6.S897/HST.956 Machine Learning for Healthcare

Lecture 10: Application of Machine Learning to Cardiac Imaging

Instructors: David Sontag, Peter Szolovits

1 Background

This lecture was a guest lecture by Rahul Deo, the lead investigator of the One Brave Idea project at Brigham and Women's Hospital. Rahul is also Adjunct Associate Professor at UC San Francisco and a member of the faculty at Harvard Medical School. He talked about how machine learning techniques are being used and can be used further to augment cardiac imaging.

2 Introduction to Cardiac Structure and Function

Before considering the applications of machine learning to cardiology, it's important to understand the field of cardiology a bit better. In particular, why should we care about cardiology in the first place? For one, coronary heart disease, or CHD (the hardening of arteries that transport blood to the heart), is the leading cause of death globally. And this is something that holds true for both developing and developed countries alike. Since cardiology is ultimately about biological diseases like CHD, it's a good idea to get a better understanding of the biology of the heart before moving on to look at how machine learning can disrupt the field.

2.1 Cardiac Function

The heart's primary function is to pump oxygenated blood throughout our circulatory system. The continual flow of blood is not only critical to deliver oxygen necessary for ATP (energy) production to all our tissues, but also to transport signaling molecules throughout our body and remove waste from cells. As a result, the volume of blood flow is quite large: the heart pumps 5 liters of blood every minute, a number that can grow to as much as 35 liters per minute during intense exercise. One crucial aspect of cardiac function is that the body must maintain extremely rhythmic beating of the heart, a not inconsequential task given that the average human heart generates a total of more than 2 billion heartbeats over a lifetime.

2.2 Structure of the Heart

Figure 1 above gives an overview of the structure of the human heart. Blood comes in through the superior vena cava into a chamber called the right atrium, from where it passes into the right ventricle. The right ventricle pumps blood to the lungs, and the newly oxygenated blood then flows via the pulmonary veins to the left atrium. Finally, the left ventricle pumps blood to the rest of the body via the aorta. Thus, the heart essentially conducts 2 circulations in series: a pulmonary circulation to pump deoxygenated blood to the lungs, and a systemic circulation to pump newly oxygenated blood to the rest of the body. In addition, the heart also has 4 different valves (mitral, tricuspid, aortic, pulmonary) that control flow of blood out of the heart's four chambers.

2.3 The Cardiac Cycle

As the heart pumps, it cycles between periods of relaxation called diastole, in which the heart is filled with blood, and periods of contraction called *systole*, in which the heart pumps out blood. This regular mechanical motion is coordinated by synchronized electrical activity, which can be visualized in an electrocardiogram (EKG). A Wiggers Diagram can be used to demonstrate the interconnectedness of these electrical and

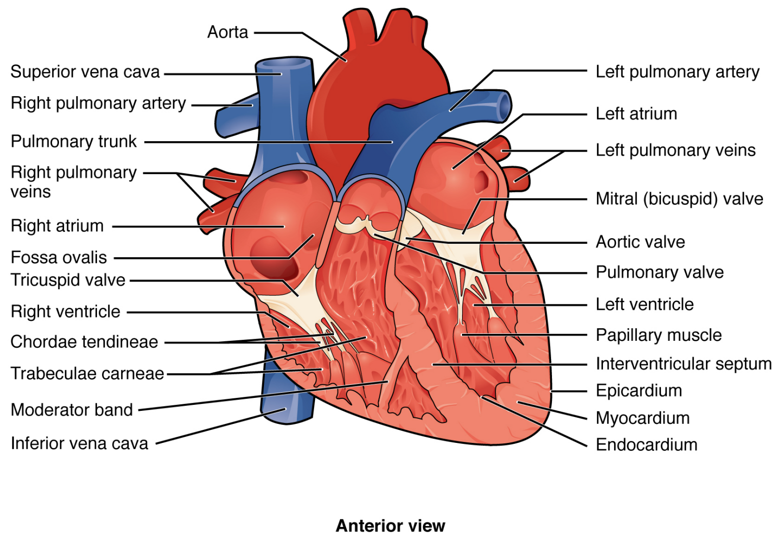

Figure 1: The major chambers, valves, and blood vessels of the human heart.

mechanical systems, allowing one to see how events in an EKG align with the physical state of the heart, as shown in Figure 2.

2.4 Cardiac Diseases

Given the complex structure of the heart, cardiac diseases are organized based on abnormalities or failures in the following different functions (with example diseases in parentheses):

- **Contractile Function** (heart failure)
- **Coronary Blood Supply** (coronary artery disease, myocardial infarction)
- **Circulatory Flow** (aortic or mitral stenosis/regurgitation)
- **Heart Rhythm** (atrial fibrillation, ventricular tachycardia)

2.5 Wrapup of Cardiac Biology

So far, our discussion of the heart has focused on examining it as a muscular organ responsible for pumping blood. However, there is a ton of biology here apart from pumping - in fact, only 31% of of cardiac cells are cardiomyocytes, the muscle cells responsible for the heart's contractions. As a result, cardiac function (and, as a result, cardiac disease) comes from the interaction of a very diverse and large group of cells, ranging from endothelial cells, fibroblasts, leukocytes, and more.

3 Major Types of Cardiac Diagnostics and How They are Used

Cardiology is by its nature extremely imaging-centric, making it an expensive field of medical study. There exist a large variety of various imaging techniques that each play critical roles in diagnosis. Here is a brief overview of some of the most important ones:

- **EKG** - An extremely cheap technique based on measuring voltage differences in the heart over time. Can be used, for example, to diagnose myocardial infarction.

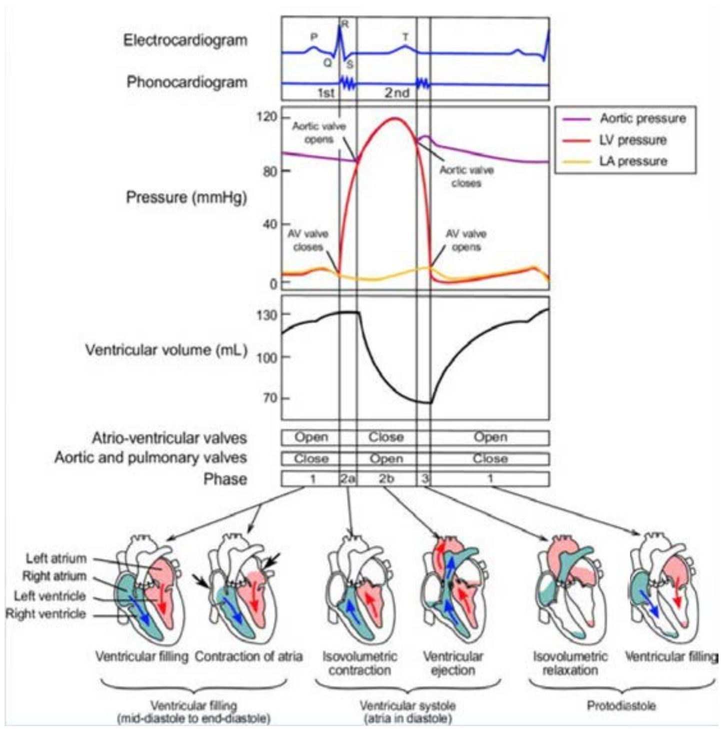

Figure 2: An example of a Wiggers Diagram.

- **Echocardiography -** A more expensive technique that uses ultrasound tech to make measurements. Can be used to get general understanding of cardiac structure and thus diagnose, for example, heart failure, valvular disease, and pulmonary hypertension.
- **MRI -** One of the most expensive imaging techniques. More expensive than echocardiography but achieves very similar diagnostic functions, so it is not used all that much in the US.
- **SPECT/PET -** Non-invasive techniques, but also extremely expensive. Can be used, for example, to infer coronary artery disease or diagnose microvascular disease.

One interesting characteristic of cardiac diagnostics is that diseases are not defined based on biology, but rather based on measurements that depart from ”normal” anatomic or phsiological values. This may be due to the limitations in our ability to study or image the human heart, but it’s not clear whether this is ultimately for the better or worse.

In cardiology, clinical decisions regarding treatments for diagnosed diseases are often, but not always, guided in part by inputs from cardiac imaging. Ultimately, a few different factors affect the role of imaging in cardiology. First, although imaging can lead to data with extremely high information content, decisions regarding treatment are affected in large part by historical studies that follow patients with the disease of interest over long time periods. Secondly, cardiologists are often stuck with the data that is already out there because someone decided it was worth paying for. The available data often controls the risk model and decision analysis you can undertake. Finally, while imaging data can be found for patients with diseases for which the imaging process is seen as an essential part of the commonly accepted management plan, it is much more difficult to obtain imaging data for other diseases or patient populations given the high cost. As a result, doctors are often stuck with the stuff that they already know something about.

4 Where's the Data?

4.1 How is Medical Imaging Data Stored

DICOM is the major international standard for storing imaging information. Image/video files are stored in a compressed DICOM format, which includes a header that contains info or characteristics about the image. Several free, open access libraries or software exist for compressing/uncompressing DICOM files, viewing the corresponding images, and reading or editing the header. Most imaging data data can be found stored in data archives.

4.2 Gaining Access to Data

Ultimately, access to imaging data can be quite limited due to the following reasons:

- Some images have burned in pixels with patient PII
- Vendors don't necessarily make it easy for users to download or de-identificate data, perhaps due to their motivation to make it difficult for users to switch vendors
- Some systems have monetized the data access pipelines, making it almost prohibitively expensive to access imaging data

Another obstacle is that labels (like physiological measurements or diagnoses) are, in many cases, stored separately in electronic health record data, forcing researchers to look for data across multiple sources.

One clear trend related to imaging data is that, as the cost of the study increases and/or the perceived utility of the data decreases, the availability of data goes down. As an example, an imaging technique like PET that is very expensive has only 8000 studies available at Brigham and Women's Hospital, whereas over 30 million EKGs can be accessed.

4.3 Characteristics of Medical Imaging Data

One of the major issues with obtaining high quality cardiac images is that, as the patient breathes, the chest wall and the heart are both continuously moving. Thus, high quality scans need to get enough temporal frequency on their data acquisition so that the movement of the heart doesn't affect the imaging. Another solution to image corruption resulting from cardiac motion can be found in the technique known as gating, in which the cardiologist lines up corresponding portions of different heartbeats from EKG data with images, allowing him or her to average the images to obtain a lower noise measurement. A summary of the characteristics of various cardiac imaging techniques can be found in Figure 3.

5 Computer Vision Topics Relevant to Cardiac Imaging

The major question to ask when trying to apply machine learning techniques to cardiology is: what physician practices can we mimic? Currently, all cardiac measurements, ranging from computing volumes of cardiac chambers to measuring ventricular thickness, are performed manually by hand. Moreover, some disease diagnoses require cardiologists manually classifying images or videos. Fortunately, many current priorities in CV are of great interest to cardiac imaging as a result.

5.1 Image Classification

In image classification, the goal is to assign a label to a given image or video. This is a ripe candidate for applying supervised machine learning techniques to cardiology. There are many simple disease recognition tasks in medicine such as identifying lung cancer, pneumonia, or breast cancer, although physicians are

Modality	Spatial Resolution	Temporal resolution
Echocardiography	2-3 mm	1-5 ms for some modes; typically 20-30ms for 2D
MRI	0.1mm	30-100 ms
Angiography (Fluoroscopy)	0.1mm	1-10 ms
Computed Tomography	0.5mm in x,y; 0.5-0.625mm in z	65-175 ms
SPECT/PET	8-10 mm for SPECT; 3-5 mm for PET	minutes

Figure 3: Comparison of Various Imaging Techniques.

generally really fast at most of these, with an experience radiologist capable of diagnosis disease based on images in less than 2 minutes.

Most of the initial successes in medical image classification have come in contexts where there are huge datasets that are already labeled - common examples include chest X-rays or mammograms, images that are collected as part of routine clinical care. Unfortunately, in many other contexts, there are bigger barriers to sharing or exporting data which has limited the size of the datasets. From a machine learning standpoint, prior to the emergence of Convolutional Neural Networks (CNNs), researchers were not doing anything in the image classification space, but recently there's been an upsurge in interest with lots of companies thinking about it. Advances like representation learning, no longer needing to handcraft features, and transfer learning (which is especially useful for training models in data-poor scenarios) have all contributed to this growing interest.

5.1.1 The Case against Automation

One challenge to automated image classification arises when one considers the question, how much benefit is there really to automating a process that takes a radiologist 2 minutes at most? Investing resources and engineering time into a task where the benefits of automation are slim seems both unnecessary and wasteful. Compounding this challenge is the fact that there's an enormous amount of liability involved with medical image classification, with radiologist being the most sued profession in medicine. This liability means that radiologists don't feel sufficiently convinced to pass the task off to a black box computer system.

Nevertheless, applications of automated image classification still exist - just not in the idealized case of a machine independently reading a patient study. The real benefits can be found in cases like triage, when a machine can analyze data to decide the highest priority or most urgent items to look at. In general, resource poor settings can skew the decision-making calculus in the favor of automation. And adding an independent matching reading of studies should, at the very least, catch some missed diagnoses.

5.1.2 Explaining the Diagnosis

Another challenge arises in the tension between predictive accuracy and descriptive accuracy. As a general rule, medicine is very demanding on descriptive accuracy while simultaneously being inflexible on predictive accuracy. As a result, in almost every case, medical decisions based on automated imaging technology have also required the support of a corresponding human confirmation. The growing use of high-dimensional, deep

CNNs have also raised concerns as to whether an explanation can even be feasibly given for an imaging-related medical decision.

The demand to provide interpretability in medicine has spurred several different research strategies in theoretical machine learning. Two ends of the spectrum can be found in:

- **Class model visualization** - In this technique, an exemplary representative image of each class is chosen by generating an image that maximizes the regularized class score of the network.

- **Image-specific class-specific saliency map** - In this technique, the derivative of the network score function for a given class is taken with respect to every single pixel and plotted in a saliency map to identify the pixels that maximally activate the given class.

5.2 Semantic Segmentation

In semantic segmentation, the goal is to assign each pixel of the image with a class label. For example, one common task in cardiology is delineating the boundaries of the heart in an image, and radiology reports often require measurements of basic structures (such as lengths or areas). In order to compute metrics like this, cardiologists are often stuck performing manual tasks like tracing lines or circles on images. As an example, a radiologist might need to manually draw a line between two ends of a structure in an image in order for the machine to measure the distance between them based on the acquired data.

From a machine learning perspective, once again research has propelled several critical advancements in this area. A certain architecture known as U-Net seems to be the favorite for the segmentation task - in fact, almost all the literature out there seems to converge to it [RFB15]. A common problem with semantic segmentation is focusing on relatively limited scales, so that classification occurs at the pixel level, although many researchers have worked on projects addressing this.

5.3 Image registration

Finally, in image registration the goal is to align or merge different images. As an example use case, a certain cardiological study may have poor spatial resolution, while another may lack some necessary functional information - in this case, merging the images produced by the two studies results in higher quality, useful images. Another use cases arises with techniques that have limitations in temporal resolution such as PET, in which case one must average images across many cardiac cycles (i.e. gating). Conditional variational autoencoders have shown to be a particularly well suited model for this task by learning geometric transformations between pairs of images.

6 A fully automated pipeline for echocardiogram interpretation

6.1 Motivation for a fully automated pipeline

Cardiovascular disease are the number 1 cause of death globally [Org17]. The current approach to cardiovascular diseases is lacking. People often delay treatment for a very long time, instead opting to see if lifestyle changes can remedy the first signs of heart disease (e.g., abnormal blood pressure, cholesterol, and blood sugar). Thus, by the time they start treatment they are far too a long the disease timeline leading to costly treatments that may not be very effective. We see that patients who die from cardiovascular diseases, die shortly after developing symptoms, such as dyspnea and angina. See Figure 4.

6.2 Goals of the fully automated pipeline

1. Low-cost quantitative metrics that are indicative of disease progression and reflect onset of tissue-level changes.

2. Specific to the disease process

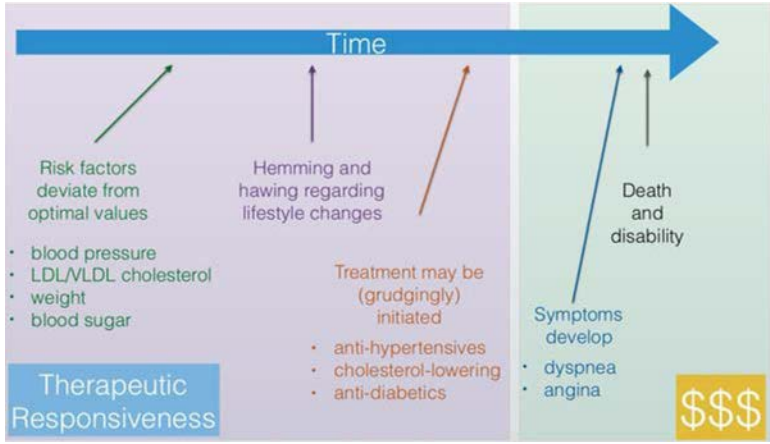

Figure 4: A rough timeline on patients with cardiovascular disease.

(a) Expressive - captures complex underlying processes (molecular, cellular, imaging ...)

(b) Multidimensional - Can't readily be "gamed"

3. Pipeline should be ameliorated with therapy (c.f. genetic risk)

The role for the fully automated pipeline for echocardiogram interpretation would be at the "low risk - high reward" portion of the current spectrum. See Figure 5

6.3 Focus of machine learning in cardiac diseases

We can use machine learning to:

1. enable much greater of volumes of data to be interpreted, so that we reduce costs of acquisition and interpretation, as well as augment interpretations of simple data.
2. augment surveillance within a hospital system, e.g. patient identification for therapies
3. perform triage, i.e. automating ECG interpretation in urgent situations in the ambulance/ER

To illustrate the need of machine learning to perform rapid triage, we provide an example. In the early 2000's, it was recognized that any delay in angioplasty and stenting would result in irreversible damage to the heart. Thus, the solution was to replace a cardiologist reviewing the ECG with a rapid triage system by ambulance personnel or ED physicians for quicker turnover. However, this resulted in an increase in false positives. Thus, there is still a need for a fast pipeline that has high quality.

6.4 Zhang, Deo, et al. approach to Automated Approach for Echo interpretation

An echo study is typically a collection of up to 70 videos of the heart taken over multiple cardiac cycles and focusing on different viewpoints. The heart is visualized from ¿10 different views, and still images are typically included to enable manual measurements. ¿7,000,000 echo studies are performed annually in Medicare population alone, and there are likely (an estimate by Deo) of about 100,000,000's of archived echo studies.

Zhang et al. built a pipeline using 14k raw echo studies and traditional computer vision algorithms for view classification and segmentation into 5 views to perform cardiac structural and functional analysis. This is an automated and low cost approach to echo interpretation. See Figure 6.

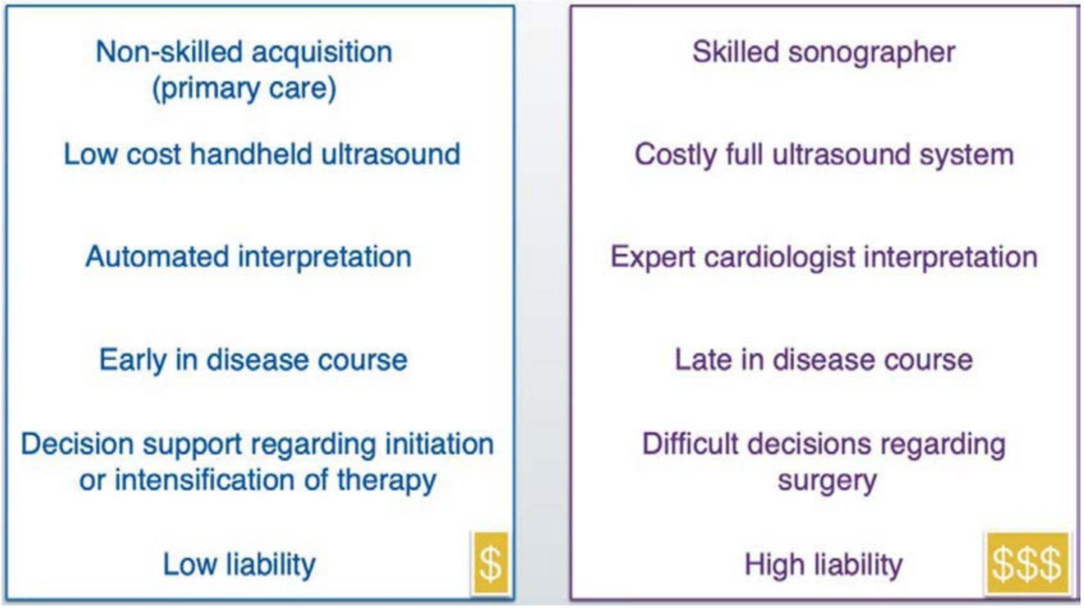

Figure 5: The left side shows the benefits of having an automated pipeline for echocardiogram interpretation in contrast to the right side without it. We see that both niches fulfill high reward, but the left is low risk and the right is high risk. For early stages, it would be beneficial to use the automated pipeline for low liability, low cost, and quick decisions for whether further analysis is needed. For late stages, it would be better to use a more trusted skilled sonographer, more expensive ultrasound equipment, and an expert interpretation.

6.5 Purpose of automated disease detection

Several rare diseases (e.g. mitral valve prolapse) would benefit from referral to cardiologist or specialty centers

Diagnoses tend to be missed at centers that see them infrequently (e.g. cardiac amyloidosis)

Automated disease detection provides another pillar of support for definitive diagnoses

7 Rethinking the future of automated interpretation: lessons

Deo's predictions for the future of cardiac imaging:

1. Routine measurements will be made in an automated way
2. Some automated diagnoses may happen at point-of-care, e.g. heart function and fluid accumulation around heart
3. Until image data acquisition is facilitated, the benefits of automated interpretation will be muted
4. Pharmaceutical companies have high motivation to perform high frequency serial imaging to assess whether there are any benefits to medications in clinical trials, and an accurate scalable quantification will be needed for this.
5. Surveillance of daily studies may be useful to enable identification of individuals who may be eligible for clincal trials or newly approved therapies.

Deo's uncertainties on automated interpretation:

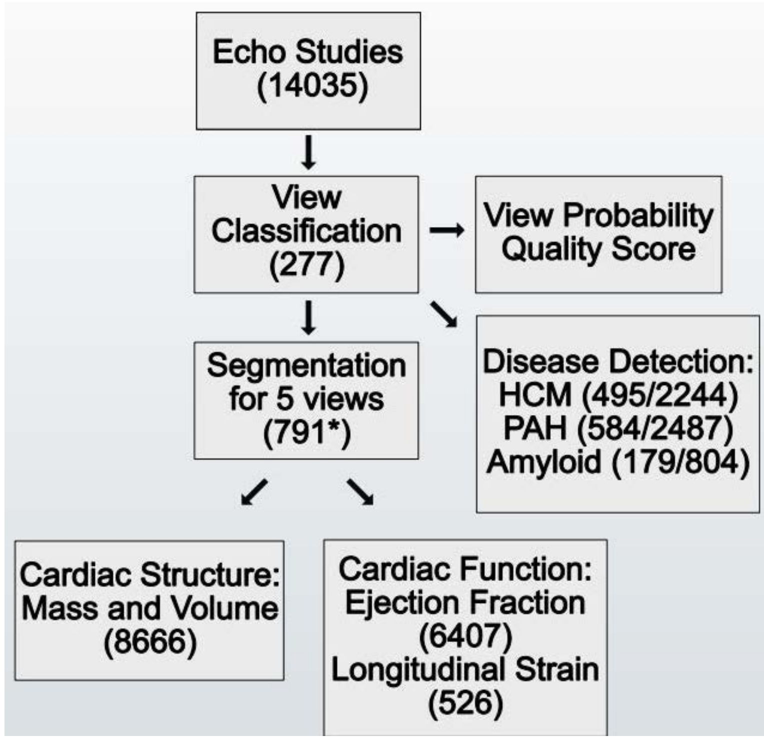

Figure 6: Zhang, Deo, et al.'s [TZDD18] approach to an automated pipeline for echocardiogram interpretation.

1. We should be using automated interpretation to elevate medicine beyond the current practice, but we need larger dataset and more images than what we currently have.
2. Disease classifications are currently crude and finer distinctions can be made between disease states.
3. Survival models are crude and better predictive models should be possible with imaging data and emerging algorithms.
4. Physicians are only interested in classifications or risk models that will change and improve practice, thus evidence is required to justify a shift.
5. There is the question of how more data will be obtained and dispersed for research.

8 Biology

There are some goals in biology that can be accomplished to facilitate cardiac machine learning. First, clinical datasets lack the scale and expressivity needed to reflect underlying biological processes. We need a data type that has the dimensionality to capture biological heterogeneity and complexity and yet can still be collected in a very scaleable manner. It is likely that we shouldn't look at expensive sequencing technologies and costly medical imaging to accomplish this. Accomplishing this would help expand the biological phenotypic space.

Another thing is that we should focus on studying individual circulating blood cells. Circulating blood cells are causally implicated in coronary heart disease (CHD) pathogensis. These blood cells are also easily accessible in a blood draw, and there are existing predictive models that use white and red blood cell characteristics for coronary artery disease. They also reflect many pathways found in diverse cell types, such as autophagy, phagocytosis, and free radical dissipation.

Third, we should foucs on cell morphology rather than genomics. This takes advantage of the computer vision advances that are able to characterize subtle distinctions between cell types and states at low cost. We can analyze tens of thousands of individual cells per patient through these methods. Fluorescent dyes can also be used to make it easier to distinguish parts of cell morphology that help us take advantage of computer vision advances more. Furthermore, it is known that cell morphology is connected to gene expression.

A summary of Deo's research approach:

1. A permissive recruitment scheme to enable rapid accrual of tens of thousands of patients per year all with expressive phenotyping and full medical records
2. Use of cell morphology/cell coutner data to massively expand phenotypic space at low cost using perturbations and diverse readouts
3. Overlapping of multiple phenotypic scales in dfiferent cohorts to convert costly, tissue-localized phenotypes (e.g. PET, CHIP sequencing) into lower cost (TTE, cell imaging) models.
4. API-based cohort identification to allow rapid identification of patients of interest.
5. Automated curation of the medical record into a vehicle for machine learning and causal inference.

9 Questions and Answers

Question 1: What should resource poor countries use for imaging? And would this imaging be useful to apply computer vision algorithms to?

Answer 1: There are cheap, portable ultrasounds which you can get for roughly $2000 nowadays and the quality of the ultrasound images is comparable enough to allow for application of modern computer vision algorithms.

Question 2: Usually people begin treatment after visiting the doctor for the first time. How do you trust the one visit when you go to the doctor to determine if you should go on medication or not?

Answer 2: There are noisy point estimates, and it's hard to determine precisely whether the timing is right (usually it is though).

Question 3: Cardiac imaging has a long history, where there were active modelers of morphologies of the heart when image scanners were coarse. These modelers of morphologies used geometric priors for reconstruction, are these geometric priors being reintroduced in some way in modern times?

Answer 3: This is not something that is widely-used anymore, and the data to do this is also unavailable.

References

[Org17] World Health Organization. Cardiovascular diseases (cvds), May 2017.

[RFB15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[TZDD18] Geoffrey H. Tison, Jeffrey Zhang, Francesca N. Delling, and Rahul C. Deo. Automated and interpretable patient ecg profiles for disease detection, tracking, and discovery, 2018.