ASPIC
ASSOCIAÇÃO PORTUGUESA DE INVESTIGAÇÃO EM CANCRO

Portrait of Cancer Research in Portugal

Retrato da Investigação em Cancro em Portugal

"la Caixa" Foundation

FUNDAÇÃO CALOUSTE GULBENKIAN

NOVARTIS

Acknowledgments

This report would not have been possible without the kind support and help of many individuals and several organizations.

We would like to thank the Aspic Members and cancer research experts that helped design and review this report, for their availability and willingness to give priceless insight:

- **Joana Paredes**, ASPIC (Associação Portuguesa de Investigação em Cancro), Presidente; i3S-Ipatimup (Instituto de Investigação e Inovação em Saúde, Universidade do Porto - Instituto de Patologia e Imunologia Molecular da Universidade do Porto)
- **Júlio Oliveira**, ASPIC (Associação Portuguesa de Investigação em Cancro), Vice-Presidente; IPO-Porto (Instituto Português de Oncologia do Porto)
- **Jorge Lima**, ASPIC (Associação Portuguesa de Investigação em Cancro), Vogal da Direção; i3S-Ipatimup (Instituto de Investigação e Inovação em Saúde, Universidade do Porto - Instituto de Patologia e Imunologia Molecular da Universidade do Porto)
- **Bruno Costa**, ASPIC (Associação Portuguesa de Investigação em Cancro), Vogal da Direção; ICVS (Instituto das Ciências da Vida e da Saúde, Universidade do Minho)
- **Cláudia Faria**, ASPIC (Associação Portuguesa de Investigação em Cancro), Vogal da Direção; IMM (Instituto de Medicina Molecular, Universidade de Lisboa), CHLN (Centro Hospitalar Universitário de Lisboa Norte)
- **Ana Preto**, ASPIC (Associação Portuguesa de Investigação em Cancro), Vogal Suplente da Direção; CBMA (Centro de Biologia Molecular e Ambiental, Universidade do Minho)
- **Carmen Jerónimo,** ASPIC (Associação Portuguesa de Investigação em Cancro), Vogal Suplente da Direção; IPO-Porto (Instituto Português de Oncologia do Porto)
- **Luísa Melo**, ASPIC (Associação Portuguesa de Investigação em Cancro), Coordenadora de Comunicação

- **Jordi Vallès**, “la Caixa” Foundation, Barcelona, Spain

We would like to deeply thank the extra effort and commitment of the following collaborators on this study:

- “la Caixa” Foundation
- Fundação Calouste Gulbenkian
- Novartis Portugal

As well as the national agencies that supported this project and kindly provide us information on their activities.

- Liga Portuguesa Contra o Cancro (LPCC)
  - **Prof. Dr. Vitor Rodrigues**- President of the board of LPCC; and **representatives** and **coordinators of the Regional Offices**: Norte, Centro, Sul, Açores and Madeira.

- INFARMED- National Authority of Medicines and Health Products
  - **Dr. Joel Passarinho**

This report was performed by the multidisciplinary team of **SIRIS Academic**, in coordination with ASPIC (Associação Portuguesa de Investigação em Cancro).

December 2021

INDEX

I. Sumário Executivo 1

I. Executive Summary 4

II. Introduction 7

III. Context and Objectives 9

IV. Brief Methodological Notes 11

V. Results 15

A. Bibliometric Production 15

Bibliometric Production (volume, specialisation and excellence) 17

Bibliometric Production per Research "Type" 22

Scientific Production by Cancer Site 25

Box 1. Portugal´s alignment with forefront topics ***28***

Impact of Published Work in Innovation and Clinical Practice 29

B. Clinical Trials 32

Volume and Growth 32

Sponsorship 34

Trial Phases 36

Intervention Types 37

Coordination and Actors (facilities and sponsors) 38

Research Focus 40

C. European Competitiveness 43

Horizon 2020 43

European Research Council (ERC) 46

Box 2. Main non-governmental funders of cancer research in the Portuguese ecosystem: Liga Portuguesa Contra o Cancro and "la Caixa" Foundation ***49***

D. National research ecosystem & Actors 51

Collaborative work, Leadership and Internationalisation 51

National Actors 52

VI. Main Messages and Insights **58**

VII. Methodology **61**

Controlled Vocabularies 61

Zero Shot Technique (classifying research projects) 65

Medical Subject Headings NLM-NIH (MeSH) 66

VIII. Annexes **67**

A. Publications and Clinical Trials according to Cancer groups **68**

B. Main Cancer MeSH Terms identified in Publications **70**

I. Sumário Executivo

O objetivo da ASPIC com a elaboração deste relatório encontra-se alinhado com uma das suas principais missões: ter um papel chave no debate sobre investigação em cancro, analisando e propondo soluções para questões relevantes para a investigação e para os investigadores em Portugal, assim como para a sua expansão e integração a nível internacional.

Assim, este trabalho pretende fornecer um **retrato da investigação em cancro a nível nacional**, atualmente inexistente, fazendo um mapeamento dos resultados da investigação e dos seus atores principais. O seu objetivo é iniciar discussões entre todos os interessados e identificar pontos fortes e alavancas para posicionar Portugal como uma voz forte na investigação em cancro a nível global.

Com este propósito, a ASPIC trabalhou em conjunto com a consultora SIRIS Academic, que concebeu e desenvolveu este estudo - **o primeiro a ser realizado a nível nacional e um primeiro passo fundamental para iniciar discussões estratégicas e políticas**. Este relatório é o resultado deste trabalho de colaboração, que visa avaliar o estado atual da investigação em cancro em Portugal na última década, bem como o potencial de Portugal para competir a nível internacional, analisando as evidências que sustentam ou dissipam as assunções atuais.

Este documento fornece uma **visão imparcial, e baseada em evidência, do país, bem como a forma como este se compara a nível internacional**, identificando os seus pontos fortes e os seus pontos fracos. Pretende-se que se constitua como uma ferramenta de trabalho, que poderá ser aproveitada pelos decisores com envolvimento ou responsabilidade em investigação em cancro a nível nacional.

Esta análise demonstra assim, de uma forma geral, que **o sistema de investigação nacional em cancro está em expansão** e é **capaz de produzir investigação de excelência que é competitiva a nível europeu**. No entanto, o conhecimento produzido ainda **carece de consolidação** e de **tradução eficaz em benefício do paciente ou da sociedade**.

O sistema de investigação biomédica em Portugal está em clara expansão, incluindo a investigação em cancro que, apesar de ainda representar apenas 10% da investigação biomédica nacional, também aumentou significativamente na última década. Prova disso é o facto de a produção bibliométrica, *per capita*, ser semelhante a sistemas como os de França, Espanha e Alemanha. O impacto científico gerado é também muito positivo com métricas de citação semelhantes a Espanha e Irlanda e acima da Alemanha e Itália. Apesar deste crescimento e impacto científico positivo, a investigação em cancro em Portugal ainda não tem a correspondente visibilidade científica e reconhecimento pelos pares, o que não é incomum num país com perfil emergente.

Portugal está a seguir a tendência global no sentido de uma intensa globalização da investigação. Nesse sentido, e paralelamente ao seu crescimento e desenvolvimento, Portugal tem perdido uma pequena parte da liderança na investigação publicada, como contrapartida ao aumento da sua internacionalização. Para que um sistema pequeno e emergente como Portugal se consolide, seguir este tipo de direção poderá ser fundamental. É mais provável que a consolidação da investigação portuguesa em cancro assente num modelo altamente dinâmico e colaborativo, o que aumentará a visibilidade, o reconhecimento pelos pares e a competitividade global em termos de desempenho.

A investigação em cancro em Portugal é também cada vez mais competitiva a nível de financiamento europeu, o que demonstra a qualidade da investigação e a evolução do ecossistema português, mas também reflete a necessidade que os investigadores nacionais têm de recorrer a fontes de financiamento extraordinárias. Quanto à obtenção de financiamento pelo ERC (European Research Council), o ecossistema português é ainda jovem e menos propenso a atrair ou estabilizar investigadores em fases mais avançadas da carreira que competem a nível Europeu.

Apesar do número de ensaios clínicos em Portugal ter aumentado significativamente na última década, e dos números, per capita, serem semelhantes aos registados na Alemanha, Portugal apresenta um ecossistema de ensaios clínicos subdesenvolvido. Os ensaios clínicos que existem são altamente dependentes do apoio privado, o que implica um perfil mais alinhado com os interesses da indústria farmacêutica (por exemplo, dedicado à aprovação de medicamentos específicos), e potencialmente menos alinhado com as necessidades clínicas específicas do país ou de cada centro hospitalar envolvido nos estudos. Seguindo esta mesma linha, os centros clínicos portugueses estão principalmente envolvidos em grandes ensaios multicêntricos e multinacionais.

Além disso, a investigação em cancro publicada por equipas nacionais não é citada em diretrizes clínicas com a mesma frequência que a de outros sistemas. No seu conjunto, a informação disponível sugere que a investigação nacional em cancro não está a ter um elevado impacto na prática clínica.

Relativamente a como Portugal transforma a investigação em cancro em tecnologias ou práticas de inovação, o sistema apresenta poucos indicadores positivos. O número de patentes biomédicas de organizações e indivíduos portugueses registadas e aprovadas pelo registo europeu é ainda bastante baixo; assim como o é a citação de publicações nacionais em cancro em patentes aprovadas por terceiros. Quanto à capacidade de captar financiamento europeu para este tipo de atividade, o sistema português apresenta ainda bastantes dificuldades e muito poucas empresas privadas portuguesas são capazes de assegurar o financiamento da UE para o desenvolvimento da inovação a partir de instrumentos de PME.

Estes dados, apontam, uma vez mais, para um sistema de inovação que não está devidamente desenvolvido. Contudo, é de salientar o facto de várias iniciativas recentes estarem a concentrar-se no desenvolvimento da capacidade do país para fazer a transferência de conhecimentos no domínio biomédico e melhorar os cuidados aos pacientes, como é o caso da AICIB (Agência de Investigação Clínica e Inovação Biomédica) ou do P.CCC (Porto Comprehensive Cancer Center).

A nível geográfico e de instituições que fazem investigação em cancro, o Porto surge em destaque, concentrando-se na Universidade do Porto, ICBAS e no IPO Porto. Segue-se Lisboa, com uma rede maior de instituições, e depois Coimbra e a região do Minho, com um crescimento recente bastante substancial. Na vertente mais clínica da investigação, os principais hospitais nacionais e IPOs têm um papel fundamental. No entanto, Portugal conta com um subconjunto de centros de investigação mais fundamentais e translacionais, que apesar de não serem especializados na investigação do cancro *per se*, contribuem grandemente para o ecossistema português e alguns são altamente competitivos nos concursos europeus (por exemplo IMM e Centros Champalimaud).

Tendo como base as principais conclusões deste trabalho, a **Direção da ASPIC sugere algumas potenciais medidas**. Estas podem servir como **base de discussão com as entidades relevantes e com poder de decisão, de forma a melhorar a estratégia nacional para a investigação em cancro**, através da otimização dos recursos, do melhoramento dos atuais pontos fortes e do preenchimento das lacunas identificadas. Estas medidas podem ser resumidas em duas principais linhas de ação:

- Criar **centros especificamente dedicados à investigação em Oncologia**, com financiamento sustentável, incluindo centros académicos, centros clínicos e indústria. Esta ação pode ser atingida através da estruturação das instituições e recursos já existentes, considerando a possibilidade de constituição de "Cancer Research Hubs".

- **Investir em instrumentos, infraestruturas e recursos humanos transversais, dedicados a promover a translação da investigação básica** em abordagens que gerem produtos inovadores e que tenham impacto clínico.

I. Executive Summary

The aim of ASPIC with this report is aligned with one of its main missions: to play a key role in the cancer research debate, analyzing and proposing solutions to research and researchers in Portugal, as well as for its expansion and integration at the international level.

Thus, this work provides a currently lacking portrait of cancer research at the national level, by mapping the outputs of cancer research as well as its main actors. It aims to initiate discussions among all stakeholders and **identify strengths and leverages to position Portugal as a strong voice in cancer research for the future and at a global level.**

In this quest, ASPIC worked together with the consultancy SIRIS Academic, which designed and developed this study - **the first to be conducted at national level and a key first step to initiate strategic and policy discussions**. This report is the result of this collaborative work that aims to assess Portugal's potential to compete internationally, analysing the evidence that supports or dispels current assumptions.

It provides an **evidence-based and unbiased view of the country and how it compares internationally**, identifying its strengths and weaknesses. We expect that this report will be a working tool to be used by the stakeholders and policy makers, with an involvement or responsibility in cancer research at the national level.

This analysis shows that **Portugal's cancer research system is expanding** and is **able to produce excellent research that is competitive at the European level**. However, the knowledge that is produced still **lacks consolidation and an effective translation into patient or societal benefit**.

The biomedical research system in Portugal is clearly expanding, including cancer research that also increased significantly in the last decade, although still representing only 10% of national biomedical research. Proof of this is the fact that production, in terms of scientific articles *per capita*, is similar to systems such as France, Spain and Germany. The generated scientific impact is also very positive, with citation metrics similar to Spain and Ireland, actually surpassing Germany and Italy. Despite this growth, and positive scientific impact, cancer research in Portugal still lacks the corresponding scientific visibility and peer recognition, which is not unusual in a country with an emerging profile.

Portugal is following the global trends towards an intense globalisation of research. In this sense, and in parallel with its growth and development, Portugal has lost some leadership in published research, as a trade-off to its increasing internationalisation. For a small and emerging system like Portugal to consolidate, following this kind of direction may be fundamental. The consolidation of Portuguese cancer research is more likely to be based on a highly dynamic and collaborative model, which increases visibility, peer recognition and global competitiveness in terms of performance.

Cancer research in Portugal is also increasingly competitive at the European level, which demonstrates the research quality and the evolution of the Portuguese ecosystem. However, this also reflects the need for national researchers to resort to extraordinary funding sources. As for obtaining funding from the ERC (European Research Council), the Portuguese ecosystem is still young and less likely to attract or stabilise researchers at later career stages who compete at European level.

Although the number of clinical trials in Portugal has increased significantly in the last decade, and the numbers, per capita, are similar to those recorded in Germany, Portugal presents an underdeveloped clinical trial ecosystem. The clinical trials that do exist are highly dependent on private support, which implies a profile more aligned with the interests of the pharmaceutical industry (for example, dedicated to the approval of specific drugs), and potentially less with the specific clinical needs of the country or of each hospital or study centre involved in the studies. Consistent with the above, Portuguese clinical centres are mainly involved in large multicentre and multinational trials.

Furthermore, cancer research published by national teams is not cited in clinical guidelines with the same frequency as that of other countries. Taken together, the analysis of the available information suggests that national cancer research is not having a high impact on clinical practice.

Regarding how Portugal transforms cancer research into innovation technologies or practices, the system shows few positive indicators. The number of biomedical patents from Portuguese organisations and individuals, registered and approved by the European registry, is still quite low, as it is the citation of national cancer publications in patents approved by third parties. As for the ability to attract European funding for this type of activity, the Portuguese system still presents many difficulties and few private Portuguese companies are able to secure EU funding for the development of innovative solutions from SME instruments.

These data point to an innovation system that is not fully developed. However, it is worth highlighting the fact that several recent initiatives are focusing on developing the country's ability to transfer biomedical knowledge to improve patient care, such as the AICIB (Agency for Clinical Research and Biomedical Innovation) or the P.CCC (Porto Comprehensive Cancer Center). Careful identification and evaluation of existing gaps and how to overcome them would be advisable.

Geographically and in terms of institutions that carry out cancer research, Porto stands out, with the University of Porto (and associated institutions), ICBAS and the IPO Porto. It is followed by Lisbon, with a larger network of institutions, and then Coimbra and the Minho region, with a fairly substantial recent growth. On the more clinical side of research, the main national hospitals and IPOs play a fundamental role. However, Portugal has a subset of more fundamental and translational research centres, which despite not being specialised in cancer research *per se*, contribute greatly to the Portuguese ecosystem and some are highly competitive in European calls (e.g. IMM and Champalimaud Centres).

Based on the main conclusions of this work, **the Board of ASPIC suggests potential measures**. These can serve as a basis for discussion with relevant entities and decision-makers, in order to improve the national strategy for cancer research, by optimizing resources, improving current strengths and filling in the identified gaps. These measures can be summarized in two main lines of action:

- **Create centers specifically dedicated to research in Oncology**, with sustainable funding, including academic centers, clinical centers and industry. This action can be achieved through the structuring of existing institutions and resources, considering the possibility of setting up “Cancer Research Hubs”.

- **Invest in instruments, infrastructures and human resources, which are transversal and dedicated to promoting the translation of basic research** into approaches that generate innovative products and have clinical impact.

II. Introduction

Cancer is the second leading global cause of mortality accounting for nearly 10 million deaths in 2020 and 19,3 million newly diagnosed cases[1]. Roughly 70% of deaths from cancer occur in low and mid-income countries, which is linked to limited accessibility to preventive strategies and/or treatment services that underlie survival rate inequalities both within Europe[2] and in the world[3]. Female breast cancer has surpassed lung cancer as the most commonly diagnosed cancer, although lung cancer remains the leading cause of death. The burden of cancer is continuously growing and the number of cases are expected to increase 47% by 2040 from that observed in 2020[4]. This increased burden of the disease, only surpassed by cardiovascular disorders, is strongly associated with population ageing and increased life expectancies, as well as a decline in mortality from other diseases such as stroke. Risk factors such as alcohol consumption, unhealthy diets and reduced physical activity are also on the rise in many countries, mostly associated with globalisation and socio-economic development [4].

In Europe, in 2020, 3.5 million people were diagnosed with cancer and 1.3 million deaths were recorded. However, over 40% of deaths are preventable[5]. In order to reverse the trend of cancer as a leading cause of mortality, research and innovation in this field are currently of high priority for the EU, deemed necessary for the study and advancement of cancer prevention, diagnosis, treatment and long-term care.

Accordingly, in 2018, the Gross Domestic Expenditure on R&D (research and development) in the EU region amounted to 294.5 billion EUR and during Horizon 2020 around 1400 cancer research projects were granted a total funding of 2 billion EUR, equivalent to 0.3 billion EUR allocated per year[6]. Looking into the next decade, the European Commission has renewed the investment on the mission on cancer by allocating at least €320 million, mostly from the Horizon Europe research programme, for the next two years as they get underway. However, and likely due to the overly ambitious aims of the first announcement, the European Commission (EC) has scaled back on the main goal of the EU mission on cancer, from its initial ambition of *saving* the lives of more than 3 million people by 2030 to the *improvement* of the lives of this same number of people[7].

In 2020, Portugal accounted for more than 60 thousand new cases and 30 thousand deaths[8]. Throughout the last decades, the incidence of cases in the country has progressively

---

[1] Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, et al. Global Cancer Observatory: Cancer Today. Lyon: International Agency for Research on Cancer; 2020

[2] Rossi et al. TheEUROCARE-5 study on cancer survival in Europe 1999-2007: Database, quality checks and statistical analysis methods. Eur J Cancer. 2015 Sep 1:51(15):2104-19.

[3] Assessing national capacity for the prevention and control of noncommunicable diseases: report of the 2019 global survey. Geneva: World Health Organization; 2020.

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

[5] https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12154-Europe%E2%80%99s-Beating-Cancer-Plan_en

[6] https://www.europarl.europa.eu/cmsdata/215166/BECA%20Background%20note%20on%20cancer%20research%2020201110.pdf

[7] https://ec.europa.eu/info/sites/default/files/research_and_innovation/funding/documents/cancer_implementation_plan_for_publication_final_v2.pdf

[8] https://gco.iarc.fr/today/data/factsheets/populations/620-portugal-fact-sheets.pdf

increased, following the trend of other European countries[9].

In a commitment to reversing this trend, Portugal, together with Germany and Slovenia, currently holding the EU presidency trio (2020-2021), have signed a declaration on cancer research during the European cancer research summit held in Porto in May of 2021[10]. The Porto declaration is the result of the work done by several researchers, scientific and clinical leaders and political decision-makers, who have reinforced the need to broaden Europe's *Beating Cancer Plan*, particularly through extending and reinforcing the European network of Comprehensive Cancer Centres (CCCs). Specifically this action calls for a joint effort throughout Europe towards a comprehensive translational cancer research approach focused on personalised and precision medicine and covering the entire cancer research continuum, from prevention to care. Within a larger context this action is integrated in the *EU Mission on Cancer*, whose Board has clearly highlighted the need for investment in innovative research[11].

European countries, such as Germany, Netherlands, UK and Spain are pioneers in cancer research, however, a positive increase trend is also observed in Belgium and Portugal. Given this, collaborative innovative initiatives across European Member States are now more crucial than ever, aiming for research excellence that can translate into an actual fight against cancer.

[9] Wild CP, Espina C, Bauld L, et al. Cancer Prevention Europe. *Mol Oncol*. 2019;13(3):528-534. doi:10.1002/1878-0261.12455

[10] https://www.2021portugal.eu/media/ritlydfz/en_porto-declaration-on-cancer-research-_final.pdf

[11] https://op.europa.eu/en/publication-detail/-/publication/b389aad3-fd56-11ea-b44f-01aa75ed71a1/

III. Context and Objectives

The Portuguese Association for Cancer Research (Associação Portuguesa de Investigação em Cancro - ASPIC), formed in 2013, fosters cancer research in all its aspects and for public benefit. It aims to promote excellence in Portuguese cancer research and help in its dissemination and internationalisation. In addition, ASPIC aims to be a key actor in the debate over cancer research in Portugal and Europe, as well as pose questions and engage in reflections relevant to its researchers and other relevant stakeholders.

There is currently a worldwide **urgency to achieve a thorough understanding of cancer, in order to improve prevention, diagnosis and treatment, and overall enhance peoples' quality of life.** A key factor in achieving this, at least at the European level, is to better integrate and align cancer research activities, from basic to applied research, and reinforce and activate networks and infrastructures.

Therefore, a contextualization of the Portuguese ecosystem in this regard is of high significance; it aims to shed light on its capacity, competitiveness and specialisation in order to better take advantage of the current context that could further boost its research excellence and impact.

In line with its mission, centred around promoting the value of cancer research and to be a forum that represents and integrates all professionals in the cancer field, **ASPIC wishes to provide such a portrait of cancer research at the national level. It aims to map all research outputs and main actors, to initiate discussions among all stakeholders and identify strengths and leverages to position Portugal as a strong voice in cancer research for the future and at a global level**.

A key objective is thus to have an as-thorough-as-possible understanding of Portugal´s cancer research, from basic to clinical research, and the development of clinical trials, its actors and networks. **It is the first study of its kind to be done at the national level and a key first step to initiate strategic and policy discussions**. In this endeavour, ASPIC had the support of important actors: "la Caixa" Foundation, Fundação Calouste Gulbenkian and Novartis Portugal. It also counted with the collaboration of Liga Portuguesa Contra o Cancro and INFARMED.

This analytical study characterizes the main research outputs, the results in terms of volume and scientific quality, the competitiveness of the Portuguese research ecosystem and its collaboration networks and national hubs. It further provides a general assessment of how these research outputs translate into innovation indicators, such as contribution to patents and SME generation. Throughout, it compares Portuguese performance with peers in an international context, and further identifies/confirms areas of excellence and deficit.

It does not intend to focus on direct analyses for policy purposes, such as those concerning funding allocation or advocacy. Analyses of this type would need an already existing evidence-based understanding of the ecosystem, which, as mentioned above, is currently non-existent. The overarching objective of this report is to **deliver data and insights for a specialized audience in the association, and to foster informed discussions with key stakeholders and decision makers over cancer research in Portugal**. In order to do so, it analyses cancer research production from different perspectives, and includes pertinent interpretations for mid-term conversations over the Portuguese research ecosystem. In addition, some reflections will be made taking into account other pertinent actors in the field, such as philanthropic funders of cancer research in Portugal.

Portugal's biomedical research ecosystem is often seen internally as one that produces (or can produce) excellent research, but that still struggles to compete at the global level. Some perceived reasons are the lack of a long-term and committed national research strategy; overdependence on national state funding and lack of continuity/predictability in funding schemes; bureaucracy; and the small size of the country (and therefore of critical mass). Specifically in cancer research, other factors should also be considered, such as sparse communication between clinical and basic researchers, low coordination of fundamental and applied research and its associated infrastructures, low cooperation between actors at the national and international level and barriers to knowledge transfer and clinical development.
With this study, we aim to assess if indeed Portugal has the potential to compete at the international level, by analysing the evidence that sustains or dispels this assumption, and discuss possible avenues for the future.
We will provide an **evidence based and unbiased view of the country, and how it compares to others. That shall permit leveraging its strengths and optimising its weaknesses by the stakeholders with an involvement and interest/responsibility in the field**.

IV. Brief Methodological Notes

In this section, some methodological concepts are introduced which are necessary to fully grasp how the work was conducted and the extent of the interpretations that can be derived from it.

Co-designing and steering an analytical report

According to SIRIS Academic philosophy, the first step of this report consisted of a co-design phase, followed by cyclic interactions and discussion with the ASPIC project team. The main steps followed are depicted below:

1. Establishing clear goals for the report: all quantitative and qualitative analysis shall be aligned with the goals of the study.
2. Defining the scope of the analysis: For example, dimensions to analyse, data procurement and data sources to be used and time window.
3. Validating key methodological aspects: identification of relevant benchmarks and the conceptual definition and perimeter of cancer research.
4. Accompanying and validating main findings and interpretations.

This process unraveled with a strong engagement of the ASPIC team and, where broader discussions took place, with additional input from key collaborators. ASPIC's project team included its president, vice-president and other representative board members, in addition to the Association´s Communication Coordinator.

Benchmarking the Portuguese research system

In order to contextualize the outputs of Portugal's cancer research, a panel of national systems used as comparators were defined *a priori* together with ASPIC's project team. Many arguments could be taken into account to define such types of panels, all depending on the objective of the intended analysis. The comparator panels for this project were chosen upon arguments based on excellency goals, positioning in the global landscape (main actors or smaller systems in volume of publications and citation indexes) and similarity of external conditions (e.g. degree of public spending in R&D, higher education model or public health system coverage).

In accordance, the resulting panel is constituted by ecosystems of established research excellence (UK, Netherlands, Germany), countries neighbouring Portugal, of similar size (Ireland, Belgium) and/or similar health systems (France, Spain, Italy), and also contemplates countries featuring a different research structure (UK, Ireland, Belgium)[12]:

---

[12] For this study, research systems are not normalized by researchers' numbers, supportive structures or available direct or competitive funding. Therefore, any comparison has to be taken cautiously, since we cannot account for *all* the different variables affecting the research outcomes of each country.

Benchmark Countries:

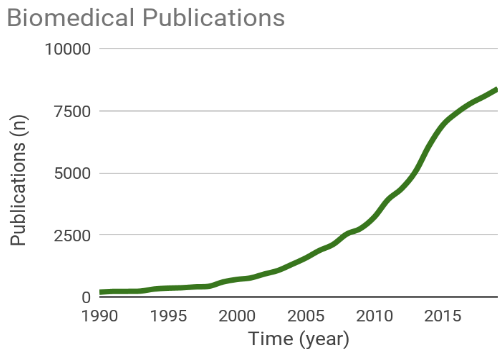
- UK
- Netherlands
- France
- Spain
- Italy
- Ireland
- Belgium
- Germany

Data sources and time window

To provide a holistic view of research in the field of cancer, the present work analyses an extremely large volume of data, which tries to cover the spectrum of R&D inputs and outputs as extensively as possible. For a global and international perspective, uniformized and open access to data sources through an API (Application Programming Interface) was preferred for comparison exercises.

Following the goals of this report, focused on better understanding Portuguese research outputs and national characteristics to assess its international positioning, information on national funding was not included. In agreement, the EU funding analysis was aimed at understanding Portugal's competitiveness amongst European partners and not in a perspective of advocating for funding reforms (as commented in Section III. Context and Objectives).

The time window of analysis was also discussed prior to the beginning of the study. The biomedical national research ecosystem in Portugal has had unprecedented growth in the last decades (see the figure below) but, in particular, especially between 2010 and 2015[13]. However, as the goals of this study were centred in the present and future of cancer research, the last decade was chosen as the relevant analysis period (2010-2020)[14].



Progression on the raw number of biomedical publications in Portugal (1990-2019)

[13] Rough estimation generated by SIRIS Academic, using a broad biomedical vocabulary and Scopus Application Programming Interface (API) resources. Therefore, the quality and precision of these results is lower than other analyses conducted in this study, but nonetheless useful for an evaluation of the appropriate time window to be used in subsequent analyses.
[14] With the exception of the time period used for the analysis of Research Projects, aligned with the H2020 Funding programme (2014-2020).

The full list of data sources and time periods analysed can be seen in the following table.

Scope	Data Source	Details	Open Access	Time Window
Publications	PubMed (National Institute of Health (NIH), USA)		Yes	2010-2020
	Scopus (Elsevier)		No	
Research Projects (competitive calls)	CORDIS (European Commission)	H2020 Framework Programmes	Yes	2014-2020
Patents	Lens.org	With data from different patent offices (e.g. EPO, WIPO)	Yes	2010-2020
Clinical Trials	Clinicaltrials.gov (National Institute of Health, USA)	Analysed through the AACT initiative	Yes	2011-2020
	National Registry_INFARMED	Data provided directly by the agency	No	2011-2020
Philanthropic Funding	“la Caixa” Foundation	Data available in their webpage	Yes	2018-2021
	Liga Portuguesa Contra o Cancro	Data provided directly by the agency	No	2011-2020

Data Sources, characteristics and time-window of the study

Identifying Cancer Research Outputs

Often research output classification systems are not useful to analyse specific and/or multidisciplinary topics; and even, if they exist, they are hardly consistent between different data sources (e.g. a specific bibliometric taxon like Scopus Subject Area is not replicated in EU frameworks or ERC funding). Therefore, it is imperative to have a strategy that enables the identification of documents for a particular area of interest across any data source: in this instance, two main semantic technologies were used (combining Natural Language Processing and Machine Learning; see section VI. Methodology for more details):

- **Controlled vocabularies (VOCs)**: for a particular area of interest, controlled vocabularies (VOCs) are created by or in close collaboration with field experts. Ideally, they are composed of unequivocal terms that fully represent a specific area (more complex VOCs may work in the intersection of two others, especially for multidisciplinary fields, such as for example Biomedical Engineering). These vocabularies are then used to identify the documents pertaining to a given research area by scanning Title, Abstracts and/or author keywords. For this study **2 Controlled vocabularies were constructed to classify publications:**
    - Cancer Research
    - Research “type”: public health, epidemiology & clinical research, and, by exclusion we identify basic and translational research[15].
- **ZeroShot**: is an unsupervised machine learning technique that automatically

[15] For this study, an *ad-hoc* controlled vocabulary by research “type” was constructed based on Portugal's publications. Since there is a good degree of grey areas in the definition of research “types” (categorisations that could also be considered artificial),- commonly: fundamental, basic, translational, preclinical, clinical, epidemiological, public health research, etc - we kept a simple structure with the 2 main groups that, are generally debated and that can, at this stage, be more accurately separated.

generates a probability score (0 to 1) of a given text belonging to a certain domain[16]. This Artificial Intelligence-based technique allows for an extra level of refinement because it does not rely on binary systems of classification (e.g. the texts having or not specific keywords), since the technique derives a probability score based on all the text provided. This technique is especially useful for analysis of research proposals[17] and, in this study, it was used in combination with the cancer research VOC to classify **EU funded projects** in a way that minimizes false positive cases.

Assigning outputs to a research system (countries, regions) or an institutional affiliation

With the constructed controlled vocabulary, we identified in PubMed around 1M publications worldwide, which were subsequently retrieved from the Scopus database (only *citable published work*[18] was included). When considering the **number of publications** attributed to a given system, country or institution, we must consider that, for this study:

- **One single publication can be attributed to multiple research systems in case of co-authorship.** This is regardless of the position/share of authorship (therefore, the number of publications of a given system is not a "share" of research, but rather the total number of published work in which any research institution from that system was involved).

- In the same line, **multiple institutional affiliations of a single author are also counted**. This means that **one single publication can be attributed to more than one institution** (or even country in case multiple affiliation includes a publication with ties to institutions in different countries), even if it relates to one single researcher.

Regarding European funded projects and clinical trials, additional critical information exists that allow us to avoid overestimating research outputs for a given system:

- **Number of projects/trials is assigned to all partners** (i.e., a same project trial will be counted in as many systems as it has participating organizations);
- Whilst **there is specific information over:**
  - **Budget allocation in EU projects** (and therefore, **funding sums are correlative to the participation of each partner in the project**)
  - Participant **role in the clinical trial** (coordinator or participant).

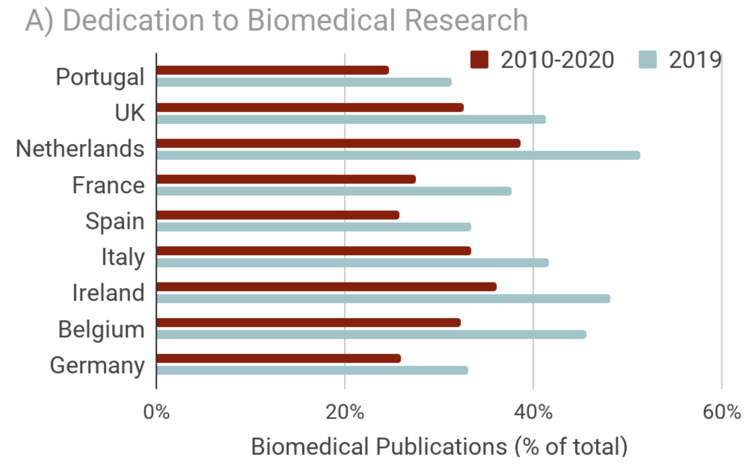
[16] Zero-shot Text Classification via Reinforced Self-training (2020): https://aclanthology.org/2020.acl-main.272.pdf
[17] When compared to Publications, Research Projects are more likely to mention major or global health challenges as well as potential "applications" or "relevancy for disease understanding". These common references may lead to higher rates of misclassification of the projects into biomedical areas. In response, more complex approaches to the identification of projects were deemed necessary for this analysis.
[18] Here, "citable" refers to original work, and includes journal articles, reviews, and conference papers.

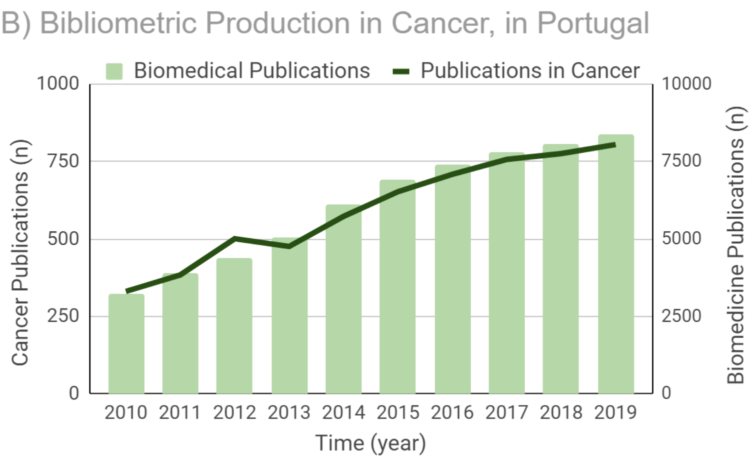
V. Results

A. Bibliometric Production

Biomedical research is one of the major areas of contemporary research, due to both its social and economical relevance, representing around 25% of H2020 funding[19]. Biomedical research currently represents more than 30% of the total citable published work in most EU27 countries analysed (*Fig. 1A*). Albeit Portugal being barely at that mark in 2020, its focus in **the biomedical domain has increased considerably in the last decade, from approximately 20% in 2010 to over 30% in 2019 of total national research** (*Fig. 1A*). **Paralleling this increase, was the rise in the number of national publications in cancer research** (*Fig. 1B*).

The slightly lower dedication to biomedical research may be due, in part, to the prominence of other areas in Portuguese research such as Engineering and Computer Sciences (*Fig. 1C*). This pattern is quite different from the case of the Netherlands, for example, where top subject areas are Medicine and Biochemistry, followed by Engineering, but where the Engineering area is almost three times smaller than Medicine.





[19] Source:analysis by SIRIS Academic, using data from CORDIS.

C) Scopus Subject Area (2019)	Portugal's Publs. (n)
Engineering	5,783
Medicine	5,414
Computer Science	4,767
Physics and Astronomy	2,954
Social Sciences	2,922
Biochemistry, Genetics and Molecular Biology	2,705
Agricultural and Biological Sciences	2,650
Environmental Science	2,646
Materials Science	2,496
Mathematics	2,296
Chemistry	2,205
Chemical Engineering	1,399
Energy	1,330
Earth and Planetary Sciences	1,299
Business, Management and Accounting	1,104
Arts and Humanities	952
Pharmacology, Toxicology and Pharmaceutics	859
Immunology and Microbiology	787
Psychology	772
Decision Sciences	654
Neuroscience	630
Multidisciplinary	560
Health Professions	509
Economics, Econometrics and Finance	472
Other (Nursing/ Veterinary/ Dentistry)	684

***Fig. 1. Biomedical and cancer publications in Portugal, registered in Scopus database : A)** share of biomedical research publications in Portugal and the panel of selected benchmark countries (2010-2020 and 2019 alone[20]); **B)** number of cancer and biomedical publications per year (2010-2019); **C)** distribution of all Portugal's publications according to Scopus Subject Area (2019).*

In order to further understand the Portuguese research ecosystem, we will discuss below the scientific outputs of Portugal's cancer research, mainly in comparison to the selected panel of comparators. The goal of this report is to position Portugal versus other research ecosystems.

[20] 2019 information is more robust and therefore used (instead of 2020) in some analysis throughout the study. Data for 2020 is less consistent due to delay in registry entries of bibliometric databases.

Bibliometric Production (volume, specialisation and excellence)

Due to its growing medical and social burden, cancer research is undoubtedly one of the main fields of biomedical research, accounting for the highest number of publications in most countries (especially so for developed countries).

The most productive country globally, both due to its size and commitment to research, is the USA. In addition, China, Japan are also large contributors, followed by several European countries (Germany, UK, Italy; and to a smaller extent, France, Netherlands and Spain).

Accordingly, the major institutions publishing research in cancer can be found in North America (USA and Canada) whereas at the European level, we find key players in the German Cancer Research Centre (DKFZ), followed by the Karolinska Institute, the Erasmus University Medical Center and Inserm[21] (*Table 1*).

Country	Publs (% world)	Institution	Country	Publs (% world)
United States	30.1	German Cancer Research Center (DKFZ)	Germany	0.7
China	18.1	Karolinska Institutet	Sweden	0.5
Japan	7.9	Erasmus University Medical Center	Netherlands	0.5
Germany	6.5	Inserm	France	0.5
UK	6.2	Università degli Studi di Roma La Sapienza	Italy	0.5
Italy	6.2	Charite - Universitätsmedizin Berlin	Germany	0.5
France	4.5	Karolinska University Hospital	Sweden	0.5
Canada	3.9	Medizinische Universität Wien	Austria	0.5
South Korea	3.6	Università degli Studi di Milano	Italy	0.5
Netherlands	3.1	University Medical Center Utrecht	Netherlands	0.4
Spain	3.1	Università degli Studi di Torino	Italy	0.4
Australia	2.9	Radboud University Nijmegen Medical Centre	Netherlands	0.4
India	2.3	Universitätsklinikum Hamburg-Eppendorf und Med. Fakultät	Germany	0.4
Taiwan	1.9	Fondazione IRCCS Istituto Nazionale dei Tumori, Milan	Italy	0.4
Switzerland	1.8	Leiden University Medical Center - LUMC	Netherlands	0.4
Sweden	1.7	Universitat Heidelberg	Germany	0.4
Brazil	1.6	Alma Mater Studiorum Universita di Bologna	Italy	0.4
Turkey	1.5	VU University Medical Center	Netherlands	0.4
Belgium	1.4	UCL	UK	0.4
Poland	1.3	Universitätsklinikum Heidelberg	Germany	0.4
Denmark	1.2	Imperial College London	UK	0.3
Austria	1.0	Université de Paris	France	0.3
Greece	0.9	Università degli Studi di Napoli Federico II	Italy	0.3
Iran	0.9	Institut de Cancérologie Gustave Roussy	France	0.3
Israel	0.9	The Netherlands Cancer Institute	Netherlands	0.3
Norway	0.8	Università degli Studi de Padova	Italy	0.3
Singapore	0.7	Universitair Medisch Centrum Groningen	Netherlands	0.3
Czech Republic	0.6	Ludwig-Maximilians Universitat Munchen	Germany	0.3
Finland	0.6	University of Cambridge	UK	0.3
Portugal	**0.6**	Institute of Cancer Research London	UK	0.3

Table 1. Major countries and institutions publishing cancer research in EU27+UK (2010-2020): *Percentage of participation in world's publications in the field.*

[21] INSERM, in France, aggregates a myriad of groups spread in organizations across the country; therefore, it cannot be compared in size to a particular research Institute, Hospital or University.

Portugal currently ranks #30 in global cancer publications production, presenting a similar trend to Finland, between 2010 and 2020. Considering the panel of benchmarks, and although far from highly performing countries such as the UK, Germany or Italy, **Portugal has increased 1.5 times its cancer publication volume in the last decade, positioning itself in a growth phase distinct from other benchmarks** (*Fig. 2A-B*). In addition, Portugal's **production per capita** (7.8 cancer publications per 100,000 inhabitants) **is already in line with countries such as Spain** (8.0) **and France** (7.3) (*Fig. 2B*). However, despite the country's visible increase of scientific publications, **its dedication to cancer (~3%) is still low** compared to other benchmark countries, as well as, to the global average (~4%) (*Fig. 2C*).

This low specialisation index in the field of cancer could be attributed to a diluted research ecosystem:

- On one hand, Portugal has not been overly specialized in biomedical research in the last decades (although it has been growing, and biomedical research represents now close to 30% of the Portuguese research) (*Fig. 1*).

- On the other hand, within the biomedical field, around **10% of Portuguese publications in the last decade were in the field of oncology**, which is on the lower tier (*Fig. 2D*). The distribution of Portuguese biomedical research per different disease groups highlights the importance of the research on Nervous System, Cardiovascular diseases and Infections, which could contribute to a lower specialisation in Cancer (*Fig. 2E*, Spain is shown for comparison)[22]. Furthermore, SIRIS has estimated the Portuguese dedication to other biomedical domains, like neurosciences (~17%, including mental disorders), infectious diseases (~14%) or immunology (~9%) (please consider that all areas have overlapping topics, e.g. meningitis will be classified as neurosciences and infectious disease)[23]; and the results agree with the fact that biomedical research in Portugal shows significant dedication to other areas.

[22] Please consider that: a) publications are more readily identified as cancer research due to the widespread use of *in vitro* models of neoplastic origin; which means it is often slightly overestimated in this specific analysis; and b) this classification per Disease Group is not equivalent to broader biomedical research domains. For example, Neuroscience is larger than Nervous Systems Diseases (since it also includes Mental Disorders and Cognitive Sciences not associated with disease); Or Immunology, for which a very significant portion of research is related to basic immunological phenomena and inflammation which are not necessarily associated with immune diseases.
[23] For this estimation, SIRIS applied controlled vocabularies for specific biomedical areas; however, this was not done with the technical sophistication utilized for the Cancer VOC, and, therefore, the precision is not in the same range as the results presented here for Cancer Research.

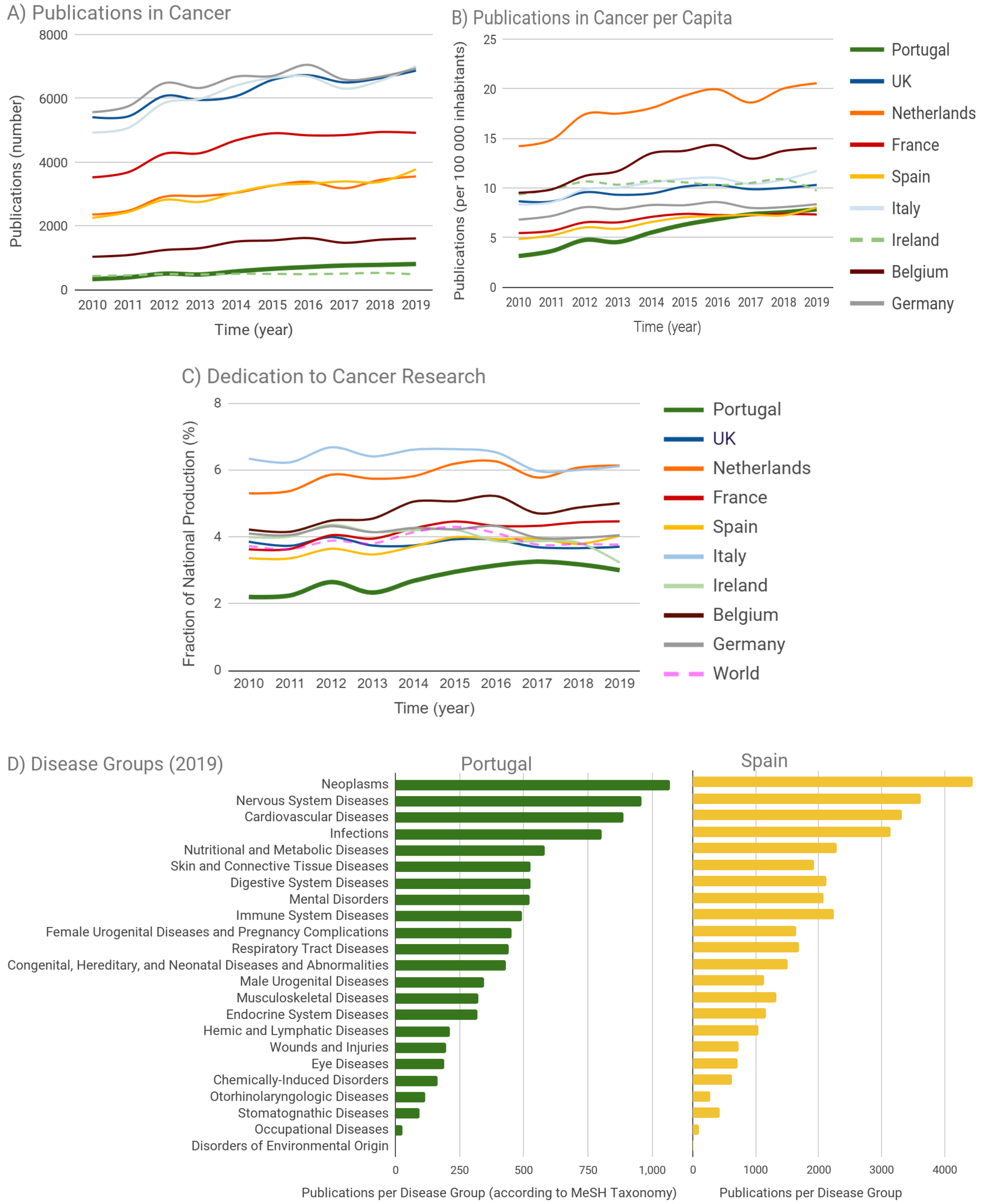


Fig. 2. Cancer publications in the panel of countries (2010-2019): A) *number of cancer Publications per year;* ***B)*** *number of cancer Publications per year per 100 000 inhabitants;* ***C)*** *percentage of cancer publications with regard to the total national production (Ireland´s dashed line, instead of a solid line has been used for easier visualisation);* ***D)*** *number of publications per disease group according to MeSH Taxonomy for Portugal and Spain (Branch [C:Diseases](#) and [F03: Mental Disorders](#); please note that this classification is not excludent, for example work in Lung Cancer will also be under Respiratory tract diseases) (2019[24]).*

[24] We used data from 2019 instead of 2020, since 2020 data reflects the redirection of several research efforts into the Infections field in relation to the COVID 19 pandemic, which would bias the results.

Furthermore, regarding scientific impact and peer recognition, measured here by the Normalized Citation Index (NCI[25]) and the proportion of publications within the top most cited journals according to Scimago[26], both Netherlands and Belgium are currently leading countries with a NCI above 1.65 and over 60% of cancer publications in top 10% journals, and 11-12% in the top 1% journals of their respective bibliometric category (*Fig. 3A-B*).

Portugal, on the other hand, **has a NCI of 1.44, which is in line with Spain and Ireland, and actually positioned above Germany and Italy** (*Fig. 3A*). At present, and despite a slight increase can be appreciated throughout the last decade, Portugal still has the **lowest rate of publications in top 10% and 1% journals** compared to the selected benchmarks: around 41% of publications in top 10% journals and just 7.7% of publications in top 1% journals (*Fig. 3B*).

It is also important to note that all research "excellence" metrics have their caveats and, thus, looking at more than one metric is useful. Although there is not necessarily a direct correlation between citations and publications in top tier journals, the data suggests that Portugal's published work is being cited at a frequency higher than its capacity to publish in top tier journals. This indicates that, **despite its successful growth the system is still striving for scientific recognition** (even if it achieves significant citation metrics).

Along this line, Portugal is not publishing in highly prestigious journals as those considered for the Nature Index[27], to the same extent as the UK or the Netherlands *(Fig. 3C*). Nevertheless, its share of published research in these journals is increasing along other major systems (see 2015-2020 values in *Fig. 3C*), while this is not the case for Italy or Ireland, supporting the idea of **growing excellence in Portuguese cancer research**.

Portugal does not have a single institution referenced in the top200 in the Nature Index Cancer 2020 either[28] (in part because it does not have many hyper specialized institutions *per se*). The rank is dominated by USA institutions, while Germany, UK and the Netherlands place several research organizations in the ranking, and Spain the Spanish National Cancer Research Centre (CNIO) (#118)[29]. We must recognize that some of the institutions in this top200 represent very large Universities and nation-wide organizations (e.g. Max Planck Society), for which Portugal also does not have a comparable organization.

Overall, critical mass, especially in smaller ecosystems, may have a great impact in research excellence. Often, there is a direct correlation between the degree of institutional/ecosystem

---

[25] The NCI takes into consideration citation patterns for a given corpus of publications (in this case a biomedical field), and can be understood as a measure of normalized "scientific impact" of the publications (versus other citation metrics based on journals like Scimago Journal Ranking and Web of Science Journal Impact Factor).
Being that it is an Index, the NCI is calculated for a specific baseline, in this case all the publications in the Scopus database identified as relating to Cancer Research (2010-2020). A NCI=1 means that a particular corpus of publications is cited, on average, the same as the baseline; A NCI around 2 would mean that they are cited twice the baseline (therefore they are plotted in a logarithmic scale).
[26] For more information see https://www.scimagojr.com/files/SJR2.pdf
[27] The Nature Index defines itself as an indicator of research performance (albeit biasing favourably institutions of a bigger size). The metrics of Count and Share used to order Nature Index listings are based on an institution's or country's publication output in 82 natural-science journals, selected on reputation by an independent panel of leading scientists in their fields. Natureindex.com, Nature Index 2020 Cancer | Supplements | Nature Index
[28] Nature Index 2020 Cancer | Supplements | Nature Index, study 2015-2019
[29] And also Barcelona Institute of Science and Technology (BIST) at #109, which represents an alliance of centers that on their own don't enter the ranking.

specialisation in a given field and in the quality achieved; e.g. the more a system is steered towards a given topic (in both human resources and infrastructures), the more likely it is to achieve excellence, since we assume it develops significant critical mass and physical resources that allows for excellence research to flourish. For instance, this can be observed in countries such as the Netherlands, in which a higher dedication to cancer research is correlated to a higher citation index (*Fig. 3A*).

However, for intrinsic reasons, systems can deviate from this broad pattern. Italy shows a negative deviation from this correlation with less impact (NCI) despite having the highest specialisation rate[30] (overall, a lower NCI in the Italian system is something we have observed for other biomedical areas). Whereas smaller systems, like Belgium and Portugal present a significant NCI despite low to moderate specialisation (there is a known dilution effect - like that observed for Germany - where notorious research may be diluted in a large number of publications). **Notably, despite a slightly lower specialisation and a very small production, Portugal displays a NCI in cancer research comparable to that of Spain** (*Fig. 3A*).

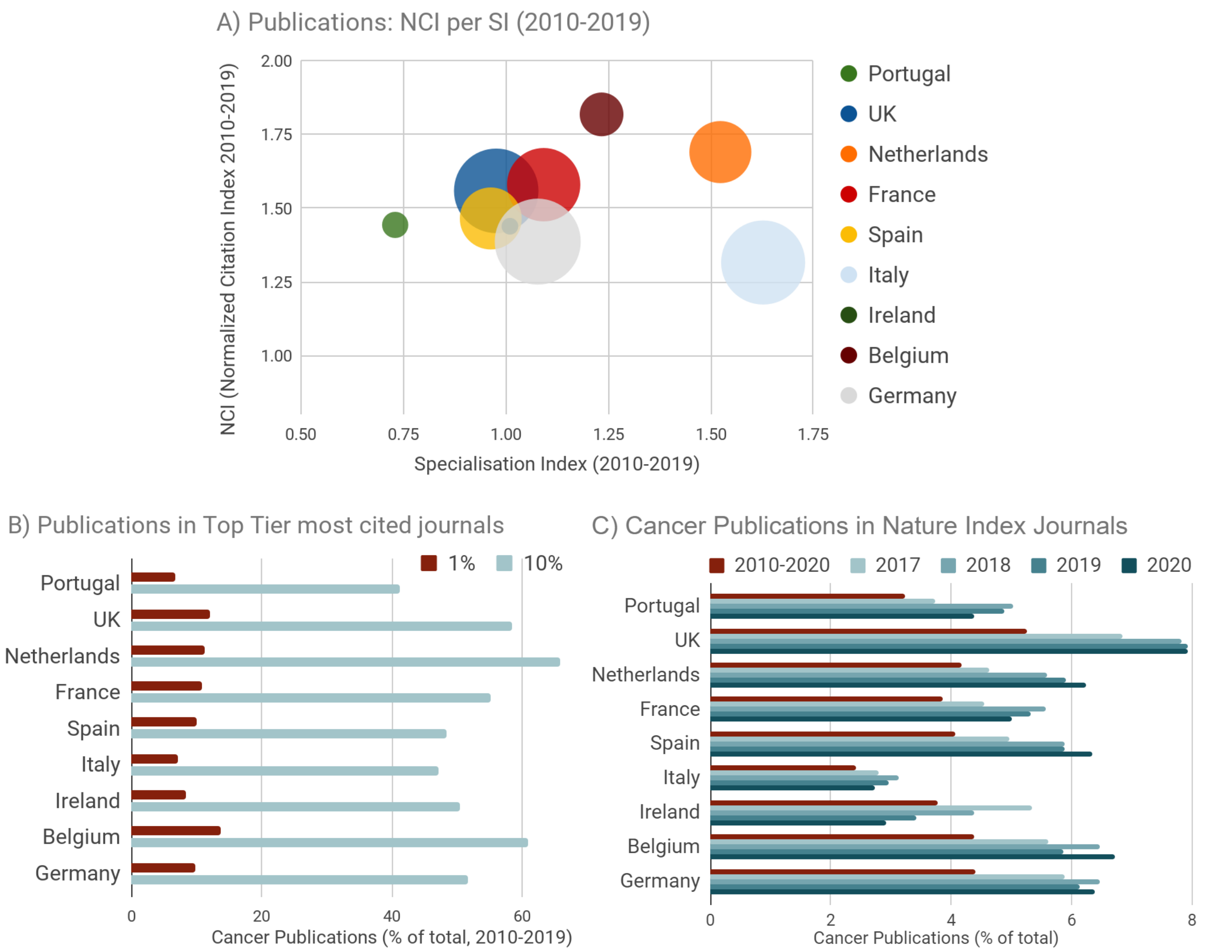


Fig. 3. Cancer publications quality in the panel of countries: A) *NCI of publications versus specialisation Index; dot size is relative to the total number of documents in cancer research for each country (2010-2019[31])* ***B)*** *percentage of publications in Top 10% and Top 1% journals (2010-2019);* ***C)*** *percentage of publications in Nature Index journals in the total number of documents in cancer research (2010-2020 and 2020).*

[30] The specialisation index is a ratio that compares the dedication to cancer research of a country with that of the world. Values above/below the unit correspond to a higher/lower degree of specialisation.
[31] Citations metrics, due to their nature, should only be analysed after a certain period of time. A 2 year buffer period between publication and analysis is ideal (see Galiani, Sebastian and Gálvez, Ramiro H., The Life Cycle of Scholarly Articles Across Fields of Research (May 2, 2017).

The biomedical research system in Portugal is in clear expansion, and its dedication to the domain increased from 20 to 30% in the last decade.

Cancer research has increased proportionally, but remains 10% of Portuguese biomedical research, while other biomedical topics also have center stage within the country.

Portugal is **not overly specialized in cancer research**.

Despite its **successful growth and significant citation metrics** in its published work, the Portuguese system is **still striving for scientific visibility and peer recognition in cancer research.**

Overall these are characteristic of an **emergent research system**, and mostly not specific to cancer research *per se*; but **Cancer research in Portugal has, certainly, the conditions to evolve**.

Bibliometric Production per Research "Type"

At present, biomedical research (as well as other domains) often struggle to transfer its findings and progress in, what is often called, the innovation path. This is a hypercomplex issue that has concentrated a lot of attention. It is certainly also true for cancer research[32], and something clearly highlighted as one of the main aims of the EU Mission on cancer[33].

Variables affecting the innovation process and leading to the famous "valley of death"[34] are often cited as being related to: **a)** funding allocation and evaluation (either specialized public funding, philanthropic funding or company-funded development); **b)** entrepreneurship skills/know-how and cultural differences and **c)** structural organization (either physical "spaces" or specialized services and "figures/roles" - as, for example, regulatory affairs).

In this study we also aimed to analyse the outputs of research across the innovation path; especially since it is perceived that the Portuguese ecosystem is in need of further development regarding biomedical knowledge transfer and innovation[35].

One way to address the cancer research continuum is to analyse the research outputs at different points in knowledge production, from its most fundamental biological understanding to its application, closer to the patient and society. As such, in this section, we will be comparing research mostly aligned with basic & translational research versus clinical research, epidemiology and public health, with a controlled vocabulary designed by SIRIS Academic (see Brief Methodological notes)[36].

[32]Cancer Clinical Trials: The Rear-View Mirror and the Crystal Ball
DOI: https://doi.org/10.1016/j.cell.2017.01.027
[33] Conquering cancer - Publications Office of the EU (europa.eu)
[34] The valley of death is where good biomedical ideas wither and die. It refers to the space between promising biomedical discoveries and its application for economy and society.
[35] Nationally, the creation of AICIB in 2018 was aimed to help bridge this gap in biomedical research in the near future. Other actors and initiatives are trying to tackle these issues, such as the Caixa Impulse programme, specifically designed and tailored to biomedical innovation projects needs (with a significant focus on capacitating researchers for conducting innovation).
[36] Although this division has its drawbacks, currently there are no other solutions that enable us to

Overall, applying the ad-hoc controlled vocabulary, we see a predominance of clinical and public health research published in all benchmark countries (around 65-75%) (Fig. 4). This may be explained by different patterns of publication and the time needed for a single publication. It might well be the case that, on average, individual dedication per paper is higher in more fundamental studies versus clinical studies, where the rate of publications per research is also, on average, lower. At present, Portugal appears to be slightly less dedicated to cancer research in the clinical setting and public health (< 60%) compared to the other systems studied (Fig. 4).

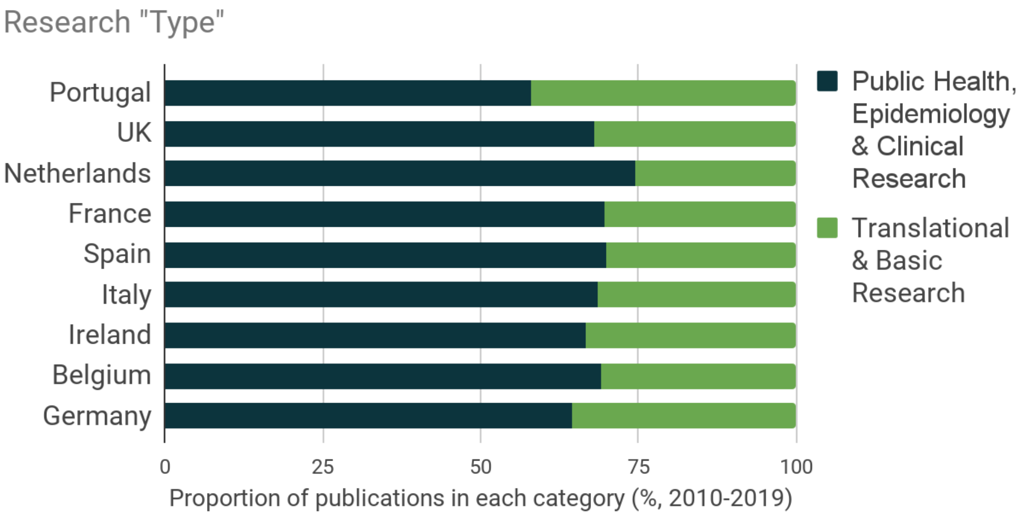


Fig. 4. Proportion of cancer publications by research type: *percentage of clinical/public health publications versus percentage of basic/translational publications, data as percentage of total publications from each country in cancer research.*

Portugal is currently growing at a fast pace compared to other benchmarks (see *Fig. 2B*); however, this result stems mainly from a more significant increase in fundamental research where, per capita, its production is now similar to Ireland, UK and Germany (*Fig. 5*). Throughout the period of 2010-2019, Portugal has increased its publications in clinical research 1.2 times (*Fig 5A*) whereas a more significant increase of roughly 1.9 times has been achieved in basic research (*Fig. 5B*). This trend, however, seems specific to Portugal and is not consistent with the benchmark countries. Overall, the Portuguese ecosystem in cancer research has been developing and growing significantly its basic and translational research.

In addition, a higher NCI is observed in clinical and public health research compared to translational and basic publications for all countries analysed, which is likely the result of different citation practices. The NCI of Portugal is quite significant in both research types

separate published research per any "research type"; as for the most part, research categories are not consensual or well-defined, and an analysis as this in cancer research, to our knowledge, has not been attempted before (specially at the publication levels, and not at journal level). Therefore, this exercise will function as a *proxy* to better understand Portuguese Cancer Research, accepting its many caveats. Along these lines, Microsoft Academic team also attempted a similar exercise, which rendered fairly similar results: a) with a similar realizations that only 2 big groups of research where potentially distinguishable: Basic vs Clinical Research, which in fact resembles ours with translational research classified in their category of basic research; while public health/ Epidemiology linked to the clinical research category; b) the share between the 2 groups is also around 1:2 (a global total of 24K publications in the basic research and 49K in clinical research).

(*Fig. 6*)[37], but more so on more fundamental research, where it exceeds systems like Germany, France and Spain. This confirms the pattern of **a greater evolution of the basic & translational research in Portugal (both in size and excellence)**.

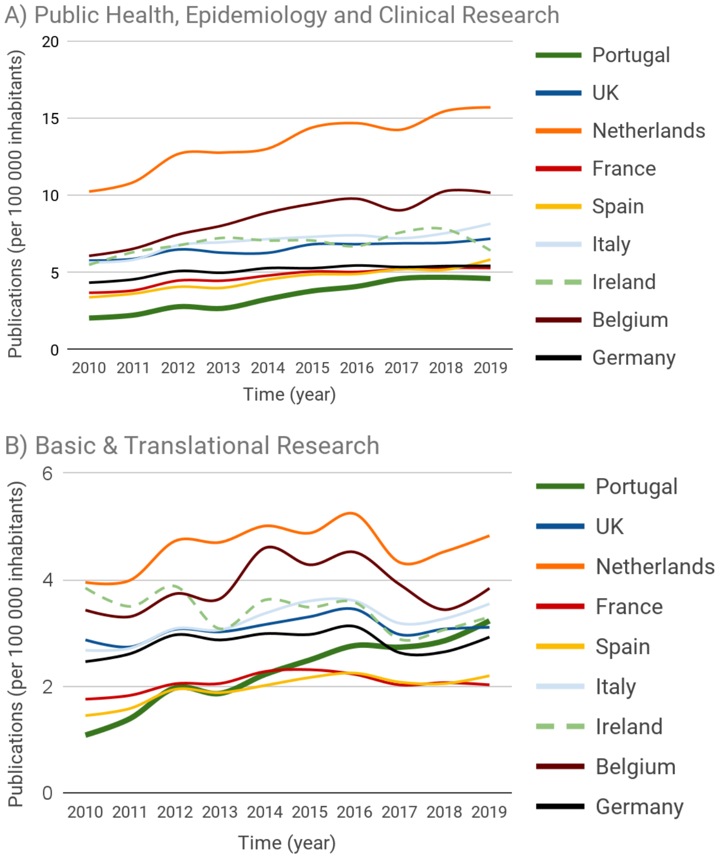


Fig. 5. Cancer publications in the panel of countries (2010-2019): A) *ratio of public health, epidemiology and clinical cancer research publications and* ***B)*** *of basic & translational cancer publications per 100 000 inhabitants.*

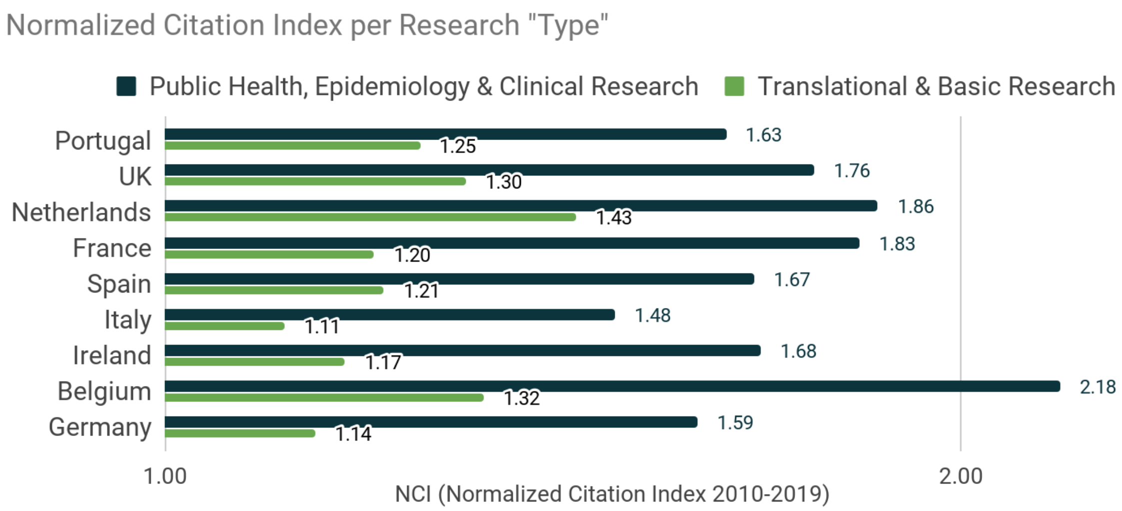


Fig. 6. Cancer research quality in the panel of countries (2010-2019): *Normalized Citation Index (NCI) of public health, epidemiology and clinical cancer research publications versus basic & translational publications (logarithmic scale); (NCI=1 means that the corpus of publications are cited, on average, the same as the baseline - worldwide publications; a NCI around 2 would mean that they are cited, on average, twice the baseline (world).*

[37] In general, we find that in smaller systems it is often easy to reach and maintain high levels of citations, due to the dilution effect which comes from being a large ecosystem.

Scientific Production by Cancer Site

Although a significant part of cancer research is dedicated to the understanding of cellular phenomena that *lead to* or *protect from* cell transformation (tumorigenesis), special attention is given to cancers per type or site, ever since it was understood that cancer is not *one* disease. In this section we show how the Portuguese ecosystem distributes its interests, and if that is aligned with the clinical relevance associated with such tumors.

Despite slight variations, for **Portugal and all the benchmarks analysed, research interests into specific cancers are globally consistent with current EU and global research patterns** (*Table 2*). For example, **Portugal shows a proportionally higher number of publications in stomach, thyroid and urinary bladder cancers.** For stomach and thyroid cancers, this is associated with the country's high incidence of these tumors[38] (see below, *Table 4*) and the outstanding seminal work of renowned researchers[39].

Dedication to specific Neoplasms* (% of national cancer publications)

MeSH Neoplasms*	**World**	**EU27+UK**	**Portugal**	**UK**	**France**	**Spain**	**Italy**	**Ireland**	**Belgium**	**Germany**	**Netherlands**
Breast Neoplasms	10.1	10.5	12.4	12.1	10.2	10.4	9.3	16.3	12.5	8.3	11.8
Lung Neoplasms	7.5	6.3	5.0	6.1	7.1	7.4	6.6	6.2	7.1	6.0	6.9
Prostatic Neoplasms	4.6	5.2	4.8	5.6	4.3	4.4	4.9	6.9	5.2	5.8	5.8
Colorectal Neoplasms	4.4	5.0	4.0	5.8	4.1	5.6	4.5	6.8	4.7	4.3	6.5
Liver Neoplasms	5.3	4.0	3.0	3.0	5.0	3.8	5.0	1.9	3.7	4.7	3.3
Skin Neoplasms	3.1	3.9	4.2	3.3	4.2	5.3	3.9	3.1	3.2	4.1	3.1
Brain Neoplasms	3.5	3.7	3.2	2.9	4.2	2.8	3.3	2.2	3.0	5.2	3.3
Pancreatic Neoplasms	2.8	2.6	1.5	2.3	2.6	2.5	2.9	1.8	1.7	3.5	2.5
Ovarian Neoplasms	2.5	2.5	1.7	2.9	2.4	2.2	2.5	2.4	3.5	2.2	2.2
Carcinoma, Non-Small-Cell Lung	2.9	2.4	1.5	2.3	2.9	3.2	2.8	2.3	3.4	2.4	3.0
Head and Neck Neoplasms	1.9	2.2	1.3	2.3	1.8	1.8	1.7	2.3	2.7	2.5	3.1
Carcinoma, Hepatocellular	3.5	2.1	1.3	1.4	2.7	2.3	3.2	0.9	1.7	2.4	1.1
Kidney Neoplasms	1.9	2.0	1.7	1.7	2.7	2.2	2.4	1.6	1.8	2.4	1.7
Uterine Cervical Neoplasms	2.1	1.9	2.4	2.0	1.9	1.5	1.4	2.2	2.8	1.1	2.4
Bone Neoplasms	1.7	1.7	1.7	1.7	1.7	1.5	1.9	1.2	2.0	1.7	1.8
Colonic Neoplasms	1.8	1.7	2.0	1.3	1.4	2.0	1.6	2.3	1.3	1.5	1.6
Stomach Neoplasms	3.0	1.6	4.8	1.4	1.1	1.4	1.9	1.3	0.9	1.7	1.4
Urinary Bladder Neoplasms	1.4	1.6	2.4	1.4	1.4	2.2	1.7	0.8	1.3	2.0	1.7
Carcinoma, Renal Cell	1.4	1.5	1.3	1.2	1.9	1.6	1.7	0.8	1.2	1.8	1.2
Thyroid Neoplasms	1.5	1.4	2.3	0.7	1.2	1.3	2.4	0.9	0.7	1.1	0.8
Esophageal Neoplasms	1.7	1.3	0.9	2.1	0.9	0.7	0.7	3.4	1.1	1.3	2.5
Rectal Neoplasms	1.1	1.3	1.2	1.3	0.9	1.3	1.1	2.2	1.1	0.9	1.9
Endometrial Neoplasms	0.9	0.9	0.7	0.9	0.7	0.9	1.0	0.5	0.9	0.5	0.9
NA	27.3	29.6	29.7	31.4	31.9	29.8	29.2	28.1	32.8	30.4	27.9

**according to the [NIH-NLM MeSH terms] attached to each publication*

Table 2. Cancer publications by Neoplasm (2010-2020): *percentage of publications of the region or country, classified by their MeSH terms that are in the C04-Neoplams branch of the MeSH Taxonomy (top 23 cancers in EU27+UK); NA relates to publications without neoplasm (and rather vague terms like Neoplasm staging, see Annex B).*

[38] https://pubmed.ncbi.nlm.nih.gov/26186469/ , 620-portugal-fact-sheets.pdf (iarc.fr)

[39] Manuel Sobrinho-Simões (thepathologist.com); Ipatimup/Sobrinho Simões; Fatima Carneiro (thepathologist.com); Ipatimup/Fatima Carneiro

In another interesting example, Ireland shows a higher dedication to breast cancer research. This could be correlated with the country's overall high cancer rates - third highest in the world in 2018 - and where breast cancer is a major concern[40]. Such prevalence may have pushed both national research and resources to be mobilized into breast cancer research.

In addition, if separated by research "type", Portuguese cancer basic and translational research presents a lower diversity (a smaller number of tumors aggregate a higher share of basic and translational research when compared to clinical and public health investigation; *Table 3*). This fact is also observed in the EU27+UK, and is most likely linked to the use and availability of research models (not all neoplasms are easy to model in laboratory conditions), and some neoplasms are indeed more frequently used for early testing in the translational research setting (e.g. breast neoplasms)[41].

	Dedication to specific Neoplasms* (% of cancer publications)			
	Portugal		EU27+UK	
MeSH Neoplasms*	**Public Health, Epidemiology & Clinical Research**	**Basic & Translational Research**	**Public Health, Epidemiology & Clinical Research**	**Basic & Translational Research**
Breast Neoplasms	7.81	4.63	7.46	3.05
Lung Neoplasms	3.46	1.49	4.63	1.67
Stomach Neoplasms	3.24	1.56	1.26	0.31
Prostatic Neoplasms	2.96	1.79	3.83	1.38
Skin Neoplasms	3.13	1.03	2.99	0.95
Colorectal Neoplasms	2.42	1.56	3.79	1.25
Brain Neoplasms	1.81	1.38	2.47	1.24
Liver Neoplasms	2.19	0.77	2.85	1.18
Urinary Bladder Neoplasms	1.12	1.31	1.22	0.34
Uterine Cervical Neoplasms	1.97	0.39	1.60	0.29
Thyroid Neoplasms	1.76	0.55	1.10	0.31
Colonic Neoplasms	0.76	1.28	0.85	0.82
Ovarian Neoplasms	1.02	0.67	1.72	0.75
Bone Neoplasms	1.03	0.65	1.25	0.48
Kidney Neoplasms	1.29	0.38	1.66	0.38
Pancreatic Neoplasms	0.94	0.59	1.73	0.88
Carcinoma, Non-Small-Cell Lung	0.96	0.52	1.85	0.59
Carcinoma, Hepatocellular	0.79	0.53	1.32	0.78
Head and Neck Neoplasms	1.06	0.24	1.80	0.44
Carcinoma, Renal Cell	0.94	0.33	1.17	0.29
Rectal Neoplasms	1.14	0.06	1.23	0.06
Bile Duct Neoplasms	0.96	0.09	0.34	0.08
Heart Neoplasms	0.91	0.08	0.28	0.03

according to the NIH-NLM MeSH terms attached to each publication

Table 3. Cancer publications by neoplasm, and by research "type" (2010-2020): *percentage of publications, classified by their MeSH terms that are "neoplasms" as percentage of the total cancer research in Portugal and in EU27+UK (top 23 cancers in Portugal's cancer publications).*

[40] Cancer rates here are third highest in the world, warns WHO; Facts&Figures – Europa Donna Ireland
[41] In addition, when classifying publications per cancer group (according to MeSH Taxonomy - *(Neoplasms by Site [C04.588])*, we can further appreciate that Digestive System tumors concentrate research (i.e., stomach, colorectal, liver, etc); followed by Urogenital and Breast cancers (see Annex A).

Cancer research is overall aligned with the clinical impact of specific neoplasms (in terms of mortality and incidence) identified in the GLOBOCAN2020 study[42]: e.g. there is a clear dedication to major "killers"[43], with some regional specificities[44].

In *Table 4* we have aligned GLOBOCAN / WHO categories of cancer sites with those present in the publications (from MeSH Taxonomy), in order to understand if Portuguese cancer research is somewhat aligned with the societal needs expressed by the clinical variables. **Overall, Portuguese research is aligned with the incidence of the different cancers in Portugal (as in the case of the Stomach and Thyroid neoplasms, much lower in other European countries)**. Of note, the significant dedication to nervous system tumors which is not fully aligned with its incidence). However, the higher research dedication to this and other specific neoplasms, appears to be in detriment of the dedication to some major cancers like lung, colorectal and pancreatic (*Table 4*).

Tumour by Cancer site*	Incidence rate (%)	Mortality rate (%)	Portugal's Cancer Publications (%)
Breast	11.6	6.2	13.0
Prostate	11.2	6.4	5.2
Lung	9	15.9	5.1
Colon	8.9	9.9	2.2
Rectum	8.2	4.3	1.7
Stomach	4.9	7.7	4.2
Bladder	4.3	4	2.3
Non-Hodgkin lymphoma	3.5	3.2	1.8
Pancreas	3	5.9	1.8
Thyroid	2.8	0.3	2.2
Liver	2.6	5	3.2
Leukaemia	2.6	3.2	3.5
Corpus uteri	2	1.2	0.7
Kidney	2	1.7	2.1
Brain, Central Nervous System	1.8	3.1	5.1
Lip, oral cavity	1.8	1.3	1.6
Melanoma of skin	1.8	1.0	3.0
Multiple myeloma	1.5	2.1	1.2
Cervix uteri	1.4	1.3	2.3
Oesophagus	1.1	1.9	0.9
Ovary	0.9	1.4	2.0

**** as categorized by WHO and the Globocan Study***

Table 4. Clinical variables and cancer research outputs aligned: *incidence and mortality in Portugal (2020)[45]; publications (% of national cancer research published work, 2015-2020) per cancer site.*

[42] GLOBOCAN 2020 is an online database providing global cancer statistics and estimates of incidence and mortality in 185 countries for 36 types of cancer. The data is part of IARC's Global Cancer Observatory, WHO.
[43] Breast, Lung, Prostate, Colorectal, Stomach, Liver, Uterus, Ovary.
[44] Of note is the high share of liver neoplasms in global research dedication (*Table 2*). Liver cancer is highly associated with chronic Hepatitis and highly prevalent in Asia (particularly due to HBV), although also worrisome in Europe due to its lethality. In accordance, there is a particular specialisation of Asian systems like China and Japan to Liver cancers; this would explain why the research on liver cancer is globally prominent, but not consistent within the panel of EU countries.
[45] https://gco.iarc.fr/today/data/factsheets/populations/620-portugal-fact-sheets.pdf

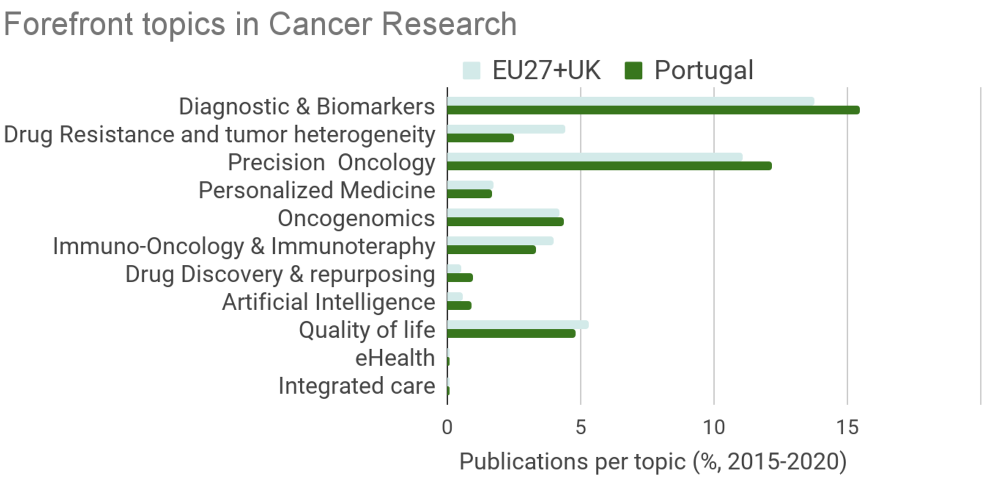
Portugal's cancer research **has grown significantly, but more so in basic and translational aspects** where it achieves significant scientific impact.

Overall, the system is **slightly less dedicated to clinical and public health research** compared to the EU countries used as comparators, although still gathering good citation metrics.

Cancer research in Portugal has a pattern of **dedication to specific neoplasms consistent with the European region**, while also showing an **alignment with its epidemiological specificities**.

Box 1. Portugal´s alignment with forefront topics

We have seen that Portugal is growing in terms of biomedical research production and that its impact (NCI) in cancer research is also high. However, does the research align to what are considered forefront topics of research in the field?[46]
Both the European Commission and experts in the field highlight a number of research topics as being at the forefront of the research that can advance diagnostic and treatment in cancer[47]. When we analysed how Portugal fares in comparison to other European countries in the dedication to these topics[48], we can see that it follows a **very similar profile**, although apparently less in drug resistance and tumor heterogeneity



Cancer Publications per forefront Topics: percentage of all cancer research publications in Portugal and EU27+UK (2015-2020), per topics selected by SIRIS Academic (see Methodology)

[46] This would indicate an alignment to funding schemes and to general global trends but in itself is not a measure of quality, nor assumes so.
[47] EC´s areas of interest and investment: molecular mechanisms of cancer; precision and personalised medicine (cancer evolution; host immune activation; tumour-host interaction); representation of a diverse population in clinical trials; genomics; diagnostic technologies (radiomics, liquid biopsies); big data and machine learning analysis; integrated diagnostics (liquid biopsies and clinical biomarkers).
[48] SIRIS Academic has generated a simple list of keywords that relate to these topics and conducted a semantic search on the set of cancer publications identified for this study (for more details see Methodology). It is not intended to be an exhaustive depiction of these topics.

Impact of Published Work in Innovation and Clinical Practice

Reliable and comprehensive innovation and technology transfer output metrics are hard to obtain. This stems from a myriad of factors: lack of quality open data national registries; the fact that private interests are not always aligned with open data policies[49]; multiple paths of innovation (products, services, protocols, processes) and different access to technology transfer support, amongst others.

Although a very indirect way to detect the impact of biomedical research, the degree of citation of publications in patents and clinical practice guidelines (a specific type of publications) can give some indication on the degree to which said research has been influential or even pivotal for innovation and the establishment/change of clinical protocols[50].

For this report, we did not study in-depth patent registries. There are several hindrances to the use of patent information as an "innovation" metric:

a. Often incentives to patent are linked to researchers prestige (and CV) and not necessarily a drive to transfer knowledge into a marketable product or service; on the other hand, they are often linked to the financial capacity to accommodate for patenting fees and/or the existence of a highly involved private sector.
b. European patents like the European Patent Office EPO are necessary for benchmark comparison, but they offer low coverage compared to national registries.
c. Patent documentation is written in a very particular manner, often intentionally avoiding discussing the actual application of the item under patent. Therefore, automatic semantic analysis techniques do not render minimum standards of quality when applied to patent documentation.

Accepting these constraints, in this study we aimed at understanding the innovation outputs that arise from cancer research, where some first insights are already discussed in the Results section Bibliometric Production per Research Type. In this section we would like to explore other data sources that can provide additional information on this matter.

Patents and Citations in Patents

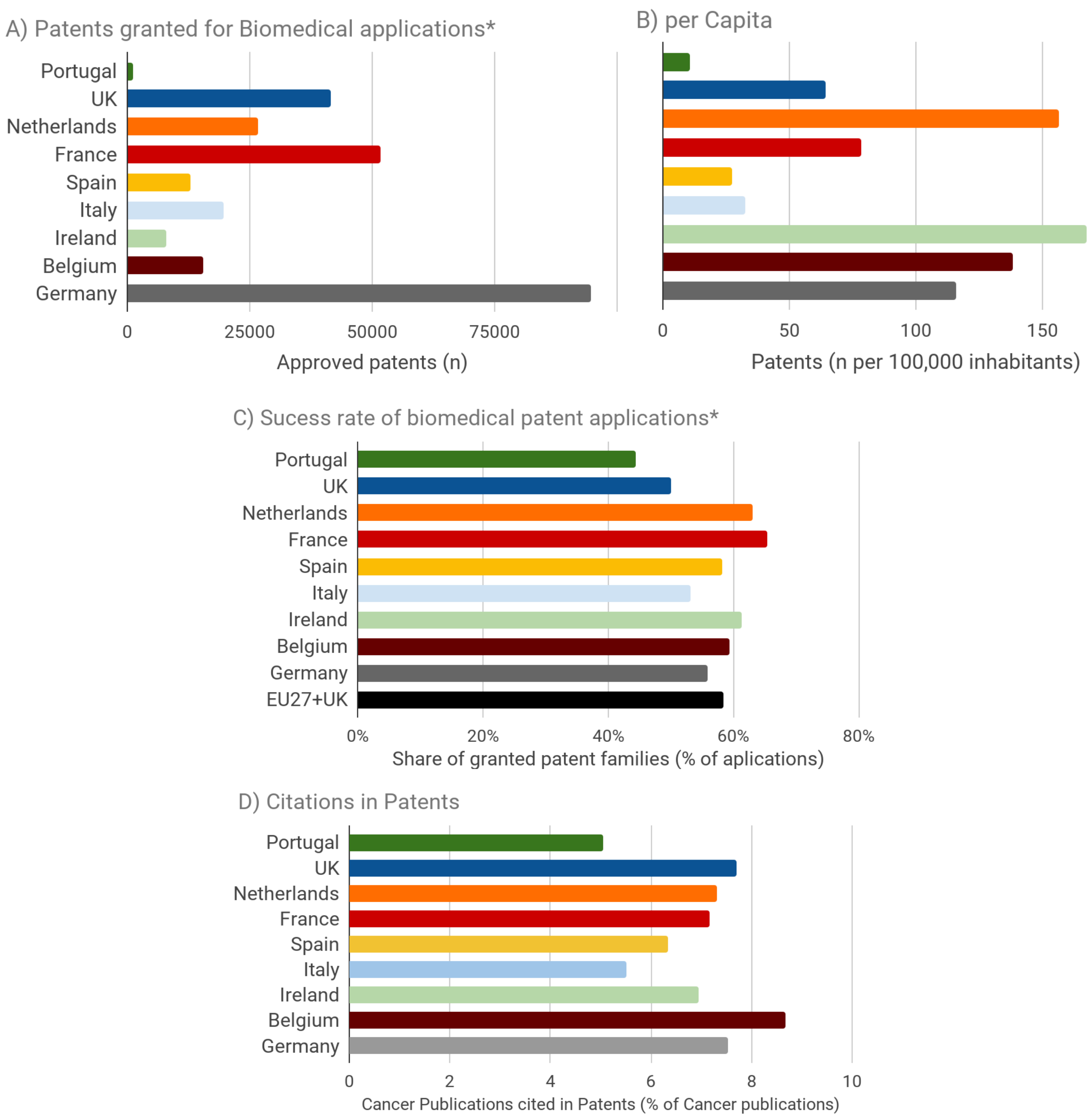
As commented above, patenting incentives, motivations and capacity may vary greatly across ecosystems. However, patents are currently still often used as an innovation indicator. In addition, when looking to identify patents that relate to a certain domain, the way these documents are written (very different from a publication or a research project), makes them virtually impossible to classify semi-automatically[51], which is a major hindrance in large-scale specialized analysis.

Nevertheless, by recurring to patent databases, we can have a sense of the overall dynamism and evolution of the national patent landscape. We can see that **the number of Portuguese accepted patents is significantly lower than the panel of benchmark countries**, in all the

[49] E.g there is a tendency to not openly disclose products' potential application/usage details on patent submission documentation for fear of leaking relevant information to competitors.
[50] We have obtained patent and clinical practice guideline citation data from PlumX Metrics, by Plum Analytics, which is integrated in Scopus. Patent citations are sourced from different patent offices (i.e. the EPO, IPO, JPO, USPTO, and WIPO); while clinical practice guideline citations are sourced from PubMed and NICE.
[51] In truth, even manually it is often very difficult to identify the application, or restrict it to a certain disease or group of diseases.

biomedical categories[52] (*Fig. 7A-B*). Germany and France are very strong systems in patentable research, although, per capita, we can also see that Ireland, Netherlands and Belgium are the most productive (*Fig. 7B*). Furthermore, the success rate of application by patent family is also lower than the rest of the panel of countries (*Fig. 7C*); however this difference is less drastic, highlighting a specially low number of applications in Portugal in comparison to the benchmark countries.

On the other hand, regarding the citation of cancer research publications in patents, it can be observed that both Portugal and Italy fall behind most countries (*Fig. 7D*). This does not necessarily correlate with a more or less active national innovation system; however, **it does not support the idea of a thriving environment, in either basic or applied research, that feeds patentable outcomes.**



[52] Patents analysed according to the WIPO technology concordance table: 1.Medical Technology, 2. Pharmaceuticals, 3. Biotechnology and 14. Analysis of Biological Materials; excluding CPC Category A61D Veterinary. Nevertheless we cannot exclude that relevant biomedical patents have been filed under other Fields and/or Categories. Source: The Lens - Free & Open Patent and Scholarly Search.

***Fig. 7. Cancer patents and publications cited in patents (2010-2019)**: **A)** number of granted patents and **B)** number of granted patents per capita, *in the technology field: 1.Medical Technology, 2.Pharmaceuticals, 3.Biotechnology and 14.Analysis of Biological Materials; excluding Category A61D Veterinary; **C)**. D) percentage of cancer publications from each country that are cited in patents according to PlumX (2010-2019).*

Citation in Clinical Practice Guidelines

When looking for cancer research publications mentioned in clinical practice guidelines (CPGs), Netherlands and Belgium have the highest impact; whereas **Portugal is the national system with the lowest influence** (*Fig. 8*). The share of cancer publications cited in CPGs is lower than other countries (*Fig. 8A*), including when assessing only clinical publications (*Fig. 8B*)

Here we might be encountering different reasons to why: a) a lower dedication to clinical research, more often cited in CPGs (already discussed in section A. Bibliometric Production by Research "type"); b) low "visibility" in the spheres where CPGs are often redacted (for example, Medical Oncology associations); c) low motivation from health professionals to publish clinical innovations.

There is obvious room to grow in this direction, regardless of which variable weighs heavier, as Portuguese cancer research is behind in impacting clinical routine.

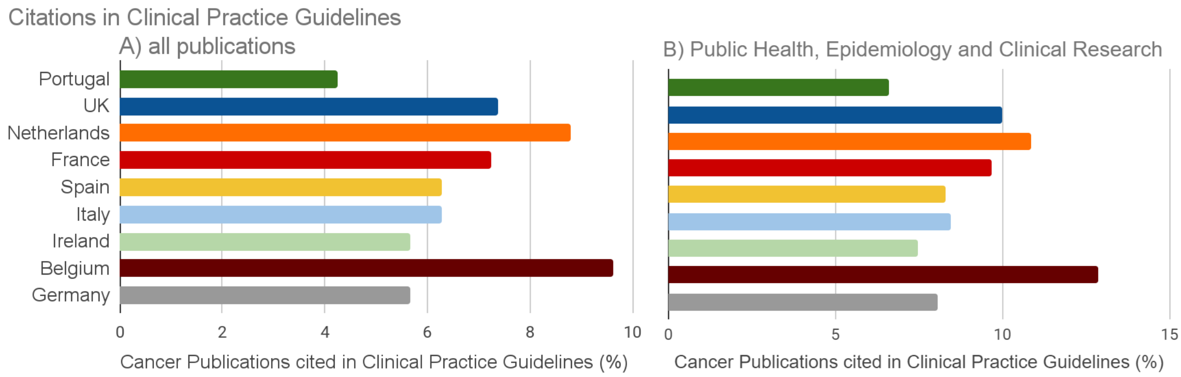


***Fig. 8. Cancer publications in clinical practice guidelines** (percentage of cancer publications from each country that are cited in clinical practice guidelines according to PlumX; 2010-2019): **A)** in all cancer publications and in **B)** cancer publications identified as clinical, Epidemiological public health (research "type").*

The number of biomedical patent applications and acceptance at the European registry from Portuguese organizations and individuals **is still rather low**. In addition, the published research is **not cited in patents and clinical guidelines** as often as in other systems.

Although the available information does not allow a fully robust and complete view, it appears as if **Portuguese cancer research is currently not imprinting substantially in innovation and clinical practice**.

In agreement, several recent initiatives are focusing on **further developing the country's capacity to do knowledge transfer in the biomedical domain and improve patient care**.

B. Clinical Trials

Clinical trials, whether they are the culmination of several years of development steps or just the refinement of clinical protocols, are critical to advance patient care. In this section we will mainly explore how Portugal clinical trials capacity compares to other systems using an international repository, as well as, the national registry data provided (which allows for complementary higher level of quality on some of the variables analysed)[53].

Clinical trials capacity varies greatly among countries, and it is linked with, among other factors: the quality of biomedical research; the connection with sponsors; quality contract research organizations (CROs) and clinical trials supporting structures; access to participants and patient pools; the efficiency of a given healthcare system to accommodate clinical studies; the quality of public policies, including at the regulatory and administrative level, and funding in support to clinical trials; and the level of citizen awareness and engagement. Therefore, each ecosystem's capacity is the result of a combination of strengths and pitfalls on all of these conditions (and potentially more).

Volume and Growth

Clinical trials, similar to other types of research, are mostly published in peer-reviewed journals. Some of the most cited biomedical journals are highly dedicated to it (e.g. Lancet, JAMA and New England Journal of Medicine[54]; amongst other clinical and epidemiological journals). Therefore, we first analysed the amount of "published clinical trials" in the Pubmed database.

Even if we can not confirm the level of coverage of this information, when looking at the proportion of publications in Pubmed that have associated a clinical trial registry identification number[55], Portugal has the lowest share of published clinical trials (2.5%), far from roughly 8% in France and Spain (data from 2020) (*Fig. 9*). The French and the Spanish ecosystem are both quite developed in respect to clinical trials management, albeit presenting very different funding models (the French system has had important governmental support since the 90's for investigator initiated trials for example[56] as well as

---

[53] International Registry: Clinicaltrials.gov
National Registry: INFARMED.
INFARMED information was provided explicitly for this study and included trials successfully authorized between 2011- May 2021, except phase 1 trials. The variables provided included exclusively: Title, promoter type, EudraCT ID, site, investigator type, number of sites per trial.
We have compared data from INFARMED (without Phase 1 trials) / EU Clinical Trials Register (EudraCT) with NIH - NLM clinical trials repository (Clinicaltrials.gov). The USA registry has comparable coverage (381 registered in INFARMED/EudraCT, and 415 in Ct.gov between 2010 and 2020, with high overlap), while providing extra resources for semantic analysis (e.g. associated MeSH terms), and was thus the registry of choice for this analysis.
[54] According to Microsoft Academic: Clinical research | Topic | Microsoft Academic
[55] Pubmed has the "databank resource" which contains, among others, registry numbers for clinicaltrials.gov, EudraCT and ISRCTN clinical trials repositories. Although, we assume it does not have full coverage, it can be a first indication.
[56] Programme hospitalier de recherche clinique - PHRC, since 1992.

the implementation of fast track programmes for trial authorization in 2018[57], whereas the Spanish system relies more on its close relationship with the private sector). The Belgium system also has an excellent reputation and standing in clinical research (due to a mixture of dedicated sites, trained health professionals at research sites and reasonably fast and affordable regulatory processes), and is the county with the higher number of trials ID registries in their publications[58].

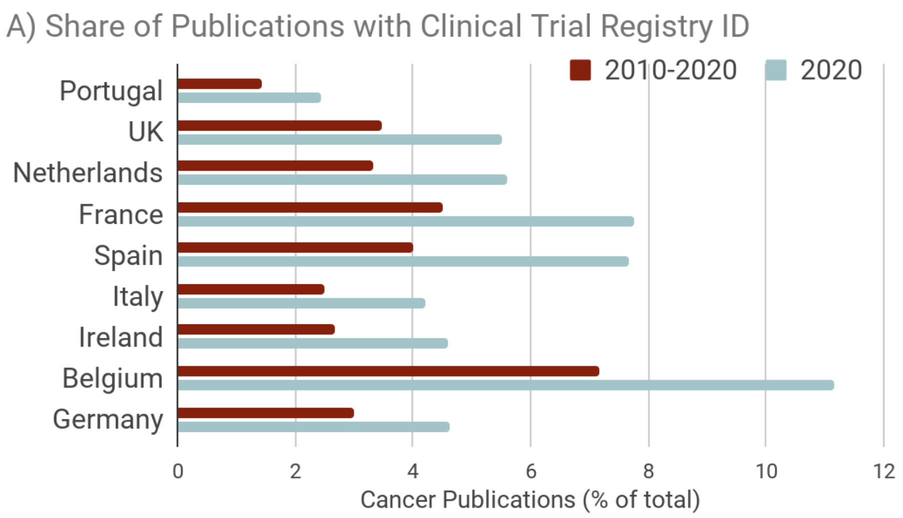


Fig. 9. Cancer trial publications: *ie, cancer publications with a clinical trial registry number according to Pubmed (percentage of cancer publications from each country that have a clinical trial registry ID) (2010-2020).*

When looking at all clinical trials registered in the NIH-NLM registry (*clinicaltrials.gov*), Portugal is positioned alongside Ireland with the lowest number, as compared to all benchmarks (*Fig. 10A*). Furthermore, per capita, the ratio of clinical trials is still low, but at the same level as Germany (*Fig. 10B*), and with a significant positive evolution (*an increase of >60% in the number of trials conducted from 2010 to 2019*). Of note, the case of France and Spain's growth (especially Spain, a system which underwent a profound transformation in clinical trials management capacity[59], and doubled its number of clinical trials between 2010 and 2019).

Cancer trials have been increasing in the last decade (nearly 30% in the sum of countries analysed here). In Portugal, despite a discrete number of cancer clinical trials (*Fig. 10C-D*), those have increased at an impressive rate (just below Spain), doubling in number between 2010 and 2019 (*Fig. 10E*). Notably enough, cancer trials represent around 30% of all trials in the panel of countries, including Portugal (*Fig. 10F*). At present, **despite the observed lower numbers, the significant growth of cancer clinical trials in Portugal and a consistent share of trials, appears indicative of a growing capacity, particularly in cancer.**

[57] Actualité - L'ANSM met en place un dispositif accéléré d'autorisation d'essais cliniques (Fast Track)
[58] Belgium is European leader in clinical trials | Invest In Flanders (flandersinvestmentandtrade.com); Clinical-Trials-in-Belgium.pdf (cromsource.com); Cancertrials.be
[59] Nine Reasons to Conduct an Oncology Clinical Trial in Spain - Sofpromed; Spain Accelerates Clinical Trial Timelines - Sofpromed

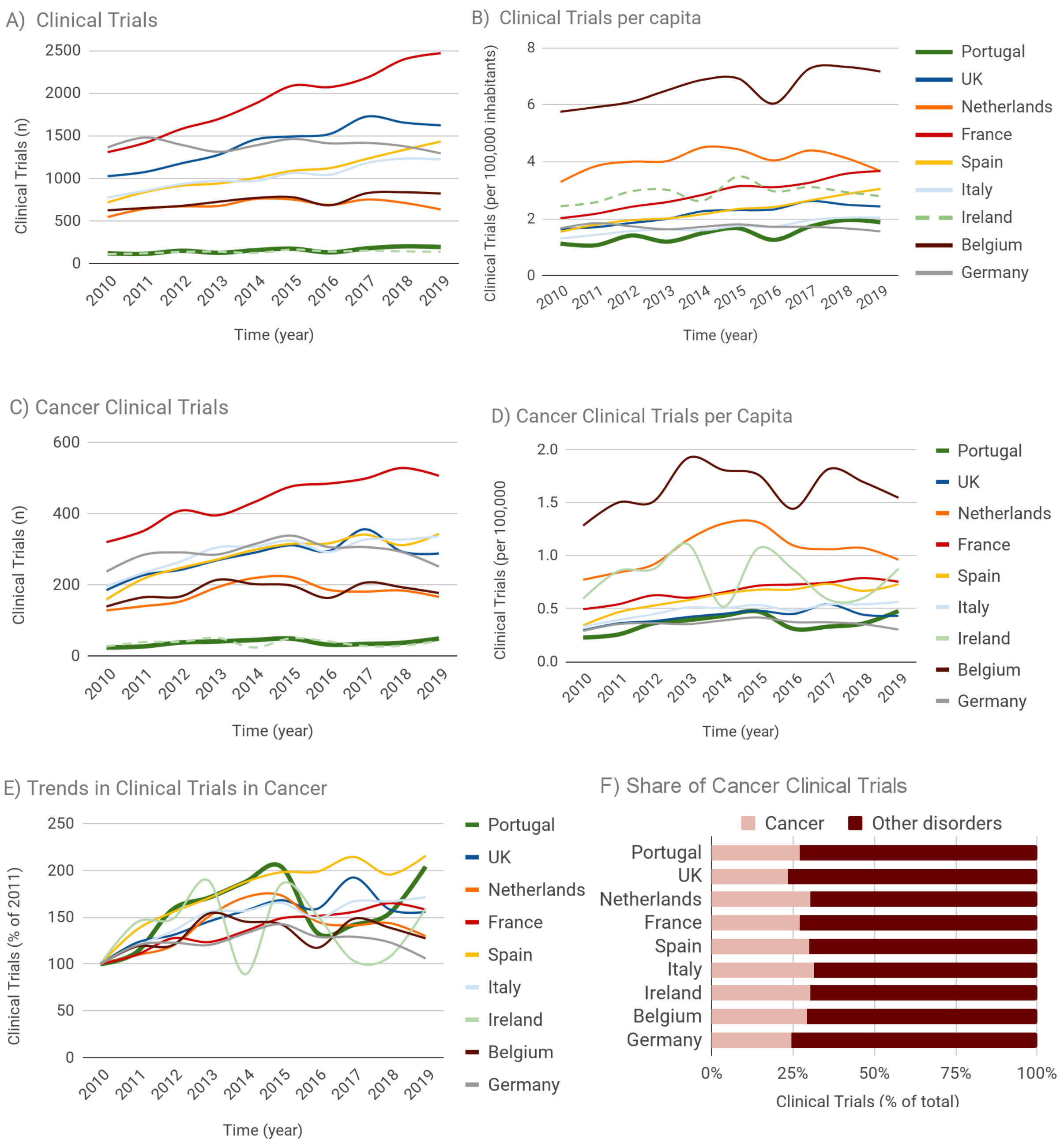


Fig. 10. Clinical trials registered in clinicatrials.gov (started between 2010-2029): *A) total number of clinical trials and B) number of clinical trials per 100,000 inhabitants; C) number of cancer clinical trials and D) number of cancer clinical trials per 100,000 inhabitants; E) growth in the number of cancer trials as percentage of the number of trials in each country versus 2010; F) share of cancer clinical trials (trials in cancer as percentage of total; 2010-2020).*

Sponsorship

In many developed countries, sponsorship of trials tilts towards private sponsorship (often around 60-70%), except in countries that dedicate relevant funding programs to independent (investigator-initiated) clinical research (such as France) (*Fig. 11A*). Furthermore, due to the globalization of clinical research and the multi-site property of many large trials (mainly industry-sponsored since they mobilize huge financial resources), in smaller slightly underperforming systems, industry sponsorship becomes more relevant. This seems to be the case for Portugal (76.7% of trials having industry sponsorship in *ct.gov*; and 85% - 90% in *ct.gov* and INFARMED registry for trials excluding phase one). Interestingly, cancer clinical trials in Portugal have increased in both sponsorship modalities since 2012 (*Fig. 11B*).

In addition, when further analysing the distribution of trials into interventional and observational studies, we find that the panel of selected countries are highly dedicated to interventional research (over 80% of clinical trials concluded are interventional) (*Fig. 11C*). This fact can indicate a registry bias against observational studies, which may constitute a higher proportion in many countries (since there are less regulatory and publishing requirements on this type of trials in the *ct.gov* registry). Noteworthy, non-industry sponsored trials have a higher share of non-interventional trials (Observational, patient registry, expanded access) (*Fig. 11D vs E*), which highlights the difference in interests between private and public entities.

The small number of Portugal's clinical trials rely on a high rate of private sponsorship, and in alignment, mainly concern interventional practices.

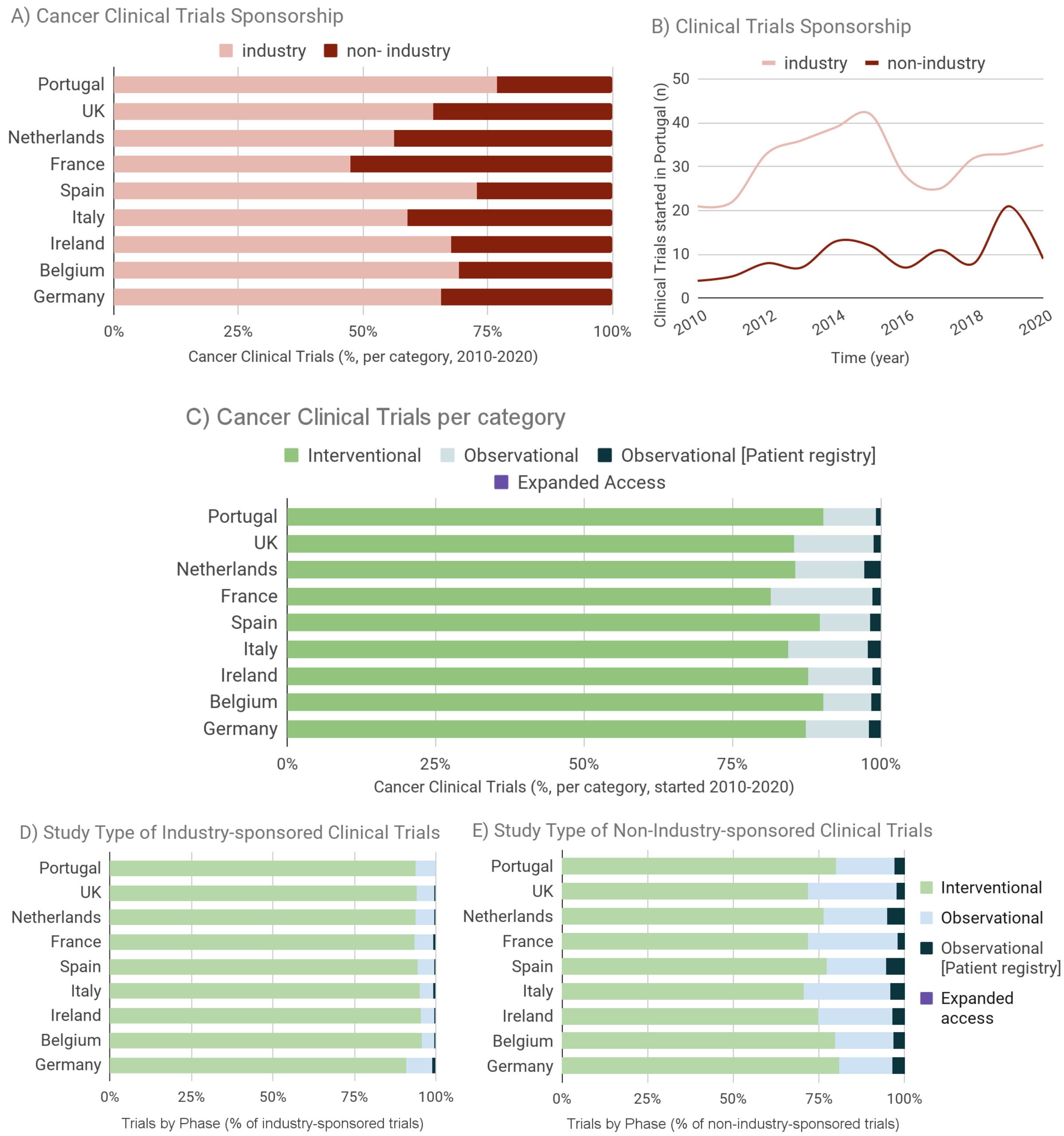


Fig. 11. Sponsorship and study type of clinical trials in cancer registered in clinicaltrials.gov (started between 2010-2020): *A) sponsor type, percentage of trials sponsored by private (industry) and non-industry institutions (NIH, Other); B) evolution on the number of trials per type of sponsor in trials started each year in which a portuguese facility participates.; C) study type, percentage of trials in each category; D) study type in percentage of trials in industry sponsored trials, and E) study type in percentage of trials in non-industry sponsored trials.*

Trial Phases

When looking at trials by stage (e.g. Phases 1 to 4), we can observe that **Portugal has a clear predominance of Phase 3 trials** (*Fig. 12A*) (63% both in clinical trials.gov and INFARMED registry) which is most likely connected to its dependency on the industry sponsorship (*Fig. 11A*). Importantly, **early clinical trials phases are often less represented in non-mandatory registries. In *ct.gov* phase 1 trials represent around 3-4% of Portuguese trials, whilst in INFARMED around 11%**. In spite of this bias, in the non-industry sponsored trials registered in *ct.gov*, there is a very small share of Phase 1 trials, which indicates a gap in the investigator-initiated trials of non-private sponsorship compared to the benchmark countries.

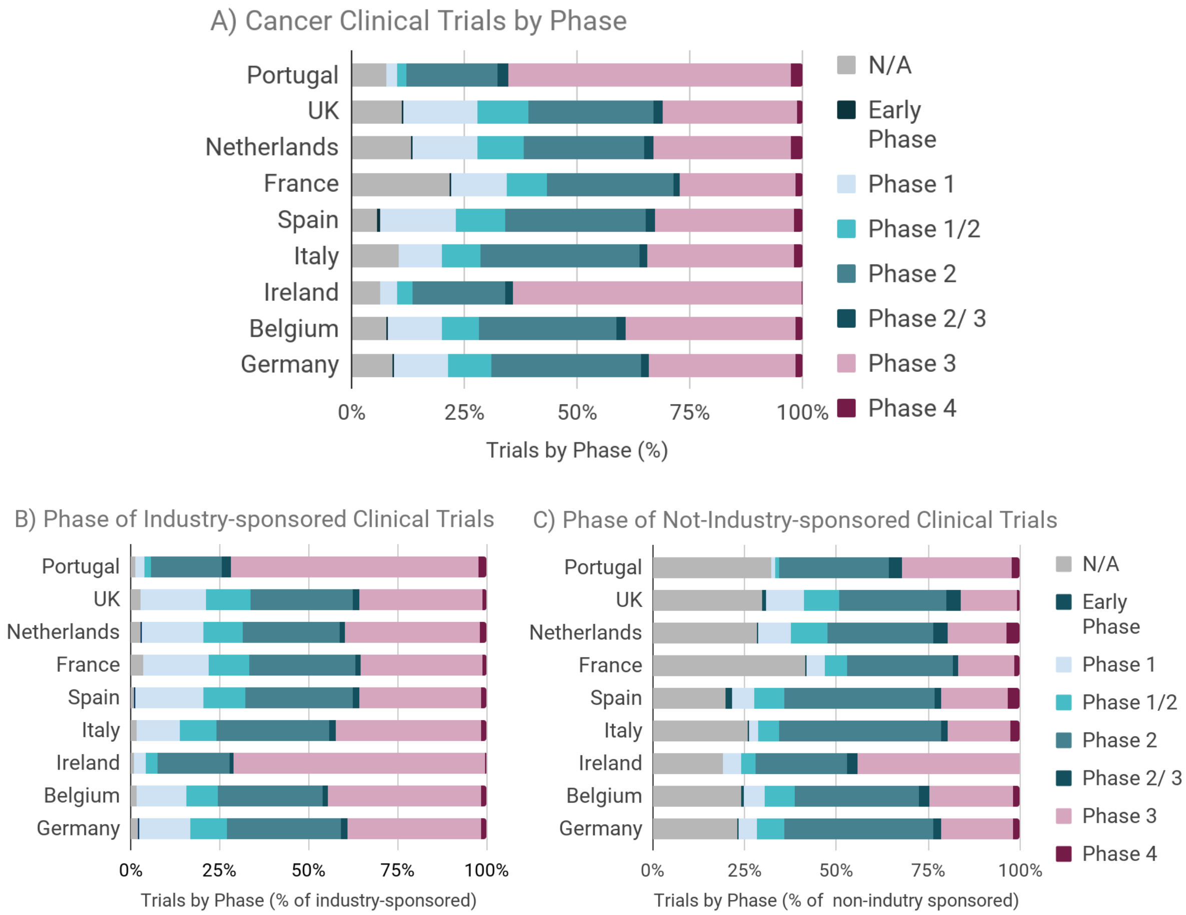


Fig. 12. Phase of cancer clinical trials in Cancer registered in clinicaltrials.gov, (started between 2010-2020): *A) percentage of clinical trials per phase (in the total number of interventional trials, excluding observational); B) and C) percentage of trials in the different phases per sponsor type.*

Intervention Types

In contrast with most benchmark countries, both in Ireland and Portugal an over-representation of phase 3 in industry-sponsored trials can be appreciated, in detriment of early phases and phase 4 trials (post-commercialization), predominantly tackled in independent clinical trials (see *Fig. 12B vs. C)*. Furthermore, there is a clear interest of the industry in pharmacological versus other types of interventions (like radiation or procedures) (*Fig. 13A)*, when compared to the wider diversity of interests of non-private sponsors (*Fig. 13B vs. C*). It should be noted that, the more the industry plays a significant role in a given ecosystem, the more divergent will the industry vs non-profit patterns of clinical trials potentially be.

Despite the dominance of industry sponsorship and drug-related interventions in Portugal, the non-industry sponsored trials conducted in the Portuguese system do include a variety of interventions similar to very distinct realities (like the UK or Netherlands) (*Fig. 13C*). This indicates that, **despite a general dependency on private funding** (and therefore, its interest)**, independent researchers in Portugal are focusing their efforts in an array of interventions** (although significantly less in behavioural studies) (*Fig. 13C*).

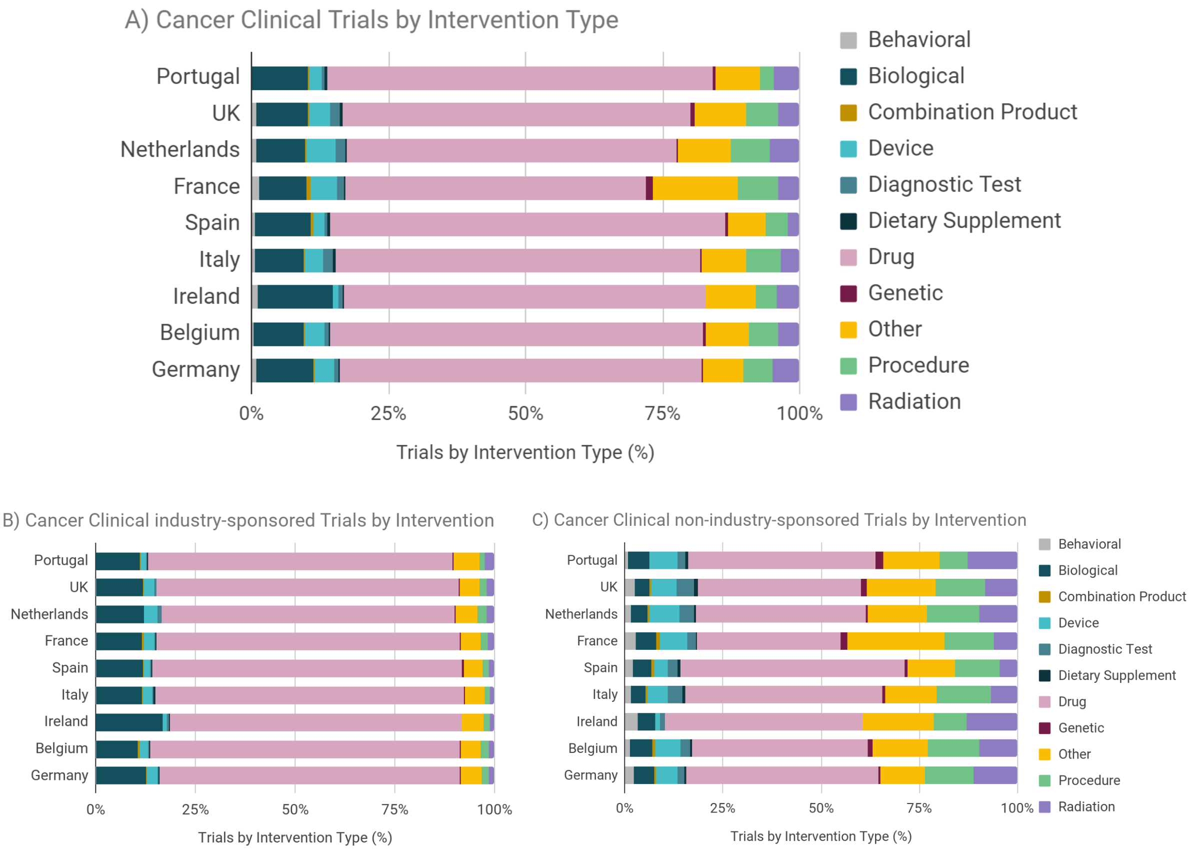


Fig. 13. Cancer clinical trials according to the Intervention, registered in clinicaltrials.gov (started between 2010-2020): *A) percentage of clinical trials per intervention type (in the total number of interventional trials, excluding observational); B) and C) percentage of trials per intervention type and sponsor type.*

Coordination and Actors (facilities and sponsors)

Portugal is the country of the panel with the lowest share of trials exclusively conducted in its territory (~7%, closer to 12% in Ireland), whereas Belgium and Spain have around 20%, UK, Netherlands and Germany just above 30%, and the Netherlands 50% (Fig. *14A*).

Even more, in Figure *14A*, we can appreciate that **Portuguese clinical sites are mainly participating in large trials, where above 10 countries are involved**; **which supports the idea of an underdeveloped system, which is attracting large private sponsored trials to a great extent**. According to the INFARMED registry, Portuguese investigators are mainly involved in multicentric studies (96%), where non-commercial sponsors are predominant (*Fig. 14B, C*).

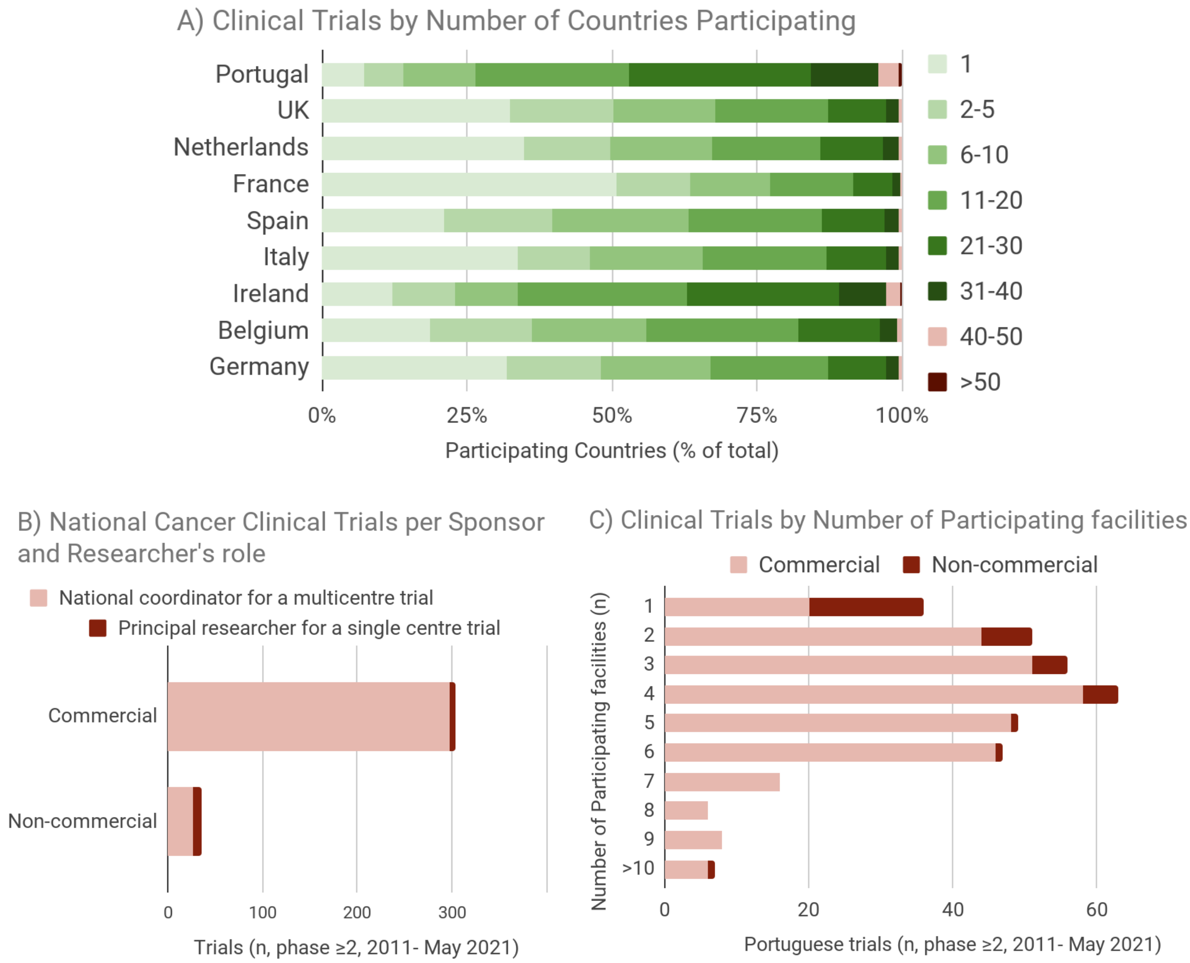


Fig. 14. Trials per number of countries involved and role of researcher: *A) percentage of trials expected to enroll patients in 1 (Portugal) to 62 different countries (source: clinicaltrials.gov; trials started between 2010-2020); B) Portuguese trials (except phase 1) according to type of sponsor and researchers role and C) number of clinical sites expected to recruit patients in each trial*[60] *(Portuguese trials approved between 2011-May 2021, source INFARMED).*

Regarding the most engaged Portuguese hubs in cancer clinical trials, we find the most populated areas of Lisboa and Porto, followed by significant contributions of Coimbra and Braga (*Table 5A*). Accordingly, the main facilities coordinating cancer trials (the vast majority multicenter) are represented by the IPOs of Porto, Lisboa and Coimbra, as well as major

[60] any trial site added to/removed from the study at a later stage than study authorization may not be reported to INFARMED, therefore this data may not correspond 100% to the reality.

hospitals (Hospital Santa Maria, Centro Hospitalar Universitário de Coimbra, Santo António Porto and Centro Hospitalar Lisboa Norte), the Champalimaud Cancer Center and a few private hospitals (Hospital da Luz and CUF) (*Table 5B*). This constitutes a set of diverse entities, all significant in the national ecosystem: National Agency on oncological therapy (IPOs), main university hospitals, philanthropic research organizations and private care facilities, which is rather interesting and may represent a strength for the Portuguese ecosystem if the entities are able to articulate and collaborate fluidly.

A) Clinical Trials (ct.gov)

City	(%)
Lisboa	94.5
Porto	83.9
Coimbra	28.0
Braga	8.4
Vila Nova de Gaia	7.7
Matosinhos	6.0
Loures	5.5
Almada	5.1
Guimarães	3.9
Santa Maria da Feira	3.6
Amadora	2.7
Vila Real	2.2
Aveiro	1.9
Setúbal	1.7
Évora	1.7
Faro	1.4

B) Clinical Trials (INFARMED)

Clinical Trials Coordinating Facilities	(%)
IPO Porto	28.9
IPO Lisboa	14.5
Centro Hospitalar Lisboa Norte, EPE, Hospital De Santa Maria	8.0
Fundação Champalimaud	7.4
Centro Hospitalar E Universitário De Coimbra, EPE	5.9
Centro Hospitalar e Universitário Lisboa Norte, EPE - Hospital Pulido Valente	4.4
Centro Hospitalar Do Porto, Epe - Hospital De Santo António	4.4
Hospital Da Luz, SA.	3.8
Hospital Cuf Porto, SA	2.7
Centro Clínico Champalimaud - Fundação Champalimaud	2.4
Centro Hospitalar Universitário De São João, EPE	2.1
IPO Coimbra	1.5
Hospital Beatriz Ângelo	1.5
Centro Hospitalar Vila Nova De Gaia/Espinho, EPE	1.5
Hospital De Braga	1.2
Centro Hospitalar De Lisboa Norte, EPEe	1.2
Centro Hospitalar De Lisboa Central, EPE. - Hospital St António dos Capuchos	1.2
IPO (*unspecified*)	0.9
Hospital Garcia De Orta, EPE	0.9
Unidade Local De Saúde De Matosinhos, EPE - Hospital Pedro Hispano	0.6
Hospital Prof. Doutor Fernando Fonseca	0.6
Hospital Cuf Descobertas, SA	0.6
Centro Hospitalar Do Baixo Vouga, EPE (Hospital Infante D. Pedro)	0.6
Centro Hospitalar De Setúbal, Hospital De São Bernardo	0.6
Centro Hospitalar De Entre o Douro E Vouga, EPE	0.6
Sociedade Gestora Do Hospital De Loures, SA	0.3
Instituto De Ciências Nucleares Aplicadas À Saúde	0.3
Hospital Senhora Da Oliveira Guimarães - EPE	0.3

***Table 5. Major Portuguese cities and facilities in cancer clinical trials (trials started between 2010-2020): A)** main cities participating in clinical trials, data as a percentage of cancer trials (note that total is ≥100% due to multi-site trials; cities involved in ≥1.4% of trials as registered in clinicaltrials.gov); **B)** Main facilities coordinating clinical trials, data as percentage of cancer trials (according to INFARMED; facilities coordinating ≥ 2 trials).*

Portugal's clinical trials are mainly sponsored by industry entities (>70%- *Fig. 11A*) and, among the most engaged companies in Portugal's clinical research, we find Hoffmann-La Roche funding around 14.7% of the trials, followed by Novartis Pharmaceuticals and Merck Sharp & Dohme Corp. (around 6%) (*Table 6*).

Regardless, some relevant non-private sponsors are also involved in a a relevant share of trials conducted in Portugal: two large international non-profit organizations for cancer research (EORTC and BIG) and one portuguese philanthropic actor (the Champalimaud Foundation, which also directly funds cancer research though its Clinical Center specialized in Oncology therapy) (*Table 6*).

Sponsor	Type	Clinical Trials (%)
Hoffmann-La Roche	Industry	14.7
Merck Sharp & Dohme Corp.	Industry	6.3
Novartis Pharmaceuticals	Industry	6.0
Janssen	Industry	4.8
AstraZeneca	Industry	4.6
European Organisation for Research and Treatment of Cancer	Other	4.1
Celgene	Industry	4.1
Bristol-Myers Squibb	Industry	3.6
Amgen	Industry	3.4
Pfizer	Industry	3.4
Bayer	Industry	2.9
Boehringer Ingelheim	Industry	2.9
Sanofi	Industry	2.9
Daiichi Sankyo, Inc.	Industry	2.2
AbbVie	Industry	1.9
Genentech, Inc.	Industry	1.9
Breast International Group	Other	1.7
Millennium Pharmaceuticals, Inc.	Industry	1.7
Fundação Champalimaud	Other	1.4
Eisai Inc.	Industry	1.4
Pharmacyclics LLC.	Industry	1.4
Astellas Pharma Global Development, Inc.	Industry	1.4
Ipsen	Industry	1.2
Gilead Sciences	Industry	1.2
MedSIR	Industry	1.2

Table 6. Major private and non-private sponsors responsible for cancer clinical trials in Portugal, as registered in clinicaltrials.gov (2010-2020): *percentage of the total clinical trials (sponsors involved in >1% of trials).*

Research Focus

Cancer research is mostly predominant versus other conditions in the clinical setting (and accounts for 30% of all trials in the benchmark panel, see *Fig. 10F*). Additionally, similar trends across the panel of countries can be further observed for specific cancers.

In terms of cancer groups, clinical trials on Digestive System, Urogenital, Breast and Thoracic neoplasms represent the four most researched neoplasms across the benchmark selection, which is aligned with the EU27+UK pattern (see Annex). When analysing per specific neoplasm (*Table 7*), we can see very similar research dedication in trials amongst EU countries, and where Portugal has strong focus in Breast and Lung Neoplasm (however, Portugal has a discrete number of trials overall, and percentage levels can be slightly less diverse than in other systems).

The fact that the pattern of clinical trials by main cancers groups is considerably consistent in Europe is inherent to the main characteristics of this type of research:

- Relatively similar epidemiology of oncology in EU countries (and therefore, clinical needs and availability of patients for recruitment in clinical trials);
- Alignment with *state of the art* global public and private drug-discovery pipelines;
- A more or less established collaborative network between EU countries (aided by EU funded alliances and clinical research networks, as well as European patient and cancer research associations); which is potentially less developed in Portugal.

	Dedication to specific Neoplasms* (% of national clinical trials)									
MeSH Neoplasms*	**EU27+UK**	**Portugal**	**UK**	**Netherlands**	**France**	**Spain**	**Italy**	**Ireland**	**Belgium**	**Germany**
Breast Neoplasms	11.2	19.8	10.0	8.5	10.6	12.5	9.4	16.4	12.6	9.3
Lung Neoplasms	7.0	9.6	8.0	9.2	8.0	10.4	8.6	10.7	7.7	8.8
Colorectal Neoplasms	6.3	3.6	4.5	5.5	4.8	5.7	5.3	5.2	5.7	4.3
Prostatic Neoplasms	6.2	5.8	8.1	6.9	5.1	4.7	4.2	8.3	5.7	5.2
Carcinoma, Non-Small-Cell Lung	6.1	11.8	7.5	9.7	7.6	11.2	8.8	9.4	8.1	9.3
Ovarian Neoplasms	2.7	1.9	3.1	2.8	2.4	2.9	2.8	3.1	4.2	2.5
Pancreatic Neoplasms	2.5	1.7	1.8	2.0	1.4	1.9	2.7	2.6	1.7	1.9
Carcinoma, Ovarian Epithelial	2.2	1.4	2.6	1.8	1.9	2.4	2.3	3.1	3.5	2.3
Carcinoma, Renal Cell	2.2	1.2	2.8	2.8	2.2	2.7	2.4	2.6	2.0	2.1
Carcinoma, Hepatocellular	2.0	1.7	1.8	1.1	2.9	2.5	2.8	1.6	2.5	2.8
Head and Neck Neoplasms	2.0	1.7	1.4	2.0	1.8	1.1	1.2	1.3	1.8	1.4
Rectal Neoplasms	1.9	0.5	1.1	1.3	1.0	0.8	1.0	0.5	0.7	0.7
Stomach Neoplasms	1.5	4.1	1.4	1.4	1.1	1.7	1.8	2.1	1.2	2.0
Urinary Bladder Neoplasms	1.5	1.9	1.7	2.0	1.2	1.5	1.5	0.8	1.3	1.3
Hematologic Neoplasms	1.4	1.0	1.3	1.2	1.6	1.2	1.5	0.5	0.8	1.0
Squamous Cell Carcinoma of Head and Neck	1.3	1.7	1.5	1.1	1.7	1.7	1.2	1.6	2.0	1.5
Esophageal Neoplasms	1.3	1.2	1.3	2.2	0.9	0.6	0.7	3.1	0.9	1.3
Colonic Neoplasms	1.2	0.5	0.5	0.8	0.8	0.7	0.6	0.3	0.4	0.3
Uterine Cervical Neoplasms	1.2	0.5	0.7	0.9	1.1	0.6	0.7	0.8	1.2	0.5
Endometrial Neoplasms	1.1	0.0	1.3	1.0	0.8	1.0	0.9	1.6	1.5	0.8
Liver Neoplasms	1.1	0.0	0.6	0.9	1.0	0.4	0.8	0.0	1.0	0.8
Triple Negative Breast Neoplasms	1.1	2.9	1.9	1.1	1.4	2.7	1.7	2.3	2.3	1.5
Leukemia, Myelogenous, Chronic, BCR-ABL Positive	0.9	1.4	0.9	1.4	1.2	0.9	1.5	0.8	1.2	1.6

**according to the NIH-NLM MeSH terms attached to each clinical trial*

Table 7. Cancer clinical trials by neoplasm (2010-2020): *percentage of clinical trials of the region or country, classified by their MeSH terms that belong to the C04-Neoplams branch of the MeSH Taxonomy (top 23 cancers in EU27+UK).*

Cancer is one of the main areas of clinical trials in developed countries, and its growth in Portugal has been very significant in the last decade (despite the small number of trials, *per capita* the number of those in cancer is often close to the German system, for example).

Nevertheless, **Portugal does appear to have an underdeveloped clinical trial ecosystem**, and the visibility that conducting clinical trials can bring to Portuguese researchers and institutions **should be clearly incentivated and supported**.

Trials conducted in Portugal are **highly reliant on private sponsorship,** which entails a profile more aligned with pharmaceutical industry interests (e.g. dedicated to specific drugs approval)**, and where Portugal facilities are mainly engaged in large multicentric and multinational trials**. Private sponsorship includes major industry partners, as well as relevant non-profit and philanthropic sponsors (national and international).

In spite of this, **non-commercial trials are also increasing and should be supported,** showing a **diverse array of interests** from the clinical researchers (in terms of trials interventions and phases, even if with potential gaps in behavioral and early phase trials) .

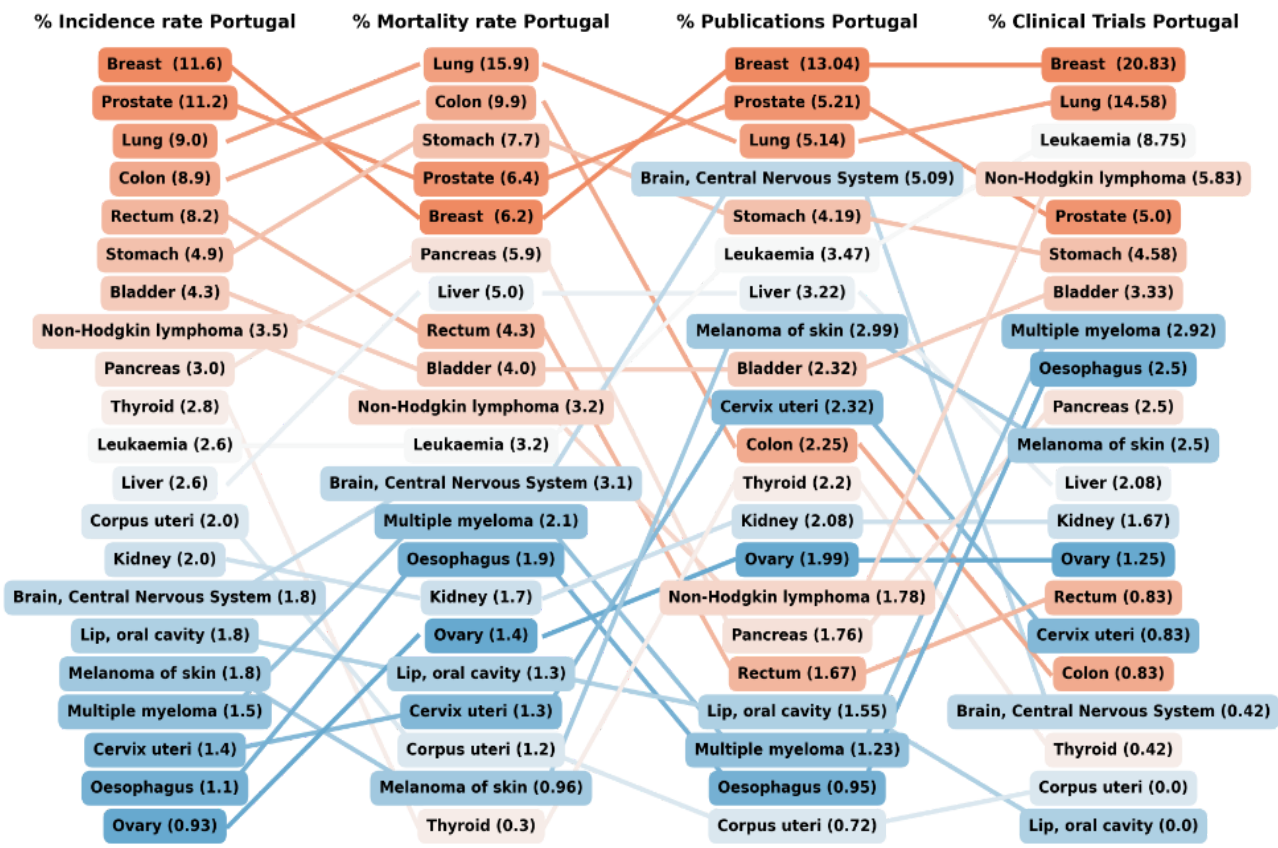


Fig. 15. ***Alignment between cancer mortality and incidence*** *(data from WHO/GLOBOCAN 2020, data in %)* ***and Published work and clinical trials in Portugal*** *(2015-2020, data as in % of all national publications and clinical trials according to clinicaltrials.gov). The alignment was made by aligning MeSH Neoplasms to WHO/GLOBOCAN Classification by Cancer site (see Annex).*

C. European Competitiveness

Horizon 2020

As previously commented at the beginning of this report, cancer is one of the five high-priority missions heavily funded by the European Commission under the Horizon 2020 program (between 225 and 310M € per year were granted to projects related to cancer research, according to our analysis). In fact, funding is expected to increase in the next few years, as new initiatives such as the *Beating Cancer Plan* fully develop[61]. In this context, understanding the current capacity of Portuguese cancer research to attract funding from European sources sheds significant light over its evolution and competitiveness[62].

In the last couple of years, **Portugal has increased the funding received for cancer research from H2020** (*Fig. 16A*). Although securing a discrete amount (41 projects for roughly 16M€) from the whole H2020 period *(Fig. 16B*), 5M€ (10 projects) were granted in 2020 alone. **Per capita, Portugal positioned above Germany, France, Italy and the UK in the 2020 calls**. The Netherlands and Spain fare especially well and, both in raw and *per capita* numbers (*Fig. 16C)*, show a highly competitive profile on the European stage.

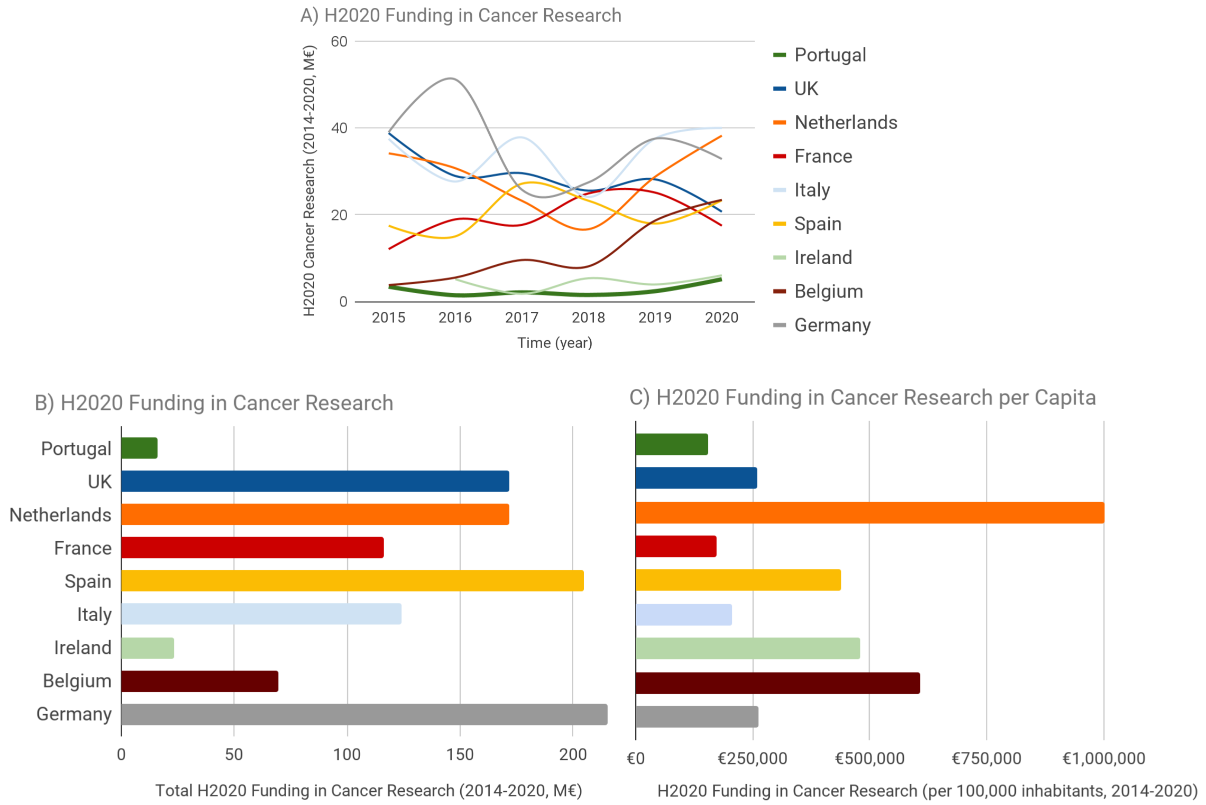


Fig. 16. European funding from H2020 granted to cancer research (2014-2020) for the panel of countries: A) *funding for cancer projects per year and* ***B)*** *in total;* ***C)*** *total funding for cancer projects per capita (100,000 inhabitants).*

[61] EU declares a €4B research war on cancer | Science|Business (sciencebusiness.net)
[62] It should be noted that some research ecosystems with insufficient or less reliable funding structures are more likely to apply and be successful in EU calls. This effect can be observed in systems like Spain and Portugal to a certain degree, since researchers are more dependent on additional funding sources. Nevertheless, the fact that the proposals evaluated by the EU funding schemes are positively selected is still highly significant.

The main H2020 instruments supporting biomedical research in general, including Cancer Research, are:

- the European Research Council and the Marie Skłodowska-Curie actions of the Excellence pillar;
- those focused on key identified Societal Challenges, and, where cancer Research is concerned, the instrument dedicated to the panel of Health[63];
- in addition to actions funding Industrial Leadership (including LEIT and SME instruments).

A similar pattern in the share of H2020 cancer funding per instrument can be observed in several systems, such as the UK, France, Netherlands, Spain and Germany (*Fig. 17A*). However, funding for SME instruments appears quite variable (where Spain and Netherlands excel) (*Fig. 17B*), and where two aspects generally intertwine: a) the actual competitiveness and b) the availability of national support for private innovation efforts .

Portuguese cancer research shows a clear gap in Industrial Leadership (LEIT) and SME instruments competitiveness[64] (*Fig. 17A-B*); and we could only detect one coordinating project from FET[65] (an instrument designed *"to turn Europe's excellent science base into a competitive advantage"*[66]).

The profile of Portuguese actors receiving H2020 cancer research funding is similar to France and Spain, with a higher involvement of research organisations than education establishments (e.g Universities), as is the case of the UK and Ireland (Fig. *17C*). This is mainly due to the different organisation of research in these systems versus those in southern Europe. The involvement of private entities is similar for most benchmark countries.

Overall, the rate of involvement of private partners in H2020 funding calls[67] granted to Portugal, being not particularly high, is currently in line with some of the benchmark countries (*Fig. 17D*). It shall be noted, however, that although involvement of private entities in EU cancer research projects is significant, it is still inferior to other research areas (which is in fact a common trait across several research ecosystems, except for the UK, Ireland and Netherlands).

---

[63] Horizon 2020 sections | Horizon 2020 (europa.eu)

[64] From 2014 to 2018, SMEs were funded as part of the H2020 consortium calls. In 2018-2020, the EU piloted the European Innovation Council (EIC), which led to changes for the SME instrument, with the creation of the EIC Accelerator.

[65] A Project coordinated by the Universidade de Aveiro, with a global budget of 3M€ (0.5M€ to UA), 2018-2022: Nanoparticles-based 2D thermal bioimaging technologies | NanoTBTech Project | H2020;

[66] Future and Emerging Technologies | Horizon 2020 (europa.eu)

[67] In a constant effort to translate R&D activities into innovation, H2020 collaborative research and innovation projects seeked to include private partners that would facilitate knowledge translation.

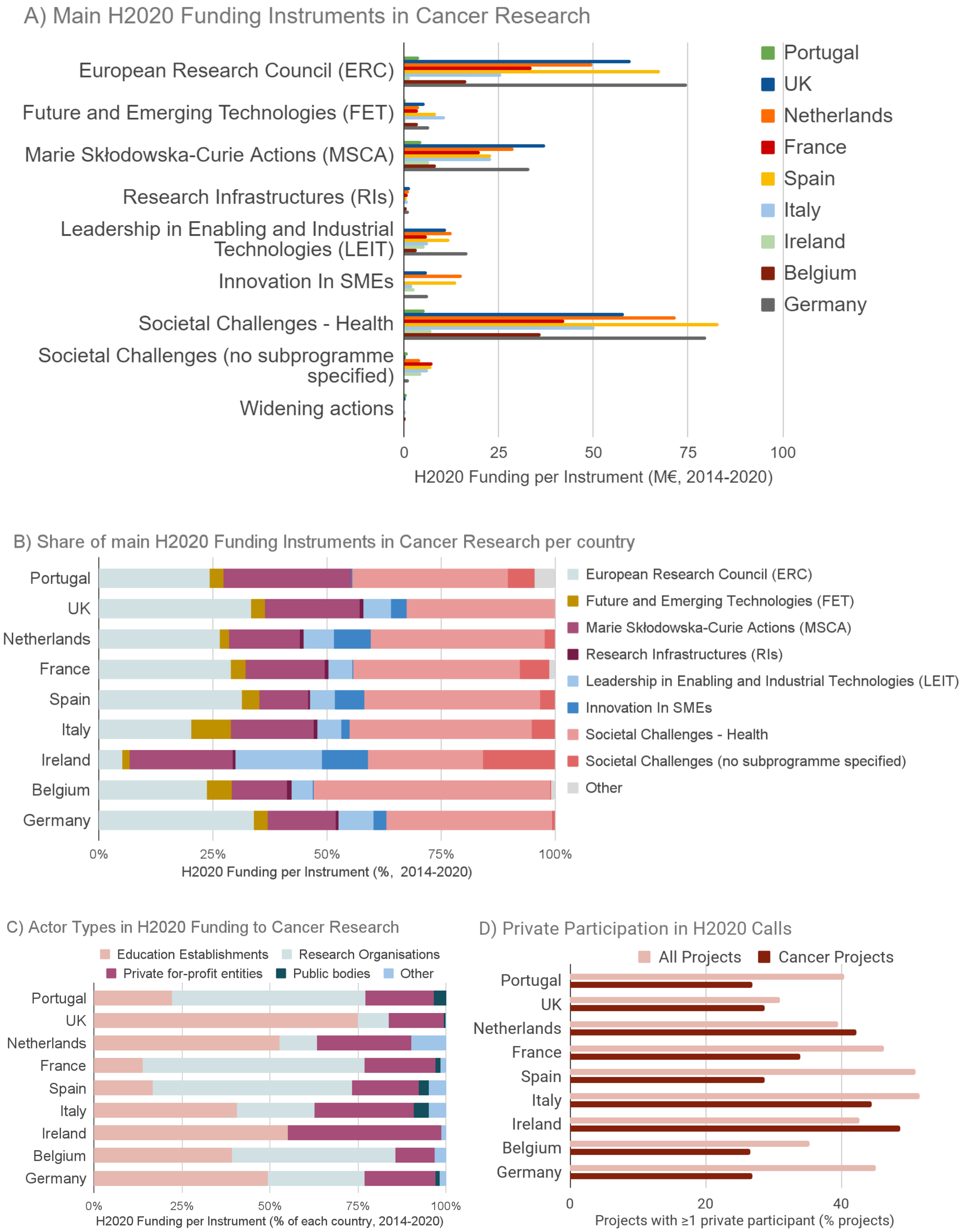


***Fig. 17. Instruments and Actors in H2020 European funding granted to cancer research (2014-2020) for the panel of countries:** **A)** H2020 Funding in Cancer per top H2020 instruments (M€); **B)** share of funding per top H2020 instruments in Cancer research (% of total funding); **C)** share of funding per H2020 Actor types in Cancer research; **D)** percentage of H2020 grants awarded with one private applicant (either coordinator or partner) for all projects in biomedical research (% relative to the total number of grants) and cancer research (percent relative to cancer-related grants).*

European Research Council (ERC)

When analysing more in detail the highly competitive European Research Council (ERC) funding, whose mission is to encourage high-quality research and support investigators on the basis of scientific excellence[68], Portugal also secures discrete funding for cancer research (3.9M€ and 5 projects in the H2020 framework programme). However, **when normalised by population** (ERC per 10M inhabitants)**, Portugal is close to the UK and Germany and surpasses France and Italy, in terms of ERC grants received for cancer research (***Fig. 18A***)** even if the funding amount is more discrete (*Fig. 18B*).

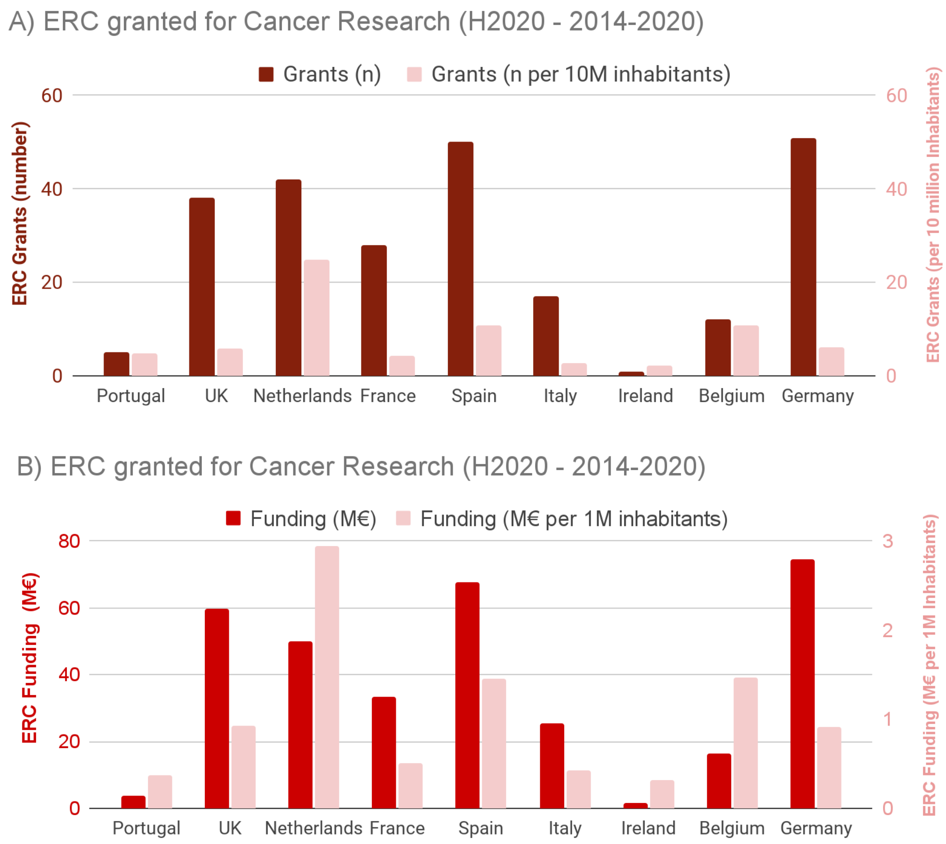


Fig. 18. European funding from ERC (2010-2020) for the panel of countries: A) *total ERC number of grants and number of grants per 10M inhabitants;* ***B)*** *total ERC funding and funding per 1M inhabitants.*

There is also considerable variability in the type of ERC grants awarded for cancer research in the countries of the panel. Systems such as the UK, Netherlands, Spain and Germany appear to be consolidated ecosystems with a balanced distribution of ERC grants across career stages (*Fig. 19A*).

Portugal does not have a balanced profile in cancer research according to the granted ERCs, despite showing a slightly more balanced profile if we take all areas under consideration (all ERCs) and in the biomedical area in general (still with only 6% of Advanced Grants) *(Fig. 19B*). Younger researchers (2-12 years post post-PhD) are mostly responsible for the competitiveness of the Portuguese ecosystem in ERC calls[69]*(Fig. 19B*), and no senior

[68] https://erc.europa.eu/about-erc/mission

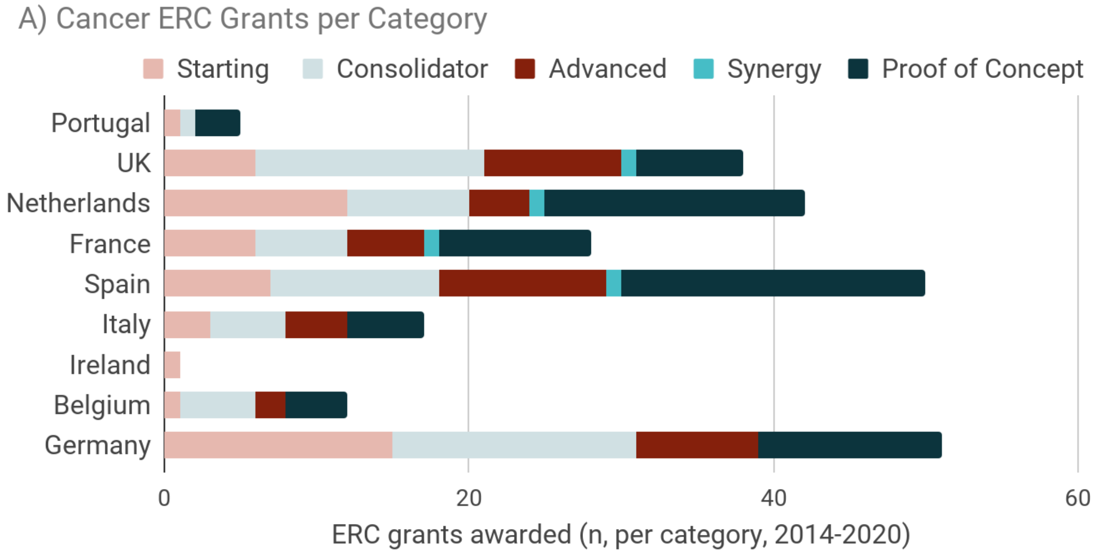
[69] *Starting grants* allocated to researchers with 2-7 years of experience since completion of PhD-;

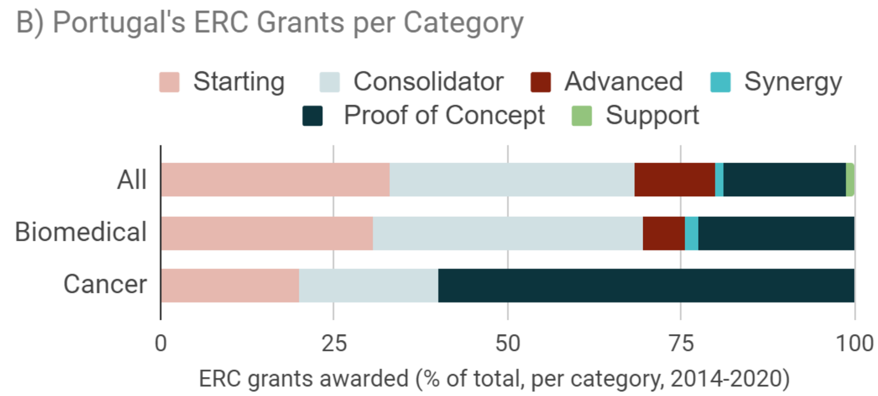
researcher has secured an ERC grant for cancer research *(Fig. 19A-B)* (either because applications are not placed or the conditions are not encountered to continue their research nationally).

Due to the recent development of the Portuguese cancer research system, the pattern shown is to be expected. Emergent systems often find it easier to attract early and mid-career researchers, versus securing or attracting senior established researchers; which is linked to the system's visibility, available infra-structure, funding opportunities, stimulating environment, among others.

Portugal has currently no projects in "Advanced" or "Synergy" profiles for cancer research; however it presents a significant proportion of "Proof of Concept" projects in line with other benchmarks (*Fig. 19A-B*).

Portugal's ERC H2020 grants also reflect the relatively smaller dedication to cancer research, since biomedical competitiveness is led by neurosciences and the broad topic of Biomaterials and Tissue Engineering (*Fig. 19C*).





Consolidator grants allocated to researchers with 7-12 years of experience since completion of PhD; *Advanced grants* allocated to active researchers -principal investigators (PIs)- who have a track-record of significant research achievements in the last 10 years; *Synergy grants* allocated to a group of 2 to 4 PIs. PIs must present an early achievement track-record or a ten-year track-record; Proof of Concept grants allocated to PIs in one of the ERC frontier research main grants (Starting, Consolidator, Advanced or Synergy). https://erc.europa.eu/funding

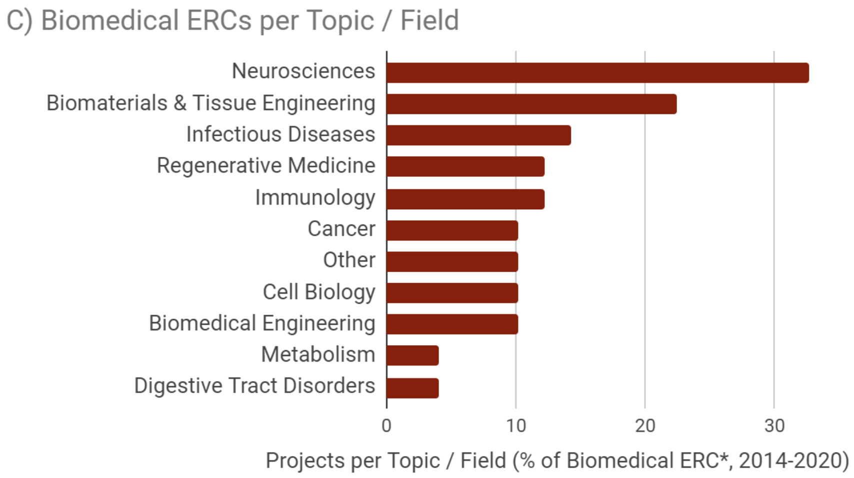


Fig. 19. European funding from ERC per type of call (2014-2020): A) *number of ERC career stage grants in cancer research for the panel of countries;* ***B)*** *% of total ERC grants awarded per category in all ERC and limited to the biomedical field, in Portugal;* ***C)*** *ERC grants per topic and/or field manually categorized (% of all biomedical ERC grants, n=51) (*please note that the total is >100%, since projects can belong to more than one category - most notably Regenerative Medicine & Biomaterials/Tissue Engineering, but also Neurosciences & Immunology for example).*

Portugal's **cancer research is increasingly competitive in EU calls**, a fact that is likely a testament of the quality of the research, the evolution of the Portuguese ecosystem (as a whole, in biomedical research, and in cancer), as well as the need to rely on extraordinary funding sources.

As discussed before, Portugal is not overly specialized in cancer research, and indeed we can see that other biomedical areas are better represented in securing additional EU competitive funding (e.g. neurosciences, a fact also observed in the philanthropic Health research calls of "la Caixa" Foundation - see below).

Considering the different H2020 funding instruments, Portugal presents a clear gap in competitiveness in innovation calls. Furthermore, **considering the ERC grantees, the Portuguese ecosystem is young and less prone to attract or stabilize highly competitive later-stage researchers**.

Regarding funding for private partners, although these are, overall, fairly well represented in EU projects (most likely as partners), **very few Portuguese private companies are able to secure EU funding for innovation developments directly from H2020 instruments**. Which again, is not supporting evidence of a properly developed innovation system.

Box 2. Main non-governmental funders of cancer research in the Portuguese ecosystem: Liga Portuguesa Contra o Cancro and "la Caixa" Foundation

Contrary to anglo-saxon systems, robust philanthropic and charity models of research funding are not overly frequent in Portugal, where there is a strong dependency on state funding, mainly through one agency (Fundação para a Ciência e Tecnologia - FCT). However, two main actors (one historic and one modern) run competitive and established funding schemes at the national level, which have very specific and complementary roles (with each other, and with the national agencies - mainly FCT). Other actors have occasionally run national funding actions in cancer research, like those promoted by Maratona da Saúde[70].

Liga Portuguesa Contra o Cancro

With first actions as early as 1931's, the Liga Portuguesa Contra o Cancro (LPCC) was officially founded in 1941, and has been a key actor in promoting cancer research in Portugal ever since[71]. Similar to other like-minded initiatives in other countries, through fundraising initiatives, the Liga started by supplementing the state's role in the advancement of cancer therapy, care conditions and cancer research at national level; including later on a strong focus on cancer prevention that includes national wide awareness and screening campaigns, notably a well established programme in breast cancer screening.

Its organisation in regional delegations is aligned with the wish to educate and create awareness in close proximity with the population and with the distinct oncology centres.

LPCC supports cancer research in direct actions that support researchers through individual calls for scholarships, mainly for applied research that shows innovative potential, and institutional support.

"la Caixa" Foundation (LCF)

LCF's mission includes supporting Iberian biomedical research, with Oncology being one of the strategic areas of interest, since 2018. As a philanthropic funder, LCF possesses instruments that support individual scholarships, project-based high-impact research, technology transfer projects and also direct institutional support[72].

CaixaResearch boosts health research to promote social progress, and operating in an excellency model, the Health Call[73] supports 20-30 projects per call (500K for an individual

[70] Maratona da Saúde uses a specific funding model named telethon (in cooperation with RTP), in which yearly televised entertainment is paired with scientific fundraising. Each year a different biomedical field is tackled. Cancer research was the focus of its 1st (2013) and 6th (2018) editions; where respectively 186K€ and 50K€ were raised, and 15 investigators were supported.
[71] https://www.ligacontracancro.pt/
[72] At a glance - "la Caixa" Foundation (fundacionlacaixa.org)
[73] Health research call - "la Caixa" Foundation (fundacionlacaixa.org)
Health research results 2021 - "la Caixa" Foundation (fundacionlacaixa.org)

project, 1M for a consortium); while the CaixaImpulse programme offers funding for different tasks necessary for product development (70-100k) as well as training in entrepreneurship for researchers.

The success of Portuguese research institutions in the LCF translational research calls is significant (16 projects); overall Portuguese organizations have secured >18% of Health Call grants (particularly in 2021, with 7 projects awarded). Within the thematics, Portugal appears more competitive in neurosciences and infectious diseases, having been awarded 5 projects in each of these thematics between 2018 and 2021.

Furthermore, and profiting from LCF thorough selection process, extra projects are awarded in collaboration with the Fundação para a Ciência e a Tecnologia, I.P. within the "Iniciativa Ibérica de Investigação e Inovação Biomédica, i4b". In total, FCT has granted support to 14 more projects highly ranked by the review panel of the LCF health call.

In the Health Calls, and considering both the panel on Oncology and Enabling technologies, Portugal has secured 3 projects for cancer research: 1 in 2019, and 2 in 2021, for a total of close to 1,5M. Two of these projects were attributed to Instituto de Medicina Molecular-Lisboa; and the remaining to i3S-Porto. Two more projects have been awarded through the i4B programme (both in 2020 and attributed to IMM).

In the CaixaImpulse programme, only one project for cancer research was granted funding to an institution in Portugal: the Laboratorio Ibérico Internacional de Nanotecnología-Braga in 2020 (300K).

D. National research ecosystem & Actors

Collaborative work, Leadership and Internationalisation

In line with global research trends, and due to the increasing complexity and interdisciplinarity of biomedical research, the degree of collaborative work in cancer research has been steadily increasing in the past decade (a trend that can be observed in all the benchmark countries) (*Fig. 20A*). In Portugal, an increase from roughly ~50% of publications with an international partner in 2010 to ~60% in 2019 can be observed (*Fig. 20A*).

Regarding leadership of bibliometric output (first or last authorship of publications), a reduction between 5 to 10% was seen across the panel of countries (Fig. *20B*). In cancer research, this decrease is strikingly correlated with the increase in international collaborative work (*Fig. 20C*)[74], which indicates that a decrease in "leadership" of research is a welcomed trade-off in contemporary research (global and increasingly transdisciplinarity).

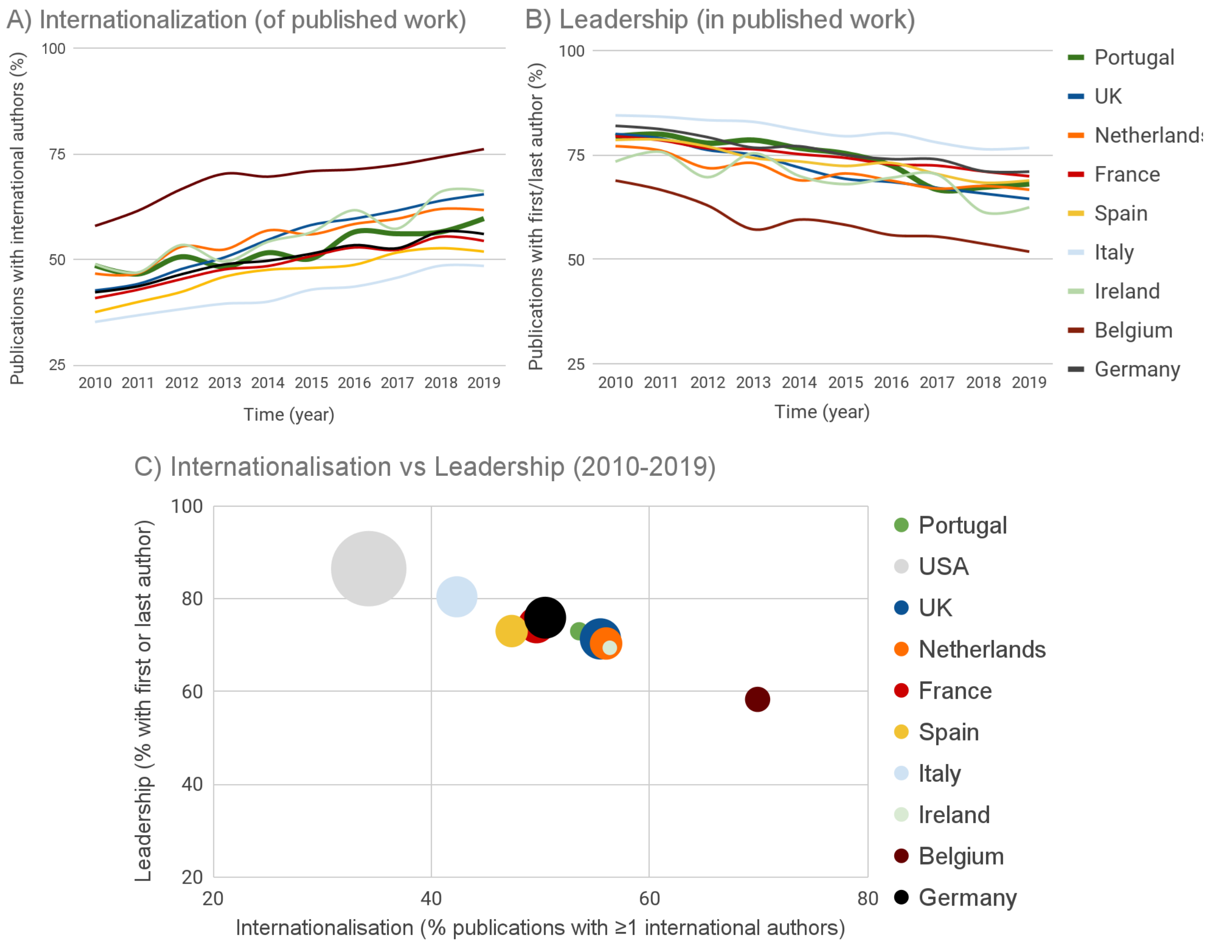


Fig. 20. International publications in cancer research (2010-2019): A) *percentage of publications in collaboration with ≥1 international partner,* ***B)*** *fraction of lead international publications (publications where the national partner is either last or first author) and* ***C)*** *percentage of "international publications" versus percentage of "lead publications" (dot size is relative to the total number of documents in cancer research for each country)[74].*

[74] USA was included although its sheer size (in research and population) makes it unfit for direct comparisons. Nevertheless, it is a good example of how size *per se* can directly affect/distort these indicators.

As outliers, we have Italy with the lowest internationalisation rate, but significant leadership; a paradigm that is not rendering the best results in terms of research excellence (in citation indexes or Nature Index recognition), although relevant specialisation (*see Fig. 2C and Fig 3*). On the other hand, as a smaller country (and one of similar size to Portugal), Belgium has the highest internationalisation rate, with the lowest leadership, and achieves high levels of peer recognition and competitiveness in EU funding for cancer research (see *Fig. 3 and 20C*).

In 2019, Portuguese entities co-authored around 60% of its scientific production in cancer research with international partners and led the research in roughly 68% of those publications. In this regard, the **Portuguese ecosystem will have to establish its model. As a country on a smaller scale, a pattern more resembling Belgium should be more fitting (where higher internationalisation of research boosts research excellence and visibility).**

The countries mainly represented in the collaborative network of Portugal cover world leaders (USA), European leaders (UK and Germany) and neighbouring countries (Spain, France, Italy) (*Table 8*). In addition, at the institutional level, we find several organizations from Brazil, Sweden and Netherlands present in the Portuguese collaboration network. Furthermore, collaboration with major highly dedicated or specialized cancer centres (*German Cancer Research Center, Memorial Sloan-Kettering Cancer Center*) are observed (*Table 8*). Overall, **we find high level partners in the Portuguese collaboration network.**

Country	share of collaborative publications
United States	18.0
Italy	15.5
Spain	15.3
UK	14.6
Germany	12.8
France	11.6
Netherlands	9.6
Brazil	8.5
Belgium	8.2
Sweden	6.6
Switzerland	6.3
Canada	6.1
Poland	5.1
Australia	4.9
Denmark	4.3
Austria	4.2
China	4.1
Norway	3.4
Japan	3.3
Finland	3.2
Greece	3.1
Israel	2.9
Chech Republic	2.9
Hungary	2.4

Institution	Country	share of collaborative publications
Erasmus University Medical Center	Netherlands	2.7
Universidade de São Paulo - USP	Brazil	2.7
German Cancer Research Center	Germany	2.7
Memorial Sloan-Kettering Cancer Center	USA	2.5
Karolinska Institutet	Sweden	2.4
Hospital de Câncer de Barretos	Brazil	2.3
Fondazione IRCCS Istituto Nazionale dei Tumori, Milan	Italy	2.1
Medizinische Universität Wien	Austria	2.1
University of Toronto	Canada	1.9
Karolinska University Hospital	Sweden	1.9
University of Cambridge	UK	1.9
KU Leuven– University Hospital Leuven	Belgium	1.9
Mayo Clinic	USA	1.9
National Cancer Institute	USA	1.8
University of Texas M. D. Anderson Cancer Center	USA	1.8
University Medical Center Utrecht	Netherlands	1.8
Leiden University Medical Center - LUMC	Netherlands	1.7
Radboud University Nijmegen Medical Centre	Netherlands	1.7
Harvard Medical School	USA	1.6
Alma Mater Studiorum Universita di Bologna	Italy	1.6
Institut Curie	France	1.6
Institute of Cancer Research London	UK	1.6
Universitat de Barcelona	Spain	1.6
UCL	UK	1.6
The Netherlands Cancer Institute	Netherlands	1.6

Table 8. Major international collaborative partners of Portugal's cancer research (2010-2020): A) *top 24 country partners, data as percentage of participation in the international publications of Portugal and* ***B)*** *top 24 collaborative institution partners, data as percentage of participation in the international publications of Portugal.*

National Actors

National Hubs

Overall, at national level, we can find four major cancer research hubs in Portugal, which are respectively: 1) Porto, the biggest hub production wise (including also Vila Nova de Gaia, Paredes, Matosinhos, Leça do Balio, Vairão), 2) Lisboa (including Oeiras, Caparica, Almada, Amadora, Carnaxide, Sacavém), 3) Coimbra and 4) Braga and Guimaraes (*Fig. 21A*). Regarding the collaborative network, Lisboa and Porto are the most connected hubs in number of co-published work, followed by Porto-Braga and Porto-Coimbra (*Fig. 21B*).

A) *HUB*	**Share of Cancer Publications (%)**
Porto	41.7
Lisboa	38.6
Coimbra	16.5
Braga and Guimaraes	9.4
Aveiro	4.2
Vila Real	2.4
Covilha	2.3
Faro	2.2
Funchal	1.1
Braganca	0.6
Viseu	0.6
Loures	0.5
Evora	0.4

B) Co-authorship between *hubs*

Hub_A	Hub_B	(n of collabs)
Lisboa	Porto	334
Braga	Porto	206
Porto	Coimbra	203
Coimbra	Lisboa	158
Porto	Aveiro	118
Vila Real	Porto	106
Coimbra	Aveiro	64
Lisboa	Braga	55
Coimbra	Covilha	53
Lisboa	Aveiro	47

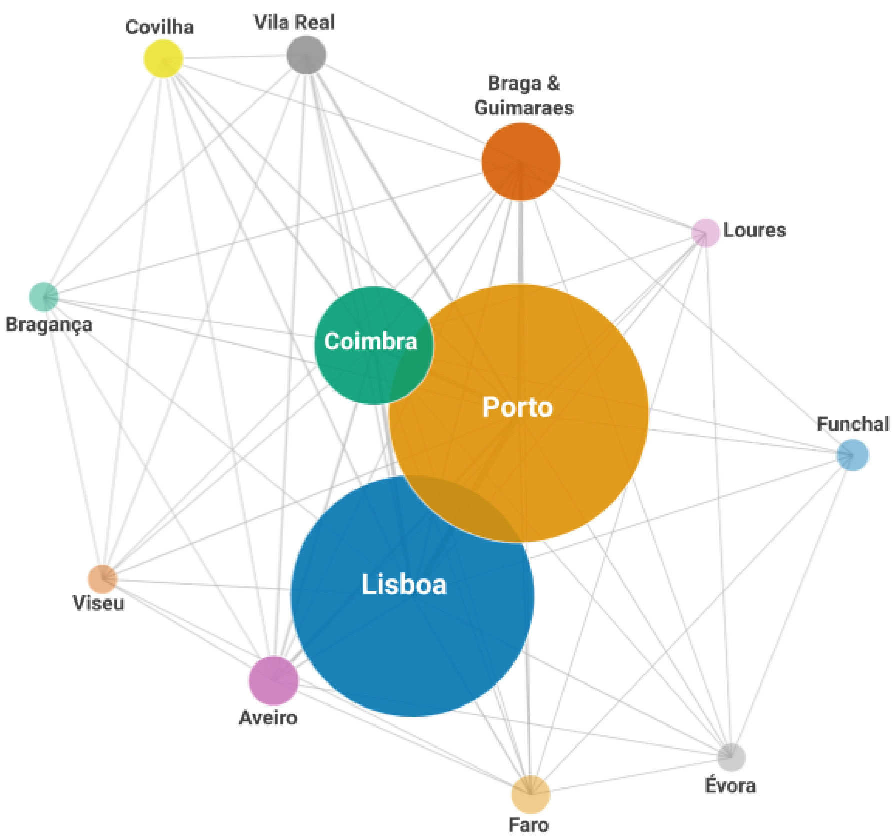


Figure 21. National hubs in cancer research (2010-2020): A) *top hubs publishing cancer research in Portugal (≥ 30 publications), data as percentage of total publications in cancer research in Portugal; and* ***B)*** *its collaborative network (connectors width is related to the number of publications co-authored between the 2 hubs; and bubble size relates to total publications in each hub).*

Research Institutions

The Porto hub, representing a share of around 41.7% of national publication on cancer research, revolves around the **Universidade do Porto**[75] **and its associated institutions: i3S, Hospital São João, Hospital Santo António, as well as IPO Porto.** The Lisbon hub, on the other hand, is constituted by a wider group of organizations, which are, at the institutional level, less centralized (*Table 9*).

Institution	City	1. Share of Cancer publications (%)	2. Public Health, Epidemiology & Clinical Research (% of institution)	3. Basic & Translational Research (% of institution)
Universidade do Porto	Porto	28.6	53.6	46.4
Instituto Portugues de Oncologia de F.G. Porto	Porto	12.4	75.6	24.4
Universidade do Minho	Braga	7.6	51.8	48.2
Hospital of São João	Porto	7.1	78.5	21.5
Centro Hospitalar e Universitário de Coimbra	Coimbra	6.7	80.8	19.2
Instituto Portugues de Oncologia de F.G. Lisboa	Lisboa	5.9	82.0	18.0
I3S- Instituto de Investigação e Inovação em Saúde	Porto	5.6	37.4	62.6
Universidade de Coimbra, Faculdade de Medicina	Coimbra	4.2	51.3	48.7
Institute of Molecular Medicine, Univ of Lisboa	Lisboa	4.1	35.2	64.8
Santa Maria Hospital, Lisboa	Lisboa	3.8	82.3	17.7
Centro De Neurociencias e Biologia Celular, Univ Coimbra	Coimbra	3.5	13.8	86.2
Universidade de Aveiro	Aveiro	3.3	26.5	73.5
Instituto Superior Técnico	Lisboa	2.6	22.0	78.0
Universidade de Lisboa	Lisboa	2.5	62.4	37.6
Universidade de Coimbra	Coimbra	2.4	35.9	64.1
Faculdade de Farmácia, Universidade de Lisboa	Lisboa	2.3	9.4	90.6
Universidade Nova de Lisboa	Lisboa	2.2	53.1	46.9
Universidade Fernando Pessoa	Porto	2.2	46.9	53.1
Universidade da Beira Interior	Covilha	2.2	37.8	62.2
Champalimaud Clinical Centre	Lisboa	2.0	85.4	14.6
Universidade de Coimbra, Faculdade de Farmácia	Coimbra	1.9	15.6	84.4
Universidade do Algarve	Faro	1.9	42.2	57.8
Hospital Curry Cabral	Lisboa	1.9	96.1	3.9
Hospital Santo António	Porto	1.5	83.0	17.0
NOVA Univ. of Lisboa, Faculty of Medical Sciences	Lisboa	1.5	64.3	35.7

Table 9. Top national institutions publishing cancer research (2010-2020): *top 25 institutions publishing cancer research in Portugal, data as:* ***1.*** *percentage of total publications in cancer research in Portugal;* ***2.*** *percentage of institutional publications in public health, epidemiology and clinical research in Portugal and* ***3.*** *percentage of institutional publications in basic and translational research.*

Furthermore, when looking at the dedication of the major actors according to the "type" of

[75] Multiple affiliation issues are probably inflanting the prevalence of University of Porto in this list (since publications from all its associated institutions are most likely also counted as Universidade do Porto); with that said, however, University of Lisboa does not appear so dedicated to Cancer Research, even if several of its institutions are.

research, we find that many of the larger actors cover the *spectrum* of research (e.g. Universidade do Porto, Universidade do Minho). However, it can be observed how clinical actors concentrate in three main portuguese cities: Porto, Lisboa and Coimbra (in the IPOs delegations and major urban centers hospitals). Clinical research depends more strictly on patient access and public health care facilities, and is therefore not surprising that more basic and translational research can be seen in a more diverse array of locations and actors (e.g. Universidade de Aveiro), even if several represent fundamental and/or translational research centers that have an undeniable connection to clinical hubs (e.g. i3S-Porto, CNBC-Coimbra, IMM-Lisboa).

In *Fig. 22* presented below, we can observe the strongest collaborations[76] between the main Portuguese institutions publishing in the cancer field. A close collaboration can be observed between **Universidade do Porto - IPO de Francisco Gentil Porto, or Universidade de Porto - Universidade do Minho**.

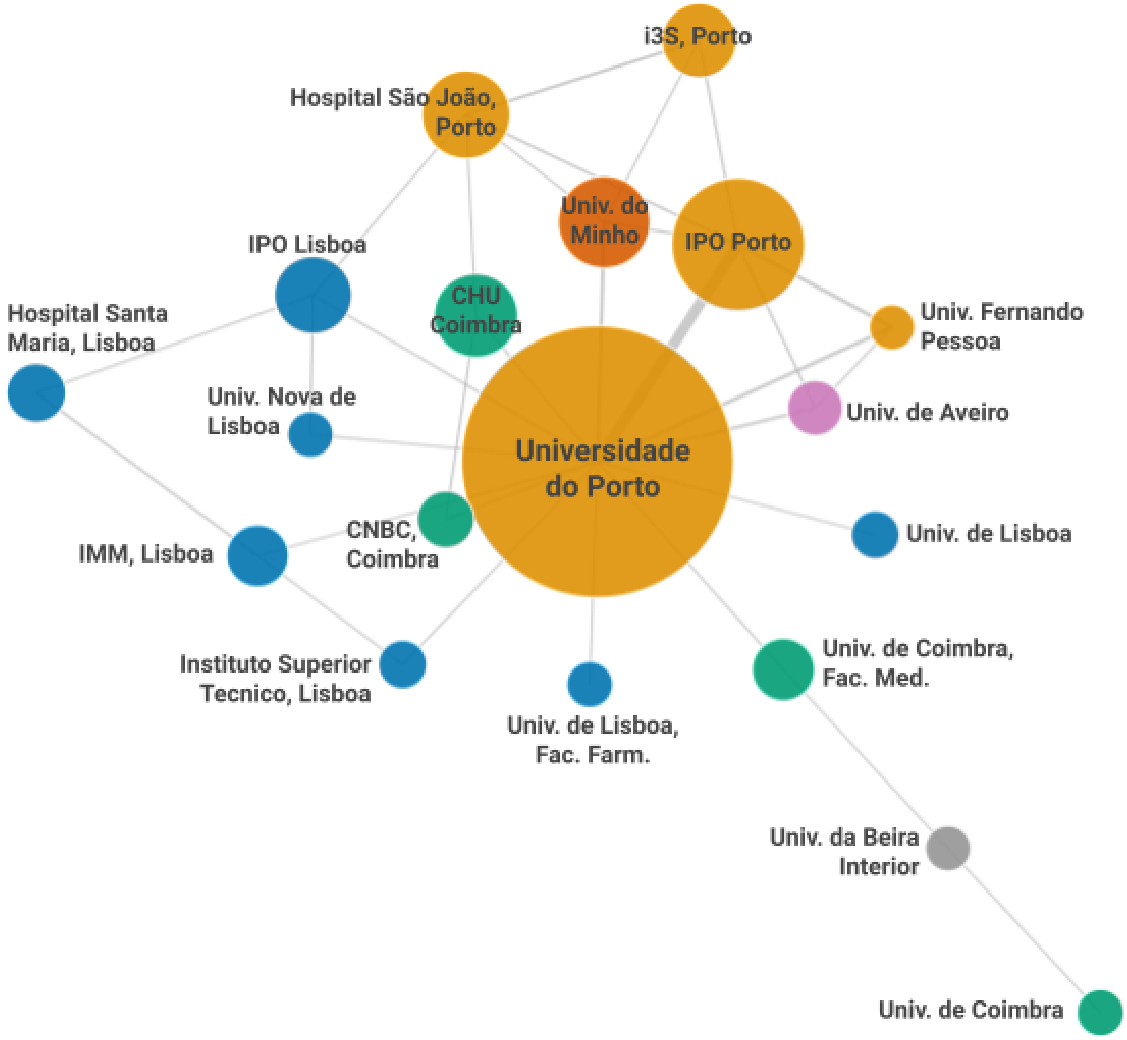


Figure 22. Major national institutions collaborative network in cancer research (2010-2020)*: (connectors width is related to the number of publications co-authored between the 2 institutions; and bubble size relates to total publications in each hub)*

[76] Virtually all institutions depicted have some degree of collaborative work. In this representation we are only depicting the more prevalent: the connections which include more than 15 publications.
For clarity, excluded from the network where obvious multiple affiliation cases (e.g. I3S Porto - Univ Porto; Hospital São João - Univ Porto; iMM - Univ Lisboa)

Interestingly, there is a **disconnection between the size of the hub and its European competitiveness**, since the main competitive national hub in Portugal is Lisboa, followed by Porto and Aveiro but the institutions that have received a larger share of Portugal's H2020 funding for cancer research were the **Instituto de Medicina Molecular, Fundação Champalimaud and the Universidade Nova de Lisboa** instead (*Fig. 23*). The majority of funding was awarded to public entities (~80%) compared to the funding distributed to private institutions (19.5%).

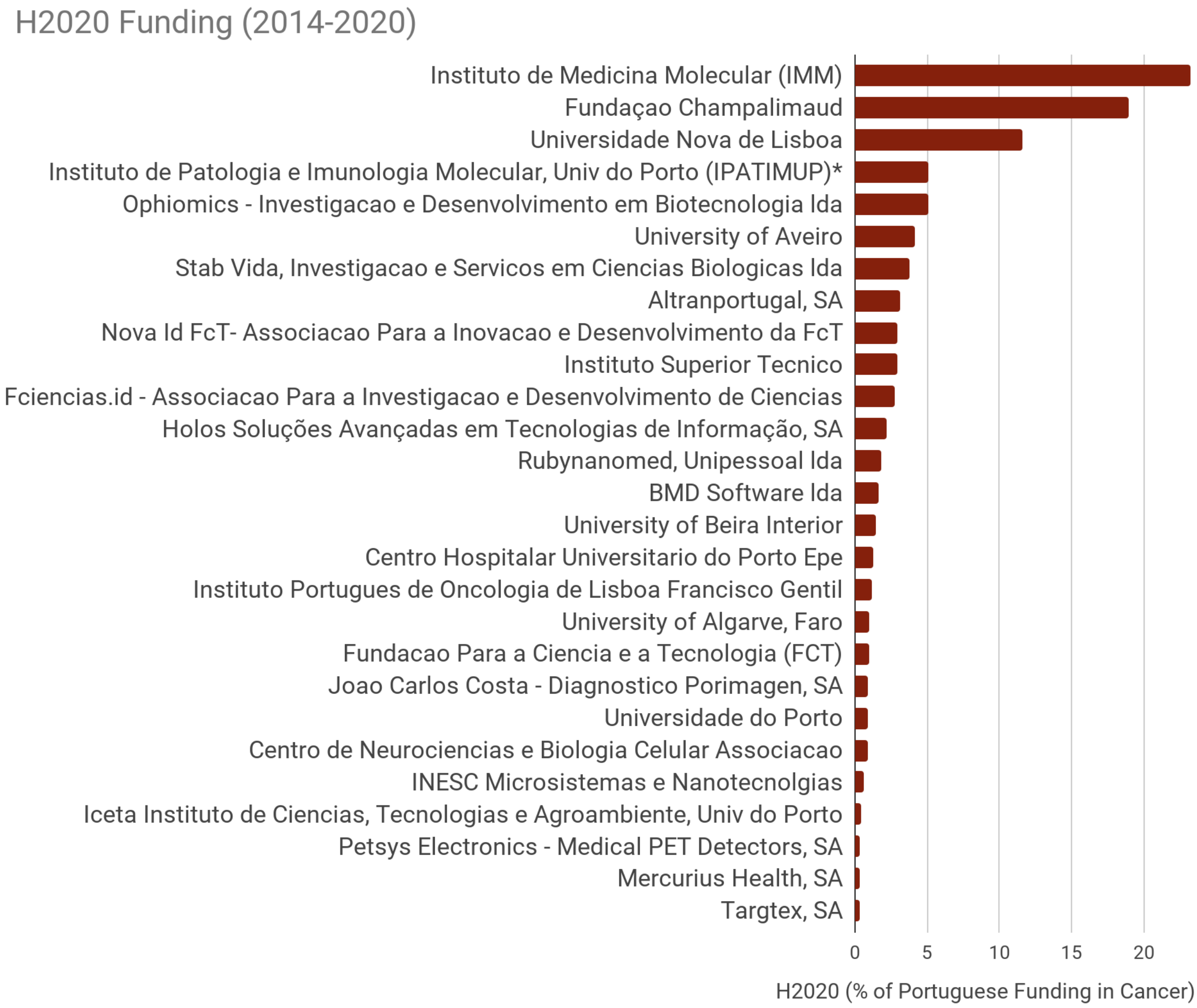


Fig. 23. Major Portuguese beneficiaries of H2020 funding in cancer research (2014-20202): *percentage of the Portuguese funding granted to cancer research in the H2020 program. * now i3S-Porto.*

Looking into individual institutions, the case of IMM versus Ipatimup/i3S is interesting, as both produce highly in cancer research and at similar levels (*Table 9*), and yet IMM shows a substantially higher European competitiveness. One possible contributor for this difference could be the existence of a dedicated office at the institutional level that support the researchers in applying and managing funding applications[77]. Indeed, IMM has a dedicated

[77] However, the existence of dedicated offices that would, in theory, be able to provide support in funding applications, do not per se guarantee increased institutional competitiveness. A more in-depth analysis would be needed to ascertain if this is indeed the cause of the difference observed.

office organised in a streamlined process in which pre-award support is linked to project management, as well as to other financial affairs, in a 360 degrees approach to institutional financial management. At i3S, although there are several transversal units and support in securing successful applications may exist, none of the units seems to be dedicated to award application and management. However, other factors are also to be considered such as the long and consolidated trajectory (both at scientific level and quality of research innovation management) of IMM versus i3S - a newer institution, which may have allowed IMM to develop and optimise its organisational model.

Of the remaining academic institutions in the top positions: Fundação Champalimaud has a dedicated office for sponsored programmes, divided into pre and post-award; Universidade Nova de Lisboa includes several individual institutions (ITQB, CEDOC, FML), some of which have a science funding office (ITQB), while others seem to have some level of *ad-hoc* support, but that is not organised as a dedicated unit (CEDOC; FML).

In Aveiro University there is a research support office, but considering the size of the university, the different lines of research and the fact that there may be other units at the different university sites, it cannot be fully compared to smaller/more specialised institutions.

Portugal is following global trends towards intense globalization of research.

In line with this and in parallel to its growth and development, **Portugal has lost a small share of leadership in published research as a trade-off to increasing its internationalisation**. For a small and emergent system like Portugal to become consolidated this might be key. On these aspects, **the consolidation of Portuguese cancer research is more likely to come in a highly dynamic and collaborative model, which increases visibility, peer recognition and overall competitiveness in performance.**

The Portuguese ecosystem in **cancer research presents main hubs around major cities and most research-intensive universities**. Important to note the prominence of the Porto hub, with a more concentrated ecosystem around Universidade do Porto and IPO Porto; followed by Lisboa with a more scattered network of different agents (Universidade de Lisboa, Universidade NOVA de Lisboa, IPO Lisboa, Fundação Calouste Gulbenkian and Fundação Champalimaud research centers). As expected, Coimbra is the third major hub, followed by the Minho region, with recent substantial growth.

On the more clinical side of research, main national hospitals and IPOs have a key role. However, Portugal counts with a subset of more **fundamental and translational research centers,** which, **despite not being specialized in cancer research *per se*, contribute greatly to the Portuguese ecosystem and some are highly competitive in EU calls** (e.g. IMM and Champalimaud research center).

VI. Main Messages and Insights

Overall, from this analysis it is clear that cancer research in Portugal is in a positive evolution in several dimensions. Some areas of knowledge creation and applicability are growing healthily, while others are still underperforming. Below, we provide a summary table of the main findings, highlighting those areas that are experiencing sustained growth and should be supported and enhanced, and those where attention should be put in.

Portugal's cancer research snapshot:

Dimension	Current status
Bibliometric Production	**Expanding**
Scientific Impact (as in citations)	**Significant** (helped by the fact of being a small ecosystem)
Scientific Recognition (as in the capacity to publish in top tier journals)	**Not aligned with research excellence (below its potential)**
"Type" of research	**Very significant evolution of basic & translational**
	To a lesser extent, clinical and public health research
European Funding	**Expanding**
Clinical trials	**Underdeveloped**
Technology transfer	**Underdeveloped**
Ecosystem	**Emergent**

In the following table we highlight the main findings and what could be possible avenues for the future for the dimensions analysed:

Main Insights and Potential Paths Forward:

Dimension	Main insights and findings	Potential future avenues of improvement
Bibliometric Production	Although significantly increasing in the biomedical domain, Portugal is not specialized in cancer, where it is still below world average. Cancer is a relatively newer area of biomedical research in Portugal, where we observe a predominance of other areas such as neurosciences and infectious diseases. However, production in Cancer research per capita is already in line with Spain and France.	A hyper specialized system is not a mandatory condition for excellency (as seen in this analysis). **Keeping a balanced distribution between different biomedical domains** (granted that they are equally supported and performing) may be preferable, as success often comes from diversity, since it increases a system adaptability to new conditions and challenges.
Scientific Impact and Recognition	Significant growth in citations metrics (despite low dedication) but still striving for visibility and recognition, as seen by a low rate of publications in top tier journals. Portugal is increasingly competitive in EU funding, especially when considering the country's small size.	The high scientific impact of cancer research is not aligned with equally high scientific recognition and visibility, which have the grounds to expand. Although it is common for an emerging ecosystem of a small size to take time to be recognised, cancer researchers, institutions and other actors should be aware of and **show the excellence of its research and aim for higher visibility and recognition.**
Balance and performance by Research "Type"	Cancer Basic and Translational research has grown and gained excellence. Clinical and Epidemiological and Public Health Research, even if under-supported, achieve significant bibliometric results.	**Keep the current support and practices** that are allowing fundamental research to thrive. **Develop new support instruments and facilities that should be put to use for clinical research**, paying attention to the balance of the system and keep supporting all research types.
Clinical trials	Despite a significant growth in the last decade, Portugal has a clear underperforming clinical trials environment, heavily reliant on industry sponsorship and interests. However, clinical trials with non-industry sponsors have been increasing in the last few years. Most clinical trials are Phase 3 multisite trials, not led by Portugal, in contrast to a more balanced profile in all benchmark countries (with the exception of Ireland). Some specific instruments are already being put in place by the government (*e.g.* AICIB). The few early phase and investigator-initiated trials do show a diverse pattern of medical interests that could be further exploited.	Support a balanced clinical trials ecosystem, with the **creation of conditions for investigator-initiated trials, early phase trials and those sponsored by non-corporate organisations**. **Support clinical research and clinical trials led nationally**, as it will increase the visibility of the Portuguese system, its clinicians and researchers, attract external investment and contribute to the development of the ecosystem as a whole.
Technology transfer	Patentable innovation is an area where Portugal's cancer research does not fare well; while SMEs in Portugal are barely able to attract EU funding.	Although many factors are likely at play, efforts should be taken in **identifying the barriers and challenges** and how to overcome these.
Ecosystem	Portugal has, overall, an emergent profile, with growing production and performance in several dimensions.	Smaller ecosystems are not necessarily at a disadvantage (concentration of citations, for example, can be higher than in larger systems

	It also follows the emergent trends towards globalization, with a corresponding decrease in leadership as trade-off of increased internationalisation	where scientific impact can suffer dilution) , however they should follow models that adjust to their size. Portugal could follow a similar model to Belgium, in which there is a **choice of international collaboration,** specific focus in nurturing strong and far-reaching collaborative and partner networks, while maintaining clear national excellence**.**

VII. Methodology

Semantic Techniques and resources used:

- ### *Controlled Vocabularies*

In this study two controlled vocabularies were mainly used to identify research related to Cancer and, inside that specific domain, research which is more on Clinical Research, Epidemiology and Public Health area was further classified.

- *What is a controlled vocabulary (VOC)?*

For a particular area of interest, controlled vocabularies are created by or with field experts and are used to identify the documents pertaining to a given research domain. **Ideally, they are composed of unequivocal terms that fully represent a specific area** (more complex VOCs may work in the intersection of two others, specially for multidisciplinary fields like biomedical engineering).

These VOCs are either based on existing thesaurus (as the MeSH Taxonomy, see next section) or keywords selected from seed corpus (ie, from texts that specifically belong to a certain domain, most often used words -which do not appear in random texts at the same rate- are chosen and manually validated).

With our VOC we **scan the documents**[78], in order to classify them as pertaining or not to the domains in study. SIRIS Academic has developed several controlled vocabularies from biomedical sciences (cancer, neurosciences, infectious diseases, etc), as well as, from other domains like AI, engineering, oceanography and cultural heritage. Some of them are more complex semantic analysis, in which we scan titles, abstract and keywords of publications and research projects.

VOCs designed specifically for this study:
- Cancer research
- Public health, epidemiology and clinical research

- *Conceptual definition of the domain*

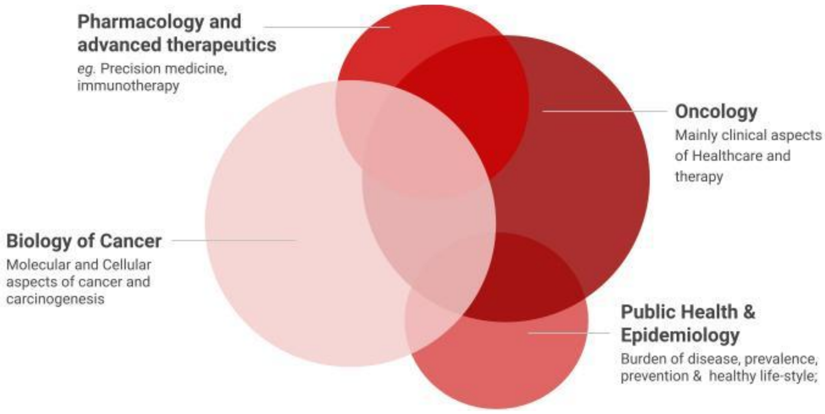
This exercise is made before looking for resources to construct a VOC. The intention is, together with experts[79], to have a first conceptual definition of which are the main areas that compose a given domain. With this established, it results easier to ensure that the VOC is designed to be the most inclusive, while being fully aware of its boundaries (e.g. processes such as cell cycle or cell division biology are not spe included into the Cancer Research VOC - unless it is specifically targeting the cellular process in cancer cells.

The controlled vocabulary designed for cancer research is relatively universal since it is based in worldwide publications. For this study, an *ad-hoc* controlled vocabulary by research "type" was constructed based on Portugal's publications which, therefore, may not apply to very particular or distant systems.

[78] More often Title+Abstracts+Keywords of publications and research projects.
[79] In this study this step was performed with the ASPIC team.

Since there is a good degree of grey areas in the definition of research "types" - commonly: fundamental, basic, translational, preclinical, clinical, epidemiological, public health research, etc.- we kept a simple structure by dividing cancer research types in two main groups:

- Basic and translational research
- Public health, epidemiology and clinical research



Cancer Research Conceptual Map

Public Health, Epidemiology and Clinical Research	Translational and Basic Research
✓ Clinical trials ✓ Research regarding patients treatment protocol ✓ Research implicating patients directly ✓ Research with patient samples as central feature (genetics of disease, biomarkers, prognostic markers, diagnostic from the application point-of-view) ✓ Epidemiology, Global Health & Public Health ✓ Medical guidelines ✓ Mental Disorders ✓ Cultural/socioeconomic impact on Health ✓ Health Literacy, Health promotion & QOL ✓ Health Education ✓ Health Policy	✓ Study of processes or diseases with the *intent to treat* ✓ Drug and vehicle development (since they have a therapeutic target) ✓ Preclinical models (including advanced ones like sheep and pig) ✓ With patients samples only as proof of concept, as in tumor samples/biobank usage which is not central to the paper; or to establish research pre-clinical models (*e.g* cell lines, organoids) ✓ Tissue, Cellular & Molecular basis of disease ✓ Tissue, Cellular & Molecular understanding of biological mechanisms ✓ Development of applied techniques - at protein, chemistry, molecular, or cellular level

Research "type" Conceptual Map

○ *Mechanism of action for document identification*

A controlled vocabulary performs under a set of filtering rules specific for each document type and aimed at including possible variants of the same concept (e.g. word permutations within the concept).
Ideally, controlled vocabularies work on a simple basis: the presence of a keyword from the VOC identifies the document as pertaining to that perimeter. However, in our experience, most sophisticated VOCs have their own particular mechanism of action, linked to the complexity and/or multidisciplinarity of the perimeter being drawn. For the VOCs in this study, keywords of interest where segregated into different categories and, a combination of these conditions was used to categorize documents (below we provide an example of what

was performed for the research "type" public health, epidemiology and clinical research, as described in Brief Methodological Notes):

- Unequivocal: the presence of **one** keyword from this list identifies the document as pertaining to the perimeter; these are responsible for the majority of the identified documents (For instance, 99% in the Cancer and 86% in the public health, epidemiology and clinical research vocabularies).
- Relevant: terms that are common, but not unequivocal for this perimeter. The presence of **two or more** keywords from this list identifies the document as pertaining to the perimeter, (e.g. when the keyword *Biopsy* and *Child* is detected in a document).
- Significant: very common and broad terms, that would lead to a very high rate of false positives by themselves. In this case, the presence of **one "relevant" and one "significant"** keyword identifies the document as pertaining to the perimeter (but 2 significant terms are not a sufficient condition); (e.g. when the keyword *Biopsy* and *Genetic predisposition to disease* is detected in a document).
- Excluding: on the reverse, the presence of **any** term for this list, **automatically excludes** that publication from analysis, regardless of other significant terms (e.g. when the keyword *animals* is identified, the publication is automatically excluded).

Sample of Useful MeSH Terms	Cancer Publications in Portugal (n, 2010-2020)	Keyword categories			
		Unequivocal	Relevant	Significant	Excluding
Humans	6922			TRUE	
Aged	1830		TRUE		
Animals	1307				TRUE
Cell Line, Tumor	924				TRUE
Aged, 80 and over	786	TRUE			
Prognosis	607	TRUE			
Retrospective Studies	564	TRUE			
Treatment Outcome	528	TRUE			
Mice	463				TRUE
Risk Factors	413	TRUE			
Neoplasm Staging	370	TRUE			
Child	310		TRUE		
Diagnosis, Differential	309	TRUE			
Antineoplastic Combined Chemotherapy Protocols	268	TRUE			
Follow-Up Studies	253	TRUE			
Neoplasm Recurrence, Local	245	TRUE			
Tomography, X-Ray Computed	241			TRUE	
Survival Rate	240	TRUE			
Genetic Predisposition to Disease	237		TRUE		
Magnetic Resonance Imaging	234			TRUE	
Time Factors	210			TRUE	
Biopsy	201		TRUE		

Example of MeSH terms used for the Public Health, Epidemiology and Clinical Research VOC, according to the four keyword categories designed

Controlled Vocabulary	*Source of Keywords*	*Categories of Terms*
Cancer Research	Medical Subject Headings (MeSH)	■ Unequivocal ■ Relevant ■ Excluding
Public Health, Epidemiology and Clinical Research	Most common keywords in *Title & Abstract & Author Keywords* of Seed Corpus (exclusively biomedical research)	■ Unequivocal ■ Relevant ■ Significant

- *Iteration and quality tests*

In order to minimize error, the corpus of publications detected by the VOCs are submitted to several rounds of testing linked to iterations (normally between five and ten) of the designed vocabularies. Random samples of 100-200 publications are manually revised for the detection of false positives and false negatives; This exercise often leads to changes in the VOC aimed at "fixing" false cases.

In the table below we show the quality achieved in this study for the identification of Scopus publications:

Corpus of documents	Text resource	VOC	% false positives	% false negatives
World Publications	Medical Subject (MeSH) terms present in PubMed Database	Cancer Research	4-5%	1-2%
World Cancer Publications	(MeSH) terms + Title + Abstracts	Public Health, Epidemiology and Clinical Research	5%	2-3%

We downloaded worldwide publications in Cancer Research between **2010 and 2020** from Scopus, Elsevier's bibliometric database (around 1M). The information available for each document includes its authors and affiliations[80], citation numbers, MeSH terms associated and classification into subject areas (through the publication's journal), among others.

- ## Restricted Keyword search for specific topics (Forefront Topics in cancer research)

In some instances a complicated vocabulary is not needed, and semantic techniques can be used, even if using a small number of keywords.

In order to identify publications related to specific forefront topics, the following list of topics and keywords was generated. It does not aim at being an exhaustive exploration of these topics, but rather an indication of how Portugal's cancer research is aligned with those. Publications containing any of the keywords listed below (present in Title/ Abstract) were

[80] Databases of this kind attempt to disambiguate any given institution by linking all of their name variants and/or sub units to a single profile in the system, in order to account for both hierarchical structures and the different ways in which authors list their affiliations, be it multiple languages, acronyms, or other sources of variation.

considered.

Forefront Topics	*keywords*
Precision Oncology	Precision|targeted *Extra keywords that need to be in the same sentence*: medicine|therapy|therapies|treatment|treatments|Oncology|therapeutic|care|drug|drugs
personalized medicine	Personalized|individualized *Extra keywords that need to be in the same sentence*: medicine|therapy|therapies|treatment|treatments|Oncology|therapeutic|care|drug|drugs
Integrated care	integrated care
Immuno-Oncology & Immunotherapy	immunotherapy|checkpoint inhibitor|CarT|PD-1|PD-L1|PDL1|immuno-oncology
Drug Discovery & repurposing	drug discovery|drug screening|drug repurposing|Drug repositioning|biosimilars
Diagnostic & Biomarkers	biomarker|digital pathology|artificial intelligence|machine learning|liquid biopsy|computacional sciences|genetic testing|genetic screening|susceptibility genes|cancer risk|genetic testing|predictive marker|predictive biomarkers|prognostic biomarker|prognostic marker|deep learning
Drug Resistance and tumor heterogeneity	drug resistance|drug resistant|tumor heterogeneity|tumor evolution|tumor microenvironment|metastatic microenvironment
Oncogenomics	next-generation sequencing|oncogenomics|genomics|mutational burden|mutation rate
Quality of life	quality of life|QOL|caregiver|cancer survivor|survivorship|long term care|palliative care
eHealth	eHealth|Telemedicine|digital health|telehealth|mobile health|mhealth
Artificial Intelligence	artificial intelligence|machine learning|deep learning

Keywords per Research Topic, to be identified in Cancer Research

- ### *Zero Shot Technique (classifying research projects)*

Due to the different tone in which research projects are written (more aspirational and mentioning global challenges, than focused on its results), their classification into a specific domain is often more complicated (versus publications). For this reason we used an additional technique to avoid mainly false positives issues: *ZeroShot*.

Zero Shot is an emerging research area of Machine Learning which allows to perform predictive functions on whatever kind of data it has never seen before[81]. Using this peculiarity in the case of Text Classification, Zero Shot promotes the creation of models able to classify general texts without having an accurate knowledge of a specific domain, but knowing something semantically similar.

The technique used in the case of the classification of research projects in the "Cancer Research domain" involves a pre-trained Zero Shot model[82] trained on billions of texts, able to provide a score (0-1) of the probability that a given project is related to Cancer Research.

[81] Zero-data Learning of New Tasks: https://www.aaai.org/Papers/AAAI/2008/AAAI08-103.pdf
[82] Zero-shot Text Classification via Reinforced Self-training: https://aclanthology.org/2020.acl-main.272.pdf

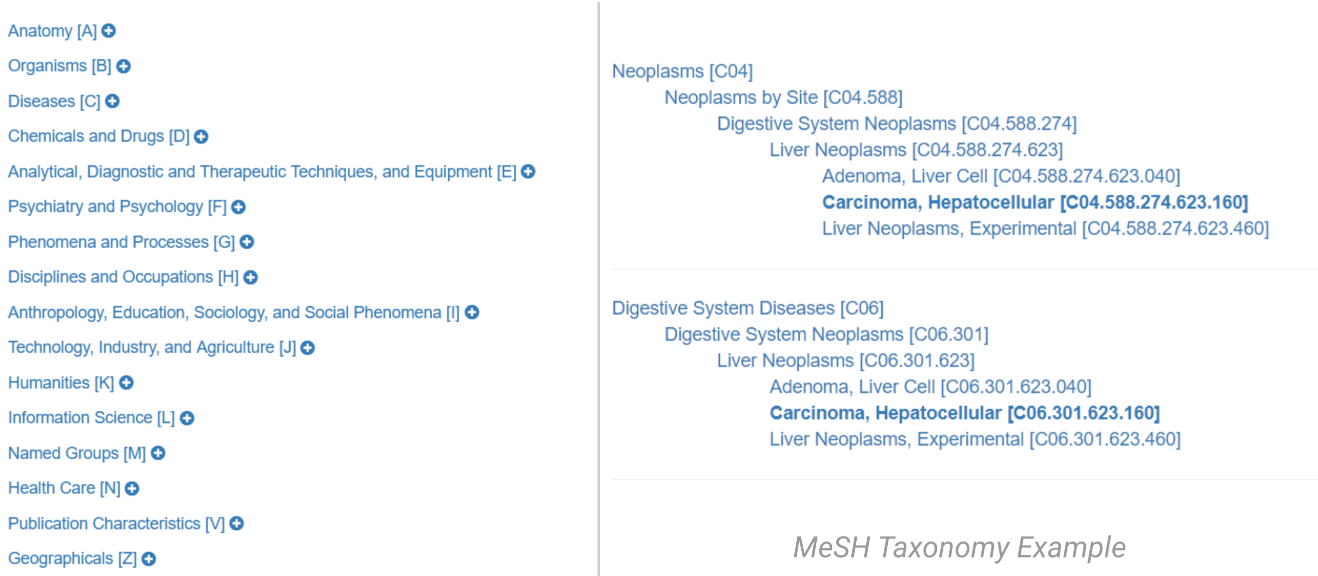
Corpus of documents	Analyse text	Classification conditions	% false positives	% false negatives
CORDIS Database (EU funded research projects)	Title + Abstract	1º have a threshold of Zero Shot Cancer Research probability >0.7	4-5%	1-2%
		2º be classified by the Cancer Research VOC		

- ## *Medical Subject Headings NLM-NIH (MeSH)*

In order to construct a biomedical controlled vocabulary that could render the best results, we identified the Medical Subject Headings (MeSH) as the most suitable terminology. Medical Subject Headings (MeSH) is a hierarchical vocabulary and taxonomy curated by the National Library of Medicine (NLM) of the NIH (National Institutes of Health) of the United States)[83]. It is used to index, catalog and search for biomedical and health-related information. MeSH includes keywords listed in the catalog (MEDLINE / PubMed) and other National Library of Medicine databases such as clinicaltrials.gov.

The MeSH taxonomy is curated by experts and is updated annually to include the latest developments in the field of health and biomedicine. Its structure includes a branch that reports on disease-related terms (branch C/ F03), which enables us to trace which diseases and disease groups are referenced in a text. Its hierarchical tree structure allows the indexing of terms in different branches of the taxonomic tree in order to better describe the domain (for example, *Hepatocellular Carcinoma* appears as a Neoplasm and as a Digestive Disease - see Figure below).

The major limitation is that not all relevant publications are indexed in the Pubmed repository and, even in those publications registered, there is often a time lag between the publication date and the availability of all variables associated with a given publication like the MeSH terms.



MeSH Taxonomy Example

[83] Medical Subject Headings - Home Page (nih.gov)

In the present study, there were 2 main usages of the MeSH taxonomy:

- As a source of validated and expert-derived keywords (terms related to cancer research that allow us to construct the VOC and identify documents referring to the field)
- Based on the MeSH terms assigned to each publication in Pubmed and clinical trials in ct.gov,, perform an analysis of predominant neoplasms in research efforts.

VIII. Annexes

A. Publications and Clinical Trials according to Cancer groups

By classifying publications by cancer sites according to the NIH-NLM MeSH Taxonomy[84] and divided per research "type", we can appreciate Portugal's pattern in cancer research interests at a different level. Overall, Digestive System tumors are those deserving more attention (by grouping stomach, colorectal, liver, etc); followed by Urogenital (cervix, uterus, prostata, urinary bladder, kidney, etc) and Breast. Overall, basic and translational research has a pattern which deviates from the clinical and public health research in a couple of instances, but more predominantly in its low dedication to endocrine and head and neck tumors (thyroid cancers are included in both categories).

Publications per Cancer Groups * (%)

	EU27+UK	***Portugal***		
MeSH Heading	***All publications***	***All publications***	**Public Health, Epidemiology & Clinical Research**	**Basic & Translational Research**
Digestive System Neoplasms	18.4	18.9	21.2	15.5
Urogenital Neoplasms	14.0	14.3	16	11.7
Breast Neoplasms	11.0	13.1	13.6	12.4
Endocrine Gland Neoplasms	8.8	7.8	9.2	5.6
Head and Neck Neoplasms	8.2	7.8	10.7	3.4
Thoracic Neoplasms	7.6	6.8	8.5	4.3
Nervous System Neoplasms	6.5	5.4	6.1	4.2
Skin Neoplasms	4.0	4.3	5.3	2.7
Bone Neoplasms	2.9	2.5	2.9	1.8
Hematologic Neoplasms	1	0.5	0.6	0.4
Abdominal Neoplasms	0.9	0.3	0.4	0.2
Eye Neoplasms	0.8	0.6	0.8	0.3
Soft Tissue Neoplasms	0.7	0.6	1	0.2
Mammary Neoplasms, Animal	0.2	0.4	0	1
Pelvic Neoplasms	0.1	0.1	0.2	0
Splenic Neoplasms	0.1	0.1	0.1	0
Anal Gland Neoplasms	0	0	0	0

**according to the NIH-NLM MeSH Taxonomy as Neoplasms by site (Neoplasms by Site [C04.588])*

Portuguese Research by cancer groups, and by clinical/public health versus basic/translational cancer (2010-2020): *publications classified by MeSH terms according to their site, overall and as percentage of the total research per "type".*

In terms of cancer groups, the Digestive System, Urogenital, Breast and Thoracic neoplasms clinical trials represent the top four researched neoplasms across the benchmark selection, which is aligned with the EU27+UK pattern.

[84] MeSH Browser (nih.gov) - (Neoplasms by Site [C04.588])

MeSH Heading	Clinical Trials per Cancer Group* (% of national trials) EU27+UK	Portugal	UK	NL	France	Spain	Italy	Ireland	Belgium	Germany
Digestive System Neoplasms	18.4	13.5	13.6	15.0	14.6	15.1	15.9	13.8	14.3	14.6
Urogenital Neoplasms	13.3	10.4	15.8	14.7	11.4	11.5	10.9	15.6	12.5	10.8
Breast Neoplasms	11.6	19.8	10.5	8.9	11.1	13.3	9.8	16.4	13.1	9.7
Thoracic Neoplasms	9.8	13.7	11.2	13.0	10.8	14.4	11.8	13.5	11.1	12.1
Endocrine Gland Neoplasms	7.2	5.8	6.5	7.3	6.0	6.4	7.6	6.0	7.2	6.3
Head and Neck Neoplasms	6.3	5.1	5.5	7.1	6.3	4.3	4.5	6.0	5.9	5.3
Nervous System Neoplasms	2.0	1.0	1.5	2.0	2.2	1.1	1.4	0.5	1.6	2.1
Myeloproliferative Disorders	1.9	2.7	2.7	2.6	2.6	2.3	3.4	3.1	2.7	3.4
Hematologic Neoplasms	1.7	1.4	2.2	1.8	2.2	1.8	2.2	2.1	1.6	1.9
Abdominal Neoplasms	0.6	0.2	0.2	0.6	0.5	0.3	0.3	0.0	0.8	0.5
Bone Neoplasms	0.5	0.0	0.2	0.6	0.6	0.2	0.5	0.5	0.2	0.2
Skin Neoplasms	0.4	0.0	0.5	0.2	0.3	0.2	0.3	0.8	0.3	0.2
Eye Neoplasms	0.3	0.0	0.2	0.4	0.4	0.3	0.1	0.0	0.1	0.3
Pelvic Neoplasms	0.1	0.0	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.1
Soft Tissue Neoplasms	0.1	0.0	0.1	0.1	0.1	0.1	0.2	0.5	0.1	0.1
NA	36.1	34.9	39.5	37.6	40.3	39.2	40.8	32.3	39.3	41.9

*according to the NIH-NLM MeSH Taxonomy as Neoplasms by site (Neoplasms by Site [C04.588])

Cancer Clinical trials by cancer group in EU27+ UK and the panel of countries (in clinicaltrials.gov, started between 2010-2020): *percentage of trials of the region or country classified by their MeSH terms into a given cancer site*.

Alignment between WHO/Globocan Cancer sites and MeSH Neoplasms used in Figure 15:

WHO Cancer Sites	MeSH Neoplasms
Breast	Breast Neoplasms [C04.588.180]
Prostate	Prostatic Neoplasms [C04.588.945.440.770]
Lung	Lung Neoplasms [C08.785.520]
Colon	Colonic Neoplasms [C06.405.249.411.307.180];
Rectum	Rectal Neoplasms [C04.588.274.476.411.307.790]
Stomach	Stomach Neoplasms [C04.588.274.476.767]
Bladder	Urinary Bladder Neoplasms [C04.588.945.947.960]
Non-Hodgkin lymphoma	Lymphoma, Non-Hodgkin [C04.557.386.480]
Pancreas	Pancreatic Neoplasms [C04.588.322.475]
Thyroid	Thyroid Neoplasms [C04.588.322.894]
Liver	Liver Neoplasms [C06.301.623]
Leukaemia	Leukemia [C04.557.337]
Corpus uteri	Endometrial Neoplasms [C13.351.500.852.762.200]; Carcinoma, Endometrioid [C04.588.945.418.948.585.124]
Kidney	Kidney Neoplasms [C12.758.820.750]
Brain, Central Nervous System	Central Nervous System Neoplasms [C10.551.240]
Lip, oral cavity	Mouth Neoplasms [C04.588.443.591]
Melanoma of skin	Melanoma [C04.557.465.625.650.510]; Skin Neoplasms [C04.588.805]
Multiple myeloma	Multiple Myeloma [C04.557.595.500]
Cervix uteri	Uterine Cervical Neoplasms [C04.588.945.418.948.850]
Oesophagus	Esophageal Neoplasms [C04.588.274.476.205]
Ovary	Ovarian Neoplasms [C04.588.322.455]

B. Main Cancer MeSH Terms identified in Publications

i) Most common MeSH Terms in Cancer Publications *(excluding Neoplasm terms)*

	Publications (n, 2010-2020)		
MeSH Term	**EU27+UK**	**World**	**Portugal**
Neoplasm Staging	13,430	45,827	237
Neoplasm Grading	4,625	14,139	82
Tumor Microenvironment	4,081	12,670	84
Xenograft Model Antitumor Assays	3,890	21,281	75
Tumor Burden	3,191	12,189	63
Antineoplastic Agents	1,144	4,463	51
Drug Screening Assays, Antitumor	1,139	4,901	49
Cancer Vaccines	834	2,972	16
Tumor Escape	519	1,431	3
Tumor Hypoxia	243	783	3
Antigens, Neoplasm	165	308	2
Antineoplastic Agents, Phytogenic	154	974	5
Carbonic Anhydrase IX	108	131	0
Biomarkers, Tumor	101	411	8
Human papillomavirus 16	70	180	0
Human papillomavirus 18	62	141	0
Topoisomerase II Inhibitors	55	167	2
Papillomaviridae	51	126	0
Angiogenesis Inhibitors	48	163	2
Neoplastic Stem Cells	47	145	3
Antibiotics, Antineoplastic	46	193	3
Receptor, ErbB-2	44	159	1
Tumor-Associated Macrophages	35	120	0
Topoisomerase I Inhibitors	34	120	1
Carcinogens	34	78	1
Antineoplastic Agents, Immunological	29	69	3
Polyomavirus	29	68	2
Herpesvirus 4, Human	27	82	0
Antigens, Viral, Tumor	27	72	0
Antimetabolites, Antineoplastic	23	66	2
BK Virus	22	84	1
Carcinogenicity Tests	22	40	1
Herpesvirus 8, Human	21	59	2
JC Virus	20	75	0
Early Detection of Cancer	20	54	0
CA-125 Antigen	19	68	2
Ki-67 Antigen	18	84	1
Proliferating Cell Nuclear Antigen	15	88	1
CA-19-9 Antigen	15	47	1
Anticarcinogenic Agents	15	42	0
Mucin-1	13	57	3
Carcinoembryonic Antigen	13	132	2
Poly(ADP-ribose) Polymerase Inhibitors	13	56	1
Antineoplastic Agents, Alkylating	13	30	1
Epithelial Cell Adhesion Molecule	13	51	0
Receptor, ErbB-3	13	44	0

Prostate-Specific Antigen	11	69	1
Topoisomerase Inhibitors	11	43	0
Alphapapillomavirus	11	25	0
Antigens, Tumor-Associated, Carbohydrate	10	31	2
Simian virus 40	10	35	0
Pancreatitis-Associated Proteins	10	24	0
Human papillomavirus 11	9	34	0
CD52 Antigen	9	18	0
Rhadinovirus	8	19	2
Human papillomavirus 6	8	36	0
Human papillomavirus 31	8	12	0
Antimitotic Agents	7	26	1
Lewis X Antigen	7	19	0
Betapapillomavirus	7	8	0
Polyomaviridae	7	8	0
Antineoplastic Agents, Hormonal	6	15	1
alpha-Fetoproteins	6	112	0
Sialyl Lewis X Antigen	6	11	0
Neprilysin	6	9	0
Receptor, ErbB-4	5	19	0
Merkel cell polyomavirus	5	12	0
Myeloablative Agonists	5	7	0
Chorionic Gonadotropin, beta Subunit, Human	4	42	0
CD24 Antigen	4	13	0
Immune Checkpoint Inhibitors	4	12	0
gp100 Melanoma Antigen	4	6	0
CD146 Antigen	4	6	0
Gammapapillomavirus	4	6	0
Antigens, Polyomavirus Transforming	3	16	0
Circulating Tumor DNA	3	10	0
Gammaherpesvirinae	3	8	0
MART-1 Antigen	3	6	0
Folate Receptor 1	2	14	1
Embryonal Carcinoma Stem Cells	2	8	1
Tumor Stem Cell Assay	2	16	0
Basigin	2	7	0
Cancer Pain	2	6	0
Carcinogens, Environmental	2	5	0
Bovine papillomavirus 1	2	4	0
Herpesvirus 2, Saimiriine	2	3	0
Mupapillomavirus	2	2	0
Ki-1 Antigen	2	2	0
Myxoma virus	1	1	1
Melanoma-Specific Antigens	1	6	0
Synaptophysin	1	5	0
Adenovirus E1A Proteins	1	3	0
Tissue Kallikreins	1	1	0
Hormones, Ectopic	0	4	0
Tumor Lysis Syndrome	0	4	0
Lymphocryptovirus	0	3	0
Cottontail rabbit papillomavirus	0	2	0

Most common MeSH Terms in the Cancer Publications identified in this study (excluding neoplasms keywords, ie. excluding MeSH heading in branch C04 of MeSH Taxonomy)

ii) Top Most common MeSH Terms in Cancer Publications *(Top 160 MeSH terms)*

MeSH Term	Publications (n, 2010-2020)		
	EU27+UK	World	Portugal
Neoplasms	26602	79958	648
Breast Neoplasms	18793	61085	520
Neoplasm Staging	13430	45827	237
Lung Neoplasms	12184	48366	208
Neoplasm Recurrence, Local	9967	30722	182
Prostatic Neoplasms	9593	26613	216
Colorectal Neoplasms	9555	28234	173
Liver Neoplasms	7556	33623	136
Skin Neoplasms	7211	18577	169
Brain Neoplasms	7091	21918	152
Neoplasm Metastasis	6478	21910	116
Melanoma	6296	15246	115
Adenocarcinoma	6101	21525	144
Carcinoma, Squamous Cell	5865	21630	114
Pancreatic Neoplasms	5125	17664	72
Neoplasm Invasiveness	4826	25179	108
Carcinoma, Non-Small-Cell Lung	4757	19429	58
Neoplasm Grading	4625	14139	82
Ovarian Neoplasms	4540	15098	76
Head and Neck Neoplasms	4285	11970	57
Lymphatic Metastasis	4239	17356	69
Carcinoma, Hepatocellular	4114	23079	60
Tumor Microenvironment	4081	12670	84
Xenograft Model Antitumor Assays	3890	21281	75
Kidney Neoplasms	3510	11762	81
Uterine Cervical Neoplasms	3366	12953	100
Tumor Burden	3191	12189	63
Bone Neoplasms	3101	10269	71
Glioblastoma	3073	9009	74
Colonic Neoplasms	2924	10228	95
Stomach Neoplasms	2876	19082	181
Urinary Bladder Neoplasms	2842	8755	100
Leukemia, Myeloid, Acute	2791	8809	28
Multiple Myeloma	2682	7237	53
Carcinogenesis	2672	11233	81
Esophageal Neoplasms	2583	10399	41
Carcinoma, Renal Cell	2558	8687	63
Glioma	2517	9229	59
Thyroid Neoplasms	2491	9310	88
Rectal Neoplasms	2379	6669	57
Cell Transformation, Neoplastic	2357	7852	57
Carcinoma	2198	8663	71
Adenoma	1837	5628	44
Endometrial Neoplasms	1694	5327	31
Neuroendocrine Tumors	1676	3760	27
Leukemia, Lymphocytic, Chronic, B-Cell	1587	3184	19
Hematologic Neoplasms	1568	4304	21
Squamous Cell Carcinoma of Head and Neck	1451	4359	12

Mouth Neoplasms	1424	6670	30
Carcinoma, Pancreatic Ductal	1411	4502	15
Precursor Cell Lymphoblastic Leukemia-Lymphoma	1406	4721	17
Sarcoma	1404	3803	24
Neuroblastoma	1372	3619	30
Triple Negative Breast Neoplasms	1233	4679	30
Neoplasm, Residual	1209	3144	10
Prostatic Neoplasms, Castration-Resistant	1208	3037	12
Lymphoma	1200	3648	24
Lymphoma, Large B-Cell, Diffuse	1175	3871	22
Peritoneal Neoplasms	1147	3196	5
Antineoplastic Agents	1144	4463	51
Drug Screening Assays, Antitumor	1139	4901	49
Leukemia	1123	3453	15
Neoplastic Cells, Circulating	1095	3160	18
Pituitary Neoplasms	1061	3160	19
Uterine Neoplasms	1053	3532	27
Hodgkin Disease	1014	2340	13
Carcinoma, Ductal, Breast	997	3187	28
Carcinoma, Basal Cell	976	2301	24
Gastrointestinal Neoplasms	969	2869	27
Precancerous Conditions	962	2827	30
Neoplasms, Second Primary	962	2649	21
Testicular Neoplasms	951	2188	28
Leukemia, Myelogenous, Chronic, BCR-ABL Positive	915	2838	31
Mesothelioma	915	2034	4
Meningioma	894	2653	14
Laryngeal Neoplasms	875	2732	20
Bile Duct Neoplasms	872	3442	61
Carcinoma, Transitional Cell	871	2299	25
Carcinoma, Ovarian Epithelial	861	2886	11
Meningeal Neoplasms	852	2555	13
Adrenal Gland Neoplasms	846	2232	20
Cervical Intraepithelial Neoplasia	838	2111	17
Soft Tissue Neoplasms	836	2391	16
Cancer Vaccines	834	2972	16
Neoplasms, Experimental	821	4117	18
Cholangiocarcinoma	793	3071	59
Adenocarcinoma of Lung	786	4631	16
Carcinoma, Papillary	771	3621	32
Osteosarcoma	765	4386	28
Lymphoma, Non-Hodgkin	715	2184	7
Leiomyoma	712	2413	15
Oropharyngeal Neoplasms	703	2015	6
Spinal Neoplasms	700	2311	6
Neoplasms, Multiple Primary	682	2353	20
Lymphoma, B-Cell	640	1783	5
Mesothelioma, Malignant	631	1418	3
Neoplasms, Germ Cell and Embryonal	627	1516	27
Neoplasms, Glandular and Epithelial	616	2134	8
Melanoma, Experimental	592	2570	12
Central Nervous System Neoplasms	587	1792	12
Genital Neoplasms, Female	577	1941	9
Intestinal Neoplasms	572	1355	18

Neoplasms, Radiation-Induced	557	1451	8
Barrett Esophagus	546	1341	11
Thyroid Cancer, Papillary	538	2884	24
Pleural Neoplasms	536	1243	6
Carcinoma, Neuroendocrine	531	1674	18
Carcinoma, Intraductal, Noninfiltrating	526	1641	19
Gastrointestinal Stromal Tumors	524	1929	13
Pheochromocytoma	520	1196	15
Tumor Escape	519	1431	3
Small Cell Lung Carcinoma	505	1988	4
Carcinoma, Lobular	490	1352	16
Lymphoma, Follicular	481	1117	13
Anus Neoplasms	476	1150	15
Thyroid Nodule	468	2016	14
Hemangioma	463	1698	11
Urologic Neoplasms	459	1353	17
Myasthenia Gravis	451	1473	15
Carcinoma in Situ	451	1200	13
Keratosis, Actinic	447	733	4
Neurofibromatosis 1	442	1210	15
Heart Neoplasms	440	1728	23
Neuroma, Acoustic	431	1120	3
Vulvar Neoplasms	431	933	11
Astrocytoma	427	1327	13
Adenocarcinoma, Mucinous	416	1610	7
Uveal Neoplasms	411	908	4
Medulloblastoma	408	1074	12
Sarcoma, Ewing	408	1053	3
Cystadenocarcinoma, Serous	403	1292	6
Neurilemmoma	380	1414	7
Paraganglioma	367	866	12
Cerebellar Neoplasms	362	1053	10
Pregnancy Complications, Neoplastic	361	1026	12
Salivary Gland Neoplasms	360	1386	18
Colorectal Neoplasms, Hereditary Nonpolyposis	359	914	15
Nevus, Pigmented	351	795	5
Adrenal Cortex Neoplasms	349	822	10
Teratoma	346	1301	15
Tuberous Sclerosis	344	942	15
Lymphoma, Mantle-Cell	343	898	8
Mammary Neoplasms, Experimental	330	1602	13
Nasopharyngeal Neoplasms	319	3149	7
Nose Neoplasms	315	1118	11
Neoplastic Syndromes, Hereditary	313	817	10
Leiomyosarcoma	310	902	9
Carcinoma, Endometrioid	296	1041	7
Penile Neoplasms	295	708	11
Precursor B-Cell Lymphoblastic Leukemia-Lymphoma	294	820	4
Skull Base Neoplasms	292	1045	0
Lymphoma, B-Cell, Marginal Zone	290	842	4
Rhabdomyosarcoma	289	804	4
Precursor T-Cell Lymphoblastic Leukemia-Lymphoma	285	741	24
Parotid Neoplasms	284	981	7
Retroperitoneal Neoplasms	282	932	5

Mediastinal Neoplasms	280	1082	5
Tongue Neoplasms	279	1372	4
Carcinoma, Merkel Cell	278	646	3
Adenocarcinoma, Follicular	276	874	13

Top 165 most common MeSH Terms in the Cancer Publications identified in this study

December 2021