risk of selection bias due to differential testing practices. While the NKBC reports the proportion of patients with EBC who had complete pathology report including all standard biomarkers to be high (95% over the entire period) (Swedish National Breast Cancer Register 2022), there is no current information about testing in patients with metastatic disease, but a study conducted in the early days of trastuzumab adoption reported high geographical variability across healthcare regions in Sweden (Wilking, Jonsson et al. 2010). If those differences remain and the patients who are not tested are significantly different from those who are, the risk of selection bias should be considered. This risk did not significantly affect our study because the effectiveness measures were taken from RCTs so, if anything, the number of patients who could have benefitted from this innovation was underestimated, rendering our results conservative. Yet, it is worth highlighting this issue as all studies in which the cohort identification relies on companion diagnostics are susceptible to the risk of ascertainment bias, a type of ascertainment bias (Delgado-Rodríguez and Llorca 2004).

3.5 TREATMENT EXPOSURE

Overall, the key exposure under study in our research was innovation in oncology, which can be defined in different ways and measured in even more. In some studies, we looked into exposure as a specific drug, in others exposure to newer drugs in general. In the latter, defining innovation was more complex and operationalising it required careful consideration. Thus, we operationalised the concept of innovation in different ways as described below.

Since treatment decisions ought to be patient-centric and the trend towards personalised medicine accelerated during the period under study, there are few indications in which the perils and fruit of innovation can be synthesised is one drug. Such is the case of HER2+ breast cancer (BC) and the reason why trastuzumab constituted a perfect case study for Paper V, where the exposure to innovation was binary, i.e., treated vs. untreated.

3.5.1 Innovation as an Index (InnovInd)

In Paper II we started by building a simple proxy that would allow for the assessment of long-term trends. An additive measure that would only tally the stock of therapies approved and/or in use for the treatment of aNSCLC in the US at the end of each year from 1973 through 2012. Each year, Innovation Index (InnovInd) would add one count to the stock in the previous year, for every new therapy identified as newly approved or in use that year. Treatments available were identified in the FDA website, the literature, clinical guidelines, the PRObels™ database (Mapi, Research et al.), and from actual usage recorded in SEER-Medicare (for the years when data was available). A detailed list of drugs as well as the year and basis for the inclusion in this variable can be found in the article. However, since this proxy did not offer the granularity to account for the speed of uptake of innovative therapies, we also developed a more granular proxy, the mean medication vintage (MMV).

3.5.2 Mean Medication Vintage

Based on a concept borrowed from the literature (Lichtenberg, Grootendorst et al. 2009, Lichtenberg 2013, Lichtenberg 2014), three proxies for innovation were constructed following the MMV concept and leveraging the inventory created for the InnovInd and treatment data from SEER-Medicare between 1991 and 2012. MMV was a discrete variable reflecting the weighted average year origin of the medications in use, estimated as described by Equation 1.

Equation 1. Estimation of MMV

$$\boldsymbol{MMV_{year\ x}} = \frac{\sum_d \boldsymbol{N_d} * \boldsymbol{Approval\ year_d}}{\sum_d \boldsymbol{N_{dx}}}$$

where *d* is drug, N_d is number of patients treated with drug *d*, $Approval\ year_d$ is the year of primary approval or evidence of off-label use (whichever came first), and N_{dx} is all the patients treated with drug *d* during *year x.*

The three levels of aggregation were:

- $MMV_{Overall}$: Estimated based on the drug usage of each cohort defined by the year of diagnosis at the national level
- MMV_{State}: Estimated in the same way but clustered geographically by state contributing data to SEER-Medicare every year
- $MMV_{Patient}$: Estimated for each patient individually (analysis restricted to patients with treatment data as no imputations were performed)

3.5.3 Wave of Innovation in mCRPC

In Paper IV, three innovative drugs introduced within a five-year period (cabazitaxel, enzalutamide, and abiraterone) were studied and the effect of treatment patterns was assessed in comparison with the then standard of care, docetaxel.

3.5.4 A Single Innovative Therapy

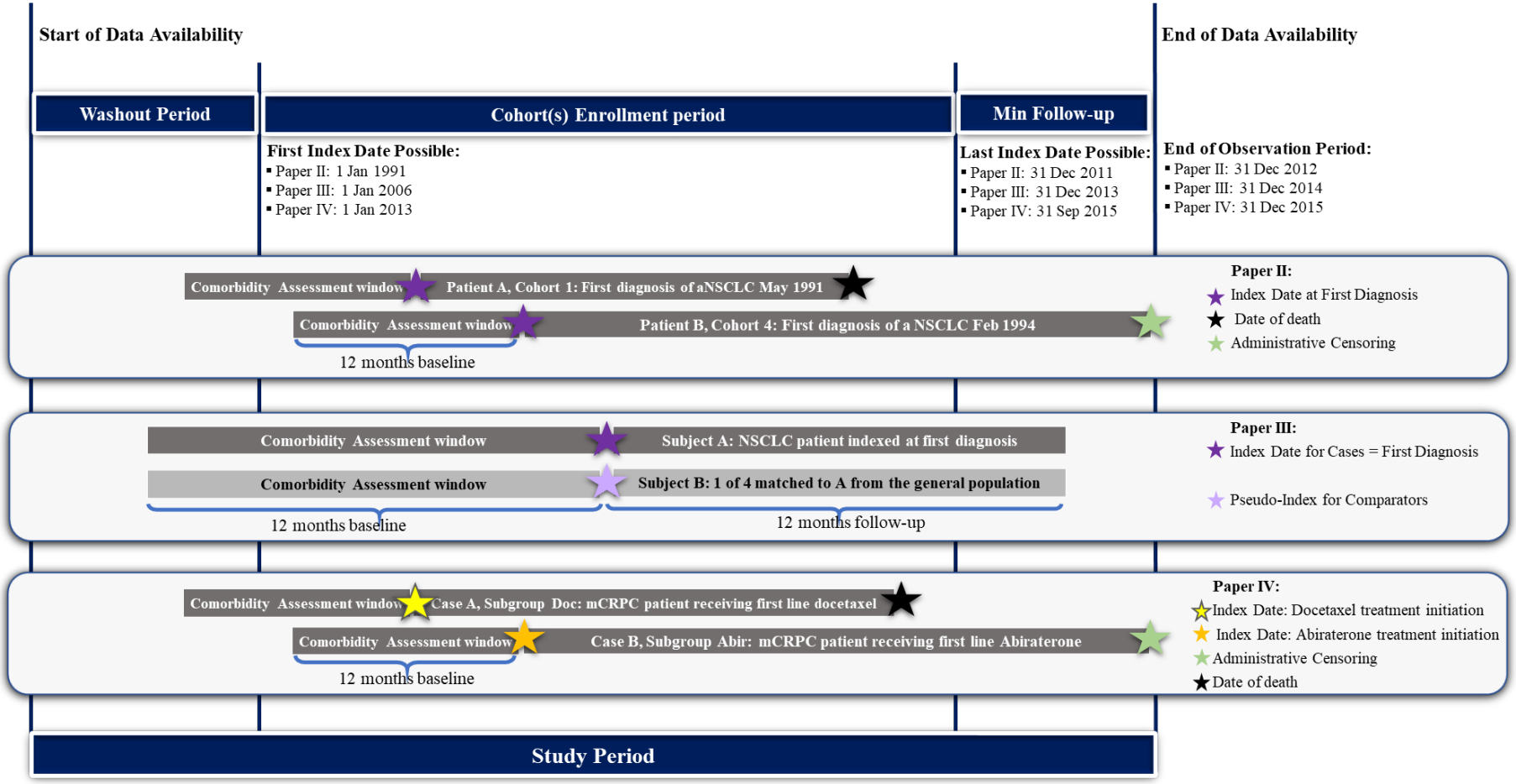
In Paper V, where the exposure to innovation was binary, and we assessed patients treated with trastuzumab vs. untreated patients.

3.6 INDEX DATE, BASELINE CHARACTERISTICS, AND CONFOUNDING

In all the studies, we aimed at indexing patients at first diagnosis and determine baseline characteristics in the year before, which also served as washout period to exclude prevalent cases, to minimise the risk of time-dependant biases (Wolkewitz, Allignol et al. 2012) and to better ascertain potential confounders (Wang, Schneeweiss et al. 2017). An exception was Paper IV, in which we indexed patients at treatment initiation due to data availability.

Figure 5 is a schematic representation of the study design from Papers II, III, and IV, depicting the relevant dates defining the periods for cohort washout, enrolment, and follow-up. Additionally, two hypothetical subjects for each study are used as examples of qualifying events start of observation, index date, and end of follow-up.

Figure 5. Study Design Diagram



Abbreviations: aNSCLC = advanced or metastatic non-small cell lung cancer; mCRPC = metastatic castration-resistant prostate cancer; NSCLC = non-small cell lung cancer

3.7 OUTCOMES, STUDY MEASURES, AND ANALYTIC APPROACHES

3.7.1 Descriptive Statistics

In the studies about advanced or metastatic NSCLC and mCRPC, individual patient-level data (IPL) was analysed; Papers II, III, and IV start with an attrition table depicting, step-by-step, the number of patients dropped and remaining after applying each inclusion and exclusion criterion from the respective protocols, to arrive at the final sample size. Following that summary specification of the analytic study population, we described the baseline characteristics assessed during the 12 months prior to the index date.

The distribution of continuous covariates such as age at diagnosis and Charlson Comorbidity Index (CCI) was generally described by the mean, standard deviation (SD), median, quartiles 1 and 3, and range (minimum and maximum). The distribution of binary covariates (sex) and categorical variables (stage at diagnosis, histology [NSCLC studies], site of metastasis [only in mCRPC study], and region [NSCLC studies] were described by the absolute (counts) and relative (percentage) frequency of each category and, if applicable, a category for missing data was added.

3.7.2 Graphic Descriptions

In Papers II, IV, and V we also resorted to graphical descriptions of covariates and outcomes to illustrate their time trajectory. For example, observing the graphical representation of the evolution of the distribution of sex, histology, and age and stage at diagnosis of aNSCLC in the US over 40 years, it was apparent that these time-varying covariates represent a challenge for traditional methods to analyse time-to-event measures. Also in Paper II, the figure overlaying the evolution over time of innovation and survival makes a visually appealing case for their close association, as demonstrated through the multivariate models. In Paper IV, the description of treatment pathways and sequences was also presented graphically with a patient-flow tree. Finally, in Paper V, overlaying the estimates of survival from the clinical trials with those estimated based on data from the NKBC was a simple yet robust approach to validation of one of the most important model parameters.

3.7.3 Comparative Assessments

In two of the papers, we used the IPL data to make explicit comparisons between groups. In Paper III, we compare patients diagnosed with advanced or metastatic NSCLC with the general population to understand excess risk in the cohort with the disease. In Paper IV, we compare patients diagnosed with mCRPC across the four index treatments, or pairwise against those initiating with docetaxel (the non-innovative therapy) to evaluate the presence of systematic differences in prescription patterns. Depending on the distribution of the covariates, appropriate tests were used to assess these differences.

To compare the distribution of continuous variables such as age, CCI, or number of hospitalisations in the year before treatment initiation (Paper IV), 95% confidence intervals (CIs) were calculated assuming a t-distribution of the means. Other characteristics were

assessed but some of them were not included in the publications. So, in in the prostate cancer study, we also used analysis of variance based "omnibus" test to compare continuous variables across treatment groups, and T-tests to perform pairwise tests against the docetaxel. These assessments provided the same information as the CIs around the means. Finally, Chi-Square tests were used to compare categorical characteristics. All tests were performed using a 5% significance level.

In Paper III, the general health of patients and comparators was not assessed based on the CCI, but with a more thorough evaluation of the prevalence of pre-existing conditions before and up-to index. The number and percentage of patients with at least one record of the diagnosis during the baseline assessment period was reported for each comorbidity in both cohorts, and a logistic regression was used to estimate the odds ratios (OR) and 95% CIs.

3.7.4 Time-to-Event Analyses

All the studies used longitudinal data so, the need to account for time influenced the analytic approach, as well as the type of measure. So, the relevant study measures were median overall survival (OS), one-year OS in Paper II, incidence rates (IR) with 95% CIs and mortality rates (MR) with 95% CIs, and median and mean duration of treatment by line with their corresponding 95% CIs in paper IV. Below, we describe the methods used to estimate these measures and to compare degrees of exposure to innovation over time (Paper II), adverse outcomes between cohorts (Paper III), and length of treatment lines within the cohort across subgroups (Paper IV).

3.7.4.1 Paper II

All patients included in the study were categorised in cohorts defined by the calendar year of diagnosis and Kaplan-Meier (KM) survival curves were derived using date of diagnosis as time zero and death as the failure event, whereby patients were censored at loss to follow-up or end of data availability. Based on the survival functions, we estimated for each cohort median OS in months and one-year OS as the proportion still alive 12 months after diagnosis. We called the plotting of these estimates against time, the Innovation Frontier (IF), as they represent the maximum survival benefit these patients did actually derive from their treatments. The IF was overlayed with the InnovInd, and presented together with the Pearson's correlation coefficient.

Then, to understand the impact of innovation expected OS, we estimated hazard ratios (HR) with 95% CIs and OR with 95% CIs, adjusted for known predictors of outcomes such as age, sex, stage, CCI, and histology. Following the graphical examination of these covariates and testing of the proportional hazard assumption using Schoenfeld residuals (Schoenfeld 1982), it became apparent that we needed an alternative to the traditional Cox proportional hazard regression model. So, we used a Royston-Parmar flexible model to account for the time-dependent effects of these covariates, by using cubic splines (Royston and Parmar 2002). Finally, given the differential follow-up time across cohorts, we run a logistic regression for one-year survival to estimate adjusted OR with 95% CIs.

3.7.4.2 Paper III

For each outcome under evaluation, we calculated the number of patients who had a new record of the corresponding diagnoses codes in any care event (primary or secondary care, and inpatient or outpatient) during follow-up. It is important to mention that, for all chronic and most acute conditions under study, those subjects with a record of the diagnosis at baseline (i.e., before index) were excluded from the set at risk. Only the assessment of recurring myocardial infarction and stroke was done on the entire cohort, regardless of pre-existing conditions. Time at risk was calculated for each cohort, as the sum of days elapsed from index date, the occurrence of the event, or end of the 12-month follow-up period, whichever occurred first. With this information, we estimated the incidence rate per 1,000 person-years and corresponding 95% CIs for each relevant outcome were estimated using Byar's approximation to the exact Poisson probabilities (Breslow and Day 1987), given the rarity of some of the outcomes. Then, to compare the incidence of each outcome in the NSCLC cohort vs the matched reference cohort, we estimated the unadjusted incidence rate ratios (IRR) and corresponding 95% CIs using Poisson regressions. All-cause mortality rates and rate ratios with corresponding 95% CIs were estimated in the same way.

3.7.4.3 Paper IV

Originally, one of the objectives of this study was to estimate survival differences across patient groups but the data was not mature enough as in two of the treatment subgroups, 50% or more were still alive at the end of follow-up. Thus, time-to-event analyses were limited to evaluate the duration of treatment per therapy and line.

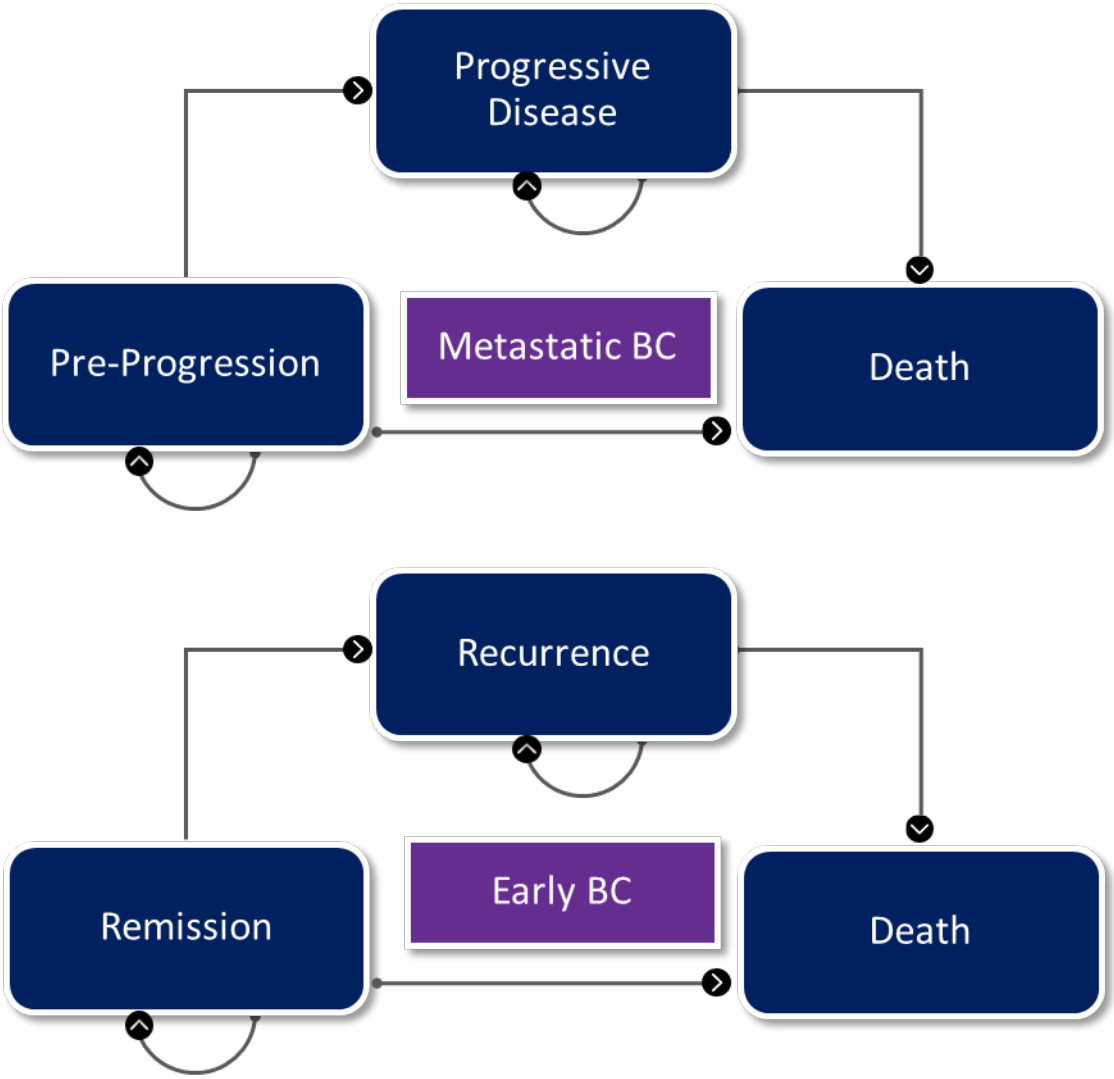
Because of the immaturity of the data, and to ensure comparability across subgroups (some of which had a small number of patients), we estimated the mean duration based on the restricted mean survival time (RMST) approach, which used the area under the survival curve from start of treatment line to the largest event-time observed in the group. The operational definitions for start and end of a treatment line as described in the article, allowed to determine time zero and the longest event-time for each group. The RMST approach has been extensively used the assessment of treatment effects in the presence of long-term survivors and proven to produce conservative effect estimates (Trinquart, Jacot et al. 2016). Additionally, the median duration on-treatment per therapy and line with corresponding 95% CIs was estimated using the KM method.

3.7.5 Health Economics Model

Paper V also leveraged IPL data for the estimation and characterisation of trastuzumab use. Two Markov models were built to simulate the accrual of costs and clinical and economic benefits of hypothetical cohorts defined by the disease stage at diagnosis (one model for MBC and another one for EBC), the age group to which they belong at the time of HER2+ BC diagnosis (in five-year intervals, pooling all patients younger than 40 years in the first cohort, and all patients older than 80 years in the last cohort), and exposure to trastuzumab vs.

no exposure. A graphical representation of the models can be seen in Figure 6**Error! Reference source not found.**.

Figure 6. Structure of the Markov Models Developed for EBC and MBC



Abbreviation: BC = breast cancer

Direct and indirect costs, quality of life (QoL) measures, and clinical outcome measures used as input in the models were drawn and adapted from published estimates from RWD studies conducted in Sweden (Lidgren, Wilking et al. 2007, Lidgren, Wilking et al. 2007, Olofsson, Norrlid et al. 2016). Costs and benefits were discounted to present values at 3% per year. In the MBC model, transition probabilities were based on exponential survival functions parametrised based on median progression-free survival and median OS from an RCT (Marty, Cognetti et al. 2005). Estimating transition probabilities for the EBC model, was more complex. We calibrated post-recurrence mortality, based on the risk of recurrence per year of follow-up published in a meta-analysis of several RCTs, to obtain the observed OS in the publication (Early Breast Cancer Trialists' Collaborative 2021). The estimates obtained from the survival calibrated in the model, were compared graphically with estimates from data hosted by the NKBC (as mentioned above), and the uncertainty around clinical

parameters drawn from RCTs, was tested in best-case worst-case scenarios sensitivity analyses, built using the upper and lower bounds of the respective parameters' CIs.

3.8 ETHICAL CONSIDERATIONS

3.8.1 Ethics Review

Our work consisted of four retrospective observational studies using registry data from Sweden, the United States and Germany on patients diagnosed with advanced aNSCLC, mCRPC, and HER2+ BC. We were required to obtain Ethics Approvals for studies for papers II, III, and V and Table 2 presents details of the submissions and decisions.

Table 2. Ethics Approvals Obtained

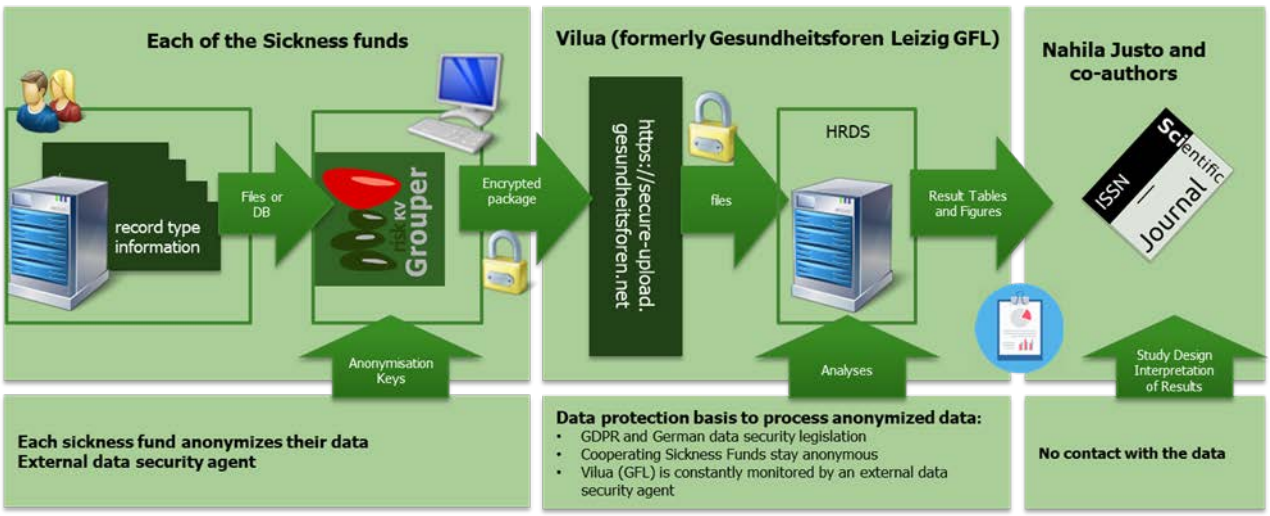
Paper	Ethics Review Body	Decision
II	In the United States, the study was approved by Quorum Review IRB	May 20, 2015 with registration number 30556/1.
	In Sweden, this study was approved by Regionala Etikprövningsmyndigheten in Stockholm	March 25, 2015 with registration number 2015/406-31/4
III	This study was approved by Regionala Etikprövningsmyndigheten in Stockholm	January 11, 2016 with registration number 2015/2012-31
V	This study was approved by Etikprövningsmyndigheten	January 25, 2022 with registration number 2021-06709-01

Abbreviation: IRB = institutional review board

Paper IV did not require an ethics submission as according to the German Federal Data Security Legislation. Figure 7 presents a schematic of the data flow in this study.

The raw claims data from the different sickness funds were fully anonymised by authorised personnel at each sickness fund and transferred to a secure database, based on a contractual agreement between Vilua (then GFL) and each participating sickness fund compliant with national law . The key for the anonymisation was kept by an external trustee (an external law firm), which also was in charge of monitoring access to the anonymised database. Access was restricted to a small number of approved personnel with Vilua (then GFL) user profiles, and the patient-level data hosted by Vilua (then GFL) was not shared with third parties. Vilua (then GFL) conducted the analyses following the study protocol and statistical analysis plan developed by the authors, and only shared summary statistics in result tables and figures. The procedures to create, host and access the database was compliant with the German law for the creation and maintenance of anonymised research databases, and these procedures were approved by the responsible Data Privacy Officer.

Figure 7. Data Flow for the Study Conduct for Paper IV



Abbreviations: GDPR = General Data Protection Regulation; GFL = Gesundheitsforen

3.8.2 Risk Assessment

All analyses were performed on historical de-identified registry data, and the results presented aggregated. Nevertheless, there are ethical considerations that ought to be considered with this type of research, given the remaining residual risks. To identify those risks, our research was reviewed against the four research ethical principles as outlined by Beauchamp and Childress in their book Principles of Biomedical Ethics (Beauchamp and Childress October 2012). Thus, we hereby discuss each risk identified in the context of our work.

3.8.2.1 Doing Good

Our research provided insights into the actual value delivered by the last generations of pharmaceutical innovation in certain oncology indications; innovation that has caused a steep raise in the cost of care for these patients. Also, understanding the current standard of care and aggregate outcomes, allows for a realistic assessment of the new innovations to come. As such, we believe that the benefits of our research outweigh the risks.

3.8.2.2 Avoiding Harm

Observational research is less prone than clinical research to directly harming subjects physically, mentally, or financially. Yet, we have identified two risks of harming subjects as follows:

- Data Integrity, Privacy, and Confidentiality

While registry data is anonymised, the risk for backwards identification in smaller populations remains a risk. The prevalence of the three cancer indications under study was sufficiently high and we took additional measures to mitigate this risk, such as data-minimisation measures (requested data only on the indispensable variables, truncation of

codes, etc.) and security measures have been put in place to safeguard the data and minimise the possibility of external parties accessing the server (file encryption, password-protected folders, etc.).

- Spurious Findings

Because of the ample geographical coverage of registries used (SEER-Medicare in the US, the Swedish National Registries, and Vilua, Healthcare Research Database in Germany), our studies count with substantial sample size. Thus, given the risk for false-positive results, no exploratory analyses for hypothesis testing were conducted, and only pre-specified analyses as per protocol and statistical analysis plan were implemented. The intention was to mitigate the risk of disseminating spurious correlations as potential causal inference that could result in policy decisions that could harm the populations under study.

3.8.2.3 Respect for Autonomy

The decision to opt out of the registries used was not within the scope of our work and as such, there was no strategy at the analytical level that could have been taken to address the potential lack of autonomous and active decision of the study subjects to participate in our research. It was not possible to gather informed consent from the registry participants because it has been replaced by the approval of a Research Ethics Committee and an administrative decision made by the authorities of each database holder. So, all the measures to mitigate the aforementioned risks (especially the risk for backwards identification) become even more important in light of the lack of informed consent.

3.8.2.4 Justice

The only risk we could identify in this concern was that our findings on certain subpopulations also could have had a correlation with socioeconomic or ethnical underserved groups. If that were the case, this correlation could have influenced the allocation of resources devoted to basic research for new treatments. To mitigate this risk, and due to the availability of data, our research analysed subpopulations according to different histological types, but we did not process data on genetic profiling or socioeconomic status and only included race as a covariate in the US but not in Europe.

4 RESULTS

In this section, we summarise the key findings of the studies; for full results, please refer to the constituting studies reproduced in this book.

4.1 PAPER I: RWE IN HEALTHCARE DECISION MAKING

The literature review resulted in 681 potentially relevant records, from which 407 unique databases were identified. Most of them were from Brazil (240), followed by Chile (62), Colombia (46), and Argentina (44). The type and diseases covered were diverse with no strong preponderance of any. The article presents an extensive description including details on geographic scope, database type, population, and outcomes captured. We validated and discussed interpretation of results with experts from each country and found that the quality of RWD varied across countries. Furthermore, RWE was not consistently used to inform health care decision making. The main use of RWE was for pharmacovigilance studies, and to lesser extent for health technology assessment, but confined to pricing decisions.

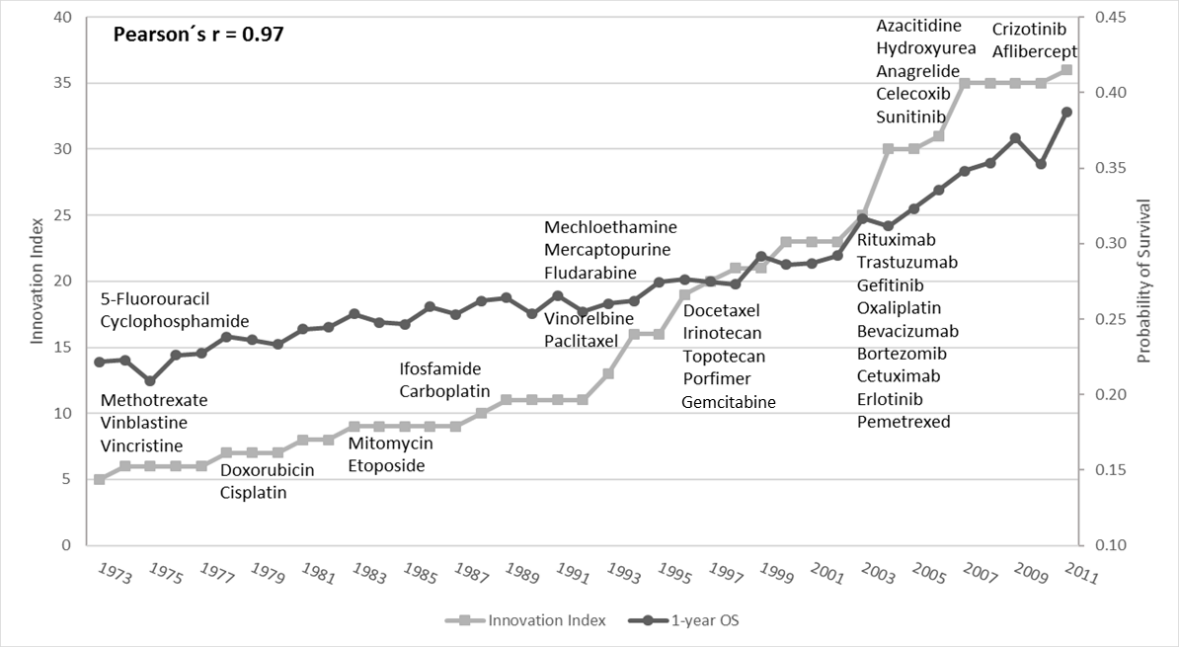
4.2 PAPER II: VALUE OF INNOVATION IN ADVANCED NSCLC

Between 1 January 1991 and 31 December 2011, 164,704 patients diagnosed with advanced NSCLC who met the protocol selection criteria were identified in SEER-Medicare and included in the study. Over the entire period, the mean age at diagnosis was 71.5 years (standard error [SE]=0.022), 43.2% were female, 79.9% were white, the most frequent histologies were adenocarcinoma (37.8%) and squamous-cell carcinoma (22.9%), and the majority of cases (56.8%) presented with metastatic disease at diagnosis (American Joint Committee on Cancer Stage IV). The population was generally comparable with a cohort extracted from the SEER Research database, except for the fact that these were younger on average (67 years) reflecting the predominance of retirees in SEER-Medicare. The long-term assessment of these covariates over time conducted on the SEER Research sample, revealed significant changes as depicted in Figure 3 of the manuscript. The series was evaluated from 1973 through 2011. In the ‘70s, the mean age at diagnosis was around 64 years, the proportion of women oscillated around 25%, and approximately 35% of patients were diagnosed with squamous cell carcinoma, followed by nearly 21% with adenocarcinoma. Almost 40 years later, average age at diagnosis was approximately 69, the proportion of women had risen to nearly 46%, and the distribution of histologies had inverted, so adenocarcinoma accounted for more than 45% of cases, compared with approximately 23% with squamous cell. The distribution of race in the sample was reflective of the race distribution of the states in which SEER is linked with Medicare, but not necessarily of the entire country.

Treatment data was reported from 60,400 patients in SEER-Medicare, which was used to derive the MMV, together with information from the scientific literature and treatment guidelines. In total, 38 chemotherapies were identified, five (13.2%) of which were FDA approved prior to their initial use in patients with advanced NSCLC. From the rest of therapies used off-label, five (13.2%) were later granted FDA approval in this indication.

When observing the graphical representation of the InnovInd over time (see Figure 8), two rounds of innovation can be spotted, one in the mid-nineties and another one in the early 2000s.

Figure 8. Innovation Index and Innovation Frontier (One-year OS) over Time



Abbreviation: OS = overall survival

Based on the SEER-Research sample, median OS increased from five months in 1973 to nine months in 2012 and the corresponding one-year OS rose from 22% to 39%, respectively. During this period, the InnovInd was well correlated with one-year OS (Pearson´s r2: 0.97).

The models run on IPL data from SEER-Medicare between 1991 and 2012, all confirmed this association, with HRs 0.97 to 0.98 (depending on the proxy) and extremely narrow CIs. The assessment of one-year OS also confirmed these findings. The predictive value of covariates was in line with the literature.

4.3 PAPER III: COMORBIDITIES AND OUTCOMES IN ADVANCED OR METASTATIC NSCLC

After implementing the protocol selection criteria, 3,834 patients diagnosed with advanced NSCLC in Skåne or Västra Götaland, and they were matched to 15,332 subjects from the general population. The matching was successful in balancing the cohorts with respect to age at diagnosis (mean 68.8 years in both cohorts) and proportion of women (47.1% in both cohorts). Most patients with advanced NSCLC were diagnosed when disease was already at metastatic stage (79.8%) and most cases were adenocarcinomas (61.4%), followed by squamous cell carcinomas (22.4%). Table 3 presents an extract of the results (refer to Tables 2 and 3 in the article for a complete list of diagnoses examined as well as number of events, time at risk, and rates).

Table 3. Prevalence of Comorbidities at Baseline and Incidence of New Diagnoses during Follow-up

Diagnoses under study	Prevalence at Baseline	Incidence during follow-up
	OR (95% CI)	**IRR (95% CI)***
ANAEMIA, NEUTROPENIA	**2.19 (1.81, 2.64)**	**11.8 (10.2, 13.6)**
Anaemia - aplastic and other	2.66 (2.12, 3.35)	15.7 (13.2, 18.6)
Anaemia – nutritional	1.59 (1.17, 2.17)	0.8 (0.5, 1.4)
Neutropenia	0.44 (0.06, 3.51)	248.2 (78.7, 782.4)
CARDIOVASCULAR DISEASE	**1.41 (1.31, 1.52)**	**2.9 (2.7, 3.1)**
Arrhythmia	1.46 (1.27, 1.68)	3.0 (2.5, 3.5)
Congestive heart failure	1.76 (1.21, 2.57)	2.7 (1.6, 4.5)
Hypertension	1.12 (1.03, 1.22)	1.3 (1.1, 1.5)
Impairment of LVEF	0.93 (0.41, 2.13)	3.3 (1.7, 6.4)
Ischaemia	1.29 (1.05, 1.58)	1.7 (1.2, 2.3)
MI	1.35 (0.96, 1.89)	2.9 (2.1, 4.0)
Stroke	2.02 (1.61, 2.53)	2.9 (2.3, 3.7)
Thromboembolism (arterial only)	4.37 (1.93, 9.92)	20.7 (9.3, 46.2)
Thromboembolism (pulmonary)	5.65 (3.81, 8.38)	39.1 (27.9, 54.9)
Thromboembolism (venous only)	4.05 (3.11, 5.28)	10.6 (8.4, 13.3)
CENTRAL NERVOUS SYSTEM	**1.50 (1.29, 1.73)**	**9.2 (8.2, 10.4)**
Brain metastases	130.12 (31.84, 531.83)	562.6 (280.0, 1130.5)
Dementia	0.52 (0.35, 0.77)	0.5 (0.3, 0.9)
Depression	1.39 (1.17, 1.66)	1.3 (1.0, 1.7)
GASTROINTESTINAL	**1.22 (0.97, 1.53)**	**2.3 (1.8, 2.9)**
Abscess	0.93 (0.51, 1.70)	3.3 (2.0, 5.6)
Dehydration	1.14 (0.38, 3.47)	4.3 (2.3, 8.2)
Hepatic impairment	1.79 (1.24, 2.59)	1.5 (0.9, 2.5)
Perforation	0.80 (0.09, 6.85)	12.0 (2.9, 50.0)
Stomatitis	1.04 (0.42, 2.56)	7.2 (3.2, 16.0)
INFECTION	**2.50 (2.25, 2.77)**	**2.3 (2.0, 2.6)**
Pneumonia	3.37 (2.35, 4.82)	18.5 (14.0, 24.5)
Upper and lower respiratory tract	2.41 (2.16, 2.68)	1.2 (1.0, 1.4)
METABOLIC	**1.09 (0.98, 1.22)**	**2.2 (1.9, 2.6)**
Diabetes	1.08 (0.96, 1.21)	2.0 (1.6, 2.3)
Hypothyroidism	1.39 (0.94, 2.06)	3.3 (2.4, 4.6)
MUSCULOSKELETAL	1.49 (1.25, 1.79)	1.3 (1.0, 1.7)
Arthritis	1.57 (1.24, 1.99)	1.4 (1.0, 2.1)
Osteoporosis	1.36 (1.05, 1.75)	1.2 (0.8, 1.7)
RESPIRATORY	**7.22 (6.46, 8.07)**	**8.4 (7.3, 9.7)**
COPD	5.76 (5.05, 6.56)	9.3 (7.7, 11.3)
Dyspnoea	8.55 (7.12, 10.26)	7.2 (5.8, 8.8)
ILD	6.16 (3.21, 11.82)	15.4 (6.3, 37.8)
SKIN	**1.22 (0.93, 1.62)**	**1.7 (1.3, 2.3)**
Rash	1.22 (0.83, 1.78)	1.9 (1.3, 2.7)
Skin hypersensitivity	1.25 (0.46, 3.41)	18.9 (8.4, 42.7)

Diagnoses under study	Prevalence at Baseline	Incidence during follow-up
	OR (95% CI)	IRR (95% CI)*
ANY DIAGNOSIS IN THE LIST	**2.44 (2.27, 2.63)**	**4.4 (4.2, 4.6)**

* Patients with a pre-existing diagnosis were excluded from the set at risk for that diagnosis, except those with recurring MI or stroke

Abbreviations: CI = confidence interval; COPD = chronic obstructive pulmonary disease; ILD = interstitial lung disease; IRR = incidence rate ratio; LVEF = left ventricular ejection fraction; MI = myocardial infarction; OR = odds ratio

At time of diagnosis, 64.8% of aNSCLC patients presented with at least one of these pre-existing conditions and, the prevalence of comorbidities was significantly higher than in the general population (OR any comorbidity = 2.44 [95% CI 2.27–2.63]). The highest contrast between cohorts was in the prevalence of respiratory diseases present in 22.7% of the patients (OR = 7.22 [95% CI 6.46–8.07]), infectious diseases present in 17% of patients (OR 2.50 [95% CI 2.25–2.77]), and cardiovascular diseases present in 35.4% of patients (OR = 1.41 [(95% CI 1.31–1.52]). While less frequently, patients with advanced NSCLC were still more prone than the general population to suffer from musculoskeletal disorders (OR = 1.49; [95% CI 1.25–1.79]) and depression (OR = 1.39; 95% [CI 1.17–1.66]). Prevalence of disorders of the gastrointestinal and metabolic systems and skin-related comorbidities, which were recorded for 2.5%, 11.6%, and 1.7% of the total population, respectively. A separate note for brain metastases and dementia is warranted. The former was recorded in only 1.7% of aNSCLC patients but the OR was 130.12; 95% CI 31.84–531.83. The latter was a surprising finding as it appeared to be less prevalent in the aNSCLC cohort than in the reference (OR = 0.52; [95% CI 0.35–0.77]).

During follow-up, patients in the aNSCLC cohort were at higher risk of receiving a new diagnosis of anaemia (IRR = 11.8 [95% CI 10.2–13.6]), diseases of the central nervous system (IRR = 9.2 [95% CI 8.2–10.4)], or respiratory disorders (IRR = 8.4 [95% CI 7.3–9.7)]. Additionally, the incidence of brain metastasis following the diagnosis of aNSCLC was significantly high (298.5 per 1,000 person-years; [95% CI 274.6–323.9]).

Certain conditions more often already present at aNSCLC diagnosis than in the general population, were also more frequently newly diagnosed after. This was the case of cardiovascular diseases, (OR = 1.41; [95% CI 1.31–1.52] and IRR = 2.9; [95% CI 2.7–3.1]), anaemia or neutropenia (OR = 2.19; [95% CI 1.81–2.64] and IRR = 11.8; [95% CI 10.2–13.6]), pneumonia (OR = 3.37; [95% CI 2.35–4.82] and IRR = 18.5; [95% CI 14.0–24.5]) and respiratory affections (OR = 7.22; [95% CI 6.46–8.07] and IRR = 8.4; [95% CI 7.3–9.7]). Finally, as expected, mortality was significantly higher in the cancer group.

4.4 PAPER IV: TREATMENT PATTERNS FOLLOWING THE INTRODUCTION OF NEW THERAPIES FOR METASTATIC CRPC

Between 1 January 2013 and 31 December 2015, 447 patients insured by the sickness funds contributing to the Vilua Healthcare Research Database (GFL back then) were diagnosed with mCRPC and met the eligibility criteria to be included in our study. Overall, mean age

was 72.99 years and metastasis location involved bone in most patients (70%), though this information was missing for 21.5% of patients. In general, patients were in poor health already at index, as the mean CCI was 8.1 out of 24 (Maximum Comorbidity Score in the version updated by Quan et al.) (Quan, Li et al. 2011) and had, on average, two hospitalisations for any cause in the year before the cancer diagnosis. A breakdown of the most common comorbidities (threshold: >15% of patients) is presented in Table 2 in the manuscript and shows that hypertension (75%), metastasis (68%), and disorders of lipoprotein metabolism and other lipidaemias (51%). In the comparison across subgroups defined by frontline therapy, we see that younger patients tended to start with docetaxel (mean age docetaxel: 69.8, [95% CI 68.4–71.2]; Mean age overall: 72.1 [95% CI 72.1–73.4]) but the comorbidity profile did not vary systematically across groups.

As shown in the treatment sequence tree in Figure 2 of the manuscript, the most common first-line therapy was abiraterone (48%), followed by docetaxel (34%), and enzalutamide (13%). Too few patients received cabazitaxel (0.5%). While most patients starting with abiraterone (61%) or enzalutamide (72%) did not progress to later lines of treatment during the observation period, most in the docetaxel group (66%) received second-line therapy (of which 64% was abiraterone and 15% was enzalutamide). Table 4 presents the results of number of patients and proportions, as well as results of the treatment duration analyses. It has been restructured and corrected as we detected errors in the CIs in the version in the manuscript. An erratum has been sent to the journal. Since most patients started treatment with abiraterone or docetaxel, their use diminished in later lines. Inversely, the proportion of patients treated with enzalutamide increased with the successive lines of therapy, as did the use of cabazitaxel, though its update was very low. In general, patients were more persistent when treated with abiraterone, particularly in first and second line, when median duration was seven and five months respectively, almost doubling duration of treatment with docetaxel. Furthermore, in first and second line, r-mean is consistently higher than the median except for patients treated with abiraterone, for whom the opposite holds true. This is due to the fact that proportionally more patients receiving abiraterone were censored than in the other groups.

Table 4. Treatment Duration (Months), per Line of Treatment and Therapy

Line	Therapy	Number of Patients*	Fraction	Median Duration	95% CI	P-value*	r-Mean Duration
1st	Abiraterone	216	48.3%	6.9	[8.3 - 10.7]	0.000	5.2
	Docetaxel	169	37.8%	3.5	[3.4 - 4.1]	1.000	3.7
	Enzalutamide	60	13.4%	4.2	[3.8 - 5.5]	0.067	4.5
2nd	Abiraterone	106	49.1%	5.0	[5.8 - 8.7]	0.000	4.8
	Cabazitaxel	13	6.0%	4.2	[3 - 5.9]	0.053	4.6
	Docetaxel	50	23.1%	2.8	[2.4 - 3.5]	1.000	3.1
	Enzalutamide	47	21.8%	2.9	[2.7 - 4.2]	0.236	3.9
3rd	Abiraterone	37	35.2%	3.1	[3.5 - 6.6]	0.050	3.9
	Cabazitaxel	13	12.4%	2.8	[1.9 - 6.1]	0.523	3.3
	Docetaxel	14	13.3%	3.7	[2.6 - 4.1]	1.000	3.5

	Enzalutamide	41	39.0%	2.9	[3.3 - 6.4]	0.078	4.3
4th	Abiraterone	16	30.8%	3.3	[2 - 6.9]	0.611	3.9
	Cabazitaxel	8	15.4%	2.1	[1.2 - 3.4]	0.221	2.6
	Docetaxel	6	11.5%	3.6	[1.3 - 6]	1.000	4.3
	Enzalutamide	22	42.3%	3.5	[2.7 - 6.3]	0.528	3.7

* Cells with fewer than 5 patients were eliminated.

**P-value from t-test against docetaxel. The omnibus p-value for the treatment length is: <0,0001

Abbreviation: CI = confidence interval

4.5 PAPER V: LIFECYCLE VALUE OF TRASTUZUMAB

According to our estimations, more than 15,000 women with HER2+ BC were treated with trastuzumab in Sweden between 2000 and 2021 (3,936 with MBC and 11,134 with EBC). Overall, 77% of patients diagnosed with HER2+ EBC received treatment, but this varied with age as trastuzumab was administered to 86% of women younger than 70 years, but only 52% of those aged 70 or older.

Table 5 is an extract of key findings from the article. Total incremental societal costs over the entire period were 4.73 billion in Swedish krona (SEK), and it was evenly split between EBC and MBC. Incremental indirect costs were negative as productivity losses were lower than they would have been without trastuzumab, so they partially offset the higher direct medical costs. The composition of direct medical costs was very different between EBC and MBC as, in the former group, over 90% were trastuzumab acquisition cost but, in the latter, the proportion of cost of breast cancer care was similar to the drug cost.

Table 5. Summary Results. Incremental Costs and Benefits, Trastuzumab vs. Standard of Care. Base Case (Worst/Best-case Scenarios)

Total Costs and Benefits in Sweden, 2000 through 2021	**MBC**	**EBC**	**Total**
Number of patients treated	3,937	11,135	15,071
Societal Costs			
Total direct medical costs (Million SEK)			
Trastuzumab acquisition cost	1,412	3,787	5,199
Trastuzumab administration cost	34	130	163
Costs of breast cancer care	1,293 (930-1,595)	216 (158-269)	1,508 (1,088-1,863)
Total cost of informal care (Million SEK)			
Trastuzumab-associated	32	118	150
Other	76 (55-94)	37 (28-46)	113 (82-140)
Total productivity costs (Million SEK)			
Trastuzumab-associated	7	25	31
Other	-596 (-433 to -729)	-1,645 (-1,286 to -1,983)	-2,241 (-1,719 to -2,711)
Total Societal Costs (Million SEK)	**2,216 (1,994-2,403)**	**2,512 (2,804-2,237)**	**4,728 (4,798-4,639)**
Health Benefits			

Total Costs and Benefits in Sweden, 2000 through 2021	**MBC**	**EBC**	**Total**
Life years gained per treated patient	1.51	1.49	1.49
QALYs gained per treated patient	0.9	0.72	0.78
Total life years gained	**6,841 (4,893-8,470)**	**19,003 (13,592-24,110)**	**25,844 (18,485-3,2579)**
Total QALYs gained	**4,075 (3,026-4,945)**	**9,363 (6,983-11,605)**	**13,437 (10,008-16,549)**
Value of Trastuzumab			
Mean cost/QALY (Thousand SEK)	544 (659-486)	268 (402-193)	352 (479-280)
Monetary value of health benefits (Million SEK)*	4,075 (3,026-4,945)	9,363 (6,983-11,605)	13,437 (10,008-16,549)
Total net monetary value (Million SEK)*	**1,859 (1,032-2,543)**	**6,851 (4,179-9,368)**	**8,710 (5,211-11,911)**
Share of value appropriated by the innovator	**0.432 (0.578-0.357)**	**0.356 (0.476-0.288)**	**0.374 (0.5-0.304)**

Abbreviations: EBC = early breast cancer; MBC = metastatic breast cancer; QALY = quality-adjusted life year; SEK = Swedish krona

The use of trastuzumab resulted in 25,844 life-years and 13,437 QALYs gained which, assuming a SEK 1 million willingness-to-pay threshold (Hultkrantz and Svensson 2008), resulted in health gains equivalent to SEK 13.4 billion. The mean cost per QALY gained was SEK 352,000 and the total net value delivered by trastuzumab over the past 20 years in Sweden was SEK 8.71 billion. Overall, the innovator appropriated 37% of the welfare surplus created by trastuzumab, and the Swedish society got the rest.

The findings were consistent in sensitivity analyses, though the magnitude changed as in the best-case scenario, net value was almost SEK 12 billion and the innovator appropriated only 30% of the surplus, and in the worst-case scenario, net value dropped to SEK 5.2 billion and the split between innovator (producer) and society (consumer) was even. Finally, OS inputs calibrated in the EBC model were validated graphically against actual survival estimates from NKBC data, showing good correspondence.

5 DISCUSSION

5.1 KEY FINDINGS IN CONTEXT

The starting point of my doctoral studies was a simple question: Has the introduction of innovative oncology therapies had a significant impact in clinically and patient-relevant outcomes when evaluated retrospectively based on real-world evidence? The short answer is yes, but the nature and magnitude of the impact differs across indications. Let us examine in which dimensions (see Section 1.3.2) these rewardable innovations have created value, and how RWE enabled this assessment.

In the US, examining RWD of therapies used in the care of patients with NSCLC over 21 years, we found that only ten out of 38 had been approved for NSCLC. Half of them were secondary indications (paclitaxel, docetaxel, gemcitabine, bevacizumab, and pemetrexed disodium) but still ranked in the top 10 in terms of uptake. Carboplatin, the backbone of NSCLC treatment, was never approved in this indication by the FDA. This was not uncommon at the time and other prominent innovations in other cancers such as paclitaxel, were never submitted to the FDA and went directly into use based on relatively small clinical trials. Both these cases proved to be effective in routine care so, over the years and around the world, their adoption was fast and their usage high. Going back to the study in NSCLC, the therapies used, in the aggregate, did deliver therapeutic benefit in terms of effectiveness as OS almost doubled, which arguably supports the assertion that secondary indications too constitute rewardable innovation. The therapies approved in the indication (mostly from 2004 and onwards) also met other criteria in the list. Five of them were targeted agents (bevacizumab targeting VEGF, erlotinib and getifinib for patients with activating EGFR mutations, and cetuximab and crizotinib for those with ALK+ tumors) which could be considered a step-change in disease management and the shift towards oral medications contributes to patient convenience (e.g., gefitinib, erlotinib, and certinib have always been ingested and, for other formerly infused therapies, oral administration is now and option, as is the case of topotecan, etoposide, or vinorelbine).

RWE could have also enabled the assessment of other attributes in our list such as public health benefits, economic benefit, and dynamic effects but they were not in scope in this study (focus of Paper V instead). Lastly, the severity of the disease, already recognised at the time of introduction of these therapies, was confirmed in our study by the persisting poor prognosis despite survival improvements. Equally, the extensive off-label during the period observed, can be considered evidence of high unmet need, as physicians felt the need to innovate on their own, even before clinical development and regulatory processes had a final say. The end of follow-up in our study (December 2012), missed the latest wave of innovation that introduced into the treatment armamentarium the new immunotherapies (such as nivolumab, pembrolizumab, atezolizumab, ipilimumab, avelumab, and durvalumab) and more targeted therapies (ALK-inhibitors such as alectinib, brigatinib, and lorlatinib, and EGFR-inhibitors such as afatinib, Osimertinib, icotinib, dacomitinib, almonertinib, olmutinib,

and rociletinib). Thus, the re-examination of the value delivered in routine clinical practice by these rewardable innovations in the care of NSCLC patients, is warranted.

Our assessment of the impact of new oncology therapies in the care of patients with mCRPC in Germany was more limited, as the short follow-up hampered the evaluation of the most important dimension of value, i.e., therapeutic benefit. Nevertheless, the patterns of adoption of these innovations that offered alternatives to the only therapy available in this indication for many years (docetaxel), qualify them as a disease management step-change. The multiplicity of regimens and sequences found in the data, suggest that physicians were searching to optimize disease management strategy given the therapeutic unmet need. If we consider that treatment and prescription patterns reveal how patients and physicians value these innovations, the perception of value differed across these new therapies. While abiraterone was adopted fast and incorporated as the standard of care in first and second line, the uptake of cabazitaxel was very low through the entire period. This may be due to the fact that cabazitaxel is very similar to docetaxel, almost a “me-too” chemotherapy that did not displace docetaxel and did not become an alternative to rechallenge because patients progressing after initial treatment may not be in a condition that allows for them to receive another cytotoxic agent. As in NSCLC, the innovative therapies with higher uptake were administered orally (abiraterone and enzalutamide, as opposed to the infused more-mature docetaxel), thus creating value in the patient convenience dimension. Finally, it is worth noting that in our study we could not assess two other innovations introduced at the time in this indication due to the lack of data. Radium-223, is a radioactive isotope that radiates alpha particle in low levels to target metastases forming or growing in the bones, limiting damage in the surrounding tissue. Sipuleucel-T, approved in Germany by the GBA in March 2015, is an autologous cell-based immunotherapy, a cancer vaccine or sorts, that activates the immune response targeted against prostatic acid phosphatase (PAP). The value of these two potentially rewardable innovations is also worth evaluating when RWD on sufficient patients followed-up for sufficient time has accrued, as well as assess the therapeutic benefit of abiraterone, cabazitaxel, and enzalutamide in terms of effectiveness and real-world safety.

Trastuzumab, as a single innovation evaluated in Sweden over 20 years, allowed for us to ascertain its value contribution in almost all the dimensions. It has contributed significant therapeutic benefit for patients with HER2+ BC, in terms of extended survival and averted or delayed relapses, which also resulted in societal gains. Decreased productivity losses offset one third of direct medical costs, fulfilling the economic benefit criterion as well. In 2000, when trastuzumab was first introduced, it became one of the first targeted agents (together with tamoxifen and aromatase inhibitors) that disrupted the way breast cancer was treated. Patient convenience also improved with the shift from intra-venous to subcutaneous administration. This actionable marker, together with the presence/absence of hormone receptor, made testing and tumour profiling part of the routine diagnostic workup and key drivers in the treatment decision. One of the most interesting aspects of this study is that it allowed for us to assess the more complex attributes of rewardable innovation, the dynamic effects. Not only was trastuzumab in the discovery path or enabled the clinical development

of follow-on and add-on innovations such as trastuzumab emtansine (Kadcyla), trastuzumab deruxtecan (Enhertu), and pertuzumab (Perjeta), but it also allowed patients to live long enough to be able to benefit from other innovations being introduced more recently, such as neratinib (Nerlynx), tucatinib (Tukysa) and margetuximab-cmkb (Margenza).

5.2 STRENGTHS

A general strength of our work was the wide geographic scope that covered countries with very different healthcare systems and processes for the assessment of innovation in oncology, including the US, two countries in Europe, and four in Latin America. Furthermore, our studies counted with large sample sizes, which represents a general advantage of RWE when compared with RCTs. And our work benefited from the use of RWD in other ways described below.

5.2.1 Neutrality

The secondary use of retrospective observational data granted our research with certain objectivity in the assessment of study measures. Capturing routine clinical practice, the conduct of the constituting studies had no influence on patient behaviour or interfere with physicians' decision, thus eliminating the risk of Hawthorne effect (Wickstrom and Bendix 2000). Similarly, it minimized the risk of compliance bias, as follow-up routines and treatment persistence were not affected by the study conduct. We also avoided the risk of recall and of reporting bias because all study measures were based on data captured in real time and by the healthcare provider (not reported by the patients).

5.2.2 Representativeness and External Validity

The cohort identification criteria in the protocols of our studies were less restrictive than the selection criteria in the respective pivotal RCTs. Thus, the study population in each case was more representative, and results more generalizable because patients were included irrespective of their age or comorbidity burden. Furthermore, we studied the value actually delivered by these therapies in real-world conditions, regardless of the circumstances in which factors external to the efficacy of a treatment, may also affect the effectiveness, such as healthcare setting, supportive care, insurance coverage, etc. This representativeness provides decision makers with a conservative, yet realistic, appraisal of value delivered in their jurisdiction; an important consideration given the issues with transferability of results from RCTs across geographies (Petrakis, Kontogiorgis et al. 2019).

5.2.3 Sensitivity Analyses

When dealing with challenges with the data or the specification of study measures, we chose the analytic approach that would yield more conservative estimates as base case and, if needed, conducted sensitivity analyses to assess alternative scenarios.

In Paper V, we combined trial efficacy data with real-world mortality through modelling, to isolate the effects of trastuzumab over time more precisely, "discounting" OS improvements

due to other advances in the healthcare system and the general extension of life expectancy in Sweden over the 20 years under study. Then, the uncertainty around these efficacy measures was tested in best-/worst-case scenarios.

A strength of data sources used from the US and Sweden, was that they capture complete follow-up until death or end of study period (i.e., practically no loss to follow-up) and in all settings (i.e., not limited to in-hospital fatalities). Thus, we leveraged these complete data on all-cause mortality (irrespective of cause of death), to account for competing risks that are in the causal path, especially adverse effects and other unwanted consequences of the disease and treatments under study. Furthermore, in Paper II, survival measures were estimated for the entire patient population, despite the fact that out of a total of 164,704 patients included, only 60,400 had records of active treatment. Thus, results were conservative, but they did also reflect the inability of treatments introduced until 2012, to offer a therapeutic alternative for most patients.

5.3 LIMITATIONS

The use of secondary data is also associated with certain limitations, of which out work was not exempt. In general, since RWD is not collected for research purposes, issues such as variable availability, coverage, data collection and quality control procedures, and coding practices are not optimized for the individual studies so they may impact the study conduct in different ways. In the Introduction, we grouped the methodological challenges derived from the use of RWD to assess value of innovation in four moments (see Section 1.5.4). Now, we discuss the limitations of our studies following the same structure.

5.3.1 Enrolment and Risk of Information Bias

Biological plausibility had been pre-established in development programs of the therapies considered in our studies, but residual risk of misclassification bias persisted in the assessment of covariates. Paper III offered additional insights in this respect. Since the study focused on advanced or metastatic disease, patients had been living with cancer inadvertently for some time before diagnosis. Furthermore, it is likely that when the disease became symptomatic and they sought care, a battery of diagnostic tests was conducted to ascertain and confirm the NSCLC diagnosis and determine the best course of action. Thus, the definition of baseline may affect the categorization of a disease or condition that was first detected relatively contemporaneously to the diagnosis of NSCLC. If we considered it a comorbidity at baseline, we would control for it in the analyses of outcomes, but if we considered it an outcome of the treatment, it would be in the causal path and not controlled for in the regressions. We conducted a sensitivity analysis evaluating the impact of different baseline assessment windows on the results, which was presented in an abstract and poster at ISPOR Annual European Congress in 2017 (Linden, Hernaez et al. 2017). The results show a statistically significant difference in the incidence unwanted outcomes in the NSCLC cohort, relative to the general population.

Furthermore, the use of secondary data to enrol patients means that the circumstances of their qualification for inclusion in the study may vary across sites or regions, which may lead to information or even selection bias. For example, in Paper V, we relied on the data collected by the NKBC on patients diagnosed with HER2+ BC with Sweden; however, a recent study found important variability of positivity across regions. This relatively wide range may be reflective of varying testing practices in the country (Acs, Fredriksson et al. 2021). In our study, the base case was built on clinical parameters on progression, relapse, and survival from clinical trials and the data from the survival estimates were validated against those obtained from the NKBC, thus avoiding any selection bias this could have caused. Yet, this data was used to estimate the numbers treated and, paucity in testing translates in fewer patients who can benefit from targeted treatment.

5.3.2 Index Date and Risk of Time-Dependant Bias

In our studies in lung cancer, patients were indexed at first diagnosis and comorbidities and other co-variates assessed at baseline to reduce the risk of wrongfully correcting for other characteristics in the causal pathway (potential mediators) (Johnson, Bartman et al. 2013, Hernan and Robins 2020). In oncology, new treatments are typically introduced in later lines of therapy so, accounting for all time since diagnosis was a data-efficient way to minimize the impact of left truncation. This was not possible in the study about prostate cancer because not all patients were observed since first diagnosis, due to the short lookback period and the switching of sickness funds.

A limitation of the study published in Paper II was the lack of patient-level data on diagnostics used over time so we could not correct fully for stage migration over time. In 2002, Positron Emission Tomography and Computed Tomography (PET-CT) scans were introduced as routine diagnostic practices and the improved precision impacted the proportion of patients diagnosed in stage IIIB or IV, which we also corroborated in the data (see Figure 3 in Paper II). While the use of a flexible semi-parametric model adjusted for time-varying covariates, we could not fully eliminate the risk of bias due to the Will Rogers Phenomenon because, part of the improvements in survival may be explained by the fact that some patients with stage IV who were wrongfully recorded as stage IIIB in the earlier years had poorer survival than those correctly staged in the same year (Singer 1990, Delgado-Rodríguez and Llorca 2004).

5.3.3 Randomization and Risk of Confounding

A key limitation of our studies was that the data sources used lacked clinical information on important clinical (certain biomarkers, performance status, etc.) and lifestyle (diet, exercise, etc.) risk factors that should have been tested as potential confounders (and dealt with if confirmed). Furthermore, the limited granularity of the data prevented us from conducting quantitative bias analysis to qualify the residual confounding.

Furthermore, in Paper II, the different aggregation levels were meant to balance the need for the proxies to represent the collective capability of a healthcare system to incorporate

innovation, against the risk of bias due to ecological fallacy (Roumeliotis, Abd ElHafeez et al. 2021). Yet, none of these proxies allowed to robustly address the inherent risks of channelling bias and confounding by indication (Petri and Urquhart 1991, Salas, Hofman et al. 1999, Wolfe, Flowers et al. 2002). In Paper IV, we did confirm the presence of channelling as patients initiating treatment with the older chemotherapy, docetaxel, were systematically younger than those receiving the newer abiraterone. This tendency could have been present in the assessment of other potential confounders, though we lacked the data to evaluate it.

5.3.4 Outcome Assessment Risk of Information Bias

In general, the completeness of the data used in our studies was a strength to minimize this risk for the assessment of survival measures. However, we lacked data on variables needed to assess other important measures in dimensions in which rewardable innovation generates value. For example, Paper III presents results of the study objectives that could be realized but the original protocol included the assessment of effectiveness and safety of new therapies introduced in Sweden for the treatment of aNSCLC. Since we could not access data on treatments administered in in-hospital setting, these analyses could not be conducted, and we only were able to assess survival and adverse effects in the overall population.

6 CONCLUSIONS

The first conclusion we draw from investigating these innovative therapies is that the way in which they were introduced in routine care and their uptake varied. In some cases, the institutional setup for their assessment was not determinant of their adoption, as in the case of carboplatin to treat aNSCLC in the US. Despite not having been evaluated by the FDA, physicians incorporated it into their armamentarium swiftly and a plethora of studies have evaluated its use in routine practice since its introduction. In other cases, therapies that were fully evaluated in all settings with large trials, approved by regulators and HTAs, and upheld in confirmatory trials and real-world studies, have been adopted but some eligible patients still go untreated, as is the case of trastuzumab to treat HER2+ BC in Sweden.

The journey of rewardable innovations from discovery to optimal use in routine care consists of several successive milestones, some are institutionalized transparent processes to assess evidence of different types, others are less formal and transparent, but they all determine the way in which we extract the value of rewardable innovation.

The next conclusion pertains to the evidence base. The research we conducted using RWD to assess the value delivered by innovative therapies in oncology demonstrates that RWE is fit to evaluate benefit in certain dimensions. RWE is preferable to RCTs to evaluate dimensions pertaining to the characterization of the target population, such as the prevalence of comorbidities to ascertain disease severity at presentation, or patient/caregiver convenience. It is also preferable to assess dimensions that require a characterization of the disease such as prevalence (to ascertain rare indications), public health priority (equity considerations) and unmet needs in routine clinical practice as evidenced in treatment patterns. Lastly, RWE is also preferable to RCTs to investigate impact of innovation on contextual dimensions such as economic benefits (cost savings, impact on productivity, etc.) or dynamic effects (especially for incremental innovation).

Regarding the assessment of clinical effects, we found that the use of RWD presents significantly more challenges and limitations than in the other dimensions. While the assessment of outcomes in routine clinical practice may provide valuable insights for the validation of comparative effects, RWE is not preferable to RCTs when it comes to evaluating therapeutic benefit.

Based on the literature, fundamentals and rationale of regulatory requirements, and our experience in the research conducted for Papers II-V, we conclude that the best evidence base for informed decision making in the assessment of the value of innovation combines both RCTs and RWE.

7 POINTS OF PERSPECTIVE

Rewardable innovation ought to deliver therapeutic benefit, but that is not enough to claim premium pricing that goes beyond conventional criteria of cost per QALY. If a new therapy only delivers therapeutic benefit, health-economic analyses allow to estimate a reward that is proportionate to the improved efficacy or safety (e.g., setting a price that maintains the new intervention under the same level of WTP threshold as standard of care). To depart from this threshold, rewardable innovation must also generate value in other dimensions. The optimal approach to measure and account for this disproportionate benefit, is still a matter for discussion.

In our research, we confirmed that several innovative therapies in oncology have delivered therapeutic benefit and value in other dimensions. Though, other new oncology treatments have failed to make good on their promise. Their rewards were not always adjusted to reflect the disproportionality in value delivered. Many more new oncology therapies were introduced since, and even more are in the pipeline. Their degree of true innovation is still to be determined but many will qualify for premium pricing under the current conditions.

Adding to the pressure of these upcoming innovations on health budgets, is the growth in cancer prevalence. So, what can be done? A fair mechanism of reward that incentivises innovation must also be affordable, to be sustainable. So, how to optimize resource allocation decisions with these constrains?

Our work was not about answering this question but, any attempt to answer it will require reliable evidence and minimal uncertainty. This is what our work was about. The use of RWE along the life cycle of potentially innovative therapies, to evaluate therapeutic benefits and other types of value actually delivered may support a fair reward mechanism that adapts to incorporate new information generated after the treatment is in use. We have seen versions and experiences of this principle in the past, though few are true success stories.

Many RWD resources currently available are not suitable for this high-stakes function. It is in the hands of their stewards to make the necessary changes to raise to the occasion. The European Commission sent to the European Parliament and the Council of the EU a proposed Regulation to create the European Health Data Space (EHDS). If approved, the EHDS would enable the generation of data with sufficient quality, completeness, reliability, and recency that might convince all the relevant stakeholders to start the conversation.

8 REFERENCES

Zehntes Buch Sozialgesetzbuch - Sozialverwaltungsverfahren und Sozialdatenschutz - (SGB X) § 80 Verarbeitung von Sozialdaten im Auftrag.

Acs, B., I. Fredriksson, C. Rönnlund, C. Hagerling, A. Ehinger, A. Kovács, R. Røge, J. Bergh and J. Hartman (2021). "Variability in Breast Cancer Biomarker Assessment and the Effect on Oncological Treatment Decisions: A Nationwide 5-Year Population-Based Study." **13**(5): 1166.

Agnafors, S., A. Norman Kjellström, J. Torgerson and M. Rusner (2019). "Somatic comorbidity in children and adolescents with psychiatric disorders." Eur Child Adolesc Psychiatry **28**(11): 1517-1525.

Agnafors, S., J. Torgerson, M. Rusner and A. N. Kjellström (2020). "Injuries in children and adolescents with psychiatric disorders." BMC Public Health **20**(1): 1273.

Ali, R., M. Hanger and T. Carino (2011). "Comparative effectiveness research in the United States: a catalyst for innovation." Am Health Drug Benefits **4**(2): 68-72.

Angelis, A., A. Lange and P. Kanavos (2018). "Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries." Eur J Health Econ **19**(1): 123-152.

Arnold, M., M. J. Rutherford, A. Bardot, J. Ferlay, T. M. Andersson, T. Myklebust, H. Tervonen, V. Thursfield, D. Ransom, L. Shack, R. R. Woods, D. Turner, S. Leonfellner, S. Ryan, N. Saint-Jacques, P. De, C. McClure, A. V. Ramanakumar, H. Stuart-Panko, G. Engholm, P. M. Walsh, C. Jackson, S. Vernon, E. Morgan, A. Gavin, D. S. Morrison, D. W. Huws, G. Porter, J. Butler, H. Bryant, D. C. Currow, S. Hiom, D. M. Parkin, P. Sasieni, P. C. Lambert, B. Møller, I. Soerjomataram and F. Bray (2019). "Progress in cancer survival, mortality, and incidence in seven high-income countries 1995-2014 (ICBP SURVMARK-2): a population-based study." Lancet Oncol **20**(11): 1493-1505.

Aronson, J. K., R. E. Ferner and D. A. Hughes (2012). "Defining rewardable innovation in drug therapy." Nat Rev Drug Discov **11**(4): 253-254.

Augustovski, F., A. Alcaraz, J. Caporale, S. Garcia Marti and A. Pichon Riviere (2015). "Institutionalizing health technology assessment for priority setting and health policy in Latin America: from regional endeavors to national experiences." Expert Rev Pharmacoecon Outcomes Res **15**(1): 9-12.

Backenroth, D. (2021). "How to choose a time zero for patients in external control arms." Pharm Stat **20**(4): 783-792.

Baird, L. G., R. Banken, H. G. Eichler, F. B. Kristensen, D. K. Lee, J. C. Lim, R. Lim, C. Longson, E. Pezalla, T. Salmonson, D. Samaha, S. Tunis, J. Woodcock and G. Hirsch (2014). "Accelerated access to innovative medicines for patients in need." Clin Pharmacol Ther **96**(5): 559-571.

Barlow, L., K. Westergren, L. Holmberg and M. Talback (2009). "The completeness of the Swedish Cancer Register: a sample survey for year 1998." Acta Oncol **48**(1): 27-33.

Beauchamp, T. L. and J. F. Childress (October 2012). Principles of Biomedical Ethics, Oxford University Press

Beletsi, A., V. Koutrafouri, E. Karampli and E. Pavi (2018). "Comparing Use of Health Technology Assessment in Pharmaceutical Policy among Earlier and More Recent Adopters in the European Union." Value Health Reg Issues **16**: 81-91.

Belloso, W. H. (2020). "On Innovation." Ther Innov Regul Sci **54**(5): 1068-1075.

Berger, M. L., H. Sox, R. J. Willke, D. L. Brixner, H. G. Eichler, W. Goettsch, D. Madigan, A. Makady, S. Schneeweiss, R. Tarricone, S. V. Wang, J. Watkins and C. Daniel Mullins (2017). "Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making." Pharmacoepidemiol Drug Saf **26**(9): 1033-1039.

Berghöfer, A., D. G. Göckler, J. Sydow, C. Auschra, L. Wessel and M. Gersch (2020). "The German health care Innovation Fund - An incentive for innovations to promote the integration of health care." J Health Organ Manag **34**(8): 915-923.

Breslow, N. E. and N. E. Day (1987). "Statistical methods in cancer research. Volume II--The design and analysis of cohort studies." IARC Sci Publ(82): 1-406.

Brooke, H. L., M. Talbäck, J. Hörnblad, L. A. Johansson, J. F. Ludvigsson, H. Druid, M. Feychting and R. Ljung (2017). "The Swedish cause of death register." Eur J Epidemiol **32**(9): 765-773.

Buchanan, J. and S. Wordsworth (2015). "Welfarism versus extra-welfarism: can the choice of economic evaluation approach impact on the adoption decisions recommended by economic evaluation studies?" Pharmacoeconomics **33**(6): 571-579.

Busse, R., M. Blümel, F. Knieps and T. Bärnighausen (2017). "Statutory health insurance in Germany: a health system shaped by 135 years of solidarity, self-governance, and competition." Lancet **390**(10097): 882-897.

Canfield, S., M. J. Kemeter, P. G. Febbo and J. Hornberger (2018). "Balancing Confounding and Generalizability Using Observational, Real-world Data: 17-gene Genomic Prostate Score Assay Effect on Active Surveillance." Rev Urol **20**(2): 69-76.

Carbonneil, C., F. Quentin and S. H. Lee-Robin (2009). "A common policy framework for evidence generation on promising health technologies." Int J Technol Assess Health Care **25 Suppl 2**: 56-67.

Claxton, K., M. Sculpher, S. Palmer and A. J. Culyer (2015). "Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients?" Health Econ **24**(1): 1-7.

Coast, J. (2009). "Maximisation in extra-welfarism: A critique of the current position in health economics." Soc Sci Med **69**(5): 786-792.

Cox, E., B. C. Martin, T. Van Staa, E. Garbe, U. Siebert and M. L. Johnson (2009). "Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report--Part II." Value Health **12**(8): 1053-1061.

Culyer, A. J. and K. Chalkidou (2019). "Economic Evaluation for Health Investments En Route to Universal Health Coverage: Cost-Benefit Analysis or Cost-Effectiveness Analysis?" Value Health **22**(1): 99-103.

de Solà-Morales, O., D. Cunningham, M. Flume, P. M. Overton, N. Shalet and S. Capri (2018). "DEFINING INNOVATION WITH RESPECT TO NEW MEDICINES: A SYSTEMATIC REVIEW FROM A PAYER PERSPECTIVE." International Journal of Technology Assessment in Health Care **34**(3): 224-240.

Delgado-Rodríguez, M. and J. Llorca (2004). "Bias." **58**(8): 635-641.

Deverka, P. A., D. C. Lavallee, P. J. Desai, J. Armstrong, M. Gorman, L. Hole-Curry, J. O'Leary, B. W. Ruffner, J. Watkins, D. L. Veenstra, L. H. Baker, J. M. Unger and S. D. Ramsey (2012). "Facilitating comparative effectiveness research in cancer genomics: evaluating stakeholder perceptions of the engagement process." J Comp Eff Res **1**(4): 359-370.

DiMasi, J. A. (2013). "Innovating by developing new uses of already-approved drugs: trends in the marketing approval of supplemental indications." Clin Ther **35**(6): 808-818.

Drummond, M., G. de Pouvourville, E. Jones, J. Haig, G. Saba and H. Cawston (2014). "A comparative analysis of two contrasting European approaches for rewarding the value added by drugs for cancer: England versus France." Pharmacoeconomics **32**(5): 509-520.

Drummond, M., M. J. Sculpher, K. Claxton, G. L. Stoddart and G. W. Torrance (2015). Methods for the economic evaluation of health care programmes. Oxford, United Kingdom, Oxford University Press.

Dutfield, G. (2017). "Healthcare innovation and patent law's 'pharmaceutical privilege': is there a pharmaceutical privilege? And if so, should we remove it?" Health Econ Policy Law **12**(4): 453-470.

Early Breast Cancer Trialists' Collaborative, g. (2021). "Trastuzumab for early-stage, HER2-positive breast cancer: a meta-analysis of 13 864 women in seven randomised trials." Lancet Oncology **22**(8): 1139-1150.

Edwards, R. T., J. M. Charles and H. Lloyd-Williams (2013). "Public health economics: a systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues." BMC Public Health **13**: 1001.

Eichler, H. G., L. G. Baird, R. Barker, B. Bloechl-Daum, F. Borlum-Kristensen, J. Brown, R. Chua, S. Del Signore, U. Dugan, J. Ferguson, S. Garner, W. Goettsch, J. Haigh, P. Honig, A. Hoos, P. Huckle, T. Kondo, Y. Le Cam, H. Leufkens, R. Lim, C. Longson, M. Lumpkin, J. Maraganore, B. O'Rourke, K. Oye, E. Pezalla, F. Pignatti, J. Raine, G. Rasi, T. Salmonson, D. Samaha, S. Schneeweiss, P. D. Siviero, M. Skinner, J. R. Teagarden, T. Tominaga, M. R. Trusheim, S. Tunis, T. F. Unger, S. Vamvakas and G. Hirsch (2015). "From adaptive licensing to adaptive pathways: delivering a flexible life-span approach to bring new drugs to patients." Clin Pharmacol Ther **97**(3): 234-246.

Eichler, H. G., K. Oye, L. G. Baird, E. Abadie, J. Brown, C. L. Drum, J. Ferguson, S. Garner, P. Honig, M. Hukkelhoven, J. C. Lim, R. Lim, M. M. Lumpkin, G. Neil, B. O'Rourke, E. Pezalla, D. Shoda, V. Seyfert-Margolis, E. V. Sigal, J. Sobotka, D. Tan, T. F. Unger and G. Hirsch (2012). "Adaptive licensing: taking the next step in the evolution of drug approval." Clin Pharmacol Ther **91**(3): 426-437.

Enewold, L., H. Parsons, L. Zhao, D. Bott, D. R. Rivera, M. J. Barrett, B. A. Virnig and J. L. Warren (2020). "Updated Overview of the SEER-Medicare Data: Enhanced Content and Applications." J Natl Cancer Inst Monogr **2020**(55): 3-13.

Erice statement on drug innovation (2008). Br J Clin Pharmacol **65**(3): 440-441.

Esmail, L. C., J. Roth, S. Rangarao, J. J. Carlson, R. Thariani, S. D. Ramsey, D. L. Veenstra and P. Deverka (2013). "Getting our priorities straight: a novel framework for stakeholder-informed prioritization of cancer genomics research." Genet Med **15**(2): 115-122.

European Medicines Agency (EMA) (2011). Assessment Report for Jevtana (cabazitaxel). Procedure No.: EMEA/H/C/002018.

European Medicines Agency (EMA) (2012). Zytiga (abiraterone) Assessment Report. Procedure no. EMEA/H/C/002321/II/0004/G

European Medicines Agency (EMA) (2013). CHMP assessment report. Xtandi (enzalutamide). Procedure No EMEA/H/C/002639.

Evans, D. (2003). "Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions." J Clin Nurs **12**(1): 77-84.

Ferlay, J., M. Colombet, I. Soerjomataram, D. M. Parkin, M. Piñeros, A. Znaor and F. Bray (2021). "Cancer statistics for the year 2020: An overview." Int J Cancer.

Fortinguerra, F., S. Perna, R. Marini, A. Dell'Utri, M. Trapanese and F. Trotta (2021). "The Assessment of the Innovativeness of a New Medicine in Italy." Front Med (Lausanne) **8**: 793640.

Frutos Perez-Surio, A., M. Gimeno-Gracia, M. A. Alcacera Lopez, M. A. Sagredo Samanes, M. D. P. Pardo Jario and M. D. T. Salvador Gomez (2019). "Systematic review for the development of a pharmaceutical and medical products prioritization framework." J Pharm Policy Pract **12**: 21.

Funk, M. J. and S. N. Landi (2014). "Misclassification in administrative claims data: quantifying the impact on treatment effect estimates." Curr Epidemiol Rep **1**(4): 175-185.

Garattini, L., M. Badinella Martini and P. M. Mannucci (2022). "Pharmaceutical patenting in the European Union: reform or riddance." Internal and Emergency Medicine **17**(3): 937-939.

Garattini, L. and A. Padula (2019). "HTA for pharmaceuticals in Europe: will the mountain deliver a mouse?" Eur J Health Econ.

Garrison, L. P., Jr. (2010). "Rewarding value creation to promote innovation in oncology: The importance of considering the global product life cycle." Oncologist **15 Suppl 1**: 49-57.

Garrison, L. P., Jr., P. J. Neumann, P. Erickson, D. Marshall and C. D. Mullins (2007). "Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report." Value Health **10**(5): 326-335.

Golan, O., P. Hansen, G. Kaplan and O. Tal (2011). "Health technology prioritization: which criteria for prioritizing new technologies and what are their relative weights?" Health Policy **102**(2-3): 126-135.

Goodliffe, J. (2003). The hazards of time-varying covariates. Proceeding of the annual meeting of the American Political Science Association.

Griffith, S. D., M. Tucker, B. Bowser, G. Calkins, C. J. Chang, E. Guardino, S. Khozin, J. Kraut, P. You, D. Schrag and R. A. Miksad (2019). "Generating Real-World Tumor Burden Endpoints from Electronic Health Record Data: Comparison of RECIST, Radiology-Anchored, and Clinician-Anchored Approaches for Abstracting Real-World Progression in Non-Small Cell Lung Cancer." Adv Ther **36**(8): 2122-2136.

Gyrd-Hansen, D. (2005). "Willingness to pay for a QALY: theoretical and methodological issues." Pharmacoeconomics **23**(5): 423-432.

Hajdu, S. I. (2011). "A note from history: landmarks in history of cancer, part 1." Cancer **117**(5): 1097-1102.

Hajdu, S. I. (2012). "A note from history: landmarks in history of cancer, part 3." Cancer **118**(4): 1155-1168.

Hajdu, S. I. (2012). "A note from history: landmarks in history of cancer, part 4." Cancer **118**(20): 4914-4928.

Hajdu, S. I. and F. Darvishian (2013). "A note from history: landmarks in history of cancer, part 5." Cancer **119**(8): 1450-1466.

Hajdu, S. I. and M. Vadmal (2013). "A note from history: Landmarks in history of cancer, Part 6." Cancer **119**(23): 4058-4082.

Hajdu, S. I., M. Vadmal and P. Tang (2015). "A note from history: Landmarks in history of cancer, part 7." Cancer **121**(15): 2480-2513.

Hanahan, D. (2022). "Hallmarks of Cancer: New Dimensions." Cancer Discov **12**(1): 31-46.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.

Hanahan, D. and R. A. J. c. Weinberg (2000). "The hallmarks of cancer." **100**(1): 57-70.

Hatswell, A. J., K. Deighton, J. T. Snider, M. A. Brookhart, I. Faghmous and A. R. Patel (2022). "Approaches to Selecting "Time Zero" in External Control Arms with Multiple Potential Entry Points: A Simulation Study of 8 Approaches." Med Decis Making: 272989x221096070.

Hernan, M. and J. Robins (2020). "Causal Inference: What if. Boca Raton: Chapman & Hill/CRC."

Hernán, M. A., B. C. Sauer, S. Hernández-Díaz, R. Platt and I. Shrier (2016). "Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses." J Clin Epidemiol **79**: 70-75.

Higgins, J. P. T. and S. Green (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011], The Cochrane Collaboration.

Hofmann, S., J. Branner, A. Misra and H. Lintener (2021). "A Review of Current Approaches to Defining and Valuing Innovation in Health Technology Assessment." Value Health **24**(12): 1773-1783.

Hofmarcher, T., P. Lindgren, N. Wilking and B. Jönsson (2020). "The cost of cancer in Europe 2018." Eur J Cancer **129**: 41-49.

Horwich, A., J. Hugosson, T. de Reijke, T. Wiegel, K. Fizazi, V. Kataja, M. Panel and O. European Society for Medical (2013). "Prostate cancer: ESMO Consensus Conference Guidelines 2012." Ann Oncol **24**(5): 1141-1162.

Hultkrantz, L. and M. Svensson (2008). "Värdet av liv." Ekonomisk debatt **2**: 5-16.

Institute for Quality and Efficiency in Health Care, I. (2017). General Methods (Allgemeine Methoden) Version 5.0. 10 July 2017. Accessed November 2019.

ISPOR Task Force (2013). "Using Real World Data for Coverage and Payment Decisions: The Ispor Real World Data Task Force Report."

Johnson, E. S., B. A. Bartman, B. A. Briesacher, N. S. Fleming, T. Gerhard, C. J. Kornegay, P. Nourjah, B. Sauer, G. T. Schumock, A. Sedrakyan, T. Stürmer, S. L. West and S. Schneeweiss (2013). "The incident user design in comparative effectiveness research." Pharmacoepidemiol Drug Saf **22**(1): 1-6.

Justo, N., M. A. Espinoza, B. Ratto, M. Nicholson, D. Rosselli, O. Ovcinnikova, S. Garcia Marti, M. B. Ferraz, M. Langsam and M. F. Drummond (2019). "Real-World Evidence in Healthcare Decision Making: Global Trends and Case Studies From Latin America." Value Health **22**(6): 739-749.

Justo, N., J. Nilsson, B. Korytowsky, J. Dalen, T. Madison and A. McGuire (2020). "Retrospective observational cohort study on innovation in oncology and progress in survival: How far have we gotten in the two decades of treating patients with advanced non-small cell lung cancer as a single population?" PLoS One **15**(5): e0232669.

Kennedy, I. (2009). Appraising the value of innovation and other benefits. A short study for NICE.

Khullar, D., J. Fisher and A. Chandra (2019). "Trickle-down innovation and the longevity of nations." Lancet **393**(10187): 2272-2274.

Kobelt, G. and Office of Health Economics (London England) (2002). Health economics : an introduction to economic evaluation. London, Office of Health Economics.

Kruse, M. and O. Thoreson (2021). "The prevalence of diagnosed specific back pain in primary health care in Region Västra Götaland: a register study of 1.7 million inhabitants." Prim Health Care Res Dev **22**: e37.

Lakdawalla, D. N., J. A. Doshi, L. P. Garrison, Jr., C. E. Phelps, A. Basu and P. M. Danzon (2018). "Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]." Value Health **21**(2): 131-139.

Lash, T., T. VanderWeele, S. Haneuse and K. Rothman (2021). Modern Epidemiology, Wolters Kluwer.

Lash, T. L., M. P. Fox and A. K. Fink (2009). Applying quantitative bias analysis to epidemiologic data. Dordrecht ; New York, Springer.

Leahy, T. P., S. Kent, C. Sammon, R. H. Groenwold, R. Grieve, S. Ramagopalan and M. Gomes (2022). "Unmeasured confounding in nonrandomized studies: quantitative bias analysis in health technology assessment." J Comp Eff Res **11**(12): 851-859.

Lichtenberg, F. R. (2013). "The Effect of Pharmaceutical Innovation on Longevity: Patient Level Evidence from the 1996-2002 Medical Expenditure Panel Survey and Linked Mortality Public-use Files." Forum Health Econ Policy **16**(1): 1-33.

Lichtenberg, F. R. (2014). "Pharmaceutical innovation and longevity growth in 30 developing and high-income countries, 2000–2009." Health Policy and Technology **3**(1): 36-58.

Lichtenberg, F. R. (2019). "How many life-years have new drugs saved? A three-way fixed-effects analysis of 66 diseases in 27 countries, 2000-2013." International Health **11**(5): 403-416.

Lichtenberg, F. R. (2022). "The effect of pharmaceutical innovation on longevity: Evidence from the U.S. and 26 high-income countries." Econ Hum Biol **46**: 101124.

Lichtenberg, F. R., P. Grootendorst, M. Van Audenrode, D. Latremouille-Viau and P. Lefebvre (2009). "The impact of drug vintage on patient survival: a patient-level analysis using Quebec's provincial health plan data." Value Health **12**(6): 847-856.

Lidgren, M., N. Wilking, B. Jonsson and C. Rehnberg (2007). "Health related quality of life in different states of breast cancer." Qual Life Res **16**(6): 1073-1081.

Lidgren, M., N. Wilking, B. Jonsson and C. Rehnberg (2007). "Resource use and costs associated with different states of breast cancer." Int J Technol Assess Health Care **23**(2): 223-231.

Linden, A. and P. R. Yarnold (2017). "Using classification tree analysis to generate propensity score weights." J Eval Clin Pract **23**(4): 703-712.

Linden, S., A. M. B. Hernaez, J. Redig, J. Nilsson and N. Justo (2017). "Impact Of Different Baseline Definitions On The Incidence Of Relevant Outcomes Associated With Cancer Following An Advanced Non-Small Cell Lung Cancer (NSCLC) Diagnosis " Value in Health **20**(9): A413-A413.

Linden, S., J. Redig, A. Banos Hernaez, J. Nilsson, D. B. Bartels and N. Justo (2020). "Comorbidities and relevant outcomes, commonly associated with cancer, of patients newly diagnosed with advanced non-small-cell lung cancer in Sweden." Eur J Cancer Care (Engl) **29**(1): e13171.

Lindskog, M., T. Wahlgren, R. Sandin, J. Kowalski, M. Jakobsson, S. Lundstam, B. Ljungberg and U. Harmenberg (2017). "Overall survival in Swedish patients with renal cell carcinoma treated in the period 2002 to 2012: Update of the RENCOMP study with subgroup analysis of the synchronous metastatic and elderly populations." Urol Oncol **35**(9): 541.e515-541.e522.

Löfgren, L., S. Eloranta, K. Krawiec, A. Asterkvist, C. Lönnqvist and K. Sandelin (2019). "Validation of data quality in the Swedish National Register for Breast Cancer." BMC Public Health **19**(1): 495.

Löfvendahl, S., M. E. C. Schelin and A. Jöud (2020). "The value of the Skåne Health-care Register: Prospectively collected individual-level data for population-based studies." Scand J Public Health **48**(1): 56-63.

Ludvigsson, J. F., C. Almqvist, A. K. Bonamy, R. Ljung, K. Michaëlsson, M. Neovius, O. Stephansson and W. Ye (2016). "Registers of the Swedish total population and their use in medical research." Eur J Epidemiol **31**(2): 125-136.

Ludvigsson, J. F., P. Otterblad-Olausson, B. U. Pettersson and A. Ekbom (2009). "The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research." Eur J Epidemiol **24**(11): 659-667.

Lyman, G. H. (2009). Oxford American Handbook of Oncology. Cary, Oxford University Press, Incorporated.

MacEwan, J. P., S. Dennen, R. Kee, F. Ali, J. Shafrin and K. Batt (2020). "Changes in mortality associated with cancer drug approvals in the United States from 2000 to 2016." J Med Econ **23**(12): 1558-1569.

MacEwan, J. P., I. Majer, J. W. Chou and S. Panjabi (2021). "The value of survival gains from therapeutic innovations for US patients with relapsed/refractory multiple myeloma." Ther Adv Hematol **12**: 20406207211027463.

MacEwan, J. P., W. Yin, S. Kaura and Z. M. Khan (2017). "The Value of Survival Gains in Pancreatic Cancer from Novel Treatment Regimens." J Manag Care Spec Pharm **23**(2): 206-213.

Maiese, E. M., K. A. Evans, B. C. Chu and D. E. Irwin (2018). "Temporal Trends in Survival and Healthcare Costs in Patients with Multiple Myeloma in the United States." Am Health Drug Benefits **11**(1): 39-46.

Makady, A., A. de Boer, H. Hillege, O. Klungel and W. Goettsch (2017). "What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews." Value Health **20**(7): 858-865.

Manchikanti, L., D. L. Caraway, A. T. Parr, B. Fellows and J. A. Hirsch (2011). "Patient Protection and Affordable Care Act of 2010: reforming the health care reform for the new decade." Pain Physician **14**(1): E35-67.

Mapi, Research and Trust PROLABELS™

Mariotto, A. B., L. Enewold, J. Zhao, C. A. Zeruto and K. R. Yabroff (2020). "Medical Care Costs Associated with Cancer Survivorship in the United States." Cancer Epidemiol Biomarkers Prev **29**(7): 1304-1312.

Marsh, K., I. J. M, P. Thokala, R. Baltussen, M. Boysen, Z. Kalo, T. Lonngren, F. Mussen, S. Peacock, J. Watkins, N. Devlin and I. T. Force (2016). "Multiple Criteria Decision Analysis for Health Care Decision Making--Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force." Value Health **19**(2): 125-137.

Marty, M., F. Cognetti, D. Maraninchi, R. Snyder, L. Mauriac, M. Tubiana-Hulin, S. Chan, D. Grimes, A. Anton, A. Lluch, J. Kennedy, K. O'Byrne, P. Conte, M. Green, C. Ward, K. Mayne and J. M. Extra (2005). "Randomized phase II trial of the efficacy and safety of trastuzumab combined with docetaxel in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer administered as first-line treatment: the M77001 study group." J Clin Oncol **23**(19): 4265-4274.

National Board of Health and Welfare [Socialstyrelsen]. (2022, 26 June 2022). "National Cancer Register [Cancerregistret] ", from https://www.socialstyrelsen.se/statistik-och-data/register/cancerregistret/

National Cancer Institute, N. (July 2021). "Cancer Trends Progress Report ", from https://progressreport.cancer.gov/after/economic_burden

Neumann, P. J. (2004). "Why don't Americans use cost-effectiveness analysis?" Am J Manag Care **10**(5): 308-312.

Neumann, P. J. and M. C. Weinstein (2010). "Legislating against use of cost-effectiveness information." N Engl J Med **363**(16): 1495-1497.

Nicod, E., K. Berg Brigham, I. Durand-Zaleski and P. Kanavos (2017). "Dealing with Uncertainty and Accounting for Social Value Judgments in Assessments of Orphan Drugs: Evidence from Four European Countries." Value Health **20**(7): 919-926.

Nicod, E. and P. Kanavos (2016). "Scientific and Social Value Judgments for Orphan Drugs in Health Technology Assessment." Int J Technol Assess Health Care **32**(4): 218-232.

Olofsson, S., H. Norrlid, E. Karlsson, U. Wilking and G. Ragnarson Tennvall (2016). "Societal cost of subcutaneous and intravenous trastuzumab for HER2-positive breast cancer - An observational study prospectively recording resource utilization in a Swedish healthcare setting." Breast **29**: 140-146.

Parker, C., S. Gillessen, A. Heidenreich, A. Horwich and E. G. Committee (2015). "Cancer of the prostate: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up." Ann Oncol **26 Suppl 5**: v69-77.

Petrakis, I., C. Kontogiorgis, E. Nena, K. Athanasakis, V. Gougoula, I. Kotsianidis and T. C. Constantinidis (2019). "Unraveling innovation potential in the real-world setting: eighteen novel agents with twenty-six approved European indications, in the management of leukemias, lymphomas, and multiple myeloma." Expert Rev Hematol **12**(12): 1063-1075.

Petri, H. and J. Urquhart (1991). "Channeling bias in the interpretation of drug effects." Stat Med **10**(4): 577-581.

Phelps, C. E., D. N. Lakdawalla, A. Basu, M. F. Drummond, A. Towse and P. M. Danzon (2018). "Approaches to Aggregation and Decision Making-A Health Economics Approach: An ISPOR Special Task Force Report [5]." Value Health **21**(2): 146-154.

Pöllinger, B., W. Schmidt, A. Seiffert, H. Imhoff and M. Emmert (2019). "Costs of dose escalation among ulcerative colitis patients treated with adalimumab in Germany." Eur J Health Econ **20**(2): 195-203.

Potosky, A. L., G. F. Riley, J. D. Lubitz, R. M. Mentnech and L. G. Kessler (1993). "Potential for cancer related health services research using a linked Medicare-tumor registry database." Med Care **31**(8): 732-748.

Quan, H., B. Li, C. M. Couris, K. Fushimi, P. Graham, P. Hider, J. M. Januel and V. Sundararajan (2011). "Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries." Am J Epidemiol **173**(6): 676-682.

Ramagopalan, S., T. P. Leahy, J. Ray, S. Wilkinson, C. Sammon and V. Subbiah (2021). "The value of innovation: association between improvements in survival of advanced and metastatic non-small cell lung cancer and targeted and immunotherapy." BMC Med **19**(1): 209.

Ramsey, S. D., D. Veenstra, S. R. Tunis, L. Garrison, J. J. Crowley and L. H. Baker (2011). "How comparative effectiveness research can help advance 'personalized medicine' in cancer treatment." Health Aff (Millwood) **30**(12): 2259-2268.

Ray, W. A. (2003). "Evaluating medication effects outside of clinical trials: new-user designs." Am J Epidemiol **158**(9): 915-920.

Regional Cancer Centrum, R. (23 March 2022). "National Breast Cancer Quality Register (NKBC) " Retrieved 2 July 2022 from https://cancercentrum.se/samverkan/cancerdiagnoser/brost/kvalitetsregister/

Rejon-Parrilla, J. C., J. Espin and D. Epstein (2022). "How innovation can be defined, evaluated and rewarded in health technology assessment." Health Econ Rev **12**(1): 1.

Rothman, K. J. (2012). Epidemiology : an introduction. New York, NY, Oxford University Press.

Rothman, K. J., S. Greenland and T. L. Lash (2008). Modern epidemiology. Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins.

Roumeliotis, S., S. Abd ElHafeez, K. J. Jager, F. W. Dekker, V. S. Stel, A. Pitino, C. Zoccali and G. Tripepi (2021). "Be careful with ecological associations." Nephrology (Carlton) **26**(6): 501-505.

Royston, P. and M. K. Parmar (2002). "Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects." Stat Med **21**(15): 2175-2197.

Salas, M., A. Hofman and B. H. Stricker (1999). "Confounding by indication: an example of variation in the use of epidemiologic terminology." Am J Epidemiol **149**(11): 981-983.

Sampat, B. and H. L. Williams (2019). "How Do Patents Affect Follow-on Innovation? Evidence from the Human Genome." Am Econ Rev **109**(1): 203-236.

Schoenfeld, D. (1982). "Partial Residuals for the Proportional Hazards Regression-Model " Biometrika **69**(1): 239-241.

Sculpher, M., K. Claxton and S. D. Pearson (2017). "Developing a Value Framework: The Need to Reflect the Opportunity Costs of Funding Decisions." Value Health **20**(2): 234-239.

Shen, Q., M. E. C. Schelin, F. Fang and A. Jöud (2021). "Diagnostic codes of cancer in Skåne healthcare register: a validation study using individual-level data in southern Sweden." BMC Cancer **21**(1): 759.

Simonds, N. I., M. J. Khoury, S. D. Schully, K. Armstrong, W. F. Cohn, D. A. Fenstermacher, G. S. Ginsburg, K. A. Goddard, W. A. Knaus, G. H. Lyman, S. D. Ramsey, J. Xu and A. N. Freedman (2013). "Comparative effectiveness research in cancer genomics and precision medicine: current landscape and future prospects." J Natl Cancer Inst **105**(13): 929-936.

Singer, R. B. (1990). "Feinsteins report of the "Will Rodgers" phenomenon: advances in diagnosis, resulting stage migration, and their impact on 6-month lung cancer mortality by stage." J Insurance Medicine **22**(2): 4.

Slothuus, U. (2000). Economic evaluation: Theory, methods & application Health Economics Papers U. o. S. D. Faculty of Social Sciences.

Spira, A., M. B. Yurgelun, L. Alexandrov, A. Rao, R. Bejar, K. Polyak, M. Giannakis, A. Shilatifard, O. J. Finn, M. Dhodapkar, N. E. Kay, E. Braggio, E. Vilar, S. A. Mazzilli, T. R. Rebbeck, J. E. Garber, V. E. Velculescu, M. L. Disis, D. C. Wallace and S. M. Lippman (2017). "Precancer Atlas to Drive Precision Prevention Trials." Cancer Res **77**(7): 1510-1541.

Statistics, C. f. M. M. S. P. Medicare Total Enrollment tables.

Stratton, M. R., P. J. Campbell and P. A. Futreal (2009). "The cancer genome." Nature **458**(7239): 719-724.

Suissa, S. (2008). "Immortal time bias in pharmaco-epidemiology." Am J Epidemiol **167**(4): 492-499.

Suissa, S., E. E. Moodie and S. Dell'Aniello (2017). "Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores." Pharmacoepidemiol Drug Saf **26**(4): 459-468.

Sung, H., J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal and F. Bray (2021). "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries." CA Cancer J Clin **71**(3): 209-249.

Swedish National Breast Cancer Register, N. (2022). Online statistics. https://statistik.incanet.se/brostcancer/. Accessed 3 March 2022. .

Thariani, R., D. L. Veenstra, J. J. Carlson, L. P. Garrison and S. Ramsey (2012). "Paying for personalized care: cancer biomarkers and comparative effectiveness." Mol Oncol **6**(2): 260-266.

Theidel, U. and J. M. von der Schulenburg (2016). "Benefit assessment in Germany: implications for price discounts." Health Econ Rev **6**(1): 33.

Therneau T (2015). A Package for Survival Analysis in S. version 2.38, .

Thokala, P., N. Devlin, K. Marsh, R. Baltussen, M. Boysen, Z. Kalo, T. Longrenn, F. Mussen, S. Peacock, J. Watkins and M. Ijzerman (2016). "Multiple Criteria Decision Analysis for Health Care Decision Making--An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force." Value Health **19**(1): 1-13.

Toumi, M., C. Remuzat, E. El Hammi, A. Millier, S. Aballea, C. Chouaid and B. Falissard (2015). "Current process and future path for health economic assessment of pharmaceuticals in France." J Mark Access Health Policy **3**.

Trinquart, L., J. Jacot, S. C. Conner and R. Porcher (2016). "Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials." J Clin Oncol **34**(15): 1813-1819.

US National Cancer Institute, D. o. C. C. a. P. S. (26 May 2022). "SEER-Medicare Linked Data Resource." Retrieved 2 July 2022, from https://healthcaredelivery.cancer.gov/seermedicare/ and https://healthcaredelivery.cancer.gov/seermedicare/overview/publications.html

Verguet, S., J. J. Kim and D. T. Jamison (2016). "Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial." Pharmacoeconomics **34**(9): 913-923.

Vernon, J. A., J. H. Golec and J. S. Stevens (2010). "Comparative effectiveness regulations and pharmaceutical innovation." Pharmacoeconomics **28**(10): 877-887.

Wagner, G. (1991). "History of cancer registration." IARC Sci Publ(95): 3-6.

Wang, S. V., S. Schneeweiss, M. L. Berger, J. Brown, F. de Vries, I. Douglas, J. J. Gagne, R. Gini, O. Klungel, C. D. Mullins, M. D. Nguyen, J. A. Rassen, L. Smeeth and M. Sturkenboom (2017). "Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0." Pharmacoepidemiol Drug Saf **26**(9): 1018-1032.

Warren, J. L., C. N. Klabunde, D. Schrag, P. B. Bach and G. F. Riley (2002). "Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population." Med Care **40**(8 Suppl): Iv-3-18.

Weinberg, R. A. (2014). The biology of cancer. New York, Garland Science, Taylor & Francis Group.

Wenzl, M. and V. Paris (2018). Pharmaceutical Reimbursement and Pricing in Germany, OECD.

Werner, A., P. Reitmeir and J. John (2005). "[Switching sickness funds and risk compensation mechanisms in the statutory health insurance system in Germany--empirical results from the cooperative health research in the region of Augsburg (KORA)]." Gesundheitswesen **67 Suppl 1**: S158-166.

Wickstrom, G. and T. Bendix (2000). "The "Hawthorne effect"--what did the original Hawthorne studies actually show?" Scand J Work Environ Health **26**(4): 363-367.

Wilking, U., B. Jonsson, N. Wilking and J. Bergh (2010). "Trastuzumab use in breast cancer patients in the six Health Care Regions in Sweden." Acta Oncol **49**(6): 844-850.

Wolfe, F., N. Flowers, T. A. Burke, L. M. Arguelles and D. Pettitt (2002). "Increase in lifetime adverse drug reactions, service utilization, and disease severity among patients who will start COX-2 specific inhibitors: quantitative assessment of channeling bias and

confounding by indication in 6689 patients with rheumatoid arthritis and osteoarthritis." J Rheumatol **29**(5): 1015-1022.

Wolkewitz, M., A. Allignol, S. Harbarth, G. de Angelis, M. Schumacher and J. Beyersmann (2012). "Time-dependent study entries and exposures in cohort studies can easily be sources of different and avoidable types of bias." J Clin Epidemiol **65**(11): 1171-1180.

Woolf, S. H. (2000). "Evidence-Based Medicine and Practice Guidelines: An Overview." Cancer Control 7 **7**(4): 362-367.

World Health Organization. (2022). "Heealth Innovation for impact." from https://www.who.int/teams/digital-health-and-innovation/health-innovation-for-impact

World Health Organization, W. "International Classification of Diseases for Oncology." Retrieved 9 July 2022, from https://www.who.int/standards/classifications/other-classifications/international-classification-of-diseases-for-oncology

Wyss, R., S. Schneeweiss, M. van der Laan, S. D. Lendle, C. Ju and J. M. Franklin (2018). "Using Super Learner Prediction Modeling to Improve High-dimensional Propensity Score Estimation." Epidemiology **29**(1): 96-106.

Yabroff, K. R., A. Mariotto, F. Tangka, J. Zhao, F. Islami, H. Sung, R. L. Sherman, S. J. Henley, A. Jemal and E. M. Ward (2021). "Annual Report to the Nation on the Status of Cancer, Part 2: Patient Economic Burden Associated With Cancer Care." Jnci-Journal of the National Cancer Institute **113**(12): 1670-1682.