# Is research responding to health needs?

**Ismael Ràfols**, research fellow at Ingenio (CSIC-UPV), Universitat Politècnica de València and visiting professor at the Centre for the Study of Science and Technology (CWTS), University of Leiden, The Netherlands (i.rafols@ingenio.upv.es)

**Alfredo Yegros**, research fellow at the Centre for the Study of Science and Technology (CWTS), University of Leiden, The Netherlands

14th September 2017

This is the original version of a short article published in the Social Observatory of the La Caixa Foundation, re-written in a form more accessible for a general readership.
Search in: https://observatoriosociallacaixa.org/en

**Highlights:**

- Health benefits from research could be significantly improved through a more systemic planning of research priorities.
- Some diseases, such as malaria or tuberculosis, that are prevalent in low income countries, deserve more research investments – and this is also the case for conditions such as depression or stroke that have a high disease burden across the world.
- The question is not only whether research is tackling the most pressing health needs, but also whether the strategies pursued are appropriate to address these needs.
- To be more responsive to social needs, priority setting should be enriched not only with information about health needs, but also with dialogue with patients and other stakeholders of health R&D.

**Abstract**

Comparisons between R&D efforts and WHO estimates of health burden for a given disease suggest substantial misalignments, both at the global and local levels. Some diseases prevalent in low and middle income countries such as malaria or tuberculosis may deserve more R&D investment given their burden. This is also the case for health conditions such as depression or stroke that have a high disease burden in both high and low income countries, yet attract relatively little research. This type of comparative approach between health needs and R&D efforts can be useful when setting priorities in R&D investment. We propose that an appraisal process for health R&D should use more healthcare data to better understand health needs, and should conduct citizen engagement so as to take into account their views on the problems to be addressed and the strategies to tackle them.

**Introduction**

People's health has greatly increased in the last century. This is largely due to improvements in living conditions, public health (e.g. access to clean water) and health care, and also (though smaller) part due to scientific discoveries. X-rays and penicillin are examples of serendipitous breakthroughs that led to great advances in medicine. Vaccines are examples of applied research that saved millions of lives. Such extraordinary successes have led to the belief that more biomedical research will eventually lead to better health. In this article, we will challenge this belief – instead, we will argue that for biomedical research to improve health it has to pay more attention to people's needs and to the




Electronic copy available at: https://ssrn.com/abstract=3106713

causes of ill-health – for example wider environmental and socio-economic conditions such as pollution or inequality (Sarewitz and Pielke, 2007).

The world spent an estimate of 240 billion US$ (purchasing power parity) in 2010 in health R&D, 90% in high income countries. As shown in Figure 1, health R&D expenditure has grown to become between 0.2% and 1.0% of GDP in high income countries (0.26% in Spain) (Røttingen et al., 2013). 60% of these R&D investments were made by the private sector, 30% by the public sector and 10% from other sources such as non-profit organisations. While acknowledging the past and potential benefits of health research, various analysts have questioned in recent years whether research efforts are currently wisely spent (Chalmers et al., 2014).

In the face of this questioning, most attention has gone to "the problem of translation", i.e. whether knowledge from fundamental science is quickly and efficiently used in applications in clinical and health care. For this purpose, many 'translational' centres and labs have been created with the aim of combining basic and clinical research – thus quickly transferring new knowledge 'from bench to bedside'.

However, taking a step back and looking from a wider perspective, there are deeper problems with current biomedical research. An alarming proportion of published research claims have been found to be false due to insufficient rigor in research methodologies (Ioannidis, 2005), while there are instances of publications unnecessarily replicating already known findings (Chalmers and Glasziou, 2010). Current evaluation systems are mainly focused on scientific visibility (e.g. publication in 'top' scientific journals), which may be leading research away from pursuing knowledge that is valuable for society (Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015). And perhaps the most serious and widely discussed issue is that private R&D (60% of total health R&D) follows market demands – which lies in diseases in rich countries, in particular in chronic conditions, while diseases or findings of greater public health relevance remain underinvested (Aaron and Siegel, 2017; Schillinger et al., 2016; Sismondo and Chloubova, 2016).

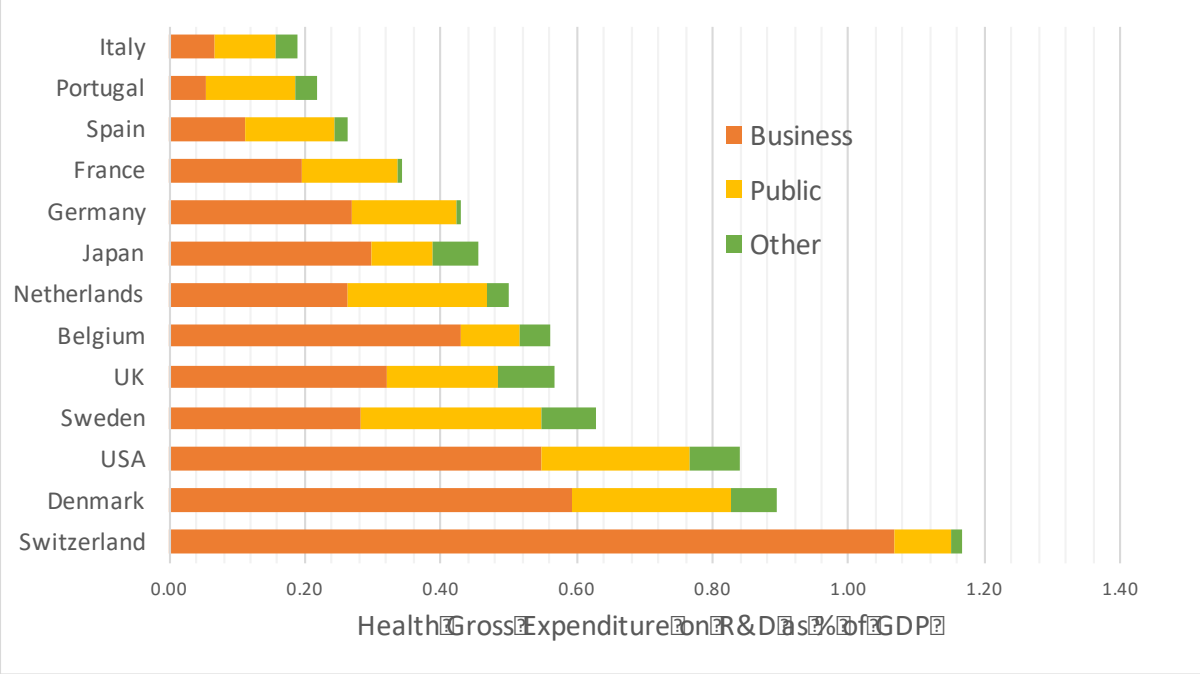
In the wake of these troubles, analysts argue that health benefits from research could be significantly improved through a more systemic planning of research priorities (Sarewitz and Pielke, 2007; Chalmers et al, 2013). Research funding is allocated according to diverse motivations (Viergever et al., 2010), generally with criteria that include perceptions on the scientific quality of projects and teams, the potential for scientific advancement in the topic, and the societal demands or needs for a given issue. In practice, while a lot of emphasis has been placed into assessing scientific quality so as to foster 'excellence', far less attention has been paid to assess if research addresses social needs. In health, various studies suggest there is lack of alignment between research priorities and health needs. Such misalignment is notorious at the national level (Gross et al., 1999; Gillum, 2011) and acute in global health research (Røttingen et al., 2013; Agarwal and Searls, 2009; Evans et al., 2014).

In this article, we will explore how priority setting in health research can be improved by asking whether research contents are aligned with societal needs. To do so, we will compare estimates of health needs with proxies of research efforts. The significant misalignments we will observe support the vision that health research can greatly benefit from making use of digital information ('big data') from healthcare, from deeper engagement with diverse social actors – a shift that is in accordance with ongoing policy discourses on 'open science' and Responsible Research and Innovation (RRI).

**Figure 1. Health Gross Expenditure on R&D by estimated funding source.** Data Source: Røttingen et al. (2013, p.16).




Electronic copy available at: https://ssrn.com/abstract=3106713




## Methods for assessing research priorities against health needs

The ultimate goal of health research is to improve the health and wellbeing of humankind and therefore, and thus it should be focused on this goal. Policy documents state that health research should be oriented 'towards humanitarian goals and solidarity', and should 'underpin every step in the innovation cycle from discovery in bench and bedside research, to implementation in healthcare, and prevention.' (EC, 2016). But how can we know if such goal is indeed fulfilled in the long term?

To try to answer this question, in 1997 the US National Institutes of Health (NIH) created a working group on priority setting. The panel observed that social health needs was one of the criteria for research allocation. However, it concluded that the NIH did not sufficiently explain how health needs were assessed and recommended to improve its health data analysis. Two years later Gross et al. (1999) published a study relating various measures of disease burden to NIH research efforts for a list of major health conditions. Research efforts were estimated in terms of funding allocated to projects related to health conditions. There are various measures of disease burden such as prevalence (number of people with a disease), incidence (frequency of new cases), hospital days, mortality, years of life lost and disability-adjusted life years (DALYs). DALYs is a measure that estimates the loss of one year of healthy life to disease, taking into consideration mortality as well as disability caused by a health condition. This means that DALY estimates sum the years not lived until life expectancy age due to deaths produced by a condition (*Years of Life Lost*) plus the time not fully enjoyed because of a health condition (*Years Lost due to a Disability*).

The study by Gross and colleagues (1999) found that the funding for a health condition was weakly correlated with mortality and years of life lost, but strongly correlated to DALYs. This meant that NIH funding was sensitive to disease burden. However, this careful analysis also showed substantial disparities in the amount of funding for the same burden across diseases. Some diseases such as AIDS, breast cancer or diabetes mellitus received much more funding per DALY than depression, colon cancer or perinatal conditions. There was thus room for improving alignment between funding and disease burden.

Subsequently, various studies at the national scale have followed this methodology to assess research funding against disease burden, for example in the US (Gillum et al., 2011), Norway (Kinge et al., 2014) and Spain (Català-López et al., 2009). Another approach has been to use the number of publications as a proxy of research effort (Agarwal and Searls, 2009; Evans et al., 2014). The bibliometric approach has the advantage of capturing research conducted in universities and hospitals by lecturers and physicians, but which is not accounted in the R&D expenditures of funding agencies.

Here, we rely on publications to estimate research efforts and DALYs as a measure of disease burden. We combined data from the databases PubMed and Web of Science to characterise publications for 2009-2013. Publications were assigned to specific disease by PubMed descriptors (known as Medical Subject Headings or MeSH) provided by experts at the US NIH National Library of Medicine. We obtained the global burden of disease estimates from a dedicated WHO website, using 2012 data.[1]

This methodology has serious limitations and results have to be interpreted with caution. First, publications are inaccurate estimate of research efforts. They do not capture most of private R&D efforts (which are confidential until patented). Not all fields publish with the same frequency due to different disciplinary traditions and incentives and some research topic become temporarily fashionable or unfashionable for reasons that are not related to either science or health.

Second, there is also a significant time lag (perhaps 2 to 5 years) between knowledge production and publications, and a much longer lag (5 to 20 years) between research findings and their use or application in societal contexts. Therefore, there is uncertainty regarding the appropriate time window for comparison. Since disease burden is also changing over time, future studies should conduct a dynamic analysis of DALYs and publications to reflect possibilities for their relative convergence or divergence.

### Lack of alignment between disease burden and research investments

The comparison between the percentage of DALYs and the percentage of publications for a given disease group allows to assess the degree of alignment between disease burden and research investments, as shown in Figure 2. We provide data for the whole world and Spain (which has a profile similar to many developed countries). Disease burden and publication patterns show striking disparities in research across disease groups. Al the global level, cancer (Malignant neoplasms) makes more than 22% of global disease publications, though cancer's disease burden is less than 10% of the total. Cardiovascular, infectious and parasitic diseases have more than 16% of burden but less than 10% of publications. Neonatal conditions are an issue with very little research in relation to the its burden. In Spain and other countries in the North, disease burdens are strikingly different to the world's. For example, cancer's share of burden is a bit higher than its share of publications, while the share of publications in infectious and parasitic diseases is much higher than their burden in Spain.

[1] http://www.who.int/healthinfo/global_burden_disease/estimates_country_2000_2012/en/

**Figure 2. Percentage of disease burden in Disability-Adjusted Life Years (DALYs) vs. Percentage of Publications for disease groups in the world (top) and Spain (bottom).** Only selected disease groups are shown. Sources: WHO Global Burden of Disease (2012). Web of Science (2009-2013).

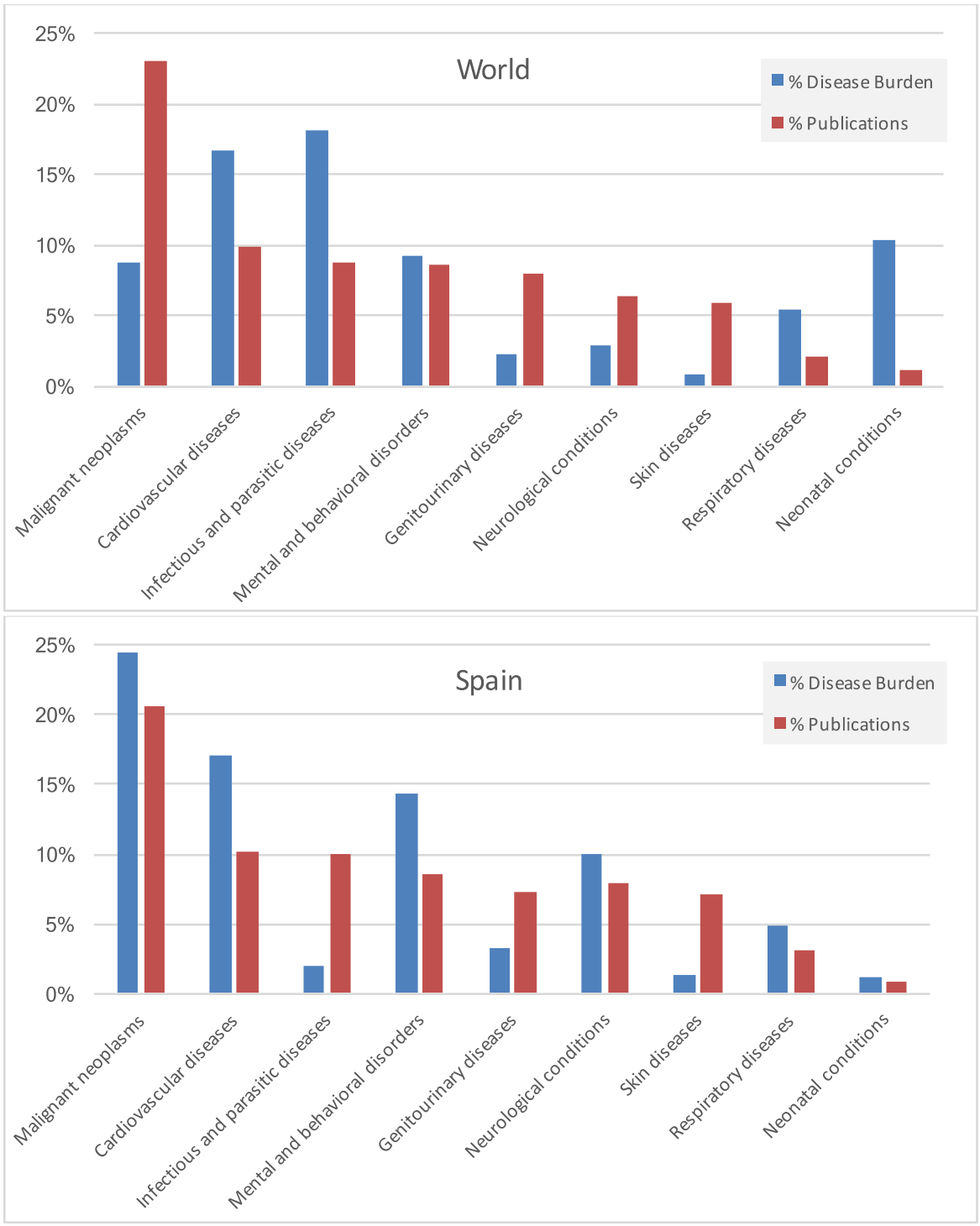


Figure 3 makes the same comparison as Figure 2, but now for specific diseases rather than disease groups, i.e. rather than cancer as a whole, we can see the break down for specific cancers. In Figure 3, the horizontal axis shows the relative disease burden (as a percentage of the total burden) and the vertical axis the number of publications (also relative to the total number of publication on diseases). The diseases that appear above the 45 degrees slope have relatively more publications than their disease burden (overstudied), whereas the disease below the 45 degrees slope have proportionally less publications (understudied).

This shows that while cancer like melanoma and breast cancer are relatively over-studied, lung cancer or colon cancer are relatively under-studied in relation to their disease burden. Similarly, within cardiovascular problems such as stroke or ischaemic heart disease (also known as coronary artery disease) have relatively few publications, but hypertensive heart disease has more. Hence, it is worth carrying out fine-grained analysis. Fine-grained perspectives (actually more detailed than shown here) are important because many diseases with small frequency of new cases (incidence) are often under-studied in relation to the burden they cause. This is particularly the case in developing countries – in these context small amounts of funding amounts might result in substantial improvements (Von Philipsborn et al., 2015).

The comparison between the patterns for the world and Spain reveals that while some diseases have similar or higher burdens in developed countries (e.g. cardiovascular or cancers), others have a much higher burden in low and middle income countries. However, these diseases are still present in high income countries (e.g. AIDS/HIV or tuberculosis), whereas some other diseases are only a major burden in countries with poor populations (e.g. malaria). Given these differences, diseases have been classified into type I (no more than 3 times higher in low and middle-income countries than in high-income countries), type II (3–35-times higher in low and middle income countries) and type III (more than 35 times higher in low and middle income countries) (Røttingen et al., 2013). Type II and III disease are shown as red triangles in Figure 3.

Figure 4 shows that most of the world health research for diseases is published in high-income countries mainly located in Europe, North America and East Asia. On the contrary, Africa, the Indian subcontinent (South Asia) and South East Asia have a small percentage of publications in relation to their populations. In high-income countries, there is an imperfect, but significant correlation between national disease burden and research efforts (as can be seen in the case of Spain in Figure 3). Since 90% of world's research is produced in developed countries, it follows that global research efforts are more aligned with the health needs of affluent societies than with the world as a whole (Agarwal and Searls, 2009; Evans et al., 2014). One can see this pattern by looking at Type II and III diseases (shown as red triangles) in Figure 3: in the world, they appear as the diseases with high disease burden and little research, whereas in Spain they appear mainly as diseases with little burden but generally more research than it is locally needed.

In an increasingly interconnected world, there are good reasons for high income countries to contribute to global health research. In the first place, equity in health is widely acknowledged as a shared value across the world – and governments can contribute to it by supporting research in global health. Secondly, as a result of globalisation and increased movement of people across continents, the 'globalisation' of lifestyles (e.g. with *junk food*) and infectious diseases, have effects in health across the globe. Under these circumstances, it is in everybody's interest to improve health conditions at a global scale. Moreover, on has to take into account that private health research (which is about 60% of global health funding) is even more narrowly focused on chronic diseases of the rich. In order to partly address unmet health needs in low-income countries, governments were recommended to spend 0.05%-0.10% or 0.15%-0.20% of GDP on health R&D, and at least 0.01% on research to meet the unmet health needs of low income countries (Røttingen et al., 2013). A few countries meet the

first criterion of total health R&D, as shown in Figure 1, but in 2010, none met the criterion of funding 0.01% of their GDP in research for neglected diseases (to cite some of the higher spenders: the US spent 0.0096%, the UK, 0.0073%, Sweden 0.0041%, in contrast only Spain, 0.0010%.

Therefore, there is a case for shifting public health research to supporting both unmet local needs and unmet global needs. As illustrated in Table 1, one may thus question the reasons for under-investment in those diseases, such as depression or chronic obstructive pulmonary disease that have relative under-investment in local and global contexts. One can also make the case for considering the resources put into research of local health priorities (such as diabetes) against those spent in global health challenges (such as malaria).

**Table 1. Comparison of research efforts against disease burden in the world and Spain.**

| | | **In Spain** | |
|---|---|---|---|
| | | **When Research effort is <u>smaller</u> than Disease burden** | **When Research effort is <u>bigger</u> than Disease burden** |
| **In the World** | **When Research effort is <u>smaller</u> than Disease burden** | Under-investment in both the world and Spain<br>*Ischaemic heart disease, depression, stroke, chronic obstructive pulmonary disease* | Global health priority<br>*HIV/AIDS, tuberculosis, malaria, lower respiratory infections (mainly type II and III diseases)* |
| | **When Research effort is <u>bigger</u> than Disease Burden** | Health priority in Spain<br>*Diabetes mellitus, lung cancer, colon cancer, Alzheimer's* | Over-investment in both the world and Spain<br>*Skin diseases, melanoma, breast cancer, hypertensive heart disease* |

**Figure 3. Percentage of Disease burden in Disability-Adjusted Years (DALYs) vs. Percentage of Publications for specific diseases in the world (top) and Spain (bottom).** Only selected diseases are shown. Blue dots show Type I disease (present across the world). Red triangles show Type II and III diseases (more prevalent in low and middle income countries). Sources: WHO Global Burden of Disease (2012). Web of Science (2009-2013).

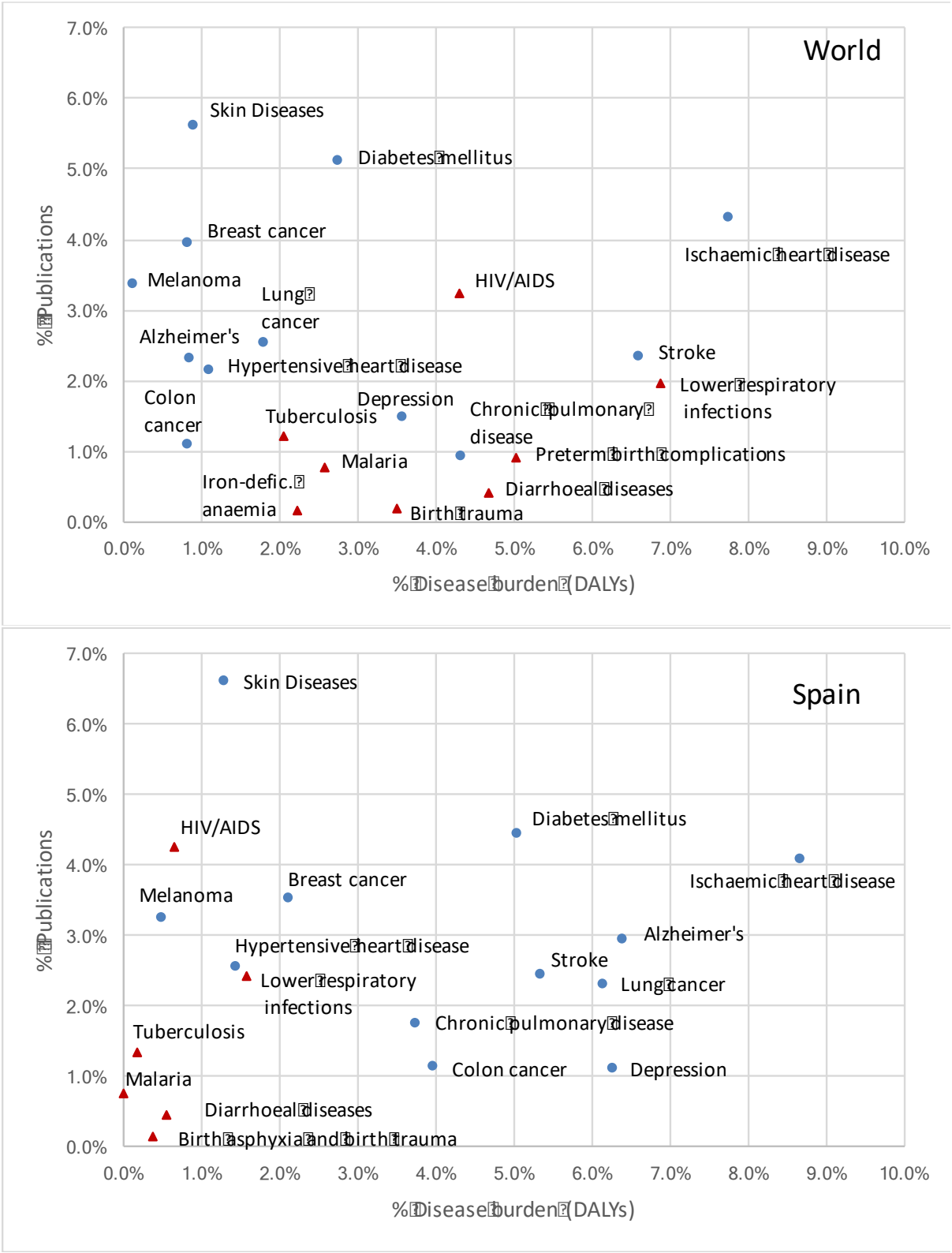
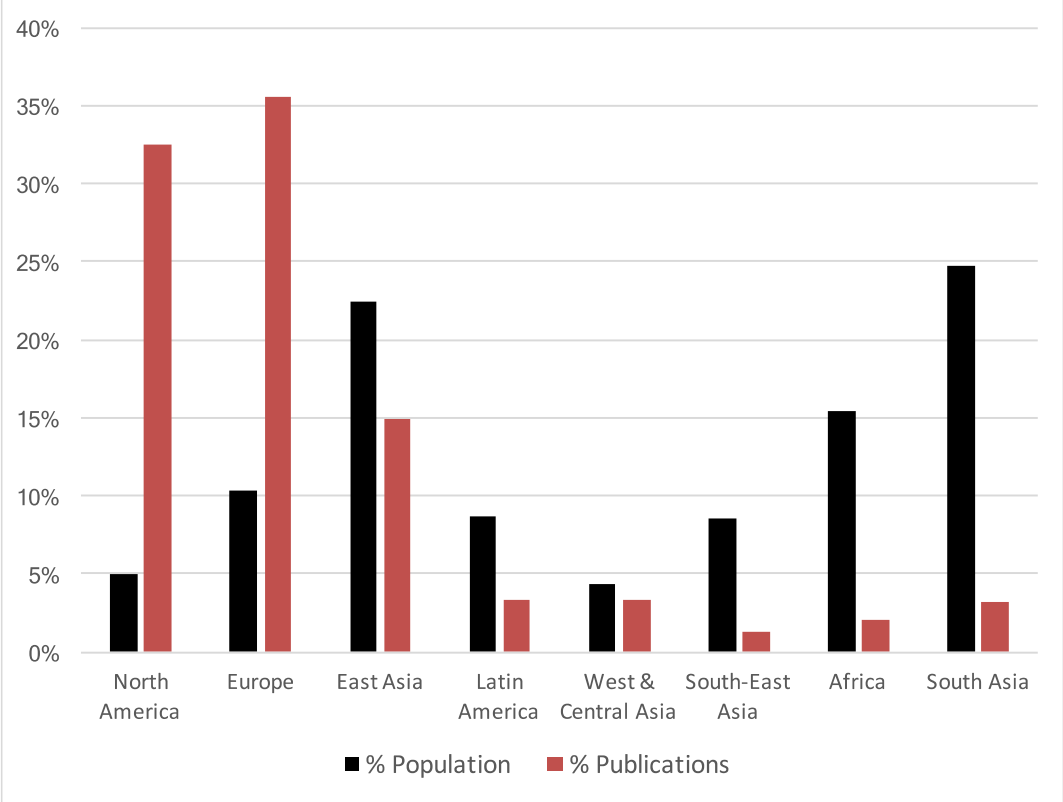


**Figure 4. Proportion across continents of population vs. proportion of disease publications.** Sources: WHO Global Burden of Disease (2012). Web of Science (2009-2013). Continents are defined following UNESCO classification. Oceania is not shown given its small size.




**Looking at the bigger picture: towards use big data and public engagement**

While the comparisons between global disease burden and number of publications may be useful, they are just one example of the type of information to consider for re-thinking research priorities. One the one hand, one can think of other data sources for comparing health needs and research efforts. For example, it is possible to collect data on funding expenditure to estimate research efforts, or data on drug expenditure in a given disease as a complementary estimate of health needs.[2]

On the other hand, R&D priority setting also has to take into account broader contextual considerations, such availability of previous therapies and expert perceptions regarding the potential for advancement in fighting a disease. For example, some may argue that there are already drugs to lower the high disease burden of infectious diseases in Africa – and hence the problem is not lack of research but inadequate healthcare services and unavailability of essential drugs due to high pricing. At the other extreme, some experts may argue that for diseases such as Alzheimer disease, one should not invest to much in research because the

[2] Interestingly, various governments are increasingly making available this type of digital records ('big data') on both funding and healthcare provision under initiatives for open government to foster transparency and accountability – thus facilitating the construction of fine-grained and diverse statistics on health needs and health R&D.

field is not yet mature enough to yield medical solutions. Indeed, it has been noted that some diseases are more difficult to address than others using approaches such as vaccines (Yaqub, 2017) and that other type of sociotechnical interventions can be more successful than conventional therapeutic approaches. For example, insecticide-treated bed nets have proven very useful in malaria.

In summary, the question is not only whether public R&D is tackling the most demanding health needs, but also whether the strategies pursued are appropriate to address these needs. Since research is highly uncertain, and various R&D strategies may contribute in different ways to improve a health condition, it is advisable to conduct various research lines in parallel. This is illustrated in Figure 5, with the case of research on obesity. This map portrays the epistemic landscape of obesity research, i.e. it shows the topics related to obesity: research on human biology and metabolism in the right-hand side; on the diseases caused by obesity at the top; on the treatments of obesity at the right hand-side, including treatments related to lifestyles (diet and exercise) towards the centre; and on social determinants of obesity at the bottom left (Cassi et al., 2017).

The size of the nodes in Figure 5 shows the proportion of publications in a topic. This proportion is a proxy of the amount of research resources spent in a given topic. Is this a wise distribution of resources? Should science policy foster more research on certain topics (e.g. diet) than on others (e.g. bariatric surgery) given their relative public health benefits? In the case of obesity, many stakeholders agree that more research is needed on the social and psychological factors related to food industry, consumption patterns and sedentary lifestyles (bottom left in Figure 5), rather than on understanding the biology or improving therapies such as surgery.

In a case such as obesity, patients and other relevant citizens have a valuable experiential knowledge and informed preferences on the type of approaches that might more important. Given that many other increasingly prevalent health conditions (such as depression or respiratory problems) are associated with environmental and social factors (such as pollution, urban lifestyles and inequality), engagement is particularly important to take into the values that citizens accord to health conditions and determinants. Data on healthcare such as global disease burden offer new opportunities in spotting the problems that deserve attention, but citizen engagement is also necessary so as to pay attention to the type of solutions that the public prefers (Stirling and Scoones, 2009).

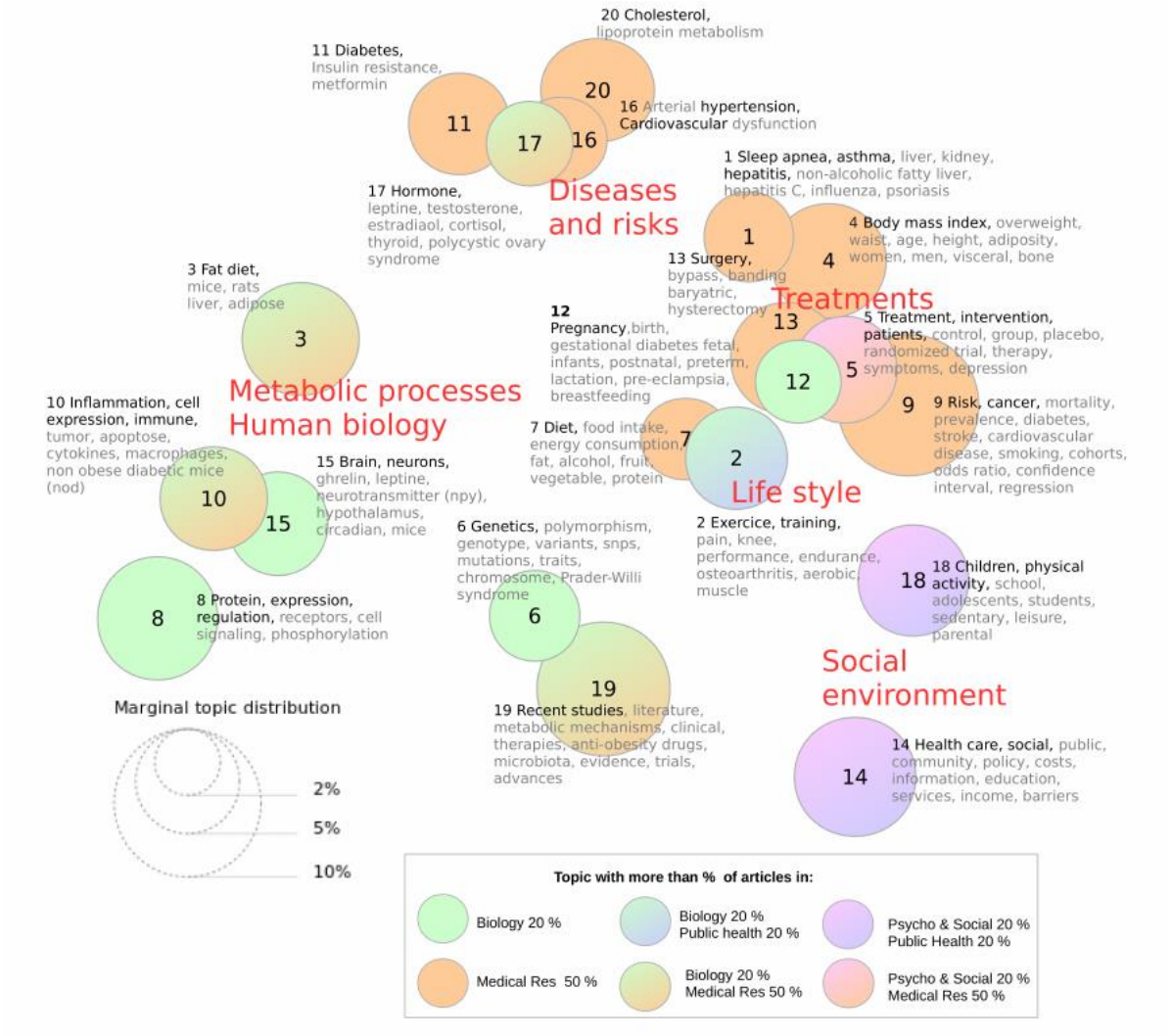
**Conclusions**

In this article, we have shown how a comparative approach can help in fostering a better alignment between R&D efforts and health needs. The data suggest that some diseases prevalent in low and middle income countries such as malaria or tuberculosis deserve more R&D investment – but this is also the case for conditions such as depression or stroke that have a high disease burden across the world. While there can be good scientific reasons for some areas having relative over- or underinvestment in relation to their disease burden, the comparison proposed can be helpful in fostering a better distribution of research resources.

The analysis can be enriched by considering the distribution of investments across research portfolios – i.e. across the diverse types of research approaches to tackle a disease. Given high uncertainty in the outcomes of research, the study of a given disease should include a variety of approaches, from basic biology to medicine, and from social studies to public health (Stirling, 2007; Nature, 2015; Wallace and Rafols, 2015).

In order for health R&D to be more attentive to the social needs, it would be commendable that the priority setting process is carried out with enriched information about health needs and in dialogue with patients and diverse stakeholders. In accordance with Open Science and Responsible Research and Innovation, priority setting should involve two main strategies: first, to gather available data on research efforts made to address a certain health problem; second to appraise stakeholder and citizen preferences on the problems to be addressed and the strategies to tackle them.

**Figure 5. Map of research topics in obesity research.** The size of the circles is proportional to the amount of research in a given topic. The proximity of circles shows the similarity between associated topics. Sources: Cassi et al. (2017)



## References


Aaron, D. G., & Siegel, M. B. (2017). Sponsorship of national health organizations by two major soda companies. American Journal of Preventive Medicine, 52(1), 20-30.

Agarwal, P., & Searls, D. B. (2009). Can literature analysis identify innovation drivers in drug discovery?. *Nature Reviews Drug Discovery*, *8*(11), 865-878.

Cassi, L., Lahatte, A., Rafols, I., Sautier, P., & de Turckheim, É. (2017). Improving fitness: Mapping research priorities against societal needs on obesity. *arXiv preprint* https://arxiv.org/pdf/1705.01151.pdf (Accessed 25-06-2017)

Catalá López, F., Álvarez Martín, E., Gènova Maleras, R., & Morant Ginestar, C. (2009). Relationship between research funding in the Spanish National Health System and the burden of disease. *Revista espanola de salud publica*, *83*(1), 137-151.

Chalmers, I., Bracken, M. B., Djulbegovic, B., Garattini, S., Grant, J., Gülmezoglu, A. M., ... & Oliver, S. (2014). How to increase value and reduce waste when research priorities are set. *The Lancet*, *383*(9912), 156-165.

Chalmers and Glasziou (2010) Avoidable waste in the production and reporting of research evidence. The Lancet.

EC (2016) 'Better research for better health. A vision for health and biomedical research from the Scientific Panel for Health'. May 2016. Available here (Accessed 18-05-2017).

Evans, J. A., Shim, J. M., & Ioannidis, J. P. (2014). Attention to local health burden and the global disparity of health research. *PLoS One*, *9*(4), e90147.

Gillum, L. A., Gouveia, C., Dorsey, E. R., Pletcher, M., Mathers, C. D., McCulloch, C. E., & Johnston, S. C. (2011). NIH disease funding levels and burden of disease. *PloS one*, *6*(2), e16837.

Glasziou, P., Chalmers, I., Altman, D. G., Bastian, H., Boutron, I., Brice, A., ... & Heneghan, C. (2010). Taking healthcare interventions from trial to practice. *BMJ*, *341*, c3852.

Gross, C. P., Anderson, G. F., & Powe, N. R. (1999). The relation between funding by the National Institutes of Health and the burden of disease. *New England Journal of Medicine*, *340*(24), 1881-1887.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, *2*(8), e124.

Kinge, J. M., Roxrud, I., Vollset, S. E., Skirbekk, V., & Røttingen, J. A. (2014). Are the Norwegian health research investments in line with the disease burden?. *Health research policy and systems*, *12*(1), 64.

NIH (2014) Report on NIH Funding vs. Global Burden of Disease. Available at: https://report.nih.gov/info_disease_burden.aspx Accessed 15-06-2017.

Røttingen, J. A., Regmi, S., Eide, M., Young, A. J., Viergever, R. F., Årdal, C., ... & Terry, R. F. (2013). Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory?. *The Lancet*, *382*(9900), 1286-1307.

Schillinger, D., Tran, J., Mangurian, C., & Kearns, C. (2016). Do sugar-sweetened beverages cause obesity and diabetes? Industry and the manufacture of scientific controversy. Annals of Internal Medicine, 165(12), 895-897.

Sarewitz, D., & Pielke, R. A. (2007). The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy*, *10*(1), 5-16.

Sismondo, S., & Chloubova, Z. (2016). "You're not just a paid monkey reading slides": How key opinion leaders explain and justify their work. *BioSocieties*, *11*(2), 199-219.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, *4*(15), 707-719.

Stirling, A., & Scoones, I. (2009). From risk assessment to knowledge mapping: science, precaution, and participation in disease ecology. *Ecology and Society*, *14*(2).

Viergever, R. F., Olifson, S., Ghaffar, A., & Terry, R. F. (2010). A checklist for health research priority setting: nine common themes of good practice. *Health research policy and systems*, *8*(1), 36.

Von Philipsborn, P., Steinbeis, F., Bender, M. E., Regmi, S., & Tinnemann, P. (2015). Poverty-related and neglected diseases–an economic and epidemiological analysis of poverty relatedness and neglect in research and development. *Global health action*, *8*(1), 25818.

Wallace, M. L., & Rafols, I. (2015). Research portfolio analysis in science policy: moving from financial returns to societal benefits. *Minerva*, *53*(2), 89-115.

WHO Global Health R&D Observatory: http://www.who.int/research-observatory/en/

WHO Global Health Estimates: http://www.who.int/healthinfo/global_burden_disease/en/

Yaqub, O. (2017). Variation in the dynamics and performance of industrial innovation: What can we learn from vaccines and HIV vaccines?. *Industrial and Corporate Change*, dtx016. https://doi.org/10.1093/icc/dtx016

## Appendix: Detailed Methodology

We obtained the global burden of disease estimates from a dedicated WHO website, using 2012 data. This data sources provides estimates of burden in terms of *Disability-Adjusted Life Years* (DALYs), which sums the years of life expectancy lost due to premature death produced by a condition (*Years of Life Lost*) and the time not fully enjoyed because of a health condition (*Years Lost due to a Disability*). These estimates are given in disease categories of the Internal Classification of Disease 10 (ICD10)[3].

We matched the ICD10 categories with medical descriptors of articles using various sources. As main source to establish the correspondence between ICD10 and the Medical Subject Headings (MeSH) we used the Unified Medical Language System (UMLS), produced by the U.S. National Library of Medicine. UMLS does not always provide the equivalence between the two classification systems and therefore in some cases we had to find the closest MeSH for a given ICD10 code. The ICD10 category of 'Endocrine, blood and immune disorders' was not considered in this study due to unclear correspondence between ICD10 and MeSH descriptors. In many cases, there is a one-to-one correspondence, but in a few cases one ICD10 category needs to be covered by various MeSH terms. We removed publications from categories to avoid counting twice a publication in two diseases. In this case, publications were left in the MeSH term closer to root of the MeSH tree (e.g. cancer related publications were removed from other disease categories and kept in 'Malignant neoplasms').

Publications were obtained from CWTS in-house version of the Web of Science (WoS) for the period 2009-2013 (5 years) on the basis of their MeSH terms. To do so, we relied in an existing matching in WoS between WoS and MEDLINE/PubMed records, which allows us the use of all the fields in both databases for the same publication.

[3] http://www.icd10data.com