

DERIVING PHYSICIANS' EXPERTISE PROFILES BASED ON ICD9-CODED ENCOUNTER NOTE LOGS

Victor Brodsky, MD

Tru V. Tran

Stephen E. Kessler II

Curtis Cole, MD

Weill Cornell Medical College



Weill Cornell Medical College

VIVO

Why? Can't we just ask?

- Medical scientists are erratic at self profiling
 - Accuracy
 - Completeness
 - Future Goals vs. Past Expertise
 - Regular updates
 - Time consuming process
- We know their specialties and billing activities
- Publication history & other sources still useful



Can someone else do it?

- The Challenge

- ▣ To accurately and efficiently determine a particular physician's areas of expertise, by disease/condition

- The Approach

- ▣ Automate the analysis of the ICD9-coded diagnostic information in the electronic medical patient records signed by physicians to determine areas of expertise

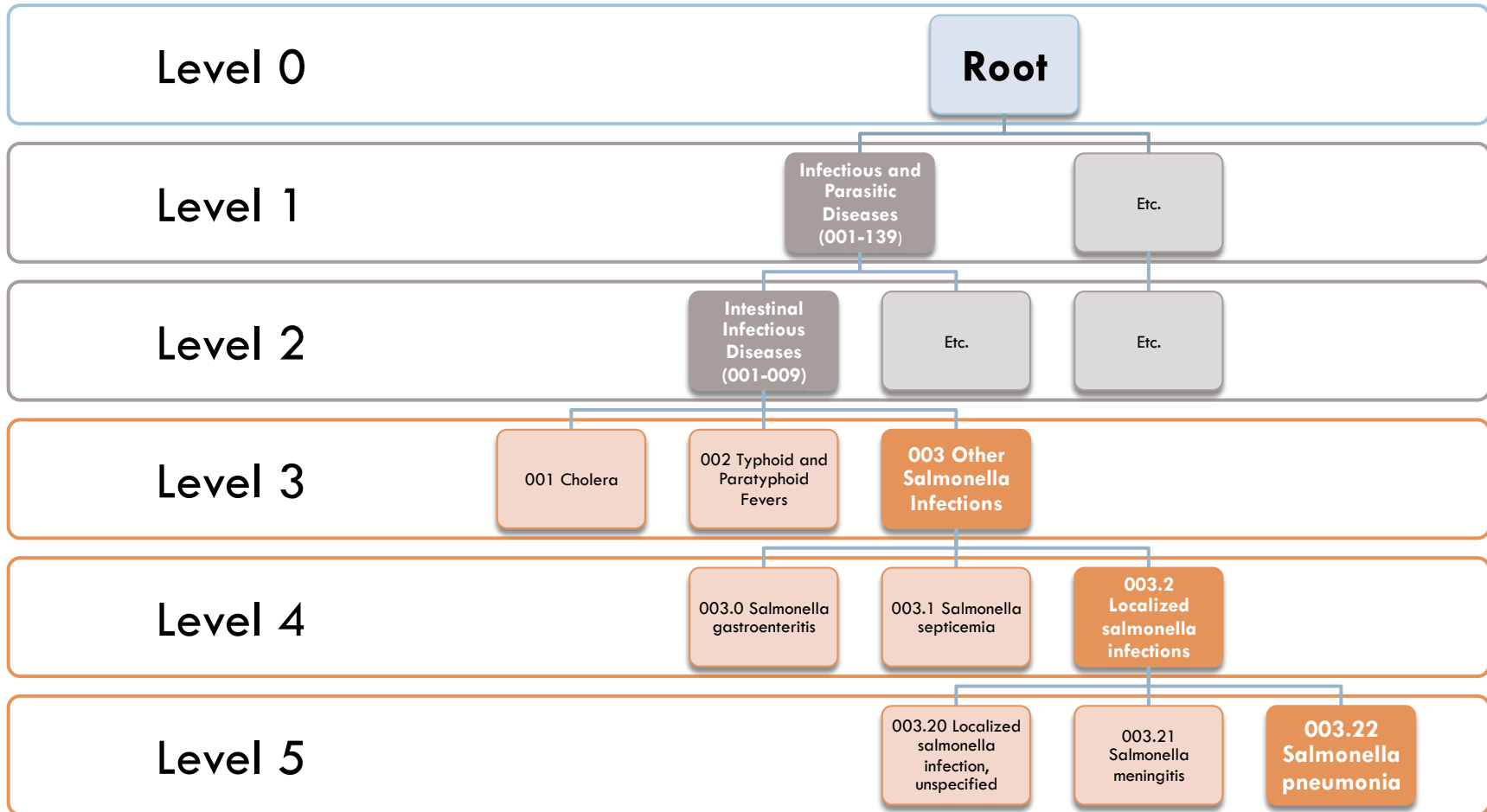


A Branch of ICD9 Hierarchical Tree

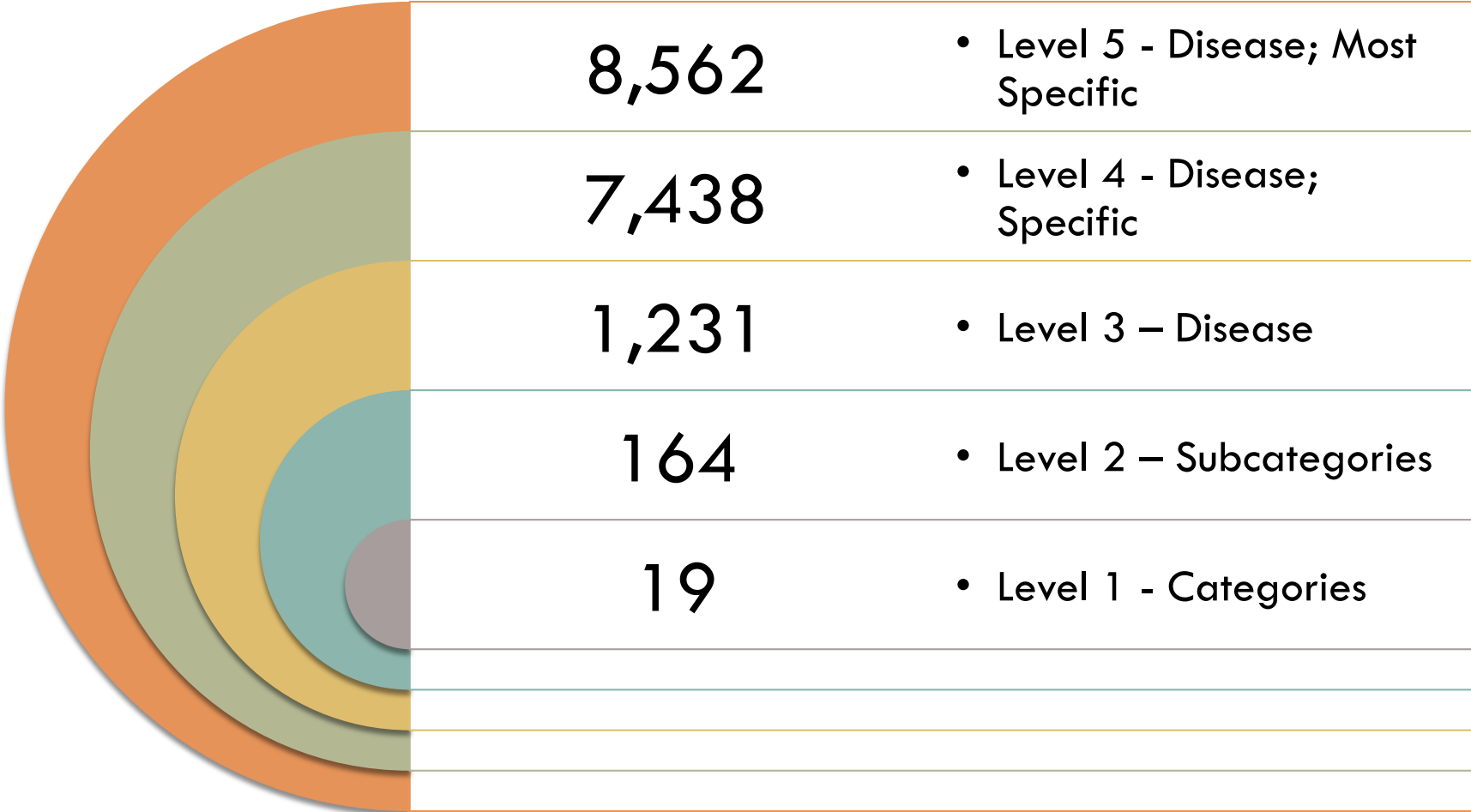
- **Infectious and Parasitic Diseases (001-139)** ← Level 1
 - **Intestinal Infectious Diseases (001-009)** ← Level 2
 - 001 Cholera
 - 002 Typhoid and Paratyphoid Fevers
 - **003 Other Salmonella infections** ← Level 3
 - 003.0 Salmonella gastroenteritis
 - 003.1 Salmonella septicemia
 - **003.2 Localized salmonella infections** ← Level 4
 - 003.20 Localized salmonella infection, unspecified
 - 003.21 Salmonella meningitis
 - **003.22 Salmonella pneumonia** ← Level 5



A Branch of ICD9 Hierarchical Tree



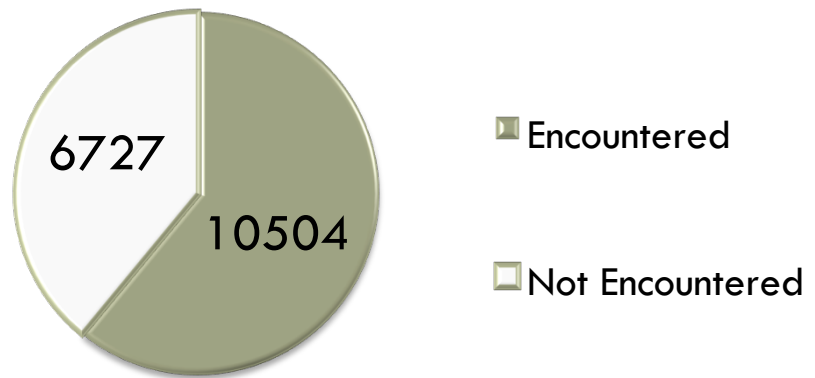
ICD9 Tree Statistics: 17,414 Items



Physician Population

- 2,443 physicians who saw > 5 patients
- Of 17,231 ICD9 Codes, each physician encountered from 1 to 1,747

Encountered ICD9 Codes



Ranking ICD9s

- Only some of logged ICD9 codes are pertinent to the physician
 - A psychiatrist may log two ICD9 codes in the EMR for a patient visit:
 - **300.3** Obsessive-compulsive disorders
 - **278.0** Overweight and obesity
- For each pair (a physician and an ICD9), we produce a score with a magnitude predicting how pertinent the given ICD9 code is to the given physician's expertise
- In a successful ranking system, ICD9 codes that are directly applicable to a physician's expertise are ranked higher than those that are incidental

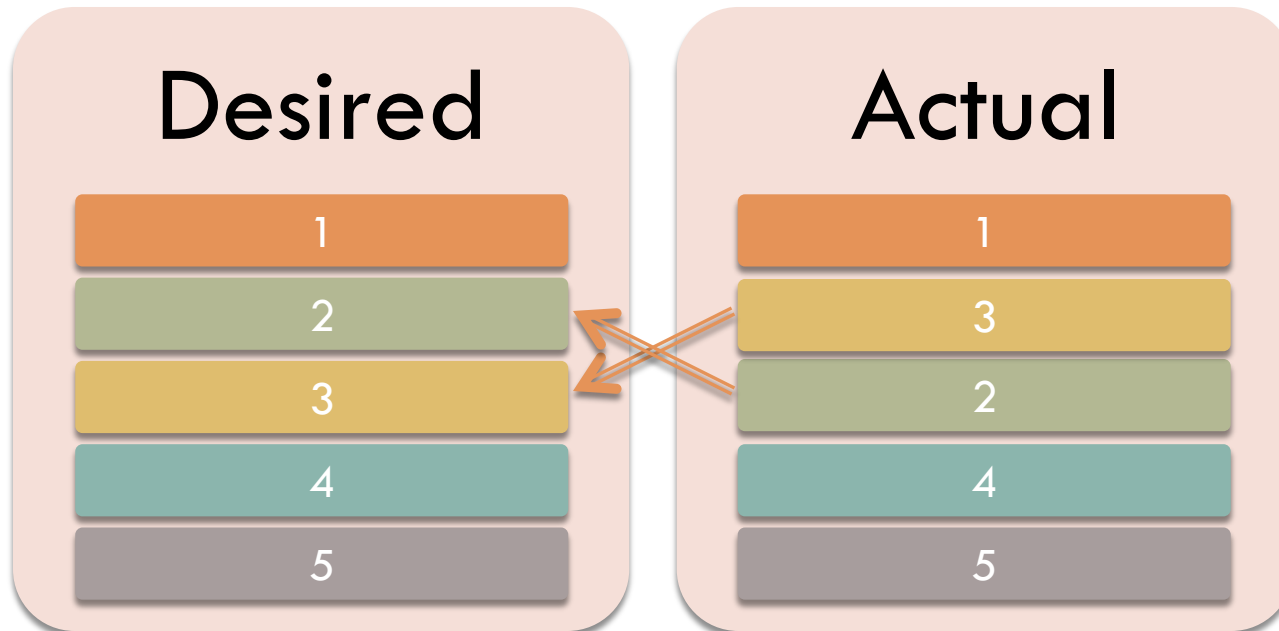


Measuring Ranking Success

- Sample physicians selected
- Each ICD9 that they had logged was subjectively marked as either relevant or irrelevant to their area of expertise based on their specialty
- Per-Physician ICD9 lists sorted by score
- Hamming distance to nearest target list found
- Will need to confirm scores by presenting profiles to the “owners” for grading



Hamming Distance



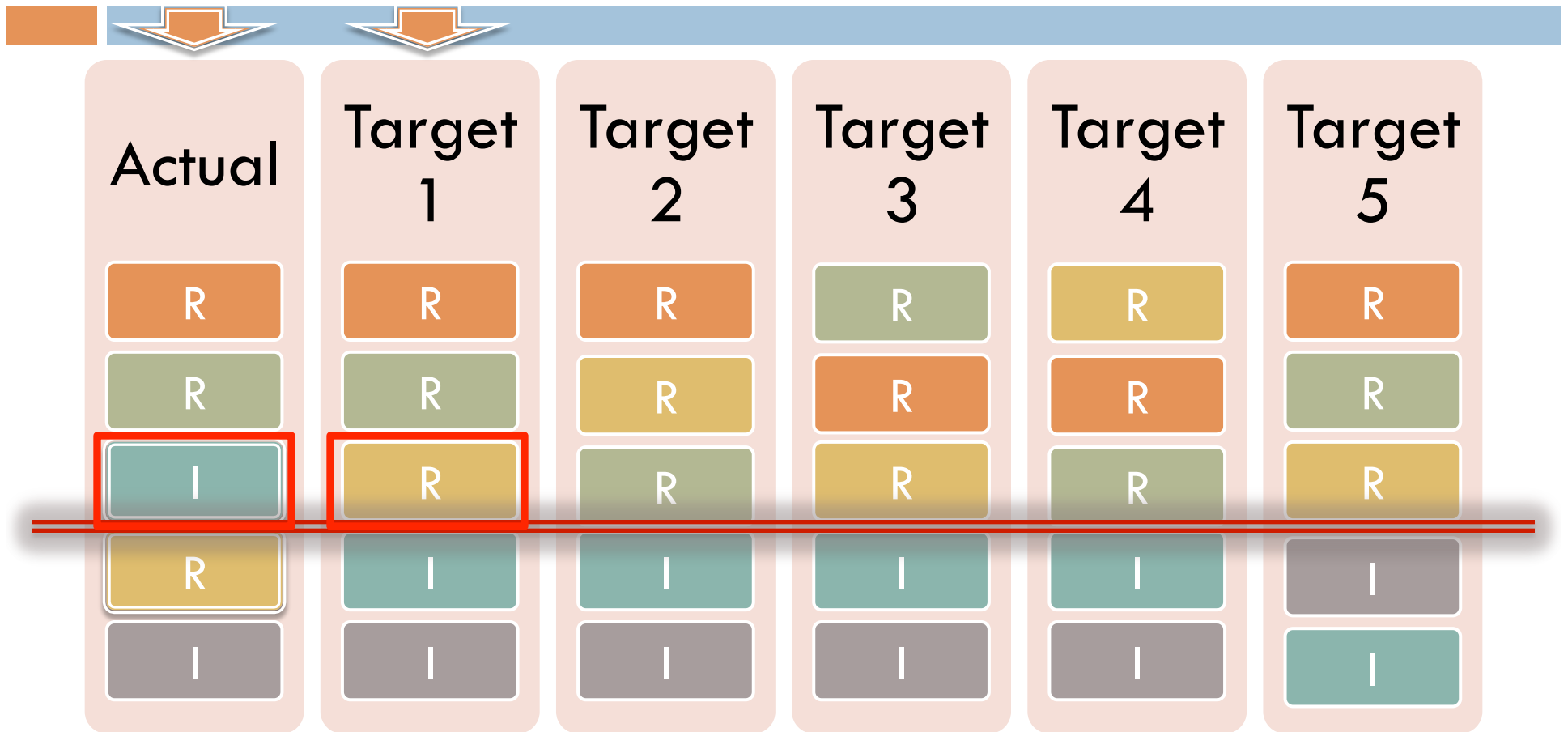
□ Hamming Distance = 2



Weill Cornell Medical College

VIVO

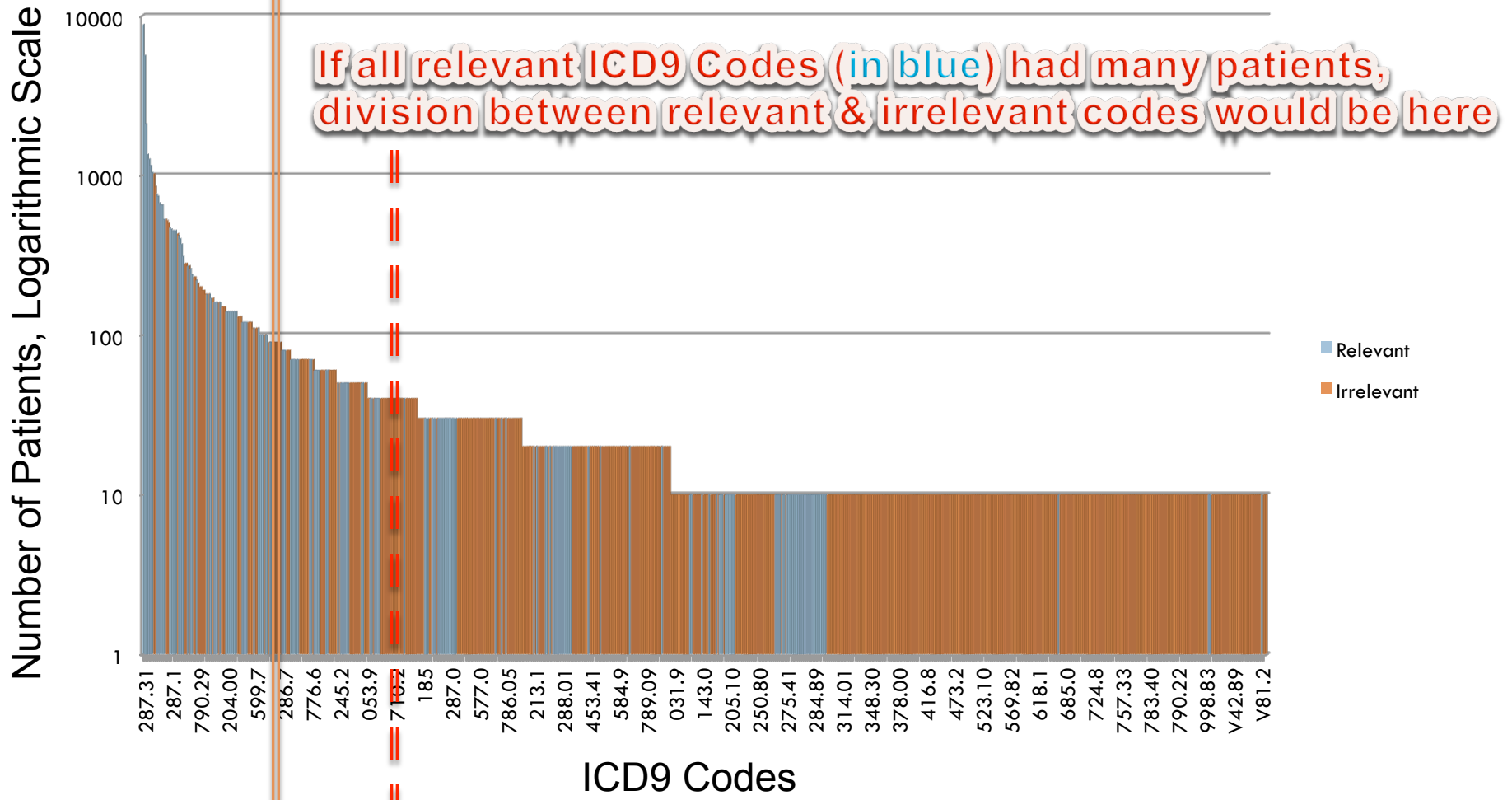
Hamming Distance, Nearest Target



- 3 Relevant, 1 is not in top 3 → HD = 1



First Efforts – Sort by # of Patients



Mean Number of Patients



Weill Cornell Medical College

VIVO

Incidence Ratios

- A physician may see one ICD9 more often than other ICD9s
- A physician may diagnose an ICD9 more often than other physicians diagnose that ICD9
- A disease may occur more often than another disease in our patient population

$$\text{Core Score} = \frac{\frac{\text{\# of patients seen for an ICD9}}{\text{total \# of patients seen}}}{\frac{\text{\# of patients seen for an ICD9}}{\text{total \# of patients seen}}}$$

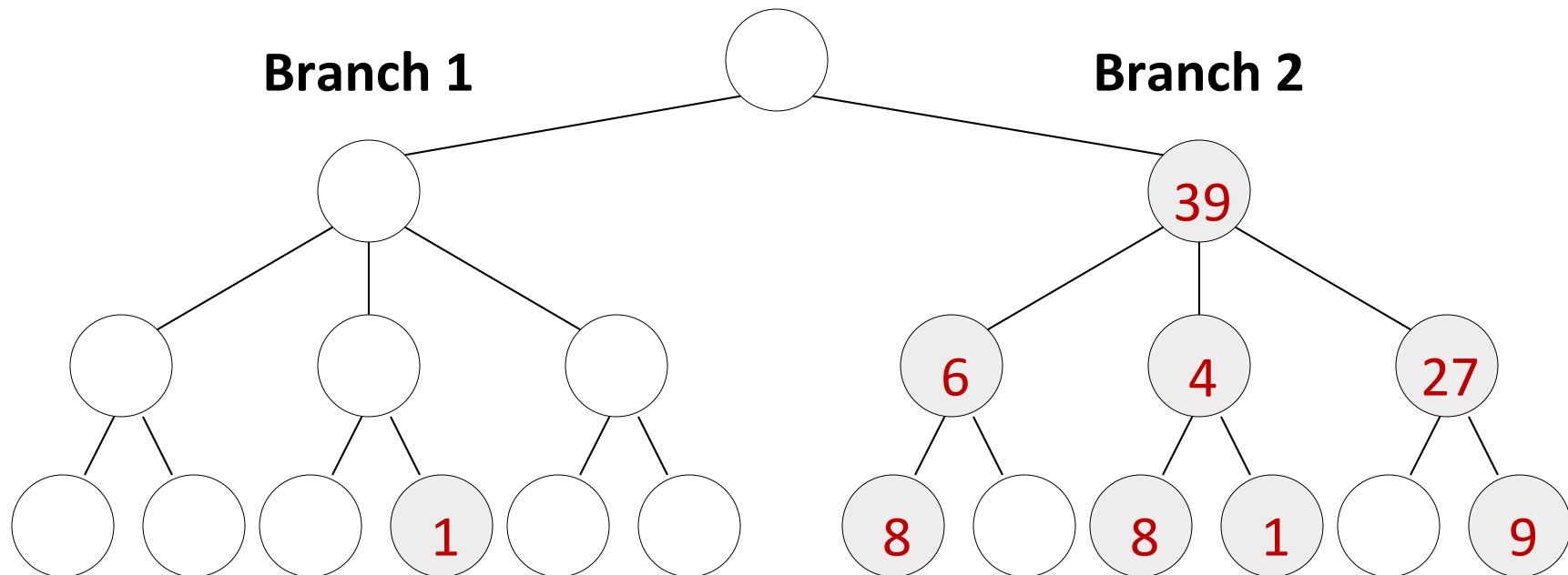
← One provider

← All providers



Utilizing the Tree Structure of the ICD9

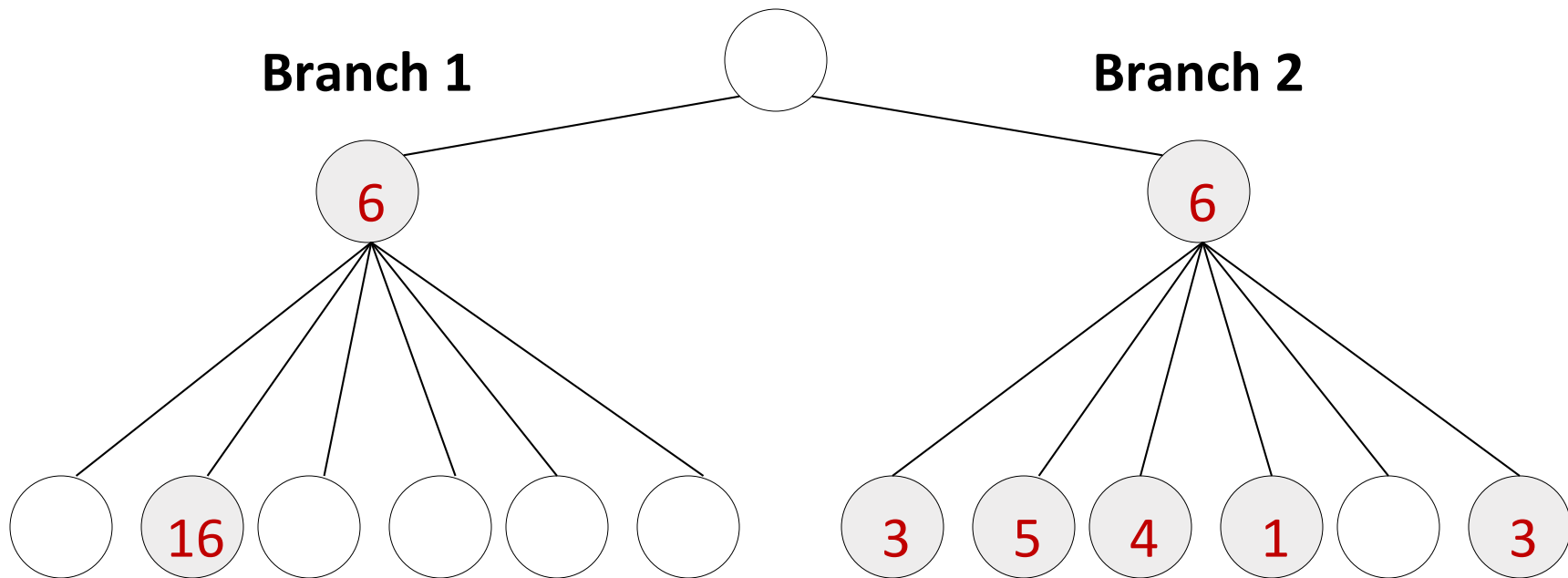
- Diagnoses by a particular physician tend to cluster within specific branches of the tree
- These clusters tend to be related to the physician's specialty
- We can utilize this info by giving these clusters greater weight



Accounting for coverage

□ 1 / 6

□ 5 / 6



Average with specialty average

100% Personal

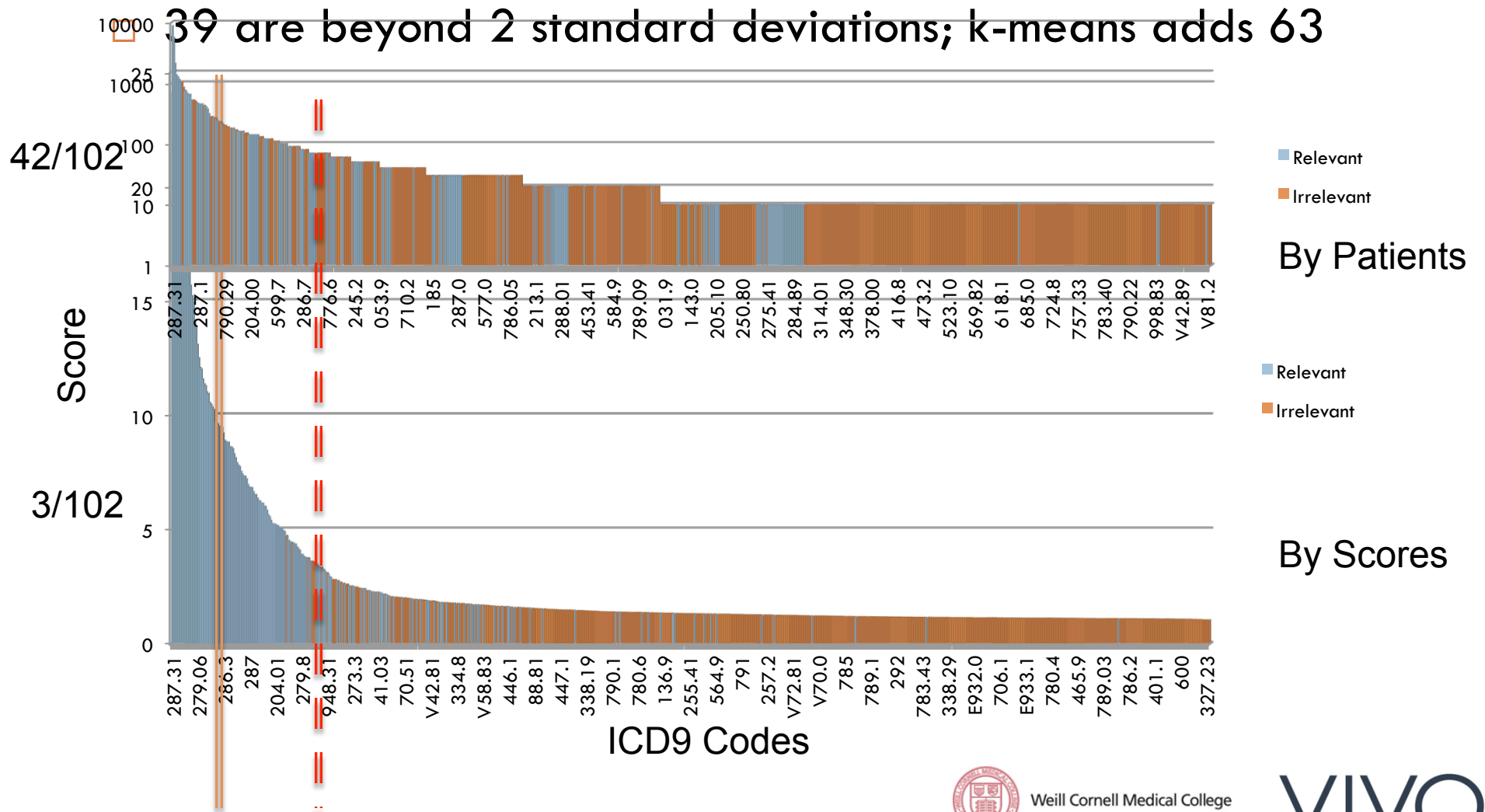
Transient Tic disorder 43
T7-T12 level w/ unspecified spinal cord injury 101
Circadian rhythm sleep disorder, jet lag type 112
Pedal cycle accident injuring unspecified person
Lethal midline granuloma
Pertussis alone (vaccination)
DTP with typhoid-paratyphoid (vaccination)
Screening examination for viral diseases
Counseling on substance use and abuse
Other abnormality of red blood cells

50% Specialty

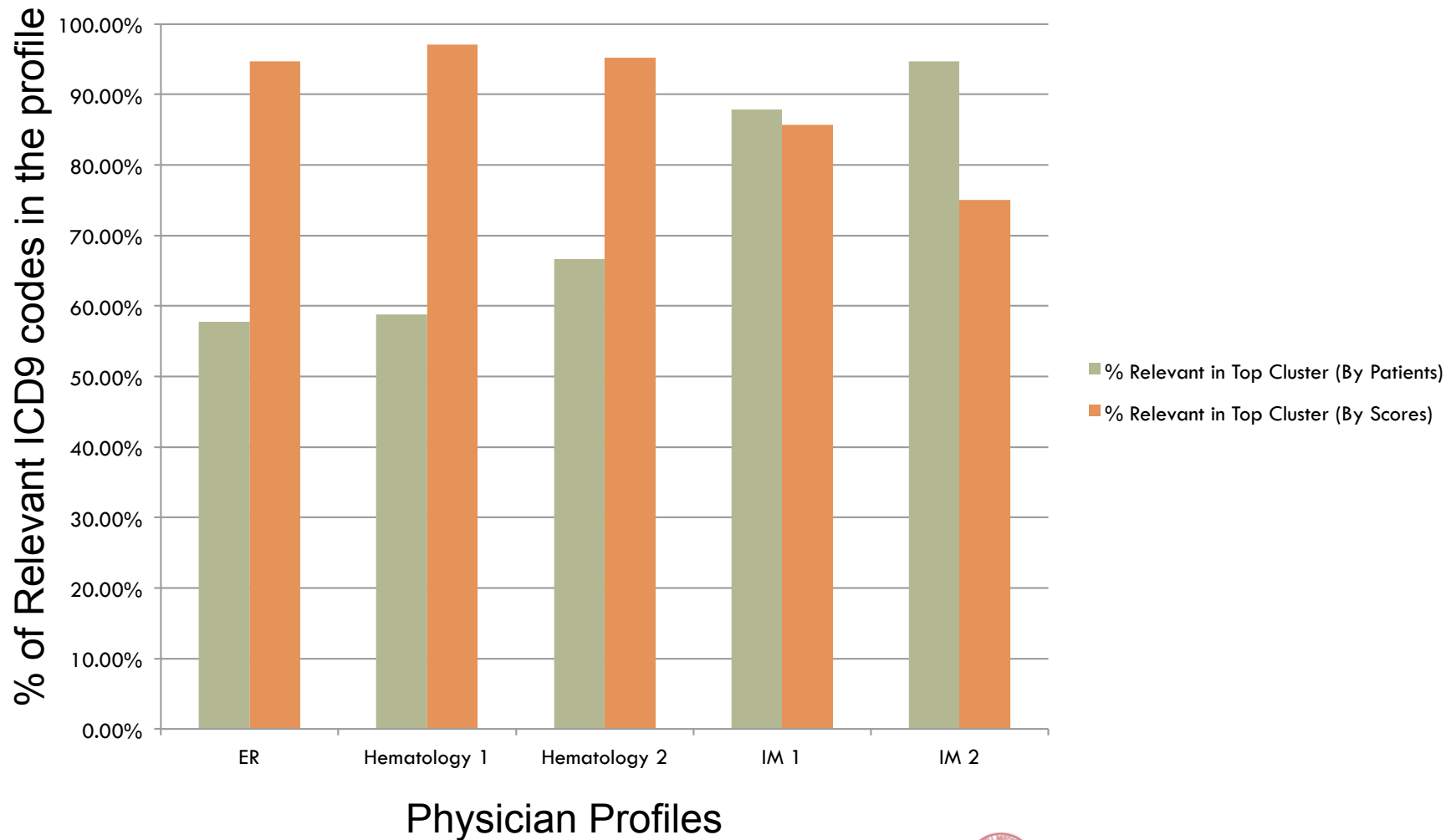
Screening for lipid disorders (cholesterol level) 23
Generalized Anxiety Disorder 39
Screening examination for viral diseases 8
Bipolar II Disorder
Mammographic microcalcification of breast
Schizophrenia, Undifferentiated Type
Shortness of Breath
Diabetes, Type II, Adult Onset
Panic Disorder without Agoraphobia
Anxiety state, unspecified



Clustering to extract profile



Resulting profile statistics



Sample profile

Immune thrombocytopenic purpura

Primary thrombocytopenia, unspecified

Other sickle-cell disease without crisis

Constitutional red blood cell aplasia

Sickle-cell thalassemia without crisis

Other thalassemia

Evans' syndrome

Constitutional aplastic anemia

Sickle-cell/Hb-C disease without crisis



Weill Cornell Medical College

VIVO

Physician ID list for a sample ICD9

ICD9 Code:157.3 – Pancreatic Cancer

12.91309	• 129219	HEMATOLOGY/ONCOLOGY
9.308263	• 113716	ONCOLOGY (MEDICAL); HEMATOLOGY; INTERNAL MEDICINE; ONCOLOGY; HEMATOLOGY; PEDIATRIC HEMATOLOGY/ONCOLOGY;
9.137091	• 118712	COLON AND RECTAL SURGERY; SURGERY
6.915659	• 119084	GASTROENTEROLOGY
6.33550	• 114357	SURGERY; GASTROENTEROLOGY
6.034863	• 115212 MEDICINE	HEMATOLOGY/ONCOLOGY - SOLID TUMOR; ONCOLOGY; HEMATOLOGY; INTERNAL
4.562387	• 115362	ONCOLOGY; HEMATOLOGY
4.494329	• 129763	ONCOLOGY; HEMATOLOGY
4.248922	• 130479	HEMATOLOGY/ONCOLOGY - SOLID TUMOR; ONCOLOGY; HEMATOLOGY

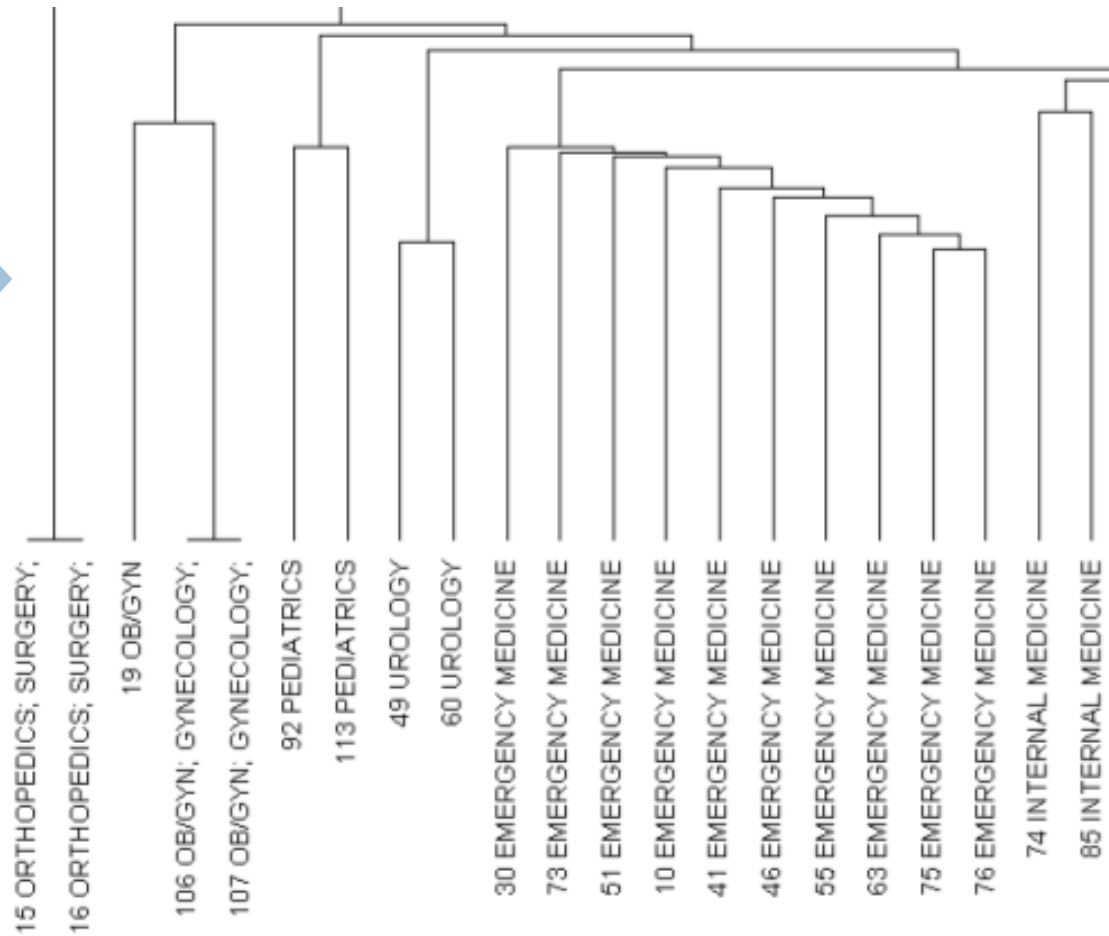


Weill Cornell Medical College

VIVO

Hierarchically Clustering Specialists

Euclidian
Distance in n-
dimensional
space



Weill Cornell Medical College

VIVO

Work in progress

- For a physician with 5 ICD9s (& patients), does each of the ICD9's constitute "20% of what she sees"
- Does the only physician who has seen one specific ICD9 become "institution's top specialist" in it?
- Are the physician's diagnosis and note correct?
- Calculating "intrinsic" ICD9 scores to separate "rare diseases" from "bicycle accidents"
- Normalizing and making scores "meaningful"



More data sources & applications

- Sources
 - ▣ Orders placed per diagnosis
 - ▣ Referrals
- Applications
 - ▣ Profiling based on medical procedures
 - ▣ Climbing meshes of terms instead of rooted trees

