6.S897/HST.956 Machine Learning for Healthcare

Lecture 2: Overview of Clinical Care

Instructors: Peter Szolovits

The lecture goes over the overview of the clinical care from the past to the present. There are many aspects concerning health care. In a general way, what are the goals of health care? What are the things that people actually do to try to achieve these goals? Paying for health care is also an interesting aspect since a lot of problems that come up and a lot of the interest that people show in the analysis of health care are motivated by money.

The lecture covers 4 main topics:

1. Goals of Medicine
2. Tasks of Medicine
3. Keeping Public Health
4. Paying for Health Care

1 Goals of Medicine

Figure 1: Interaction between shaman and his patients.

In general, one may say that the goals of medicine are to cure the sick and to keep people healthy. Health care helps cure the sick through the healer. In the left photo of Figure 1, shaman uses his knowledge and experience he had with other patients to cure the sick. This is an effective form of health care, but there is no much record keeping. On the right is the traditional shaman who uses hands-on interaction with patient. Both emphasize the importance of the interaction between the healer and the patient. Asking a futurist doctor about what the medicine will be like in the future, they will emphasize that the role of healer is not an automaton who just figures out what is the right thing but to persuade the patient to trust to follow what he or she is suggesting. This can also be more beneficial through the placebo effect.

But before we evaluate how good we are doing to reach these goals, we first need to consider the notion of "health". What is "health" exactly? According to the World Health Organization (WHO), "health" is

> a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity

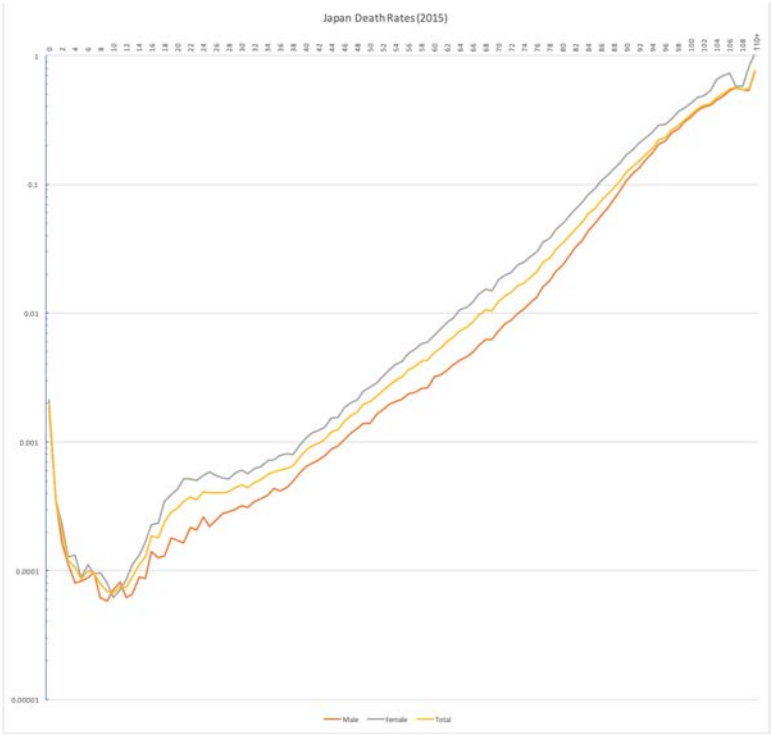
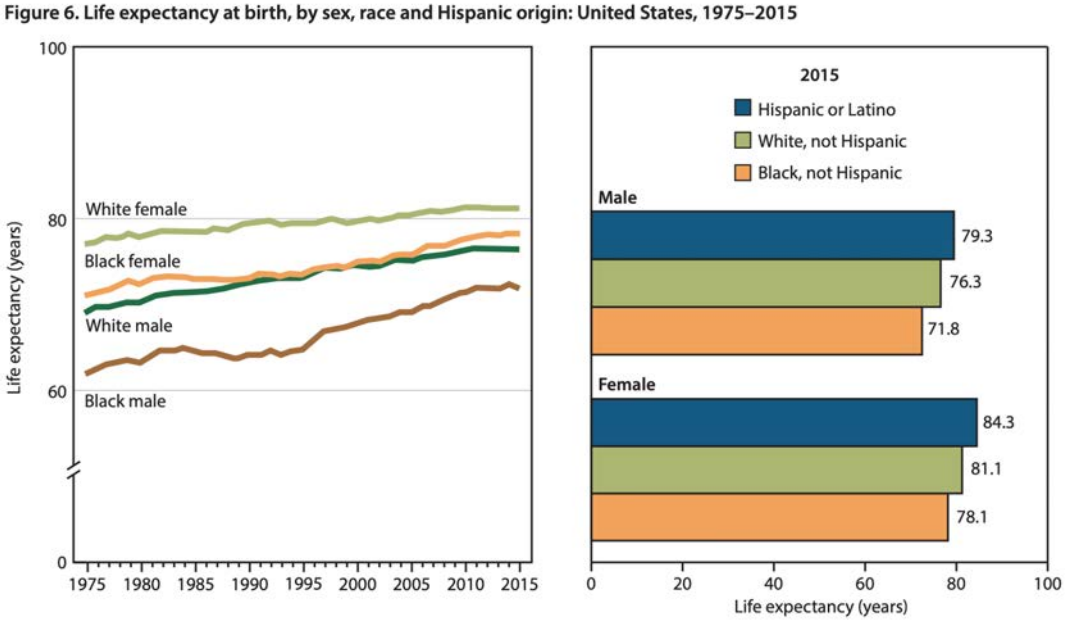


Figure 2: Mortality rate by age in Japan, 2015.



Courtesy of the CDC. Image is in the public domain.

Figure 3: Life expectancy at birth by ethnicity.

In addition to the definition of "health", one of the health notions that is easy to measure is life expectancy or longevity (how long people live.) In general, human's life expectancy in each country is increasing. Even when comparing between the short period of time, life expectancy is also significantly increasing, e.g., a Rwandan male in 2018 is expected to live 62.6 years as opposed to a Rwanda male in 2001 with life expectancy of 38.35 years because of the genocide.

Both age and ethnic also play a crucial role in longevity. Figure 2 depicts the mortality rate by age in Japan, 2015. The shape of the curve, however, is generally the same in other countries. Infants have a high mortality rate and the rate drops as they get older. However, when they become teenagers, they are more risk seeking, and the mortality rate increases. Figure 3 also shows the disparity of mortality in each ethnicity. In the United States from $1975 - 2015$, white people tend to live longer than black people, while Hispanic are likely to live longest for both male and female.

The causes of death also have a close relationship to the life expectancy trend. In the USA 2014, the top causes of death are chronic diseases such as heart disease or cancer. Only very few causes are infectious. This is interesting because the causes of death of people in the past, say 1850s, are very different. They are mostly infectious diseases and people generally do not live long enough to die from chronic diseases. People also develop more chronic conditions such as arthritis, hypertension or hearing impairment when they live longer.

While, in general, the longer people live the better, we have to take life quality into account too when we evaluate the outcome. A model in 1 represents the quality of life as the integral of the product between the life quality function over time $q(t)$ and the discount function $g(t)$. The discount factor takes into account the fact that people prefer to experience pain later in their life. There is no way to know what is the right discount rate, but one can borrow the financial discount factor as the rough approximation of the discount factor g. One way to measure the quality of life q is to take into account the daily activities that people could perform. These include basic activities such as showering, dressing, or self-feeding or instrumental activities such as moving within the community, preparing meal, or using the telephone.

$$V_N = \int_{t=0}^{T} q(t)g(t-N)dt \tag{1}$$

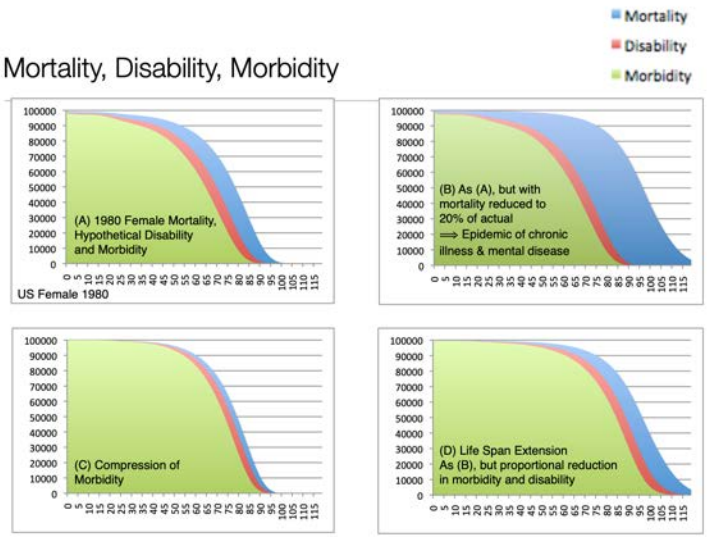


Figure 4: Mortality, Disability and Morbidity curves of US female in 1980.

Since life expectancy is not the only factor, we must consider not only the notion of mortality, but also disability and morbidity. In general, people start from being healthy and well, then they move on to

morbidity phase, the state of being diseased. This then leads to disability and mortality respectively. The top left figure in Figure 4 depicts the morbidity curve (between green and red regions), disability curve (between red and blue regions) and the mortality curve (outer edge of blue region.) If, however, we could change the characteristic of these curves, how should we aim to change them? Decreasing the rate of mortality as in the top right figure in Figure 4 may allow people to live longer but people will also have a longer disability life. On the other hand, in the bottom left figure, we have the same mortality rate but people are being healthy longer and have a very short morbidity and disability phase.

In addition to the individual quality of life, we also need to consider the social quality of life. This is a hard question since we do not know what is the best scenario for a high social quality of life. If the social quality of life is just an aggregation of quality of life of all people, we could just increase the number of people to increase indefinitely. But that does not seem to be the case as having more people leads to problems and lower quality of life overall. It is unclear too whether having less people would increase the social quality of life as many current developing countries are suffering from under-population problem.

The longer longevity also leads to a change in notion of medicine. In the past, people go to doctor to get cured. They want the doctor to cure acute illness such as infection or broken arms. Currently, as we live longer, we focus more on managing long-term chronic disease. We want to understand more how these long-term diseases develop. In the future, the notion of medicine will move on to prevention and prediction of diseases.

2 Tasks of Medicine

In general, the interaction between the doctor and the patient involves tasks such as diagnosis, prognosis, and therapy.

- Diagnosis: having a doctor figuring out what is wrong with the patient
- Prognosis: predicting what will happen to the patient if the doctor does nothing
- Therapy: figuring out what to do and how to cure the patient

To execute these three tasks, doctor usually goes through the medical cycle, or even larger, the enterprise-level clinical process. This finally leads to a question of how the health system learns.

2.1 The Medical Cycle

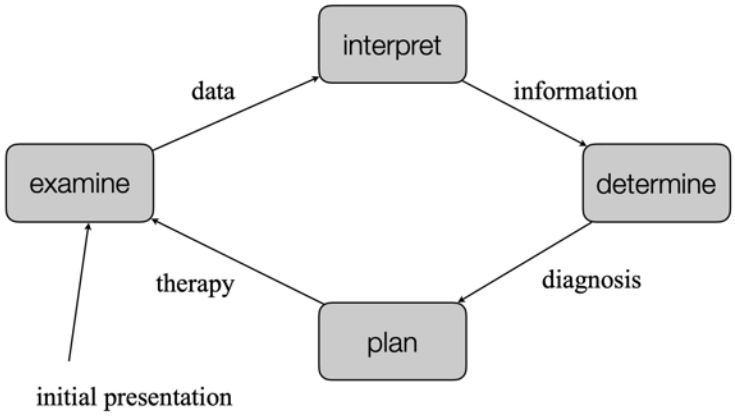


Figure 5: Diagram representing the medical cycle.

Figure 5 depicts the cyclic process of care. The process starts with the initial presentation where the doctor interacts with the patient, asking the patient about the conditions. The doctor then examines the patient and generates data which is mostly measurement. The measurement can also be based purely on the observation. After interpreting the data and generating information, the doctor then determines the diagnosis. It is crucial to note that it may not be so much of determining the diagnosis as making a guess. Based on the instructor experience, doctors are quite willing to make a guess because it is useful to believe that they understand what is going on and concentrate on the potential next steps. Otherwise, doctors have to explore the vast number of possible things, which does not provide a good guidance. **Machine Learning on the other hand does not have such limitation and can potentially explore the vast number of possibilities.** This is one of the important benefit Machine Learning has over doctors.

Having made the diagnosis, the doctor plans a therapy and executes it. If the patient gets better, then the diagnosis is right. If the patient does not get better, the doctor reexamines how the conditions of the patient are different from what the doctor expects it to be and makes a revision, starting the cycle again.

This concept of reinterpreting the entire process is very critical to the medical process. If we want to apply machine learning to the same process, we must also follow its cyclic nature. Machine Learning in health care goes all the way back in 1950s when Alan Turing mentioned health care as one of the interesting area of Artificially Intelligence. He predicted that once there was enough data, many models for health care would be built. **However, a lot of the early approach of Artificial Intelligence in health care took a one-shot approach where people tried to build a model that gets in the snapshot data of the patient, gives out the diagnosis, and ends the process. Such approach does not work very well because it does not obey the cyclic nature of the process of providing health care.** Only until 1980s when people started to realize that the model has to be in the process for the long run, not for just a single interaction.

2.2 Enterprise-level Clinical Process

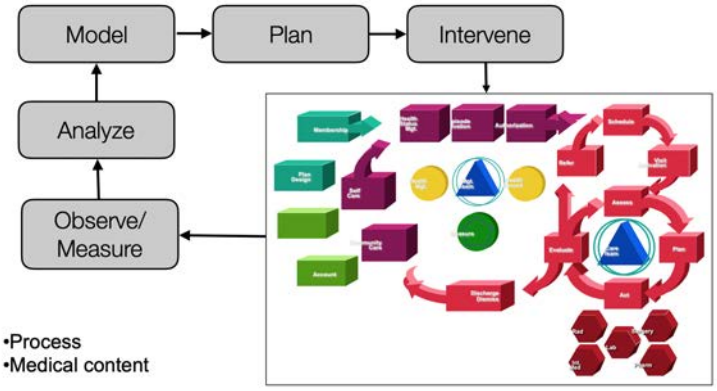


Figure 6: Diagram representing the enterprise-level clinical process.

The enterprise-level process of care is a larger cycle. The medical cycle is just a part of the whole process. To improve the process, we need to make the observation about it, analyze it, model it, plan how to improve it and intervene the system. This class of task is central to the improvement of health care since it observes how health care system works and figures out how to make it better. In general, we can improve health care in any of these 3 level.

- More acute phase level: dealing with the patients that are in the middle of illness.
- Medical cycle level: dealing with patients that are in the cycle of being well and being sick

- System level: Designing health care, potentially on the enterprise-level, that works better for the population.

Learning such health systems and how to improve them are studied extensively. More study can be read in [FAD+17].

2.3 How Does the Health System Learn?

Health system recently shifted from a "traditional-based medicine" to an "evidence-based medicine" where you make your analysis based on the evidence. Such concept is intrinsic to the Engineers, but was a novel concept for the doctors. Specifically, the idea is to use the Randomized Controlled Clinical Trial (RCT.)

Let's imagine that a drug company comes up with a new drug A and wants to prove that it is more effective than drug B for condition X. Ideally, RCT should involve the following steps.

- Find patients that is suffering from X, ideally just X and not suffering from anything else.
- Consult with a statistician and design an experiment of how to collect data from the patients.
- Pre-define metrics to measure which drug is doing better.
- Have the statistician determines how many patients should enroll in the trial to achieve a small p-score.

However, such ideal steps are not achieved in practice because of several reasons. It is very hard to find patients that fit the trial criteria. The drug company that involves also wants the trial to be quick and have smallest number of patients as possible.

Another alternative approach to learn health system was suggested by the Institute of Medicine or the National Academy of Medicine as follow:

> one in which progress in science, informatics, and care culture align to generate new knowledge as an ongoing, natural by-product of the care experience, and seamlessly refine and deliver best practices for continuous improvement in health and health care

This is, however, hard to do in practice since people are not treated by experimental protocol. There is no randomization since the doctors have to treat the patients in the way that they think is the best for the patients. We also need a technical infrastructure for truly big data that records historical data of numerous patients. It needs to be organizes in a way that easy to make analysis.

3 Keeping Public Health

In order to keep people healthy, we also need to consider how to maintain public health. We will explore 3 aspects involving public health: tracking disease, tracking infections, and quarantine.

3.1 Tracking disease

The attempt to track down and classify disease traces all the way back. Here are the historical attempts in the past.

- In 17^{th} century, John Graunt estimated that the mortality before age 6 in that time was around 36%.
- In 18^{th} century, Sauvage, Linnaeus and Cullen made first attempt at systemic classification of disease.
- In 1853, William Farr introduced in the first International Statistical Congress the first disease classification system. The system divided diseases into categories: epidemic diseases, constitutional diseases, local diseases, developmental diseases, and diseases from violence. Note that this was before Louis Pasteur and his theory of the cause of disease.

- After Louis Pasteur's discovery, in 1890s, Bertillon introduces a more comprehensive and hierarchical system of disease classification.
- In 1920s-40s, the International Classification of Diseases (ICD) is introduced.
- In 2015, the 10^{th} revision of ICD (ICD-10) came out. Now ICD is regulated by WHO.

Furthermore, Centers for Disease Control and Prevention also collects from all over the place death certificates that record the direct and indirect causes of death. The certificate can be used to determine the immediate cause of death, the approximate cause of death and the underlying cause of death of each person. This is the statistical data that we now have available for improving disease classification and for learning other medical task.

3.2 Tracking Infections

There has been many examples of tracking infections in the past. In 1854, there was an attempt to find the cause of Cholera. People came to the wrong conclusion that the cause of Cholera was miasma (bad air). John Snow instead plotted on the map the position of infected patient and discovered that the cause was the unsanitary of the city water pump.

In 2003, there was also an attempt to study the cause of West Nile Virus (WNV), a nasty mosquito-borne disease. With very similar same technique as John Snow, they discovered that WNV came from the airplane through JFK airport.

3.3 Quarantine

Quarantine is the practice to isolate sick people with a contagious disease from people who are not sick. This controversial practice used to be used a lot by public health official. Here are some interesting examples.

- If you are sick when you arrive at Ellis Island, you will be subjected to a quarantine, isolating you in a building to observe how sick you are.
- Typhoid Mary was a carrier of typhoid fewer, but was not herself affected. she was a cook, and whenever she was employed, people were sick. As a result, she was jailed in a sanatorium to prevent her from infecting other people.
- Jesse Helms suggests to have all people with AIDS quarantined to prevent the spread of the disease.
- Quarantine of people with Ebola in Africa. The report suggests that quarantine is a very controversial practice that conflicts with human rights and should only be used as a last resort. The technical conclusion also suggests that quarantine in urban area is really hard and involves a proper waste disposal problem.

4 Paying for Health Care

The fundamental cause of the high cost of health care is that in general, nobody wants to die, and when people is sick, they want the best possible care no matter how much it costs. As a result, many drug therapies are very expensive. There is no party that tries to push the cost back except for the insurance company, who just passes the cost to the client.

However, there is a lot of waste in the process. High cost sometimes results from an unnecessary procedure. For some patients, Half of health expenses are spent in the lat year of life. The procedure also involves a bad medicine that leads to an unnecessary death. It is approximated that there are 48-98K unnecessary deaths per year in the US.

The health care in the United States also performs not so efficiently. Compared to the other developed countries, the US spends much higher percentage of GDP on health care. The life expectancy of people in the US is, however, lower than other developed countries that spend less.

There have been many attempts from the past to present to control the cost of health care and make the system more efficient. These includes managed care, ObamaCare, and Hospital Readmissions Reduction Program.

References

[FAD+17] CP Friedman, NJ Allee, BC Delaney, AJ Flynn, JC Silverstein, K Sullivan, and KA Young. The science oflearning health systems: Foundations for a new journal. *Learn Health Sys*, 2017.

6.S897 / HST.956 Machine Learning for Healthcare

Spring 2019