Artificial Intelligence in Medical Imaging
--- Pearl and Pitfall

Zhen Li, MD, PhD,
Professor, Chief physician, Vice Chairman, Department of Radiology
Tongji Hospital, Huazhong University of Science and Technology
AI Change Life
Medical Crisis---need AI

- Dependency Ratio increase --- Not enough money
Medical Crisis all over the World

### Share of Population Age 65 And Over

<table>
<thead>
<tr>
<th>Country</th>
<th>2014</th>
<th>2050</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>25.8%</td>
<td>40.1%</td>
</tr>
<tr>
<td>Germany</td>
<td>21.1</td>
<td>30.1</td>
</tr>
<tr>
<td>France</td>
<td>18.3</td>
<td>25.8</td>
</tr>
<tr>
<td>U.K.</td>
<td>17.5</td>
<td>23.6</td>
</tr>
<tr>
<td>Canada</td>
<td>17.3</td>
<td>26.3</td>
</tr>
<tr>
<td>Poland</td>
<td>15.0</td>
<td>31.7</td>
</tr>
<tr>
<td>U.S.</td>
<td>14.5</td>
<td>20.9</td>
</tr>
<tr>
<td>Russia</td>
<td>13.3</td>
<td>25.7</td>
</tr>
<tr>
<td>China</td>
<td>9.6</td>
<td>26.8</td>
</tr>
<tr>
<td>Brazil</td>
<td>7.6</td>
<td>21.1</td>
</tr>
<tr>
<td>India</td>
<td>5.8</td>
<td>14.7</td>
</tr>
</tbody>
</table>

Not enough doctors
Fewer want to become doctors

- Less US young generation want to became a doctor
10-year trends in the production and attrition of Chinese medical graduates: an analysis of nationwide data

Selina S Lien, Russell O Kosik, Angela P Fan, Lei Huang, Xudong Zhao, Xiaojie Chang, Yuhwa Wang, Qi Chen

Abstract

Background Over the past decade, China’s systems of medical education and health care have undergone unprecedented reform. Despite these reforms, a trend of declining interest among medical graduates in pursuing careers in clinical practice has persisted. The aim of this analysis is to examine physician workforce production and attrition rates and use current trends to predict whether an adequate and equitable supply of physicians will exist in the future.

Methods We analysed data about Chinese medical graduates and physicians in clinical practice between 2005 and 2015 using the Chinese National Health and Family Planning Commission reports and yearbooks published by the government of China and the Peking Union Medical College. These sources covered all of China (31 provinces; Hong Kong and Macao were not included). The international review board of National Yang-Ming University approved this study.

Findings From the beginning of 2005 to the end of 2014, China produced 4 314 791 5-year clinical medical graduates and 413 186 5-plus-2 programme (5-year medical education programme plus 2-year graduate programme) medical graduates: total 4 727 977 clinical medical graduates. However, during this period, there was an increase of only 752 233 (15·91%) in the total number of clinical physicians registered in practice. Using demographic data from this 10-year period, we found that the proportion of physicians aged 25–34 years had decreased from 31·3% to 22·6%, and the proportion of physicians aged 60 years and older had increased from 2·5% to 11·6%. Meanwhile, 5-plus-2 programme graduates increased from 4·3% to 11·2%, and rural areas had a shortfall of over 500 000 physicians.

• Only 15.91% medical students registered as physicians
### Which Countries Worry The Most About Healthcare?

Share worried about healthcare in selected countries in January 2018

<table>
<thead>
<tr>
<th>Country</th>
<th>Worry Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungary</td>
<td>72%</td>
</tr>
<tr>
<td>Poland</td>
<td>62%</td>
</tr>
<tr>
<td>Brazil</td>
<td>46%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>42%</td>
</tr>
<tr>
<td>China</td>
<td>38%</td>
</tr>
<tr>
<td>United States</td>
<td>38%</td>
</tr>
<tr>
<td>Russia</td>
<td>27%</td>
</tr>
<tr>
<td>Australia</td>
<td>25%</td>
</tr>
<tr>
<td>Japan</td>
<td>17%</td>
</tr>
<tr>
<td>Peru</td>
<td>17%</td>
</tr>
<tr>
<td>France</td>
<td>14%</td>
</tr>
<tr>
<td>India</td>
<td>14%</td>
</tr>
<tr>
<td>Germany</td>
<td>13%</td>
</tr>
<tr>
<td>South Korea</td>
<td>6%</td>
</tr>
<tr>
<td>Turkey</td>
<td>3%</td>
</tr>
</tbody>
</table>

n=20,202 adults ages 16–64 in 27 participating countries

Source: Ipsos
Medical Crisis

• Dependency Ratio increase --- Not enough money
• More and More aged people --- Need more
• Not enough doctors
• Fewer want to became a doctor
• --- Worry about Healthcare
Discussion---AI Can?

- Cut Medical cost?
- Improve work efficiency?
- Replace doctor’s work?
- Standardized treatment of patients?
- ----AI doctors?
Medical imaging Fit for AI

• Cost increase --- There are approximately 30,000,000 MRI procedures performed each year, with an annual 9% growth rate - more than twice that of general medical expenditures (4.1%) in USA.
• Relatively “simple”
• Digit data
• Relatively standardized data collection
• Big data

<table>
<thead>
<tr>
<th>Testing Facility Location</th>
<th>Test Type</th>
<th>Average Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orlando, FL</td>
<td>MRI</td>
<td>$2,229</td>
</tr>
<tr>
<td>Dallas, TX - MRI Testing Facility A</td>
<td>MRI</td>
<td>$3,624</td>
</tr>
<tr>
<td>Dallas, TX - MRI Testing Facility B</td>
<td>MRI</td>
<td>$2,172</td>
</tr>
<tr>
<td>San Diego, CA</td>
<td>MRI</td>
<td>$2,826</td>
</tr>
<tr>
<td>Salt Lake City, UT</td>
<td>MRI</td>
<td>$1,694</td>
</tr>
<tr>
<td>Detroit, MI</td>
<td>MRI</td>
<td>$3,461</td>
</tr>
<tr>
<td>New York, NY - MRI Testing Facility A</td>
<td>MRI</td>
<td>$1,785</td>
</tr>
<tr>
<td>New York, NY - MRI Testing Facility B</td>
<td>MRI</td>
<td>$2,199</td>
</tr>
<tr>
<td>Raleigh, NC</td>
<td>MRI</td>
<td>$3,001</td>
</tr>
<tr>
<td>Omaha, NE</td>
<td>MRI</td>
<td>$2,502</td>
</tr>
</tbody>
</table>
Pearl: AI is really HOT in Radiology --- RSNA 2018

- Topic “Tomorrow Radiology Today”—— AI
• President of the Radiological Society of North America (RSNA), Vijay M. Rao: "How Emerging Technology Will Empower Tomorrow's Radiologists to Provide Better Patient Care."
• Pro. Michael P. Recht: Artificial Intelligence, Analytics, and Informatics: The Future is Here
Where dreams come true!
Many new AI companies emerged
• ELEVATING RADIOLOGY: Provide important information to radiologists to help guide diagnosis, speed up treatment and improve patient satisfaction

• Snapshot Freeze 2: Full-heart motion tracking freezing platform based on deep learning SSF 2.0

• Bone VCAR: Spine assessment technique based on deep learning

• True Fidelity Images: Deep learning image reconstruction system
Articles in WOS

- Artificial Intelligence
- AI AND medical
- AI AND Imaging
• Radiomics

• Texture  AND CT

• Texture  AND MR
Main AI Methods

• Texture Analysis
• Radiomics
• Machine Learning
• Depend on Principle of tumor heterogeneity
• Main Clinical Application
Normal Tissue: low phenotypic heterogeneity

- Noise: low
- Genotypes: homogeneous
- Microenvironment: structured

Tumor: high phenotypic heterogeneity

- Noise: high
- Genotypes: heterogeneous
- Microenvironment: disorganized

Polyak, Kornelia TI, et al. Intra-tumour heterogeneity: a looking glass for cancer?
• Metrics quantify the perceived feature of an image

Tumor Heterogeneity and Texture Analysis
- Grayscale texture analysis method
  - Fractal Analysis
- GLCM, RLM..
- Wavelet Analysis
Pre-processing

Select images or ROIs

Gray level quantization

Feature estimation

1st order gray-level (mean, variance), gradient

2nd order GLCM, RLM

Spectral

Feature selection

Statistical

1st order gray-level features of same type

PCA

Step-wise discriminant analysis

Classification

LDA

Neural networks

Bayes decision

SVM

Logistic regression

Decision trees

K-nearest-neighbor

Evaluation

Test/training set + cross-validation

Leave-one-out

K-means or hierarchical clustering
Changes in radiomic features on CT between baseline and two weeks post-nivolumab treatment are predictive of response in non-small cell lung cancer (NSCLC) as defined by RECIST: Preliminary findings

Dr: Mohammadhadi Khorrami

- A recent landmark development in lung cancer therapy is the increasing use of immune checkpoint inhibition (ICI) therapy primarily due to their:
  - potential to induce durable and clinically meaningful responses
  - excellent toxicity profile
- Nivolumab acts by blocking a negative regulator of T-cell activation and response (PD-1:PD-L1 axis), thus allowing the immune system to attack the tumor.
- **BUT:** the response rate to nivolumab is around ~ 20 %, 3M
Motivation

- **Radiomics**: the analysis of quantitative features extracted from medical images.
- *We have previously used radiomics on pre-treatment CT to characterize response to Chemo- and trimodality (surgery + chemoradiation) therapy in NSCLC.*

**Objective**: Would quantifying the change in radiomic features before and after therapy be leveraged to predict response to therapy in NSCLC?
Patients with stage III unresectable or stage IV NSCLC who received Nivolumab (n=73)

15 patients who responded to Nivolumab
14 Nivolumab Not-responsive tumors
Blinded for Validation (N = 44)

27 Responders*
17 Not-Responders*

*Patients who did not receive Nivolumab after 2 cycles due to lack of response or progression as per RECIST were classified as ‘non-responders’.
Step 1: Manual segmentation of lung nodules

Step 2: Feature Extraction
- Gabor (48)
- Law (25)
- Haralick (13)
- Collage (13)

Step 3: Classifier Construction
Feature Selection
- Cross validation
- Repeat at least 100 times
- Sort based on AUC

Building Classifier
- LDA

Step 4: Independent validation of classifier
Testing the built classifier on the independent and blinded validation cohort and reporting results as Area Under ROC curve
Intratumoral Radiomic Features

- **Biological Markers**
  - Haralick and Gabor features represent the relative tumoral heterogeneity (more aggressive tumors have are more heterogeneous)
  - Laws’ features are a quantitative measure of how the texture varies as we move around the tumor

4 top selected features
1. Skewness of Haralick
2. Mean of Haralick
3. Median of Laws $L_5 \times W_5$
4. Kurtosis of Gabor
Peritumoral Radiomics

• Peritumoral region is defined as the space immediately surrounding the mass
• Peritumoral features can possibly capture Tumor infiltrating lymphocytes (TIL) in responder patients

5 top annular peritumoral features
  • Kurtosis of CoLIAGe (ring 8)
  • SD of Gabor, f=0, $\theta=\pi/8$ (ring2)
  • Median of Laws $R5 \times S5$ (ring 12)
  • Skewness of Laws $R5*R5$ (ring 14)

Hypothesis: Change in intratumoral and peritumoral radiomic texture features between baseline and two weeks post-nivolumab treatment can predict response to Nivolumab.
Results

Training set (N = 29) → LDA Classifier → Validate on independent validation set (N = 44)

Top selected Delta-Radiomic features

The LDA classifier trained with top features resulted in AUC of 0.92 on the validation set in distinguishing responders from non-responders.
Correlation with Lymphocyte density

Responder

Non-Responder

Pink regions show the tumor infiltrating lymphocytes (TILs). Green regions are the non-TIL structures.

Lymphocyte density was found to be negatively ($\rho = -0.5$) correlated with a change in peritumoral Gabor feature between pre- and post-treatment CT scans from the first annular ring.
Take Home Message

• Combination of intra- and peri-tumoral radiomic features enabled prediction of response to nivolumab treated NSCLC patients from pre- and post-treatment non-contrast CT scans.

• The prediction performance on a validation set (N=44) had an AUC of 0.92.

• Lymphocyte density was found to be negatively (ρ=-0.5) correlated with a change in peritumoral Gabor feature from the first annular ring.

• We thus hypothesize that peritumoral textural features have an biological correlation with Tumor Infiltrating Lymphocytes (TILs) which provides a deeper understanding of imaging phenotypes and immune response.
A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study

Lancet Oncol 2018
Published Online
August 14, 2018
http://dx.doi.org/10.1016/S1470-2045(18)30413-3
See Online/Comment/
http://dx.doi.org/10.1016/S1470-2045(18)30429-7

Clinical Application
Summary

Background Because responses of patients with cancer to immunotherapy can vary in success, innovative predictors of response to treatment are urgently needed to improve treatment outcomes. We aimed to develop and independently validate a radiomics-based biomarker of tumour-infiltrating CD8 cells in patients included in phase 1 trials of anti-programmed cell death protein (PD)-1 or anti-programmed cell death ligand 1 (PD-L1) monotherapy. We also aimed to evaluate the association between the biomarker, and tumour immune phenotype and clinical outcomes of these patients.

Data input

<table>
<thead>
<tr>
<th>Training set</th>
<th>Validation sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOSCATO (n=135)</td>
<td>The Cancer Genome Atlas (n=119)</td>
</tr>
<tr>
<td>CT scans (DICOMs) RNA-seq (n=135)</td>
<td>CT scans (DICOMs) RNA-seq (n=119)</td>
</tr>
<tr>
<td>CT scans (DICOMs) RNA-seq (n=119) Pathology slides (n=77)</td>
<td>Immune phenotype cohort (n=100)</td>
</tr>
<tr>
<td>Estimation of CD8 cell infiltrate by CD8 gene expression signature</td>
<td>CT scans (DICOMs) (n=100) Clinical data (n=100)</td>
</tr>
<tr>
<td>Estimation of CD8 cell infiltration by CD8 gene expression signature (primary endpoint)</td>
<td>Association with tumour immune phenotype</td>
</tr>
<tr>
<td>Radiomic signature of CD8 cells</td>
<td>Predictive and prognostic responses to immunotherapy</td>
</tr>
</tbody>
</table>

Immunotherapy-treated dataset From phase 1 trials of anti-PD-1/PD-L1 monotherapy (n=137)
<table>
<thead>
<tr>
<th></th>
<th>All patients (n=135)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median age (IQR)</strong></td>
<td>55.6 (41.4-64.7)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>73 (54%)</td>
</tr>
<tr>
<td>Male</td>
<td>62 (46%)</td>
</tr>
<tr>
<td><strong>Type of cancer</strong></td>
<td></td>
</tr>
<tr>
<td>Adenoid cystic carcinoma</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Breast</td>
<td>17 (13%)</td>
</tr>
<tr>
<td>Colorectal</td>
<td>8 (6%)</td>
</tr>
<tr>
<td>Gynaecological</td>
<td>18 (13%)</td>
</tr>
<tr>
<td>Head and neck</td>
<td>14 (10%)</td>
</tr>
<tr>
<td>Kidney</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Liver</td>
<td>9 (7%)</td>
</tr>
<tr>
<td>Lung</td>
<td>30 (22%)</td>
</tr>
<tr>
<td>Neuroendocrine</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>Other*</td>
<td>4 (3%)</td>
</tr>
<tr>
<td>Other epidermoid carcinoma†</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>Prostate</td>
<td>7 (5%)</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>6 (4%)</td>
</tr>
<tr>
<td>Upper gastrointestinal tract</td>
<td>8 (6%)</td>
</tr>
<tr>
<td>Urothelial</td>
<td>8 (6%)</td>
</tr>
<tr>
<td><strong>Volumes of interest location</strong></td>
<td></td>
</tr>
<tr>
<td>Adenopathy</td>
<td>22 (16%)</td>
</tr>
<tr>
<td>Liver</td>
<td>54 (40%)</td>
</tr>
<tr>
<td>Head and neck</td>
<td>6 (4%)</td>
</tr>
<tr>
<td>Abdominopelvic or subcutaneous</td>
<td>13 (10%)</td>
</tr>
<tr>
<td>Lung</td>
<td>40 (30%)</td>
</tr>
<tr>
<td><strong>Median genomic-based CD8 cell score (IQR)</strong></td>
<td>1.60 (0.89-2.47)</td>
</tr>
</tbody>
</table>

Data are n (%), unless otherwise indicated. * Includes spindle epithelial tumours with thymus-like differentiation, peritoneal desmoplastic small round cell tumours, hepatoblastomas, and nephroblastomas. † Includes anal cancer and penile cancer.

Table 1: Characteristics of patients in the training dataset (MOSCATO trial)
Figure 2: Radiomics workflow

PD-L1 - programmed cell death ligand 1
A MOSCATO training set

**AUC = 0.74 (95% CI 0.66–0.82)**

TCGA validation set

**AUC = 0.67 (95% CI 0.57–0.77)**

Assumed tumour immune phenotype

**AUC = 0.76 (95% CI 0.66–0.86)**

B Wilcoxon p value = 0.013

Radiomics-based CD8 cell score

- High
- Low

Overall survival (%)

Hazard ratio = 0.58 (95% CI 0.39–0.87)

p value = 0.0081

Number at risk

- High: 68 (0), 62 (1), 54 (1), 48 (2), 46 (2), 42 (2), 37 (2), 34 (2), 30 (4), 27 (4), 25 (5), 21 (6)
- Low: 69 (0), 55 (0), 41 (1), 31 (2), 24 (2), 20 (2), 20 (2), 17 (2), 15 (3), 13 (3), 11 (6)

Response to anti-PD-1 or anti-PD-L1 at 6 months

Stable disease and disease control

Stable disease, partial response, and complete response
Figure 4: Evaluation of agreement between radiomic score of CD8 T cells and pathology
(A) Radiomic score of CD8 T cells as a function of the pathologist’s semi-quantitative quantification of tumour-infiltrating lymphocytes in the TCGA cohort, including patients with bladder cancer, liver hepatocellular carcinoma, lung adenocarcinoma, and lung squamous cell carcinoma. (B) Radiomic score of CD8 T cells by tumour type in head and neck tumours. Head and neck squamous cell carcinoma data are from the MOSCATO cohort (n=6) and the TCGA cohort (n=76), and adenoid cystic carcinoma data (n=4) are from the immune phenotype cohort. Seven patients with undifferentiated nasopharyngeal carcinomas from our institute (known to be very immune-inflamed) were added for this post-hoc analysis. TCGA=The Cancer Genome Atlas.
Findings:

• The radiomic predictor provided a promising way to predict the immune phenotype of tumours and to infer clinical outcomes for patients with cancer who had been treated with anti-PD-1 and PD-L1.

• Our imaging biomarker could be useful in estimating CD8 cell count and predicting clinical outcomes of patients treated with immunotherapy, when validated by further prospective randomised trials.
Discussion:

- **How do you think AI will change medicine?**
  - Rapid screening patients
  - Classification of patient
  - Remind possible errors
  - Quantitative analysis and prediction
  - Reduce medical error
  - Improve efficiency
  - Cut medical cost
  - Fully automated diagnosis
  - Reduce medical errors
  - Working as a doctor
  - Unmanned hospital
Pitfall

• No breakthrough in basic technical principles
• Data consistency
• Data reliability
• Sample limitation in machine learning
• Unknowns and uncertainties in Medicine
• Business, Business, Business
1 No breakthrough in basic technical principles

- Based on **Least squares (1950’s)**,
- "Least squares" means that the overall solution minimizes the sum of the squares of the errors made in the results of every single equation.
Artificial intelligence is actually statistics, but it uses a very gorgeous rhetoric, which is actually statistics. A lot of formulas are very old, but we say that all artificial intelligence uses statistics to solve problems.

• The AI neural network is a machine that simulates human neural networks. Many algorithms have evolved. The massive operation of massive artificial neurons (nodes) is itself a statistically optimal choice, such as AlphaGo’s "Deep Neural Network" and "Enhanced Neural Network".

• The enhanced learning of the AlphaGo "Enhanced Neural Network", the so-called left and right mutual struggle, is actually one of the best applications of "statistical optimization". The machine is self-study according to a certain model of procedures and mathematical formulas, which is completely different from human self-study. If it is a self-study of people, it is not necessarily summarized as statistical optimization.
Principle of AI
Sample Statistical Extracting features
Create model (LASSO)
Model verification
World Miss 2018 Winners
Depend on Sample and Model

- Classical China
- Classical Euro
- Modern USA
- Modern African American
Who do you choose among these candidates?

- It’s time to elect a leader, and your vote is crucial. Here are some facts about the three candidates:
  - Candidate A: There are contacts with some dishonest politicians, and there will be astrology for divination. He has an affair, an old smoker, and drinks 8 to 10 cups of martini a day.
  - Candidate B: He has had two records of dismissal in the past, sleeps until noon, gets poppies at college, and drinks a large quart of whiskey every evening.
  - Candidate C: He is a war hero who is honored, vegetarian, does not smoke, only occasionally drinks a little beer. There has never been an extramarital affair.
Depend on Data sample and Model

- A
- B
- C
IBM Dr Watson Failure

- Based on Sloan Kettering’s data
- But not fit for all hospitals
2 Data consistency

- Standardization
- Repeatability
- Consistency

- CT values are relative values (calibrated every morning?)
- Different equipment, KVp, mAs, contrast agent type, injection method, Scanning Method, etc.
HOW About MR?

- 1.5 T, 3T, manufacturer, TR, TE, TI, Nex,.....
3 Data reliability

- Patient History is reliable?
- Is the algorithm of data consistent and reliable?
- ROI?

European Journal of Cancer (2012) 48, 441–446

Radiomics: Extracting more information from medical images using advanced feature analysis

Philippe Lambin\textsuperscript{a,a,e,f}, Emmanuel Rios-Velazquez\textsuperscript{a,e}, Ralph Leijenaar\textsuperscript{a,e}, Sara Carvalho\textsuperscript{a,e}, Ruud G.P.M. van Stiphout\textsuperscript{a,e}, Patrick Granton\textsuperscript{a,e}, Catharina M.L. Zegers\textsuperscript{a,e}, Robert Gillies\textsuperscript{b,c}, Ronald Boellard\textsuperscript{c,e}, André Dekker\textsuperscript{a,e}, Hugo J.W.L. Aerts\textsuperscript{a,d,e}
4 Sample limitation

• For example: Rare disease

• According to the definition of the World Health Organization (WHO), rare diseases account for 0.65 ‰ to 1 ‰ of the total population.

• More than 300 million rare patients are expected worldwide, with more than 16.8 million in China. More than 7,000 rare diseases have been identified.

• ---- Is there enough sample for rare diseases?

• What is the 4.3% error rate mean? (0.3/7 Billion)
Difficult to establish unified data platform

- Standardization
- Different hospitals, different doctors
- Guideline change each year
- Law, culture
5 Unknowns and uncertainties in Medicine

• We know few about disease and human being

• Dermatologist-level classification of skin cancer with deep neural networks
  • Andre Esteva, *NATURE*, 542(7639) P115-+
  • DOI: 10.1038/nature21056, FEB 2 2017, 825 cites
Deep convolutional neural networks (CNNs) show potential for general and highly variable tasks across many fine-grained object categories. We test its performance against 21 board-certified dermatologists on biopsy-proven clinical images with two critical binary classification use cases: keratinocyte carcinomas versus benign seborrheic keratoses; and malignant melanomas versus benign nevi.

Outfitted with deep neural networks, mobile devices can potentially extend the reach of dermatologists outside of the clinic. It is projected that 6.3 billion smartphone subscriptions will exist by the year 2021 and can therefore potentially provide low-cost universal access to vital diagnostic care.
Figure 3: Skin cancer classification performance of the CNN and dermatologists.

- The deep learning CNN outperforms the average of the dermatologists at skin cancer classification using photographic and dermoscopic images. Our CNN is tested against at least 21 dermatologists at keratinocyte carcinoma and melanoma recognition.

• Daniel S. Kermany
(A–F) Comparisons were made for pneumonia versus normal (A) with cross-entropy loss plotted against the training step (B), as well as comparisons between bacterial pneumonia and viral pneumonia (C) and the associated cross-entropy loss (D). Plots were normalized with a smoothing factor of 0.6 in order to clearly visualize trends. The area under the ROC curve for detecting pneumonia versus normal was 96.8% (E). The area under the ROC curve for detecting bacterial versus viral pneumonia was 94.0% (F). Training dataset: orange. Validation dataset: blue.
How do you Think the two great papers?

• We test its performance against 21 board-certified dermatologists on biopsy-proven clinical images with two critical binary classification use cases: keratinocyte carcinomas versus benign seborrheic keratoses; and malignant melanomas versus benign nevi.

• Comparisons were made for pneumonia versus normal, as well as comparisons between bacterial pneumonia and viral pneumonia in Chest X-Ray Images, AI is better than Radiologist.
Are you Kidding me?---As a doctor

- How do you know this patient MUST be keratinocyte carcinomas OR benign seborrheic keratoses, malignant melanomas OR benign nevi? bacterial pneumonia OR viral pneumonia?
- Do you Know how many diseases in skin or Lung?
- How do you know that this case is the disease you listed? Can't it be other diseases? Even unknown diseases?
- Form doctor’s side: Exclude--less possible---possible—find
- Most of diseases are unclear until now---Black Box
- Not AS auto driver, For GPS
6 Business, Business, Business

• A: AI software developed by Big Data, who is the benefit?
• Did the patients agree?
• Is the data owned by the hospital?

• Increase medical cost or decrease?
C: Rare disease

- What is the 4.3% error rate mean? (0.3/7 Billion)
- Doctor make mistake: Insurance, Guideline
- AI error: Who pay? At least 5% error, Insurance Fee?
D: Security

• Private Data Security --- Personal Privacy
• National Security
• Race Specific---Gene Editing Events (Rare CCR5 Gene Mutations, AIDS Resistance in Some Populations in Western Europe)
Form a Radiologist side

- Health Management
- Rapid screening patients
- Help for Standardization
- Classification of patient
- Remind possible errors
- Reduce medical error
- Quantitative analysis and prediction
- More efficiency
- -----make a doctor’s life easier!
Ideal AI need:

- Standard scanner and methods
- Standard and reliable history and clinical data
- Private Data Security
- All data is safely shared, No copyright problem
- Reliable Model
- Most relative diseases can be acted as Standard models
- Cheaper than doctors
- Pay for Copyright
Summary

• Doctors and patients both need AI
• AI is really promising---good tool
• Wrong Way---diagnosis and treatment
• Right ---*screening, exclude* and *statistics*
• Data Security
• Business
Most Important

• Respect nature, Respect life
  ------We know few about nature, The Earth is no more than dusty
• Remember baseline
• ------Google：Don’t Be Evil
Acknowledgement: Our Team

Thanks for Attention!