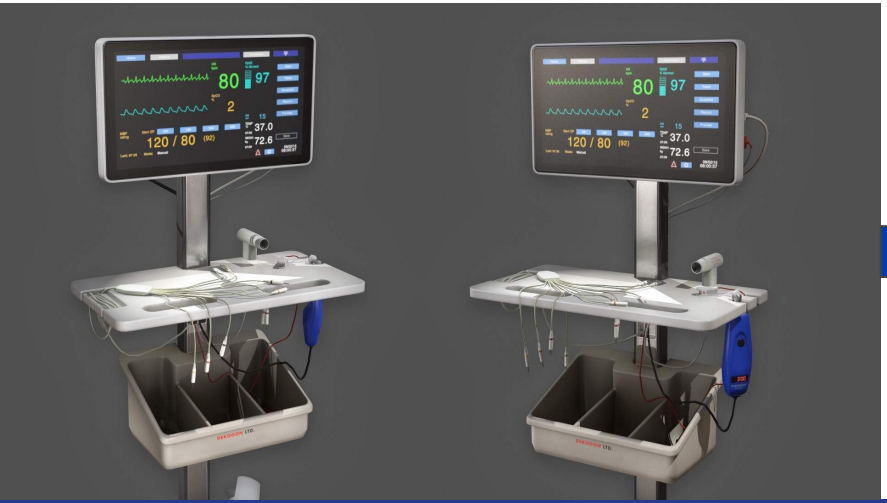
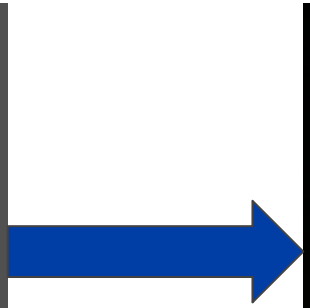
Case Study: Disease Detection and Diagnosis with ML

ter.ps/389iweek4

Apple

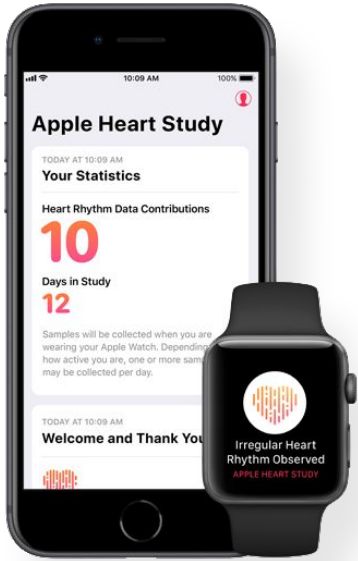
- Apple Watch Series 4
- FDA Approved (Class II) EKG (electrocardiogram)
- Third party EKG watch straps (alivecor)





Stanford Medicine Heart Study

- Tracks Atrial fibrillation
- Open enrollment to anyone with Apple Watch & Iphone 5s+
- Around 600 total participants
- November 2017 - January 2019
- https://med.stanford.edu/appleheartstudy.html



Atrial fibrillation

- Irregular Heartbeat
- 25% of all individuals experience this over the course of their lifetime[1]
- Many symptoms & potential causes
- High blood pressure, heart failure
- Early warning sign for other diseases

[1] Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. Circ Res. 2014;114(9):1453-1468.

Challenges

- Bias (Participants are educated, tech savy, live in the U.S)
- Small sample size (600) vs millions who will buy a device
- Short duration (15 months) - long term health outcome impacts unclear
- High variability in patient behavior

Viz.ai

- Deep learning to detect large vessel occlusions (LVOs) on CTA imaging
- Alerts doctors instantly

- Rapid stroke detection
  - Life threatening, time sensitive
- Analyzes brain scan images
- Kleiner Perkins, GV

Grail

- Raised **> $1.5 Billion** for early cancer detection research
- Faint signals of ctDNA in the blood
- High intensity sequencing → ~1 TB of data / individual
- Using ML (e.g. hierarchical neural networks)
  - classify patients by **presence**, **type**, and **severity** of cancer
- Eventual goal to open source their datasets
- Two major public studies:
  - Circulating Cell-Free Genome Atlas Study (CCGA)
  - STRIVE (Breast Cancer)
  - Link

GRAIL

Scale

- Larger sample size: potential for more representative data
- Circulating Cell-Free Genome Atlas Study (CCGA)
  - 15,000 individuals for DNA Genome analysis
  - Context: HGP took 10+ years and $3 Billion in funding
  - Cancer + healthy mutations from cFDNA of white blood cell genomes
- STRIVE (Breast Cancer)
  - 120,000 women over course of 5 years
  - Largest long term study for breast cancer

Freenome

- Cell-free biomarker patterns
  - Fragments of DNA, RNA, and protein
  - Real-time snapshot of cancer’s activity
- AI genomics platform to learn biomarker patterns
  - cancer stage
  - cancer type
  - most effective treatment pathways
- Predict response to (immuno)therapy
  - personalized medicine
- Colon cancer

Challenge with Liquid Biopsies

- Expensive to obtain samples
- Processing data is complex and time consuming
- Pressure to publish positive results
- Very little cfDNA compared to other DNA in the blood!

- But if it works, it would allow:
  - real-time monitoring of tumor
  - MUCH cheaper than imaging e.g. PET-MRIs, thus more frequent
  - catching cancer when it is curable

Put into Practice: Breast Cancer

Breast Cancer

Early Warning Signs

https://www.webmd.com/breast-cancer/understanding-breast-cancer-symptoms

- Physical factors such as size and shape
- Color and differences in texture
- Changes underneath the skin, discharge

Examining the Dataset

- What Source is the data from?
- What information does the data contain?
- Dataset

Further Reading

- https://www.theverge.com/2018/9/12/17850660/apple-watch-series-4-ekg-electrocardiogram-health-2018
- https://qz.com/1389202/the-apple-watch-series-4-health-features-arent-all-that-impressive/
- https://www.freenome.com/blog-freenome/2018/6/19/how-blood-tests-can-be-used-to-detect-cancer?p
- https://www.linkedin.com/pulse/what-you-need-know-ai-cancer-paul-walsh/
- https://www.freenome.com/blog-freenome/2018/5/30/plmgmyg0j8er7mz7gtige0tt5lx69ep?p
- https://www.the-scientist.com/news-opinion/artificial-intelligence-to-boost-liquid-biopsies-64380
- https://openreview.net/forum?id=H1DkN7ZCZ
- https://drive.google.com/file/d/0B1T58bZ5vYa-QlR0QlJTa2dPWVk/view