(In)-equality of opportunity in the allocation of R&D resources for rare diseases

Preliminary version – Please do not quote without authors' permission

Setti Raïs Ali[1]

Hospinnomics, Fondation Imagine for rare diseases, Paris 1 Panthéon-Sorbonne (France)

Sandy Tubeuf

Université Catholique de Louvain (Belgium) and University of Leeds (UK)

Abstract

This paper investigates the existence of inequalities of opportunity in the allocation of R&D resources within rare diseases and identifies the characteristics of rare diseases that appear to lead R&D investments. Rare diseases affect less than 1 in 2,000 citizens. With over 7,000 recognized rare diseases and 350 million people affected worldwide, rare diseases are not so rare when considered collectively. Rare diseases are generally underserved by drug development because pharmaceutical industries consider R&D investments in rare diseases too costly and risky in comparison with the low expected returns due to the small population involved. We use data on rare diseases research from Orphanet and academic publications from MEDLINE and test the existence of inequalities using dominance tools and bilateral tests. We show that rare diseases in children and with a smaller prevalence, such as ultra-rare diseases, are underserved by R&D. R&D efforts appear to be concentrated in more profitable research areas with potentially larger sample size for trials design and adult population.

Keywords: R&D; Equality of opportunity; FOSD; Rare Diseases; Orphan Drug
JEL codes: I11- I14- I18 - L65

[1] Corresponding author contact details: Setti Raïs Ali – Hospinnomics (PSE/AP-HP) – 1, Parvis Notre-Dame, 75004 Paris – Phone number: 0668303070 – Email: setti.rais@psemail.eu

I. Introduction

A disease is characterized as rare if it affects less than 1 in 2,000 citizens, which represents 250,000 or fewer patients in the European Union (Drummond & Towse, 2014). While over 7,000 recognized rare diseases with 80% of them being genetic, a total of 350 million people are affected worldwide, and so patients with rare diseases are not very rare when considered collectively (Giannuzzi et al., 2017). The diagnostic of rare diseases may be very challenging, and often the causes and features of rare diseases remain elusive. The course of the disease is often unpredictable, and most of the recognised rare diseases are debilitating and/or life threatening (Field & Boat, 2010). Rare diseases can affect anyone, at any age and are associated with significant health needs (Schieppati et al., 2008). Patients with rare diseases generally face a poor health status because of the disease itself but also because their health care pathway to accessing appropriate diagnosis and treatment for their condition can be lengthy and complicated. The costs of drug development targeting rare diseases are particularly high as industries have difficulties in recruiting patients in clinical trials (Gericke et al., 2005). Pharmaceutical industries consider R&D investments in rare diseases too costly and risky in comparison with the low expected returns due to the small population involved. Consequently, patients with rare diseases are underserved by drug development. The pharmaceutical sector is an highly regulated sector from the very first step of translational research to the market authorization of the drug and marketing (Scott Morton & Kyle, 2011). While pharmaceutical firms naturally pursue a revenue maximization exercise, the regulator is in position to endorse ethical considerations and impact the allocation of R&D investments by increasing firms' profitability in underserved research areas. Despite governmental initiatives providing incentives for pharmaceutical firms to invest in rare diseases enacted in 2000 with the European Union Orphan Drug regulation[2], it is estimated by the National Center for Advancing Translational Sciences[3] that 95% of rare diseases do not have treatment options in 2018. Given that disparities in investment decisions determine patients' access to treatments, the allocation of R&D resources is a determinant of inequalities in access to care in the whole population (Williams & Cookson, 2000). There have been considerable discussions in the philosophical and political economy literature about the role of the welfare state in promoting equity in the provision of certain goods and services (Cookson & Dolan, 2000; Hughes et al., 2005; Martin et al. 2002; Temkin, 2003). The regulation schemes in pharmaceutical markets directly impact a fairer distribution of R&D investments across diseases in need of appropriate treatment and indirectly impact treatment and care opportunities and ultimately health status of patients with rare diseases. Moreover,

[2] Regulation (EC) No 141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products (OJ L 18, 22.1.2000, p.1), last amended by Regulation (EC) No 596/2009 (OJ L 188, 18.07.2009, p. 14)
[3] See: https://ncats.nih.gov/

several studies that were conducted on the relationship between pharmaceutical innovations and mortality, suggest that the launching of new drugs decreases mortality in various contexts and therapeutic areas (Lichtenberg, 2001, 2014a, 2014b, 2017). In this context, the social justice literature can offer a pertinent framework to delineate the responsibility of the public health sector in tackling health inequities and setting priorities that benefit disadvantaged groups. Following Daniels (1985), there is a moral right of health care according to which health care is a concern of justice and promotes equal opportunity by preventing and curing diseases. The standpoint for equal opportunity is to identify an equal distribution of natural assets and if not, asks whether the disease threatens opportunity (Dworkin, 1981a, 1981b; Roemer, 1998). Justice requires compensating those with less desirable natural assets to overcome a disadvantaged natural lottery. It seems wrong that a person's opportunities should be limited by genetic endowments (Buchanan et al., 2000).

Our approach adopts a less standard framework for inequality of opportunity where we are not interested in individuals' outcome but we consider diseases as being the observations of importance here. We assess whether there are inequalities of opportunities in R&D distribution within rare diseases. We aim to uncover which diseases characteristics appear to encourage R&D within rare diseases in addition to the population size to benefit when the sample of affected patients is actually very small. To the best of our knowledge, no academic work has yet explored the fairness of the distribution of pharmaceutical R&D resources across diseases in general and across rare diseases more specifically.

The objectives of this paper are twofold. Firstly, it will investigate whether opportunities are equal within in the allocation of R&D resources in rare diseases using cumulative distribution functions and stochastic dominance tests. Secondly, it will identify the characteristics of rare diseases that appear to lead R&D resources. R&D resources are successively measured using five alternative proxies: the number of clinical trials per rare disease, the number of research projects per rare disease, the number of approved drugs with marketing authorisation at the European level, the number of orphan drugs designations and the number of published articles per rare disease on MEDLINE. We appraise rare diseases characteristics with the Orphanet data using condition-specific mean age at death, mean age at first onset, disease prevalence, and two binary characteristics: uncertain disease evolution versus certain, and immediate danger of death versus non-immediate.

Our results suggest that R&D investments underserve rare diseases that occur in infancy and that affect a smaller number of patients; this is observed for most of our R&D proxies. R&D efforts are possibly concentrated in rare diseases where there are higher chances of finding an effective treatment because when a disease affects more patients, more patients will be able to join a clinical trial. The other characteristics that appear to lead R&D resource allocation for rare diseases include the mean age at symptoms appearance, a larger market size, a lower level of uncertainty regarding the disease presentation

and progression, and non-immediate danger of death.

The paper is structured as follows. Section 2 presents the framework of inequality of opportunity. Section 3 presents the data, and section 4 the method. Section 5 introduces the empirical application on rare diseases. Discussion and concluding remarks are in Section 6.

II. The inequality of health opportunity: a conceptual framework

a. How to allocate R&D investments across diseases?

The question of how much resources should be invested in R&D diseases, especially rare diseases, which are complex and concern a heterogeneous group and a small proportion of people, is a moral dilemma for policymakers (McCabe et al., 2005, Paulden et al., 2014). It relates to whether policymakers should consider the distribution of the number of drugs brought to the market when considering the level of inequity in health care faced by patients with rare diseases while they also regulate the process to drug development and create opportunities of drug development for patients with rare diseases. One approach, called *substantive justice* refers to the allocation of outcomes within the patients population, while another approach referred as *procedural justice* considers the process to the outcome (Cookson & Dolan, 2000). As the level of R&D is likely to impact future health attainments, the allocation of R&D resources can be considered as one of the instruments to increase fairness in health care and equity in health itself.

Justice principles that draw upon social justice literature and embody different views on fairness may be categorized into three broad relevant groups in the context of R&D resources allocation. First, according to the maximization principle (Bentham & Mill, 2004), policymakers should aim at maximizing the total sum of health within the population. Hence, a particular attention should be given to capacity to benefit from public resources. No extra weight would be given to any particular patients group, whatever the level of their health needs and the severity of the disease. Second, equality is the most important feature of egalitarianism (Nord et al., 1995; Temkin, 2003), which considers that people ought to be treated as equals, regardless of their characteristics. Finally, prioritarian principles (Otsuka, 2013; Temkin, 2003) give emphasis to health needs; they stipulate that the most severely ill categories of patients should receive priority according to the "Rule of Rescue" (McKie & Richardson, 2003) and regardless of their capacity to benefit from public funding. The World Health Organization[4] recommends prioritizing "those with the greatest need", even in settings where resources are substantially constrained. Similarly the consideration

[4] Human rights and health: http://www.who.int/news-room/fact-sheets/detail/human-rights-and-health

of patients with needs for highly specialized treatments is emphasized in the European Commission[5], which explicitly mentions the right of patients with a rare disease to be entitled to the same quality of treatment as any other patients. While those principles are likely to guide decision makers, they shall be considered in conjunction with the trade-offs decision makers inevitably have to make between the advantages and disadvantages of each health care decision. The use of public resources needs to be justified towards the general public and taxpayers while including equitable considerations. For example, decision makers consider the diseases' characteristics, such as the burden of illness and the severity of the condition, as well as the population size to benefit from the treatment within reimbursement decisions in health care (Thébaut & Wittwer, 2017). Finally, studies conducted on general population regarding priority-setting in the allocation of public resources suggest that society supports social justice considerations by expressing preferences for the distribution of public resources in favour of deprived categories of patients, regardless of the opportunity cost in healthcare provision and how priority-setting may divert resources away from other categories of patients (Brazier et al., 2013).

All these elements acknowledge that the approach of fairness of policymakers promotes equity. Since politics care about "who gets what and when", the allocation of pharmaceutical R&D resources should be a major concern for policymakers as well as a substantial instrument to promote health of patients with rare diseases. More importantly, while policymakers explicitly endorse ethical considerations in the decision-making process, they do not clearly disclose the characteristics of diseases that should be prioritized. For example, distribution of public resources may have the objective of prioritizing diseases presented with a higher probability of premature death in patients. Another possible way to address equity considerations within rare disease could be to aim at equalizing the life expectancy across these rare diseases.

b. *The concept of equality of opportunity*

In the past decade, the analysis of health inequalities and health related outcomes has adopted a normative understanding of those inequalities and their health determinants calling upon the philosophical literature regarding social justice and especially the concept of equality of opportunity (Arneson, 1989; Barry, 2005; Cohen, 1989; Dworkin, 1981a, 1981b; Fleurbaey, 2008; Roemer, 1998). According to this concept, some sources of inequality are more objectionable than others and could represent priorities for policies aiming at reducing health inequalities. Equality of opportunity draws a distinction between "legitimate" and "illegitimate" sources of differences in health disparities. While legitimate differences can be attributed to determinants within people's control, illegitimate differences are related to determinants beyond people's

[5] European Commission memo: http://europa.eu/rapid/press-release_MEMO-14-141_en.htm

control often called circumstances and provide evidence of inequality in opportunity. Typically, parental socioeconomic status or country of birth, are determinants beyond people's control because individuals did not choose them. Similarly individuals do not choose their gender at birth, ethnicity, social background, and any other characteristics that are exogenous to them in the sense that they cannot be influenced by their actions (Ramos & Van De Gaer, 2015; Roemer & Trannoy, 2015). On the other hand, there are variables that can be subject to individual choice, such as health care preferences and lifestyles. The partition between the variables that should belong to the illegitimate or the legitimate determinants is debatable and largely depends on ethical views. Moreover characteristics under individual control are challenging to observe and measure, and the extent that an individual should be held responsible for his preferences is controversial among health economists (Fleurbaey & Schokkaert, 2009; Jusot et al., 2013).

While empirical investigations of the equality of opportunity in health and income put the emphasis of inequality of opportunity in relation to socioeconomic status (Bourguignon et al., 2007; Ferreira & Peragine, 2015; García-Gómez et al., 2014; Jusot et al., 2013), the equality of opportunity framework can accommodate very diverse investigations in outcome distribution. We propose here to adapt this framework to the distribution of R&D investments across rare diseases. The unit of analysis in this context is not individuals but rare diseases and we study whether opportunities in R&D investments are equal within rare diseases according to specific characteristics of rare diseases.

From our ethical perspective, specific characteristics of rare diseases, such as the average age at symptoms appearance, the prevalence or the average age at death, each represent one circumstance. Equality of opportunity in health care would require for R&D investments in rare diseases to be distributed independently of the characteristics of the disease. In the context of R&D investments allocation for rare diseases, differences in outcomes are the result of deterministic and random factors. We do not argue that our empirical investigation will provide estimates of the magnitude of inequality of opportunity in R&D investments for patients with rare diseases; neither provides a comprehensive set of the determinants of inter-individual differences in R&D investments across rare diseases. Our analysis is meant to identify the sub-groups of rare diseases that are under-served by R&D and that could be targeted by policymakers in search of more equitable distribution of R&D investments across rare diseases.

III. Data

We investigate the inequity in the allocation of R&D resources using data from Orphanet, which is the reference portal providing information about orphan drugs and rare diseases. Orphanet was established in France by the INSERM (French National Institute for Health and Medical Research) in 1997. This

initiative then became European from 2000 and gradually grew to a Consortium of 40 countries within Europe and across the globe[6]. The Orphanet dataset comprises about all rare diseases, granting them a unique Orphanet identification number to facilitate sharing information on each disease.

a. R&D resources outcome measures

Orphanet provides us with four different outcomes variables that can be used to proxy the R&D resources allocated to each of the rare diseases at the European level. We first use an inventory of clinical trials activities targeting rare diseases. Clinical activities include interventional studies treating or preventing a rare disease using drugs, combination of drugs and biological products. Second, we use the list of research projects targeting each rare disease. Research projects are projects that have been selected through a competitive process established by a scientific committee, or issued from a national research funding. Clinical trials activities and research projects include both single-centre and national and international multicentre research projects at the European level. Third, Orphanet provides us with the number of orphan drugs designations that qualify for the financial incentives provided by the EU Orphan Drug legislation. Finally, we consider the number of drugs with marketing authorisation at the European level per rare disease (we refer to them as orphan drugs).

The four outcomes proxies for R&D investments are completed with an outcome of published research on rare diseases, which is measured by the number of scientific publications per disease. We accessed MEDLINE using PubMed search engine in July 2017 from its inception date to present using the MEDGEN unique identifier of the 8755 diseases classified as rare diseases and we counted the number of scientific publications for each rare diseases. MEDLINE is the largest database of academic references on life sciences and biomedical topics and our search was based on an algorithm coded in Python.

Table 1 displays the descriptive statistics of the R&D resources outcome measures. There is a total of 9220 rare diseases and most of them attract almost no R&D resources. The mean number of research projects, clinical trials, orphan designation and orphan drugs appears to be very low, ranging between 0.12 and 0.72, the median being 0 for each of the outcome. The fifth quartile is equal to 0 for research projects, clinical trials, orphan designation and orphan drugs, suggesting the absence of any investments for a vast majority of rare diseases. The number of academic publications captures the knowledge currently built on each rare disease; this includes for example the natural history of the disease, information on diagnostic criteria, and the impact of the disease on quality of life and health status. The mean number of academic publications per rare diseases is approximately 578 [median=85], while the maximum reached for one of the rare diseases is 177,430 articles. We present in Table 2 the linear correlation coefficient between all

[6] See: https://www.orpha.net/consor/cgi-bin/index.php

the R&D resources outcome measures. Correlations range from 0.16 to 0.69; this suggests that R&D resources outcome measures capture different aspects of R&D but are positively correlated. In particular, some R&D resources represent investments corresponding to different phases of drug development, which are related. For example, the number of clinical trials is correlated with the number of orphan drugs with a linear correlation coefficient is 0.63 and this is explained by clinical trial activities being a prerequisite for market approval.

b. Rare diseases characteristics

Rare disease characteristics were provided by the Orphanet dataset and include the following variables: the average age at first symptoms appearance, the average age at death, and the prevalence in the population.

The average age at symptoms appearance for each disease was not provided as a single age but as a category among a choice of four categories: *Infancy, Childhood, Adults & Elderly* and *All ages*. The average age at death for each disease was also available as a category including five possible categories: *Infancy, Childhood, Adults & Elderly, Normal Life Expectancy* and *All ages*. The prevalence of each rare disease in the population was sometimes provided as a value (25%) but most of the time provided as an interval (75%); the latter mainly happens because the uncertainty around the number of patients with the condition is high. We homogenised the values and intervals using intervals overlaps and the mid-point of each intervals to construct a discrete variable of prevalence in 4 categories (<1 over 1,000,000; 1 to 9 over 1,000,000; 1 to 9 over 100,000; 1 to 9 over 100,000). We then created two binary variables using the same data. First, we created a dummy variable representing an *Immediate Danger of Death* equals to one when the age of first symptoms appearance equals the average age of death category. Second, we constructed a dummy variable measuring the *Uncertainty on Disease Evolution* equals to one if the age of symptoms appearance and/or the mean age at death in classified as unpredictable.

Table 3 presents the distribution of the average *age at symptoms onset* and suggests that the four categories *Infancy, Childhood, Adults & Elderly* and *All ages* are balanced in our sample, which means that rare disease do not affect an age group disproportionally and so may appear at any point in life . On the contrary, the average *age at death* show great discrepancies in distribution across the age groups (Table 4); almost half of the rare diseases are characterised with an average age at death that is unpredictable (*All ages*) and only 22% of the rare diseases are given a *normal life expectancy*. Figure 2 displays the frequency distribution for the *prevalence* variable and suggests that rare diseases prevalence is highly skewed toward 0. For 77% of rare diseases in the sample, the prevalence is under one case for 1,000,000 individuals. This suggests that rare diseases are mainly ultra-rare.

In Table 3, we investigate the relationships between all the rare disease characteristics using the Cramer's

V[7] statistics. The age of symptoms onset is by construction related to the mean age at death in the sense that the patient cannot be at risk of death before symptoms' appearance and so the Cramer's V is 0.46. The relationships are weaker between the other variables: the association between the mean age at death and the prevalence is 0.19 while it is 0.24 for the prevalence and the age at symptoms' onset.

All diseases characteristics were not always available for each rare disease in the Orphanet dataset. We studied more specifically the attrition in the dataset. The shared missing pattern for all variables are visually described in Figure 1. All the R&D investments (research projects, clinical trials, orphan designation and orphan drugs) variables for the 9,220 rare diseases are non-missing since they are directly provided in Orphanet and the count is equal to 0 in the absence of R&D investments. The search for academic publications provided us with 95% of correspondence between the Orphanet identification number and the MEDGEN unique identifier. These 5% missing values are shared with all the rare disease characteristics. Regarding the rare diseases characteristics, the average age at symptoms appearance and average age of death share most of their missing values, while prevalence is the rare disease characteristic with the lowest level of missing values. We further investigated missing values by comparing the average number of our R&D resources outcome measures for missing values versus non-missing values. Results are reported in Table 5 and suggest that missing values have on average a significantly lower number of research projects, academic publications, clinical trials, orphan designation and orphan drugs. Still, the average difference between missing and non-missing values is substantially low, and most of the rare diseases characteristics share the same missing values.

IV. Methods

We are especially interested in the share of R&D investments devoted to rare diseases and how it is distributed across rare diseases.

Loosely following Lefranc et al. (2009) and Lefranc and Trannoy (2016), we detect inequality of opportunity comparing cumulative distribution functions (CDF) of the R&D investments devoted to rare diseases conditioned on a set of variables representing diseases characteristics. These diseases characteristics represent the so-called 'circumstances' according to Roemer (1998).

The CDF of the number of academic publications in a rare disease with an average age at symptom onset classified in *Adult & Elderly* describes the distribution of opportunities in R&D investments of this rare disease. If on the one hand, this CDF is clearly different than the one of another rare disease with an

[7] The Cramer's V statistics indicates how strongly two categorical variables are associated (Sheskin, 2003). The statistics is ranging between 0 and 1, the maximum value indicating perfect relationship.

average age at symptom onset classified in *Infancy* and if, on the other hand, this difference is such that the rare disease has a higher chance of being invested and researched when the average age at symptom onset is classified in *Adult & Elderly*, one can reasonably associate this result to a difference in opportunities in R&D investments related to average age at symptom onset. This example is a typical situation of stochastic dominance at first order.

Let us consider two distributions A and B with respective cumulative distribution functions $F_A(y)$ and $F_B(y)$, and A dominates at first order B, written $A \geq_{FSD} B$ if and only if $F_A(y_j) \leq_{FSD} F_B(y_j)$, where y represents one of the five proxies of R&D investments for each rare disease and $y_j=\{ y_1,y_2,\ldots,y_k\}$.

It means that R&D investments is higher in distribution A than in distribution B and this is true at every points of the distribution. Graphically, the cumulative distribution function of R&D investments of the sub-group of rare diseases in B is always above that of rare diseases in A at any point of the distribution.
We compare the cumulative distribution functions of each five proxy of R&D investments. The five proxy variables are (a) the number of research projects, (b) the number of academic publications, (c) the number of clinical trials, (d) the number of orphan designations and (e) the number of orphan drugs with marketing authorization across age class of the disease symptoms. These variables are inherently discrete. Empirically, the inference procedure relies on tests of stochastic dominance at first order, such as unilateral Kolmogorov-Smirnov (KS) tests of equality of distribution, which are appropriate with discrete variables.
For each characteristic, we test the null hypothesis of equality of the distributions in pairs. Then, we test the null hypothesis of first-order stochastic dominance of the distribution of A over B, and the distribution B over A. If the test accepts dominance of one distribution over the other but not the other way round (e.g. $F_A(y_j) \leq_{FSD} F_B(y_j)$, and $F_B(y_i) \not\leq_{FSD} F_A(y_j)$), we consider that equality of opportunity is violated.
The same approach can be proposed when comparing sub-groups of rare disease according to any characteristic such as the average age at symptoms appearance, average age at death, prevalence, and two binary characteristics: uncertainty or not on disease evolution, and immediate or not danger of death.
It is important to underline that this approach remains relevant even when all disease characteristics are not observed or cannot be combined. According to Lefranc et al. (2006; 2009), equality of distributions conditional on circumstances is a necessary condition for equality of opportunity even if circumstances are not fully described. As a result, if the KS test shows significant differences between CDFs then we can say that equality of opportunity is violated if we had the opportunity to measure perfectly circumstances. This provides a rationale to perform first the non-parametric test separately on the CDF conditional on each disease characteristics individually, which is helpful because of the relative small size of the sample. We

then considered combining rare disease characteristics together in order to generate a set of diseases circumstances, however this was only possible with the prevalence level. We weighted the rare diseases according to their frequency in the population of patients with rare diseases along with each of the other disease characteristics.

V. Results

a. Non-parametric tests on each diseases characteristic

We compare the distributions of R&D investments as measured by five alternative proxies according to different rare disease characteristics and use the significance level of the differences between distributions using Kolmogorov Smirnov (KS) tests to conclude on the existence of stochastic dominance.

Average age at symptoms appearance - Results comparing the distribution of the five different proxies of R&D investments for rare diseases according to the four categories of age at symptoms appearance are presented in Table 6. They suggest that the distribution of all proxies of R&D investments targeting rare diseases occurring during *Infancy* are dominated by the distribution of any R&D investments of rare diseases with an average age at symptom onset classified in *All Ages* and in A*dult & Elderly*. All five proxies of R&D investments appear to favour rare diseases in older age groups. When rare diseases in *Infancy* are compared with rare diseases in *Childhood*, the distributions of the number of research projects, clinical trials and academic publications all favour rare diseases in *Childhood* (p-values respectively 0.006, 0.012, 0.061) however we cannot conclude on dominance when comparing the distribution of the number of orphan designations and the distribution of number of orphan drugs with marketing authorisation (p-values respectively 0.234, 0.701).

The distribution of most of the R&D proxies for rare diseases in *Childhood* and *Infancy* are dominated by the distributions for rare diseases in *Adult & Elderly* and *All Ages*, except for the distribution of the number of research projects with *All Ages* where the Kolmogorov Smirnov test is inconclusive (p-value=0.696). When considering *Adult & Elderly* versus *All Ages,* we find that for the distribution of two of the R&D outcomes (the number of research projects and academic publication) in *Adult & Elderly* dominate the distribution in *All Ages*, and the distribution of clinical trials in *All Ages* dominates the one in category *Adult & Elderly*. The KS tests remain inconclusive for the number of orphan designations and of orphan drugs (p-value respectively 0.771 and 0.990).

Average age at death - Results for the paired KS tests comparing the distribution of R&D investments for rare diseases over the five categories of the average age at death are presented in Table 7. They suggest that the distributions of R&D investments targeting diseases with an average age at death in *Infancy* are

dominated by the distributions of R&D investments for higher categories of average age at death (*Adult & Elderly, All Ages, Normal Life Expectancy*). This result holds for all R&D proxies, except for the distribution of the number of orphan drugs (p-values respectively 0.272, 0.417, 0.184). When rare diseases in *Infancy* are compared to rare diseases in *Childhood*, the distribution of the number of academic publications is in favour of diseases with mean age at death in *Childhood* (p-value=0.036). The dominance tests are inconclusive when we compare the distributions of the number of research projects, clinical trials, orphan designations and orphan drugs with marketing authorization (p-values respectively 0.136, 0.742, 0.832, 1.000). When considering rare diseases with an average age at death in *Childhood* versus rare diseases in *Adult & Elderly* or in *Normal Life Expectancy*, the distributions of all R&D investments, except the number of orphan drugs for the category *Adult & Elderly* (p-value=0.156), favour diseases in categories *Adult & Elderly* and *Normal Life Expectancy*. When considering rare diseases with an average age at death in *Childhood* versus *All Ages*, the distribution of academic research favours the category *All Ages* (p-value=0.065). We cannot conclude on dominance for the distribution the number of research projects, clinical trials, orphan designations, and orphan drugs.

When considering rare diseases with an average age at death in *Adults & Elderly* versus those in *All Ages*, results suggest that the distributions of most proxies of R&D for the category *Adults & Elderly* dominate the distributions of R&D for rare diseases with an unpredictable mean age of death. However the test cannot conclude regarding dominance between *Adults & Elderly* versus *All Ages* in the distribution of the number of orphan drugs (p=0.136). When considering rare diseases with an average age at death in *Adults & Elderly* versus those with *Normal Life Expectancy*, results suggest that the distributions clinical trials for the category *Adults & Elderly* dominate the distributions of R&D for rare diseases with *Normal Life Expectancy*. When considering rare diseases with an average age at death in *All Ages* versus those with *Normal Life Expectancy*, results suggest that the distributions all proxies for R&D for rare diseases with *Normal Life Expectancy* dominate the distributions of R&D for disease in the category *All Ages*.

Prevalence in the population - Results for the two tailored KS tests comparing the distribution of R&D investments for rare diseases over the four prevalence categories are presented in Table 8. They suggest that the distributions of most proxies of R&D targeting diseases in higher prevalence categories dominate the distributions of R&D investments of diseases in lower prevalence categories. When considering rare diseases with a prevalence <1,000,000 versus rare diseases in higher prevalence categories, all distributions of R&D investments favour diseases in higher prevalence categories (p-value=0.000 in all cases). When considering rare diseases with a prevalence 1-9 over 1,000,000 versus rare diseases in higher prevalence categories, the distributions of academic research and clinical trial activities favour diseases in higher prevalence categories. When considering rare diseases with a prevalence 1-9 over 1,000,000 versus

rare diseases in 1-9 over 10,000, the distributions of orphan designations favour diseases in 1-9 over 10,000. We cannot conclude on dominance when we compare the distributions of the number of research projects, orphan designations and orphan drugs (respectively research projects, and orphan designations) for rare diseases with a prevalence 1-9 over 1,000,000 versus 1-9 over 100,000 (respectively 1-9 over 10,000). When considering rare diseases with a prevalence 1-9 over 10,000 versus rare diseases with a prevalence of 1-9 over 100,000, the distributions of academic research, clinical trials, and orphan designations favour diseases in the higher prevalence category. The KS tests remain inconclusive for the number of research projects and of orphan drugs (p-value respectively 0.296 and 0.263).

Immediate danger of death - We now partition rare diseases between those with an immediate danger of death versus the other rare diseases by combining the average age at symptoms' onset and the mean age at death. We compare the distribution of the five proxies of R&D investments for those two groups of rare diseases. Results are presented in Table 9. They suggest that the distributions of R&D investments targeting diseases with an immediate danger of death are dominated by the distributions of R&D investments of diseases with non-immediate danger of death across most proxies of R&D investments, except for the distribution of the number of research projects and orphan drugs where the test is inconclusive (p-value respectively 0.886, 0.121).

Uncertainty on Disease Evolution - We now compare rare diseases according to whether there is uncertainty about their evolution. We consider that diseases for which both the average age at symptoms' onset and the average age at death are classified in "*All Ages*" category in the dataset are uncertain. The binary comparisons presented in Table 10 show that the distributions of two proxies of R&D investments (academic research, and orphan designations) of diseases with uncertainty on disease evolution are dominated by the distributions of the R&D investments of diseases with lower uncertainty (p-values respectively 0.006 and 0.001). The KS tests remain inconclusive for the distribution of the number of research projects, clinical trials, and orphan drugs.

b. Non-parametric tests on each diseases characteristic weighted by disease prevalence

We performed the same analysis accounting additionally for the prevalence category of the rare diseases using weights. Most of the results still hold in the weighted analysis.

Average age at symptoms appearance - Results displayed in Table 11 suggest that the distribution of most of R&D investments targeting diseases with a lower category of average age at symptoms onset (*Infancy* and *Childhood*) are dominated by the distributions of R&D investments for all other categories of average age at symptoms' onset (*Adults & Elderly* and *All ages*). The distribution of most of the R&D

proxies for rare diseases in *Infancy* are dominated by the distributions for rare diseases in *Adult & Elderly* and *All Ages*, except for the distribution of the number of research projects with *Adult & Elderly* where the KS test is inconclusive (p-value=0.191). When rare diseases in *Childhood* are compared with rare diseases in *All Ages*, the distribution of all R&D outcomes both favour rare diseases in *All ages*, however we cannot conclude on dominance when considering the number of research projects. When rare diseases in *Childhood* are compared with rare diseases in *Adult & Elderly*, the distribution of the number of academic publications and the distribution of the number of orphan designations are both in favour of diseases occurring in *Adult & Elderly* (p-value respectively 0.000 and 0.011). However we cannot conclude on dominance when considering the number of research projects, clinical trials and orphan drugs. When considering *Adult & Elderly* versus *All Ages,* we find that the distribution of three of the R&D outcomes (clinical trials, orphan designation and orphan drugs) over five for rare diseases in category *Adult & Elderly* are dominated by rare diseases in *All ages* (p-value respectively 0.001, 0.001, 0.044) . The KS tests remain inconclusive for the number of research projects and academic publications.

Average age at death - Results for the paired KS tests comparing the distribution of R&D investments for rare diseases over the five categories of the average age at death are presented in Table 12. They suggest that the distributions of R&D investments targeting diseases with an average age at death in *Infancy* are dominated by the distributions of R&D investments for higher categories of average age at death (*Adult & Elderly, All Ages, Normal Life Expectancy*). This result holds for the five R&D proxies, except for the distribution of the number of orphan drugs when considering the categories *All Ages,* and *Normal Life Expectancy* (p-value respectively 0.366, 0.184). When rare diseases in *Infancy* are compared to rare diseases in *Childhood*, the distributions of the number of research projects and clinical trials are in favour of diseases occurring in *Childhood* (p-values respectively 0.000 and 0,000). However, the distribution of the number of academic research, orphan designations and orphan drugs are in favour of diseases with mean age at death in *Infancy* (p-values respectively 0.000; 0,000 and 0.032). The distribution of most of the R&D proxies for rare diseases with mean age at death in *Childhood* are dominated by the distributions for rare diseases with mean age at death in *Adult & Elderly*, *All Ages* and *Normal Life Expectancy*.
When considering rare diseases with an average age at death in *Adults & Elderly* versus those in *All Ages*, results suggest that the distributions of all proxies of R&D for the category *Adults & Elderly* dominate the distributions of R&D for disease with an unpredictable mean age of death. When considering rare diseases with an average age at death in *Adults & Elderly* versus those with *Normal Life Expectancy*, results suggest that the distributions of all proxies of R&D for the category *Adults & Elderly* dominate the distributions of R&D for disease with an average age at death in *Adults & Elderly*.

Immediate danger of death - When combined with disease prevalence, results suggest that the distribution of all R&D investments targeting diseases with an immediate danger of death are dominated by the distributions of R&D investments of diseases without immediate danger of death. Results are displayed in Table 13.

Uncertainty on Disease Evolution - Results in Table 14 compare rare diseases according to whether there is uncertainty about their evolution. The results differ from the one computed in the absence of weights. More specifically, they suggest that the distributions of R&D investments targeting diseases with lower uncertainty are dominated by the distributions of R&D investments of diseases with higher uncertainty, when considering the following proxies: research projects, clinical trials, orphan designations (p-values respectively 0.000, 0.041, and 0.007). The KS tests are inconclusive for all the number of academic publications and orphan (p-values respectively 0.971 and 0.396)

VI. Discussion

We investigated the distribution of R&D investments across rare diseases as measured by the number of research projects, academic publications, clinical trials, orphan designations and orphan drugs with marketing authorization. When comparing the distribution of these five proxies of R&D investments across rare diseases with different average age at symptoms' appearance, it appeared than the life stages at which the disease occurs is associated with different levels of R&D investments. Results suggest that diseases with symptoms appearing during *Infancy* and *Childhood* are dominated in terms R&D investments by rare diseases with symptoms appearing among *Adult & Elderly*. When considering the average age at death of rare diseases, the same age groups of *Adult & Elderly* is favoured. Results suggest that diseases with an average age at death in *Infancy*, and in *Childhood* are dominated in terms R&D investments by diseases with an older average age at death. This result is robust to the inclusion of frequency weights accounting for the prevalence levels in our sample. While it is known that rare diseases are generally underserved by drug development in comparison with other diseases, our study shows that within rare diseases there are sub-groups of rare diseases that are worse-off regarding R&D. Rare diseases that affect younger patients are the most deprived in terms of drug development among rare diseases. Epidemiology studies conducted on rare diseases show that up to 75% of rare diseases are paediatrics (Bavisetty et al., 2013). One reason may be that developing therapies for children is more challenging. Children are a very heterogeneous group with different physiological, developmental, psychological and pharmacological characteristics (Joseph et al., 2015). The consideration of growth and puberty is also crucial issue, and therapies must embody the impact they may have on the reproductive system (Lathyris

et al., 2014). The metabolization of drugs is heterogeneous across age groups within childhood and it makes it difficult to evaluate the optimal dosage for the therapy whilst it is necessary to prevent toxicity. Overall, the development of therapies for children is more costly and not attractive to pharmaceutical firms. Furthermore R&D in therapies for children raise important ethical concerns as parents must provide consent in place of their child and may be reluctant to expose their child to the likelihood of adverse effects and newly developed treatments (Joseph et al., 2015).

Our results also confirms that market share is a driver of R&D activities, which is in line with previous evidence (Dubois, 2015) as rare diseases in high prevalence categories are favoured by R&D investments. As drug development entails large fixed costs that are decreasing with market size since recruitment in clinical trials is far more costly for ultra-rare diseases, a larger market size gives the opportunity to pharmaceutical firms to recover their fixed costs.

We also compared the distribution of R&D activities when rare diseases are associated with an immediate danger of death after the first symptoms, and when rare diseases show a high level of uncertainty in terms of rate of progression or disease presentation. Our results suggest that rare diseases with an immediate danger of death and rare diseases that embody a high level of uncertainty are more deprived by drug development than other rare diseases. In the analysis with frequency weights based on prevalence levels, diseases with high level of uncertainty are favoured, but the risk of death surrounding rare diseases still do not foster further R&D investments.

This study presents limitations, especially regarding the dataset we used. All the disease characteristics were not available for all the rare diseases in the sample. This limited number of data availability prevented us from aggregating rare diseases characteristics in the analysis. It would have been interesting to combine these disease characteristics to generate a "type" according to Roemer (1993). We faced dramatic reductions in sample size due to missing data when building a complete balanced data. Still, we studied the missing data patterns and found that the difference in the mean number of R&D resources of missing values compared to the non-missing values is negative and quite low. Another limitation important to underline is that R&D investments are likely to increase the availability of some disease characteristics and vice versa if some disease characteristics are available R&D is likely to be stimulated.

We summarised the average value for each of the proxies of R&D investments in Figure 4. The hierarchy in disease characteristics is rather stable across the proxies of R&D investments. The most deprived category over all R&D investments is the group of rare diseases with an average age at first symptoms

during *Infancy* and *Childhood*. The second most deprived characteristic is uncertainty about rare diseases evolution then comes the group of diseases with an immediate danger of deaths. While the difference in average R&D investments is very low, it is somewhat dependent on disease characteristics. This points out the existence of inequalities of opportunity in R&D across rare diseases that are not currently addressed at the European level. The health promotion of the most deprived sub-groups of rare disease could be a desirable form of compensation to prevent long-term discrepancies in health technologies availability and ultimately discrepancies in patients' opportunities to access care.

VII. Acknowledgments

Authors are grateful to Clémence Thébaut, Aki Tsuchiya and the EuHEA 2018 participants for their useful comments and suggestions. We also received detailed and useful reports from two reviewers and the editors of this special issue that helped us substantially improve the paper. Any remaining errors are the responsibility of the authors.

VIII. References

Arneson, R. J. (1989). Rawlsian Theory of Justice: Recent Developments. *Ethics*, *99*(4), 695–710.

Barry, B. (2005). Why social justice matters. Polity. ISBN: 978-0-745-62993-3

Bavisetty, S., Grody, W. W., & Yazdani, S. (2013). Emergence of pediatric rare diseases. *Rare Diseases*, *1*

Bentham, J., & Mill, J. S. (2004). *Utilitarianism and Other Essays*, Penguin UK.

Bourguignon, F., Ferreira, F. H. G., & Menéndez, M. (2007). Inequality of opportunity in Brazil. *Review of Income and Wealth*, *53*(4), 585–618.

Brazier J, Rowen D, Mukuria C, Whyte S, Keetharuth A, Rose Hole A, Tsuchiya, Shackley P (2013). Eliciting societal preferences for burden of illness, therapeutic improvement and end of life for value based pricing: a report of the main survey Available from www.eepru.org.uk

Buchanan, A. (2000). Rawls's law of peoples: Rules for a vanished Westphalian world. Ethics, 110(4), 697-721.

Cohen, J. (1989). Democratic Equality. *Ethics*, *99*(4), 727–751.

Cookson, R., & Dolan, P. (2000). Principles of justice in health care rationing. Journal of medical Ethics, 26(5), 323-329.

Daniels, N. (1985). *Just health care*. Cambridge University Press. ISBN: 9780511624971

Drummond, M., & Towse, A. (2014). Orphan Drugs Policies: A Suitable Case for Treatment: Editorial. *European Journal of Health Economics*, *15*(4), 335–340.

Dubois, P. (2015). Market size and pharmaceutical innovation. *The RAND Journal of Economics*.

Dworkin, R. (1981a). What is Equality? Part 1: Equality of Welfare. *Philosophy and Public Affairs*, *10*(3), 185–246.

Dworkin, R. (1981b). What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs*. Retrieved from https://philpapers.org/rec/DWOWIE-2

Ferreira, F. H. G., & Peragine, V. (2015). *Equality of Opportunity: Theory and Evidence*. The World Bank.

Field, M. J., & Boat, T. F. (2010). *Profile of Rare Diseases*. National Academies Press (US). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK56184/

Fleurbaey, M. (2008). *Fairness, Responsibility, and Welfare*. Oxford University Press.

Fleurbaey, M., & Schokkaert, E. (2009). Unfair inequalities in health and health care. *Journal of Health Economics*, *28*(1), 73–90.

García-Gómez, P., Schokkaert, E., Ourti, T. V., & d'Uva, T. B. (2014). Inequity in the Face of Death. *Health Economics*, *24*(10), 1348–1367.

Gericke, C. A., Riesberg, A., & Busse, R. (2005). Ethical issues in funding orphan drug research and development. *Journal of Medical Ethics*, *31*(3), 164–168.

Giannuzzi, V., Conte, R., Landi, A., Ottomano, S. A., Bonifazi, D., Baiardi, P., … Ceci, A. (2017). Orphan medicinal products in Europe and United States to cover needs of patients with rare diseases: an increased common effort is to be foreseen. *Orphanet Journal of Rare Diseases*, *12*(1).

Hughes, D. A., Tunnage, B., & Yeo, S. T. (2005). Drugs for exceptionally rare diseases: do they deserve special status for funding? *QJM: An International Journal of Medicine*, *98*(11), 829–836.

Joseph, P. D., Craig, J. C., & Caldwell, P. H. Y. (2015). Clinical trials in children. *British Journal of Clinical Pharmacology*, *79*(3), 357–369.

Jusot, F., Tubeuf, S., & Trannoy, A. (2013). Circumstances and efforts: how important is their correlation for the measurement of inequality of opportunity in health?. Health economics, 22(12), 1470-1495.

Lathyris, D., Panagiotou, O. A., Baltogianni, M., Ioannidis, J. P. A., & Contopoulos-Ioannidis, D. G. (2014). Safety of Medical Interventions in Children Versus Adults. *Pediatrics*, *133*(3), e666–e673.

Lefranc, A., Pistolesi, N., & Trannoy, A. (2009). Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France. *Journal of Public Economics*, *93*(11–12), 1189–1207.

Lefranc, A., & Trannoy, A. (2016). Equality of Opportunity: How to encompass Fifty Shades of Luck, 36, ECINEQ Working Paper Series.

Martin, D. K., Giacomini, M., & Singer, P. A. (2002). Fairness, accountability for reasonableness, and the views of priority setting decision-makers. *Health Policy*, *61*(3), 279–290.

Lichtenberg, F. R. (2001). The effect of new drugs on mortality from rare diseases and HIV (No. w8677). National Bureau of Economic Research.

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

Lichtenberg, F. R. (2014b). The impact of pharmaceutical innovation on longevity and medical expenditure in France, 2000–2009. *Economics & Human Biology*, *13*, 107–127.

Lichtenberg, F. R. (2017). *The Impact of Public and Private Research Support on Premature Cancer*

Mortality and Hospitalization in the U.S., 1999-2013 (Working Paper No. 23241). National Bureau of Economic Research.

McCabe, C., Claxton, K., & Tsuchiya, A. (2005). Orphan drugs and the NHS: should we value rarity?. BMJ: British Medical Journal, 331(7523), 1016.

McKie, J., & Richardson, J. (2003). The Rule of Rescue. *Social Science & Medicine*, *56*(12), 2407–2419.

Nord, E., Richardson, J., Street, A., Kuhse, H., & Singer, P. (1995). Maximizing health benefits vs egalitarianism: An Australian survey of health issues. *Social Science & Medicine*, *41*(10), 1429–1437.

Otsuka M. (2013). Prioritarianism and the Measure of Utility. *Journal of Political Philosophy*, *23*(1), 1–22.

Paulden, M., Stafinski, T., Menon, D., & McCabe, C. (2015). Value-based reimbursement decisions for orphan drugs: a scoping review and decision framework. Pharmacoeconomics, 33(3), 255-269.

Piraino, P. (2015). Intergenerational Earnings Mobility and Equality of Opportunity in South Africa. World Development, 67, 396–405.

Ramos, X., & Van De Gaer, D. (2015). Approaches to Inequality of Opportunity: Principles, Measures and Evidence. *Journal of Economic Surveys*, *30*(5), 855–883.

Roemer, J. E. (1998). Theories of distributive justice. Harvard University Press.

Roemer, J. E., & Trannoy, A. (2015). Chapter 4 - Equality of Opportunity. In A. B. Atkinson & F. Bourguignon (Eds.), *Handbook of Income Distribution* (Vol. 2, pp. 217–300). Elsevier.

Schieppati, A., Henter, J. I., Daina, E., & Aperia, A. (2008). Why rare diseases are an important medical and social issue. The Lancet, 371(9629), 2039-2041.

Sheskin, D. J. (2003). Handbook of parametric and nonparametric statistical procedures. crc Press.

Scott Morton, F., & Kyle, M. (2011). Chapter Twelve - Markets for Pharmaceutical Products. In M. V. Pauly, T. G. Mcguire, & P. P. Barros (Eds.), *Handbook of Health Economics* (Vol. 2, pp. 763–823). Elsevier.

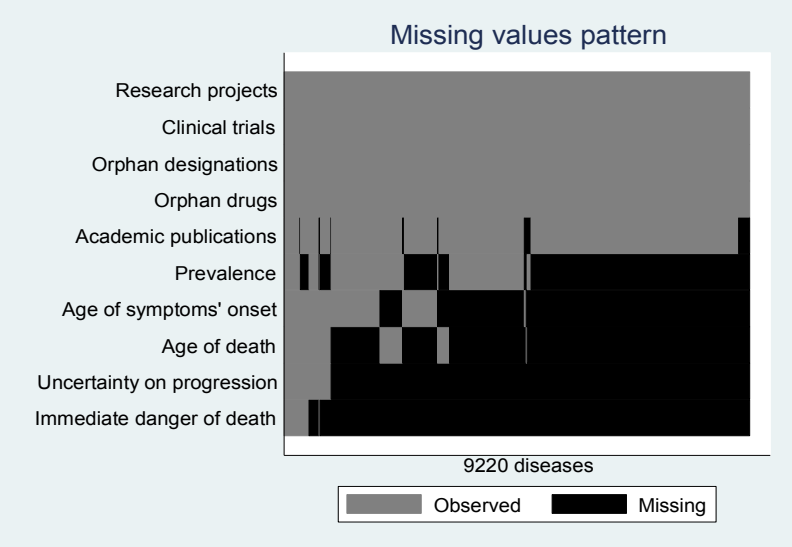
Temkin, L. S. (2003). Equality, priority or what?. Economics & Philosophy, 19(1), 61-87.

Thébaut, C., & Wittwer, J. (2017). L'évaluation économique en santé au prisme de l'économie normative : principes allocatifs et règles de priorisation, Taking redistributive principles into account in the economic evaluation of health care: a review of available methods. *Revue française des affaires sociales*, (3), 169–191.

Williams, A., & Cookson, R. (2000). Equity in health. Handbook of health economics, 1, 1863-1910.

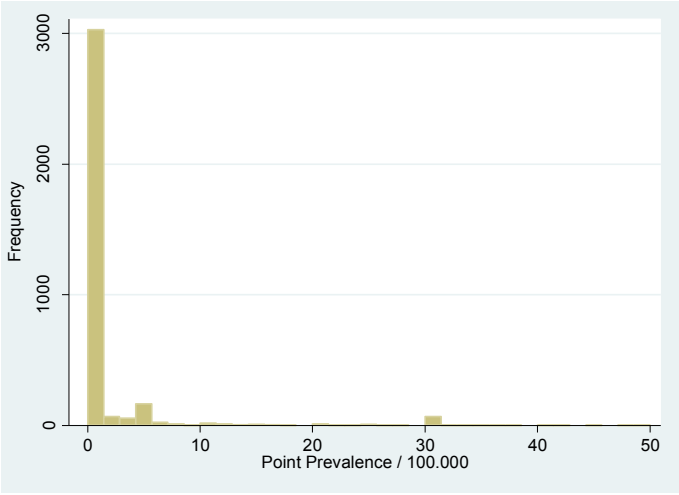
IX. Figures

Figure 1 – Missing values pattern in terms of all variables of interest



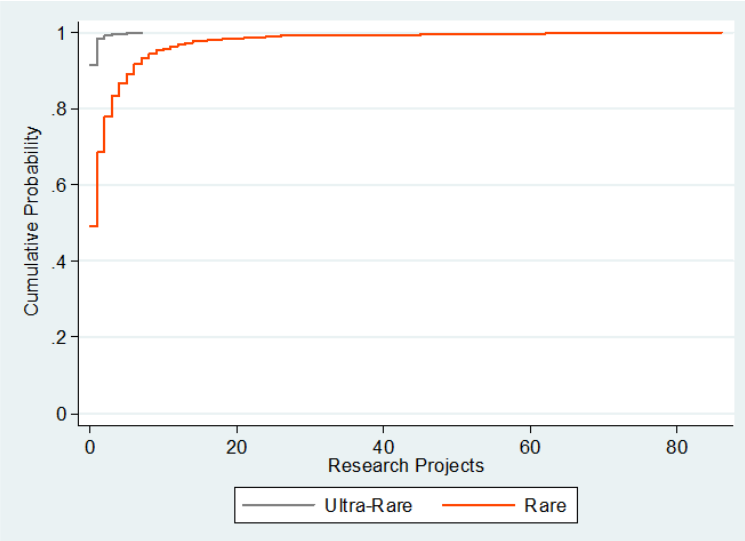
Note: This graph provides visual investigation of shared missing values between all variables considered in the analysis.

Figure 2: Frequency distribution of rare disease prevalence



Source: Orphanet Dataset

Figure 3: cumulative distribution of research projects of ultra-rare versus non ultra-rare diseases



Source: Orphanet Dataset