Lisbon School of Economics & Management
Universidade de Lisboa

MASTER

ACTUARIAL SCIENCE

MASTER'S FINAL WORK

INTERNSHIP REPORT

MODELLING AND FORECASTING THE SNS ONCOLOGICAL PHARMACEUTICALS EXPENDITURE

(A MULTIPLE LINEAR REGRESSION APPROACH)

TOMÁS TABORDA MARTINS

SUPERVISION

JOSÉ ANTÓNIO PALMA E SANTOS ALVES CARPINTEIRO
ONOFRE ALVES SIMÕES

JUNE - 2025

ABSTRACT

There is a growing concern in the European Union (EU) over the financial sustainability of public healthcare systems and how this relates to ageing populations. Portugal is perhaps affected by the two problems more severely. This study addresses one component of this complex situation: national expenditure on Group 16 pharmaceuticals. This is because they have a significant impact on total hospital pharmaceutical expenditure within the Portuguese National Health Service (Serviço Nacional de Saúde, SNS).

Group 16 pharmaceuticals are primarily used to treat oncological issues, so it is important to analyse the determinants of oncological expenditure within its three subgroups: cytotoxics, hormones and anti-hormones, and immunomodulatings. It is also crucial to predict the future evolution of the indicator as accurately as possible, in both absolute terms and as a percentage of total public expenditure (TPE) and gross domestic product (GDP).

Following the approach of Kadkhodamanesh et al. (2021), explanatory models for each subgroup using multiple linear regressions were constructed. These models were based on explanatory variables related to demography and disease prevalence, innovation, prevention, and pharmaceutical market. Due to the limited statistical power of these models, when predicting future oncological expenditure, direct projections were employed instead of using the models.

The results show that the variables related to prevention contribute to an acceleration of the oncological expenditure. Other variables related to demography and disease prevalence, and pharmaceutical market, also contribute to explain the growth of the oncological expenditure on some subgroups, though the effects of some of these variables diverged from theoretical expectations.

Based on projections, total expenditure on oncological pharmaceuticals in the SNS is expected to grow until 2030 representing 0.40% of the GDP, and the proportion of TPE allocated to oncology is projected to be 0.91% in 2030.

Keywords: Oncological Pharmaceuticals Expenditure, Portuguese National Health Service, Ageing, Multiple Linear Regression, Financial Sustainability

RESUMO

Existe uma preocupação crescente na União Europeia (UE) relativamente à sustentabilidade financeira dos sistemas públicos de saúde e à sua relação com o envelhecimento populacional. Portugal é, possivelmente, um dos países mais afetados por estes dois problemas. Este estudo aborda um dos componentes desta situação complexa: a despesa nacional com os medicamentos do Grupo 16. Isto deve-se ao facto de estes medicamentos terem um impacto significativo na despesa total com medicamentos hospitalares no âmbito do Serviço Nacional de Saúde (SNS).

Os medicamentos do Grupo 16 são utilizados, maioritariamente, no tratamento de patologias oncológicas, sendo importante analisar os determinantes da despesa oncológica nos seus três subgrupos: citotóxicos, hormonas e anti-hormonas, e imunomoduladores. É igualmente crucial prever, com o maior grau de precisão possível, a evolução futura deste indicador, tanto em termos absolutos como enquanto percentagem da despesa pública total (TPE) e do produto interno bruto (GDP).

Seguindo a abordagem de Kadkhodamanesh et al. (2021), foram construídos modelos explicativos para cada subgrupo com recurso a regressões lineares múltiplas. Estes modelos baseiam-se em variáveis explicativas relacionadas com demografia e prevalência de doença, inovação, prevenção, e mercado farmacêutico. Devido ao reduzido poder estatístico destes modelos, optou-se por utilizar projeções diretas, em vez dos modelos, para a previsão da despesa oncológica futura.

Os resultados mostram que as variáveis relacionadas com a prevenção contribuem para uma aceleração da despesa oncológica. Outras variáveis, relacionadas com a demografia e prevalência da doença e o mercado farmacêutico, também contribuem para explicar o crescimento da despesa oncológica em alguns subgrupos, embora os efeitos de algumas destas variáveis tenham divergido das expectativas teóricas.
Com base em projeções, estima-se que a despesa oncológica do SNS continue a crescer até 2030, representando 0.40% do GDP, sendo que a proporção da TPE alocada à oncologia deverá atingir 0.91% em 2030.

Palavras-chave: Despesa com Medicamentos Oncológicos, Serviço Nacional de Saúde, Envelhecimento, Regressão Linear Múltipla, Sustentabilidade Financeira

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere gratitude to all the individuals and institutions that, in one way or another, played a fundamental role in the completion of this work.

To my supervisor, Professor Onofre Simões, for his attentive guidance and wise advice which were essential to the improvement and successful conclusion of this work.

To *Direção-Geral do Tribunal de Contas*, for the opportunity to undertake my internship in Department V, with special thanks to the Honourable Judge Luís Cracel Viana, Councillor of the *Tribunal de Contas.*

To my supervisor at *Tribunal de Contas*, Senior Auditor Coordinator Dr. José Carpinteiro, for the knowledge and insight generously shared throughout the internship, and to the Chief Auditor Iva Maia, for her availability. I would also like to extend my appreciation to all remaining members of Department V, for the kindness with which they welcomed and treated me.

To Professor José Dias Curto, for his insightful advice during a key stage of this work.

I am also very grateful to *Registo Oncológico Nacional* (RON) and *Autoridade Nacional do Medicamento e Produtos de Saúde* (INFARMED), for providing additional data that proved extremely valuable to the progress of this work.

Last but not least, to my family, for their unwavering support throughout this journey and my entire academic and personal life. Specially, to my sister, Matilde, and my parents, Maria José and Paulo. Their presence, along with their unconditional love, understanding and support, has been the strength that has never let me fall. For always being there and continuing to stand by my side, this work is, with all my love, dedicated to them.

CONTENTS

LIST OF ACRONYMS AND ABBREVIATIONS

ACSS - *Administração Central do Sistema de Saúde*

ARIMA - Autoregressive Integrated Moving Average

DGS – *Direção-Geral da Saúde*

EC – European Commission

EU – European Union

GDP – Gross Domestic Product

INE – *Instituto Nacional de Estatística*

INFARMED - *Autoridade Nacional do Medicamento e Produtos de Saúde*

OECD - Organisation for Economic Co-operation and Development

OLS - Ordinary Least Squares

RMSE – Root Mean Squared Error

RON – *Registo Oncológico Nacional*

SNS – *Serviço Nacional de Saúde*

SSE – Sum of Squares due to Residuals

SSR – Sum of Squares due to Regression

SST - Total Sum of Squares

TPE – Total Public Expenditure

VIF – Variance Inflation Factor

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Context and Objectives

This report has been prepared as part of a five-month curricular internship at the Audit Department of the Portuguese Court of Auditors which controls the Social Affairs, *Departamento V – Setor Social e de Saúde da Direção-Geral do Tribunal de Contas* (TdC), within the scope of a Performance Audit on Access to and Expenditure on Pharmaceuticals in the Portuguese National Health Service, *Serviço Nacional de Saúde* (SNS), as stipulated in the Annual Program of the Second Section of the TdC for 2024.

Throughout the internship, I had the opportunity to participate in meetings, events, and activities that proved to be highly enriching. These experiences allowed me:

- to gain insight into the perspectives of experts and researchers in the fields of health and ageing in Portugal;
- to understand that this issue is a shared concern among other European Union countries, but also to deepen my knowledge of the healthcare sector, which was essential for the development of this work.

Among the various actions that took place it is important to highlight:

- Remote international meeting joining Israel, Albania, Slovakia, North Macedonia, Paraguay and Portugal, on the topic "Preparedness of Governments for an Ageing Population" – Audit designation and registration (January 20, 2025);
- Launch of the 2023 Reports: Ageing & Access to Healthcare – Cátedra BPI | "la Caixa" Foundation of Health Economics (January 22, 2025);
- Monitoring of the Performance Audit on Access to and Expenditure on Pharmaceuticals in the SNS;
- Development of datasets for major expenditure aggregates;
- Seminar "Insurer Payments, Hospitals Price Markups, and Manufacturer Revenues for Oncology and Specialty Drugs in the US", Nova SBE (May 29, 2025);
- Conference "Oncology 2030 – Innovation Challenges in the Fight against Cancer" – APIFARMA (June 4, 2025).

After a preliminary study, the audit focused on hospital pharmaceuticals, as these have increasingly contributed to total pharmaceuticals expenditure in recent years, compared to retail (ambulatory); another key aspect is that pharmaceuticals commonly used to treat oncological diseases (antineoplastic and immunomodulatings) has become a key driver of hospital pharmaceutical expenditure in recent years given their therapeutic complexity

and the frequent introduction of high-cost innovative pharmaceuticals, particularly for oncological indications. These pharmaceuticals fall within Group 16 of the Pharmacotherapeutic Classification, which includes antineoplastic and immunomodulatings agents and comprises pharmaceuticals used primarily in the treatment of cancer and in the modulation of the immune system. It is subdivided into three main therapeutic classes: Cytotoxics, which act directly on cancer cells to inhibit their growth or destroy them; Hormones and anti-hormones, used in hormone-dependent cancers such as breast and prostate cancer; Immunomodulatings, which either stimulate or suppress the immune response and are used in both oncology and autoimmune conditions.

Therefore, the audit sought to address two questions:

1. What are the primary determinants of the evolution of expenditure on oncological pharmaceuticals in the hospital setting within the SNS?
2. Has the state ensured the sustainability of expenditure on oncological pharmaceuticals in the hospital setting within the SNS?

Considering that the increase in expenditure on these type of pharmaceuticals coexists with an ageing population in the European Union (EU) and Portugal, we will be particularly interested, when finding the answer to question 1, in answering at the same time a more specific question: Is the ageing of the population the most determining factor in the growth of SNS expenditure on oncological pharmaceuticals?

In fact, life expectancy has increased and the proportion of the population in older age groups now exceeds the proportion in younger age groups. These demographic changes require continuous monitoring and analysis – as exemplified by the Ageing Report from the European Commission (EC) and the Health at a Glance from the Organisation for Economic Co-operation and Development (OECD) – since their impact can span multiple sectors of society, including the healthcare sector, see European Commission (2024) and OECD (2024).

1.2. Literature Review

Given its relevance in public expenditure, public healthcare expenditure - including pharmaceuticals expenditure - has been one of the major concerns of healthcare payers over the years. This is largely due to its complexity and volatility caused by a wide range

of dynamic factors that includes demographic changes in the population (European Commission, 2024).

Recognizing the critical importance of healthcare expenditure, the EC publishes the Ageing Report every three years, providing long-term projections for public expenditure on pensions, healthcare, long-term care and education in the European Union, up to 2070. The report is based on multiple scenarios accounting for different factors, supporting its Member States on healthcare planning and resource allocation. In addition, the OECD releases the Health at a Glance report every two years, in cooperation with the EC, offering a comprehensive assessment of the performance of EU health systems.

Due to the inherent unpredictability associated to healthcare and pharmaceuticals expenditure, numerous studies using different approaches have been conducted at a national level, aiming to model and to forecast future expenditure. For instance, Odnoletkova et al. (2025) give a quite exhaustive literature review on the topic.

One commonly implemented approach for modelling and forecasting future expenditure are Autoregressive Integrated Moving Average (ARIMA) models, see Kleiber and Zeileis (2008) for general reference. ARIMA models allow to forecast future values based on past observations of the expenditure in a easy way. Klazoglou and Dritsakis (2018) applied ARIMA to model total healthcare expenditure as a percentage of Gross Domestic Product (GDP) in the United States, and Bindel and Seifert (2025) used it to model and forecast antibacterial pharmaceutical consumption in Germany until 2040.

Linear regression has also been widely implemented, allowing to study the relationship between certain indicators and variables such as public healthcare and pharmaceutical expenditure, as well as income and profits in the pharmaceutical industry. It has also enabled to construct predictive models to forecast future values of these variables.

Di Matteo and Di Matteo (1998) applied a time-series cross-section regression analysis in order to examine the determinants of Canadian provincial government health expenditures and concluded that the proportion of people aged over 65 had a more significant impact than real per capita GDP or real per capita federal transfers. Seven years later, Di Matteo (2005), in a similar study using state-level data for United States and province-level data for Canada, concluded that the proportion of people aged over 65 was a significant determinant of real per capita health expenditures.

In Ecuador, Pérez et al. (2018), using a linear regression approach studied the impact of GDP and healthcare expenditures on the income and profits of pharmaceutical industry and, in India, Vilas et al. (2019) used the same method to explore the impact of some macroeconomic indicators on the growth of public expenditure. Wettermark et al. (2010), in a study developed in the metropolitan health region of Stockholm, applied a linear regression to hospital sales and dispensed pharmaceuticals in ambulatory care of each pharmacological group based on four observations, from 2006 to 2009, and then, taking into account the factors that could increase or decrease future utilization and expenditure, performed predictions for the two next years, concluding that the expenditure was expected to increase.

Sala (2020), aiming to explore the degree of influence of age and income on health expenditures per capita in Romania, applied a multiple linear regression and concluded that the proportion of population aged over 80 years contributed to an increase in the health expenditure.

Kadkhodamanesh et al. (2021), which is the main methodological reference for our work, performed a multiple regression analysis using data from 1999 to 2017, with the aim of estimating pharmaceuticals expenditure as a percentage of GDP in Iran. The independent variables were divided into five categories – macroeconomic factors, health system structure, disease profiles, pharmaceuticals system structure and innovation – and the authors concluded that, after the percentage of healthcare expenditure (health system structure), disease prevalence (disease profiles) was the variable with the greatest effect on the dependent variable.

1.3. Structure of the Report

The document is structured as follows for the development of the work: Chapters 2 and 3 discuss the weight of oncological pharmaceutical expenses in the global expenditure of the National Health Service (SNS) and the relevance and possible impact of ageing on these expenses. Chapter 4 contains theoretical elements and a description of the used variables. Chapter 5 presents the application and the results obtained, followed by an analysis. Finally, Chapter 6 draws conclusions and suggests future research.

2. HEALTH EXPENDITURE IN PORTUGAL AND IN THE EU

2.1 Overview

The first feature to consider when examining current health expenditure is the need to distinguish between private sector expenditure (private health expenditure) and public administration expenditure (public health expenditure). The former includes household out-of-pocket expenditure and voluntary health insurance, while the latter includes compulsory public and contributory schemes (Conselho das Finanças Públicas, 2024).

According to the OECD (2024a) and as shown in Figure 1, in 2022, public health expenditure in Portugal was higher than private health expenditure, similarly to the EU.

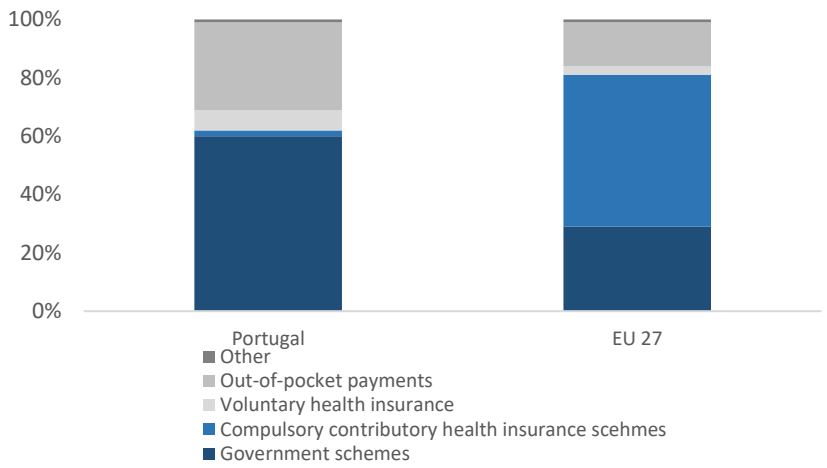

Figure 1: Components of health expenditure in 2022

Source: Health at a Glance 2024

Note: Other category refers to charities, employers, foreign and undefined schemes

In fact, in Portugal, public health expenditure accounted for 62% of total health expenditure – below the EU average - with government schemes representing 60%. On the other hand, private health expenditure represented 38% - above the EU average - with household out-of-pocket representing 30%. Although public health expenditure in Portugal is below the EU average, the health sector is heavily funded by public resources, highlighting the importance of a thorough analysis.

2.1.1 SNS expenditure

The SNS was established on September 15, 1979, with the aim of guaranteeing universal access to healthcare, irrespective of citizens' economic or social status[1]. Its establishment was of paramount importance as before its inception families predominantly depended on their own means or private institutions for healthcare services.

Among the healthcare services provided, two main levels of care stand out: primary healthcare, which includes general and family medicine consultations, and hospital care, which encompasses specialist consultations, inpatient care, and complementary diagnostic procedures. These levels of care are further complemented by the provision of continuous integrated care to individuals in a dependent situation through the *Rede Nacional de Cuidados Continuados Integrados*[2] (Conselho das Finanças Públicas, 2024).

As a model based on the public provision of healthcare services, the SNS is funded through allocations from the Government Budget[3]. According to data from the *Administração Central do Sistema de Saúde* (ACSS), the Budget transfers accounted for approximately 95.3% of the SNS total revenue of 13 625.6 M€, in 2023. The percentage is higher than that recorded in previous years, making Portugal one of the European countries with the highest share of healthcare system financed through public revenues (Conselho das Finanças Públicas, 2024).

Regarding its expenditure, which is included in public health expenditure, it is primarily composed of three elements – personnel costs, supplies and external services and inventory purchases. The expenditure with pharmaceuticals sold in community pharmacies covered by the SNS contributes to supplies and external services, while the SNS expenditure on hospital pharmaceuticals contributes to inventory purchases (Conselho das Finanças Públicas, 2024).

In 2023, according to data from ACSS and *Instituto Nacional de Estatística* (INE), the total expenditure of the SNS has been increasing over the years, amounting to 14 060.7 M€ and accounting for 12.5% of total public expenditure and 82.5% of public health expenditure, as shown in Figure 2. This increasing trend has been followed by household out-of-pocket expenditures as well.

[1] *Lei nº 56/79, de 15 de setembro* (Lei n.º 56/79 | DR)
[2] *Decreto-Lei n.º 101/2006, de 6 de junho* (Decreto-Lei n.º 101/2006 | DR)
[3] *Nº1 da Base 23 da Lei de Bases da Saúde aprovada pela Lei n.º 95/2019, de 4 de setembro* (Lei n.º 95/2019 | DR)

According to ACSS and INE, household out-of-pocket and SNS expenditure have been representing a significant portion of total health expenditure. In 2023, they correspond to 82.7% of total health expenditure, see Figure 3.

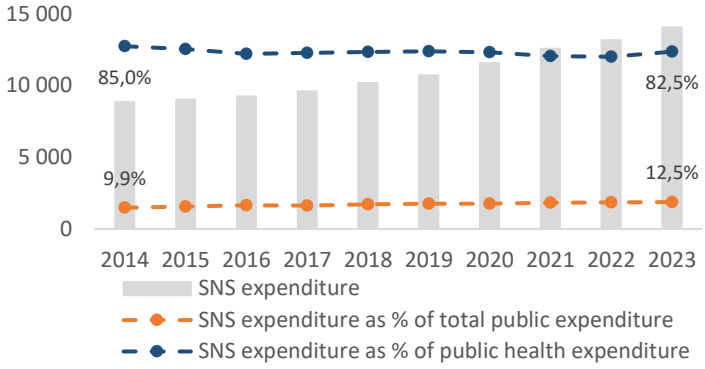

Figure 2: SNS expenditure (M€) - share of total public and health expenditure (2014-2023)

Source: ACSS *(Conta do Serviço Nacional de Saúde)* and INE (*Conta Satélite da Saúde* and *Principais Agregados das Administrações Públicas*)

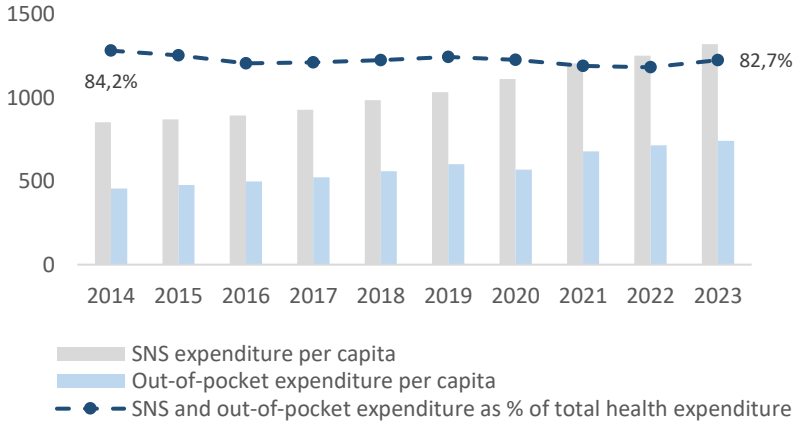

Figure 3: SNS and out-of-pocket expenditure (€) per capita - share in health expenditure (2014-2023)

Source: ACSS *(Conta do Serviço Nacional de Saúde)* and INE (*Conta Satélite da Saúde* and *Estatísticas Demográficas – 2023*)

Indeed, the rising trend in SNS expenditure has been accompanied by an increase in household out-of-pocket expenditure. However, without overlooking the significance of household out-of-pocket expenditure, it is essential to focus the analysis on SNS given its social relevance and its role as the main public health financing agent.

2.2 Pharmaceuticals Expenditure in Portugal and in the EU

The majority of EU Member States, as part of their obligation to uphold the right to health, must ensure the availability of pharmaceuticals to the population. While in many

countries outside of Europe pharmaceuticals are provided in public sector facilities that belong to the welfare, in Europe, pharmaceuticals are often supplied through private channels and largely publicly funded through a social health insurance or a national health service. Expenditure on pharmaceuticals covered by public funds includes those administered or dispensed in hospital settings during an episode of hospital care (hospital market), as well as those provided outside hospital care prescribed by doctors and dispensed in community pharmacies (retail or ambulatory market) (Vogler et. al 2011).

According to the OECD (2024a), in Portugal in 2022, expenditure on retail pharmaceuticals covered by public resources was greater than expenditure covered by private resources, see Figure 4. This is similar to what is observed in the EU; however, expenditure on retail pharmaceuticals only provides a partial explanation of the total pharmaceutical costs in the health system, since expenditure on hospital pharmaceuticals can be significant.

Regarding pharmaceuticals used in hospitals, there is no database that allows for detailed comparisons across countries, as few countries report information on this aspect (García-Goñi, 2022). However, according to the OECD (2023) and as shown in Figure 5, in Portugal, similar to some other countries, the average annual growth rate of expenditure on hospital pharmaceuticals was several times higher than the average annual growth rate of retail expenditure, between 2011 and 2021, specially due to the introduction of new treatments in oncology and immunology.

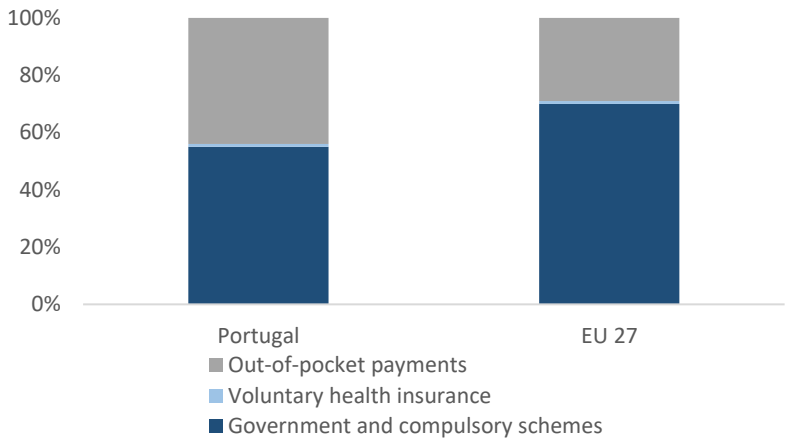

Figure 4: Expenditure on retail pharmaceuticals by type of financing in 2022

Source: Health at a Glance 2024

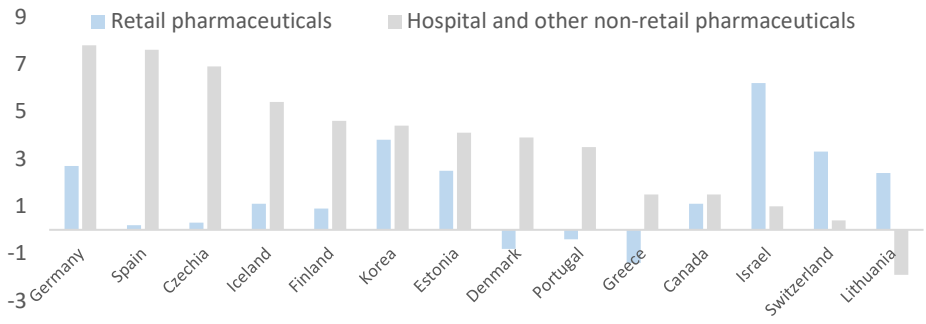

Figure 5: Annual average growth in pharmaceuticals expenditure in real terms (2011-2021)

Source: Health at a Glance 2023

In Portugal, in 2022, approximately 55% of retail pharmaceutical expenditure was publicly funded, reflecting the substantial role of the state in ensuring access to treatments. Furthermore, the significantly higher average growth rate of hospital pharmaceutical expenditure over the past decade highlights the need to focus attention on this category, given its increasing weight in total pharmaceutical expenditure and its implications for sustainability of the healthcare system.

2.2.1 SNS expenditure on pharmaceuticals

In Portugal, pharmaceuticals are subject to the processes of Reimbursement[4] and *Avaliação Prévia Hospitalar*,[5] which determine the state funding for their acquisition. The first process concerns the support provided by the State to patients for the purchase of pharmaceuticals in community pharmacies, while the second applies to pharmaceuticals intended to be acquired by entities overseen by the member of the Government responsible for the health sector – namely, SNS hospitals – establishing the conditions for their acquisition.

According to the *Autoridade Nacional do Medicamento e Produtos de Saúde* (INFARMED) and as shown in Figure 6, SNS expenditure on pharmaceuticals has shown a rising trend since at least 2014, reaching 3 553 M€ in 2023, mirroring the evolution of public expenditure, public health expenditure and SNS expenditure. This trend is observed for both retail and hospital pharmaceuticals, but expenditure on hospital pharmaceuticals has grown more significantly in recent years, having surpassed retail

[4] *Portaria 195-D/2015, de 30 de junho*
[5] *Decreto-Lei nº 97/2015, de 1 de junho* (Decreto-Lei n.º 97/2015 | DR)

pharmaceuticals expenditure in its contribution to the total expenditure in 2021 (Figure 7).

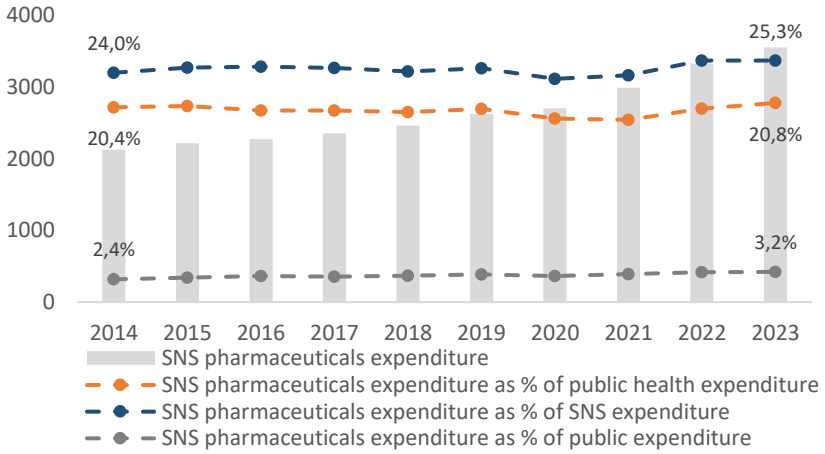

Figure 6: SNS pharmaceuticals expenditure (M€) - share in the public, public health and SNS expenditure (2014-2023)

Source: INFARMED monthly reports (*Análise de consumo de medicamentos em meio ambulatório* and *Análise de consumo de medicamentos em meio hospitalar*), ACSS *(Conta do Serviço Nacional de Saúde)* and INE (*Conta Satélite da Saúde* and *Principais Agregados das Administrações Públicas*)

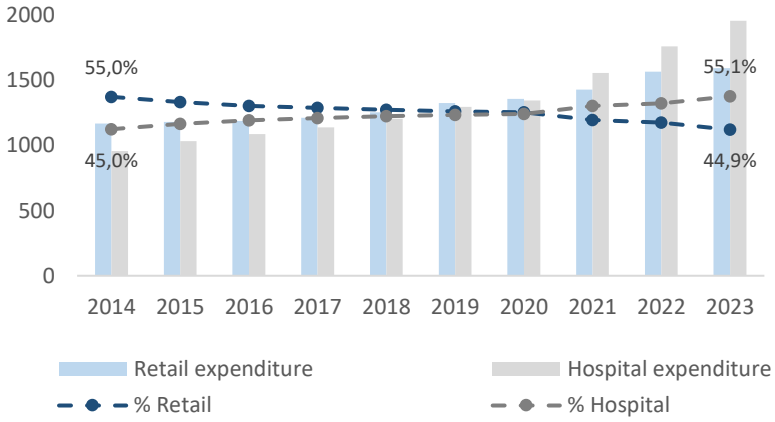

Figure 7: Retail and Hospital pharmaceuticals expenditure (M€) - share in SNS pharmaceuticals expenditure (2014-2023)

Source: INFARMED monthly reports

2.2.1.1 SNS expenditure on Group 16 pharmaceuticals

As with retail pharmaceuticals, hospital pharmaceuticals in Portugal are subject to a pharmacotherapeutic classification system that aligns with the World Health Organization's Anatomical Therapeutic Chemical classification system. [6] This system

[6] *Despacho nº 4742/2014, de 2 de abril* (Despacho n.º 4742/2014 | DR)

facilitates the quicker and more accurate identification of pharmaceuticals based on their intended therapeutic use. In this classification, pharmaceuticals or active substances are divided into different groups and subgroups according to the organ or system they are intended to treat, as well as their chemical, pharmacological, and therapeutic properties. Among the various groups, one requires particular attention: Group 16 - referred to as "Antineoplastic and Immunomodulating Pharmaceuticals" - that includes pharmaceuticals or active substances that are extensively used in the treatment of oncological diseases and in the stimulation of the immune system. This group is further divided into three subgroups - cytotoxics, hormones and anti-hormones, and immunomodulatings - that can have other therapeutic indications besides the oncological indication.

According to INFARMED, SNS expenditure on Group 16 pharmaceuticals has been increasing since at least 2014, rising from a share of 39% to nearly 50% of total hospital pharmaceuticals expenditure, see Figure 8, and the expenditure with oncological cytotoxics, hormones and anti-hormones and immunomodulatings has been also increasing representing 64,8% of the total Group 16 expenditure in 2023.

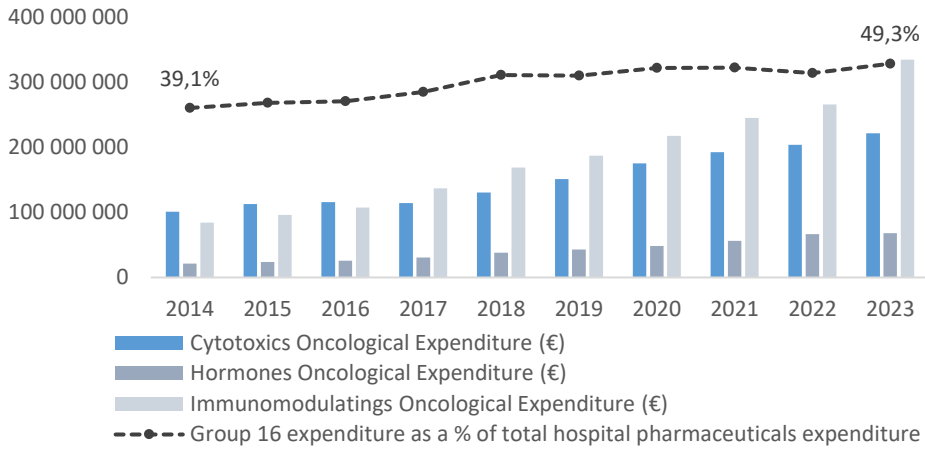

Figure 8: Contribution of Group 16 to the hospital pharmaceuticals expenditure (€) (2014-2023)

Source: INFARMED monthly reports and INFARMED database

Since this group alone accounts for nearly half of SNS hospital pharmaceutical expenditure, the need of a detailed analysis is quite evident. Being composed of three subgroups of pharmaceuticals that can have different therapeutic uses, but are commonly used in oncological diseases, and their expenditure has an increasing trend, it is important to analyse the expenditure with oncological use for each subgroup.

3. AGEING IN PORTUGAL AND IN THE EU

3.1 Overview

Ageing in Portugal and in the EU has transitioned from being merely a perception to becoming an undeniable reality. According to the OECD (2024a), the Union has been undergoing, and continues to undergo, a transformation in its demographic profile - and Portugal is no exception. As stated by the European Commission (2024), the population of older age groups is expected to continue to increase, while the population of younger age groups is projected to decline.

The ageing process in Europe has not been homogeneous, as each country has experienced the increase in the proportion of the elderly population (people over 65) and the decline in the young population (people under 15) at its own pace, since 1950. Northern and Western European countries aged earlier, with a more gradual evolution, whereas Southern European countries experienced ageing later and in a more pronounced manner. In the 1980s, Portugal began its ageing trajectory, ceasing to be one of the youngest countries. Between 1991 and 2001, the number of elderly individuals surpassed the number of young people, indicating that the ageing process in Portugal intensified during the 1990s (Gomes & Moreira, 2014). What could explain this evolution?

Population growth can be understood through the relationship between the natural balance (difference between the number of births and deaths) and the migratory balance (difference between immigration and emigration). In Portugal, since the 1950s, the population growth rate has undergone fluctuations, particularly during the 1970s and the second half of the 1980s. These shifts were driven by the 1974 Revolution and Portugal's accession to the European Economic Community, which impacted migratory balance—first, through the return of Portuguese citizens from former colonies, and later, by the reactivation of emigration (Moreira & Gomes, 2014). From mid 1980s until today, the population growth rate alternates between positive and negative, becoming negative between 2010 and 2018 due to the impact of the financial crisis, but becoming positive after 2018. According to INE, the current effective growth rate has reached its highest level, driven by a sharp increase in migratory balance, in spite of a continuous decline in the natural balance. However, the effects that mortality and fertility have on population growth greatly impacts the age structure of populations.

In Portugal, the crude mortality rate (number of deaths per 1000 inhabitants) has shown a generally declining (though irregular) trend, since the 1950s. Meanwhile, the

infant mortality rate (number of deaths of children under 1 year per 1000 live births) has decreased sharply, and life expectancy at birth has increased (Azevedo & Baptista, 2014). This pattern aligns with the epidemiological transition model proposed by Omran (2005), which defined three phases of mortality evolution - "The Age of Pestilence and Famine", "The Age of Receding Pandemics" and "The Age of Degenerative and Man-Made Diseases" – which corresponded to shifts in the predominant causes of death.

From 1950 onward, Portugal entered the third phase of the epidemiological transition, moving away from a second phase characterized by increased neonatal resistance to prevailing diseases. This transition to the third phase led to an overall decline in mortality, and a rising in life expectancy at birth (Azevedo & Baptista, 2014). This evolution in mortality and the subsequent rise in life expectancy at birth has led to an increase in the elderly population and an ageing of the top of the age pyramid (Moreira & Henriques, 2014).

The decline in fertility, which can be assessed by the total fertility rate (average number of children per woman in reproductive age), began in the 1960s and intensified after the 1974 Revolution, as there was greater participation of women in society. Since 2000, fertility rates have remained low, insufficient to ensure generational renewal, with a gradual smooth declining trend. By 2012, fertility levels reached critically low values, 1.29 according to INE. This trend of very low fertility levels has led to a progressive reduction in the young population and a consequent ageing of the base of the age pyramid (Moreira & Henriques, 2014).

The ageing of the top combined with the ageing of the base has resulted in a progressive increase in the ageing index (ratio between the elderly population and the young population) and the old-age dependency ratio (ratio between the elderly population and the working-age population), indicating that the elderly population is surpassing the young and working-age populations (people aged between 15 and 64 years). According to EUROSTAT, in 2023, the elderly population represented 23.9% of the resident population, while the young population accounted for 12.9%, leading to an ageing index of 185.3 and an old-age dependency ratio of 37.8. It is expected that this trend will continue until 2050, when both indices are projected to stabilize as shown in Figure 9, where historical data end in 2023.

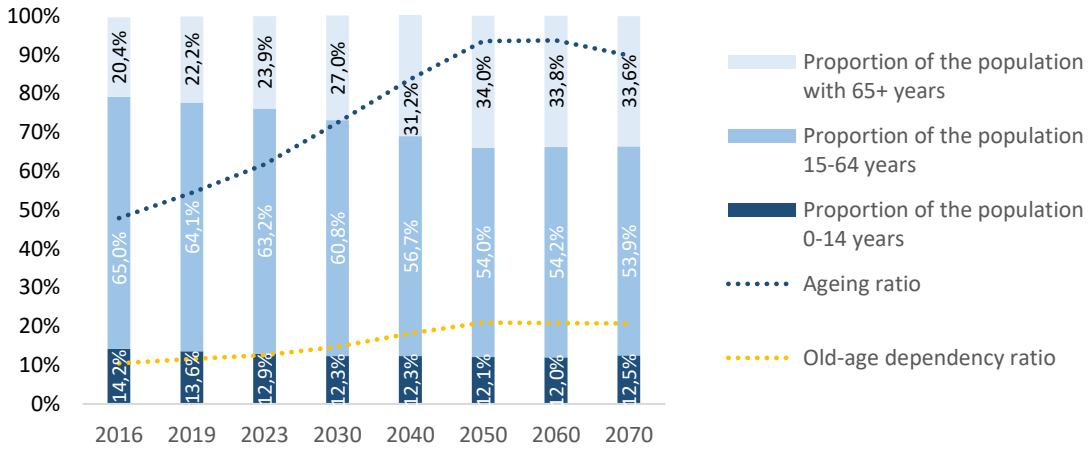

Figure 9: Age groups, ageing index and old-age dependency ratio in Portugal (2026-2070)

Source: EUROSTAT

In a clear way, Rodrigues (2018) defines the problem that Portugal faces saying "…People are dying less, but fewer are being born…There is a lack of young people, and the number of elderly continues to rise" (p. 19-20), and according to the projections, ageing will remain a problem in the future.

3.2 Ageing and Health Expenditure

Population ageing imposes fiscal challenges concerning government expenditure and revenue, with expenditure including pension and healthcare costs (Amaglobeli & Shi, 2016). According to European Commission (2024), public expenditure on pensions and health as a percentage of the GDP is expected to increase in the EU and Portugal until 2050. Regarding healthcare costs, in Portugal they are expected to increase until 2060 reaching 8.2% of the GDP which is a significant value despite being below EU average that is 9.7%.

This behaviour may be partially explained by the size and structure of the populations, since there is a relation between the age of individuals and their use of healthcare services. In fact, in 2022, according to the age-related expenditure profiles of publicly financed healthcare provision per capita as a percentage of GDP for fourteen EU countries (including Portugal), the expenditure surpassed 10% of the GDP for individuals above 65 years (European Commission, 2024).

However, beyond the impacts caused solely by changes in demographic structure, the increase in life expectancy and its relationship with disease prevalence and incidence is a crucial factor to consider (Scott, 2023).

According to Gruenberg (1977) and Verbugge (1984), an expansion of morbidity may occur, where the reduction in mortality is not due to a decrease in disease incidence but rather to a reduction in its lethality. In this sense, a reduction in mortality and an increase in longevity would be accompanied by an increase in morbidity. On the other hand, according to Fries (1989), a compression of morbidity could occur, where morbidity would be concentrated in the last years of life, leading to longevity being accompanied by better levels of health in the population. The way in which ageing will occur is, therefore, unpredictable, with possible impacts in the healthcare expenditure.

The European Commission (2024), see Figure 10, projected healthcare expenditure as a percentage of GDP for the 27 EU Member States, considering seven scenarios, including the following ones: the "baseline" scenario (which primarily captures the impact of demographic factors and moderately the impact of non-demographic factors), the "pure demographic" scenario (which captures only the impact of demographic factors), the "healthy ageing" scenario (which captures the impact of compression of morbidity), and the "no healthy ageing" scenario (which captures the impact of morbidity expansion). The study concluded that in all four scenarios, healthcare expenditure as a percentage of GDP, in Portugal, is expected to increase from 2022 to 2070, with the "no healthy ageing" and "baseline" scenarios showing the largest variation.

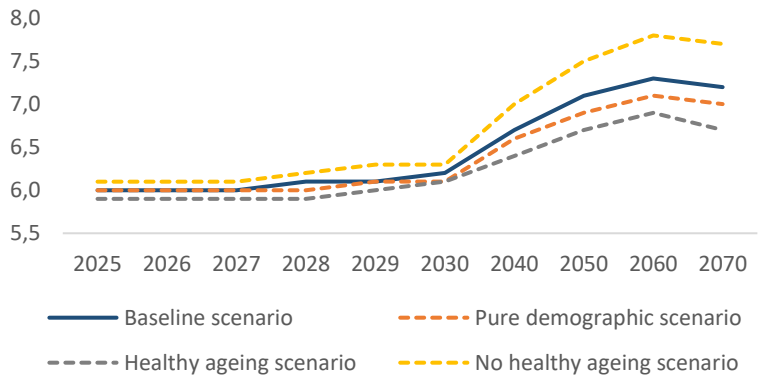

Figure 10: Healthcare expenditure as a percentage of GDP in Portugal (2025-2070)

Source: Ageing Report 2024

In summary, while it is reasonable to expect that the demographic structure of the population will place increasing pressure on healthcare expenditure, it is likely that the most significant impact will stem from the specific conditions under which the ageing process unfolds.

3.3 Pharmaceuticals Expenditure

Regarding pharmaceuticals expenditure, this is also affected by ageing. However, there are other factors that can have an impact on the expenditure.

According to Belloni et al. (2016), factors influencing growth in pharmaceuticals expenditure may be divided into three categories: demand for healthcare and pharmaceuticals, market dynamics and pharmaceuticals policies. The size and structure of the population is related to changes in demand since the tendency to develop diseases that require pharmaceuticals to their treatment increases with age.

Market dynamics may influence expenditure through the introduction of new pharmaceuticals or new indications for existing ones and both contribute to the increase of treatment options (Belloni et al. 2016). Additionally, when the patent of a pharmaceutical expires, there is the possibility of the introduction of generics and biosimilars which are bioequivalent to the original pharmaceuticals and promote price competition (Garcia-Goñi, 2022).

Pharmaceuticals policies may influence expenditure positively or negatively. While cost-containment policies, such as External Reference Pricing, may contribute to a decrease in the expenditure, an increase on coverage of health and pharmaceuticals may contribute to an increase in the expenditure (Belloni et al. 2016). External Reference Pricing is the practice of regulating the price of a pharmaceutical in a given country by comparing it with the price of the same pharmaceutical in a set of reference countries.

3.3.1 Oncological pharmaceuticals expenditure

From 1995 to 2022, the incidence of cancer in Europe increased by 58% (Manzano et al. 2025). Population growth, although moderate, and ageing are the most important factors explaining this increase, but other reasons may also explain it:

- Advancements in healthcare, as people are now surviving diseases that were previously fatal and sometimes they reach advanced ages, which increases the number of individuals at risk of developing cancer;
- Screenings, as they allow the detection of cases that might never be symptomatic;
- Cancer registration, as a more complete registration of cancer cases has allowed to register cases that previously were not recorded (Manzano et al. 2025).

In Portugal, similarly to what happens in the EU, there has been a consistent pattern of early detection of cancer and, over the years, cancer has ceased to be solely a disease

affecting older age groups and has increasingly begun to affect younger populations. Despite this trend, the older age groups continue to be the most affected (OECD, 2025).

According to the Registo Oncológico Nacional (2024), in 2021, in Portugal, individuals aged 65 and over were responsible for 36 759 new cases, accounting for 60.5% of the total and regarding the incidence rate, it exceeded 1 000 cases per 100 000 people from the age of 60. It is estimated that, in 2022, due to higher incidence rates among older age groups and the declining cancer mortality rate, individuals over the age of 70 recorded a higher prevalence compared to younger cohorts (OECD, 2025).

In fact, according to Barros and Santos (2024a), it is among the elderly population that there is a higher probability of being a chronic patient, with this probability increasing for both women and men since 2017. Moreover, the number of elderly individuals seeking care through the SNS is eight times higher than those who turn to the private sector (Barros & Santos 2024b).

According to the OECD (2024a), large inequalities in health status may exist among older people due to their socio-economic status, which impacts the exposure to risk factors and lesser access to health services during their life. In fact, older people with a lower level of education are on average in poorer health than those with higher level of education. Also, income inequalities may affect healthcare expenditure, particularly the one related to health insurance, which is expensive, see European Commission (2024),.

The greater predisposition of older age groups to develop oncological diseases aligns with the main causes of cancer. According to Manzano et al. (2025), the causes of cancer may be divided into three main components – environmental or behavioural factors, hereditary factors, and random mutations – and, as individuals age, the cellular repair mechanisms that are supposed to detect and destroy mutated cells become less effective, rendering the elderly population more exposed and vulnerable to cancerous cells. Additionally, certain risk behaviours, such as physical inactivity – which, according to the OECD (2024a), was responsible in 2023 for the onset of conditions such as depression, dementia, diabetes, cardiovascular diseases, and cancer – tend to increase with age. In 2019, 78% of people aged over 65 in the EU did not meet the minimum recommended level of physical activity per week, and in Portugal, approximately 90% failed to reach that level (OECD, 2024a).

According to the OECD (2024b), as populations continue to age in the future, governments will face an increase in cases of cancer and, naturally, an increase in its associated costs and, based on population projections, the number of cancer cases is

expected to grow by an average of 44% over the next 30 years in the OECD and 30% in the EU, with Portugal showing a lower growth rate of 17%.

Associated to this expected increase in cancer cases it is expected that healthcare expenditure in cancer treatment per capita increases by 59% for the EU between 2023 and 2050, with Portugal showing similar values (European Commission, 2025). This expectation can be extrapolated to the SNS since, as can be observed in Figure 11, the unitary price of Group 16 pharmaceuticals has been rising indicating that as the consumption of these pharmaceuticals increase, the costs to the SNS rise even further.

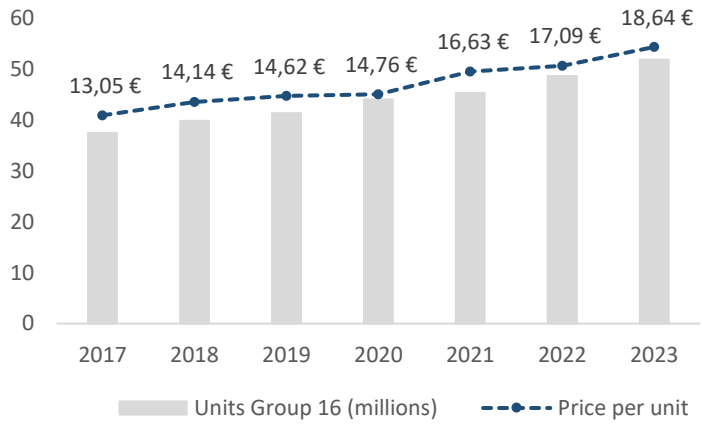

Figure 11: Number of units (millions) of Group 16 and price per unit

Source: INFARMED monthly reports

Therefore, although cancer is a disease that does not discriminate by age, older age groups tend to exhibit higher incidence and prevalence due to biological, behavioural and social factors. Consequently, an increase in longevity and the resulting predominance of older age groups in Portugal is expected to have impact on healthcare and pharmaceuticals expenditure and, particularly, on SNS expenditure on oncological pharmaceuticals. Besides explanatory variables associated to population ageing and prevalence, it is also important to consider other possible explanatory variables.

4. APPLICATION

4.1 Methodology

Kleiber and Zeileis (2008) and Wooldridge (2013) are the main theoretical and methodological references for this application. For further details refer to these works. Appendix 1 presents the basic concepts used in this work.

As already said, our main reference in this part is Kadkhodamanesh et al. (2021). Like these authors, we use multiple linear regressions to explore which variables have impact on the oncological expenditure of each subgroup in Group 16, and, at the same time, we construct predictive models that may be useful to predict future expenditure. Other works previously cited, see Wettermark et al. (2010), Pérez et al. (2018), Vilas et al. (2019) and Sala (2020), were also source of inspiration.

Kadkhodamanesh et al. (2021) applied a multiple linear regression using annual data and a set of different type of variables – GDP per capita, government health expenditure, disability-adjusted life-years, generic pharmaceuticals, as a share of the total value of the pharmaceuticals market, price index, new pharmaceuticals entering the market each year and number of patients in the total population – in order to study their impact on the Iran's pharmaceuticals expenditure. In this work, their model is adapted to the available data.

Due to the limited number of observations, which may reduce the statistical power of the regression models, a direct projection of the dependent variables may be considered to estimate the share of future oncological expenditure, in both total public expenditure and gross domestic product, rather than relying on projections based on the explanatory variables.

For the projections, given the limited number of observations, the approach used in a study conducted by Ovchynikova and Dupliak (2021) is followed, in which certain indicators related to the forestry industry were projected through trend-based extrapolation or average annual growth rates. For the projection of the total public expenditure and gross domestic product, which were not selected as explanatory variables, as will be explained below, ARIMA models were used.

4.2 Variables

Unlike Kadkhodamanesh et al. (2021), whose study has the purpose of explain Iran's total pharmaceutical expenditure based on the drivers of that expenditure, the aim of this work

is much more limited, as it only seeks to understand the relationship and the impact of some variables in the SNS expenditure on oncological pharmaceuticals. Therefore, their model was adapted considering the context of hospital pharmaceuticals, the availability of relevant data and the data that could be collected on a reasonable period of time. The explanatory variables selected to the models result from this process.

In this context, it is important to note that using variables specifically related to oncological patients treated in SNS institutions would have been more appropriate. However, due to time constraints, the analysis relied on data for new cases of cancer diagnosed in mainland SNS institutions, which may be less directly linked to oncological expenditure. In fact, is not possible to determine if all diagnosed cases of cancer were treated in SNS institutions or how many were treated with a specific type of pharmaceutical.

In addition to variables related to demography and disease prevalence, pharmaceutical market, and innovation, other variables related to prevention were included, due to their relevance in the SNS. More specifically, new cases of cancer by age group are linked to demography and disease prevalence, screenings to prevention, orphans and biosimilars to pharmaceuticals market, and new pharmaceuticals to innovation.

The variables defined for each of the three models, for the three subgroups, are listed below (for more details about the variables, see Table 19 of Appendix 2):

- Model 1 – Oncological cytotoxics expenditure (CTE)

  New cases of cancer by age group (NC014, NC1564, NC65);

  Breast cancer screenings (BSC);

  Cervical cancer screenings (CSC);

  Colorectal cancer screenings (CRSC);

  Orphans (CTO) - They are intended for the diagnosis, prevention, or treatment of conditions that affect no more than five in every ten thousand people in the EU and for which there are no satisfactory methods of diagnosis, prevention, or treatment;

  New pharmaceuticals (NP).

- Model 2 – Oncological hormones and anti-hormones expenditure (HE)

  For this model, we selected the same variables as in Model 1, with the exception of CTO; it was not included since there is no consumption of orphans in this subgroup.

- Model 3 – Oncological immunomodulatings expenditure (IME)

For this model, we selected the same variables as in Model 1, with the exception of CTO, which is now replaced with Orphans (IMO), and we added another variable, Biosimilars (IMB) - a biological pharmaceutical developed with the aim of being similar to another already approved biological pharmaceutical. For this model, the variable Biosimilars was added since immunomodulatings is the only subgroup of Group 16 that includes biological pharmaceuticals and, consequently, has biosimilars market.

4.3 Data

After the definition of the variables, the annual data for each variable was collected from various sources and this process posed significant challenges due to the limited availability of annual data related to the healthcare and hospital pharmaceuticals sector.

The annual data on SNS oncological expenditure on cytotoxics, hormones and anti-hormones and immunomodulatings, orphans, biosimilars and new pharmaceuticals were only available for the most recent years and, as a result, data for earlier years had to be formally requested to INFARMED. The data collected corresponds to the period from 2007 to 2023, except new pharmaceuticals that correspond to the period from 2007 to 2024.

Regarding data on new cases of cancer, it had to be formally requested to RON since information was publicly available only for some years. Due to the limited reliability of 2023 data, the collected data corresponds to the period from 2007 to 2022.

The data on different type of screenings were obtained from *Direção-Geral da Saúde* (DGS) reports on *Estatísticas da Saúde* and correspond to the period from 2009 to 2023.

Before fitting the regressions, a preliminary analysis of the variables was performed:

- Between 2007 and 2014, the oncological expenditure of the SNS on each of the Group 16 subgroups, as well as on the Group 16, experienced fluctuations. However, since 2014, this expenditure has grown steadily. In the early years, spending on cytotoxics was the main contributor to the group's total expenditure, but from 2017 onwards, immunomodulatings became the primary contributor. This shift can be explained by the fact that, over the past 16 years, oncological expenditure on immunomodulatings recorded an average annual growth rate of 13.06%, whereas cytotoxics and hormones and anti-hormones grew at average rates of 6.18% and 4.28%, respectively. The evolution of the oncological

expenditure on the three subgroups of Group 16 may be seen in Figure 17, Figure 18 and Figure 19 of the Appendix 2.

- The number of new cases of cancer by age group is higher for the age group above 65 years, however, across all three age groups, the trend between 2009 and 2022 has been inconsistent, showing both periods of increase and decrease; on average, the number of new cases of cancer in people under 15 years and between 15 and 64 years have decreased and have increased in people above 65 years. In 2020, the new cases of cancer were less than in previous years for people aged between 15 and 64 years and people above 65 years, as less diagnostics were made due to COVID-19 pandemic.
- With regard to the three types of cancer screening, the SNS has made a strong commitment to promoting early detection of certain cancers. Breast cancer screenings have recorded the highest number of screenings, with participation rates ranging from 50.5% to 64.7% over the past 15 years, resulting in an average participation rate of 59.79%. This suggests that the increase in the number of screenings has been driven by a rise in the number of women invited. In the case of colorectal cancer screening, the number of screenings has also increased; however, the fact that the participation rate from 2009 to 2017 was higher than that from 2018 to 2023 indicates that the increase in screenings is mainly due to a greater number of people being invited. Cervical cancer screenings show an average participation rate of 76.95%, strongly influenced by the past four years, during which rates exceeded 88%, suggesting that the rise in screenings may be more closely linked to improved participation rates. In 2020, participation rates for all three types of screenings remained within normal ranges; however, there was a decline in the number of screenings due to the COVID-19 pandemic as a consequence of periods of screening suspension.
- The consumption of oncological orphan pharmaceuticals within the cytotoxics subgroup was high between 2007 and 2011, followed by an abrupt decline between 2012 and 2015 and, from 2016 onwards, consumption returned to higher levels. The low values observed between 2012 and 2015 can be attributed to a decrease in the use of one of the components within the cytotoxics subgroup, namely tyrosine kinase inhibitors. In the immunomodulatings subgroup, the consumption of orphan pharmaceuticals showed a significant decline between 2020 and 2023, in contrast to previous years when consumption remained

relatively high and stable. These decreases may be explained by the fact that certain pharmaceuticals ceased to be classified as orphans once the condition they target is no longer considered rare.

- The consumption of oncological biosimilars within the immunomodulatings subgroup began in 2017 and has grown significantly, with an average growth rate of 253.16%. This sharp increase was largely driven by the years 2018 and 2019, which is consistent with the typically higher demand for such pharmaceuticals following their introduction in the market.
- The number of new oncological pharmaceuticals approved for reimbursement has fluctuated over the years; however, in recent years, more pharmaceuticals have been approved compared to earlier periods. This trend may be closely related to the emergence of increasingly rare diseases within the field of oncology, which require innovative treatments, as well as to a possible reduction in the average evaluation time.

Following a preliminary analysis of the data and variables, the logarithm was applied to each variable and the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests were conducted to assess their stationarity before performing the regressions to prevent spurious results and it is important to note that, when the series are stationarized prior to applying the regression model, the model primarily captures short-term relationships. One of the results of these tests (see Table 20 in Appendix 3) indicated that the logarithm of the variables NC1564, CSC, and NP were stationary at level with constant mean, that is, integrated of order zero, I(0). For the remaining variables, first differencing was applied before reapplying the stationarity tests (see Table 21 in Appendix 3). The results showed that the logarithm of HE, IME, NC014, NC65, BSC, CRSC, CTO, and IMO were integrated of order one, I(1). Regarding the logarithm of the variables CTE and IMB, none of the tests indicated that they were I(1), so Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was applied to their first differences. The KPSS test did not reject the null hypothesis of stationarity for the variable IMB but did reject for the variable CTE. However, given the fact that the series of the logarithm of CTE is almost detrended with an approximated constant mean after first differencing, its first difference was used in the analysis (see Figure 20 in Appendix 3).

After an evaluation of the stationarity of the variables, it is important to look at the correlations not only between the dependent variables and the explanatory variables but also between the explanatory variables in the three models:

- Model 1

Table 1 shows a significant correlation between the dependent variable DLog(CTE) and the explanatory variable Log(NP). Despite the insignificance of the correlations between DLog(CTE) and Log(NC1564), DLog(NC65) and DLog(BSC), a positive correlation would be expected, what may be explained by the limited number of observations. Regarding the correlations between the explanatory variables, the correlation of DLog(NC65) with Log(NC1564) and DLog(BSC) is high indicating that multicollinearity problems may arise (Woolridge, 2013).

	DLog(CTE)	DLog(NC014)	Log(NC1564)	DLog(NC65)	DLog(BSC)	Log(CSC)	DLog(CRSC)	DLog(CTO)	Log(NP)
DLog(CTE)	1.000								
DLog(NC014)	0.261	1.000							
Log(NC1564)	-0.226	-0.162	1.000						
DLog(NC65)	-0.151	-0.220	0.753 ***	1.000					
DLog(BSC)	-0.154	-0.139	0.615 **	0.859 ***	1.000				
Log(CSC)	0.369	-0.116	0.108	0.223	0.418	1.000			
DLog(CRSC)	0.185	0.0707	0.625 **	0.639 **	0.521 *	0.463	1.000		
DLog(CTO)	0.438	-0.303	-0.0987	0.0933	-0.0190	0.128	-0.267	1.000	
Log(NP)	0.600 **	-0.268	-0.381	-0.234	-0.105	0.583 **	-0.0937	0.606 **	1.000

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 1: Correlations between variables for Model 1

- Model 2

Regarding Model 2, and as shown in Table 2, there is a significant correlation between the dependent variable DLog(HE) and the explanatory variables Log(CSC) and Log(NP). As in Model 1, the negative correlation between DLog(HE) and the same two variables related to cancer cases is contrary to the expected. Regarding the correlations between explanatory variables, there is a high correlation between the same explanatory variables as in Model 1.

	DLog(HE)	DLog(NC014)	Log(NC1564)	DLog(NC65)	DLog(BSC)	Log(CSC)	DLog(CRSC)	Log(NP)
DLog(HE)	1.000							
DLog(NC014)	0.0136	1.000						
Log(NC1564)	-0.174	-0.162	1.000					
DLog(NC65)	-0.105	-0.220	0.753 ***	1.000				
DLog(BSC)	0.0636	-0.139	0.615 **	0.859 ***	1.000			
Log(CSC)	0.699 ***	-0.116	0.108	0.223	0.418	1.000		
DLog(CRSC)	0.109	0.0707	0.625 **	0.639 **	0.521 *	0.463	1.000	
Log(NP)	0.632 **	-0.268	-0.381	-0.234	-0.105	0.583 **	-0.0937	1.000

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 2: Correlations between variables for Model 2

- Model 3

Table 3 shows that the dependent variable DLog(IME) is significantly correlated with the explanatory variables DLog(IMB) and Log(NP). The negative correlation

between DLog(IME) and all variables related to cancer cases is contrary to the expected and the same happens with the variables DLog(BSC), DLog(IMO) and DLog(IMB). Regarding the correlations between the explanatory variables, there is a high correlation between the same explanatory variables as in Models 1 and 2, and, additionally, the explanatory variable DLog(IMO) is highly correlated with Log(NC1564), DLog(NC65) and DLog(BSC).

	DLog(IME)	DLog(NC014)	Log(NC1564)	DLog(NC65)	DLog(BSC)	Log(CSC)	DLog(CRSC)	DLog(IMO)	Log(NP)	DLog(IMB)
DLog(IME)	1.000									
DLog(NC014)	-0.142	1.000								
Log(NC1564)	-0.237	-0.162	1.000							
DLog(NC65)	-0.0741	-0.220	0.753 ***	1.000						
DLog(BSC)	-0.106	-0.139	0.615 **	0.859 ***	1.000					
Log(CSC)	0.347	-0.116	0.108	0.223	0.418	1.000				
DLog(CRSC)	0.0748	0.0707	0.625 **	0.639 **	0.521 *	0.463	1.000			
DLog(IMO)	-0.184	-0.256	0.766 ***	0.689 ***	0.747 ***	0.339	0.337	1.000		
Log(NP)	0.601 **	-0.268	-0.381	-0.234	-0.105	0.583 **	-0.0937	-0.202	1.000	
DLog(IMB)	0.588 **	-0.462	0.0755	-0.0006	0.0138	0.201	-0.0936	0.105	0.289	1.000

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 3: Correlations between variables for Model 3

4.4 Models and Results

4.4.1 Model 1 - Oncological cytotoxics expenditure

As a first step the full model was estimated using annual data from 2009 to 2022, with the following equation:

$$\Delta\log(CTE_t) = \beta_0 + \beta_{NC014} * \Delta\log(NC014_t) + \beta_{NC1564} * \log(NC1564_t) + \beta_{NC65} * \Delta\log(NC65_t) + \beta_{BSC} * \Delta\log(BSC_t) + \beta_{CSC} * \log(CSC_t) + \beta_{CRSC} * \Delta\log(CRSC_t) + \beta_{CTO} * \Delta\log(CTO_t) + \beta_{NP} * \log(NP_t) + \varepsilon_t, t = 1, 2, \ldots, T \quad (1)$$

The results of the estimation are in Table 4 and show that, although the model explains a substantial percentage (72.65%) of the sample variance in the DLog(CTE), its overall explanatory power is limited when all the explanatory variables are accounted for, as reflected by the small value of the adjusted R^2. Moreover, the F-statistic is not statistically significant suggesting that the model, as a whole, does not provide a reliable fit to the data. These findings may be affected by the limited number of observations and the number of variables and suggest that several of the included explanatory variables may not meaningfully contribute to explain the dependent variable. Additionally, the variables

DLog(NC65) and DLog(BSC) have a VIF higher than 10, something expected due to their high degree of collinearity.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	3.395	8.051	0.422	0.695		
DLog(NC014)	0.209	0.256	0.816	0.460		1.879
Log(NC1564)	-0.301	0.752	-0.400	0.710		3.528
DLog(NC65)	-0.787	1.299	-0.606	0.577		20.401
DLog(BSC)	0.0567	0.203	0.279	0.794		10.392
Log(CSC)	-0.0196	0.102	-0.192	0.857		3.861
DLog(CRSC)	0.145	0.120	1.211	0.293		7.046
DLog(CTO)	0.147	0.129	1.138	0.319		6.189
Log(NP)	0.0110	0.0589	0.187	0.860		8.323
R-squared	0.7265					
Adjusted R-squared	0.1795					
F statistic	1.328					
p value	0.4166					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 4: Regression results of the full Model 1

As a next step, a backward regression based on p-values was applied to remove non-statistically significant variables and mitigate multicollinearity problems. After the estimation, shown in Table 5, the model became statistically significant explaining 61.19% of the sample variance in the dependent variable and the adjusted R^2 improved. Regarding the retained explanatory variables, we only retained explanatory variables with a VIF lower than 10, as it is stated in the literature.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	0.0222	0.0185	1.197	0.262		
DLog(NC65)	-0.868	0.318	-2.731	0.0232	**	1.937
DLog(CRSC)	0.147	0.0493	2.979	0.0155	**	2.0680
DLog(CTO)	0.149	0.0458	3.250	0.0100	***	1.233
R-squared	0.6119					
Adjusted R-squared	0.4825					
F statistic	4.729					
p value	0.03019					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 5: Backward regression results of Model 1

Regarding model adequacy, the standard statistical tests did not reject the null hypothesis of no autocorrelation, homoscedasticity and normality at a significance level of 5% as shown in Table 6. Figure 12 displays a graphical representation of the observed and the fitted values of the oncological cytotoxics expenditure (CTE) corresponding to a Root Mean Squared Error (RMSE) of 5 494.063 thousand euros.

Test	Statistic	p value	dL	dU
Durbin-Watson	1.387		0.715	1.816
Breusch-Godfrey	6.800	0.450		
Breusch-Pagan	3.208	0.361		
Shapiro-Wilk	0.908	0.174		

Table 6: Tests for Model 1

The final model may be written in the following way:

$$\Delta\log\widehat{(CTE_t)} = 0.0222 - 0.868 * \Delta\log(NC65_t) + 0.147 * \Delta\log(CRSC_t) + 0.149 * \Delta\log(CTO_t) \quad (2)$$

From equation (2), some conclusions about each explanatory variable assuming other variables remain constant, may be drawn:

- A 1 percentage point increase in the growth rate of new cases of cancer diagnosed in individuals aged over 65 years is associated with an approximated expected 0.868 percentage points decrease in the growth rate of oncological cytotoxics expenditure;
- A 1 percentage point increase in the growth rate of colorectal cancer screenings is associated with an approximated expected 0.147 percentage points increase in the growth rate of oncological cytotoxics expenditure;
- A 1 percentage point increase in the growth rate of consumption of oncological cytotoxics orphans pharmaceuticals is associated with an approximated expected 0.149 percentage points increase in the growth rate of oncological cytotoxics expenditure.

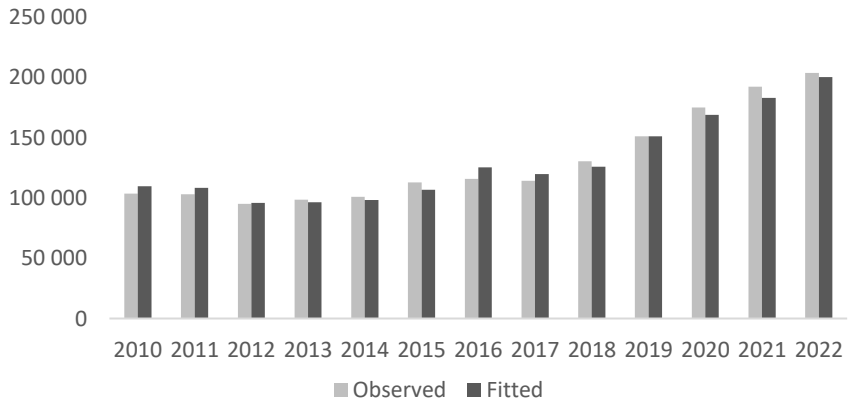

Figure 12: Observed vs Fitted values of CTE (thousand €) of Model 1

4.4.1.1 Model 1 - Oncological cytotoxics expenditure (in level)

Given the high number of variables compared to the number of observations and the reduced meaning of the R^2, this subsection adopts a less restrictive approach concerning the risk of spurious regression: the regression was estimated using the variables in level, with the objective of comparing the results to those obtained from the regression based on the stationarized series and study the long run relationships between the variables.

The results of the estimation for Model 1 are in Table 7 and show that the full model is statistically significant, however, the variable CSC has a VIF value greater than 10.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	221 700	26 670	8.312	0.000412	***	
NC014	53.48	18.56	2.882	0.0345	**	2.35
NC1564	-8.069	0.565	-14.278	3.04e-05	***	3.72
NC65	3.169	0.306	10.363	0.000144	***	6.88
BSC	-0.0319	0.0212	-1.505	0.193		8.41
CSC	0.0102	0.0254	0.404	0.703		13.73
CRSC	0.373	0.0161	23.197	2.77e-06	***	5.04
CTO	0.0368	0.00471	7.813	0.000551	***	1.97
NP	-315.0	139.5	-2.258	0.0735	*	5.60

R-squared	0.9991
Adjusted R-squared	0.9977
F statistic	703.2
p value	3.386e-07

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 7: Regression results of the full Model 1 in level

As a next step, the variable CSC was removed and a new regression was fitted. After the estimation, shown in Table 8, the model remained statistically significant explaining around 99.9% of the sample variance in the CTE and all the explanatory variables had a VIF value lower than 10, as stated in the literature.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	216 300	21 370	10.120	5.41e-05	***	
NC014	55.46	16.60	3.341	0.0156	**	2.19
NC1564	-7.998	0.498	-16.050	3.72e-06	***	3.36
NC65	3.200	0.275	11.644	2.42e-05	***	6.46
BSC	-0.0276	0.0170	-1.622	0.156		6.31
CRSC	0.376	0.0139	26.983	1.71e-07	***	4.38
CTO	0.0377	0.00389	9.698	6.90e-05	***	1.55
NP	-288.2	113.8	-2.533	0.0445	**	4.33

R-squared	0.9991
Adjusted R-squared	0.998
F statistic	933.9
p value	1.111e-08

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 8: Regression results of the Model 1 in level without the variable CSC

The standard statistical tests did not reject the null hypothesis for no autocorrelation – the value of the Durbin-Watson statistic is between the lower and upper values - homoscedasticity and normality at a significance level of 5% and the final model fits quite well to the observed data with the difference between the observed and the fitted values ranging from 33.249 and 2 269.719 thousand euros, corresponding to a Root Mean Squared Error (RMSE) of 1 076.240 thousand euros.

The final model may be written in the following way:

$$\widehat{CTE_t} = 216\,300 + 55.46 * NC014_t - 7.998 * NC1564_t + 3.200 * NC65_t - 0.0276 * BSC_t + 0.376 * CRSC_t + 0.0377 * CTO_t - 288.2 * NP_t \quad (3)$$

4.4.2 Model 2 - Oncological hormones and anti-hormones expenditure

In this model the approach was similar to the previous model. The full model was estimated using annual data from 2009 to 2022, with the following equation:

$$\Delta\log(HE_t) = \beta_0 + \beta_{NC014} * \Delta\log(NC014_t) + \beta_{NC1564} * \log(NC1564_t) + \beta_{NC65} * \Delta\log(NC65_t) + \beta_{BSC} * \Delta\log(BSC_t) + \beta_{CSC} * \log(CSC_t) + \beta_{CRSC} * \Delta\log(CRSC_t) + \beta_{NP} * \log(NP_t) + \varepsilon_t, t = 1, 2, \ldots, T \quad (4)$$

The results of the estimation, see Table 9, show that the model, as a whole, does not provide a reliable fit to the data. Regarding multicollinearity problems, all explanatory variables have a VIF lower than 10, contrary to what happened in Model 1.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	-4.369	13.271	-0.329	0.755		
DLog(NC014)	0.220	0.405	0.544	0.610		1.501
Log(NC1564)	0.151	1.242	0.122	0.908		3.075
DLog(NC65)	0.128	1.460	0.088	0.933		8.243
DLog(BSC)	-0.0678	0.279	-0.243	0.817		6.240
Log(CSC)	0.231	0.180	1.282	0.256		3.860
DLog(CRSC)	-0.0536	0.141	-0.379	0.720		3.441
Log(NP)	0.0362	0.0602	0.601	0.574		2.782
R-squared	0.6163					
Adjusted R-squared	0.07924					
F statistic	1.148					
p value	0.456					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 9: Regression results of the full Model 2

As a next step, a backward regression was fitted whose results are shown in Table 10, indicating that the explanatory variable Log(CSC) is statistically significant explaining 48.92% of the sample variance in the DLog(HE), which is not a desirable value, and the value of the adjusted R^2 improved.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code
(Intercept)	-2.722	0.857	-3.177	0.00881	***
Log(CSC)	0.232	0.0714	3.246	0.00779	***
R-squared	0.4892				
Adjusted R-squared	0.4428				
F statistic	10.54				
p value	0.007794				

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 10: Backward regression results of Model 2

As in Model 1, the standard statistical tests did not reject the null hypothesis of no autocorrelation, homoscedasticity and normality at a significance level of 5% as shown in Table 11 and, in Figure 13, it is the graphical representation of the observed and the fitted values of the oncological hormones and anti-hormones expenditure (HE) corresponding to a RMSE of 2 829.364 thousand euros.

Test	Statistic	p value	dL	dU
Durbin-Watson	1.308		1.010	1.340
Breusch-Godfrey	7.434	0.190		
Breusch-Pagan	1.877	0.171		
Shapiro-Wilk	0.919	0.245		

Table 11: Tests for Model 2

The final model may be written in the following way:

$$\Delta\widehat{\log(HE_t)} = -2.722 + 0.232 * \log(CSC_t) \quad (5)$$

From equation (5) we can derive the following conclusion:

- A 1% higher level of cervical cancer screenings is associated to an approximated expected 0.232 percentage points increase in the growth rate of oncological hormones and anti-hormones expenditure.

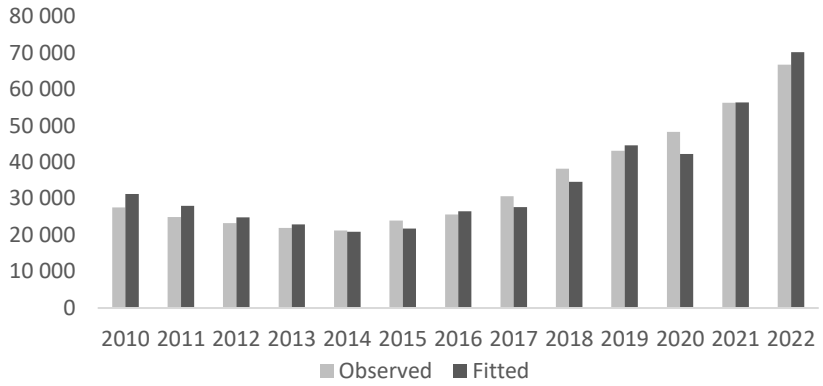

Figure 13: Observed vs Fitted values of HE (thousand €) for Model 2

4.4.2.1 Model 2 - Oncological hormones and anti-hormones expenditure (in level)

As in Model 1, we estimated the regression with the variables in level. The results of the estimation, see Table 12, show that the model is statistically significant.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	163 600	55 560	2.944	0.0258	**	
NC014	-45.44	39.52	-1.150	0.294		2.17
NC1564	-2.671	1.251	-2.136	0.0766	*	3.71
NC65	-0.496	0.630	-0.787	0.461		5.94
BSC	-0.00778	0.0442	-0.176	0.866		7.44
CSC	0.0778	0.0500	1.557	0.170		10.85
CRSC	0.139	0.0355	3.917	0.00783	***	4.99
NP	-135.7	308.9	-0.439	0.676		5.59
R-squared	0.9647					
Adjusted R-squared	0.9234					
F statistic	23.39					
p value	0.0005961					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 12: Regression results of the full Model 2 in level

As in the first model, the variable CSC was removed and then a regression was fitted whose results are shown in Table 13, indicating that the model remained statistically significant explaining around 95% of the sample variance in the HE.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	127 600	55 420	2.302	0.0548	*	
NC014	-35.49	42.78	-0.830	0.434		2.12
NC1564	-1.944	1.273	-1.527	0.171		3.19
NC65	-0.402	0.688	-0.584	0.577		5.89
BSC	0.0206	0.0442	0.466	0.655		6.18
CRSC	0.164	0.0347	4.735	0.00212	***	3.96
NP	132.6	281.2	0.471	0.652		3.85
R-squared	0.9504					
Adjusted R-squared	0.9078					
F statistic	22.34					
p value	0.0003115					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 13: Regression results of the Model 2 in level without the variable CSC

The standard statistical tests did not reject the null hypothesis for no autocorrelation, homoscedasticity and normality at a significance level of 5% and the final model fits quite well to the observed data with the difference between the observed and the fitted values range from 7.985 and 7.414.577 thousand euros, corresponding to a RMSE of 3 039.624 thousand euros.

The final model may be written in the following way:

$$\widehat{HE}_t = 127\,600 - 35.49 * NC014_t - 1.944 * NC1564_t - 0.402 * NC65_t + 0.0206 * BSC_t + 0.164 * CRSC_t + 132.6 * NP_t \quad (6)$$

4.4.3 Model 3 - Oncological immunomodulatings expenditure

Similarly to the previous two models, the full model was estimated using annual data from 2009 to 2022, with the following equation:

$$\Delta\log(IME_t) = \beta_0 + \beta_{NC014} * \Delta\log(NC014_t) + \beta_{NC1564} * \log(NC1564_t) + \beta_{NC65} * \Delta\log(NC65_t) + \beta_{BSC} * \Delta\log(BSC_t) + \beta_{CSC} * \log(CSC_t) + \beta_{CRSC} * \Delta\log(CRSC_t) + \beta_{IMO} * \Delta\log(IMO_t) + \beta_{NP} * \log(NP_t) + \beta_{IMB} * \Delta\log(IMB_t) + \varepsilon_t, t = 1, 2, \ldots, T \quad (7)$$

The results of the estimation are shown in Table 14 and indicate that the model, as a whole, is not statistically significant as happened in Models 1 and 2.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	22.886	18.761	1.220	0.310		
DLog(NC014)	0.305	0.349	0.874	0.447		1.734
Log(NC1564)	-2.017	1.658	-1.217	0.311		8.538
DLog(NC65)	0.695	1.179	0.590	0.597		8.376
DLog(BSC)	-0.152	0.231	-0.659	0.557		6.671
Log(CSC)	-0.150	0.194	-0.771	0.497		6.957
DLog(CRSC)	0.174	0.163	1.070	0.363		7.084
DLog(IMO)	0.115	0.142	0.812	0.476		8.654
Log(NP)	0.0748	0.0529	1.415	0.252		3.341
DLog(IMB)	0.0365	0.0166	2.206	0.115		1.511
R-squared	0.7871					
Adjusted R-squared	0.1485					
F statistic	1.233					
p value	0.481					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 14: Regression results of the full Model 3

As a next step, similarly to what was done in the previous two models, a backward regression was fitted. After the estimation, shown in Table 15, the model became statistically significant, explaining 58.82% of the sample variance in the DLog(IME).

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	13.868	6.227	2.227	0.0529	*	
Log(NC1564)	-1.314	0.593	-2.218	0.0537	*	1.692
DLog(CRSC)	0.120	0.0639	1.883	0.0923	*	1.698
DLog(IMB)	0.0346	0.0110	3.136	0.0120	**	1.0398
R-squared	0.5882					
Adjusted R-squared	0.451					
F statistic	4.286					
p value	0.03882					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 15: Backward regression results of Model 3

The standard statistical tests did not reject the null hypothesis of no autocorrelation, homoscedasticity and normality at a significance level of 5% as shown in Table 16; however, if considering a significance level of 10%, the standard statistical test for homoscedasticity did not reject the null hypothesis. From Figure 14 we can see that the difference between the observed and the fitted values of the oncological immunomodulatings expenditure (IME) corresponds to a RMSE of 6 283.751 thousand euros.

Test	Statistic	p value	dL	dU
Durbin-Watson	1.157		0.715	1.816
Breusch-Godfrey	11.767	0.109		
Breusch-Pagan	6.674	0.0831		
Shapiro-Wilk	0.57	0.713		

Table 16: Tests for Model 3

The final model equation may be written:

$$\Delta\log\widehat{(IME_t)} = 13.868 - 1.314 * \log(NC1564_t) + 0.120 * \Delta\log(CRSC_t) + 0.0346 * \Delta\log(IMB_t) \quad (8)$$

Similarly to what was done with the other two models, from equation (8) we can conclude about each significant explanatory variable:

- A 1% higher level of new cases of cancer in individuals aged between 15 and 64 years is associated to an approximated expected decrease of 1.314 percentage points in the growth rate of oncological immunomodulatings expenditure;
- A 1 percentage point increase in the growth rate of colorectal cancer screenings is associated with an approximated expected 0.120 percentage point increase in the growth rate of oncological immunomodulatings expenditure;
- A 1 percentage point increase in the growth rate of consumption of oncological immunomodulatings biosimilars pharmaceuticals is associated with an approximated expected 0.0346 percentage point increase in the growth rate of oncological immunomodulatings expenditure.

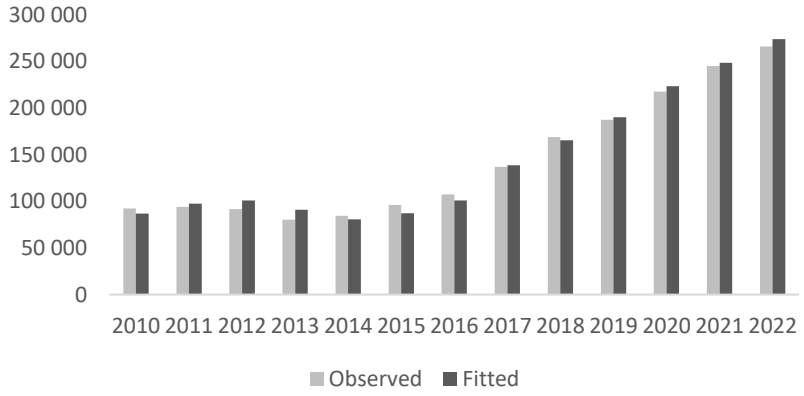

Figure 14: Observed vs Fitted values of IME (thousand €) for Model 3

4.4.3.1 Model 3 - Oncological immunomodulatings expenditure (in level)

Likewise to what was done in the previous two models, the results of the estimation are shown in Table 17 and indicate that the model was statistically significant; however, since all the explanatory variables with the exception of NC014 have a VIF value greater than 10, with particular attention to IMB and CRSC with notably high VIFs, we can conclude that the marginal power of explanation of each of these variables, when the other ones are included, might be quite insignificant.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	428 800	543 100	0.790	0.474		
NC014	-133.7	227.6	-0.587	0.589		4.59
NC1564	-17.07	12.97	-1.316	0.259		25.39
NC65	7.618	4.578	1.664	0.171		19.99
BSC	-0.366	0.215	-1.705	0.163		11.17
CSC	0.286	0.300	0.953	0.394		24.86
CRSC	0.505	0.852	0.593	0.585		182.95
IMO	0.357	0.423	0.845	0.446		48.25
NP	-1 238	1 918	-0.646	0.554		13.71
IMB	1.148	2.987	0.384	0.720		310.28
R-squared	0.9828					
Adjusted R-squared	0.9441					
F statistic	25.39					
p value	0.003518					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 17: Regression results of the full Model 3 in level

As a next step, before fitting a new regression, between the two variables with larger VIFs, the one with a lower Pearson correlation with the dependent variable was removed. This process was repeated until all the explanatory variables had a VIF lower than 10 and resulted in the removing of the variables CRSC, IMO and CSC. After the estimation, shown in Table 18, the model remained statistically significant, explaining around 97.3% of the sample variance in the IME.

Coefficient	Estimate	Std. Error	t value	p value	Sign. Code	VIF
(Intercept)	55 490	210 100	0.264	0.799		
NC014	-70.36	148.6	-0.474	0.650		2.15
NC1564	-3.242	5.039	-0.643	0.540		4.21
NC65	4.897	2.369	2.068	0.0775	*	5.88
BSC	-0.179	0.156	-1.151	0.288		6.45
NP	986.8	956.3	1-032	0.336		3.75
IMB	1.952	0.345	5.663	0.000764	***	4.54
R-squared	0.9726					
Adjusted R-squared	0.9491					
F statistic	41.42					
p value	4.032e-05					

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 18: Regression results of the Model 2 in level without the variables CRSC, IMO and CSC

The standard statistical tests did not reject the null hypothesis for no autocorrelation, homoscedasticity and normality at a significance level of 5% and the final model fits quite well to the observed data with the difference between the observed and the fitted values ranged from 886.472 and 23 187.600 thousand euros, corresponding to a RMSE of 10 478.879 thousand euros.

The final model equation may be written:

$$\widehat{IME_t} = 55\,490 - 70.36 * NC014_t - 3.242 * NC1564_t + 4.897 * NC65_t - 0.179 * BSC_t + 986.8 * NP_t + 1.952 * IMB_t \quad (9)$$

4.4.4 Main conclusions on the three models

Analysing the results, it is evident that all three models exhibit limited explanatory power - particularly the second model - and some variables in the models contribute to an acceleration of the oncological expenditure while others to a slowdown.

Regarding the variables related to demography and prevalence of cancer, new cases of cancer among individuals over 65 years old and new cases of cancer among individuals aged between 15 and 64 contribute to a slowdown of the oncological expenditure on cytotoxics and immunomodulatings, respectively. The limited number of observations and the fact that new cases of cancer by age group do not follow a clear trend may explain this results. Additionally, as already said, it is not possible to determine if all diagnosed cases of cancer were treated in SNS institutions or how many were treated with a specific type of pharmaceutical. Moreover, since some time series were differenced to achieve stationarity, the estimated relationships mainly capture short-term dynamics rather than long-term effects, which may have influenced the results.

Concerning the variables related to screenings, results were as expected. Colorectal cancer screenings contribute to the acceleration of the oncological expenditure on cytotoxics and immunomodulatings and cervical cancer screenings to the acceleration of the oncological expenditure on hormones and anti-hormones.

Regarding variables related to the pharmaceutical market, the contribution of orphan cytotoxics pharmaceuticals consumption to the acceleration of the oncological expenditure on cytotoxics was as expected. However, the contribution of biosimilar immunomodulatings pharmaceuticals consumption to the acceleration of the oncological

expenditure on immunomodulatings was contrary to expectations. This may be explained either by the fact that biosimilar consumption is directly linked to the prices of pharmaceuticals, or by the possibility that their current level of consumption is still insufficient to significantly reduce expenditure.

With respect to the additional analysis conducted using the variables in level - specifically focusing on those that significantly influenced the three categories of oncological expenditure - the findings revealed both expected and unexpected results, similar to the analysis based on stationarized series.

When focusing on new cancer cases, the results indicate that cases among individuals aged 15 to 64 years were associated with a decrease in oncological expenditure on cytotoxics. However, the findings for new cases among individuals under 15 years of age and those aged 65 and over were more aligned with the existing literature, showing a contribution to the increase of the oncological expenditure on cytotoxics and immunomodulatings. These results may suggest that new cases of cancer among individuals under 15 years of age have a long-term increasing effect on oncological expenditure despite a no short-term impact. A similar long-term effect is observed for cases among individuals aged 65 and over, despite their decreasing short-term impact.

4.5 Projection of the Oncological Expenditure

In this section, the oncological expenditure of the SNS for each subgroup was projected until 2030, to assess its evolution in this 5-year period, in absolute terms and also as a percentage of the total public expenditure (TPE) and of the GDP. The initial idea was to use the previous models to project the three oncological expenditures through projections of the regressors; however we decided to act in a different way and the three oncological expenditures were projected directly. Still, for the purpose of comparison, projections based on Models 1, 2 and 3 are in Appendix 5.

For the reasons already described, we followed Ovchynikova and Dupliak (2021). Accordingly, for the direct projection of the three expenditures, a trend line was fitted for the period from 2007 to 2022. Among those with a R^2 value greater than 0.90, the one with the best predicting performance for the year 2023 was selected. The dependent variables were projected as follows (for the results of the projections, see Table 22 of Appendix 4):

- For the variable CTE, a second-degree and third-degree polynomial trending lines were fitted with R^2 values of 0.951 and 0.975, respectively. The second-degree polynomial performed better when predicting 2023 value;
- For the variable HE, a second-degree and third-degree polynomial trending lines were fitted with R^2 values of 0.992 and 0.993, respectively. The second-degree polynomial performed better when predicting 2023 value;
- For the variable IME, an exponential, second-degree and third-degree polynomial trending lines were fitted with R^2 values of 0.937, 0.960 and 0.972, respectively. The third-degree polynomial performed better when predicting 2023 value.

For the projection of TPE and GDP (non-explanatory variables, as said before), data from 1960 to 2024 was collected from PORDATA and ARIMA modelling was applied. The best model was identified using "auto.arima()" function from R software. The obtained model for TPE was an ARIMA (2,2,0) and for GDP was an ARIMA (0,2,2). The results are in Table 23 of Appendix 4.

Looking at the projections, see Figure 15, it is expected that the SNS oncological expenditure on the three subgroups will grow at an average annual rate of 12.72% until 2030, with particular attention on hormones and anti-hormones and immunomodulatings, which will grow at average annual rates of 12.15% and 14.97%, respectively.

Additionally, based on the projections of TPE - which is projected to grow at an average annual rate of 4.79% until 2030 - and of GDP - which is projected to grow at an average annual rate of 4.48% until 2030 -, the percentages of SNS oncological expenditure in terms of the TPE and the GDP are expected to increase by approximately 0.28 and 0.13 percentage points, respectively, in the 5-year period, see Figure 16.

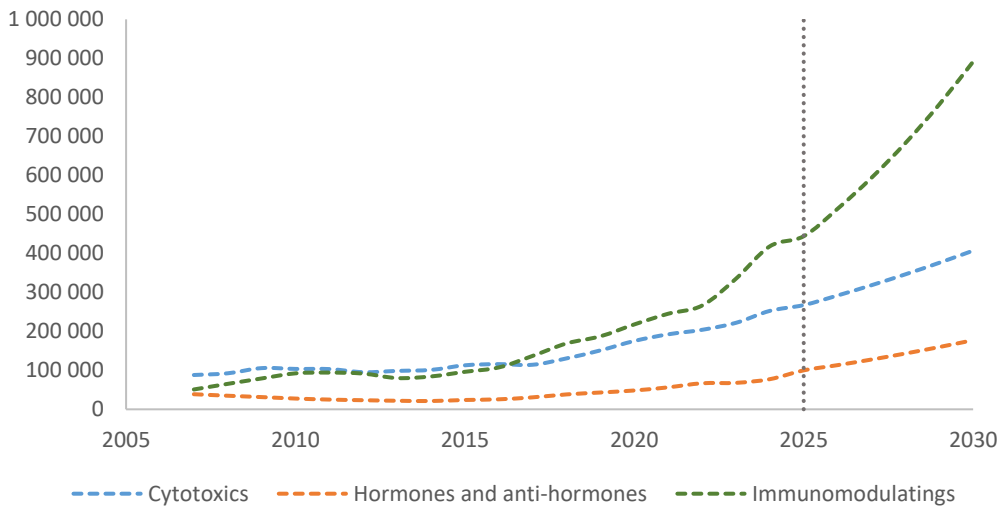

Figure 15: Projection of oncological expenditure (thousand €) of each subgroup (2025-2030)

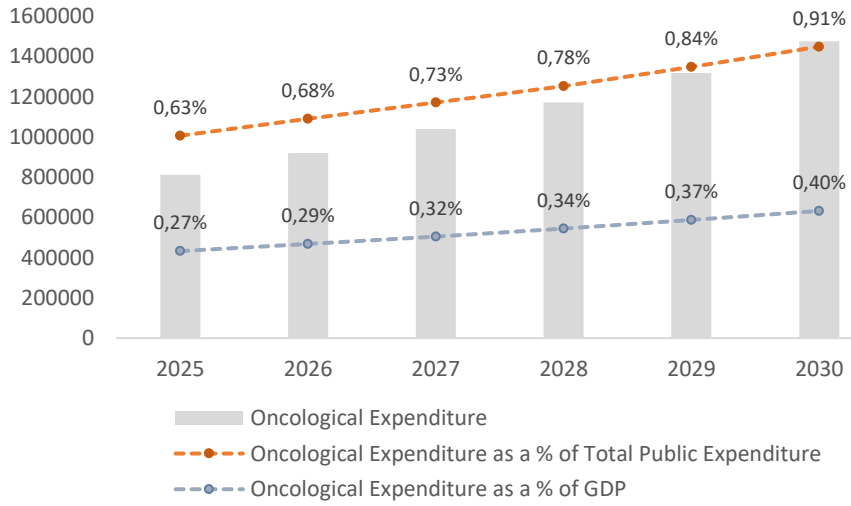

Figure 16: Projection of oncological expenditure (thousand €) and % of TPE and GDP (2025-3030)

5. CONCLUSIONS

This study aimed to develop models to explain the oncological pharmaceuticals expenditure of the SNS using a multiple linear regression approach, focusing on each subgroup of pharmaceuticals included in Group 16. The models were constructed using a set of explanatory variables that reflect demography and prevalence of cancer, pharmaceutical market, prevention and innovation. This approach aligns with previous studies in the field of health economics and pharmaceutical policy, which have increasingly relied on statistical modelling to better understand and forecast healthcare spending patterns.

Once the models were implemented, the goal was to understand the impact of the variables related to demography and prevalence of cancer on each oncological expenditure. Additionally, a prediction of the future oncological expenditure was applied to assess its share in both total public expenditure and gross domestic product. However, due to the limited capacity of explanation of the models, a direct projection of the expenditure was applied instead of using the models.

The findings concerning the impact of new cases of cancer on oncological expenditure revealed a certain degree of inconsistency. Specifically, the emergence of new cases does not appear to contribute to an acceleration of the oncological expenditure, meaning it does not significantly influence the rate of expenditure increase. Nonetheless, the analysis indicates a long run increasing relationship between new cases of cancer - particularly associated with new cancer diagnoses among individuals aged 65 years and older - and oncological expenditure. This pattern suggests that demographic ageing may exert substantial pressure on healthcare costs related to oncology. It is important to note, however, that these inconsistent results may be partially explained by data limitations, including the selection of variables that may capture the phenomenon only indirectly, the restricted number of available observations, and the use of stationary series that capture short run relationships.

Despite some contradictory results, the study showed some relevant findings. Variables related to prevention, such as colorectal and cervical cancer screenings, contribute to an acceleration of the oncological expenditure while the consumption of oncological cytotoxics orphans pharmaceuticals contribute to an acceleration of the oncological expenditure on cytotoxics.

These results may suggest that, as cancer prevention becomes a more widespread reality in our country, oncological expenditure is expected to increase at a faster rate, potentially due to a rise in cancer incidence. In particular, the contribution of colorectal cancer screenings to this acceleration may indicate that, as the population continues to age, expenditure will likely grow more rapidly, given that these screenings are targeted at individuals aged between 50 and 74.

Moreover, it is expected that the oncological expenditure will grow until 2030 at an annual average rate of 12.72%, which means that it will have increasing weights regarding its representation in the total public expenditure and in the gross domestic product, suggesting a potential unsustainability situation in the medium term or in the long term. This trend is specially concerning given the demographic transition happening in Portugal and many EU countries, where ageing populations are likely to place additional pressure on healthcare systems. This reinforces the need for sustainable planning, as demographic ageing may amplify long-term financial pressure on the SNS.

Despite the valuable insights provided by the models, the study was not without limitations. A major constraint was the limited number of observations available for analysis, since some of the explanatory variables were only reported on an annual basis, and the data collection proved challenging for some variables, as their availability was limited in public databases, requiring formal requests to institutions. This led to delays and restricted the scope of the analysis.

For future research, beyond the screening or the new cases of cancer variables already employed, the analysis may be further refined through the inclusion of indicators relevant to hospital production, such as the number of oncological patients treated, number of medical consultations, or number of day hospital sessions - all of which are directly associated with the prescription and consumption of oncological pharmaceuticals. The collection of these data poses significant challenges, as they are dispersed across multiple hospital units specialized in oncological care, resulting in a more complex and time-intensive data collection process. As such, this constitutes a complementary development that necessitates a longer timeframe for implementation. It would also be valuable to conduct an analysis based on price and quantity, that is, to study which variables may impact these two factors and their impact on the oncological expenditure.

Additionally, the study could be adapted to explore the possibility of increasing the number of observations. Potential solutions include redefining the type of variables to

enable the use of monthly or quarterly data or conducting a panel data regression based on hospital-level information.

Despite the limitations of this study and the possible future research, it provided a first attempt to explain the evolution of SNS expenditure on oncological pharmaceuticals in terms of some explanatory variables and predict the future expenditure. This may support policymakers on more informed decision-making processes regarding resource allocation. Moreover, this study highlights that, although Portugal is included in studies conducted by the European Commission (EC) on overall healthcare expenditure, there is a need to focus the analysis not only on overall SNS expenditure but also on specific expenditure aggregates of the SNS.

REFERENCES

ACSS. (n.d.). *Serviço Nacional de Saúde - Portal da Transparência*. Retrieved February 14, 2025, from *Conta do Serviço Nacional de Saúde*: conta do serviço nacional de saúde — transparência

Amaglobeli, D., & Shi, W. (2016). How to assess fiscal implications of demographic shifts: A granular approach. *IMF How to Notes 2016, 2*. doi:https://doi.org/10.5089/9781475536072.061

Azevedo, A., & Baptista, M. I. (2014). A mortalidade em Portugal, 1950-2011. In *Dinâmicas demográficas e envelhecimento da população portuguesa (1950-2011): evolução e perspetivas* (pp. 227-402). Lisboa: Fundação Francisco Manuel dos Santos.

Barros, P., & Santos, C. (2024). *Acesso a Cuidados de Saúde, 2023.* Fundação "la Caixa", Banco BPI e Nova SBE.
relatorio_acesso_cuidados_saude_dez_2024.pdf

Barros, P., & Santos, C. (2024). *Relatório de Envelhecimento.* Fundação "la Caixa", Banco BPI e Nova SBE.
relatorio_de_envelhecimento_12_jul_2024.pdf

Belloni, A., Morgan, D., & Paris, V. (2016). OECD Health Working Papers. *Pharmaceutical Expenditure and Policies: Past Trends and Future*(87). doi:http://dx.doi.org/10.1787/5jm0q1f4cdq7-en

Bindel, L., & Seifert, R. (2025). Long-term forecast for antibacterial drug consumption in Germany using ARIMA models. *Naunyn-Schmiedeberg's Archives of Pharmacology*. doi:https://doi.org/10.1007/s00210-024-03721-4

Conselho das Finanças Públicas. (2024). *Evolução do Desempenho do Serviço Nacional de Saúde em 2023.*
evolução do desempenho do serviço nacional de saúde em 2023

DGS. (n.d.). *Estatísticas da Saúde*. Retrieved April 10, 2025, from estatísticas da saúde

Di Matteo, L. (2005). The Macro Determinants of Health Expenditure in the United States and Canada: Assessing the Impact of Income, Age Distribution and Time. *Health Policy, 71*(1), 23-42. doi:https://doi.org/10.1016/j.healthpol.2004.05.007

Di Matteo, L., & Di Matteo, R. (1998). Evidence on the Determinants of Canadian Provincial Government Health Expenditures. *Journal of Health Economics, 17*(2), 211-228. doi:https://doi.org/10.1016/S0167-6296(97)00020-9

European Commission. (n.d.). Retrieved February 12, 2025, from EUROSTAT: statistics | eurostat

European Commission. (2024). *2024 Ageing Report. Economic and Budgetary Projections for the EU Member States (2022-2070).* Brussels.

European Commission. (2025). *EU Country Cancer Profiles Synthesis Report 2025.* Paris: OECD Publishing. doi:https://doi.org/10.1787/20ef03e1-en

Fries, J. F. (1989). The Compression of Morbidity: Near or Far? *The Milbank Quarterly, 67*(2), 208-232. doi:https://doi.org/10.2307/3350138

Fundação Francisco Manuel dos Santos. (n.d.). *PORDATA*. Retrieved June 11, 2025, from contas públicas | pordata

Fundação Francisco Manuel dos Santos. (n.d.). *PORDATA*. Retrieved June 18, 2025, from produto interno bruto (pib) | pordata

García-Goñi, M. (2022). Rationalizing Pharmaceutical Spending. *IFM Working Papers*.

Gomes, C., & Moreira, M. (2014). Evolução da população portuguesa. In *Dinâmicas Demográficas e Envelhecimento da População Portuguesa: Evoluções e Perspetivas (1950-2011)* (pp. 29-109). Lisboa: Fundação Francisco Manuel dos Santos.

Gruenberg, E. M. (1977). The Failures of Success. *The Milbank Memorial Fund Quarterly. Health and Society, 55*(1), 3-24. doi:https://doi.org/10.2307/3349592

INE. (n.d.). Retrieved January 27, 2025, from Principais Agregados das Administrações Públicas - Estatísticas da Despesa Pública: portal do ine

INE. (n.d.). Retrieved January 27, 2025, from *Conta Satélite da Saúde*: portal do ine

INE. (2024). *Estatísticas Demográficas 2023. Retrieved from* portal do ine

INFARMED. (n.d.). *Análise de consumo de medicamentos em meio ambulatório.* Retrieved January 27, 2025, from análise de consumo de medicamentos em meio ambulatório - ambulatório - infarmed, i.p.

INFARMED. (n.d.). *Análise de consumo de medicamentos em meio hospitalar.* Retrieved January 27, 2025, from análise de consumo de medicamentos em meio hospitalar - hospitalar - infarmed, i.p.

Kadkhodamanesh, A., Varahrami, V., Zarei, L., Peiravian, F., Hadidi, M., & Yousefi, N. (2021). Investigation the determinants of pharmaceutical expenditure share of GDP in Iran and selected OECD countries. *Journal of Pharmaceutical Policy and Practice, 14*(82), 1-10. doi:https://doi.org/10.1186/s40545-021-00371-2

Klazoglou, P., & Dritsakis, N. (2018). Modeling and Forecasting of US Health Expenditures Using ARIMA Models. In ICOAE, *Advances in Panel Data Analysis in Applied Economic Research* (pp. 457-472). doi:https://doi.org/10.1007/978-3-319-70055-7_36

Kleiber, C., & Zeileis, A. (2008). *Applied Econometrics with R.* Springer New York, NY. doi:https://doi.org/10.1007/978-0-387-77318-6

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models* (5 ed.). McGraw-Hill. Retrieved from (pdf) applied linear statistical models

Manzano, A., Svedman, C., Hofmarcher, T., & Wilking, N. (2025). *Comparator Report on Cancer in Europe 2025 - Disease Burden, Costs and Access to Medicines and Molecular Diagnostics.* Lund, Sweden: The Swedish Institute for Health Economics.

Moreira, M. J., & Gomes, C. (2014). Dinâmicas demográficas do envelhecimento:. In *Dinâmicas Demográficas e Envelhecimento da População Portuguesa (1950-2011): Evoluções e Perspetivas* (pp. 111-168). Lisboa: Fundação Francisco Manuel dos Santos.

Moreira, M. J., & Henriques, F. C. (2014). Mudanças demográficas e estado de saúde em Portugal entre 1970 e 2013. In T. F. Rodrigues, & M. O. Martins, *Envelhecimento e Saúde. Prioridades Políticas num Portugal em Mudança* (1º ed., pp. 71-101). Instituto Hidrográfico.

Odnoletkova, I., Chalon, P., Devriese, S., & Cleemput, I. (2025). Projections of public spending on pharmaceuticals: A review of methods. *Pharmaco Economics*, 375-388. doi:https://doi.org/10.1007/s40273-024-01465-w

OECD. (2023). *Health at a Glance 2023: OECD indicators.* Paris: OECD Publishing. doi:https://doi.org/10.1787/7a7afb35-en

OECD. (2024). *Health at a Glance: Europe 2024.* Paris: OECD Publishing. doi:https://doi.org/10.1787/b3704e14-en

OECD. (2024). Tackling the Impact of Cancer on Health, the Economy and Society. *OECD Health Policy Studies*. doi: https://doi.org/10.1787/85e7c3ba-en

OECD. (2025). *Perfil sobre cancro por país: Portugal 2025.* Paris: OECD Publishing. doi: https://doi.org/10.1787/ffdcd7a9-pt

Omran, A. R. (2005). The Epidemiologic Transition: A Theory of the Epidemiology of Population Change. *The Milbank Quarterly, 83*(4), 731-757. doi:https://doi.org/10.1111/j.1468-0009.2005.00398.x

Ovchynikova, O., & Dupliak, O. (2021). Analysis, forecasting, and management of indicators of the forest industry of the region. *SHS Web of Conferences, 107*(10005). doi:https://doi.org/10.1051/shsconf/202110710005

Pérez, C., Fernández, C., Méndez, V., Méndez, P., & Fernández, A. (2018). Evolution of GDP and its impact on the pharmaceutical sector of Ecuador (2007-2016). *Journal of International Studies, 11*(1), 288-296. doi:http://doi.org/10.14254/2071-8330.2018/11-1/22

Registo Oncológico Nacional. (2024). *Registo Oncológico Nacional de Todos os Tumores na População Residente em Portugal, em 2021.* Porto: Instituto Português de Oncologia do Porto Francisco Gentil.

Rodrigues, T. (2018). *Envelhecimento e Políticas de Saúde.* Lisboa: Fundação Francisco Manuel dos Santos.

Sala, L. (2020). The relationship between population ageing and medical expenditures in Romania. *Facta Universitatis: Economics and Organization, 17*(2), 157-172. doi:https://doi.org/10.22190/FUEO191223012S

Scott, A. J. (2023). The Economics of Longevity - An introduction. *The Journal of Economics of Ageing, 24*. doi:https://doi.org/10.1016/j.jeoa.2022.100439

Verbugge, L. M. (1984). Longer Life but Worsening Health? Trends in Health and Mortality of Middle-Aged and Older Persons. *The Milbank Memorial Fund Quarterly. Health and Society, 62*(3), 475-519. doi:https://doi.org/10.2307/3349861

Vilas, B., Khandare, V., Abdul, S., & Maheboob, A. (2019). Trends of public expenditure in India: An empirical analysis. *International Journal of Social Science and Economic Research, 4*(5).

Vogler, S., Zimmermann, N., Leopold, C., & Joncheere, K. (2011). Pharmaceutical policies in European countries in response to the global financial crisis. *Southern Med Review, 4*(2), 69-79.

Wettermark, B., Wilking, N., Kalin, M., Korkmaz, S., Hjemdahl, P., Godman, B., . . . Persson, M. (2010). Forecasting drug utilization and expenditure in a metropolitan health region. *BMC Health Services Research, 10*(128). doi:http://doi.org/10.1186/1472-6963-10-128

Wooldridge, J. (2013). *Introductory Econometrics: A modern approach* (5 ed.). Cengage Learning.

APPENDIX 1 - BASICS OF MULTIPLE LINEAR REGRESSION

Regression analysis is a statistical technique to understand the relationships between a dependent variable and a set of explanatory variables, whose final purpose is to generate a model that can be used to forecast future values of the dependent variable given specified values of the explanatory variables.

The general form of a multiple linear regression model for time series data can be written as follows:

$$y_t = \beta_0 + \beta_1 x_{t,1} + \beta_2 x_{t,2} + \cdots + \beta_k x_{t,k} + \varepsilon_t, \quad t = 1, 2, \ldots, T \tag{A1}$$

where y_t is the dependent variable associated to year t, $x_{t,k}$ is the explanatory variable k associated to year t, β_0 is the intercept, β_k is the coefficient associated to the explanatory variable k and measures the effect that a change on the explanatory variable k has on the dependent variable given all the other variables constants, and ε_t is the random error and corresponds to possible unknown effects on the dependent variable.

Usually, the unknown parameters β_k are estimated using the method of Ordinary Least Squares (OLS) that chooses the estimates to minimize the sum of the squared residuals. After the estimation of the parameters, (1-α)% confindence intervals may be constructed in order to assess the marginal effect of a explanatory variable using the formula:

$$\hat{\beta}_k \pm t_{\frac{\alpha}{2}, n-p'} * se(\hat{\beta}_k), \tag{A2}$$

where $t_{\frac{\alpha}{2}, n-p'}$ is the 1-$\frac{\alpha}{2}$ quantile of a $t_{n-p'}$ distribution and p' is the number of explanatory variables plus one.

When applying a multiple linear regression with time series data, it is important to analyse the stationarity of the series to avoid the possibility of a spurious regression. In fact, in the case of a spurious regression, the model is most possibly capturing a common stochastic trend between the variables and not a causal relationship. Additionally, the residuals of the regression may be non-stationary and strongly autocorrelated and, in these circumstances, the usual t and F tests carried out on OLS estimation do not follow since the t and F distributions are meaningless.

The test for significance of regression determines whether the relationship in (A1) holds or not and involves an analysis of variance of partitioning of the total sum of squares (SS_T) that represents the total variability in the dependent variable:

$$SS_T = \sum_{t=1}^{T} (y_t - \bar{y})^2 = SS_R + SS_E, \tag{A3}$$

where $\bar{y}$ is the mean of the observed dependent variable, SS_{R} is the sum of squares due to the regression, that is, the portion of the total variability in the dependent variable explained by the regression model, and SS_{E} is the sum of squares due to residuals, that is, the portion of the total variability in the dependent variable that is not explained by the regression. The hypothesis of the significance test are:

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \quad vs \quad H_1 : at\ least\ one\ \beta_k \neq 0 \tag{A4}$$

The test is based on a F-test statistic, F, that compares the model with just the intercept with the model with all the explanatory variables and if the test rejects H_0, the model is statistically significant. The significance of each coefficient given that all the other variables are included in the model is determined based on a T-test statistic, T, given by the quotient between the expected value and the standard error.

Additionally, to test how well the OLS regression fits the data, the statistic coefficient of multiple determination, R^2, given by the quocient between SS_{R} and SS_{T}, is used, measuring the amount of reduction in the variability of the dependent variable obtained by using the explanatory variables in the model. A value of R^2 close to zero indicates a poor fit of the OLS line and a value close to one indicates a good fit.

However, a large value of R^2 does not necessarily imply that the regression model is good, since adding variables to the model never decreases R^2 leading to a possible situation of overfitting. The adjusted R^2, $adj - R^2$, that adjusts for the number of parameters in the model, will not always increase as variables are added to the model and is a more reliable measure of the goodness-of-fit of the model.

One problem that may arise in a multiple linear regression and requires special attention is multicollinearity, that is, when there is a high correlation between two or more explanatory variables. One way to measure multicollinearity in a regression is through the Variance Inflation Factor (VIF), that can be calculated in the following way:

$$VIF_k = \frac{1}{1-{R_k}^2} \tag{A5}$$

where ${R_k}^2$ is obtained by fitting a regression involving only the explanatory variables and considering x_k as dependent variable. This problem does not affect the predictive power but may affect the precision of the estimation of the impact of the explanatory variables in the dependent variable since it directly affects the variance of the coefficients. In fact, the higher the correlation between the explanatory variable x_k and the other variables, the

higher the VIF and the variance of the coefficient $\hat{\beta}_k$, and the T-test statistic of the coefficient tends to be statistically insignificant.

Normally, if the VIF value is greater than 10 (which means a value of ${R_k}^2$ greater than 0.90), then multicollinearity may be a problem. Besides collecting more data, among different methods, one of the simplest methods to deal with multicollinearity is to remove highly correlated variables before fitting the regression or remove the explanatory variables with a VIF value greater than 10, prioritizing the variables that are more important to the research. Additionally, although backward elimination does not directly detect multicollinearity, it may reduce its effects by eliminating variables with weak marginal significance, which often arise due to high intercorrelation among explanatory variables (Kutner et al., 2005).

An important part of model-building procedure is checking the adequacy of the model and regression model residuals are very useful on checking how well the regression model assumptions of normally and independently distributed model errors with constant variance are satisfied.

Residuals plots are the primary approach to check model adequacy, however some statistical tests may be employed to test the following assumptions:

- Testing for autocorrelation

  Autocorrelation refers to the correlation of the regression residuals across observations, typically in time series data. The presence of autocorrelation violates one of the classical assumptions of the OLS method and may lead to inefficient estimates and underestimated standard errors. To test for autocorrelation, the Durbin-Watson test is commonly applied, particularly to detect first-order serial correlation. Values of the DW statistic close to 2 suggest no autocorrelation, while values significantly below or above 2 indicate positive or negative autocorrelation, respectively. Many textbooks still recommend using tabulated upper and lower bounds of critical values (dL and dU) depending on the significance level, number of observations and number of regressors. For higher-order autocorrelation, the Breusch-Godfrey test is more appropriate, as it allows for testing serial correlation beyond the first lag and in models that include lagged dependent variables.

- Testing for heteroskedasticity

  Heteroscedasticity arises when the variance of the error terms is not constant across observations, violating another key assumption of OLS. This can lead to biased

standard errors, resulting in unreliable hypothesis testing. Several tests are available to detect heteroscedasticity. One widely used method is the Breusch-Pagan test, which assesses whether the squared residuals from a regression are related to the independent variables.

- Testing for non-normality

Normality of the error terms is a desirable property in linear regression, particularly for conducting valid inference when sample sizes are small. Deviations from normality can affect the reliability of confidence intervals and hypothesis tests. To assess whether residuals are normally distributed, formal statistical tests such as the Shapiro-Wilk test are typically employed.

APPENDIX 2 - VARIABLES IN THE MODELS

Variable	Symbol	Definition	Years	Source
Dependent variables				
Oncological cytotoxics expenditure	CTE	Annual SNS expenditure on oncological cytotoxics pharmaceuticals (thous €) at mainland level	2007-2023	INFARMED**
Oncological immunomodulatings expenditure	IME	Annual SNS expenditure on oncological immunomodulatings pharmaceuticals (thous €) at mainland level	2007-2023	INFARMED**
Oncological hormones and anti-hormones expenditure	HE	Annual SNS expenditure on oncological hormones and anti-hormones pharmaceuticals (thous €) at mainland	2007-2023	INFARMED**
Independent variables				
New cases of cancer	NC014	Annual new cases of cancer diagnosed in the mainland SNS institutions in individuals aged under 15 years	2007-2022	RON**
	NC1564	Annual new cases of cancer diagnosed in the mainland SNS institutions in individuals aged between 15 and 64 years	2007-2022	RON**
	NC65	Annual new cases of cancer diagnosed in the mainland SNS institutions in individuals aged over 65 years	2007-2022	RON**
Breast cancer screenings	BSC	Number of annual breast cancer screenings performed by asymptomatic women aged between 50 and 69 in the mainland	2009-2023	DGS*
Cervical cancer screenings	CSC	Number of annual cervical cancer screenings performed by asymptomatic women aged between 25 and 60 in the mainland	2009-2023	DGS*
Colorectal cancer screenings	CRSC	Number of annual breast cancer screenings performed by asymptomatic men and women aged between 50 and 74 in the mainland	2009-2023	DGS*
Cytotoxics orphans	CTO	Annual consuption of oncological cytotoxics orphan pharmaceuticals (unities)	2007-2023	INFARMED**
Immunomodulatings biosimilar	IMB	Annual consumption of oncological immunomodulatings biosimilar pharmaceuticals (unities)	2007-2023	INFARMED**
Immunomodulatings orphans	IMO	Annual consuption of oncological immunomodulatings orphan pharmaceuticals (unities)	2007-2023	INFARMED**
New pharmaceuticals	NP	Annual number of new innovative oncological pharmaceuticals with financing decision	2007-2024	INFARMED**

Note: * indicates data publicly available and ** indicates data obtained in collaboration with institutions

Table 19: Information about the variables

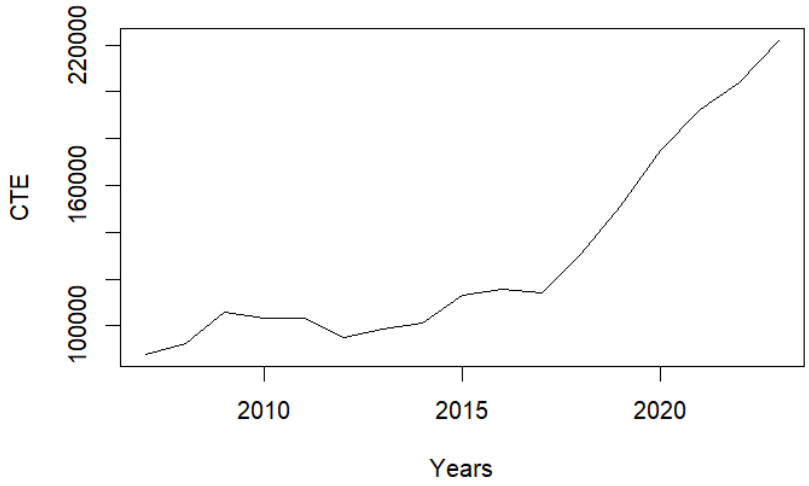

Figure 17: Oncological cytotoxics expenditure (thousand €) (2007-2023)

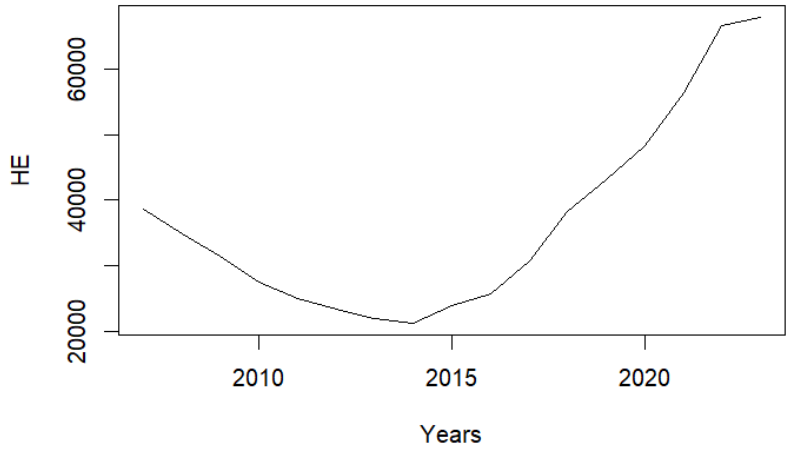

Figure 18: Oncological hormones and anti-hormones expenditure (thousand €) (2007-2023)

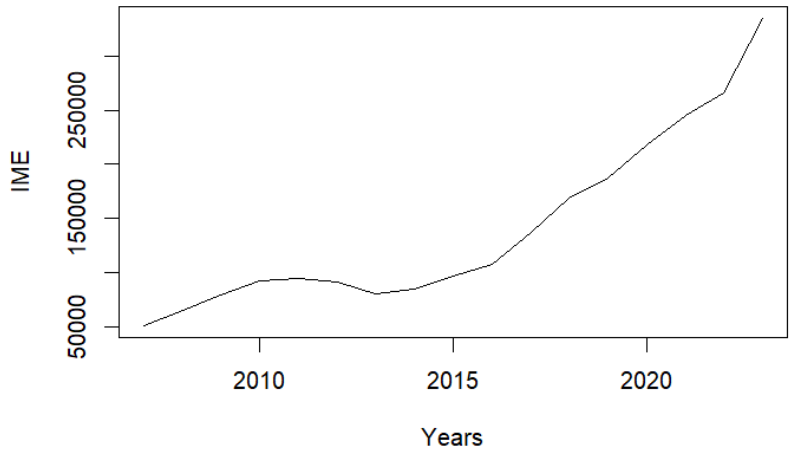

Figure 19: Oncological immunomodulatings expenditure (thousand €) (2007-2023)

APPENDIX 3 – STATIONARITY TESTS

Variable	ADF test	PP test
Log(CTE)	0.1464	1.3479
Log(HE)	0.2243	0.7356
Log(IME)	0.3518	0.4849
Log(NC014)	0.2803	-2.4534
Log(NC1564)	-1.6876	-3.8596 **
Log(NC65)	-1.6884	-2.6172
Log(BSC)	-1.9331	-2.2858
Log(CSC)	-2.757 *	-2.6181
Log(CRSC)	-0.2655	-0.367
Log(CTO)	-2.2152	-1.7772
Log(IMO)	-2.4412	-0.5348
Log(IMB)	-0.8718	-0.3345
Log(NP)	-8.178 *	-1.5309

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 20: Stationarity tests for the logarithm of the variables

Variable	ADF test	PP test	KPSS test
DLog(CTE)	-0.534	-1.9579	0.3979 *
DLog(HE)	-3.0895 **	-1.3743	
DLog(IME)	-7.787 ***	-1.927	
DLog(NC014)	-3.664 **	-4.5179 ***	
DLog(NC65)	-3.9168 ***	-5.6463 ***	
DLog(BSC)	-2.2365 **	-6.0641 ***	
DLog(CRSC)	-2.1149	-3.8867 **	
DLog(CTO)	-1.7904	-3.5049 **	
DLog(IMO)	-3.8119 ***	-3.3545 **	
DLog(IMB)	-1.9037	-2.5017	0.1503

Note: * indicates significance at 0.10 level, ** indicates significance at 0.05 level and *** indicates significance at 0.01 level

Table 21: Stationarity tests for the first differences of the logarithm of the variables

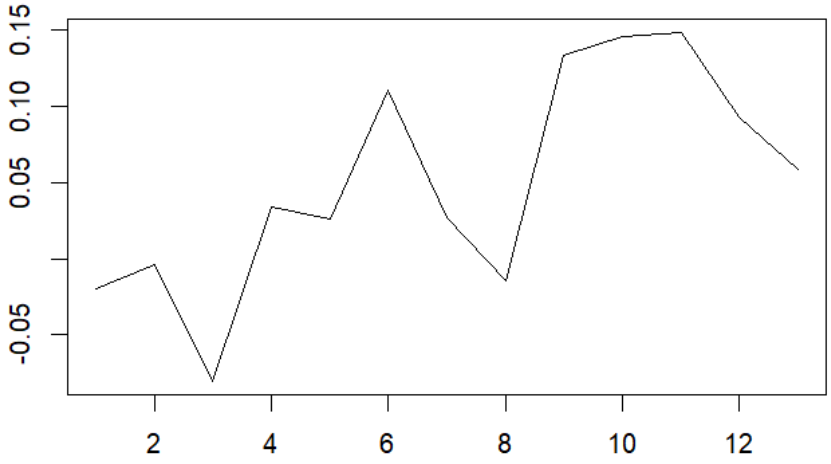

Figure 20: First difference of the variable Log(CTE)

APPENDIX 4 – RESULTS OF THE PROJECTION

Year	CTE	HE	IME	Total
2025	267238,850	99898,860	443837,769	810975,479
2026	291901,000	113248,000	514233,000	919382,000
2027	318192,250	127654,260	593622,751	1039469,261
2028	346112,600	143117,640	682580,568	1171810,808
2029	375662,050	159638,140	781679,997	1316980,187
2030	406840,600	177215,760	891494,584	1475550,944

Table 22: Direct projection of the oncological expenditures (thousand €) (2025-2030)

Year	TPE	GDP
2025	128 953 605	299 998 216
2026	134 892 235	314 681 632
2027	142 068 879	329 365 048
2028	149 616 882	344 048 465
2029	156 314 721	358 731 881
2030	162 973 168	373 415 297

Table 23: Projection of TPE and GDP (thousand €) (2025-2030)

APPENDIX 5 – ATTEMPT OF PROJECTION BASED ON THE REGRESSION

Despite the limited power of the regression models, an attempt of projection based on the explanatory variables was also conducted to compare it with the previous projection and give some insights for the future researches. The explanatory variables were also projected according to the study conducted by Ovchynikova and Dupliak (2021) (for the results see Table 24):

- For the variables NC1564 and NC65 trending lines were fitted using values from 2007 to 2021. Among the lines with a R^2 higher than 0.80, the one performing better on predicting 2022 value was chosen. For NC1564 it was not possible to choose a trending line, so it was used the annual average growth rate of 1.09% and for NC65 it was chosen a logarithmic trending line;
- For the variables CSC and CRSC trending lines were fitted using values from 2009 to 2019 to avoid the value from 2020 that was affected by COVID-19 pandemics. Among the lines with a R^2 higher than 0.80, the one performing better on predicting 2021, 2022 and 2023 values was chosen. For CSC it was chosen a linear trending line and for CRSC it was chosen a second-degree polynomial trending line;
- For the variable CTO, it was not possible to fit a trending line with R^2 higher than 0.80, so it was used the average annual growth rate of 4.11% from 2007 to 2023;
- For the variable IMB, a second-degree polynomial trending line was chosen.

The results of the projections that can be seen in Table 25, show a more optimistic scenario for the expenditure, corresponding to an annual average growth rate of 8.28%. The oncological expenditures on cytotoxics and immunomodulatings are projected to grow at an annual average rate of 4.30% and 3.19%, respectively, and hormones and anti-hormones expenditure at 26.93%. These results seem unreasonable and reinforce the limited reliability of the projections based on the explanatory variables due to the limited statistical power of the regressions.

Year	NC1564	NC65	CSC	CRSC	CTO	IMB
2025	36795	55310	305292	300363	440566	118713
2026	37197	55667	319740	348379	458661	138294
2027	37603	56008	334188	399960	477500	159293
2028	38014	56332	348636	455105	497112	181708
2029	38430	56642	363084	513814	517530	205542
2030	38850	56939	377532	576087	538787	230793

Table 24: Projection of the explanatory variables (2025-2030)

Year	CTE	HE	IME	Total
2025	235300,473	101708,791	386778,3	723787,564
2026	245934,821	126567,385	411752,093	784254,299
2027	256742,215	159124,886	431422,351	847289,453
2028	267747,595	202031,405	444995,442	914774,442
2029	278973,181	258935,117	451928,903	989837,201
2030	290439,166	334884,192	451968,889	1077292,247

Table 25: Projection of the oncological expenditures based on the regression (thousand €) (2025-2030)