

# Can Machine Intelligence turn us into healthier people?

Rise of Digital Signals and their  
Potential for the eHealth Industry

**Dr. Stefan Ebener,**  
Manager CE Specialists,  
Machine Learning - EMEA

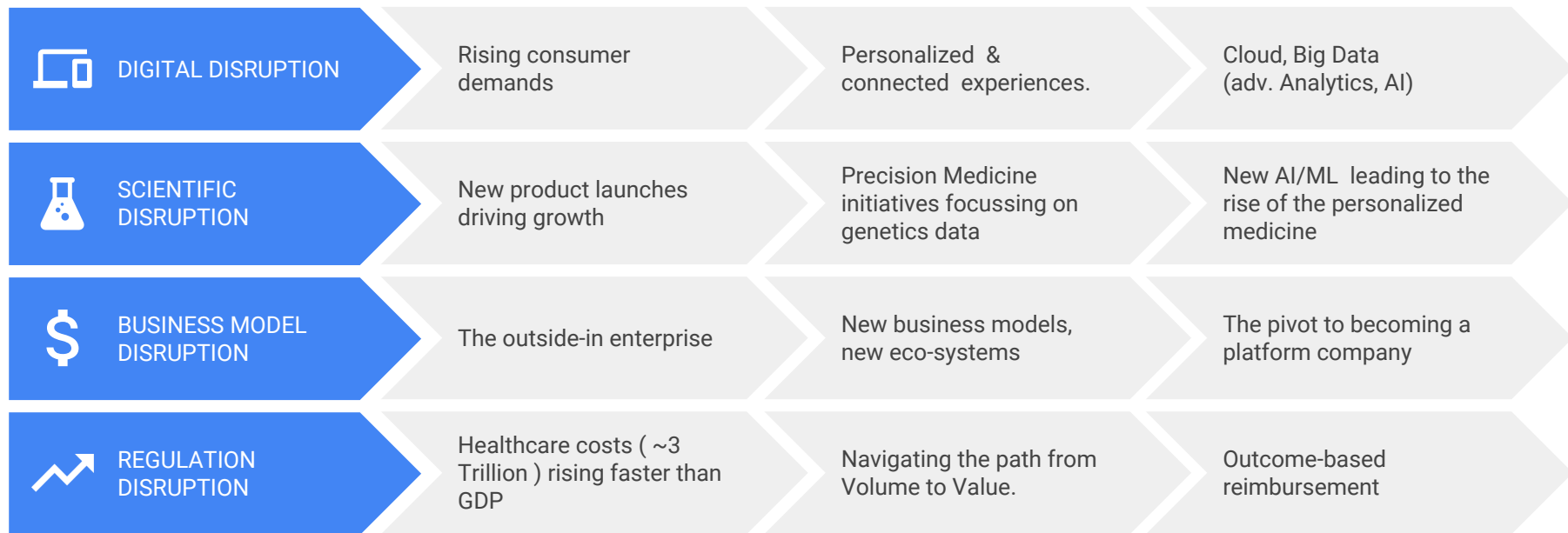
February 2019



GOOGLE CLOUD



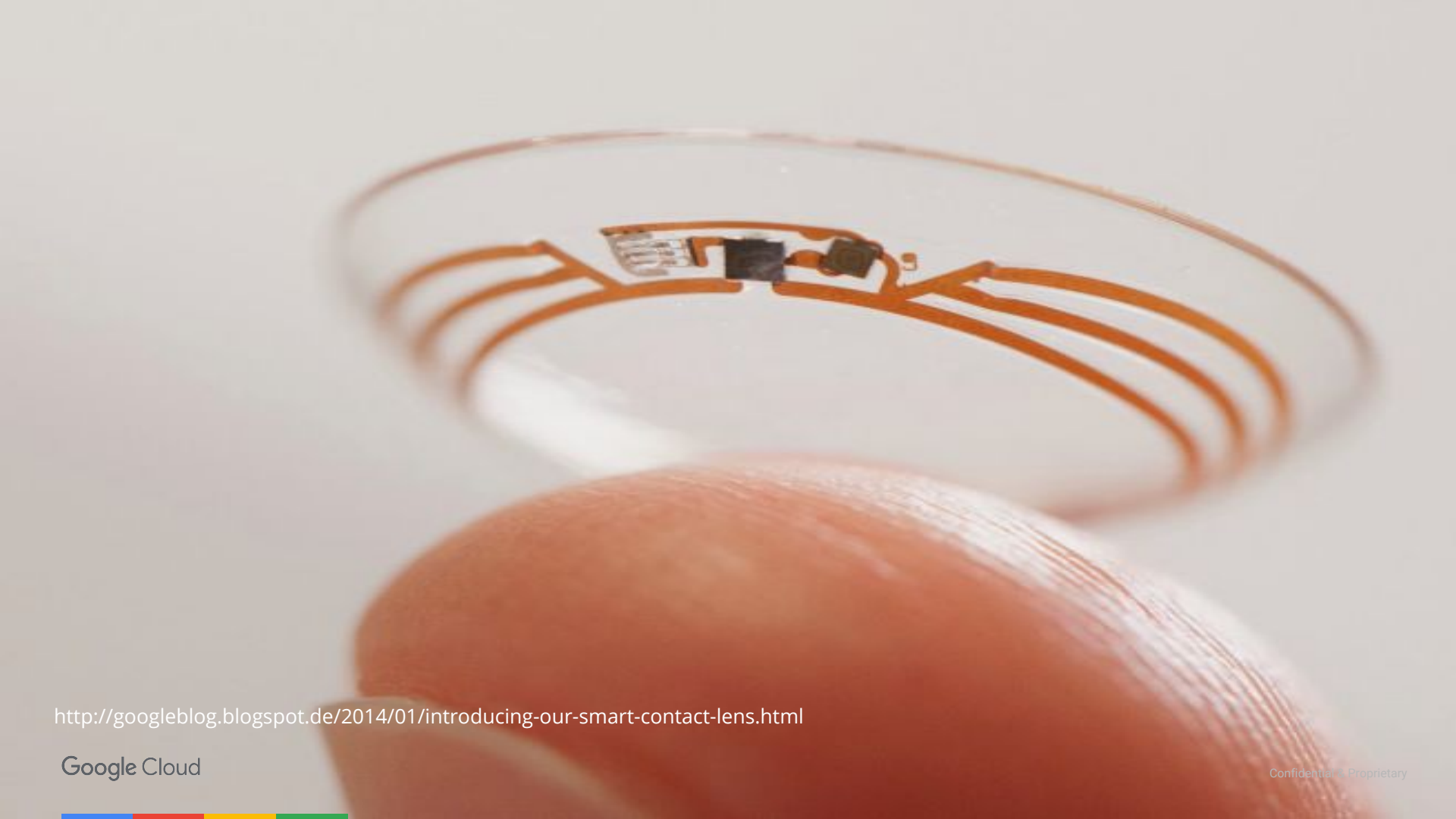
# Healthcare industry's unique challenges



# Moonshot Thinking

A photograph of a Space Shuttle Columbia launching from the launchpad. The shuttle is white with a red nose cone and is surrounded by a massive plume of white smoke and fire. The launchpad structure is visible to the left of the shuttle. The sky is a clear blue.

10 X improvements, not 10%



<http://googleblog.blogspot.de/2014/01/introducing-our-smart-contact-lens.html>

Google Cloud

Confidential & Proprietary





Pick it up



Start eating



Spill less

**LIFT**ware



## Multiple Attachments

Fork, soup spoon, keyholder, and more are coming soon.



## Stabilizing Technology

Cutting-edge electronics will work to actively detect and stabilize your tremor.

## Battery

Ultra-thin rechargeable battery will last for several days on a charge.

[google.com/liftware](https://google.com/liftware)

# Moonshot Thinking

MOST INNOVATIVE COMPANIES

162 SHARES

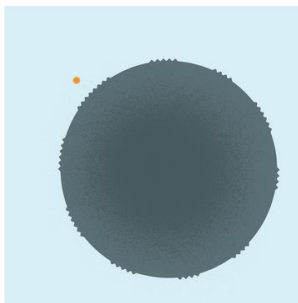


RE

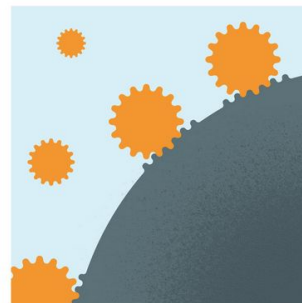
## GOOGLE X MOONSHOT WANTS TO MAP A PERFECTLY HEALTHY HUMAN BODY

**BASELINE STUDY, A NEW GOOGLE X PROJECT, W  
HIDDEN IN OUR BIOLOGY.**

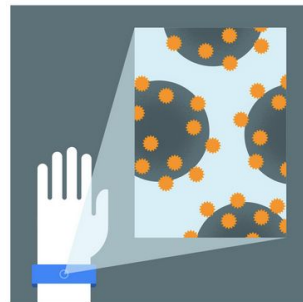
### Detecting disease early with nanoparticles



Nanoparticles are really small: more than 2000 nanoparticles could fit inside a red blood cell.



Nanoparticles circulate in the blood and can be built to attach to particular types of cells, such as circulating cancer cells.



A device worn on the outside of the body can detect the nanoparticles and provide useful information to physicians.

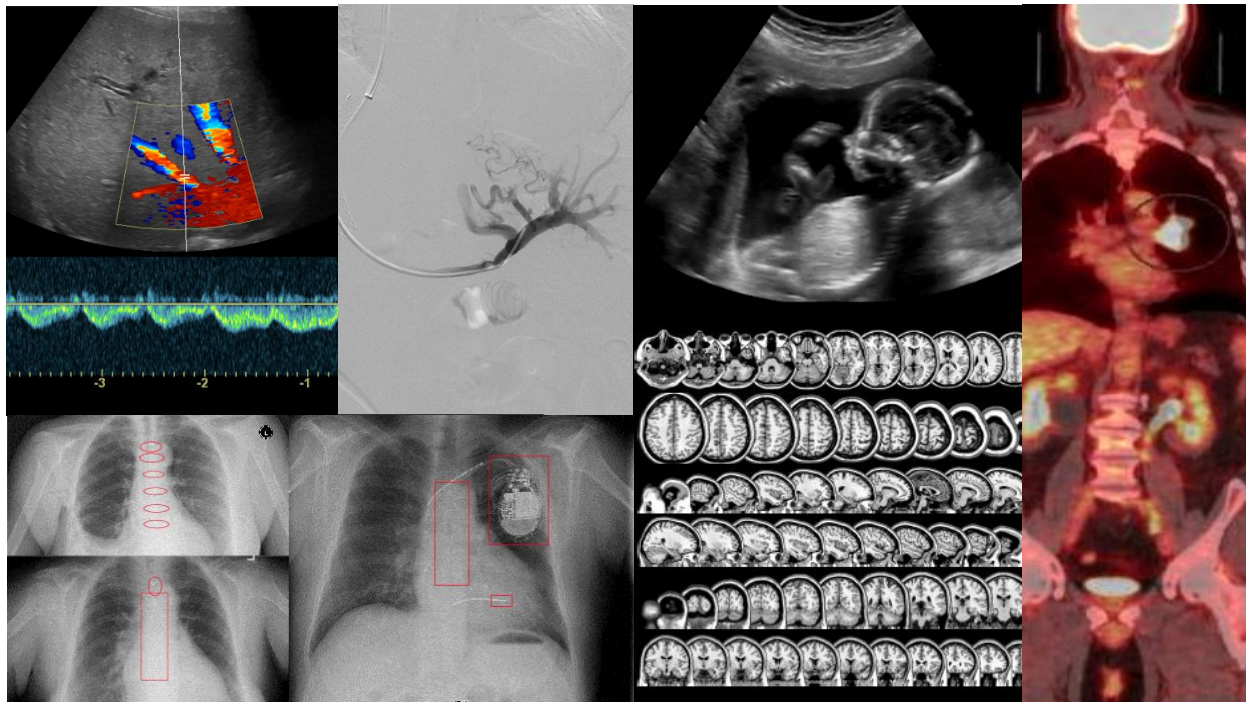
# The Power of Medical Imaging

Google Cloud





# Medical Imaging: Radiology

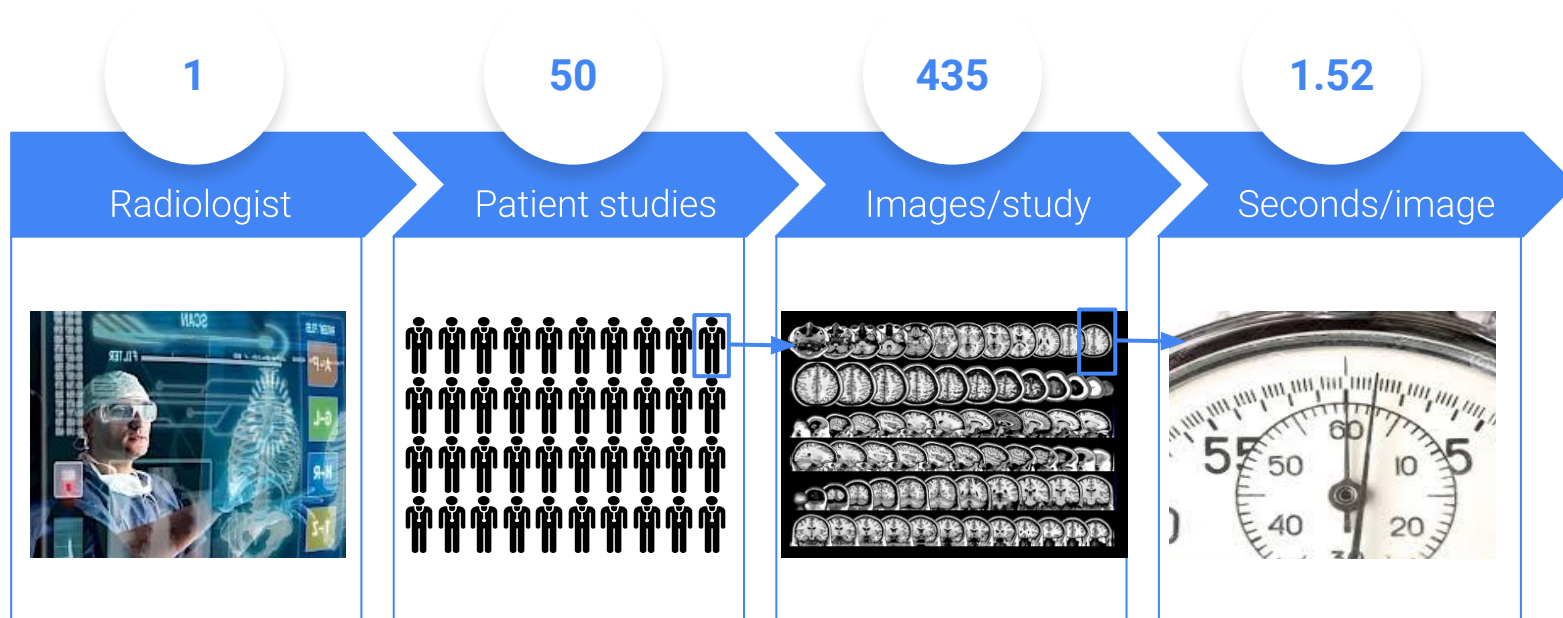




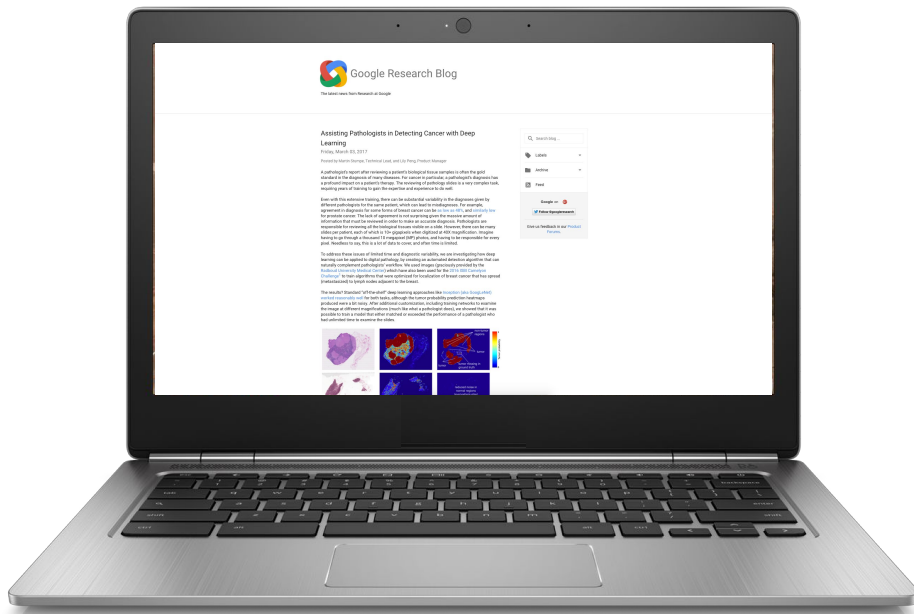
# How big is medical imaging data?



# Diagnostic radiology at human limits



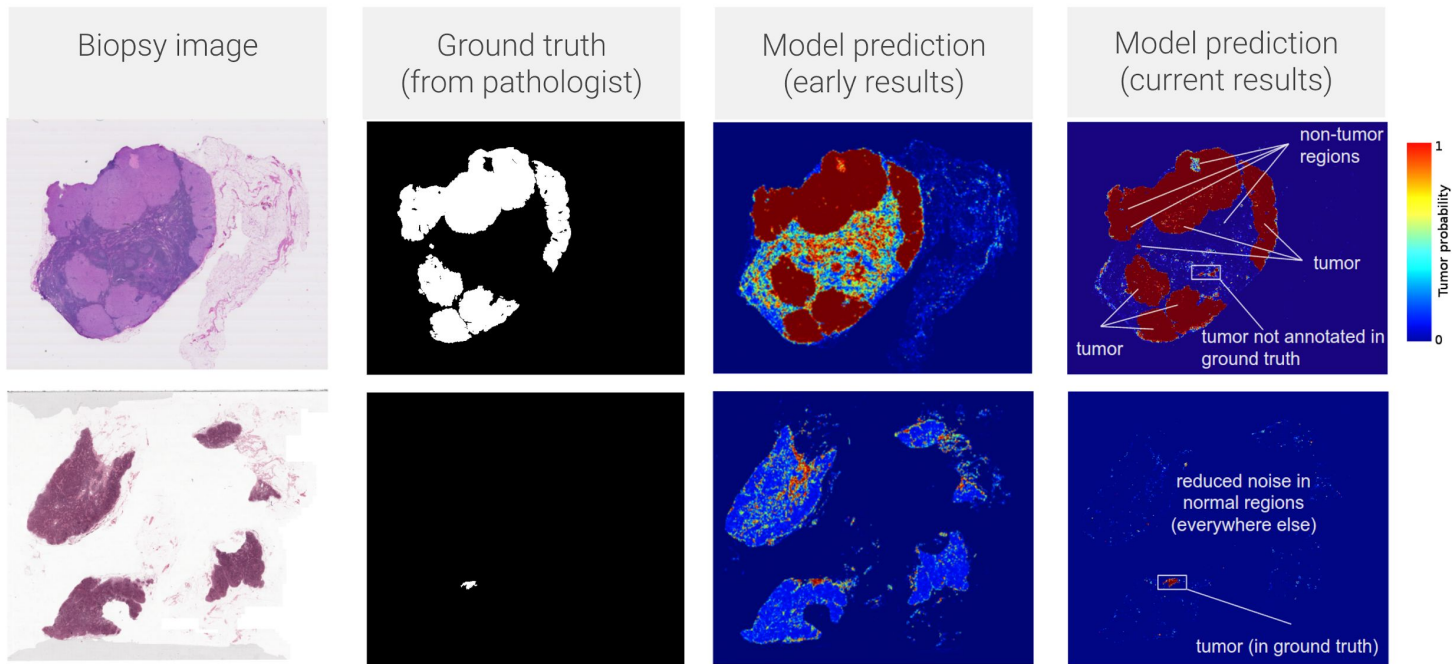
# Detecting Cancer Metastases on Gigapixel Pathology Images



Blog:  
<https://research.googleblog.com/2017/03/assisting-pathologists-in-detecting.html>

Paper:  
<https://arxiv.org/abs/1703.02442>

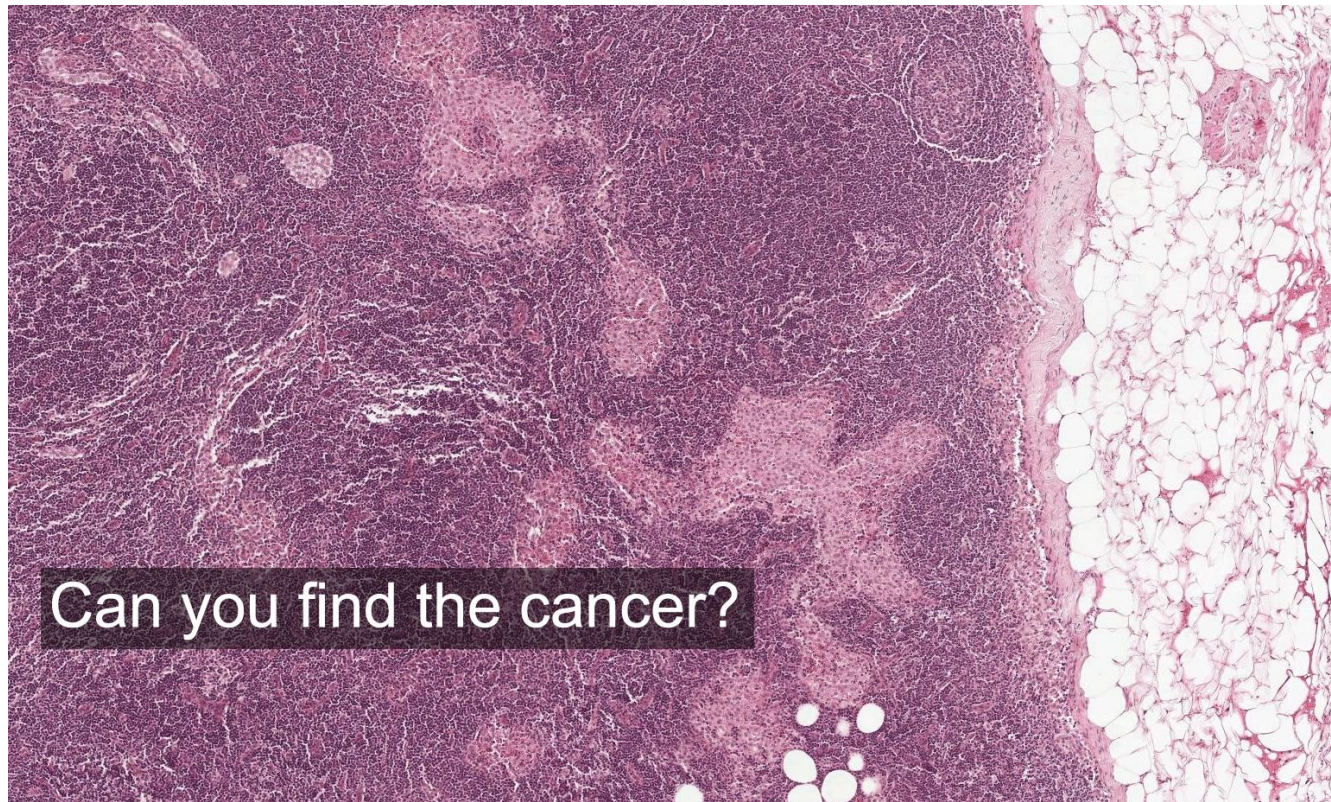
# Detecting breast cancer metastases in lymph nodes



Gulshan & Peng et al. JAMA 2016

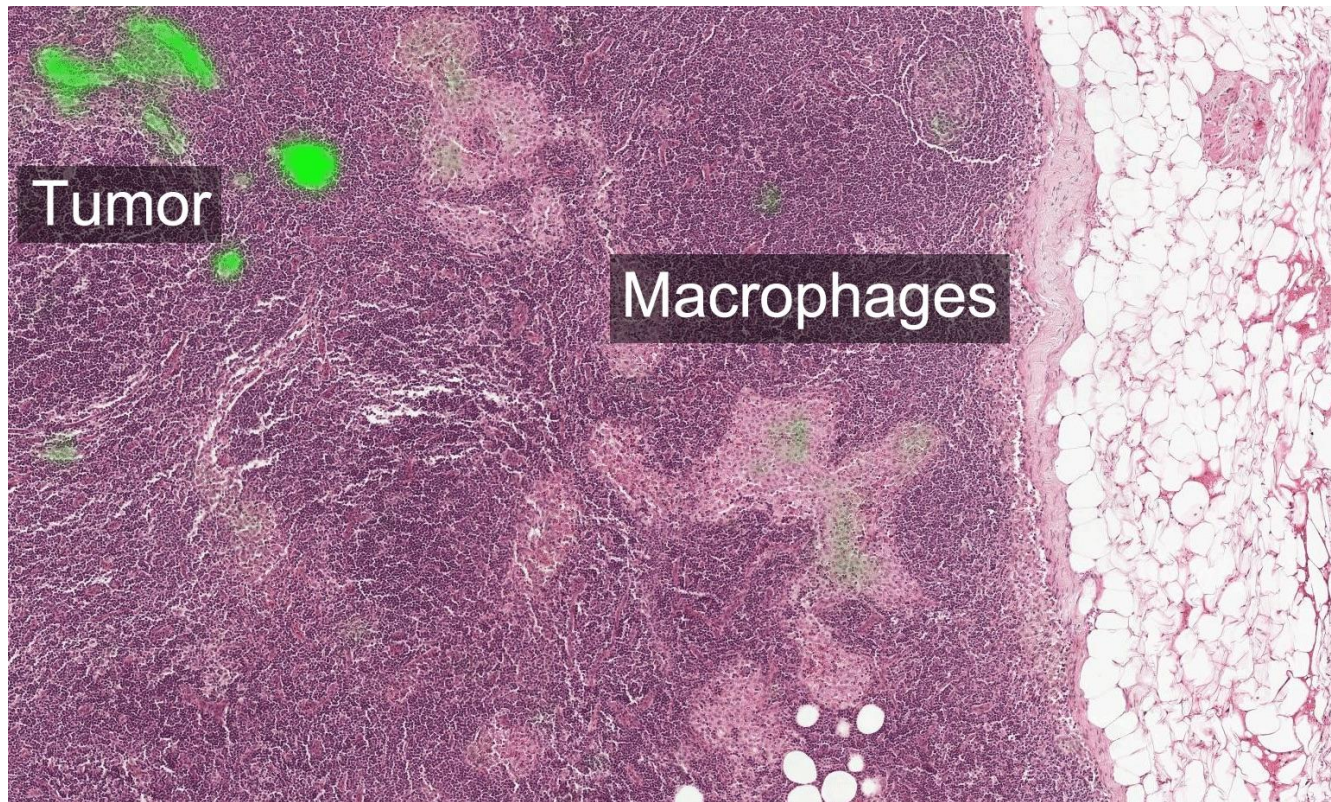


# Preliminary results in digital pathology





# Preliminary results in digital pathology



# Model performance compared to pathologist

	Our model	Pathologist*
Tumor localization score (FROC)	0.89	0.73
Sensitivity at 8 FP	0.92	0.73
Slide classification (AUC)	0.97	0.96

\* pathologist given infinite time per image (reaching 0 FPs)

Evaluated using Camelyon16 dataset **(just 270 training examples!)**

# Medical imaging global perspective: Scarcity and error



~**Two thirds** of world's population lack access to medical imaging (WHO)



> **43 million** people world affected by medical error annually, greatest burden in low income countries



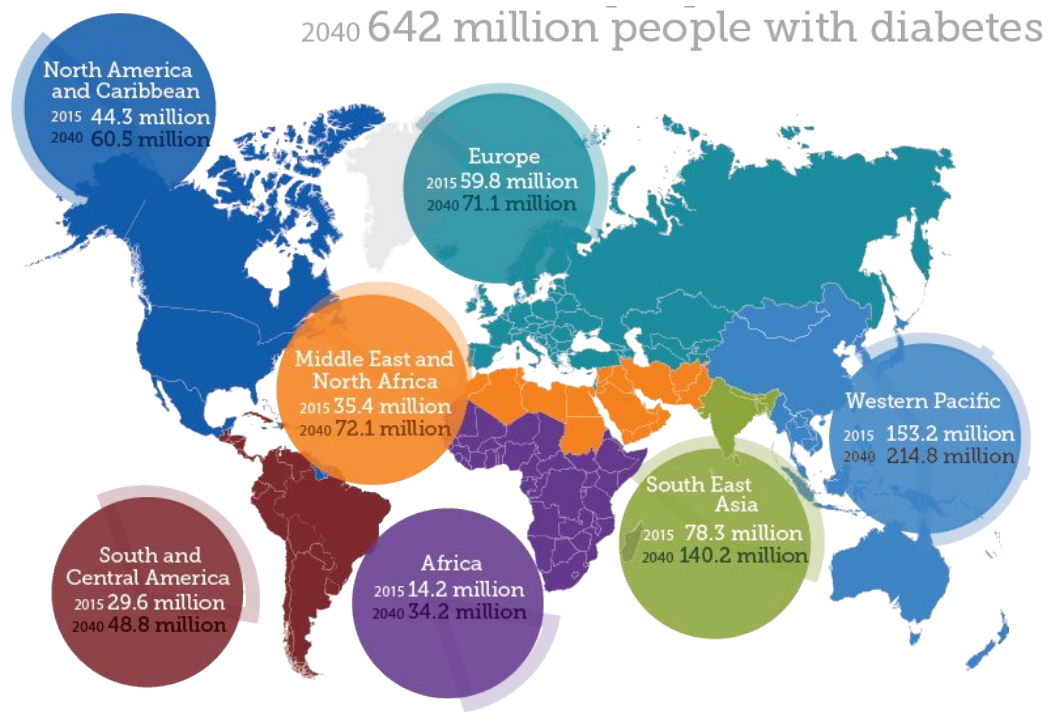


## INDIA

Shortage of **127,000** eye doctors

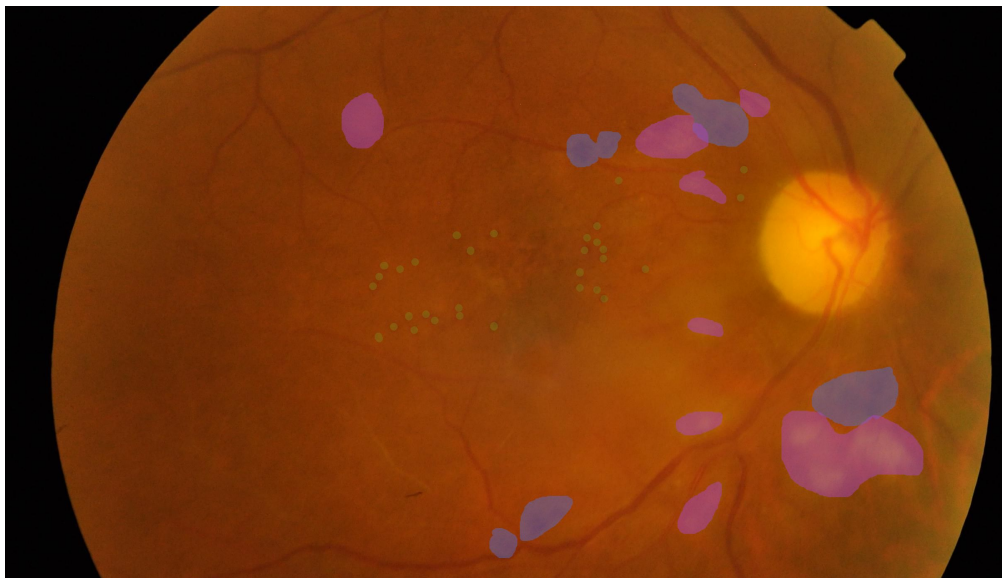
**45%** of patients suffer vision loss before diagnosis

# Diabetic retinopathy: Fastest growing cause of blindness





# Google's first medical imaging project: **Diabetic retinopathy**



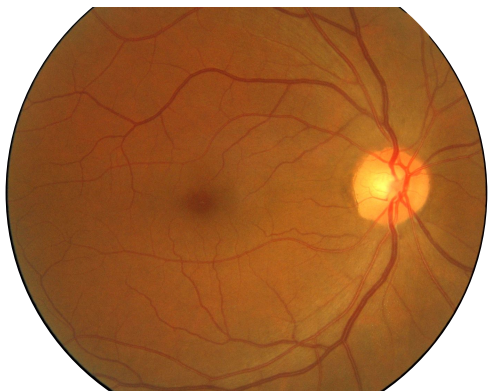
JAMA | **Original Investigation** | INNOVATIONS IN HEALTH CARE DELIVERY

## Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs

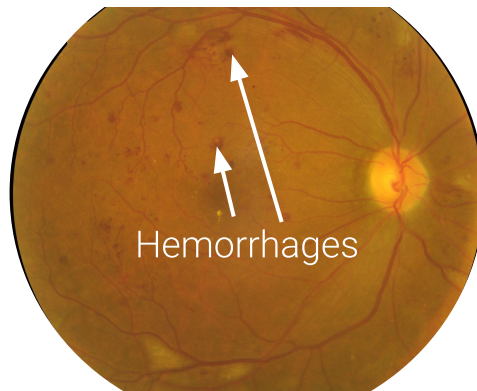
Varun Gulshan, PhD; Lily Peng, MD, PhD; Marc Coram, PhD; Martin C. Stumpe, PhD; Derek Wu, BS; Arunachalam Narayanaswamy, PhD;  
Subhashini Venugopalan, MS; Kasumi Widner, MS; Tom Madams, MEng; Jorge Cuadros, OD, PhD; Ramasamy Kim, OD, DNB;  
Rajiv Raman, MS, DNB; Philip C. Nelson, BS; Jessica L. Mega, MD, MPH; Dale R. Webster, PhD

# How DR is diagnosed:

## **Retinal fundus Images**



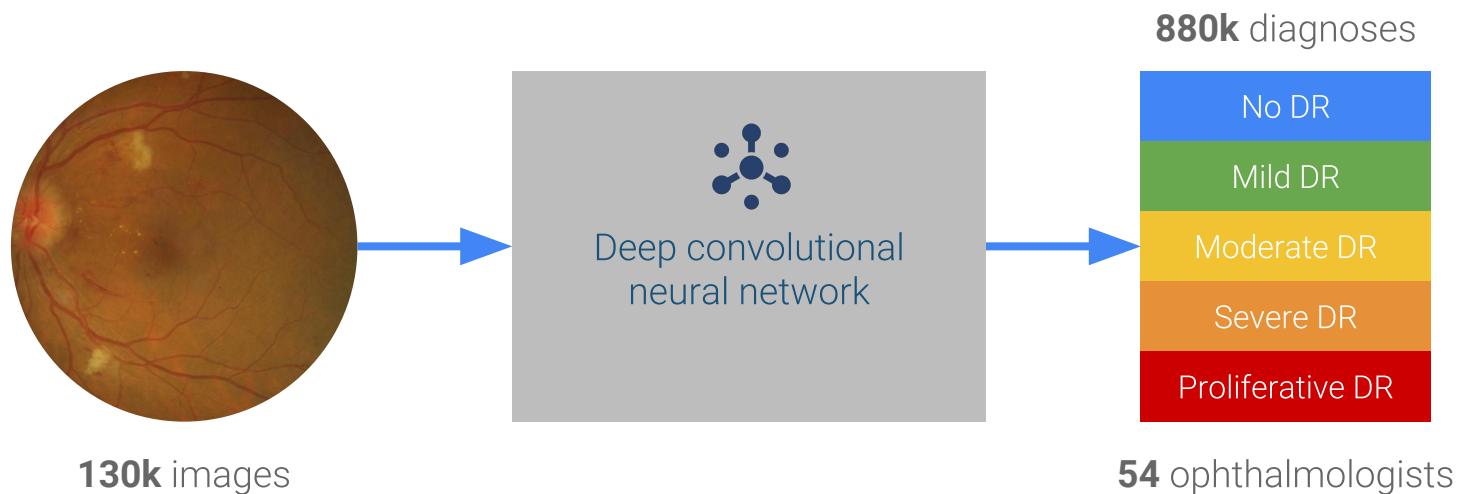
Healthy



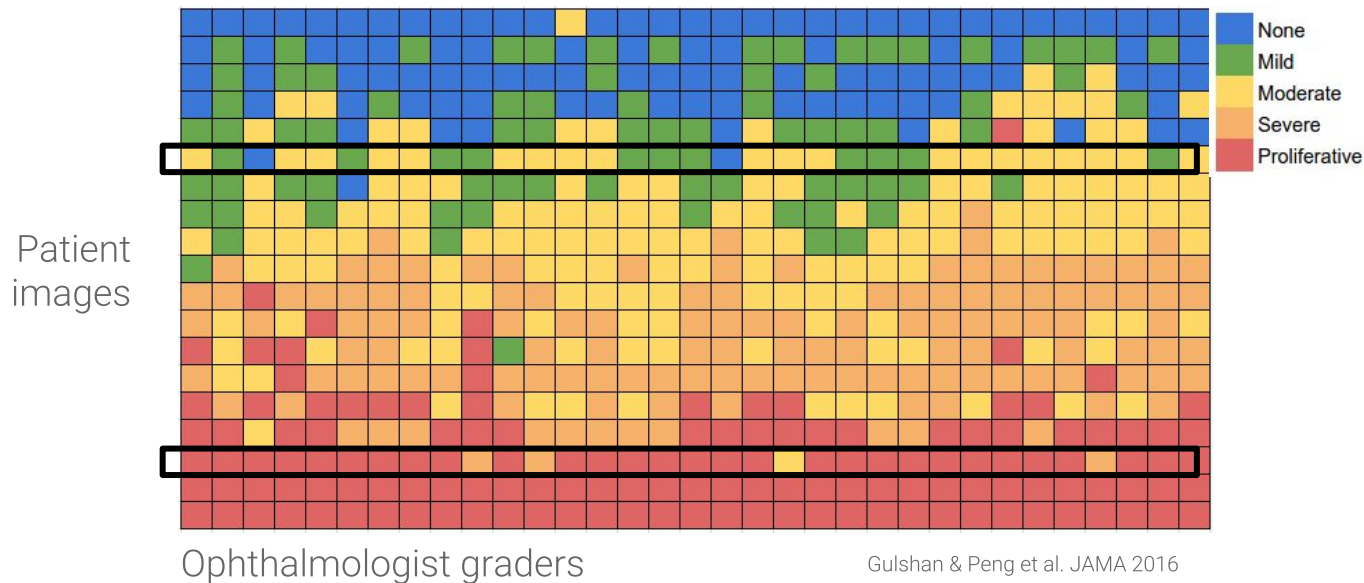
Diseased

No DR
Mild DR
Moderate DR
Severe DR
Proliferative DR

# Adapt a Google deep neural network to read fundus images

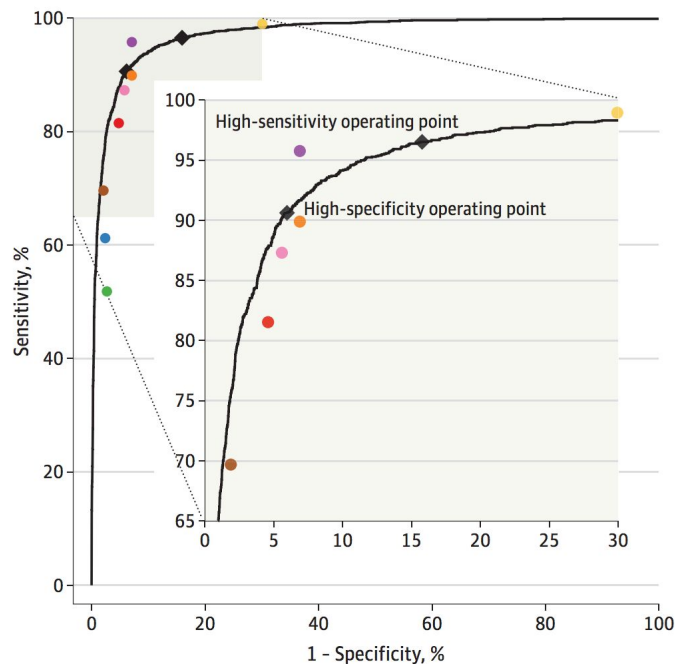


# As we would expect, doctors' opinions differ.



# On par with human ophthalmologists

Figure 3. Validation Set Performance for All-Cause Referable Diabetic Retinopathy in the EyePACS-1 Data Set (9946 Images)



Median of 8 ophthalmologists

**0.91**


F-score

**Google ARDA Algorithm**

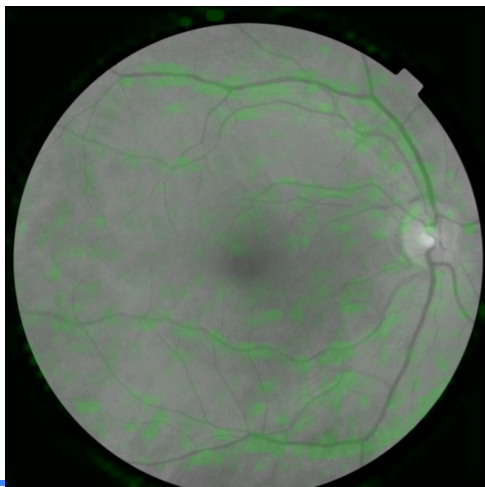
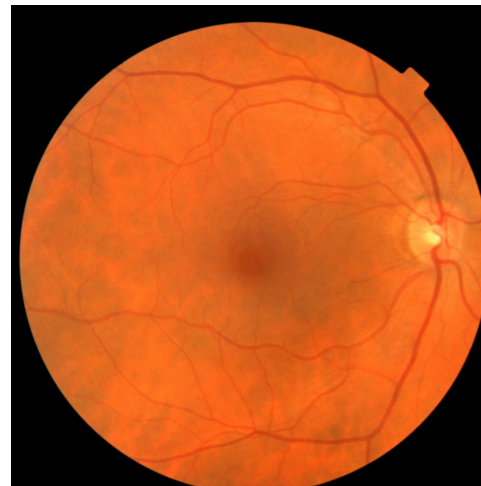
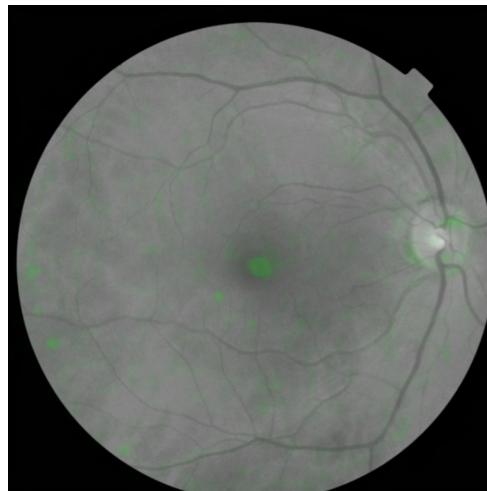
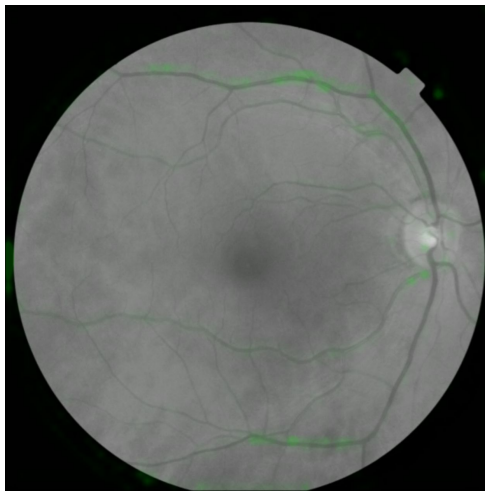
**0.95**

F-score





But that is not  
how the story  
ends!



Age  
Actual: 57.6 years  
Predicted: 59.1 years

Self-reported sex  
Actual: Female  
Predicted: Female

Current smoker  
Actual: Nonsmoker  
Predicted: Nonsmoker

BMI  
Actual: 26.3 kg/m<sup>2</sup>  
Predicted: 24.1 kg/m<sup>2</sup>

# Machine Learning in Genomics

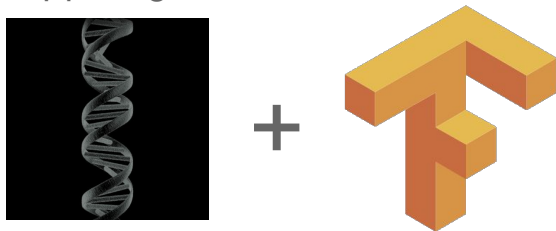
Google Cloud



# The Genomics' team mission

1

Extend TensorFlow to better support genomics data



2

Develop deep learning models for generalizable genomics problems

**DeepVariant**

sequence (ex.  
DNA/RNA)  $\rightarrow$  label

**Biomarkers**

3

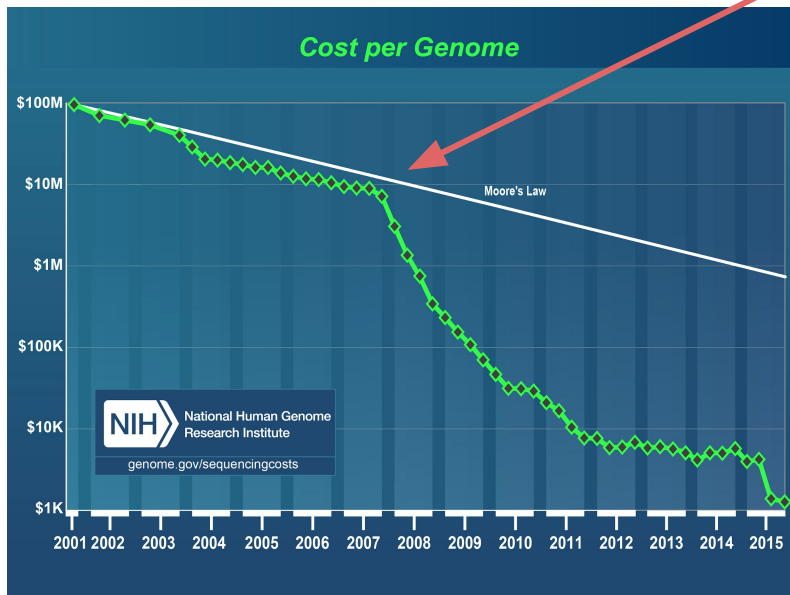
Release tools and capabilities as open source software





# Next-generation DNA sequencers (NGS) traded data quality for quantity

Massively parallel sequencing technology arrives



True genome sequence: 3 billion bases in 23  
contiguous chunks (chromosomes)

..... cttgggttga tattgtcttg gaacatggag gttgtgtcac cgtaatggca caggacaaac cgactgtcga catagagctg gttacaacaa  
cagtcagcaa catggcggag gtaagatcct actgctatga ggcataata tcagacatgg ctctggagac .....



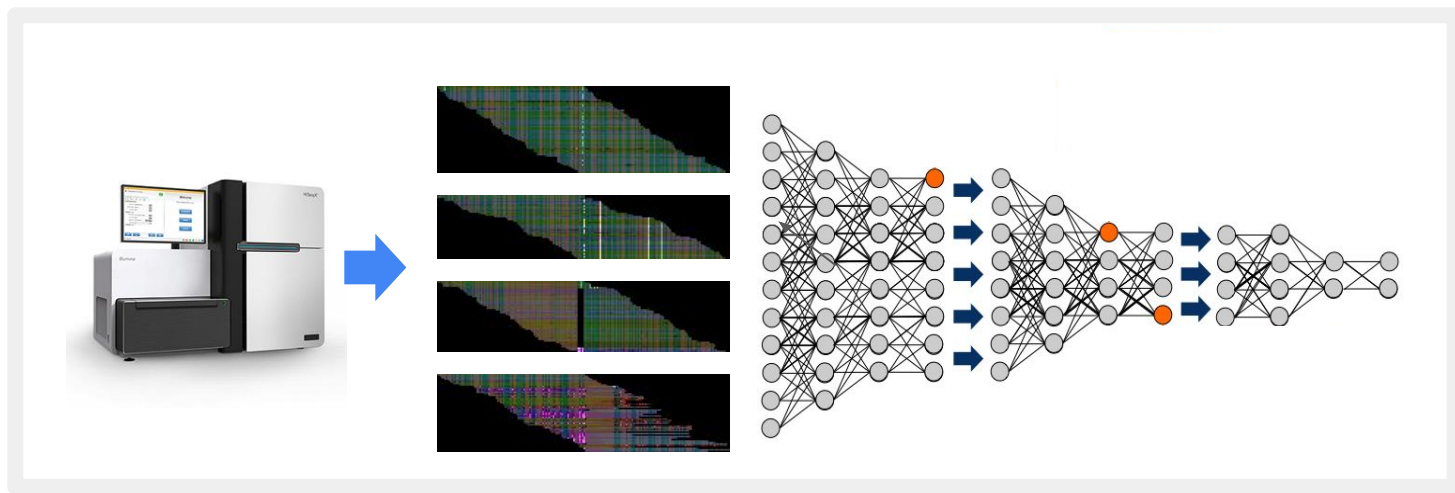
Actual sequencer output: ~1 billion ~100  
basepair long DNA reads (30x coverage)

Read1: cttgggtgatattgtcttggaaacatggagttgtgtcacccgtaatggcacaggacaaacc  
Read2: gatatgtcttggaaacatggagttgtgtcacccgtaatggcacaggacaaaccgactgtcg  
Read3: tggaaacatggagttgtgtcacccgtaatggcacaggacaaaccgactgtcgacatagagct  
Read4: ggttgtgtcacgtaatggcacaggacaaaccgactgtcgacatagagctggttactgtcg  
....  
Read 1,000,000,000: ....aactgtcgacatagagctggttactgtcgacatagagctggt

No location info in the  
reads. Base error rates  
between 0.1-10%.

# DeepVariant – ML-based variant caller

- Developed by Verily and Google Brain
- Convolutional neural network approach



# DeepVariant open sourced Dec 4th, 2017

google / deepvariant

Watch 90 Star 818 Fork 183

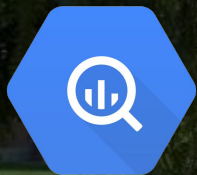
<> Code Issues 3 Pull requests 0 Projects 0 Insights

DeepVariant is an analysis pipeline that uses a deep neural network to call genetic variants from next-generation DNA sequencing data.

tensorflow deep-neural-network genomics science dna sequencing genome bioinformatics deep-learning ngs deepvariant machine-learning

104 commits 3 branches 3 releases 7 contributors BSD-3-Clause

Top tool on GitHub across many genome informatics categories



# BigQuery



Using BigQuery as their enterprise data warehouse to join patient, diagnosis, treatment, outcome and genome data to create personalized medicinal care for each patient.



**6 million**

UC Denver uses information about a patient's DNA from their genomic profiles to predict the risk of developing diseases and to develop targeted treatments. That requires examining the genetic makeup, health history and treatments of over 6 million patients.



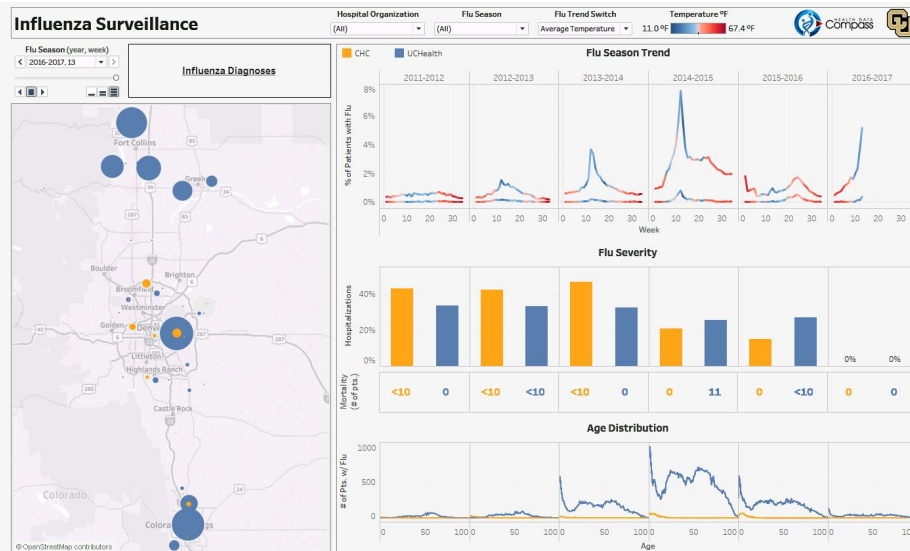
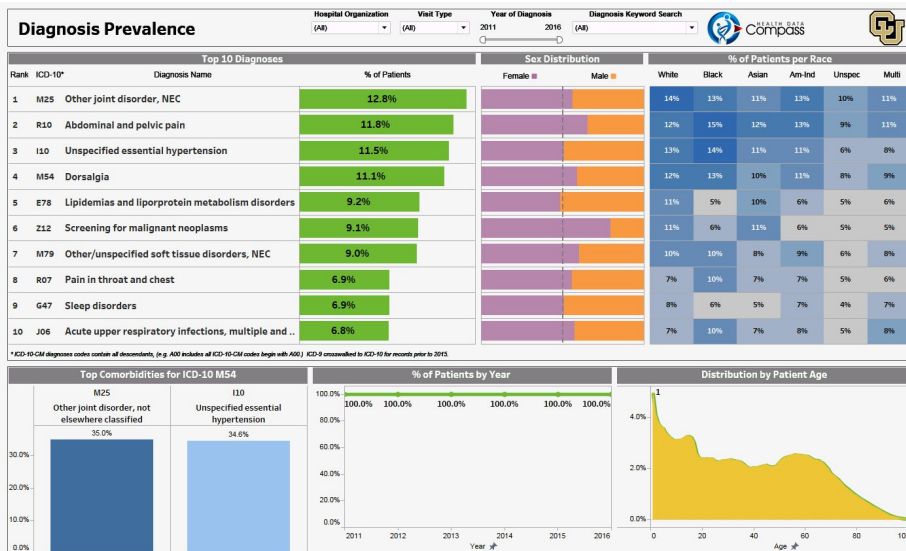
**8 hrs → 15 min**

Previously, it took eight hours for the Oracle system to execute complex analyses on the entire set of patient records. In BigQuery, it takes just 15 min.





# Insights



# Three ways Google Cloud can help you Benefit from Machine Learning

Build your own models:  
your data + your model

Cloud-ML



Train our implemented models:  
your data + our models

AutoML



Call our trained models as APIs:  
our data + our models



Cloud  
Vision API



Cloud  
Speech API



Cloud  
Jobs API



Cloud  
Translation API



Cloud Natural  
Language API



Cloud Video  
Intelligence API

# GCP research credits program



Move from bold ideas to breakthrough discoveries in a fraction of the time with Google Cloud Platform. Try it with free credits.

Faculty researchers in eligible countries can apply today for GCP credits to access the power and flexibility needed to advance their research and scale with ease.

# GCP education grants

## Prepare Your Students for Cloud Computing

Give students the ability to learn on the leading cloud platform. On Google Cloud Platform, students will find all the tools necessary to build a wide range of applications and have access to the same infrastructure, data analytics, and machine learning that Google uses.

University faculty can apply for Google Cloud Platform Education Grants for their computer science courses. Approved applicants will receive a grant for free credits which their teaching staff and students can use on any Google Cloud Platform product

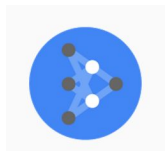




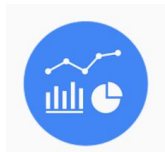
# What do you need to start?



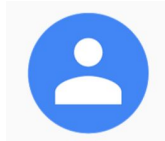
A **C-level sponsor** inside your organization



A defined **machine learning use case** that can generate significant value



**Data** and the ability to move it to the Google Cloud



**Team** with business people and engineers

**Dr. Stefan Ebener**

sebener@google.com



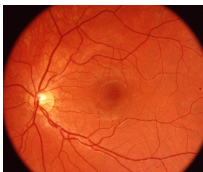
GOOGLE CLOUD





# Machine Learning

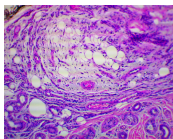
## Retina



Diabetic retinopathy  
Diabetic macular edema  
Glaucoma

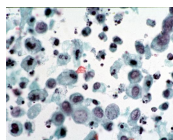
Age-related macular degeneration  
Retinopathy of prematurity  
Hypertensive retinopathy

## Digital pathology



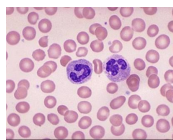
### Histopathology

Cancer biopsies  
Infectious disease



### Cytopathology

Pap smear  
FNA



### Hematology

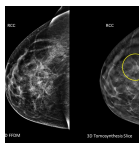
Blood smear

## Radiology



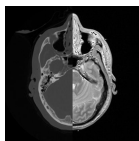
### X-ray

PTX, PNA  
Lung Cancer Screening



### Mammography

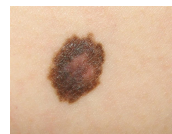
Breast Cancer Screening



### CT

Cancer Screening  
Stroke  
Pulmonary Embolism

## Other

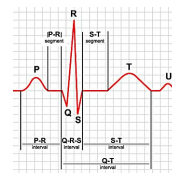


### Skin

Cancer, General  
Rashes



### Other clinical photos & signals



Ear infections  
EKGs, EEGs, EMGs  
Sleep studies, ICU  
monitoring