Using Credence Calibration for Everything

1. Prediction-based medicine (PBM)
2. Preventing overcharging by prosecutors
3. Prediction-based-Medicine instead of ~~Evidence-based-Medicine~~ Authority-based-Medicine

Prediction-based medicine (PBM)

We need a new paradigm for doing medicine. I make the case by first speaking about the problems of our current paradigm of evidence-based medicine.

The status quo of evidence-based medicine

While biology moves forward and the cost of genetic-sequencing dropped a lot faster than Moore's law the opposite is true for the development of new drugs. In the current status quo the development of new drugs rises exponentially with Eroom's law. While average lifespan increased greatly about the last century in Canada the average life span at age 90 increased only 1.9 years over the last century. In 2008 the Centers for Disease Control and Prevention reported that life expectancy in the US declined from 77.9 to 77.8 years. After Worldbank data Germany increased average lifespan by two years over the last decade which is not enough for the dream of radical lifespan increases in our lifetime.

When it costs 80 million to test whether an intervention works and most attempts show that the intervention doesn't work we have a problem. We end up paying billions for every new intervention.

Eric Ries wrote "The Lean Startup". In it he argues that it's the job of a startup to produce validated learning. He proposes that companies that work with small batch sizes can produce more innovation because they can learn faster how to build good products. The existing process in medicine doesn't allow for small batch innovation because the measuring stick for whether an intervention works is too expensive.

In addition the evidence-based approach rests on the assumption that we don't build bespoke interventions for every client. As long as a treatment doesn’t generalize about multiple different patients, it’s not possible to test it with a trial. In principle a double-blind trial can't give you evidence that a bespoke intervention that targets the specific DNA profile of a patient and his co-morbidity works.

The ideal of prediction-based medicine

The evidence-based approach also assumes that practitioners are exchangeable. It doesn't model the fact that different physical therapist or psychologists have different skill levels. It doesn't provide a mechanism to reward highly skilled practitioners but it treats every practitioner that uses the same treatment intervention the same way.

Its strong focus on asking whether a treatment beats a placebo in double-blind studies makes it hard to compare different treatments against each other. In the absence of an ability to predict the effect sizes of different drugs with the literature the treatment that wins on the market is often the treatment that's best promoted by a pharmaceutical company.

How could a different system work? What's the alternative to making treatment decisions based on big and expensive studies that provide evidence?

I propose that a treatment provider should provide a patient with the credence that the treatment provider estimates for treatment outcomes that are of interest to the client.

If Bob wants to stop smoking and asks doctor Alice whether the treatment Alice provides will result in Bob not smoking in a year, Alice should provide him with her credence estimation. In addition Alice's credence estimations can be entered in a central database. This allows Bob to see Alice's Brier score that reflects the ability of Alice to predict the effects of her treatment recommendations.

In this framework Alice's expertise isn't backed up by having gotten an academic degree and recommending interventions that are studied with expensive gold-standard studies. Her expertise is backed by her track record.

This means that Alice can charge money based on the quality of her skills. If Alice is extremely good she can make a lot of money with her intervention without having to pay billions for running trials.

Why don't we pay doctors in the present system based on their skills? We can't measure their skills in the present paradigm, because we can't easily compare the outcomes of different doctors. Hard patients get send to doctors with good reputations and as a result every doctor has an excuse for getting bad outcomes. In the status quo he can just assert that his patients were hard.

In prediction-based medicine a doctor can write down a higher credence for a positive treatment outcome for an easy patient than a hard patient. Patients can ask multiple doctors and are given good data to choose the treatment that provides the best outcome for which they are willing to pay.

In addition to giving the patient a more informed choice about the advantages of different treatment options this process helps the treatment provider to increase his skills. They learn about where they make errors in the estimation of treatment outcomes.

The provider can also innovate new treatments in small batches. Whenever he understands a treatment well enough to make predictions about its outcomes he's in business. He can easily iterate on his treatment and improve it.

The way to bring prediction-based medicine into reality

I don't propose to get rid of evidence-based medicine. It has its place and I don't have any problem with it for the cases where it works well.

It works quite poorly for body work interventions and psychological interventions that are highly skill based. I have seen hypnosis achieve great effects but at the same time there are also many hypnotists who don't achieve great effects. In the status quo a patient who seeks hypnosis treatment has no effective way to judge the quality of the treatment before he's buying.

A minimal viable product might be a website that's Uber for body workers and hypnotists. The website lists the treatment providers. The patient can enter his issue

and every treatment provider can offer his credence of solving the issue of the patient and the price of his treatment.

Before getting shown the treatment providers, a prospective patient would take a standardized test to diagnose the illness. The information from the standardized test will allow the treatment providers make better predictions about the likelihood that they can cure the patient. Other standardized tests that aren't disease specific like the OCEAN personality index can also be provided to the patient.

Following the ideas of David Burn's TEAM framework, the treatment provider can also tell the patient to take tests between treatments sessions to keep better track of the progression of the patient.

When making the purchasing decision the patient agrees to a contract that includes him paying a fine, if he doesn't report the treatment outcome after 3 months, 6 months and 1 year. This produces a comprehensive database of claims that allows us to measure how well the treatment providers are calibrated.
Various Quantified Self gadgets can be used to gather data. Many countries have centralized electronic health records that could be linked to a user account.

The startup has a clear business model. It can take a cut of every transaction. It has strong network effects and it's harder for a treatment provider to switch because all his prediction track record is hosted on the website.

Thanks to various people from the Berlin Lesswrong crowd who gave valuable feedback for the draft of this article.

Preventing overcharging by prosecutors

In criminal cases in the United States prosecutors often add a lot of charges for a defendant to have ammunition for coercing the defendant into a plea deal. This is toxic because the defendant doesn't know which of those charges are likely to hold up in court if the case would be decided by a jury. Given that there's little cost to the prosecutor for adding additional charges, defendants are often overcharged.

I propose that whenever a prosecutor files a charge for a defendant, the prosecutor should state the likelihood that in the absence of a deal the court will find the defendant guilty of the charge. The ability of the prosecutor to accurately access the likelihood can be measured via the Briers score or a Log score.

The current score should be publicly accessible on the website of the court. This allows the defendant to know whether they can trust the likelihood values the prosecutor gives. The score should also be printed on ballots when the prosecutor seeks reelection to create much higher incentives for the prosecutor to give the correct likelihood then convicting a lot of people.

After the prosecutor provides the likelihood for the charges it's much easier for a defendant to make a good decision about whether taken a given plea deal is in their interest. Prosecutors with a good Briers score will be able to make more plea deals to reduce their overall workload because it's easier for the defendant to know that a deal is in their interest.

While this reform wouldn't fix all problems with plea deals, as some plea deals are due to the defendant being given charges that would actually hold up in court given the existing criminal code, the reform will provide defendants with fairer plea deals. Defendants getting fair plea deals is good for the system given that it keeps overall legal costs down.

I would expect that many juries will automatically throw out a 10% or 20% charges because there will be people on the jury who would argue that the prosecutor thinking that there's only a 20% chance that a charge holds means that there's reasonable doubt against the charge.

This reform is both in the interests of citizens who care about law and order and citizens who care about reducing sentencing overall as everybody should be interested in prosecutors providing fair plea deals.

Prediction-based-Medicine instead of Evidence-based-Medicine Authority-based-Medicine

This article explores how evidence-based Medicine currently works in 2021 with ivermectin as a case study. I don't make any argument for or against ivermectin but look at the deeper issues about how we decide which drug and expert to trust. First I explore the status quo and then I will present Prediction-based Medicine as an alternative to the status quo.

The idea of Evidence based Medicine

In the last three decades the idea of Evidence-based Medicine arose. A large part of what Evidence-based Medicine is about is doctors reading the medical literature and making treatment decisions based on the evidence available in the literature.

There's no study that showed any benefit of doctors reading the medical literature, so instead of providing evidence that Evidence-based Medicine works based on the standards of evidence that Evidence based Medicine proposed, the framework gets accepted without empiric evidence and is only backed up by intellectual arguments. For a long-time that wasn't a real problem and everybody more or less agreed on what the dogma of Evidence based Medicine is supposed to mean.

Ivermectin

At the moment we have two meta-analyses about Ivermectin who went through competent peer review before getting published. With competent peer review I mean review by people who are competent and familiar enough with the subject to spot when control and placebo get exchanged in the reporting of one trial.

The two competently peer reviewed meta-analyses are published in Reviews in Medical Virology and American Journal of Therapeutics who have impact factors of 4.221 and 1.955 which suggests that both normal journals and not obscure ones that are willing to publish anything. Given that more prestigious journals are more likely to publish wrong data, calling for meta analysis is good enough and it doesn't make sense to require it to be published in more prestigious journals.

While the two meta-analyses conclude that there's moderate-certainty evidence for ivermectin working, authorities like the FDA, CDC, WHO and EMA reject ivermectin as a treatment.

As a doctor deciding how to treat a patient, the doctor has to make a decision what to trust. While we can make theoretical arguments about whom we should trust, it's essentially about the question in which authority you want to believe. There's no empiric evidence for a heuristic that's proven to help us decide whom to trust. This is bad because one side of the debate is likely right and given that some doctors follow either side this means that a lot of patients get bad treatment.

Deciding about Ivermectin is about believing in heuristics of authorities

I could make an argument here for the decision making heuristics of one authority or the other, but the argument I could make wouldn't be based on good empiric evidence. This has the great advantage that even when people read my article in a few years, I run no risk of being seen as having made the wrong call.

Instead, I argue that this case shows that medicine currently is essentially about believing the heuristics of authorities without any evidence which heuristics are the best. When people argue that they make a treatment decision because the CDC, FDA, WHO or EMA holds a position, the decision essentially is about trust in authority.

Prediction-based Medicine

Once you accept that treatment decisions are about believing in authorities, the logical conclusion is to ask how to determine which authority is better than another. This is similar to the problem of determining which political analyst to trust. Philip Tetlock found a way to evaluate the experts. To evaluate them he asked them to predict political outcomes in the future. Modern statistics like the Briers scoring rule or the Logarithmic scoring rule then allow us to evaluate the predictions.

The same process can be used in medicine to decide whom to trust. This can be done at two levels. At one level we could ask individual scientists and institutions to predict the outcome of clinical trials. We can also ask doctors to predict the outcomes of individual patients.

While the first is also valuable, I will focus here on the second. When dealing with a pandemic of a new virus, in the beginning no trials are available about what the best treatments are but there are still patients that need to be treated.

The FLCCC is a group of Intensive Care Unit (ICU) doctors who came up with treatment guidelines that differed from the official guidelines. One way to evaluate whether to trust them when Dr. Pierre Kory of the FLCCC gave his first senate testimony in May 2020 would be to ask: "How good are you at predicting the outcomes for your COVID-19 patients compared to ICU doctors who follow the standard protocol." If Dr. Pierre Kory has a bad score at predicting outcomes for COVID-19 patients he should be ignored. If he however advocates a different treatment and is good at making predictions that suggests that other ICU doctors are well advised to learn from him.

In a pandemic where good fast response to treat the first wave of patients is very important, this process allows expertise to travel much faster through the medical community because it's easy for doctors to identify the people who understand treating COVID-19 well and learn from them.

Outside of the pandemic we can give those doctors with the best prediction scores in their field seats at the table where the treatment guidelines are written because they are the people who are best at synthesising all available evidence to make treatment decisions for a particular disease.

Conclusion

We need to use the empiric method to find out how to best interpret the available evidence. Asking doctors to predict outcomes of their clinical cases is an empiric method that allows for scoring their prediction, which gives us a good heuristic to decide whom to trust.