

Healthcare III: The Global Burden of Disease Study

December 15, 2016

- Guest speaker: Abraham Flaxman (University of Washington, School of Global Health)
<http://globalhealth.washington.edu/faculty/abraham-flaxman>
- Focus: Big Data Science in the **Global Burden of Disease (GBD) Study**: Data Sources, Computational Methods, and Remaining Challenges in Estimating Disease Burden
<http://thelancet.com/gbd>

All images included in these notes are from the slides of Abraham Flaxman.

1 Introduction

- The Global Burden of Disease Study (henceforth GBD) is a massive, collaborative, comprehensive worldwide observational epidemiological study funded by the Gates Foundation.
- It is an effort to quantify the burden of these diseases, injuries, and risk factors across countries from 1990 to the present.
- This study allows us to ask specific questions and make comparisons over different places, times, and different types of diseases.
- They want to understand things like, how bad is Malaria? How is it in Ghana? Is it getting better or worse? How does it compare to countries near other specific countries? In addition, road traffic injuries are a big issue. How do road traffic injuries compare to malaria?
- There are two things that could be of use or interest in this area:
 - What goes in.
 - What comes out.
- Towards the beginning of this study, one of the first questions computer scientists would ask was: what are the units for measuring global burden? We need a way to figure out the value. The take-away from health economics is: do not measure health loss in dollars. Instead, the GBD measures in “years,” or DALYs:

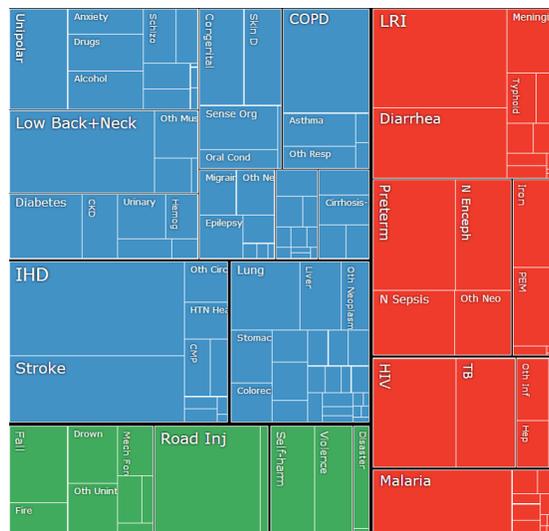
$$\text{DALYs} = \text{YLL} + \text{YLD}$$

(disability adjusted life years = years of life lost + years lived with disability)

That is, DALYs are a summary health metric which account for premature mortality (first term) and non-fatal health outcomes (second term). Note that YLDs are weighted: a weight of 1 is death, 0 is negligible. For instance, a year with depression might have weight 0.15.

There is another popular metric QALYs, which are similar but measure quality-adjusted life years, and are good (whereas DALYS are bad, although DALYS averted are good). There are ongoing discussions about which metric might be better.

- Examples: Suppose someone has a heart attack and dies at age 62. They had a theoretical maxima of age 82. Then YLLs from the heart attack are 20. On the other hand, if the same person is diagnosed with depression, then a year with depression would add 0.15 YLDs.
- Hundreds of people went through tens of thousands of papers to put all of this information together, giving the following tree map:



- Red: Infectious diseases. These are often more prevalent in low-income or less developed communities.
- Blue: Non-communicable diseases. These are often more prevalent in affluent communities.
- Green: Injuries. These are not diseases, but are worth our consideration as well. They may also be indicative of other conditions.
- There is a visualization tool “GDB Compare” which is an interactive version of this here: <http://thelancet.com/gbd/gbd-compare-visualisation>. Gates called this an overly complicated but very useful website.

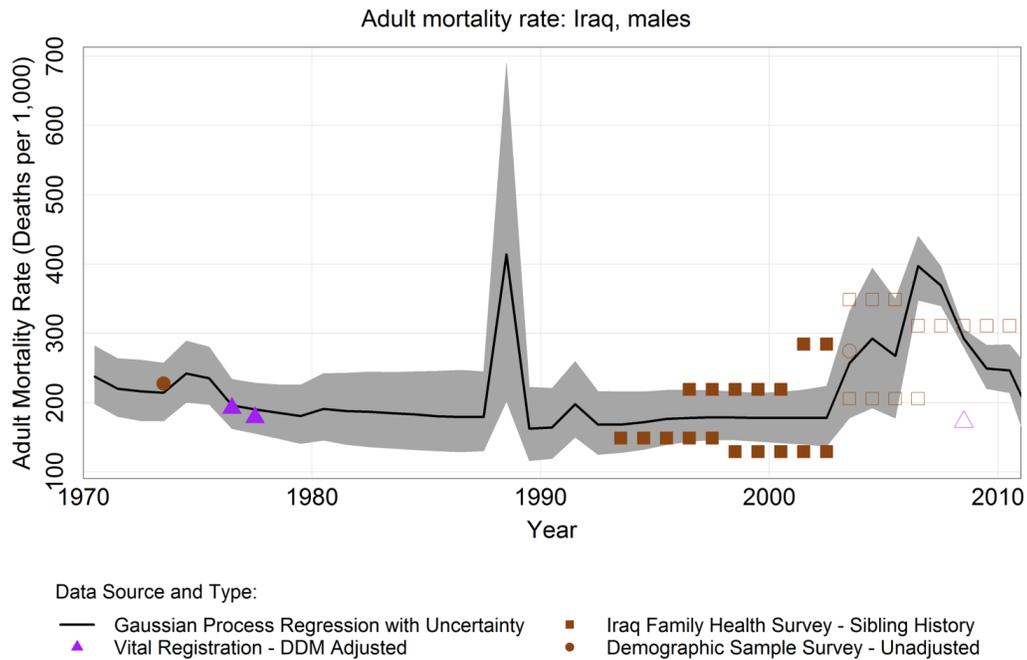
Notes on a demo of this tool:

- Switch from level 3 to 1 and play around with the year. Go to “explore” — it is clear that the US is later in the epidemiological shift.
- Note that the numbers are consistent for each type of cause, but the rate is decreasing as the population increases.

- The epidemiological profile of China looks a lot more like the US than, say, Zimbabwe.
- Note that in 1990, infectious diseases were the majority of the burden. That has been changing both in developing and developed countries.

2 Data Collection and Incentive Issues

- After establishing such a metric, we run into the problem that people care so much about these metrics that they want them to come out a certain way (e.g., lower to show that a particular project is working or higher to attract more funding/attention/effort toward a particular disease.) We have similar issues in different domains such as in education.
- To help curb this problem, GBD has some practices in place. For instance, they independently estimate the total number of deaths, then ensure that the deaths due to the 200 diseases and injuries add up to that total. This assumes that each death is caused by exactly one disease or injury.
- One issue is that we need to be able to properly estimate the total number of deaths. This is an easier task in countries such as the US where the Census Bureau keeps careful records, but might be more challenging in countries such as Iraq where such data is not readily available and researchers have to come up with models to estimate the number of deaths while sometimes missing years of data. In the example below, they did not have data after 2006:



*Hollow points indicate data excluded from the analysis

- There are opportunities for mechanism design to help incentivize better records to be kept. Currently, the largest source of this data is from death certificates. Under the status quo, death certificates must indicate that a single death has a single cause. This can cause errors; for example, self-harm (an injury) might be recorded as the cause of death, and depression, which may have been a reason for the self-harm, would not get recorded as the cause of death because there can only be one cause of death.

INTERNATIONAL FORM OF MEDICAL CERTIFICATE OF CAUSE OF DEATH		
Cause of death		Approximate interval between onset and death
I Disease or condition directly leading to death*	(a) Pulmonary embolism due to (or as a consequence of)	hours
Antecedent causes Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last	(b) Pathological fracture of Femur due to (or as a consequence of)	2 days
	(c) Secondary malignant neoplasm (femur) due to (or as a consequence of)	2 months
	(d) Malignant neoplasm of breast (nipple)	1 year
II Other significant conditions contributing to the death, but not related to the disease or condition causing it	Essential hypertension	5 years
	Obesity	10 years
*This does not mean the mode of dying, e.g. heart failure, respiratory failure. It means the disease, injury, or complication that caused death.		
If means the disease, injury, or complication that caused death.		
This does not mean the mode of dying, e.g. heart failure, respiratory failure.		
condition causing it		
not related to the disease or contributing to the death, but		
Other significant conditions		
II		

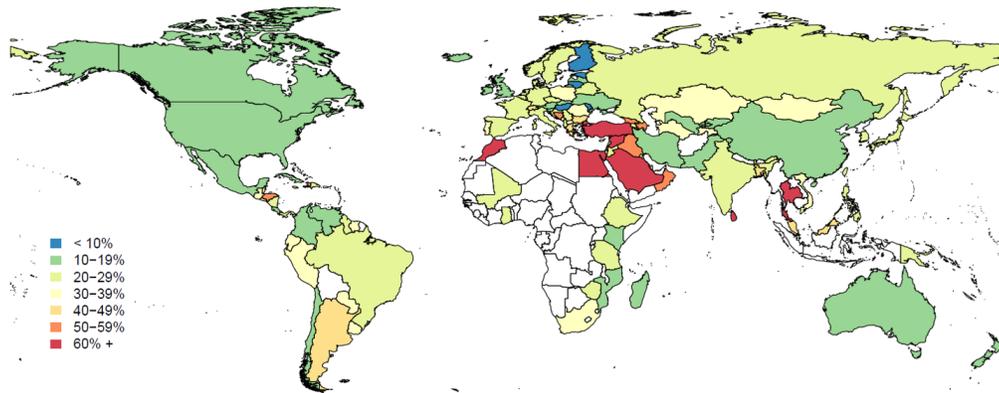
As pictured above, doctors are supposed to list the single cause of death and then any antecedent conditions specifically relating to the cause of death, in order of most recent to least recent.

There are also many examples of death certificates which contain obviously incorrect data.

INTERNATIONAL FORM OF MEDICAL CERTIFICATE OF CAUSE OF DEATH		
Cause of death		Approximate interval between onset and death
I Disease or condition directly leading to death* Antecedent causes Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last	(a) Acute myocardial infarction due to (or as a consequence of)	1 hour
	(b) Essential Hypertension due to (or as a consequence of)	1 year
	(c) Diabetes mellitus 2 due to (or as a consequence of)	14 years
	(d) Hypothyroidism	24 hours
II Other significant conditions contributing to the death, but not related to the disease or condition causing it		
<i>*This does not mean the mode of dying, e.g. heart failure, respiratory failure. It means the disease, injury, or complication that caused death.</i>		
<small>If means the disease, injury, or complication that caused death. This does not mean the mode of dying, e.g. heart failure, respiratory failure.</small>		
condition causing it not related to the disease or contributing to the death, but Other significant conditions II		

Causes should be listed in order of most recent to least recent, but 24 hours is more recent than the two conditions before it.

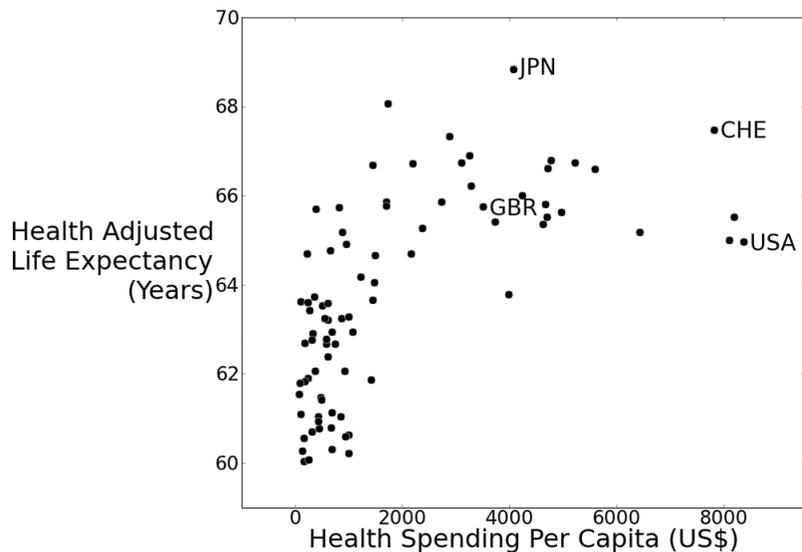
- As it stands, many countries do not have death certificate data that is usable, if at all. In the map below, countries in blue and green are those where at least 80 percent of the death certificates have valid causes of death, while those in orange and red have more than half of the death certificates with an underlying cause of death which cannot be used for burden estimates. The countries that are not colored in have no death certificate data at all!



- Furthermore, in some countries, these death certificates are filled out via a verbal autopsy survey through a structured interview with a relative of the deceased who is aware of the health experience, but this is complicated.
- There are a couple of high-level issues with recording deaths:
 - Issue #1: Doctors or survey collectors currently have no incentive to do a good job filling out death certificates, and getting the information to fill them out correctly is hard. There are no rewards and no verification, and many doctors would prefer to spend their time treating people instead.
 - Issue #2: The procedure specifies that only conditions related to the single cause of death should be listed. This ignores other possibly related causes.
- This may not be a huge problem for the global estimates because most inaccuracies are happening in the older patients. In the measure of DALYs, older patients are weighted less because the theoretical remaining years are fewer.

3 What Comes Next?

- There is real demand for higher precision.
 - Now that we have some answers for what is going on, the next question is what are we going to do about it.
 - The US spends the most per capita on health care compared to most countries, but the health adjusted life expectancy is 65, as opposed to Japan which is about the same life expectancy but spends about half.
-



- Incidence vs prevalence DALYS: The key question in incidence vs prevalence DALYS is do you “charge” the years of burden to the moment the disease begins (incidence DALYS) or do you spread the charge across the duration of the disease (prevalence DALYS)? For GBD headline estimates, we currently mix them: incident years of life lost are counted at the moment of the death, while prevalent years lived with disability are counted across the duration of the disease. This is confusing, but it gets confusing no matter how you do it.
- Men and women not evaluated separately: data shows that differences in life expectancy by gender are not due to differences in health, but rather due to health behaviors (like smoking or not).

4 Questions

- The DALY metric is imperfect. Take for instance, someone who dies early in a year X, then all the cost in terms of DALYs gets allocated to year X. So in principle, a natural disaster or war that killed many people in year X would create a lot of DALYs in that year, but then potentially reduce DALYs in years X+1, X+2, and so forth, since the people who would have contributed to the DALYs in X+1 onward died in year X, and year X absorbed the full cost. As a result, in principle, a reduction in DALYs in a given year X+1 might reflect something bad in the recent past, rather than something good in year X+1. This was an extreme example – not something that’s likely to occur in practice, but more to motivate the question of why an early death in year X doesn’t contribute its DALYs equally across X, X+1, X+2, and so forth up to the life expectancy of the person; it wasn’t clear why it was the right accounting to charge it all to year X.
- What else is the same data used for? Are there incentives that might bias reporting?
 - Countries are interested in their national rankings, and that pressure is passed down through health systems to try to do better (to perform better, but this is bad if it leads to misreporting better performance).

- Having a high GDP makes a country look successful; similarly with health outcomes.
 - One way of gaming the system: what’s a neonatal death and what’s a still birth? Misreporting would reduce the “disease.”
 - Do we know anything about how the data is collected and what it’s used for?
 - There is some information published on this. It is somewhat sensitive: the government agencies that are responsible do not want to show that they are doing a bad job.
 - When you’re trying to figure out spending how do you account for confounding factors? E.g., Japan has low birth rate. From a mechanism design perspective is just using this DALY approach
 - Where are the biggest holes? The most interesting investigation comes from Deacon: hacking death by finding security exploits in registration perspective. This describes how you can gain access and kill people on paper.
 - Take-away message to pass on for future mechanism design: use DALYs to measure health (as opposed to money). At minimum, everyone agrees that increasing DALYs is good.
 - Are we worried that addressing the objective as DALYs could not actually address the issue we care about, that it’s a bad proxy and could have other negative impacts?
 - There is an expectation that HIV is underreported because there’s a stigma associated with it: the person filling out the report doesn’t want to acknowledge it, so they don’t list it as the main cause. As a result, there is a rule that specifically says if it says HIV anywhere for someone, that’s the cause of death (to account for this underreporting).
 - What are the penalties for misreporting? What are the rules around this space?
 - There’s no money attached to filling out death certificates correctly. They is no pay for performance.
 - There is actually a huge pot of money coming from the Bloomberg Foundation to try to improve the quality of this data.
 - There is a list of causes that are notifiable conditions: the world wants to know if anyone shows up with some disease (i.e. smallpox). These things actually get tracked carefully.
 - How do you allocate your finite time as a doctor to reduce DALYs? Is this ethical?
 - No, people are pretty uncomfortable choosing priority based on DALYs when choosing between three different people in the ER and who to treat first.
 - Is there any sort of optimization done? For example: here’s our finite budget, how should we allocate funding to maximize DALYs saved? How is the data actually used in allocation?
 - There are some efforts from the NIH to use this data for funding allocation: How do we fund compared to what’s burden-like for the US vs. globally?
-

- The disability weights are getting a lot of attention, and they make a huge difference. The first pass had doctors around a table making estimates on the weight of burden of a disease. They didn't match up with the weights from the study. The study's disability weights are universal, but a doctor's personal estimate is really specific and not universal. This first pass had a small number of doctors around a table with specific experiences. Instead, a sample from all around the world is needed.
 - Now they ask: What do you think a migraine is? How does one live with it? Should we invest in program X or prevent Y many people from death? Using these questions, we can get a rank ordering. These are pretty consistent across different cultures. thus we're not on too shaky ground.
 - Previously, funding was allocated based on advocacy. One reason the GBD study is so powerful is because every special interest group wants to say how important it is for limited resources to go to some specific cause, and this data allows more objective allocation.
- In reference to image in Section 3: Is there a belief that there should be a straight line between DALYs and spending? What if the return on investment isn't the same for different illnesses?
 - Yes, there should be a lot more to the story than just the two propositions.
 - For example, maybe the research for a certain condition isn't there, so it's a burden we don't have the care to spend the money on. Or, maybe we're spending money on something that isn't a large burden because preventative care is money well spent.
-