



**HAL**  
open science

## Inequities in cancer drug development in terms of unmet medical need

Eliana Barrenho, Réka Halmai, Marisa Miraldo, Iván Tzintzun, Setti Raïs Ali, Léa Toulemon, Jean-Claude Dupont, Lise Rochaix

### ► To cite this version:

Eliana Barrenho, Réka Halmai, Marisa Miraldo, Iván Tzintzun, Setti Raïs Ali, et al.. Inequities in cancer drug development in terms of unmet medical need. *Social Science & Medicine*, 2022, 302, 10.1016/j.socscimed.2022.114953 . halshs-03672739

**HAL Id: halshs-03672739**

**<https://shs.hal.science/halshs-03672739v1>**

Submitted on 22 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## Inequities In Cancer Drug Development in Terms of Unmet Medical Need

Eliana Barrenho\*<sup>1,2</sup>, ORCID: 0000-0002-0084-2959

Réka Halmai<sup>3</sup>, email: [rekahalmi@gmail.com](mailto:rekahalmi@gmail.com);

Marisa Miraldo<sup>1,4</sup>, email: [m.miraldo@imperial.ac.uk](mailto:m.miraldo@imperial.ac.uk); ORCID:0000-0002-5772-7740

Iván Tzintzun<sup>3</sup>, email: [ivan.tzintzun@psemail.eu](mailto:ivan.tzintzun@psemail.eu); ORCID: 0000-0002-7388-5941

Setti Raïs Ali<sup>3</sup>, email: [setti.rais@univ-paris1.fr](mailto:setti.rais@univ-paris1.fr); ORCID: 0000-0002-3371-4769

Léa Toulemon<sup>3,5</sup>, email: [lea.toulemon@ipp.eu](mailto:lea.toulemon@ipp.eu); ORCID: 0000-0002-0424-5924

Jean-Claude K. Dupont<sup>3,6</sup>, email: [jean-claude.dupont@psemail.eu](mailto:jean-claude.dupont@psemail.eu); ORCID: 0000-0002-8951-3704

Lise Rochaix<sup>3,7</sup>, email: [lise.rochaix@psemail.eu](mailto:lise.rochaix@psemail.eu); ORCID: 0000-0002-6467-0332

\*Corresponding author. E-mail: email: [eliana.barrenho@imperial.ac.uk](mailto:eliana.barrenho@imperial.ac.uk). Address: Imperial College Business School, Exhibition Road, SW72AZ London. T: +44 (0)20 7594 9764. Tel: +33 1 45 24 74 46

### Contribution Statement and guarantor information

All authors contributed to the: i) conceptual development of the research question; ii) research design and methodological approach; iii) drafting of the paper. The statistical analyses have been ran by Reka Halmai, Marisa Miraldo and Iván Tzintzun. All authors have seen and approved the final version of the manuscript for publication. All authors had full access to all of the data and can take responsibility of the data and the accuracy of the data analysis. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at [http://www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: all authors had financial support from the INCa (Institut National du Cancer) (grant number #9580) for this work and Setti Raïs Ali also received from the Imagine Institute for rare diseases a PhD scholarship during this period; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

**Funding sources and statement:** This work was supported by a grant on “Rare Cancers & Orphan Drugs, innovation and Inequality” from INCa (Institut National du Cancer) (grant number #9580) and by the European Union Pilot Projects and Preparatory Actions Programme (2014- 2020), as pilot project PP-1-2-2018-Rare 2030, as well as the Imagine Institute for rare diseases. Authors are grateful to Zoé Fontier, Thelma Arcelin and Philippe Gorry for their help and useful comments. The funders had no role in the study design, in the

---

<sup>1</sup> Department of Economics and Public Policy, Business School, Imperial College London

<sup>2</sup> Organisation for Economic Co-operation and Development (OECD)

<sup>3</sup> Hospinnomics (PSE–Ecole d’Economie de Paris, Assistance Publique des Hôpitaux de Paris–AP-HP)

<sup>4</sup> Centre for Health Economics and Policy Innovation (CHEPI), Imperial College London

<sup>5</sup> Institut des Politiques Publiques, PSE–Ecole d’Economie de Paris

<sup>6</sup> Centre de Recherche des Cordeliers, Sorbonne Université, Université de Paris, Inserm, Laboratoire ETREs, F-75006, Paris, France

<sup>7</sup> Université Paris 1 Panthéon-Sorbonne, Paris-Jourdan Sciences économiques (UMR 8545), Paris, France

collection, analyses and interpretation of the data, in the writing of the report and in the decision to submit the article for publication. All researchers are independent from the funders. The opinions and arguments expressed herein are those of the authors and do not necessarily reflect the official views of the OECD or its member countries.

**Data sharing statement:** Data from Orphanet are available via signature of a Data Transfer Agreement for academia. Clinical trials data: <https://www.clinicaltrials.gov/>

**Ethical approval:** not required.

## **Inequities In Cancer Drug Development in Terms of Unmet Medical Need**

21<sup>st</sup> February, 2021

### **Abstract**

This study measures inequality and inequity in the distribution of clinical trials on cancer drug development between 1996 and 2016, comparing the number of clinical trials with cancer need, proxied by prevalence, incidence, or survival rates for both rare and non-rare cancers. We leverage a unique global database of clinical trials activity and costs between 1996 and 2016, constructed for 227 different cancer types to measure for rare and non-rare cancers: i) inequalities and inequity of clinical trial activity, considering all trials as well as split by R&D stage; ii) inequalities and inequity in R&D investment proxied by trial enrollment and duration; iii) evolution of inequity over time. Inequalities are measured with concentration curves and indices and inequities measured with the health inequity index. We find four important results. First, we show *pro-low need inequity* across cancer types for both rare and non-rare cancers, for all need proxies. Second, we show inequity differs across R&D stages and between rare and non-rare cancers. The distribution of clinical trials for non-rare cancers disproportionately favors low-need non-rare cancers from earlier to later stages of R&D, whilst for rare cancers this only occurs in Phase 2 trials. Third, inequity analyses in R&D investment show that only trial enrollment for rare cancers and trial duration for non-rare cancers are disproportionately concentrated among low-need cancers. Finally, while *pro-low need inequity* has persisted between 1996 and 2016 for non-rare cancers, it has faded for rare cancers post-EU orphan drugs' legislation.

***Keywords:*** Equity; Rare and Non-Rare Cancers; Clinical trials; Unmet Cancer Needs.

# 1. Introduction

Rare and non-rare cancers rank as a leading cause of death and an important barrier to increasing life expectancy in every country of the world (Bray et al, 2021). According to estimates from the World Health Organization (WHO) in 2019 (WHO, 2020), cancer is the first or second leading cause of death before the age of 70 years in 112 of 183 countries. In many countries the lifetime risk of developing some form of cancer exceeds 50% (Tanday, 2015). The importance of rare cancers is high, representing about 22% of all cancers diagnosed worldwide, with an occurrence of 6 per 100,000 individuals approximately (Pillai and Jayasree, 2017).

To address cancer unmet need, the last decades have witnessed substantial investments by the pharmaceutical industry directed at the development of oncological medicines (Scannell et al., 2012; Pammolli et al., 2011) now representing more than 30% of the total new market launches (Iervolino and Urquhart, 2018). Pharmaceutical investment has led to novel therapies that improved survival and quality of life of millions of patients globally. For example, the vaccine against the human papillomavirus, has been found to decrease the risk of developing cervical cancer by 88% in girls and women vaccinated before the age of 17 (Lei et al., 2020). Tyrosine kinase inhibitors have improved 5-year survival rates in patients diagnosed with myeloid leukemia by 40.4% (Di Felice et al., 2018). However, incentives for pharmaceutical innovation vary across the different disease areas. With larger market sizes, non-rare cancers exhibit higher prospects of profitability and higher probability of success (POS) on clinical trials than rare cancers (Wong et al., 2019). Therefore, although a disproportionately large amount of understanding of cancer biology comes from the study of rare cancers, most R&D activity has historically focused on non-rare cancers (Boyd, 2016).

A key policy concern is whether the distribution of R&D efforts across the different types of cancers is equitable, i.e., distributed according to need rather than to purely profitability considerations. This motivated several jurisdictions to design Orphan Drug (OD) legislations and policies to foster innovation for rare-cancers. Notably, the US and EU enacted, respectively, the Orphan Drug Act in 1983 and the Orphan Drug legislation of the European Parliament in 2000. These innovation policies provided incentives to encourage R&D for drugs aimed at treating, preventing, or diagnosing rare diseases. Patent extensions and/or financial incentives are granted if the target population in drug development is considered as suffering from a rare disease under each jurisdiction prevalence-based definition. However, the incentive award criteria apply across all drugs that meet the threshold, without additional conditions. This implies that targeted population size remains a key criterion for pharmaceutical firms' R&D decisions. Even if OD policies have succeeded in directing research efforts towards rare diseases, these incentives may not materialize equally across the different diseases since the criteria for benefiting from supply-side incentives is mainly the targeted population size. OD policies may also have spillover effects on non-rare cancer innovation. For example, it is plausible that a firm selectively applies for OD designations for therapies that can also be used for non-rare cancers where market size is larger. These incentives may have equity implications that warrant the formal assessment of the distributional characteristics of drug development.

Therefore, the main purpose of this study is to assess the extent to which R&D for oncological medicines is equitable, i.e. distributed across disease areas according to their relative need. Need is proxied here by either prevalence, incidence, or 1-year and 5-year survival rates. To do so, we

leverage a unique global database of clinical trials activity and costs between 1996 and 2016, constructed for 227 different cancer types. Using the different need proxies, we measure for rare and non-rare cancers: i) inequalities and inequity of clinical trial activity, considering all trials as well as split by R&D stage; ii) inequalities and inequity in R&D investment proxied by trial enrollment and duration; iii) evolution of inequity over time. We use concentration curves (CCs) and concentration indices (CIs) to measure inequalities and assess whether the distribution of R&D activity and investment is equally distributed across disease areas, concentrated among low need diseases (henceforth *pro-low need* inequality), or concentrated among high need diseases (henceforth *pro-high need* inequality). Complementary, we explore the equity dimension of the distribution of R&D activity and investment by estimating the Horizontal Inequity Index (HII), which shows whether R&D activity is distributed in proportion to need.

There are four main findings from our analyses. First, we show *pro-low need* inequity across cancer types for both rare and non-rare cancers. These results hold for all cancer need proxies. Second, we show inequity differs across R&D stages and between rare and non-rare cancers, with the distribution of clinical trials for non-rare cancers disproportionately concentrated among low-need non-rare cancers from earlier to later stages of R&D, and for rare cancers, only in Phase 2 trials. Third, inequity analyses in R&D investment show that trial enrollment is disproportionately concentrated on low-need rare cancers and that trial duration (for non-rare cancers) are disproportionately concentrated among low-need cancers. Fourth, while inequity *pro-low need* cancers has persisted between 1996 and 2016 for non-rare cancers, the inequity *pro-low need* cancers shown for rare cancers pre-legislation fades after the introduction of OD legislation.

Overall, our results show the presence of misalignment between R&D investments in anticancer



drugs and unmet need. If underlying social preferences are pro-equality, our results suggest that health systems need better incentives and resource allocation frameworks, for a better alignment between cancer drug innovation and cancer need.

The remainder of the paper is organized as follows: Section 2 discusses the relevant literature. Sections 3 and 4 present the data sources and methods supporting the analyses. Section 5 presents the results and Section 6 presents the discussion and main conclusions of the study.

## **2. Literature Review**

There is a small growing set of contributions that addresses inequalities in pharmaceutical innovation and attempts to assess whether observed inequalities are warranted by differences in unmet need across diseases. However, the methods used in the literature are mostly descriptive and therefore do not allow assessing the distributional impacts of innovation, nor the magnitude and direction of inequity. Most of the existing studies perform correlation analyses and suggest contradictory results on the direction of the mismatch of therapeutic innovation across large disease groups. A first group of studies shows no correlation between health need and pharmaceutical innovation (Catalá-López et al., 2011; Neumann et al., 2005). A second group, however, suggests the intensity of pharmaceutical innovation and R&D output by disease area is correlated with need, but this correlation is small in magnitude and there may be a disproportionate concentration of innovation in some disease areas (Atal et al., 2018; Viergever, Terry, and Karam, 2013; Catalá-López et al., 2010).

There is only one study that formally measures the magnitude and direction of health inequalities of pharmaceutical innovation both across and within broad disease areas (Barrenho et al., 2019).

The study suggests that new drug approvals are globally relatively well-aligned with disease burden between 1990 and 2010, but that there may be disproportionate concentrations of innovation in some disease areas, while unmet need remains in others. The results show pro-need inequalities in anticancer drug development across 39 types of cancer.

While some of these studies attempt to infer the existence of inequity in the distribution of innovation, based on observed inequalities in innovation across disease areas, they fail to systematically benchmark these inequalities with the distribution of need across disease areas.

Establishing the existence of inequalities is not sufficient to assess how equitable R&D activity is. An unequal distribution of R&D activity across disease areas may be warranted by differences in the magnitude of unmet need across disease areas. Therefore, the assessment of equity requires the explicit comparison between the distribution of R&D activity and the distribution of cancer need, which is the main focus of this paper. We use different proxies for cancer need and add to the existing literature by measuring inequity in innovation for both rare and non-rare cancers. Finally, we enhance these studies by using broader proxies of R&D activity which measure clinical trials activity during the R&D process by looking at the number of clinical trials in total and by R&D stage, as well as proxies of R&D effort, namely trial enrollment (number of patients enrolled) and duration. This is in contrast with existing studies on drug innovation that use partial measures of innovation, such as the number of published HTA assessments (Catalá-López et al., 2011; Neumann et al., 2005), published clinical evidence (Kaplan et al., 2004; Kaplan, 2015), or self-reported data on innovation (Martino et al., 2012), that do not fully capturing the extent of innovation or suffer from publication reporting bias. Other studies improve on these proxies by considering R&D stages such as clinical R&D stage (Viergever and

Li, 2015) and market registration and launch (Catalá-López et al., 2010; Lichtenberg, 2005; Barrenho et al., 2019). However, by selectively focusing on single stages of the R&D process, they fail to capture the innovation process in its entirety. Yet risk, complexity and returns on investment differ substantially across the different R&D phases, and it is likely to affect rare and non-rare cancers differently. It is therefore important to analyze the distribution of clinical trials for each R&D phase and to assess whether advanced phase clinical trials present a more pronounced inequity than early phase trials.

Furthermore, the number of clinical trials may not reflect the actual magnitude of the research efforts. Martinez (2016) estimates the Phase 2 and Phase 3 oncology clinical trial costs at \$11.2 million and \$22.1 million respectively. Much of these costs are driven by the expenses associated with participants' enrollment (patients' recruitment and retention costs), others being explained by medical staff expenses (site monitoring, retention, and recruitment). Therefore, we also add to this literature in that we analyze differences in clinical trials investments.

Finally, the literature relies on data that does not have sufficient granularity to assess inequalities or inequities for rare cancers. While Barrenho et al. (2019) assess inequalities in cancer innovation, the authors do not stratify the analyses to distinguish rare and non-rare cancers, and the study uses burden of disease (BoD) to proxy cancer need, which is available for less than 30 types of cancer, and largely unavailable for rare cancers, thus not allowing accounting for heterogeneity in need between rare and non-rare cancers. Importantly, there is yet no formal measure of inequity in this paper or in the related literature. Therefore, while our study relates to this literature, we build on it in three significant ways: 1) we are the first to measure inequity in R&D activity, and R&D investment (proxied by trial enrollment and duration) for different types

of cancer, including rare and non-rare cancers, using different need proxies; 2) we assess how inequity has evolved over time; 3) we make use of original data on the different R&D stages to assess inequity both within and between early and later stages of clinical trials.

### **3. Data Sources**

We build a unique dataset that links clinical trials targeting rare and non-rare cancers with epidemiological data per type of cancer and over time. We do so by merging data from three different sources: i) the Clinicaltrial.gov database that provides registry data on the number of clinical trials targeting each type of rare and non-rare cancer between 1996 and 2016; ii) Orphanet that provides data on clinical trials targeting rare cancers between 1996 and 2016; iii) Rarecarenet.eu that provides epidemiological data on prevalence, incidence, one- and five-year survival rates for the year 2007, per cancer type.

Clinicaltrials.gov is a publicly available database provided by the U.S. National Library of Medicine at the National Institute of Health (NIH), containing privately and publicly funded clinical trials globally since 1970.

The data provides the number of clinical trials for 20 non-rare cancer types. To do so, we use medical subject headings (MeSH) (MeSH is a comprehensive and controlled vocabulary for the purpose of indexing journal articles and books in the life sciences; it serves as a thesaurus that facilitates searching) to characterize clinical trials by types of cancer in terms of their site of origin. Commonly, one MeSH term corresponds to more than one type of cancer, given that one clinical trial might target several cancer types simultaneously. We consider all multiple cancer types targeted by each clinical trial.

Using this information, we construct two datasets with a total of 31,081 observations: one dataset that contains yearly counts of the number of clinical trials targeting each cancer type assigning equal weights to all observations— henceforth labelled as non-rare clinical trial dataset (NRD) - and a second dataset where observations are weighted according to the number of cancer types targeted by each trial - henceforth labelled the non-rare weighted clinical trial dataset (NRWD)- with weights inversely proportional to the number of cancer types in each trial. This enables accounting for trials targeting a large number of cancers being less effective in developing cancer-specific treatments. The NRWD allows us to perform the analysis correcting for possible over-representation of trials with many cancer types associated. Figure 1 shows each step of the dataset construction.

*Insert Figure 1 here*

For rare cancers, we also collect the number of clinical trials from [clinicaltrials.gov](http://clinicaltrials.gov). As in Bell and Smith (2014), we identify clinical trials targeting rare cancers if at least one submitted MeSH annotation in the conditions was found in the disaggregated data. Trials primarily targeting non-rare cancer types, with a possible application to a rare cancer, will tend to report all possible applications, even though research is dominantly non-rare cancer targeted. This may lead to an overestimation of the number of rare cancer trials. Therefore, to identify clinical trials truly targeting rare cancers, we use Orphanet database which provides information on clinical trials between 1996 and 2016, classified into rare types of cancer by clinical experts. Orphanet classification of rare diseases is based on the definition European Medicines Agency (EMA) as life-threatening or chronically debilitating conditions that affect no more than 5 in 10,000 people in the EU (European Medicines Agency, 2013).

We consider only those clinical trials registered on [clinicaltrials.gov](http://clinicaltrials.gov) that target rare cancers according to Orphanet. Naturally, the question arises as to whether we lose information by using only a subset of the Orphanet's database. While merging the two registers would enable a more comprehensive characterization of trials targeting rare disease, most of the clinical trials included in the Orphanet dataset do not have a trial number (NCT) which would enable us to identify them easily in the [clinicaltrial.gov](http://clinicaltrial.gov) dataset. This makes it very difficult to match the trials on the basis of year and pathology as rare cancers could be named differently (differences in the precision of the reported clinical subtype, for example) in Orphanet versus [clinicaltrials.gov](http://clinicaltrials.gov). Therefore and to avoid biases that affect differently the measurement of trials for rare and non-rare diseases, we limit our data to [clinicaltrials.gov](http://clinicaltrials.gov), as having only one source in the case of the non-rare cancers but several in the case of rare cancers might bias our analyses.

Furthermore, as already mentioned, we assume that cancer types are uncorrelated with the proportion of clinical trials registered. We call one observation a trial addressing one cancer type; therefore, we construct two rare cancer clinical trials dataset according to the way we assign weights to each observation. We refer to rare cancer clinical trials data (RD) as the dataset of those observations with equal weights, and to rare weighted clinical trials data (RWD) the dataset with weights that are inversely proportional to the number of entries of the trial in the dataset. The process of dataset construction for rare cancers is presented in Figure 2.

*Insert Figure 2 here*

Besides trial activity measured by number of entries, we assess inequalities and inequity in R&D investment proxied by the number of enrolled patients per year and the duration of clinical trials. Albeit being imperfect measures of total R&D efforts (Joe Martinez, 2016), trials' duration and

costs are a significant part of overall expenditure in R&D. According to Sertkaya et al. (2016), around 87% of the required investment to cover identifiable costs is related to trial enrolment (46%) and duration (41%). Thus, while trial enrolment and duration do not fully capture R&D costs, they capture a substantial part and are thus likely to correlate meaningfully with total R&D costs. Key costs that are not captured (such as site overhead costs) are likely to be common to all disease areas and therefore are likely to be allocated (in accounting terms) to each trial based on trial duration and trial size, thus reinforcing the validity of using these cost items as proxies for R&D efforts.

The number of patients enrolled in trials and trial durations are only considered for Phase 2 clinical trials, with already completed enrollment. We also choose to consider Phase 2 clinical trials as they recruit patients more uniformly. Many oncological drugs are accessing the market without completing Phase 1 and/or Phase 3 trial because of ethical considerations (Wright et al., 2002; Agrawal and Emanuel, 2003). Duration of the trial is measured as the number of days that occur between the date the trial was firstly posted and its completion date.

Finally, to measure need, we use estimates of European crude, 15-year prevalence, age-specific incidence, as well as 1-year and 5-year age-adjusted relative survival rates for cancers, the latter being obtained from the rarecancernet.eu. Relative survival was estimated by the EUROCORE-5 study according to the Hakulinen method. Their data is drawn from the EUROCORE-5 study, the largest, European collaboration of research efforts to provide information on the epidemiology of rare cancers. They provide estimates of the prevalence, incidence and survival rate based on data obtained from 92 Cancer Registries, in 22 European countries, including more

than 21 million cancer diagnoses until 2007. The list of rare tumors is hierarchically structured in three layers based on the ICD-O morphology and topography.

The final dataset contains each cancer's epidemiological data matched with the number of clinical trials/weighted clinical trials targeting the given type of cancer, divided by year/period/phase. The question arises whether using epidemiological data estimated at a given point in time while clinical trials are collected for the whole period could bias our analysis. While this is a hypothetical limitation of the analyses due to data unavailability, it is unlikely to bias our analyses given that the ranking of cancers in terms of epidemiological variables is likely to remain constant over time. Based on IHME Global Burden of Disease data, when comparing prevalence of non-rare cancers in 1996 and 2016, we find a very small difference in rankings. As the concentration curves and indices measure relative inequality and inequity, they are invariant to equi-proportional changes in the variables.

Given that the classification of rare cancers clinical trials differs between Orphanet and EUROCORE-5 study (which uses ICD-O morphology and defines 194 rare cancers), we harmonize by matching manually the Orphanet classification onto the EUROCORE classification. The accuracy of harmonization of the two classifications is then verified by an expert specialized in genetics.

The final dataset consists of 32,535 observations, with 31,081 observations for non-rare cancers and 1,454 for rare cancers. One trial may target several cancers and thus count as several observations. Considering a trial only once, we obtain 26,948 clinical trials observations: 801 trials targeting rare cancers and 26,147 targeting non-rare cancers.



## 4. Methods

### 4.1 Concentration curves and concentration indices

We use concentration curves (CCs) and concentration indices (CIs) to measure inequalities in cancer R&D for the various types of cancer. Concentration curves plot the cumulative share of R&D allocated to various cancer types, ranked by need — i.e. prevalence, incidence and, 3-year or 5-year survival rates — against the cumulative share of R&D activity targeting rare and non-rare types of cancer. If each cancer type were to benefit from the same volume of clinical trials, concentration curves would lie on the 45° diagonal line described as the line of equality.

Inequality is therefore represented as the area that lies between the concentration curve and the equality line. If the concentration curve lies below the line of equality, there is an inequality favoring high-need (defined either by prevalence, incidence, or survival) cancer (hereinafter *pro-high need* inequality), whereas if the concentration curves lie above the line of equality, R&D will be concentrated on low-need cancers, i.e. *pro-low need* inequality.

Although the concentration curves offer a visual representation of the direction of the inequality, they do not allow assessing its magnitude. Therefore, for each distribution, we compute a standard concentration index that corresponds to twice the area between the concentration curve and the line of equality. There is a family of concentration indices, and the choice of concentration index is dependent on the characteristics of the outcome variable as well as normative judgements on the measurement of inequality (O'Donnell et al., 2016). As the

number of clinical trials for each cancer type is not upper bounded and the measurement scale is fixed, the standard CI is an appropriate measure of the divergence from the line of equality and is defined as:

$$CI = 1 - 2 \int_0^1 C_x dx \quad (1)$$

where  $C_x$  is the concentration curve. It can also be calculated as:

$$\frac{1}{n} \cdot \sum_{i=1}^n \left( \frac{T_i}{\underline{T}_i} - (2E_i - 1) \right) \quad (2)$$

where  $T_i$  is the outcome variable for which inequality is measured (in our case, the number of clinical trials for cancer  $i$ ), while  $E_i$  is the fractional rank of the ordering variable (prevalence, incidence, 1-year and 5-year survival rate for cancer  $i$ ). The CI ranges between -1 (i.e. maximum *pro-low need inequality* with all research concentrated on the cancer with the least need) to 1 (i.e. maximum *pro-high need inequality*, i.e. all research concentrated on the cancer with the highest need). As the index measures relative inequality, it is invariant to proportional changes in the variable of interest and can be an appropriate measure even when concentration curves coincide, thus the inequality assessment cannot be done by visual inspection of the curves.

## 4.2 Horizontal Inequity Index

Concentration curves and indices can only be used to assess inequalities. While these analyses are useful to assess how trials are distributed across cancers, inequalities observed through the concentration curves and indices cannot be directly interpreted as inequity. Variations in R&D activity due to differences in need might be a legitimate source of inequality. Therefore, a judgement on whether these inequalities are warranted or unwarranted when one accounts for

differences in need across diseases requires the assessment of inequity in the distribution of trials by benchmarking the distribution of trials against the distribution of need (O'Donnell et al, 2008). We therefore measure equity by computing the Horizontal Inequity Index (HII) that captures the extent to which R&D activity is distributed in proportion to need (O'Donnell et al., 2008; Wagstaff and Van Doorslaer, 2000) and is given by the area between the CC of R&D activity and the CC of need. Mathematically, it is the difference between the two CIs, calculated as shown in (2), i.e.:

$$HII = CI_{R\&D\ outcomes} - CI_{need} \quad (3)$$

The HII ranges between -2 and 1; in the case where R&D is distributed in proportion to need (e.g. high-need diseases receiving the same trials in proportion to their associated need than low-need diseases or when both distributions of need and R&D activity overlap with the equality line) the index is zero. A negative (positive) index represents a concentration of R&D activity in high-need (low-need) diseases, that are targeted with more activity than the fair share their relative need would warrant when compared to other disease areas.

We leverage these methods to run several analyses: i) inequalities and inequity of clinical trials, and R&D costs for rare and non-rare cancers; ii) inequity for all trials pooled together and by R&D phase comparing early with late-stage trials; iii) evolution of inequity over time and how it changed from the period pre- to post-Orphan drug European legislation, by comparing the period 1996-2000 with 2001-2016. The baseline analyses are run for both unweighted and weighted data and results do not change. Results are available in the supplementary material.

## 5 Results

### 5.1 Comparing rare and non-rare cancers' characteristics

Table 1.a summarizes the main differences between rare and non-rare cancers in our database. The average prevalence among non-rare cancers is 93.52 per 100,000, while for rare cancers it is 4.64; the incidence is 30.22 for non-rare cancers and 0.58 for rare cancers ( $p < 0.01$ ). Although non-significant, the survival rate is slightly higher for rare cancers (51.45%) compared to non-rare cancers (48.46%). The average number of clinical trials is 1554 for each non-rare cancer and 7.02 for the rare ones ( $p < 0.01$ ). The reason for such a difference is that, during the period of the study, only 102 rare cancers have at least one clinical trial, which represents 49% of all the rare cancers identified in Orphanet. Average enrollment and duration were calculated by only considering cancers with at least one trial. The mean enrollment is equal to 176.05 individuals for rare cancers and 103 for non-rare cancers; lastly, the duration appears slightly larger for non-rare cancers (3.88 years) compared to rare ones (3.76). Nonetheless, the differences in enrollment and duration are not significant.

*Insert table 1a here*

Table 1.b shows the number of clinical trials (both unweighted and weighted) for rare and non-rare cancers. More than 40% of non-rare clinical trials target only four types of cancer: breast cancer, leukemia and tracheal, bronchus and lung cancer, and prostate cancer. For rare-cancers 19% of all trials target one type of cancer, the acute myeloid leukemia, while the 20 most researched cancer types account for 75% of all rare cancer trials.

*Insert Table 1b here*

## **5.2 Inequalities and inequity in trial activity**

### **5.2.1 Inequalities and inequities of total clinical trials**

Regarding inequalities, Table 2 presents the CIs of clinical trials across cancer types ranked by prevalence, for rare and non-rare cancers separately. Results show *pro-high need* inequality for both rare and non-rare cancers, with trials concentrated on higher prevalence cancers, and this inequality being greater for rare cancers (CI=0.541, p-value<0.01) than for non-rare cancers (CI=0.261, p-value<0.05). The results are qualitatively similar when using incidence as a proxy for cancer need (Table 2). They differ when using survival rates where we find no evidence of such concentration. Results can be found in the supplementary material (Table 2).

*Insert Table 2 here*

Yet the unequal distribution of clinical trials can be partly explained by the differences of need across the different types of cancer. This is confirmed by the fact that concentration curves for prevalence (Figure 3) envelop those curves for clinical trials, resulting in larger concentration indices, both for rare (CI=0.788, p-value<0.01) and non-rare cancers (CI=0.538, p-value<0.01).

*Insert Figure 3 here*

The extent to which the concentration of trials in high-need cancers is inequitable depends on the relative need across cancers. Results (Table 3) show that total trials are disproportionately concentrated among low-prevalence cancers (inequity *pro-low need*) for both rare (HII=-0.247, p-value<0.10) and non-rare (HII=-0.277, p-value<0.005). Results are qualitatively similar when

using incidence as a proxy for cancer need (Table 3). Also here, results differ when using survival rates where we find no evidence of such concentration. These results are available in the supplementary material.

*Insert Table 3 here*

Results from the inequity analysis (Table 3) for total clinical trials show *pro-low need* inequity and results hold for prevalence, incidence and 5-year survival rates, while for 1-year survival rates, inequity only remains significant for rare cancers (HII=-0.253, p-value<0.10).

### **5.2.2 Inequalities and inequities over time**

We next assess the evolution of inequity over time and compare inequity in the period pre- and post-Orphan drug European legislation adopted in 2000. While the legislation seized incentives to foster investment in rare diseases, it is unclear whether it benefited equally all rare diseases and to what extent it changed incentives to invest in non-rare cancers. To assess this, we compare levels of inequity pre- and post- legislation. Table 3 shows that in the period pre-legislation, clinical trials were disproportionately concentrated among low-prevalence cancers (inequity *pro-low need*) for both rare (HII=-0.447, p-value<0.05) and non-rare (HII=-0.323, p-value<0.05). After the introduction of the legislation, *pro-low need* inequity only persists for non-rare cancers (HII=-0.273, p-value<0.05), provided the results for rare cancers are statistically non-significant (HII=-0.242, p-value>0.10), indicating that trials are distributed in proportion to need for rare cancers. Results are qualitatively similar when measuring need with cancer incidence (Table 3). For both 1-year and 5-year survival rates, results only differ with regards to rare cancers. The results are available in the supplementary material.

Figures 4a and 4b show the yearly HII for the distribution of clinical trials for both rare and non-rare cancers separately, with cancer types ranked by prevalence. Results suggest a *pro-low need* inequity for non-rare cancers, given the negative and statistically significant HII between 1996 and 2016 (Table 4) with the level of inequity remaining constant over time. For rare cancers, while results suggest inequity *pro-low need* in 2000, post-2000 the HII decrease in magnitude but are also not statistically different from zero in most years. Results are qualitatively similar when measuring need with other proxies and relatively lower in magnitude for non-rare cancers (Table 4).

*Insert Table 4 here*

*Insert Figure 4a and 4b here*

While no data is available pre-2000, these results provide some evidence of a change in the distribution of clinical trials for rare cancers during the period of the introduction of the European OD legislation, with the inequity *pro-low need* shown in 2000 fading after the introduction of the legislation with HHIs not statistically different from zero for most years post-2000, suggesting a distribution of clinical trials proportional to prevalence. However, these results should be interpreted with caution as the study design is not suited to assessing causality: other simultaneous policy effects could have been at work, along with the introduction of European OD legislation.

### **5.2.3 Inequalities and inequities by R&D phase**

During the period of analysis, the number of clinical trials varies across R&D stages and the

distribution differs between rare and non-rare cancers: respectively 16%, 43% and 41% of clinical trials for rare cancers are in Phase 1, Phase 2 and Phase 3 stages, compared with 28%, 39% and 11% for non-rare cancers.

Figure 5 shows the concentration curves for Phase 1, Phase 2 and Phase 3 stages for cancer types ranked according to prevalence both for rare and non-rare cancers. Whilst we found evidence of inequity *pro-low need* (Table 3), there is heterogeneity across the different R&D stages: inequity *pro-low need* is observed in Phase 2 trials for rare and non-rare cancers (rare: HII=-0.284, p-value<0.05; non-rare: HII: -0.294, p-value<0.05), and only for non-rare cancers in Phase 1 (rare HII=-0.197, p-value>0.10; non-rare HII: -0.373, p-value<0.01). No inequity is observed in Phase 3 trials for rare cancers (HII=-0.223, p-value>0.10) while for non-rare cancers, the HII is negative but only marginally significant (HII: -0.185, p-value<0.10). Results are qualitatively similar when using incidence as a proxy for need, except the results for Phase 3 trials for non-rare cancers that become statistically non-significant.

*Insert Figure 5*

### **5.3 Inequalities and inequities in R&D costs**

The distribution of the number of clinical trials does not necessarily accurately capture the R&D cost, nor the magnitude of investment dedicated to the various cancer types. Thus, we assess inequality and inequity in the distribution of R&D costs using two proxies, namely, trial enrollment and trial duration.

Starting with inequalities, we see in Table 2 the concentration curves and indices of both trial



enrollment and trial duration across cancer types ranked by prevalence, for rare and non-rare cancers separately. Results suggest a *pro-high need* inequality for both trial enrollment and trial duration for both rare cancers (trial enrollment CI = 0.551, p-value < 0.01; trial duration CI = 0.435, p-value < 0.05) and non-rare cancers (trial enrollment CI = 0.543, p-value < 0.01; trial duration CI = 0.222, p-value < 0.05).

Results from the inequity analysis (Table 3) show that only trial enrollment for rare cancers (trial enrollment: HII=-0.237, p-value<0.10; trial duration: HII=-0.245, p-value>0.10) and trial duration for non-rare cancers (trial size: HII=-0.104, p-value<0.05; trial duration: HII=-0.316, p-value<0.05) are disproportionately concentrated among low-need cancers, showing *pro-low need* inequity.

## 5.4 Robustness checks

As described in the data section, a trial might target multiple cancers and therefore our analyses are at cancer-trial level, with each trial targeting on average 1.8 and 1.18 cancer types respectively for rare and non-rare cancers. To ensure that duplicate entries with equal weights do not distort our analysis, as a robustness check, we recalculate table 2, 3 and 4 using the weighted dataset. Results, available in the supplementary material, show very similar patterns after estimating the CIs and HIIs with the weighted data and confirm our findings in the previous section.

## 6 Conclusion

This paper measures inequality and inequity in cancer drug development conducted between 1996 and 2016 for the treatment of rare and non-rare cancers. We find four important results. First, while the inequality analysis suggests a concentration of clinical trials *pro-high need* cancer types, there is inequity *pro-low need* when considering the relative need across cancer types for both rare and non-rare cancers. In other words, while there are more trials targeting diseases with higher associated need, that extra activity is not sufficient in light of their relative need when compared to other disease areas. These results hold when measuring cancer need in terms of prevalence, incidence, or 1-year and 5-year survival rates.

Second, we find inequity changes in the distribution of clinical trials for rare cancers after the introduction of the OD legislation. While *pro-low need* inequity has persisted between 1996 and 2016 for non-rare cancers, the *pro-low need* inequity also shown for rare cancers pre-OD legislation fades after 2000. However, this descriptive analysis cannot accurately attribute this effect to the introduction of the OD legislation. Additional analysis and data are needed to assess the impact of this legislation.

Third, we find that inequity of R&D activity differs across R&D stages and between rare and non-rare cancers. The distribution of clinical trials for non-rare cancers disproportionately favors low-need non-rare cancers across all R&D stages, whilst for rare cancers this only occurs in Phase 2 trials. These results may be partially explained by the lower probability of success of non-rare cancers with high prevalence, where scientific knowledge and/or for which diagnostics to develop new treatments are already consolidated (Austin et al., 2018). For rare cancers however, there seems to be a disproportionate investment in trials targeting very rare cancer types where competition is fierce and large profits are at stake (McKinsey & Co, 2021).

Fourth, we show there is inequity in the distribution of R&D investment (proxied by trial enrollment and duration). Results show that only trial enrollment for rare cancers and trial duration for non-rare cancers are disproportionately concentrated among low-need cancers, showing *pro-low need* inequity.

Our results should be interpreted in the light of data-related limitations. While we contribute to the literature by considering activity in different R&D phases, the use of clinical trials as a proxy for innovation has also limitations. Clinical trial activity data is an imperfect proxy for investment allocation across the different disease areas. We do not observe R&D expenditure across all trial phases and therefore imperfectly capture trials costs. Additionally, our data sources do not offer a complete mapping all R&D activity. For rare diseases Orphanet data may provide a more comprehensive account of clinical trial activity for rare diseases. We have opted to use clinicaltrials.gov data instead to ensure comparability across rare and non-rare diseases, but also to mitigate any biases that could arise from different levels of reporting across the two registries. While it is highly plausible that trial underreporting in Clinicaltrials.gov is unrelated with cancer type, we cannot rule out bias arising from a potential association of the level of reporting with the development of orphan drugs. Also, while we attempt to use the most disaggregated level of data to characterize cancer types by relying on trial data, these data may still fail to capture the full characterization of each disease type. The extent to which this biases our findings is unclear because in the same way may be allocating more trials to a group of cancer types, we may also attach to that group a higher disease need (i.e. need across all different genetic types of the same cancer). In addition, while we consider prevalence, incidence, and

survival rates as proxies for need, these are still imperfect measures of unmet cancer need since they do not capture the overall burden associated with disease, nor the fact that some diseases may be treated with non-drug treatments. Moreover, our analysis does not factor efficiency nor unpacks the reasons behind inequality nor inequity in the distribution of clinical trials' activity and investment. We cannot rule out that differences across rare and non-rare inequities relate to differences in underreporting of trials across cancer types. Furthermore, even if such biases would not exist, there are multiple reasons that may warrant differences in innovation across cancer types, such as differences in technical complexity and understanding of disease pathogenesis that may imply that innovation in one disease area is more challenging than in another. Inequities in drug development may also reflect technical failure of the projects rather than purely industry strategic behavior. Therefore, while the measurement of inequity of the distribution of innovation is an essential first step into identifying the need for better incentive policies, the identification of the type and extent of incentives required would require further research into determinants of inequity. To that end, better and more granular data is required to deepen our understanding of inequity drivers.

Finally, we assess equity on the grounds of a principle of justice that assumes that proportionality is equitable, i.e., disease areas should receive innovation in proportion to need. While these principles guide policy-making in most health systems and therefore should guide incentive schemes' design and allocation of R&D efforts across disease areas, other principles of justice could have been considered.

Despite these caveats, our results suggest that health systems need to refine incentives and

resource allocation frameworks, for a better alignment between cancer drug innovation and medical need, if underlying social preferences are pro-proportionality. There is therefore a pressing need to further evaluate the impact of existing incentive mechanisms targeted at promoting innovation in unmet need areas.

## 7 References

- Agrawal, M., & Emanuel, E. J. (2003). Ethics of phase 1 oncology studies: reexamining the arguments and data. *Jama*, 290(8), 1075-1082. 2.
- Austin, C.P., Cutillo, C.M., Lau, L.P., Jonker, A.H., Rath, A., Julkowska, D., Thomson, D., Terry, S.F., de Montleau, B., Ardigò, D., Hivert, V., Boycott, K.M., Baynam, G., Kaufmann, P., Taruscio, D., Lochmüller, H., Suematsu, M., Incerti, C., Draghia-Akli, R., Norstedt, I., Wang, L., Dawkins, H.J. and (2018), Future of Rare Diseases Research 2017–2027: An IRDiRC Perspective. *Clinical And Translational Science*, 11: 21-27. <https://doi.org/10.1111/cts.12500>
- Auvin, S., Irwin, J., Abi-Aad, P., & Battersby, A. (2018). The problem of rarity: Estimation of prevalence in rare disease. *Value in Health*, 21(5), 501–507.
- Barrenho, E. and Miraldo, M. (2018). R&D success in pharmaceutical markets: A duration model approach', *health econometrics (contributions to economic analysis, volume 294)*.
- Barrenho, E., Miraldo, M., & Smith, P. C. (2019). Does global drug innovation correspond to burden of disease? The neglected diseases in developed and developing countries. *Health economics*, 28(1), 123-143.
- Bell, S. A. and Smith, C. T. (2014). A comparison of interventional clinical trials in rare versus non-rare diseases: an analysis of clinicaltrials.gov. *Orphanet journal of rare diseases*, 9(1):170.
- Belloni, A., Morgan, D., and Paris, V. (2016). Pharmaceutical expenditure and policies.
- Boyd, N., Dancey, J. E., Gilks, C. B., & Huntsman, D. G. (2016). Rare cancers: a sea of opportunity. *The Lancet Oncology*, 17(2), e52-e61
- Catala-Lopez, F., Garcia-Altes, A., Alvarez-Martín, E., Genova-Maleras, R., and Morant-

- Ginestar, C. (2010). Does the development of new medicinal products in the European Union address global and regional health concerns? *Population health metrics*, 8(1):34.
- Catalá-López, F., García-Altés, A., Álvarez-Martín, E., Gènova-Maleras, R., Morant-Ginestar, C., & Parada, A. (2011). Burden of disease and economic evaluation of healthcare interventions: are we investigating what really matters?. *BMC health services research*, 11(1), 1-10.
- Chapman, S., V. Paris and R. Lopert (2020), "Challenges in access to oncology medicines: Policies and practices across the OECD and the EU", OECD Health Working Papers, No. 123, OECD Publishing, Paris, <https://doi.org/10.1787/4b2e9cb9-en>.
- Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P. C., and Sculpher, M. (2015). Methods for the estimation of the national institute for health and care excellence cost-effectiveness threshold. *Health technology assessment (Winchester, England)*, 19(14):1.
- Di Felice E, Roncaglia F, Venturelli F, et al. The impact of introducing tyrosine kinase inhibitors on chronic myeloid leukemia survival: a population-based study. *BMC Cancer*. 2018;18(1):1069. Published 2018 Nov 6. doi:10.1186/s12885-018-4984-3
- Exarchakou, A., Rachet, B., Belot, A., Maringe, C., and Coleman, M. P. (2018). Impact of national cancer policies on cancer survival trends and socioeconomic inequalities in England, 1996-2013: population-based study. *BMJ*, 360: k764.
- Fitzmaurice, C., Allen, C., Barber, R. M., Barregard, L., Bhutta, Z. A., Brenner, H., Dicker, D. J., Chimed-Orchir, O., Dandona, R., Dandona, L., et al. (2017). Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis

- for the global burden of disease study. *JAMA oncology*, 3(4):524–548.
- Gatta, G., Van Der Zwan, J. M., Casali, P. G., Siesling, S., Dei Tos, A. P., Kunkler, I., Otter, R., Licitra, L., Mallone, S., Tavilla, A., et al. (2011). Rare cancers are not so rare: the rare cancer burden in Europe. *European journal of cancer*, 47(17):2493–2511.
- Hofmarcher, T. et al. (2019), Comparator Report on Cancer in Europe 2019 - Disease Burden, Costs and Access to Medicines, IHE, Lund, Sweden, <https://ihe.se/en/publicering/comparatorreport-on-cancer-in-europe-2019/>.
- Iervolino, A. and Urquhart, L. (2018). Evaluate pharma world preview 2017, outlook to 2022.
- Islami, F. et al. (2019), “National and State Estimates of Lost Earnings From Cancer Deaths in the United States”, *JAMA Oncology*, <http://dx.doi.org/10.1001/jamaoncol.2019.1460>.
- Martinez J (2016). Driving drug innovation and market access: Clinical trial cost breakdown. *Center Point clinical services*. [2016 Sep 27]. Available at : <https://www.centerpointclinicalservices.com/blog-posts/driving-drive-drug-innovation-and-market-access-part-1-clinical-trial-cost-breakdown/>.
- Kaplan, W. (2015). Background paper. In *Transforming Health Care Scheduling and Access: Getting to Now*. National Academies Press (US).
- Kaplan, W., Laing, R., et al. (2004). *Priority medicines for Europe and the world*. World Health Organization Geneva.
- Lei J, Ploner A, Elfström KM, Wang J, Roth A, Fang F, Sundström K, Dillner J, Sparén P. HPV Vaccination and the Risk of Invasive Cervical Cancer. *N Engl J Med*. 2020 Oct 1;383(14):1340-1348. doi: 10.1056/NEJMoa1917338. PMID: 32997908.
- Lichtenberg, F. R. (2005). Pharmaceutical innovation and the burden of disease in developing and developed countries. *Journal of Medicine and Philosophy*, 30(6):663–690.



- Martino, O. I., Ward, D. J., Packer, C., Simpson, S., and Stevens, A. (2012). Innovation and the burden of disease: retrospective observational study of new and emerging health technologies reported by the Euroscan network from 2000 to 2009. *Value in Health*, 15(2):376–380.
- McKinsey&Co (2021). Precision medicine in practice: Strategies for rare cancers. <https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/precision-medicine-in-practice-strategies-for-rare-cancers>
- Neumann, P. J., Divi, N., Beinfeld, M. T., Levine, B.-S., Keenan, P. S., Halpern, E. F., and Gazelle, G. S. (2005). Medicare’s national coverage decisions, 1999–2003: quality of evidence and review times. *Health Affairs*, 24(1):243–254.
- Nestler-Parr, S., Korchagina, D., Toumi, M., Pashos, C. L., Blanchette, C., Molsen, E., Morel, T., Simoens, S., Kaló, Z., & Gatermann, R. (2018). Challenges in research and health technology assessment of rare disease technologies: Report of the ISPOR rare disease special interest group. *Value in Health*, 21(5), 493–500.
- O’Donnell O, van Doorslaer E, Wagstaff A, Lindelow M. Analyzing Health Equity Using Household Survey Data. Washington, DC: The World Bank, 2008 DOI:10.1596/978-0-8213-6933-3.
- O’Donnell, O., O’Neill, S., Van Ourti, T., and Walsh, B. (2016). conindex: Estimation of concentration indices. *The Stata journal*, 16(1):112–138.
- Pammolli, F., Magazzini, L., and Riccaboni, M. (2011). The productivity crisis in pharmaceutical r&d. *Nature reviews Drug discovery*, 10(6):428.
- Paris, V., Slawomirski, L., Colbert, A., Delaunay, N., and Oderkirk, J. (2017). *New Health Technologies Managing Access, Value and Sustainability*. Paris: OECD Publishing.

- Pillai, R. K. and Jayasree, K. (2017). Rare cancers: challenges & issues. *The Indian journal of medical research*, 145(1):17.
- Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., & McNamee, C. (2019). Drug repurposing: Progress, challenges and recommendations. *Nature reviews Drug discovery*, 18(1), 41–58.
- Scannell, J. W., Blanckley, A., Boldon, H., and Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature reviews Drug discovery*, 11(3):191.
- Sertkaya, A., Wong, H. H., Jessup, A., and Beleche, T. (2016). Key cost drivers of pharmaceutical clinical trials in the United States. *Clinical Trials*, 13(2), 117-126.
- Tanday, S. (2015), “Lifetime risk of cancer is more than 50% in Britain”, *The Lancet Oncology*, Vol. 16/3, p. e109, [http://dx.doi.org/10.1016/S1470-2045\(15\)70036-7](http://dx.doi.org/10.1016/S1470-2045(15)70036-7).
- Viergever, R. F. and Li, K. (2015). Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. *BMJ open*, 5(9): e008932.
- Wagstaff A, Van Doorslaer E. Equity in health care finance and delivery. In: Culyer AJ, Newhouse JP. (eds.) *Handbook of Health Economics*. vol. 1, Amsterdam: Elsevier 2000; p. 1803-1862.
- Wright, J. R., Crooks, D., Ellis, P. M., Mings, D., & Whelan, T. J. (2002). Factors that influence the recruitment of patients to Phase III studies in oncology: the perspective of the clinical research associate. *Cancer*, 95(7), 1584-1591.
- Wong, C. H. (2017). *Estimation of clinical trial success rates and related parameters*. PhD thesis, Massachusetts Institute of Technology.
- Wong, C. H., Siah, K. W., and Lo, A. W. (2019). Estimation of clinical trial success rates and

related parameters. *Biostatistics*, 20(2):273–286.

World Health Organization (2018), Pricing of cancer medicines and its impacts: A comprehensive technical report for the World Health Assembly Resolution 70.12 Operative paragraph 2.9 on pricing approaches and their impacts on availability and affordability of medicines for the prevention, World Health Organization, <https://apps.who.int/iris/bitstream/handle/10665/277190/9789241515115-eng.pdf?ua=1>

## Figures

Figure 1: Process of database construction for non-rare cancer clinical trials

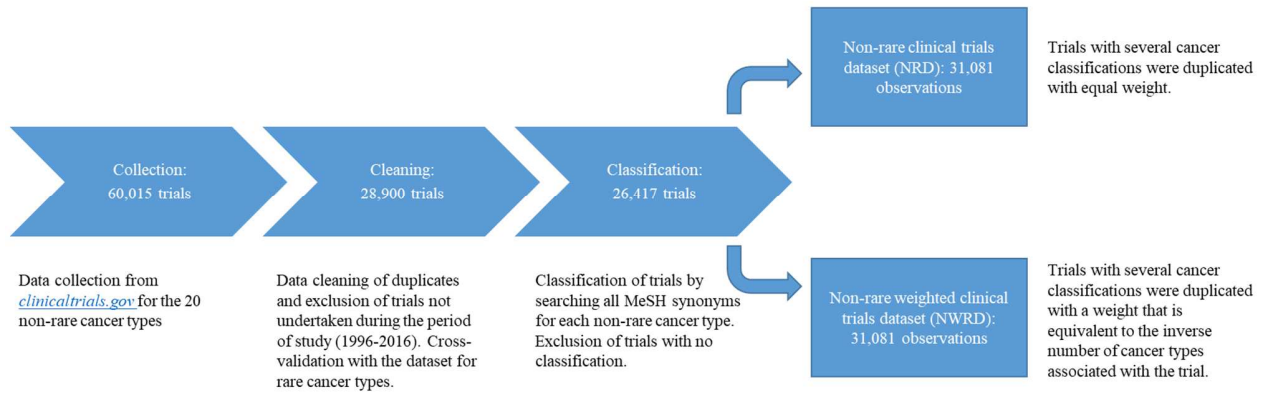


Figure 2: Process of database construction for rare cancer clinical trials

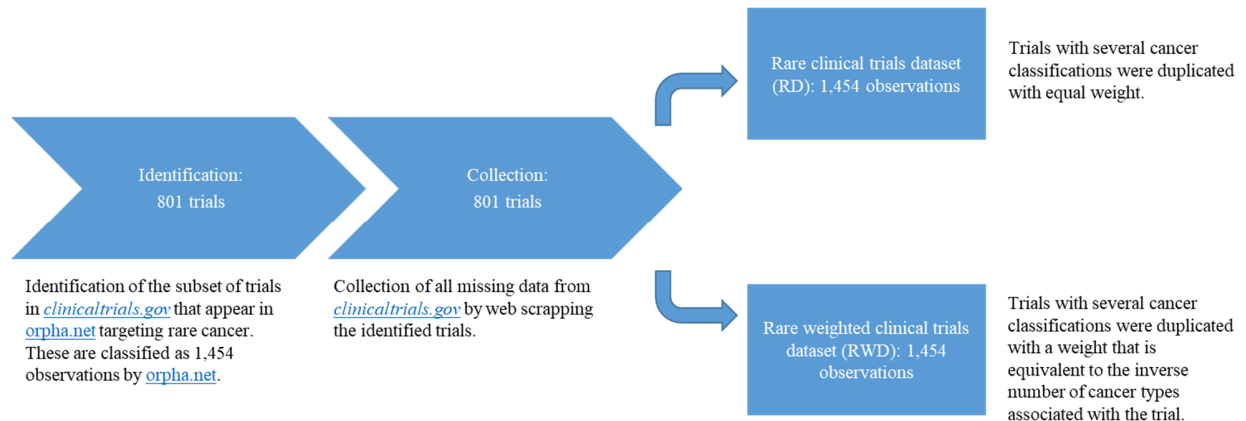
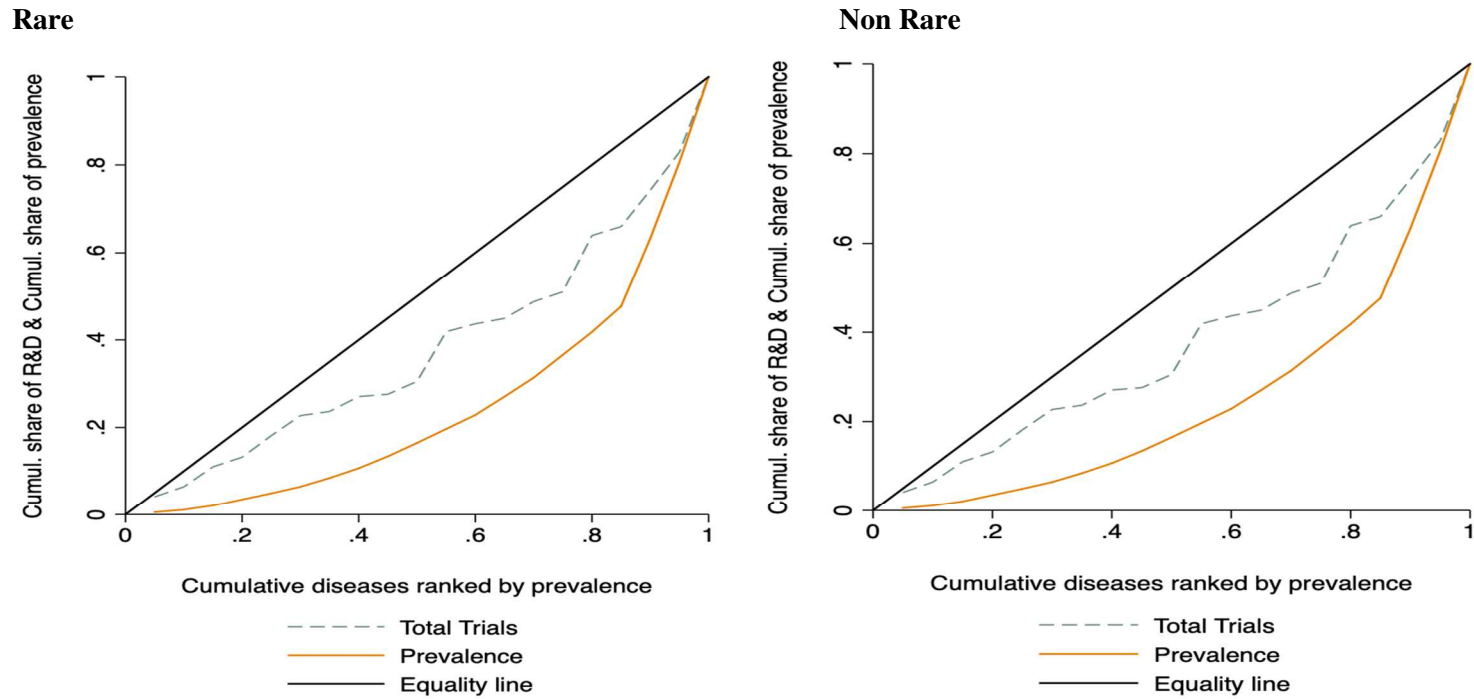
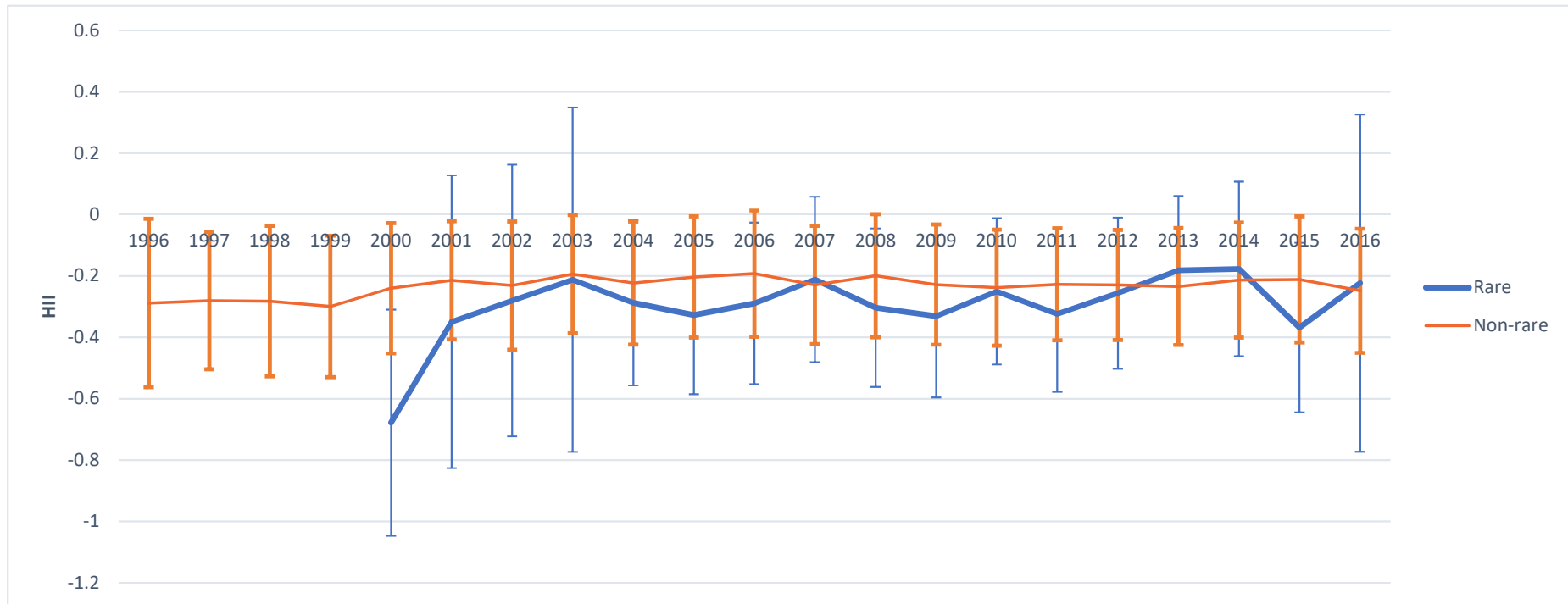


Figure 3: Concentration curves for total trials and prevalence, using non-weighted data split for rare and non-rare.



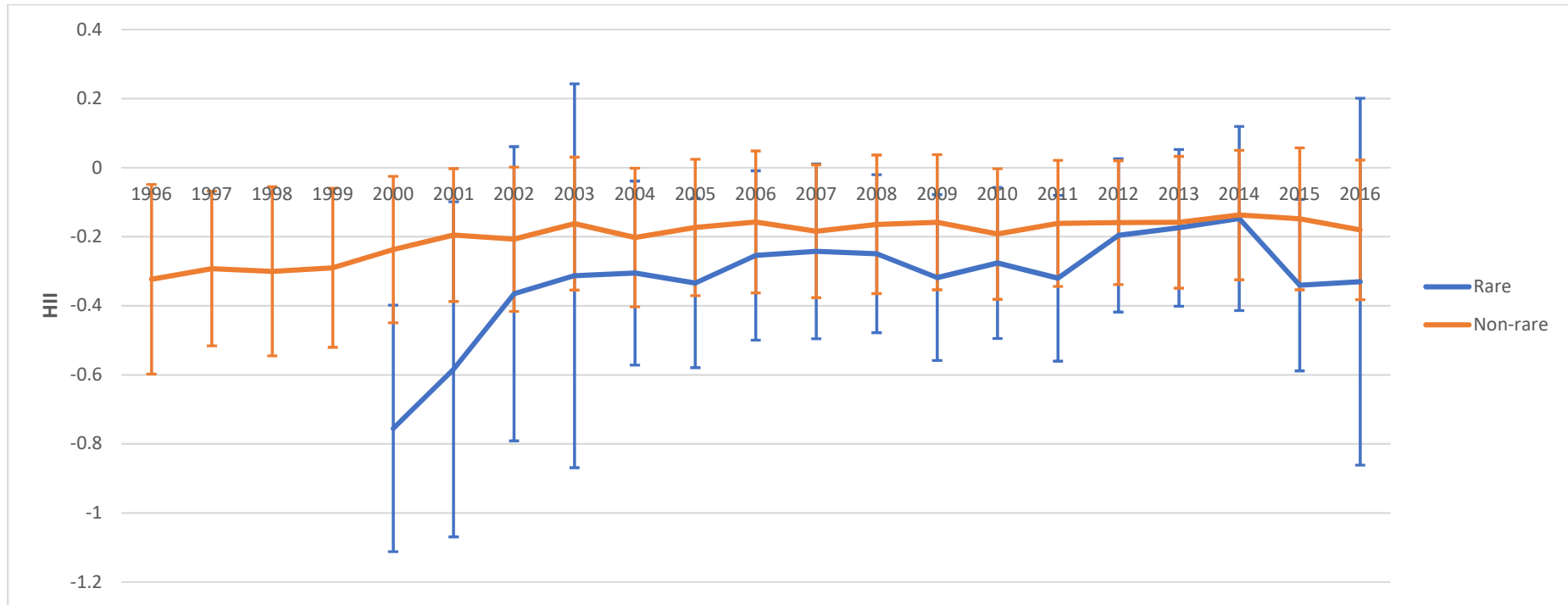
Note: The curves plot the cumulative percentage of the total (unweighted) number of clinical trials and the cumulative share of prevalence, split by rare and non-rare cancers, against the cumulative percentage of cancer types ranked by prevalence.

Figure 4a: Horizontal Inequity indices with error bars for total unweighted trials over time, calculated for rare and non-rare cancers separately, using prevalence as need variable.



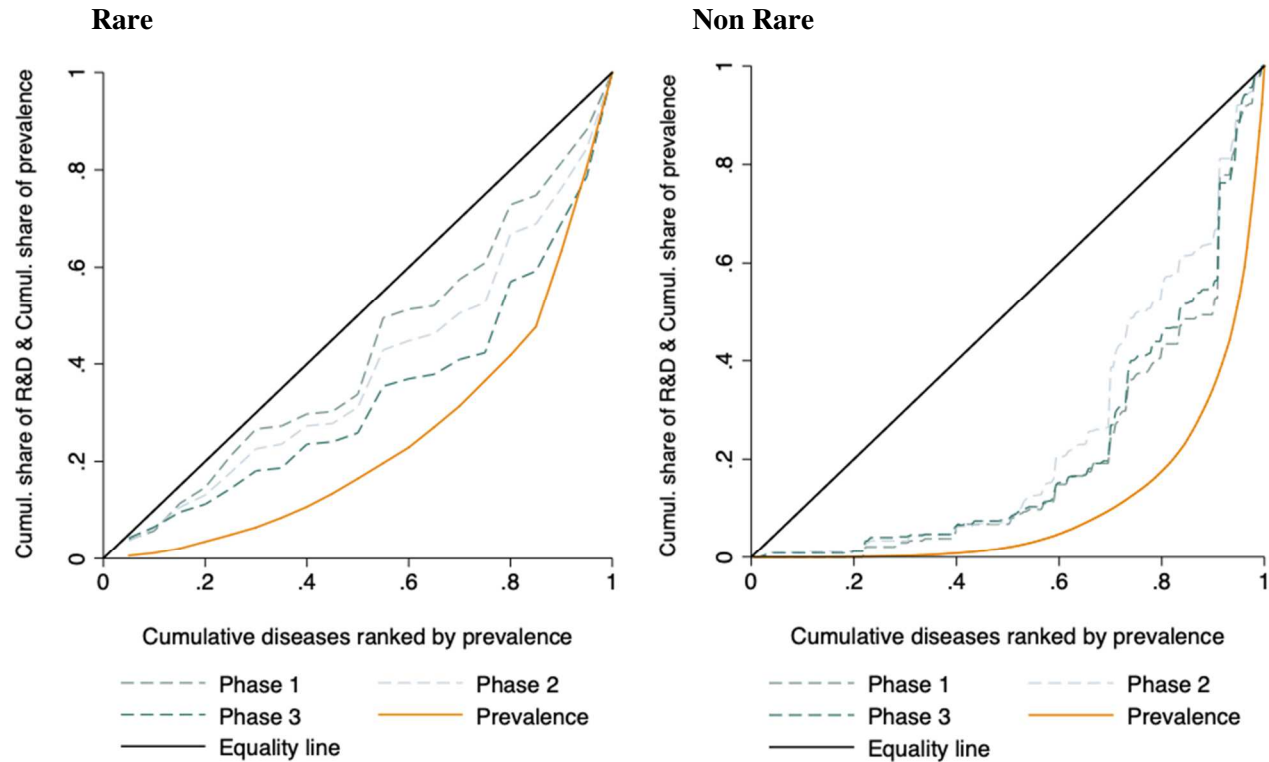
Note: HIIs are calculated for rare and non-rare cancers separately by plotting concentration curves of the cumulative percentage of the number of total clinical trials against the cumulative percentage of both the prevalence and incidence. Error bars calculated using 90% confidence intervals. Due to a very small number per year of trials in rare cancers (i.e., 99% rare cancers without any trial), point estimates are not computed.

Figure 4b: Horizontal Inequity indices with error bars for total unweighted trials over time, calculated for rare and non-rare cancers separately, using incidence as need variable.



Note: HIIs are calculated for rare and non-rare cancers separately by plotting concentration curves of the cumulative percentage of the number of total clinical trials against the cumulative percentage of both the prevalence and incidence. Error bars calculated using 90% confidence intervals. Due to a very small number per year of trials in rare cancers (i.e., 99% rare cancers without any trial), point estimates are not computed.

Figure 5: Concentration curves for prevalence and total trials in Phase 1, Phase 2 and Phase 3 for rare and non-rare cancers separately, using prevalence as need variable.



Note: The curves plot the cumulative percentage of unweighted clinical trials by Phase 1, Phase 2 and Phase 3 trials and the cumulative share of prevalence, split by rare and non-rare cancers, against the cumulative percentage of cancer types ranked by prevalence.



Table 1.a) T-tests of the epidemiological variables and R&D proxies for the 1996-2016 period.

Outcome	Non-rare	Rare	Difference
Prevalence	93.52 (23.48)	4.64 (0.72)	88.87*** (7.51)
Incidence	30.22 (6.29)	0.58 (0.08)	29.65*** (1.93)
Survival rate 5 years	48.46 (1.88)	51.45 (6.51)	-3.01 (6.39)
Clinical trials	1554 (309.09)	7.02 (1.65)	1547.03*** (94.21)
Enrollment	103.22 (14.44)	176.05 (3.73)	-72.83 (57.06)
Duration	3.88 (0.02)	3.76 (0.08)	0.11 (0.10)

Note: Standard errors in parentheses. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

Table 1.b) Number of clinical trials and weighted clinical trials of the top 20 non-rare cancers and rare cancers types, between 2000 and 2016.

<b>Non-Rare Cancers</b>	<b>No. of trials</b>	<b>No. of weighted trials</b>	<b>% of non-rare trials</b>	<b>% of non-rare weighted trials</b>
Lip, oral and larynx cancer	169	125	0.52%	0.48%
Thyroid cancer	296	266	0.95%	1.02%
Uterine cancer	393	307	1.26%	1.17%
Non-melanoma skin cancer	554	358	1.78%	1.37%
Bladder cancer	630	501	2.03%	1.92%
Chronic lymphoid leukemia	683	311	2.20%	1.19%
Malignant skin melanoma	695	472	2.24%	1.81%
Esophageal cancer	719	506	2.31%	1.93%
Kidney cancer	1058	854	3.40%	3.27%
Stomach cancer	914	704	2.94%	2.69%
Non-Hodgkin lymphoma	1189	956	3.83%	3.66%
Liver cancer	1283	1096	4.13%	4.19%
Pancreatic cancer	1420	1218	4.57%	4.66%
Ovarian cancer	1421	1138	4.57%	4.35%
Brain and nervous system cancer	1533	1310	4.93%	5.01%
Colon and rectum cancer	2570	2175	8.27%	8.32%
Prostate cancer	2687	2491	8.65%	9.53%
Leukemia	3529	2974	11.35%	11.37%
Tracheal, bronchus, and lung cancer	4031	3559	12.97%	13.61%
Breast cancer	5314	4827	17.10%	18.46%
<i>Total</i>	<i>31081</i>	<i>26147</i>	<i>100%</i>	<i>100%</i>
<i>Sample Total</i>	<i>31081</i>	<i>26147</i>		
<b>Rare Cancers</b>	<b>No. of trials</b>	<b>No. of weighted trials</b>	<b>% of rare trials</b>	<b>% of rare weighted trials</b>

Other myeloproliferative neoplasms	18	8	1.24%	1.00%
Neuroblastoma and ganglio- neuroblastoma	20	4	1.38%	0.50%
Astrocytic tumors of CNS	21	3	1.44%	0.87%
T cutaneous lymphoma	23	7	1.58%	0.87%
Rhabdomyosarcoma	25	2	1.72%	0.25%
Hepatocellular carcinoma of liver and IBT	29	10	1.99%	1.25%
Hodgkin lymphoma classical	32	7	2.20%	0.87%
Diffuse B lymphoma	33	4	2.27%	0.50%
Endocrine carcinoma of pancreas & digestive system	36	11	2.48%	1.37%
Mantle cell lymphoma	41	10	2.82%	1.25%
Precursor B/T lymphoblastic leukemia/lymphoma	42	14	2.89%	1.75%
Salivary gland type tumors of head and neck	42	11	2.89%	1.37%
Follicular B lymphoma	47	14	3.23%	1.75%
Malignant/immature teratomas of ovary	50	14	3.44%	1.75%
Chronic myeloid leukemia	53	20	3.65%	2.50%
Other T cell lymphomas & NK cell neoplasms	66	20	4.54%	2.50%
Plasmacytoma/multiple myeloma	101	30	6.95%	3.75%
Soft tissue sarcoma of head and neck	133	25	9.18%	3.12%
Acute myeloid leukemia	269	139	18.50%	17.35%
<i>Total</i>	<i>1104</i>	<i>583</i>	<i>75.93%</i>	<i>44.94%</i>
<i>Sample Total</i>	<i>1454</i>	<i>801</i>		

Table 2. Concentration Indices calculated for rare and non-rare cancers separately using four measures to proxy need: 1) prevalence; 2) incidence; 3) 1-year survival; and 4) 5-year survival

	<b>Prevalence</b>		<b>Incidence</b>		<b>1-year survival rate</b>		<b>5-year survival rate</b>	
	Rare	NonRare	Rare	NonRare	Rare	NonRare	Rare	NonRare
<i>Total</i>	0.541*** (0.131)	0.261** (0.108)	0.544*** (0.131)	0.245** (0.110)	-0.047 (0.136)	-0.022 (0.124)	-0.025 (0.136)	-0.054 (0.123)
<i>Pre</i>	0.341** (0.170)	0.215* (0.107)	0.223 (0.171)	0.158 (0.112)	0.168 (0.171)	-0.035 (0.118)	0.207 (0.171)	-0.066 (0.117)
<i>Post</i>	0.546*** (0.133)	0.265** (0.109)	0.553*** (0.132)	0.252** (0.111)	-0.052 (0.138)	-0.021 (0.126)	-0.031 (0.138)	-0.053 (0.125)
<i>Phase 1</i>	0.591*** (0.161)	0.165 (0.108)	0.611*** (0.161)	0.155 (0.109)	-0.089 (0.166)	-0.096 (0.112)	-0.029 (0.166)	-0.113 (0.111)
<i>Phase 2</i>	0.504*** (0.123)	0.245** (0.106)	0.485*** (0.123)	0.228** (0.108)	0.006 (0.128)	-0.046 (0.120)	0.018 (0.128)	-0.073 (0.119)
<i>Phase 3</i>	0.565*** (0.139)	0.354** (0.126)	0.590*** (0.138)	0.334** (0.129)	-0.096 (0.144)	0.021 (0.151)	-0.082 (0.144)	-0.021 (0.151)
<i>Person Years</i>	0.551*** (0.120)	0.435** (0.163)	0.566*** (0.120)	0.398** (0.169)	-0.027 (0.126)	0.152 (0.190)	-0.006 (0.126)	0.093 (0.192)
<i>Duration</i>	0.543*** (0.132)	0.222** (0.104)	0.551*** (0.132)	0.194* (0.107)	-0.044 (0.138)	-0.049 (0.116)	-0.020 (0.138)	-0.075 (0.115)

Note: Standard errors in parentheses. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

Table 3. Horizontal Inequity Index (HII) calculated for rare and non-rare cancers separately using four measures to proxy need: 1) prevalence; 2) incidence; 3) 1-year survival; and 4) 5-year survival

	Prevalence		Incidence		Survival rate 1 year		Survival rate 5 years	
	Rare	NonRare	Rare	NonRare	Rare	NonRare	Rare	NonRare
<i>Total</i>	-0.247* (0.148)	-0.277** (0.101)	-0.229* (0.133)	-0.231** (0.104)	-0.253* (0.136)	-0.219 (0.126)	-0.348** (0.136)	-0.382*** (0.126)
<i>Pre</i>	-0.447** (0.190)	-0.323** (0.125)	-0.550*** (0.184)	-0.318** (0.124)	-0.038 (0.171)	-0.232* (0.118)	-0.117 (0.171)	-0.394*** (0.118)
<i>Post</i>	-0.242 (0.150)	-0.273** (0.100)	-0.220 (0.134)	-0.224** (0.104)	-0.258* (0.138)	-0.218 (0.128)	-0.354** (0.138)	-0.381*** (0.127)
<i>Phase 1</i>	-0.197 (0.174)	-0.373*** (0.125)	-0.162 (0.162)	-0.321** (0.117)	-0.295* (0.166)	-0.293** (0.113)	-0.352** (0.166)	-0.441*** (0.113)
<i>Phase 2</i>	-0.284** (0.143)	-0.294** (0.110)	-0.288** (0.130)	-0.248** (0.105)	-0.200 (0.128)	-0.243* (0.122)	-0.305** (0.128)	-0.401*** (0.121)
<i>Phase 3</i>	-0.223 (0.154)	-0.185* (0.105)	-0.183 (0.136)	-0.142 (0.117)	-0.302** (0.144)	-0.175 (0.154)	-0.405*** (0.144)	-0.349** (0.154)
<i>Person-years</i>	-0.237* (0.133)	-0.104 (0.134)	-0.207* (0.119)	-0.078 (0.151)	-0.233* (0.126)	-0.045 (0.191)	-0.329*** (0.126)	-0.235 (0.192)
<i>Duration</i>	-0.245 (0.150)	-0.316** (0.118)	-0.222 (0.135)	-0.282** (0.109)	-0.250* (0.138)	-0.246** (0.117)	-0.343** (0.138)	-0.403*** (0.116)

Note: Standard errors in parentheses. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01.

Table 4. HII over time calculated for rare and non-rare cancers separately using prevalence and incidence as measures of need

Year	Prevalence		Incidence	
	Rare	NonRare	Rare	NonRare
1996		-0.313*		-0.352**
		(0.152)		(0.153)
1997		-0.337***		-0.322**
		(0.100)		(0.113)
1998		-0.336**		-0.337**
		(0.135)		(0.132)
1999		-0.341**		-0.330**
		(0.131)		(0.127)
2000	-0.499*	-0.298**	-0.565**	-0.285**
	(0.280)	(0.122)	(0.273)	(0.120)
2001	-0.349	-0.299***	-0.584**	-0.268**
	(0.290)	(0.098)	(0.295)	(0.103)
2002	-0.280	-0.280**	-0.365	-0.251**
	(0.269)	(0.115)	(0.259)	(0.116)
2003	-0.212	-0.249**	-0.313	-0.219*
	(0.341)	(0.113)	(0.338)	(0.111)
2004	-0.263	-0.272**	-0.286*	-0.253**
	(0.161)	(0.110)	(0.161)	(0.109)
2005	-0.291**	-0.256**	-0.328***	-0.232**
	(0.125)	(0.105)	(0.122)	(0.110)
2006	-0.261	-0.241**	-0.236	-0.211*
	(0.163)	(0.108)	(0.158)	(0.112)
2007	-0.327*	-0.276**	-0.399**	-0.232**
	(0.194)	(0.102)	(0.187)	(0.105)
2008	-0.223	-0.259**	-0.179	-0.226*
	(0.204)	(0.107)	(0.189)	(0.109)
2009	-0.276*	-0.282**	-0.250**	-0.223**
	(0.144)	(0.106)	(0.124)	(0.105)
2010	-0.223	-0.292***	-0.203	-0.254**
	(0.145)	(0.097)	(0.127)	(0.104)
2011	-0.288*	-0.281***	-0.263*	-0.221**
	(0.163)	(0.096)	(0.150)	(0.100)
2012	-0.258	-0.290***	-0.218	-0.230**
	(0.160)	(0.088)	(0.145)	(0.098)
2013	-0.198	-0.270**	-0.151	-0.204*
	(0.151)	(0.096)	(0.137)	(0.104)
2014	-0.175	-0.274***	-0.150	-0.201*
	(0.204)	(0.094)	(0.191)	(0.101)
2015	-0.147	-0.269**	-0.063	-0.205*
	(0.340)	(0.104)	(0.326)	(0.106)
2016	-0.003	-0.281**	0.038	-0.218*
	(0.442)	(0.108)	(0.432)	(0.108)

Note I: Standard errors in parentheses. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01. Note II: Point estimates for years 1996 to 1999 were omitted since only 13 clinical trials occurred in this period for all

207 different rare cancers and thus the HII's yearly estimates may be unreliable due to the lack of clinical trials.