VANDERBILT

Creating Interpretable Collaborative Patterns for Auditing

You Chen

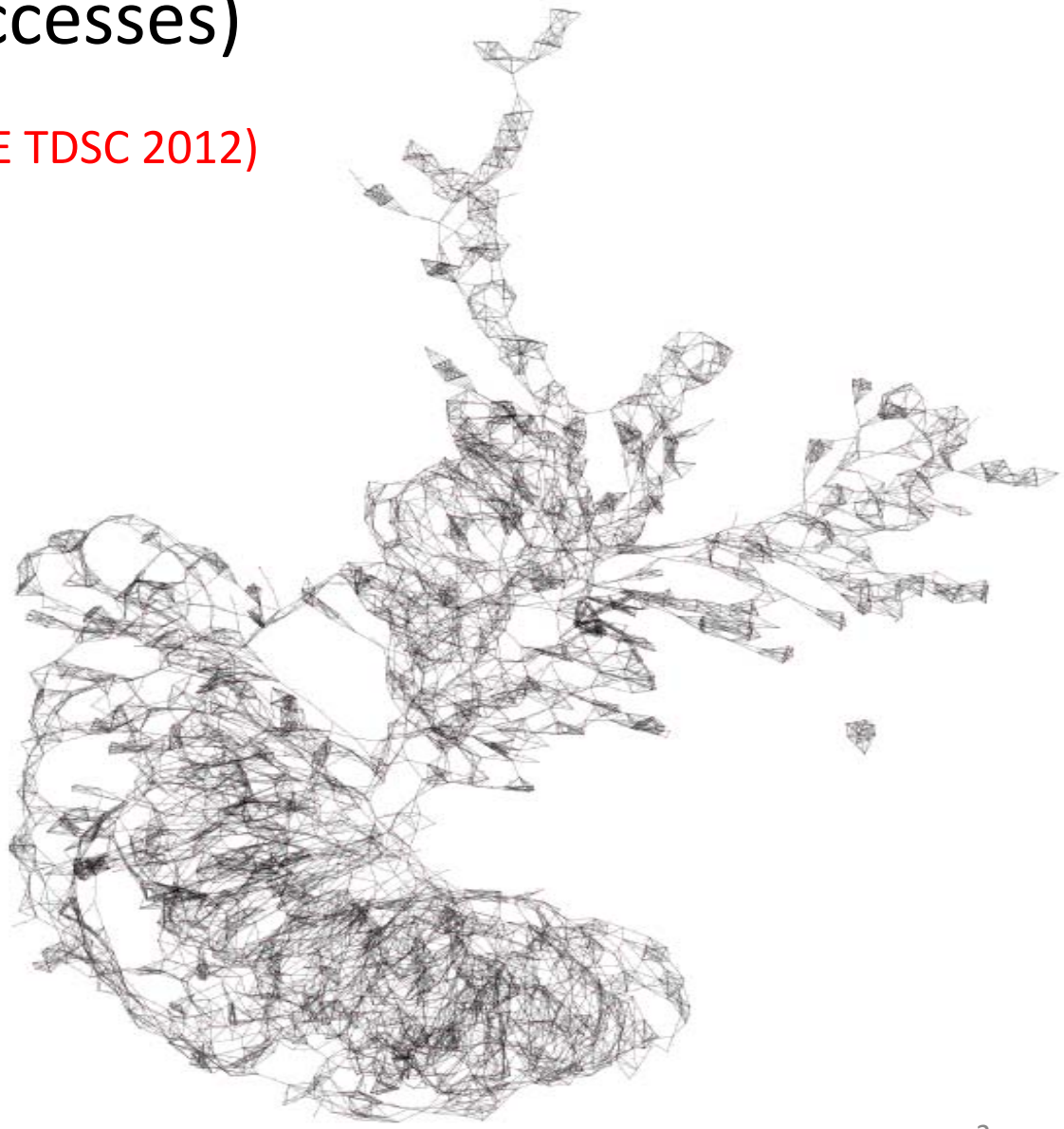
Department of Biomedical Informatics,
Vanderbilt University

You.chen@vanderbilt.edu

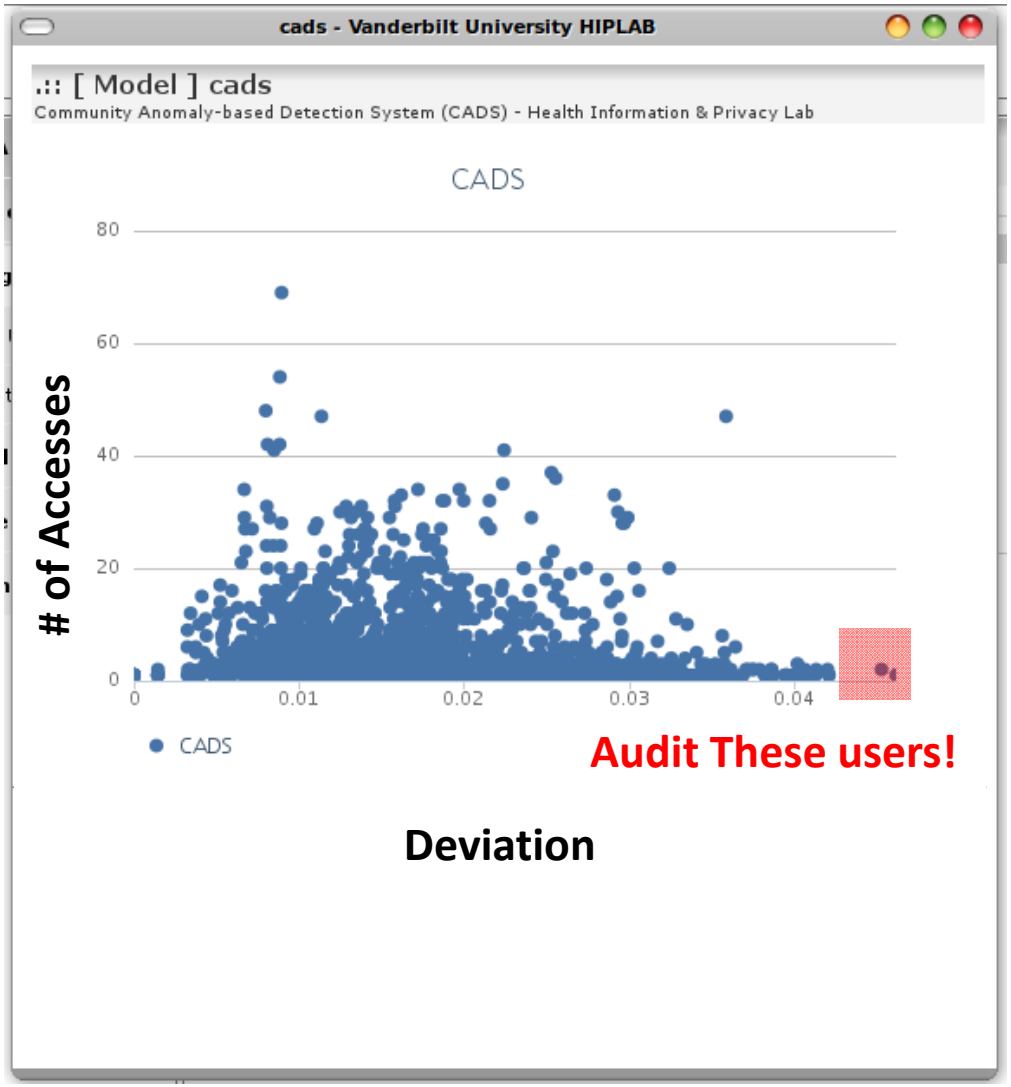
http://hiplab.org/~ychen

6-Nearest Neighbor Network-Vanderbilt Medical Center (1 day of accesses)

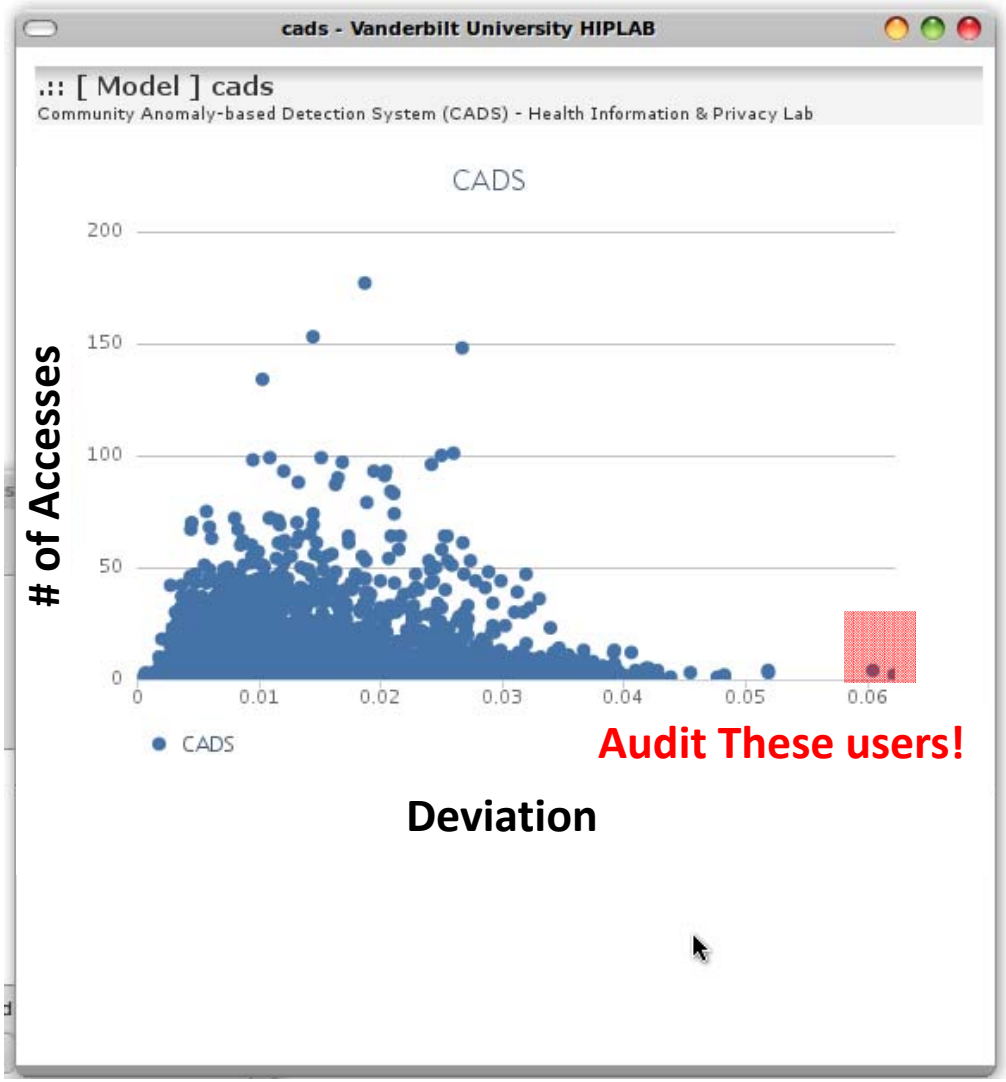
(Chen, Nyemba, & Malin – IEEE TDSC 2012)



CADS on Vanderbilt Dataset

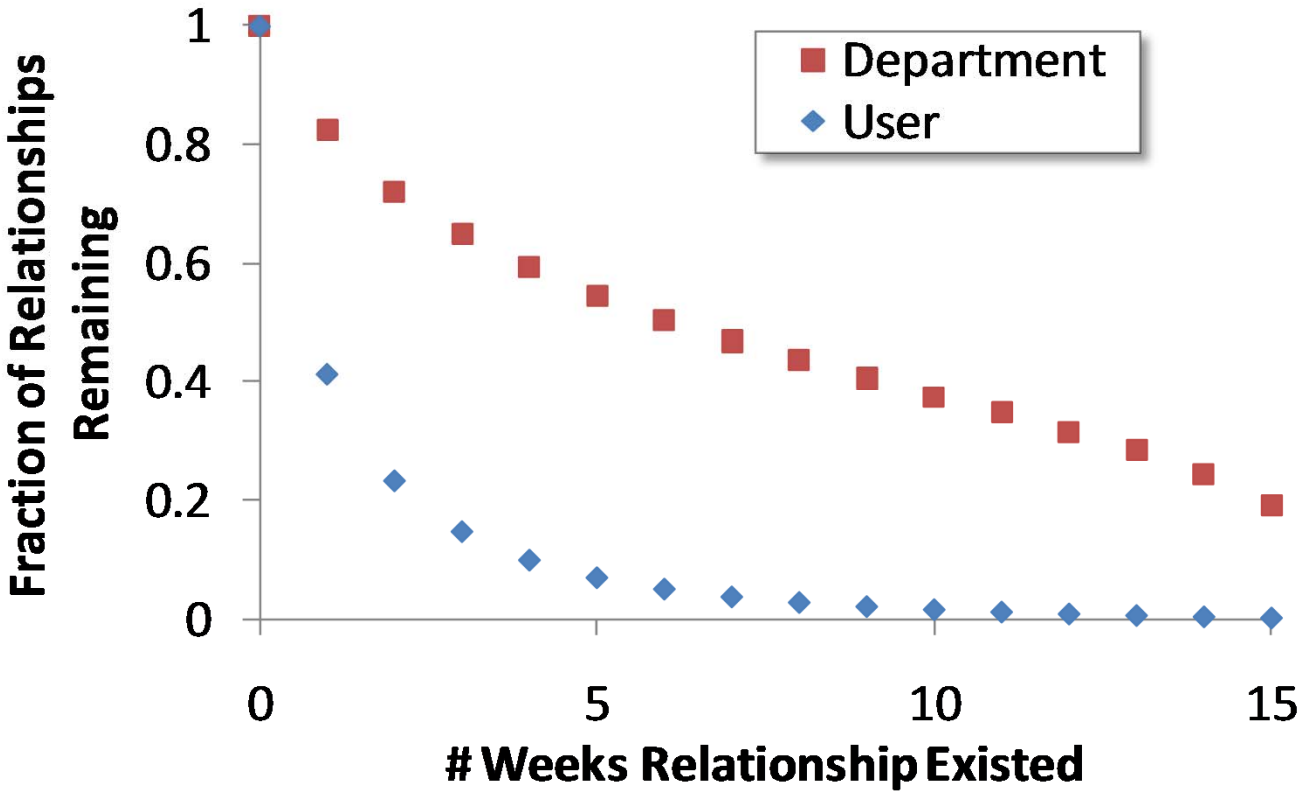


CADS on Northwestern Dataset



But Relationships Decay…

(Malin, Nyemba, & Paulett – JBI 2011)



- EMR <user, user> relationships
- EMR <department, department> relationships

Department Level

Auditing Medical Record Accesses through Department Interactions

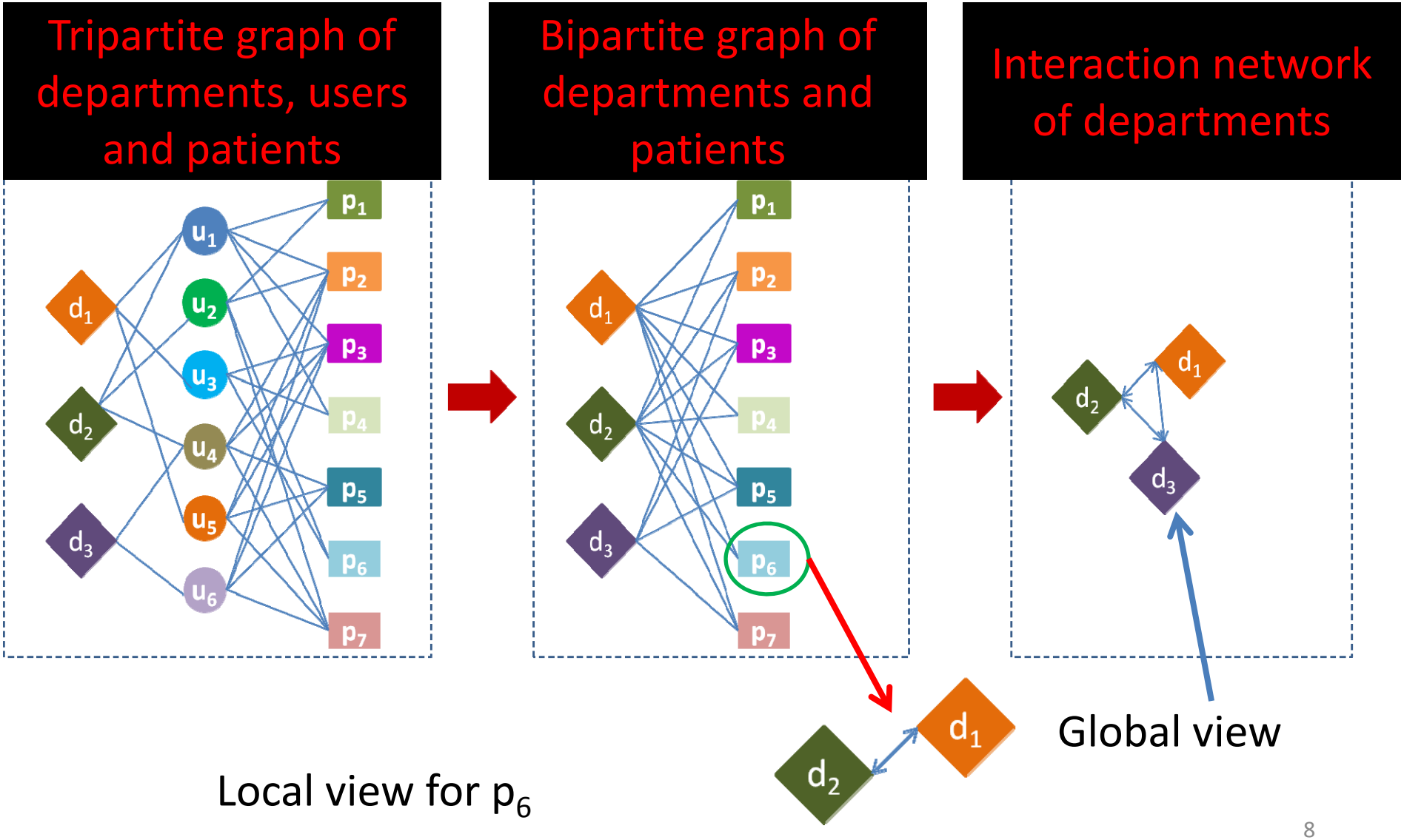
Hospital Departmental Relations Can Be Inferred

(Chen, Nyemba, & Malin - AMIA 2012)

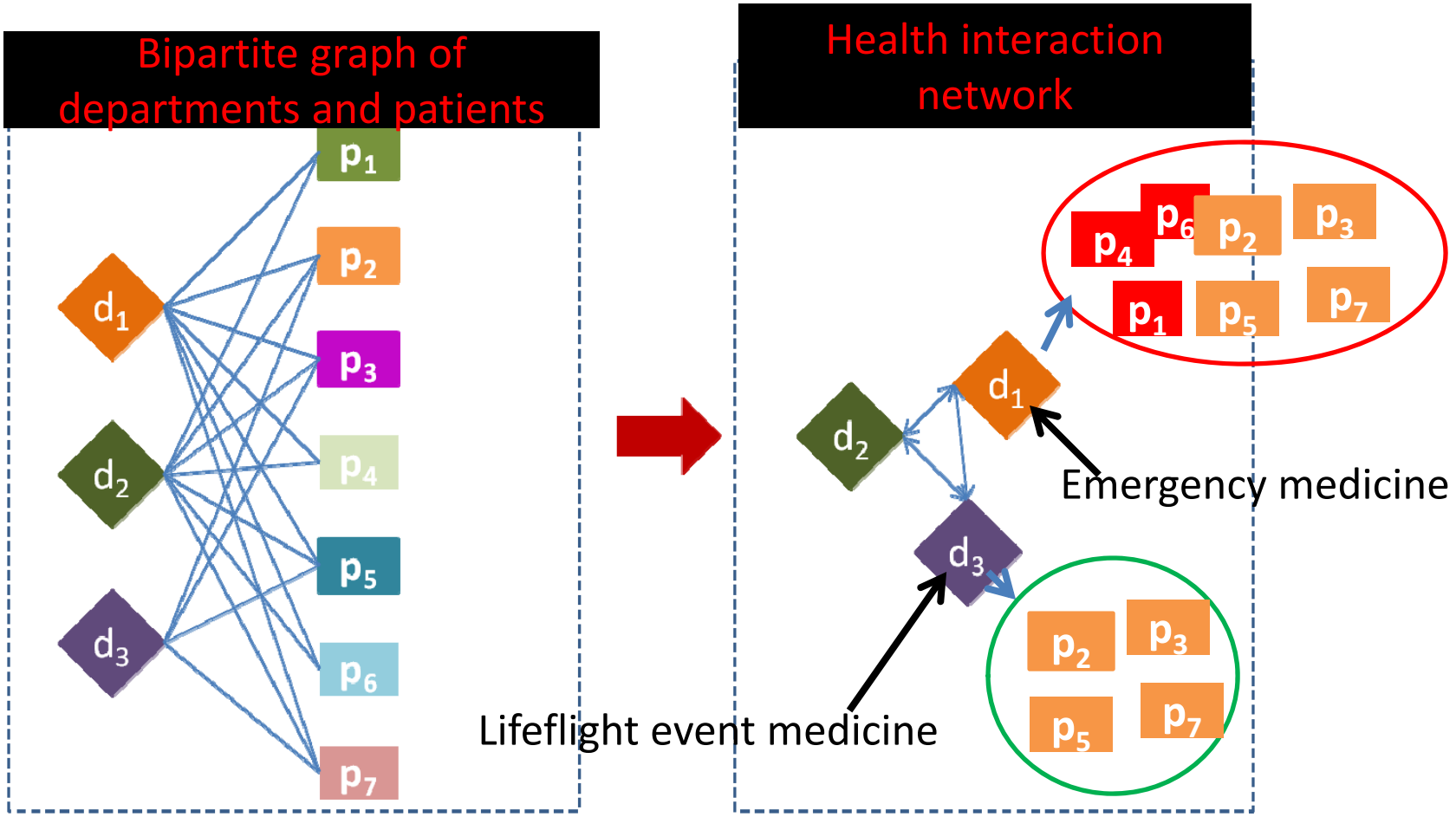
- Probability department d_i accesses a patient's record, given department d_j accessed the record.

Department (d_i)	Department (d_j)	Min Certainty	Max Certainty
Intradepartmental Relations			
4East OB/GYN	4East OB/GYN	0.74319	0.7669
Adult Emergency Medicine	Adult Emergency Medicine	0.74024	0.78453
Cancer Infusion Center	Cancer Infusion Center	0.73171	0.844
8N Inpatient Medicine	8N Inpatient Medicine	0.7197	0.80909
Newborn Nursery	Newborn Nursery	0.70406	0.72727
Interdepartmental Relations			
DOT Radiology	Orthopaedics	0.99621	1
Nursing Education and Development	Medical Information Services	0.95833	1
Main OR - Trauma/Renal	Medical Information Services	0.94444	1
Life Flight Event Medicine	Emergency Medicine	0.90805	1
Emergency Medicine Admin	Adult Emergency Medicine	0.91489	0.94186

Organization Level-Department

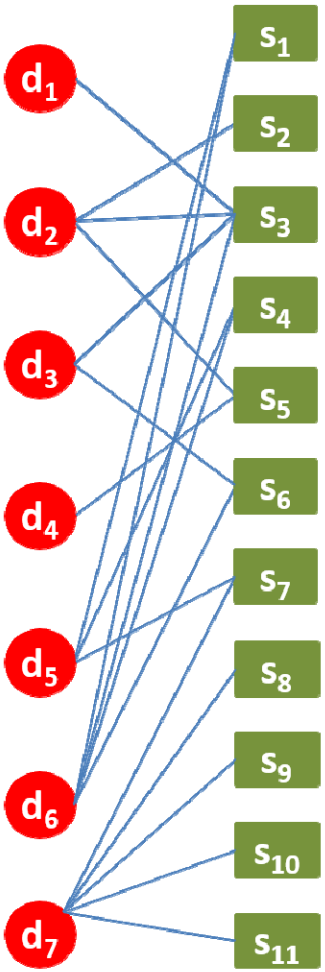


Certainty to Model Relationship Among Departments

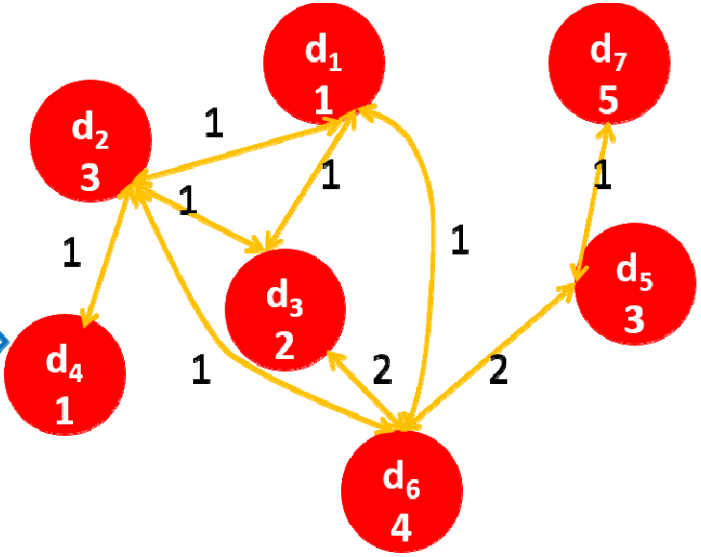


Cert(Emergency medicine (d_1)->Lifeflight event medicine (d_3)) = 4/7

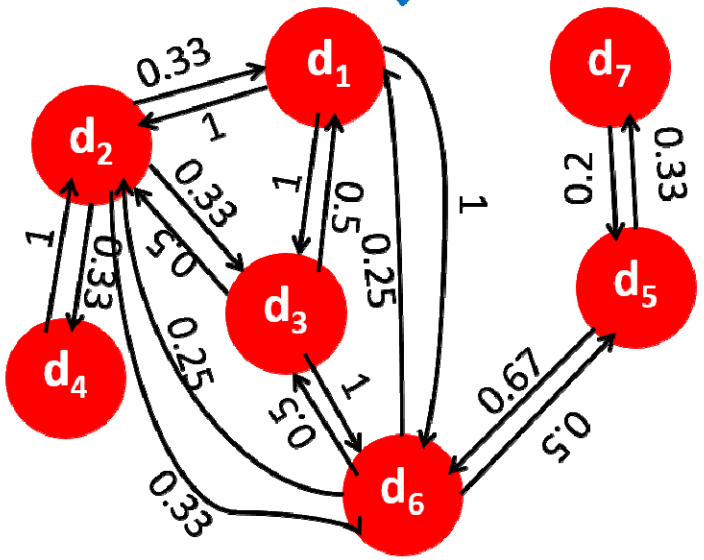
Lifeflight event medicine (d_3)-> Cert(Emergency medicine (d_1)) = 4/4



(a) A bipartite graph of departments and subjects

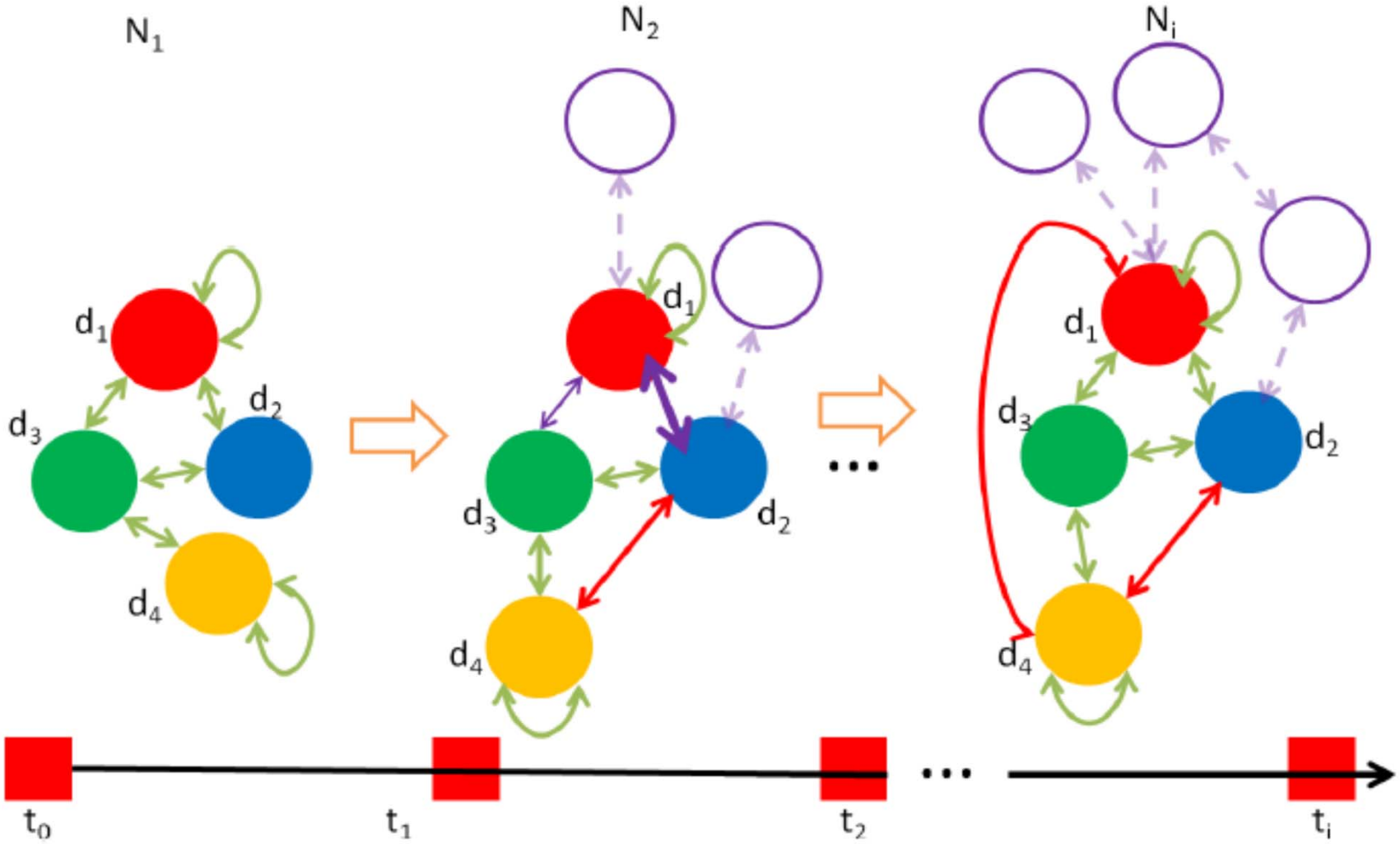


(b) A social graph of departments based on common subjects accessed

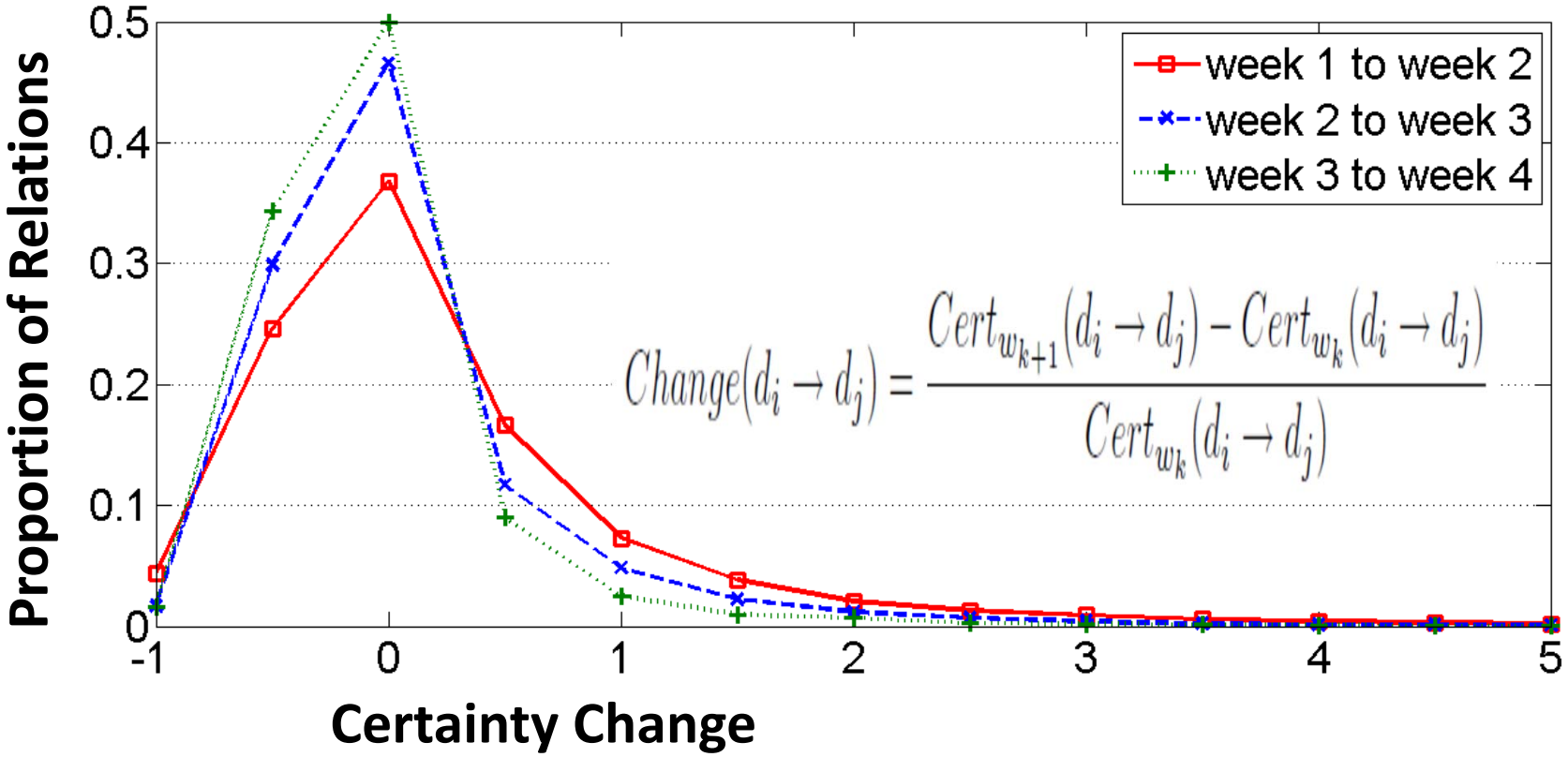


(c) A confidence social network of departments

Evolution of A Global Network Over the Time



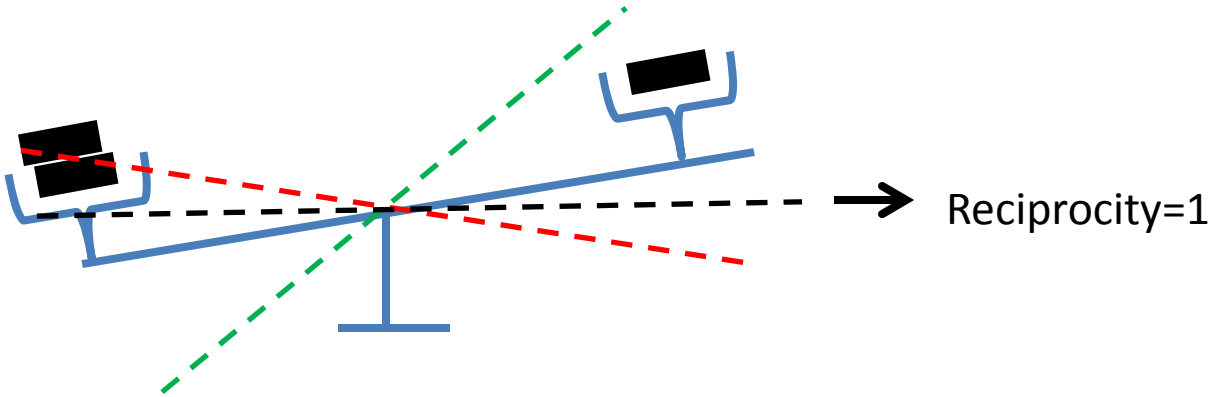
The changes become smaller over time (centralization: green > blue > red)



Degree of relations between departments changes little over time

>82.5% of the change resides in [-0.25, 0.25]

Using reciprocity to characterize the mutual interaction between all pairs of departments in the global network



Inpatient Admin -> VUH Admitting 0.75

VUH Admitting -> Inpatient Admin 0.12

VANDERBILT UNIVERSITY

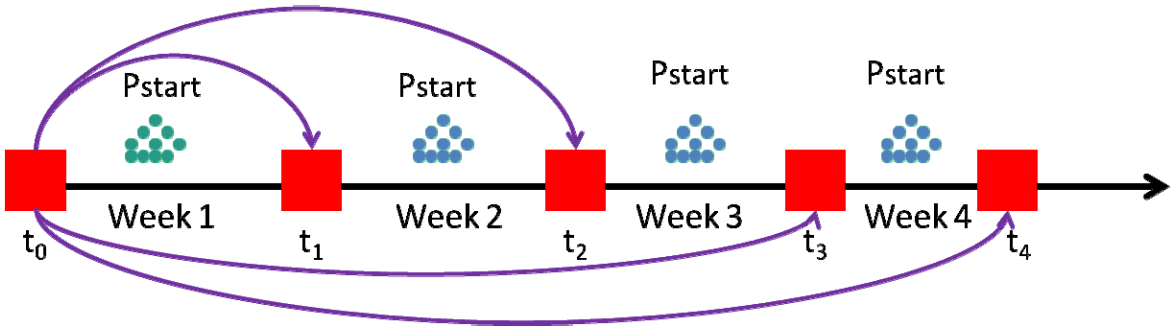
Although the relations of the network are very unbalanced, the unbalance is stable over time

Time	Week 1	Week 2	Week 3	Week 4
Reciprocity	0.267	0.2814	0.2858	0.2871

$$Reciprocity = \frac{\sum_{\forall d_i, d_j \in D, i \neq j} \left| (Cert(d_i \rightarrow d_j) - a) \times (Cert(d_j \rightarrow d_i) - a) \right|}{\sum_{\forall d_i, d_j \in D, i \neq j} (Cert(d_i \rightarrow d_j) - a)^2}$$

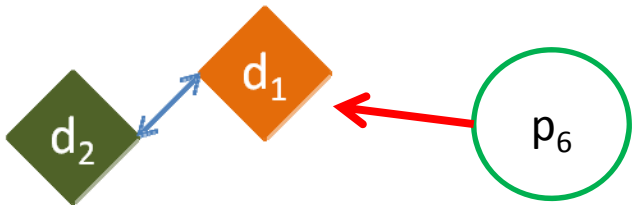
$$a = \frac{\sum_{\forall d_i, d_j \in D, i \neq j} Cert(d_i \rightarrow d_j)}{2 \times |E|}$$

Evolution of Local Network Relations Can be Used Detect "Strange" Behavior



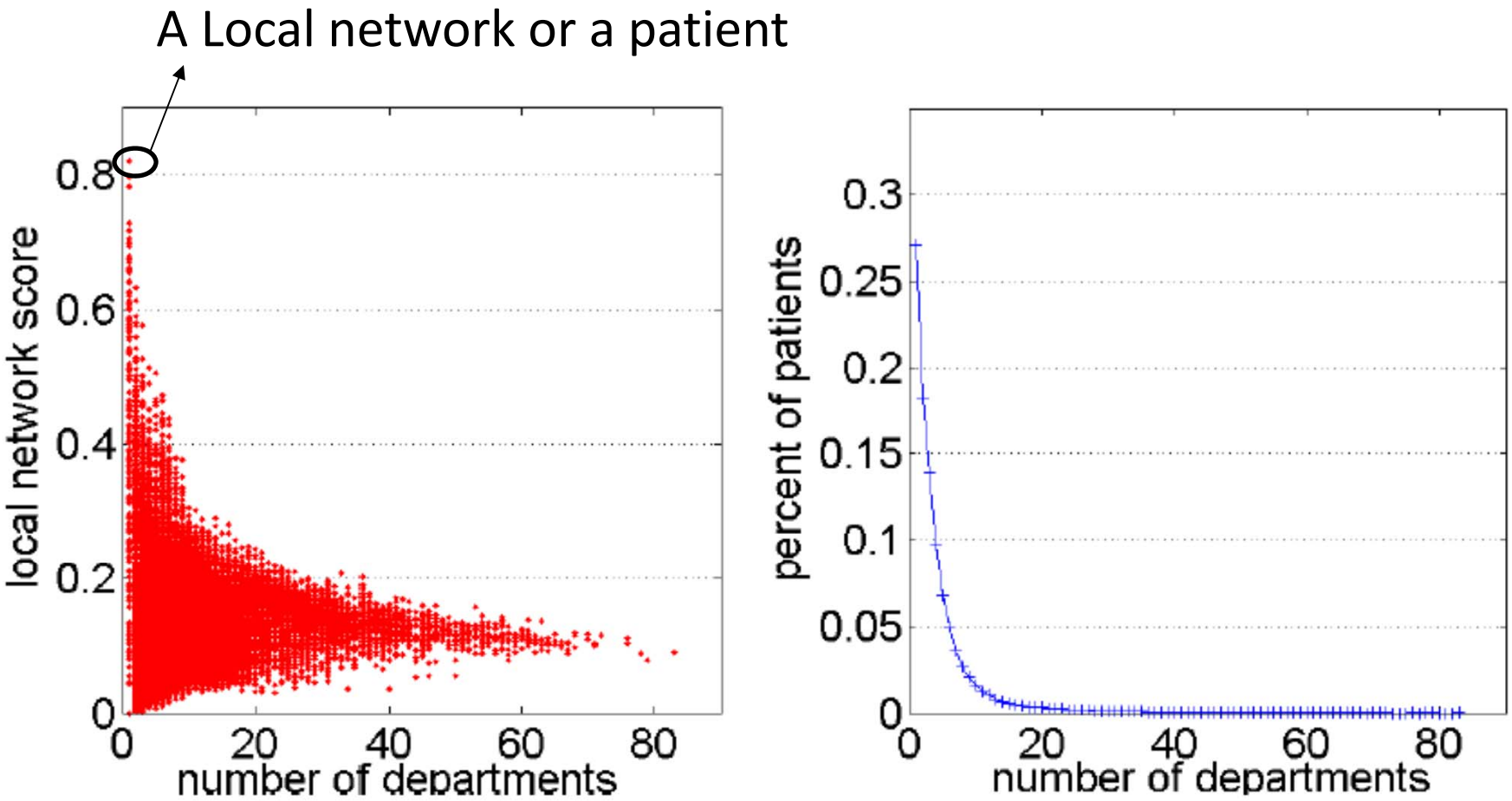
Each point in the P_{start} corresponds to a local network

A Local network for p_6

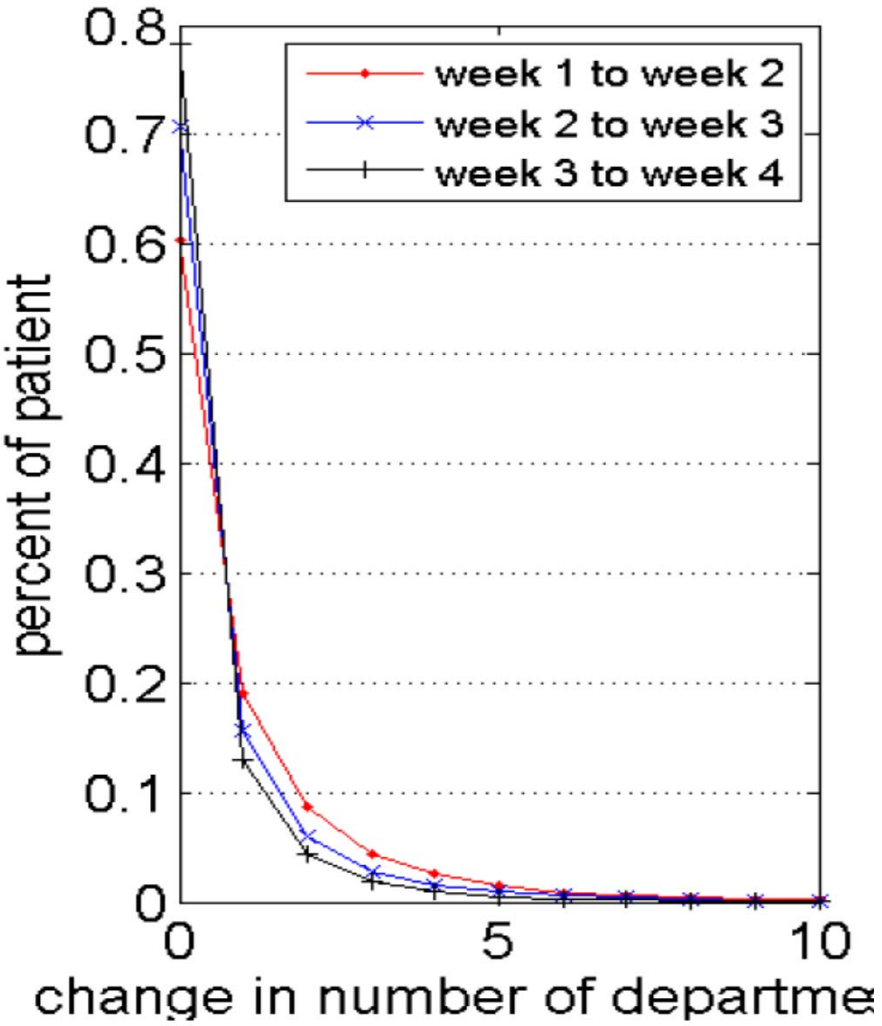


$$Score(p_k) = \frac{\sum_{\forall d_i, d_j \in D_k} Cert(d_i \rightarrow d_j)}{2 \times |E_k|}$$

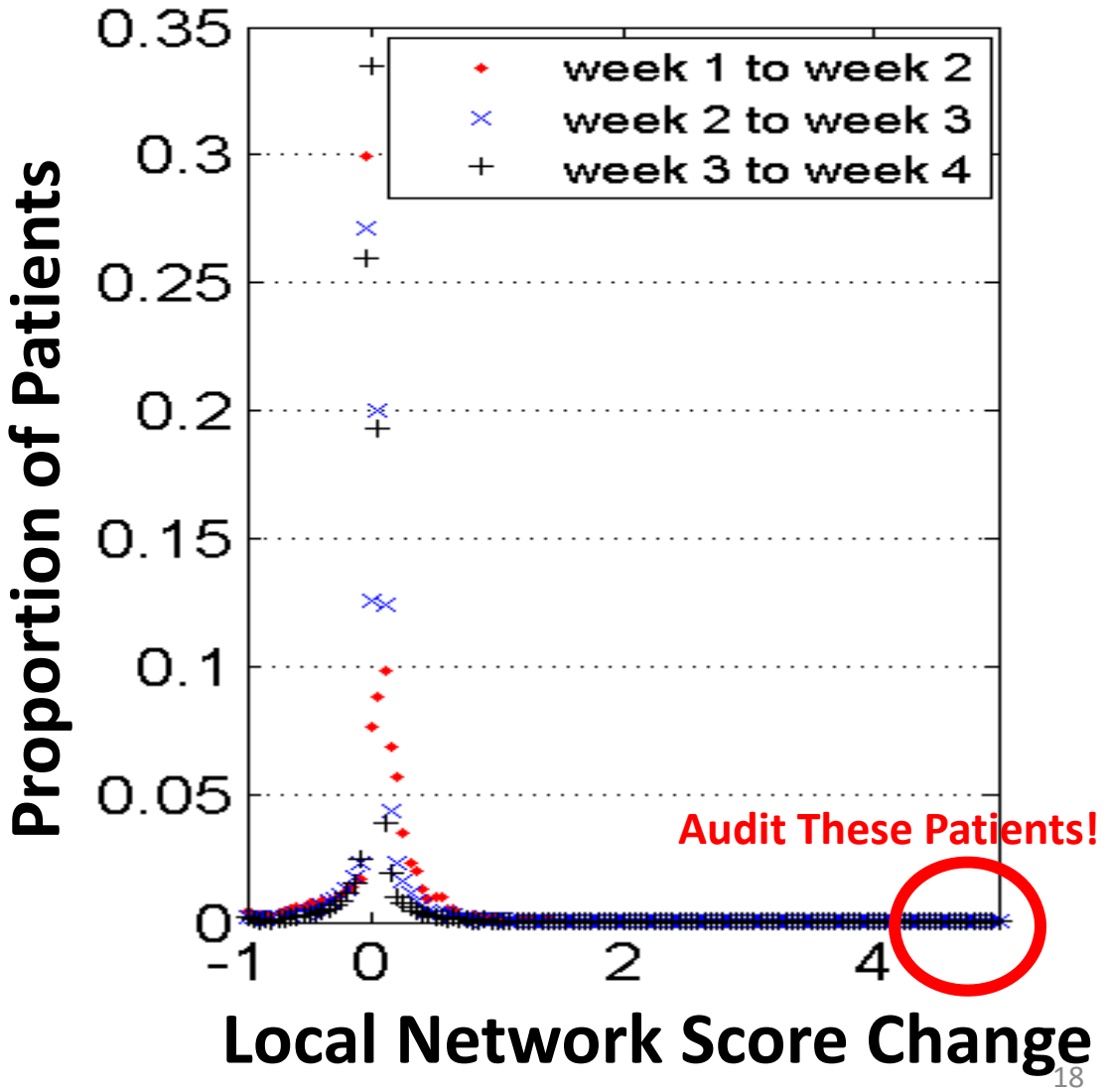
Over 80% of local networks whose size is less than 5



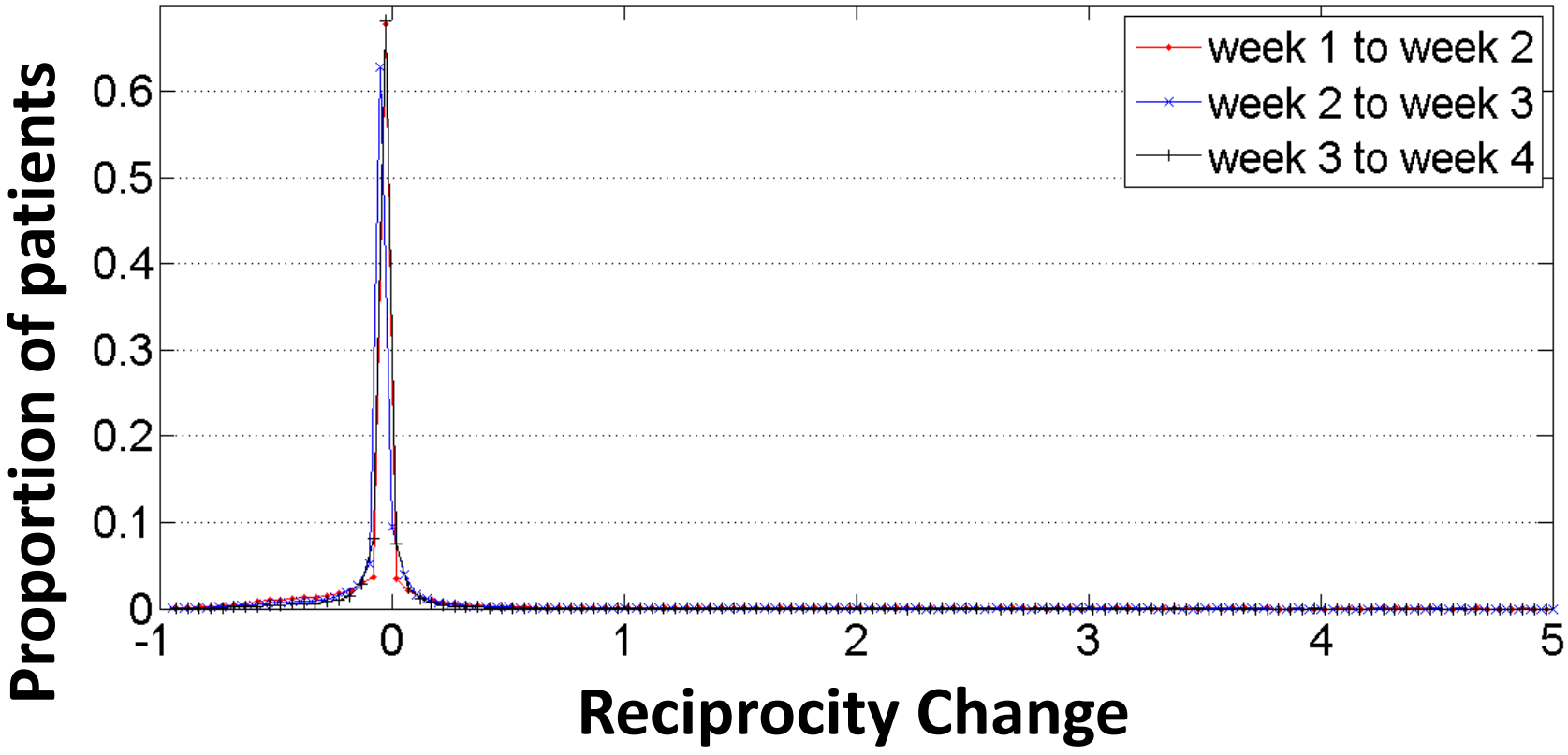
Over 80% of local networks has number of departments change less than 2



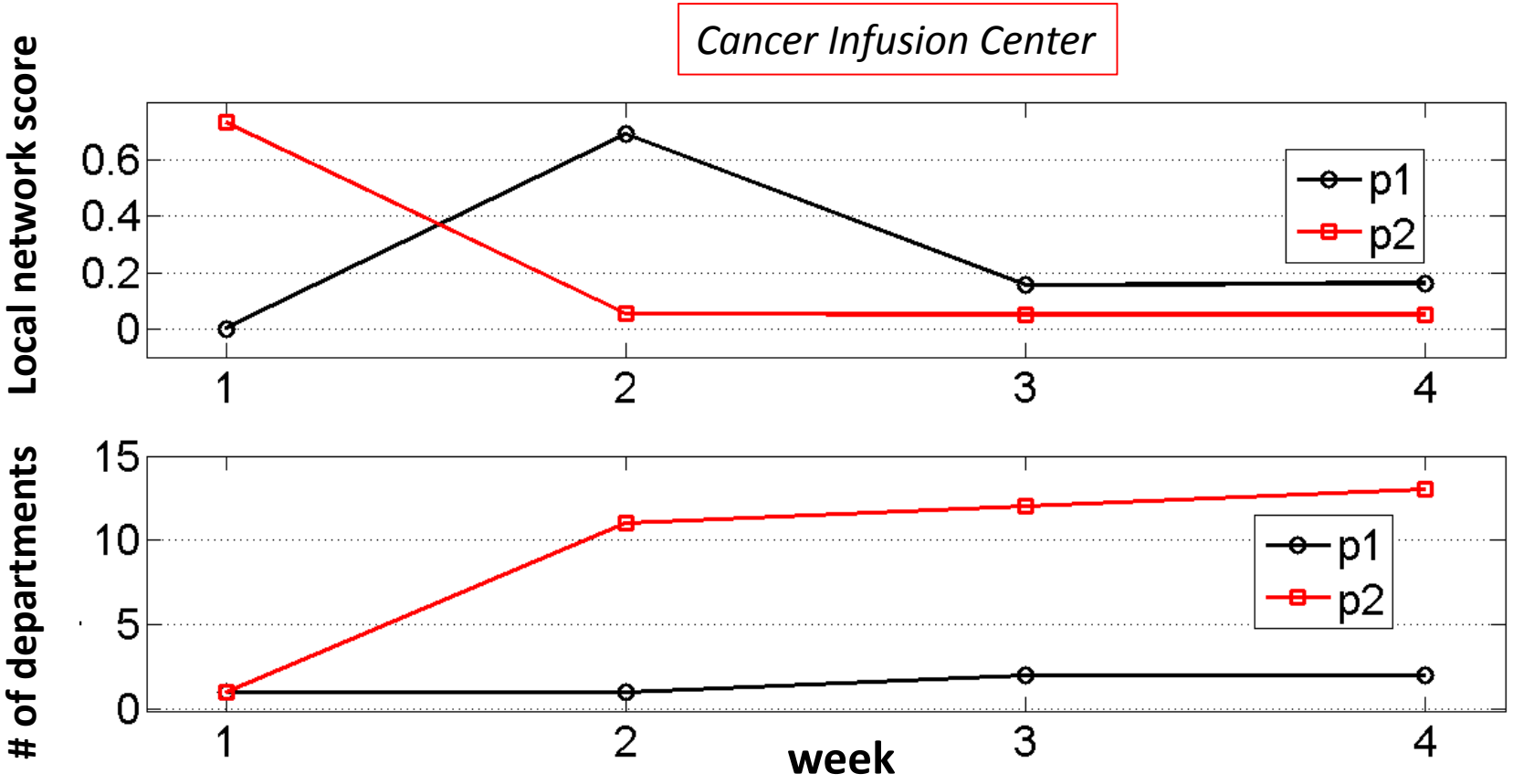
Most Patients Network Suggest They Are "Normal"



Approximately 99% of patients are normal because they have a change of reciprocity <0.1



p2 has -0.93 change of local network score and -0.79 change of local reciprocity from the 1st to the 2nd week



Breast Center, [Anonymized Street Location], Care/Eskind Diab Acces, Disease Management Service, Eskind Diabetes - Adult, Free Stipends, Internal Medicine, VIM, VMG Physician Billing Services, Vanderbilt Home Care Primary

Findings

- We hypothesized an HCO would exhibit strong stability

  → **confirmed by our experiments**

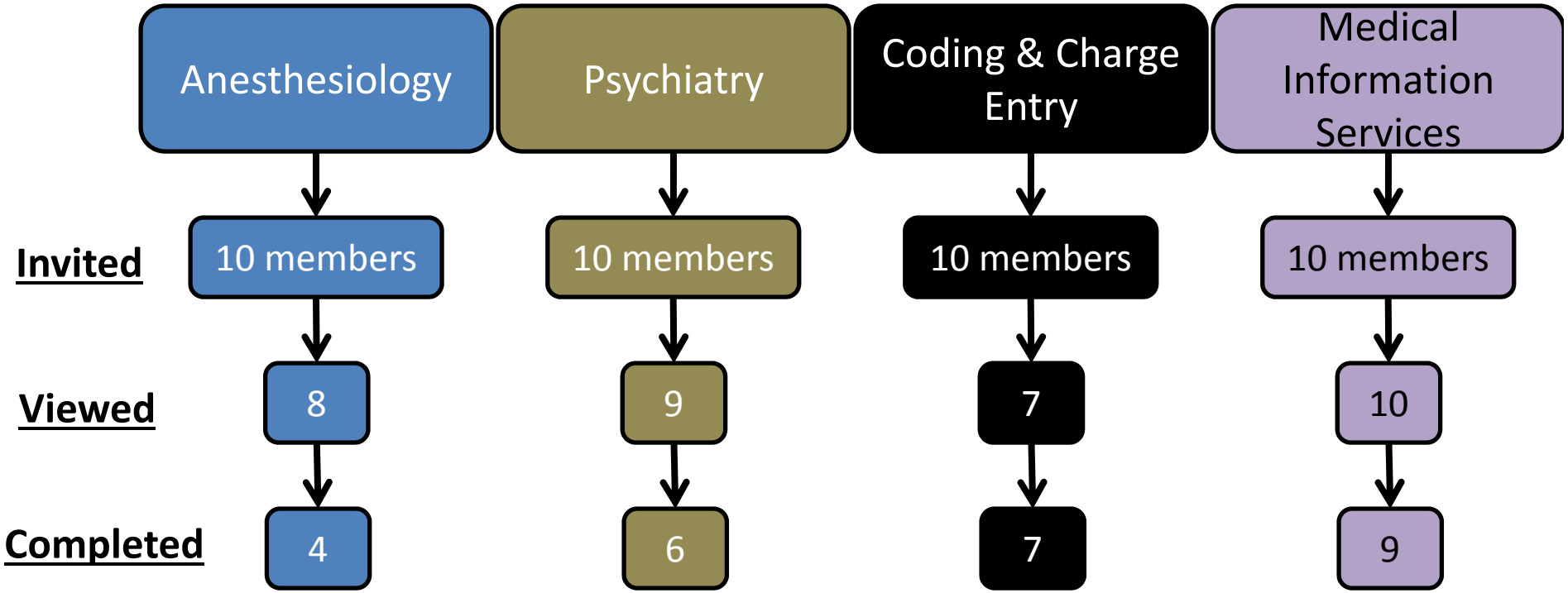
- The changes in the score of local networks do not justify the claim that the patient has been intruded upon, but may provide a reason for an investigation that incorporates more nuanced domain knowledge

But Do You Believe the Data?

Survey Population

(Chen, Lorenzi, Nyemba, Schildcrout & Malin – IJMI 2014)

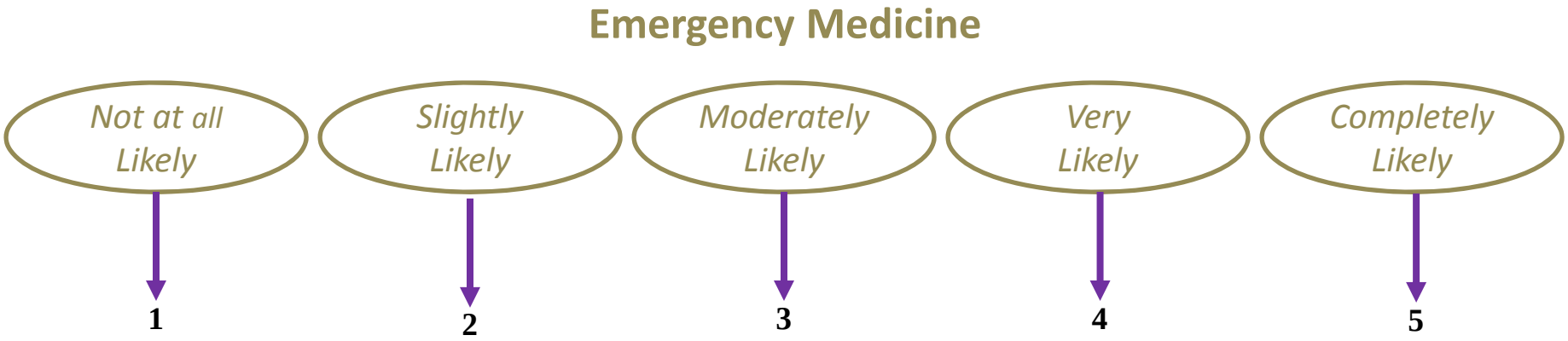
- Vanderbilt University Medical Center areas



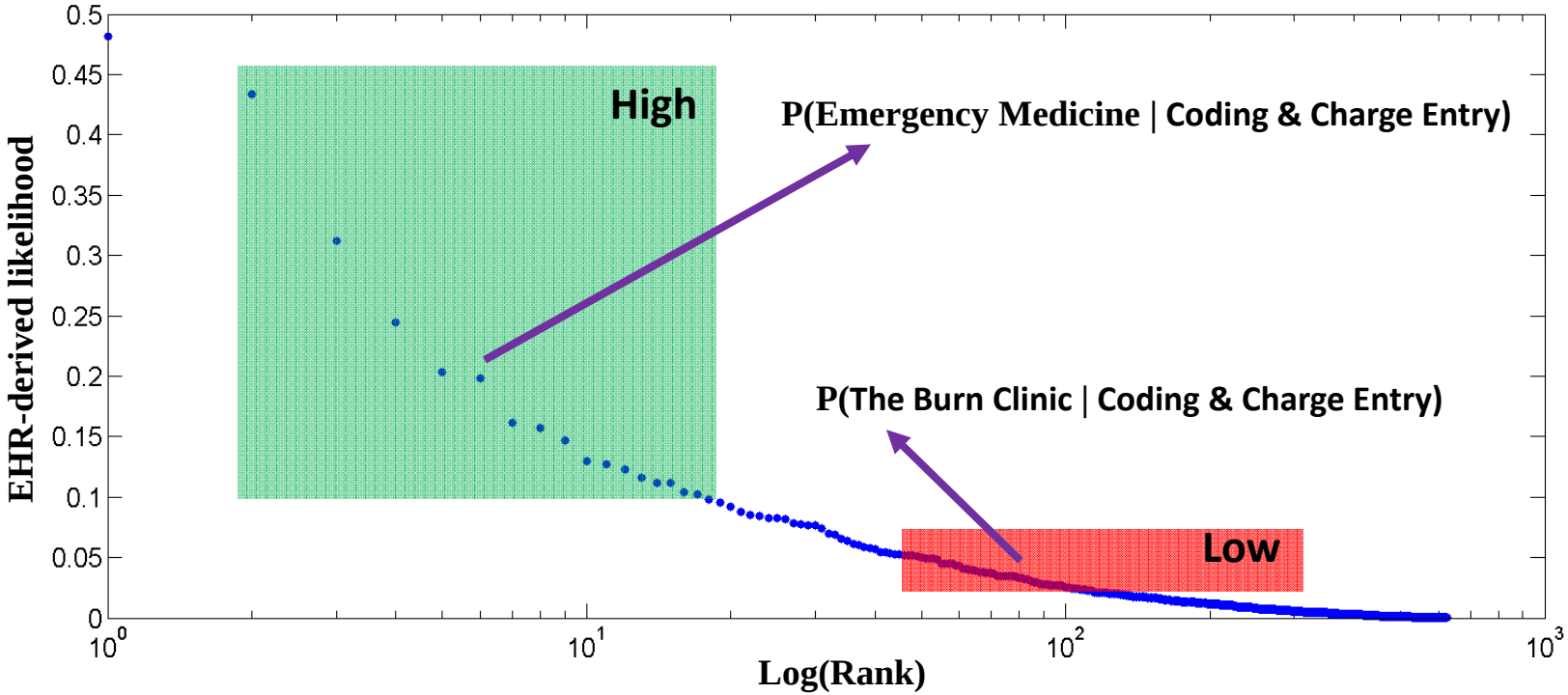
34 respondents did the survey and 26 of them are valuable

Survey Questions

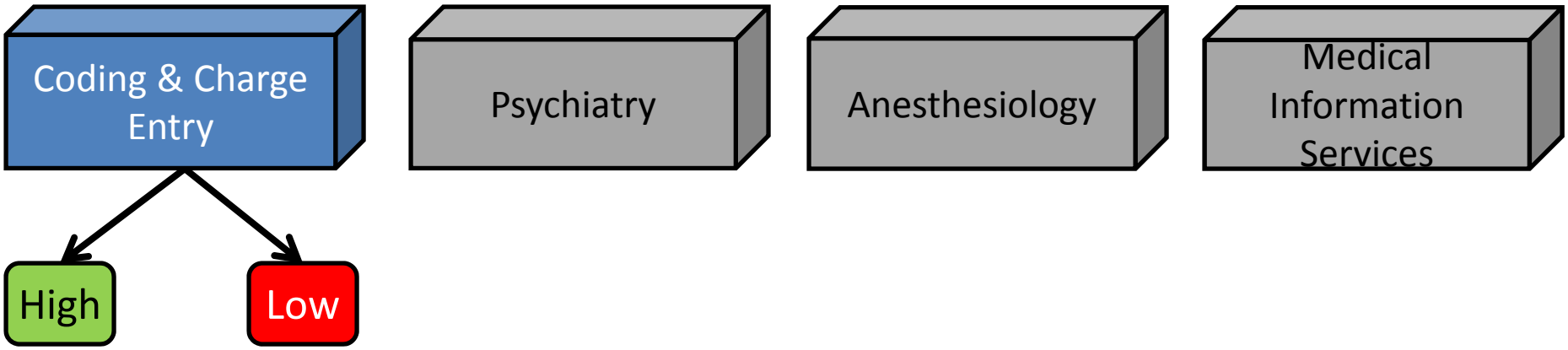
- Departmental interactions
- Conditional probabilities of accessing a record (conditioned on the HCO area)
- "Given someone from Coding & Charge Entry accessed the record, *what's the chance someone from the following Area accessed the record*?



Coding & Charge Entry Interactions
(one week, ~620 points)

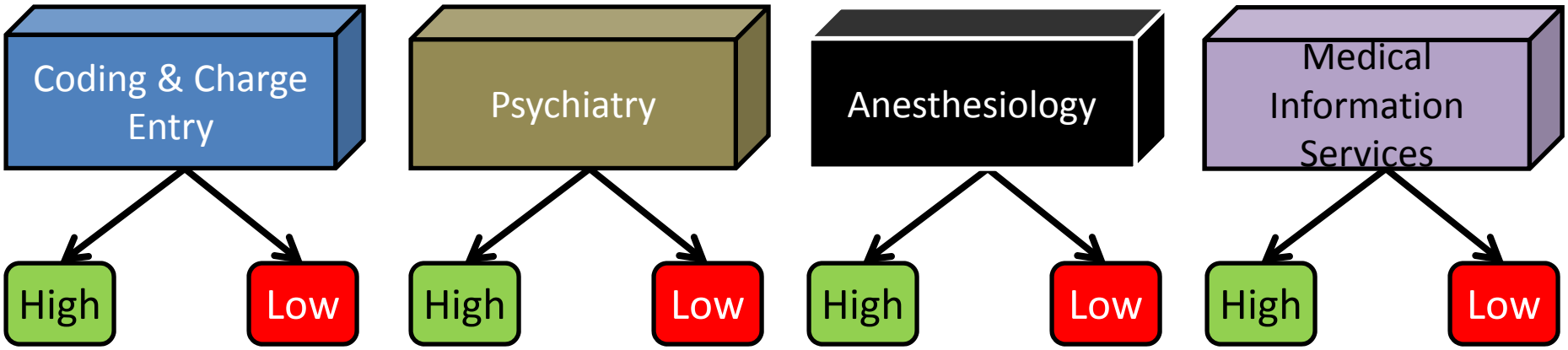


Survey Questions

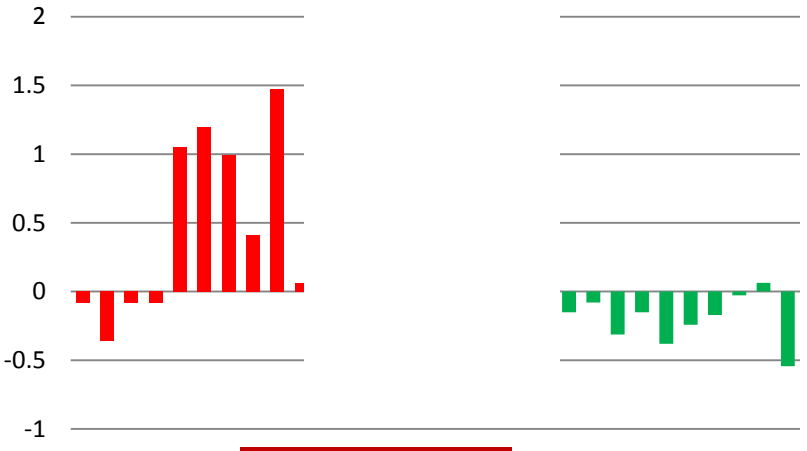


20 rules

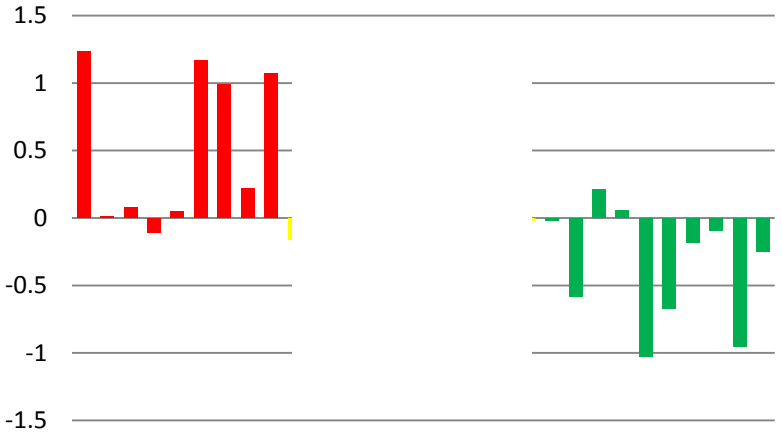
Survey Questions



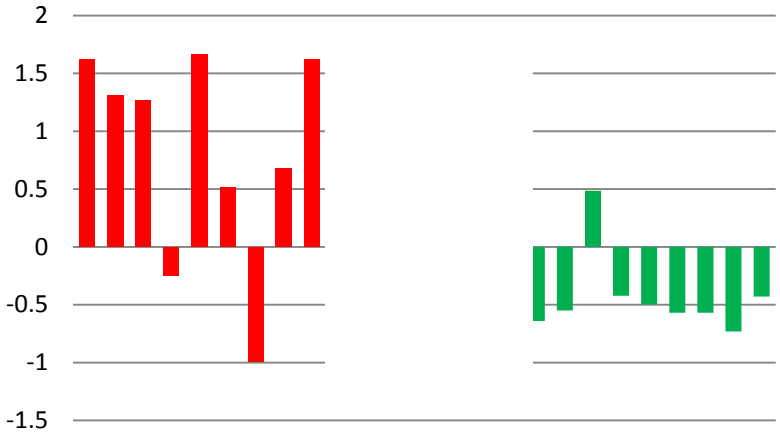
MIS RESPONDENT(10)



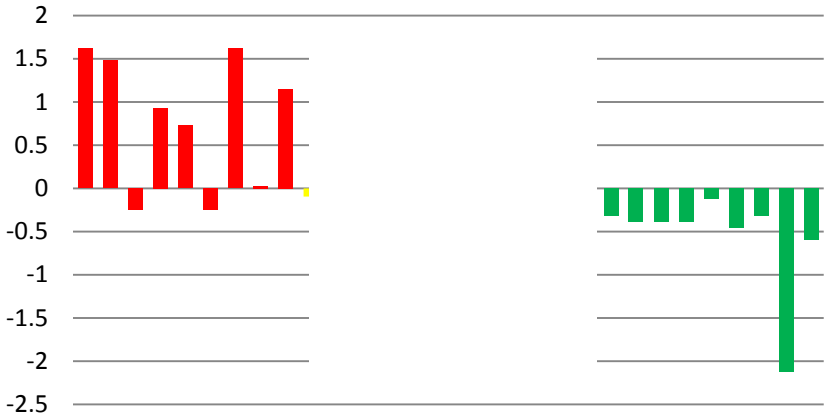
MIS Rules



CODE Rules

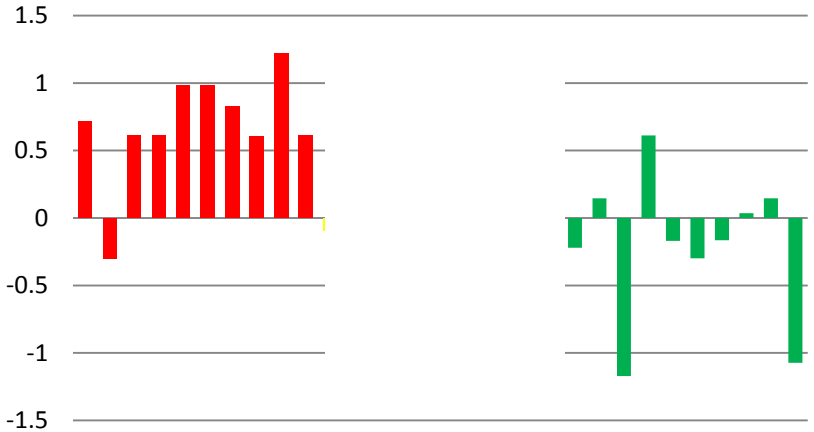


PSY Rules

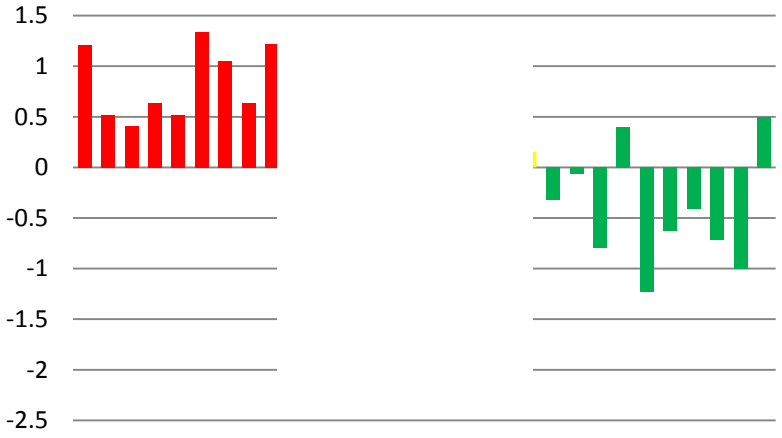


ANE Rules

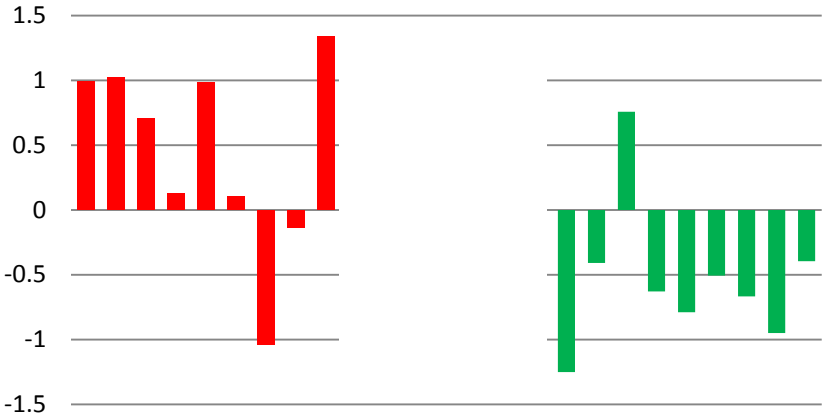
CODE RESPONDENT(7)



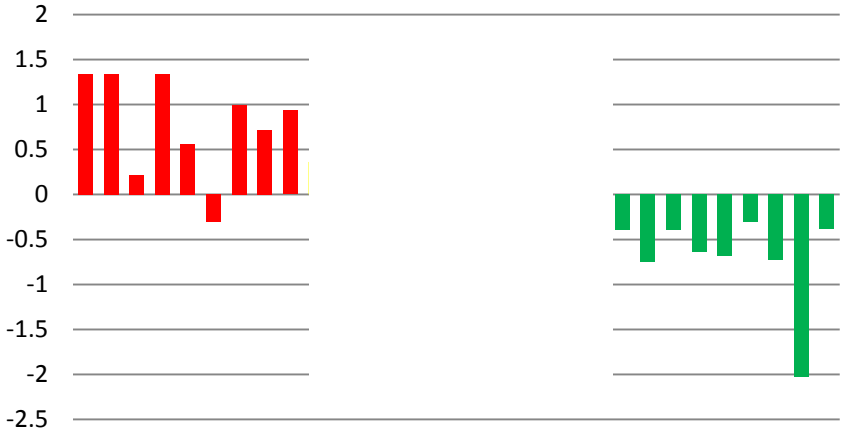
MIS Rules



CODE Rules

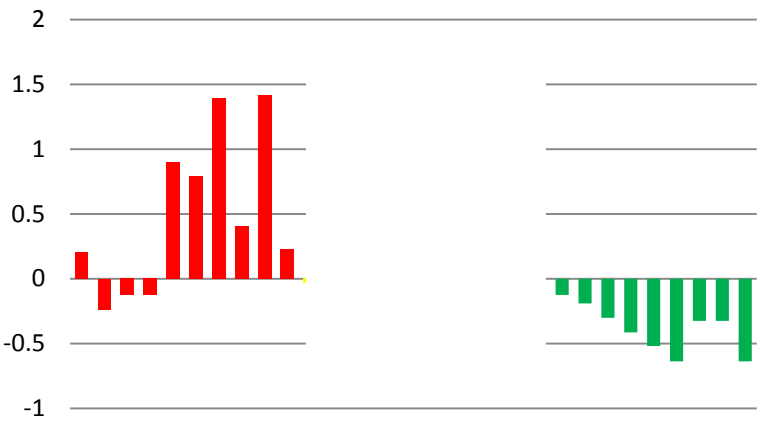


PSY Rules

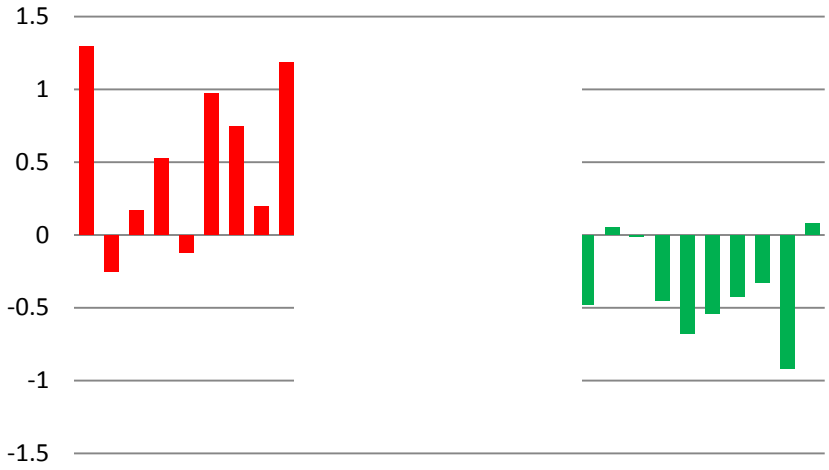


ANE Rules

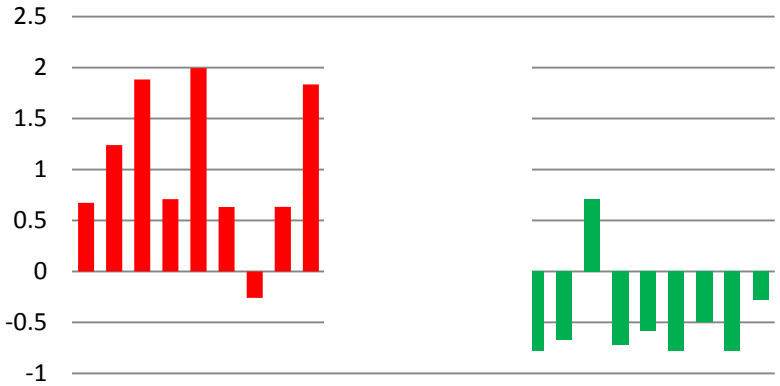
PSY RESPONDENT(9)



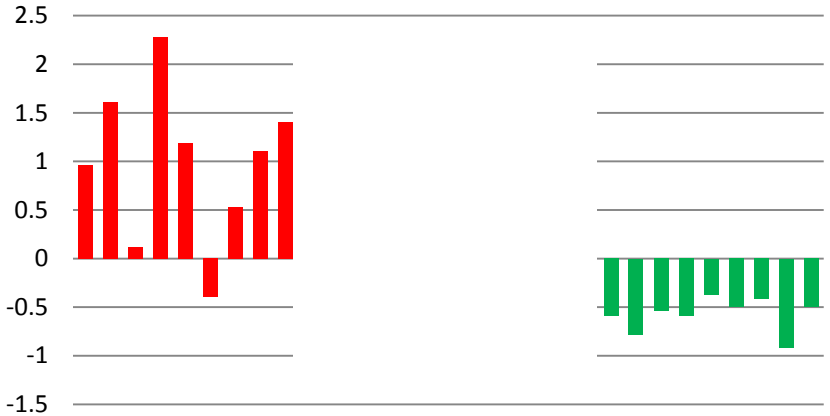
MIS Rules



CODE Rules

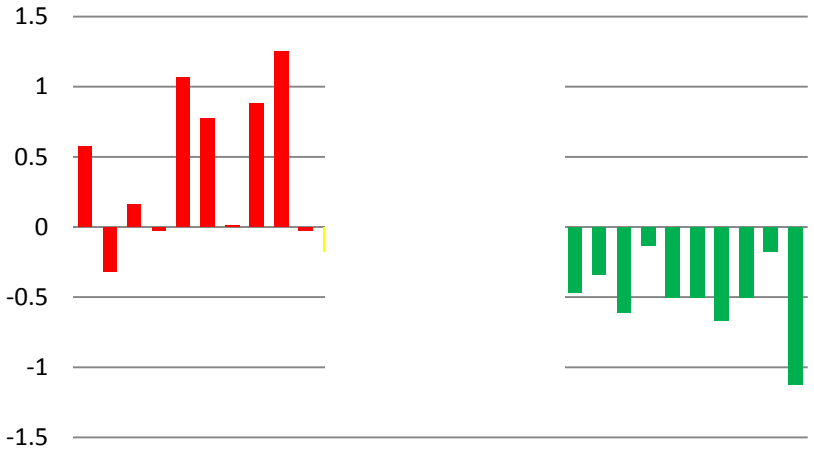


PSY Rules

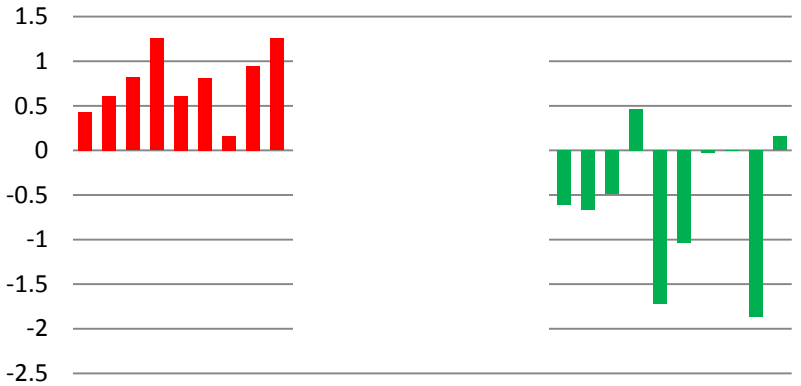


ANE Rules

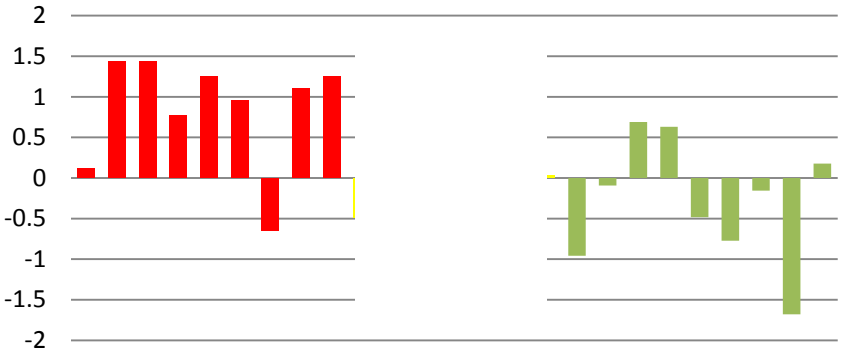
ANE RESPONDENT(8)



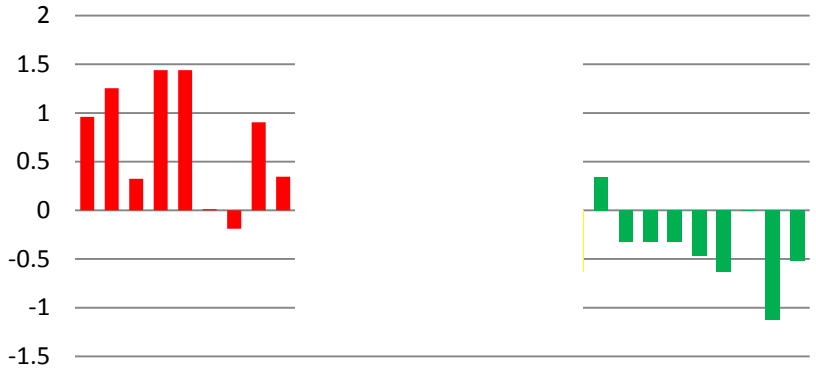
MIS Rules



CODE Rules



PSY Rules



ANE Rules

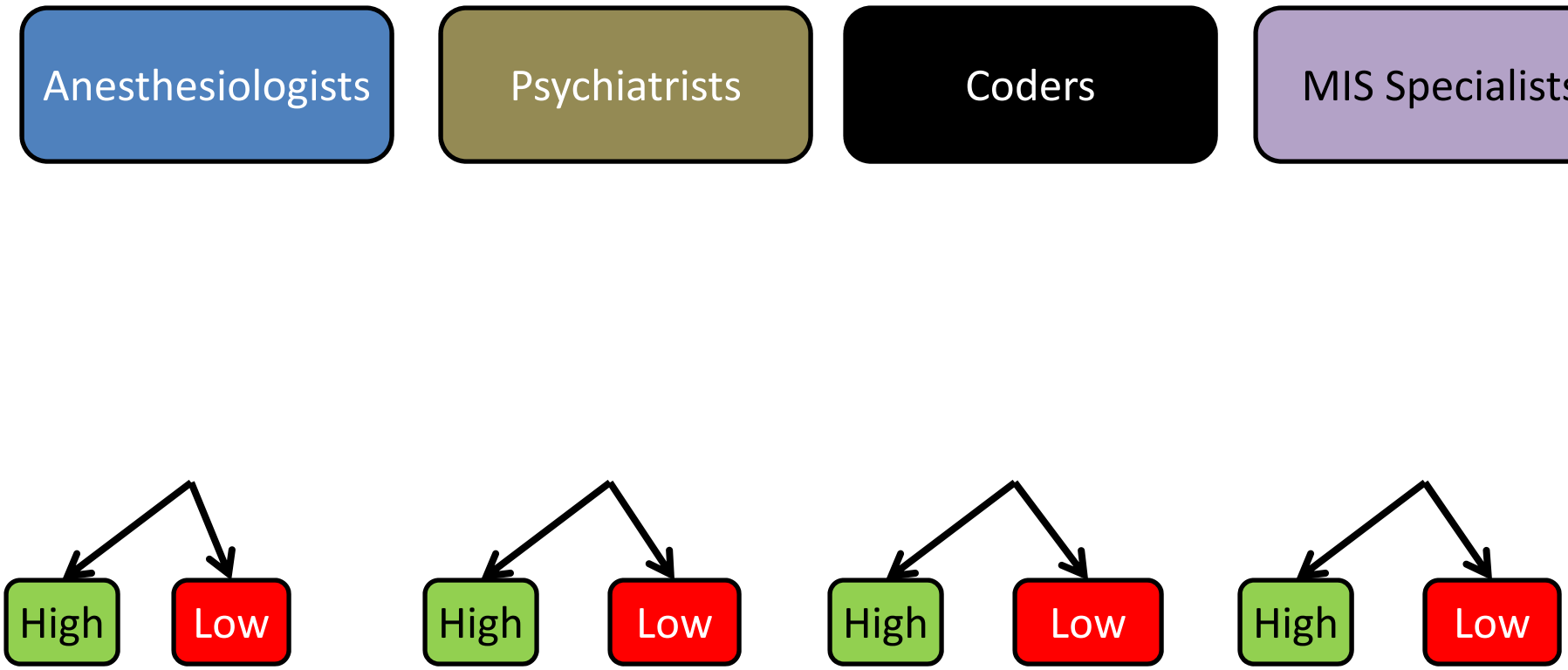
Hypothesis

- 1) Employees can distinguish between high, and low likelihood rules for all HCO areas
- 2) Employees can distinguish between high and low likelihood rules for their own HCO area
- 3) employees can distinguish between high, and low likelihood rules in their own HCO area better than they can in other HCO areas

One respondent has 8 observations
The total number of observations is 8*26=208

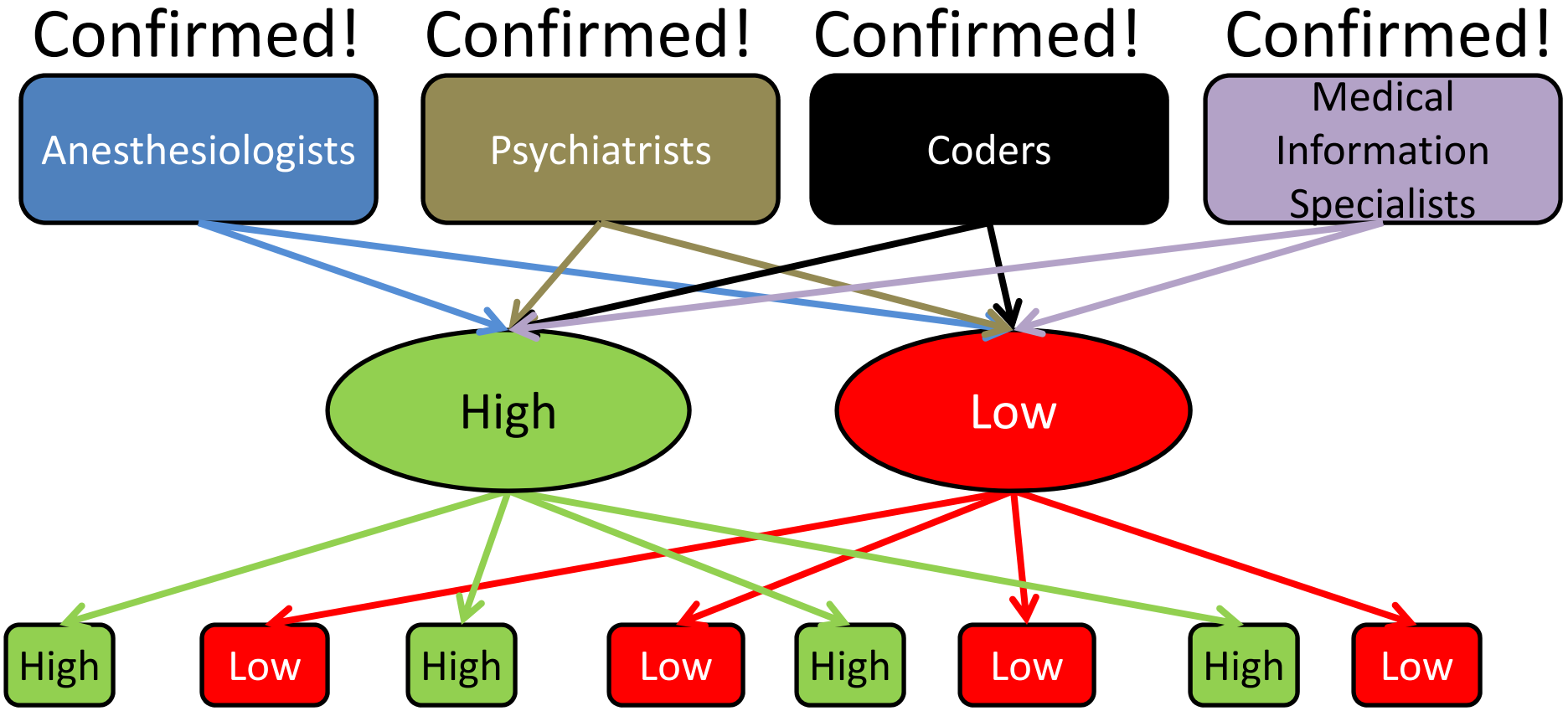
Respondent (ID)	Respondent Type (P)	Rule Type (R)	Rule Class (C)	Average Score of Responses
1	MIS	ANE	High	3
1	MIS	ANE	Low	2
1	MIS	CODE	High	3.3
1	MIS	CODE	Low	2.1
1	MIS	MIS	High	3.1111
1	MIS	MIS	Low	2.125
1	MIS	PSY	High	2.9
1	MIS	PSY	Low	2.05

Hypothesis Test 1 – Rules of All HCO Areas:



Hypothesis Test1 – Rules of All HCO Areas:

One-sided t-test, 95% confidence



Linear Mixed Effects Model

Rule class: high or low

Respondent type: MIS, CODE, PSY, ANE

Rules type: MIS, CODE, PSY, ANE

$$lemr\,(aveScore \sim h \times p + r + (1 \mid id), data)$$

Random effect of each participant

$$Y = \beta_0 + \beta_1(h=1) + \beta_2(p=code) + \beta_3(p=psy) + \beta_4(p=ane)$$
$$+ \beta_5(r=code) + \beta_6(r=psy) + \beta_7(r=ane)$$
$$+ \beta_8(p=code)(h=1) + \beta_9(p=psy)(h=1) + \beta_{10}(p=ane)(h=1)$$

How MIS respondents distinguish high and low likelihood of rules for all HCO areas

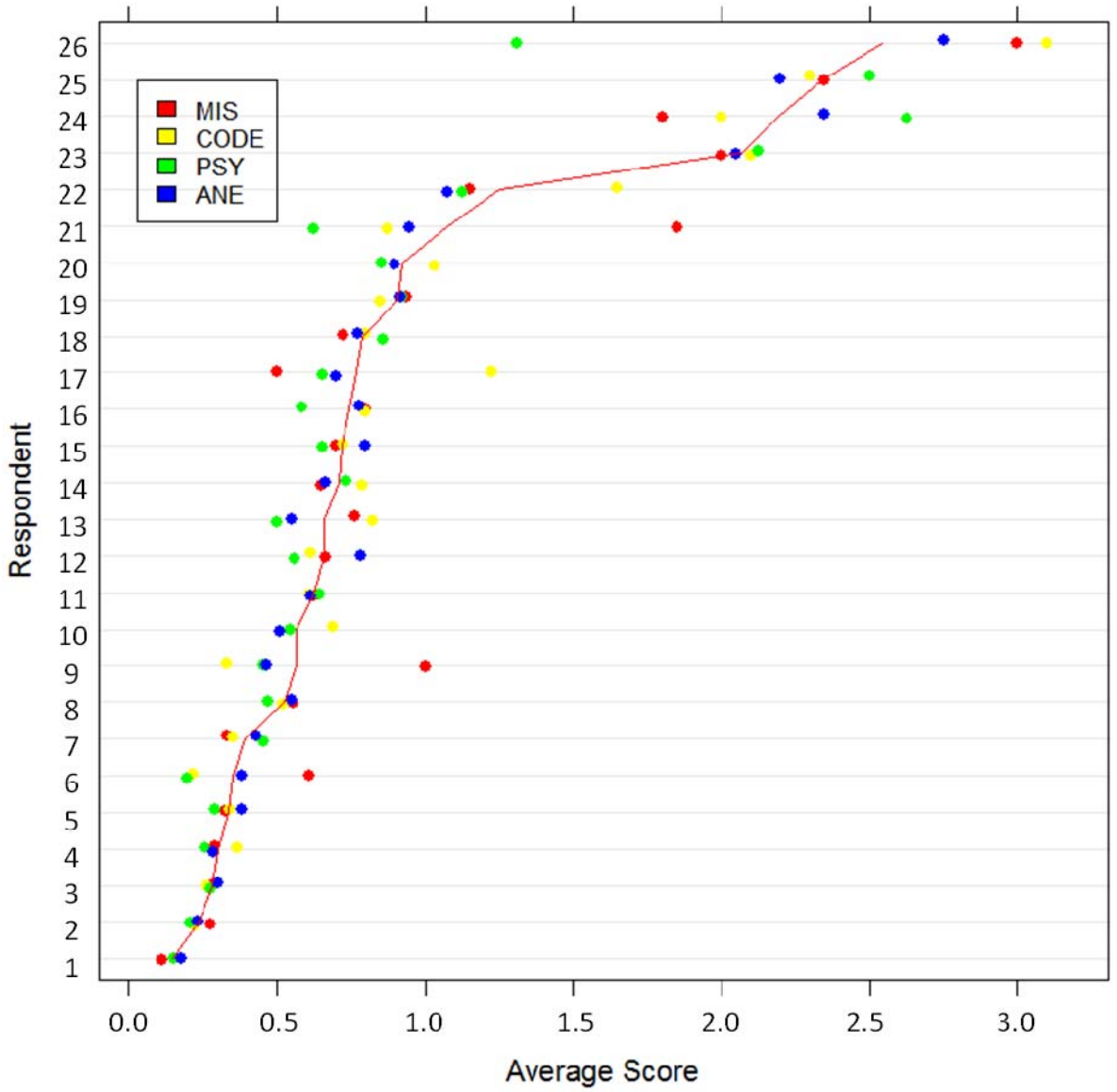
E(y | p=mis, h=0, r=ALL) = $\beta_0 + \beta_5 + \beta_6 + \beta_7$;
E(y| p=mis, h=1, r=ALL) = $\beta_0 + \beta_1 + \beta_5 + \beta_6 + \beta_7$;
E(y | p=mis, h=1, r=ALL)-E(y | p=mis, h=0, r=ALL) =β_1

How about CODE, PSY and ANE?

Distribution of respondents' average score for low rules

Certain respondents are inclined to assign large likelihoods (upper right section of the plot),

while others are included to assign small likelihoods (the lower left section of the plot).



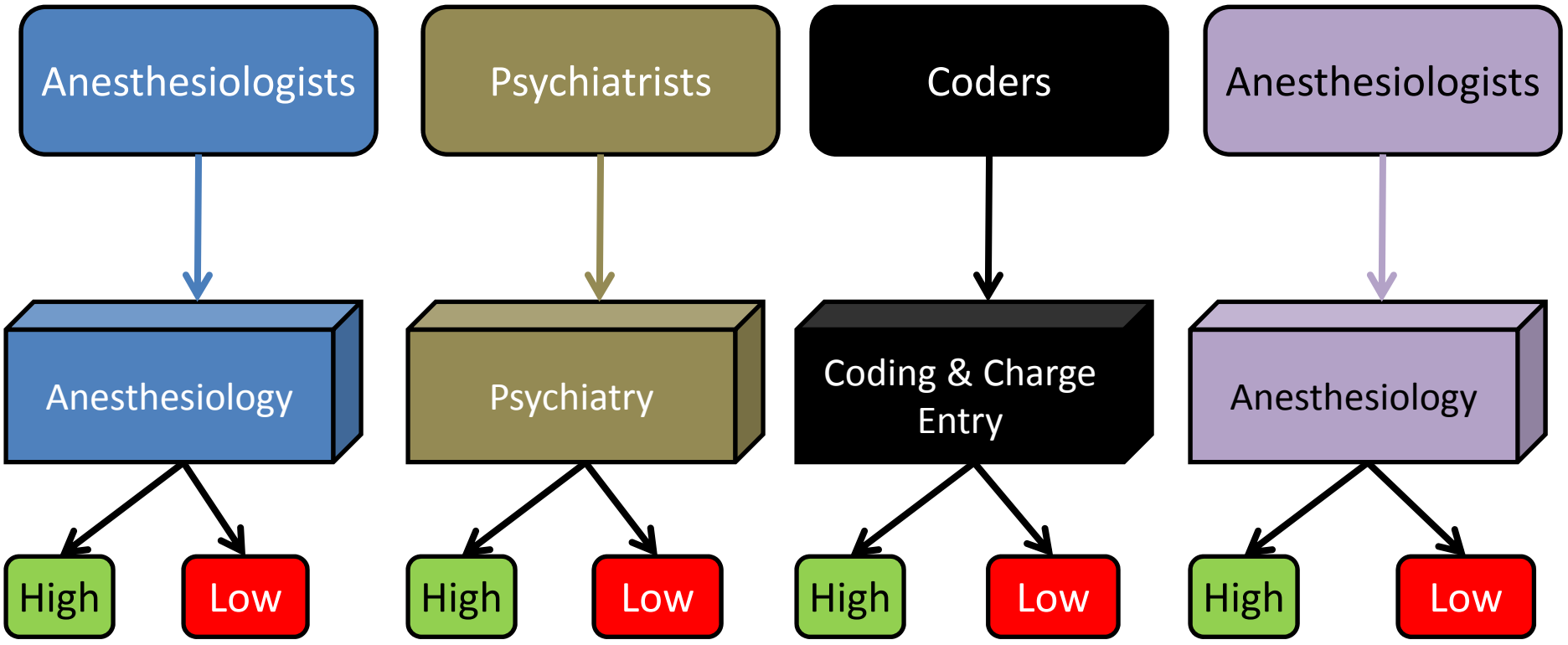
β	β values	description	p value
β_1	0.351557	MIS respondents distinguish high and low likelihood rules for all HCO areas	$1.91*10^{-9}$
$\beta_1+\beta_8$	0.521492	CODE respondents distinguish high and low likelihood rules for all HCO areas	$1.11*10^{-6}$
$\beta_1+\beta_9$	0.677858	PSY respondents distinguish high and low likelihood rules for all HCO areas	$9.33*10^{-8}$
$\beta_1+\beta_{10}$	0.691166	ANE respondents distinguish high and low likelihood rules for all HCO areas	$1.22*10^{-8}$

Respondents from four areas can distinguishing between high and low likelihood rules for all HCO areas

Hypothesis Test 2– Self Assessment:

Linear Mixed Effects Model

One-sided t-test, 95% confidence



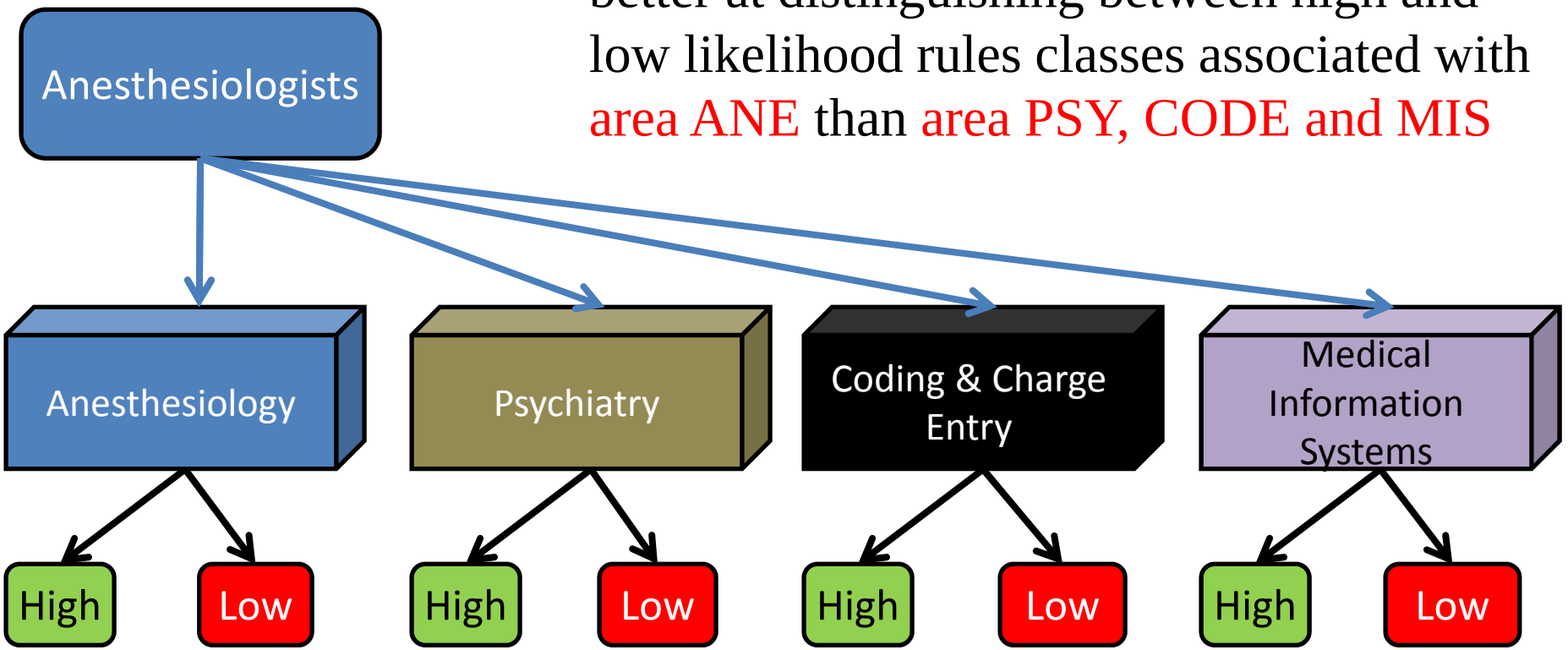
Confirmed! Confirmed! Confirmed! Confirmed!

Hypothesis Test 3– Bias Toward Own Rules

Linear Mixed Effects Model

One-sided t-test, 95% confidence

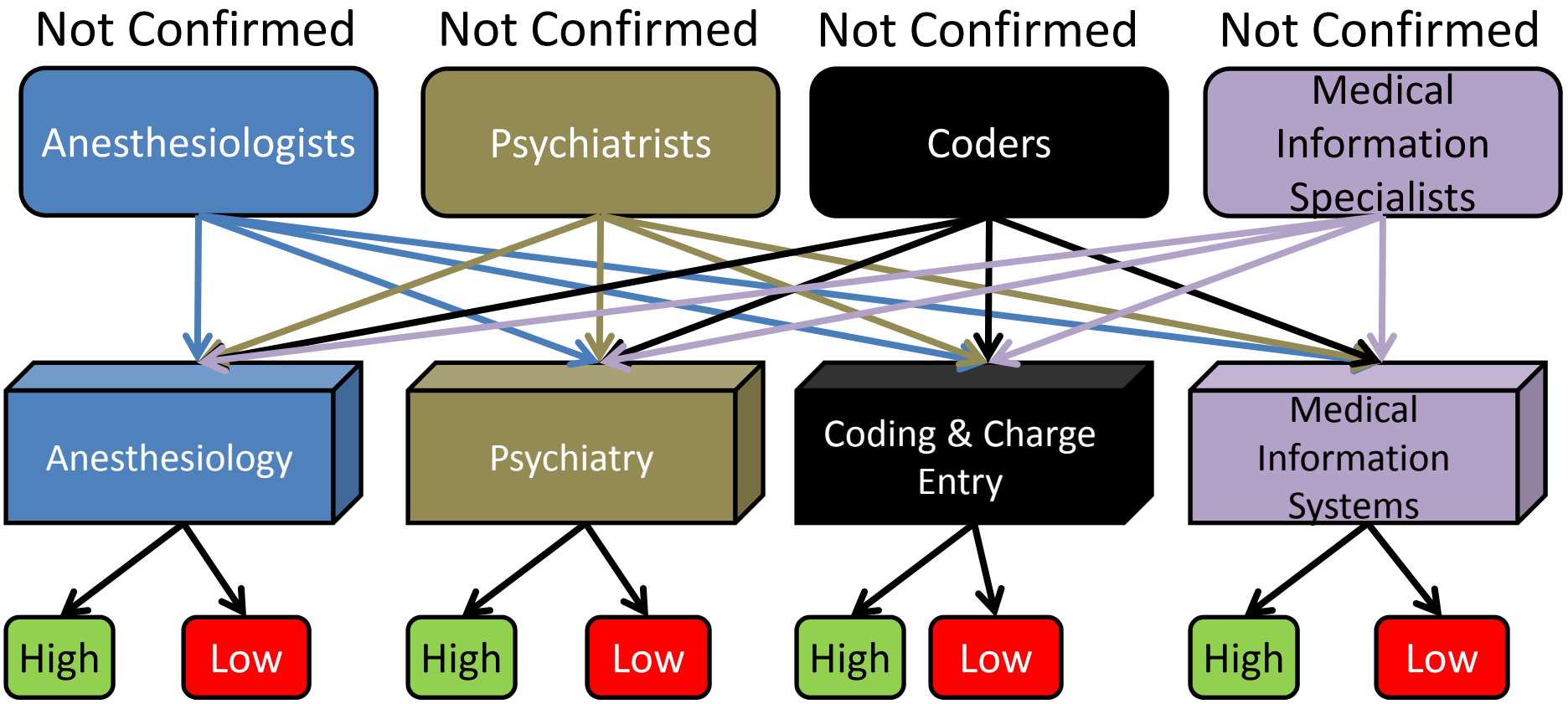
Respondents from HCO area ANE are better at distinguishing between high and low likelihood rules classes associated with area ANE than area PSY, CODE and MIS



Hypothesis Test 3 – Bias Toward Own Rules

Linear Mixed Effects Model

One-sided t-test, 95% confidence



Conclusions

- Healthcare organization employees generally understand what goes on around them…

  … and for other sections of the organization as well!

- Automated healthcare organizational modeling may be possible.

- Anomalies detection through collaborative patterns may be reliable!

Q&A

Thanks!