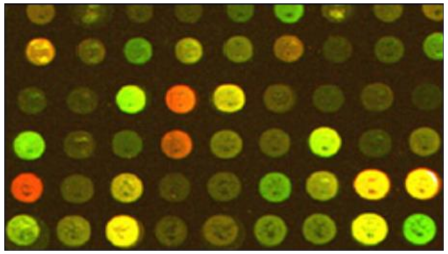
Smartphones Medicalized, with Data Analytics for Complex Diseases Management

Shuai Huang, Ph.D.
Associate Professor
Industrial and Systems Engineering
University of Washington

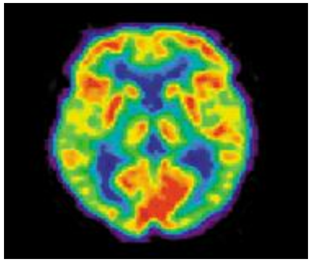
Research in Healthcare

Bench → Bedside → Community → Policy

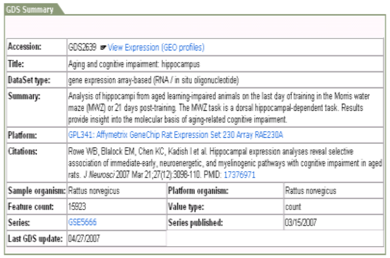
Microarray data



Neuroimaging



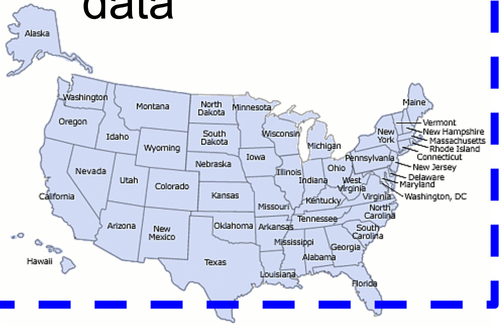
Electronic medical records

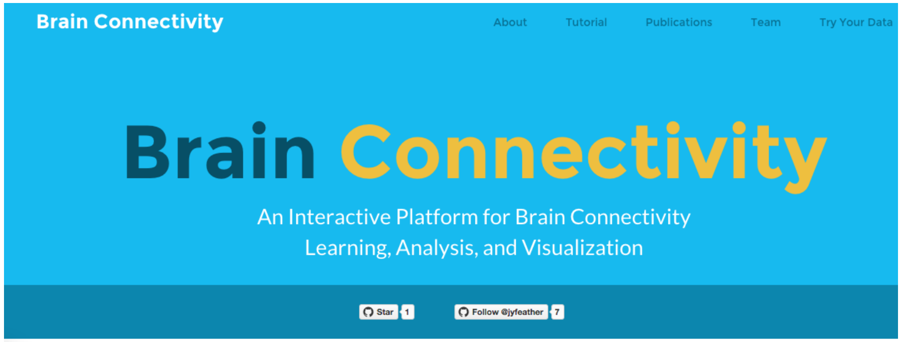


GDS Summary

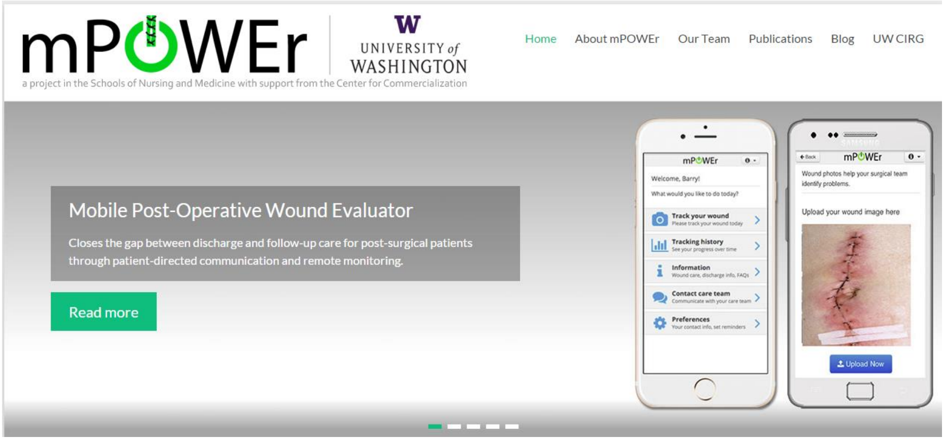
Accession:	GDS2639 View Expression (GEO profiles)		
Title:	Aging and cognitive impairment: hippocampus		
DataSet type:	gene expression array-based (RNA / in situ oligonucleotide)		
Summary:	Analysis of hippocampi from aged learning-impaired animals on the last day of training in the Morris water maze (MWZ) or 21 days post-training. The MWZ task is a dorsal hippocampal-dependent task. Results provide insight into the molecular basis of aging-related cognitive impairment.		
Platform:	GPL341: Affymetrix GeneChip Rat Expression Set 230 Array RAE230A		
Citations:	Rowe WB, Blalock EM, Chen KC, Kadish I et al. Hippocampal expression analyses reveal selective association of immediate-early, neuroenergetic, and myelinogenic pathways with cognitive impairment in aged rats. *J Neurosci* 2007 Mar 21;27(12):3098-110. PMID: 17376971		
Sample organism:	Rattus norvegicus	**Platform organism:**	Rattus norvegicus
Feature count:	15923	**Value type:**	count
Series:	GSE5666	**Series published:**	03/15/2007
Last GDS update:	04/27/2007		

Epidemiological data



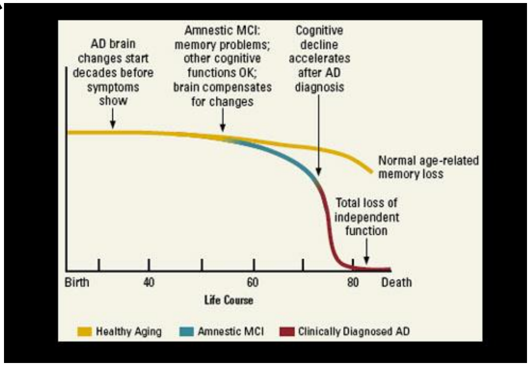


http://brainconnectivity.cc/

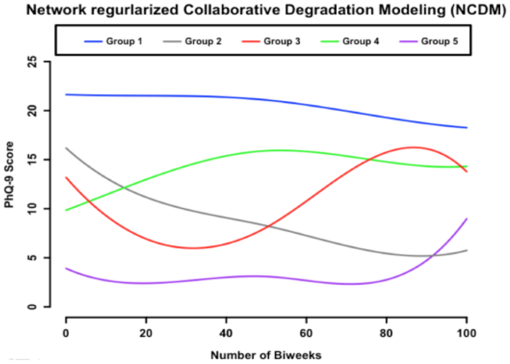


From Reactive Care to Preventative Care

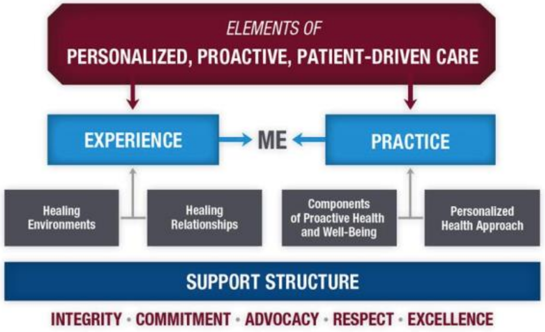
Massive individuals having heterogeneous disease trajectories



Management of the Aging population



Management of mental health such as Depression



VA cohort with complex chronic conditions

Process, limited resource, preventative care

Prognostics challenges

- **Heterogeneity of the degradation processes**
- **Time-varying nature of the degradation**

Monitoring challenges

- **Need smart schedules for proactive health monitoring**
- **Able to update individuals' dynamic condition**

Intervention challenges

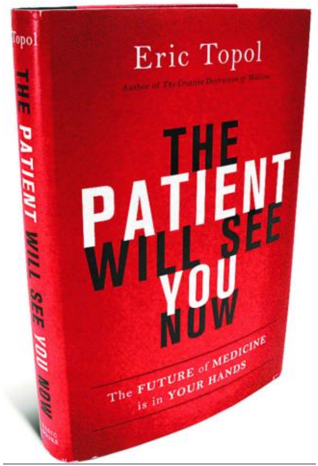
- **Difficult to identify the top high risk individuals**
- **Smart intervention resource allocation**

https://www.va.gov/patientcenteredcare/about

Medicalized Smartphones & More Patient Engagement

Smartphones

- Built-in sensors: GPS, accelerometer, wifi-related, audio, proximity sensors, etc.;
- This sensor base continuously enhanced with more sensors, medical apps …



292 THE PATIENT WILL SEE YOU NOW

FIGURE E.1: The medicalized smartphone of the future. Check marks indicate functions that are now operational, at least in part. Adapted from xkcd.com.

Patients

- Today, over 81% US adults own smart devices, 69% track at least one health indicator (e.g., weight, sleep), and 59% sought health information online in the last year.
- Patients increasingly seek ways to engage in their healthcare using the emerging technologies such as smartphone

From Sensor Data to Assessment of Health Conditions

Remote Measurement of Cognitive Stress via Heart Rate Variability

Daniel McDuff[1], Sarah Gontarek[2] and Rosalind Picard[1]

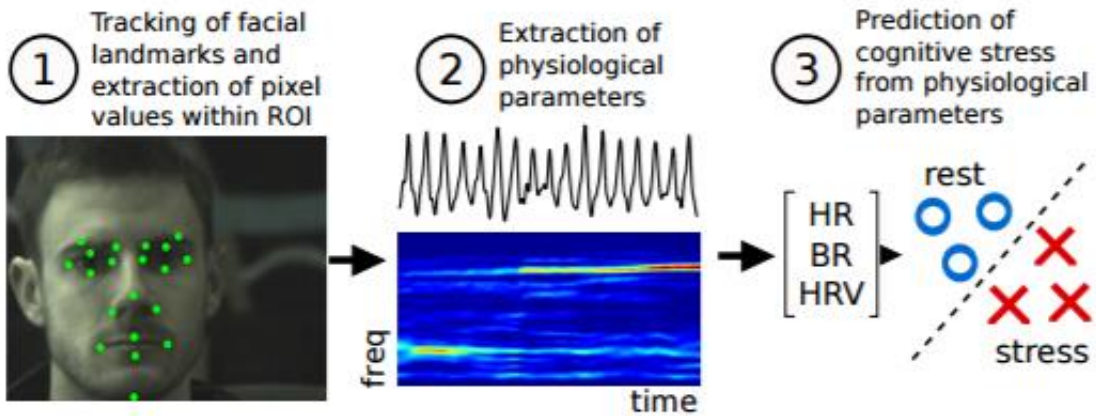


Fig. 1. Overview of the automated system for prediction of cognitive stress from remotely measured physiology. 1) Facial landmarks detected and color channel information extracted from the ROI, 2) BVP extracted from color channel signals and HR, BR and HRV parameters calculated, 3) physiological features used to predict restful state or cognitive stress state.

The new technologies not just provide a way to collect existing data; they actually create new data, and challenge our concepts of "health" and "diseases".

Google May Know the Diagnosis …

CORRESPONDENCE

. . . And a Diagnostic Test Was Performed

N Engl J Med 2005; 353:2089-2090 | November 10, 2005 | DOI: 10.1056/NEJM200511103531923

To the Editor:

At a recent case conference with a distinguished visiting professor, a fellow in allergy and immunology presented the case of an infant with diarrhea; an unusual rash ("alligator skin"); multiple immunologic abnormalities, including low T-cell function; tissue eosinophilia (of the gastric mucosa) as well as peripheral eosinophilia; and an apparent X-linked genetic pattern (several male relatives died in infancy). The attending physicians and house staff discussed several diagnostic possibilities, but no consensus was reached. Finally, the visiting professor asked the fellow if she had made a diagnosis, and she reported that she had indeed and mentioned a rare syndrome known as IPEX (immunodeficiency, polyendocrinopathy, enteropathy, X-linked). It appeared to fit the case, and everyone seemed satisfied. (Several weeks later, genetic testing on the baby revealed a mutation in the *FOXP3* gene, confirming the diagnosis.)

"How did you make that diagnosis?" asked the professor. Came the reply, "Well, I had the skin-biopsy report, and I had a chart of the immunologic tests. So I entered the salient features into Google, and it popped right up."

"William Osler," I offered, "must be turning over in his grave. You googled the diagnosis?"

Where does this lead us? Are we physicians no longer needed? Is an observer who can accurately select the findings to be entered in a Google search all we need for a diagnosis to appear, as if by magic? The cases presented at clinicopathological conferences can be solved easily; no longer must the discussant talk at length about the differential diagnosis of fever with bradycardia. Even worse, the Google diagnostician might be linked to an evidence-based medicine database, so a computer could e-mail the prescription to the e-druggist with no human involvement needed. The education of house staff is morphing into computer-search techniques. Surely this is a trend to watch.

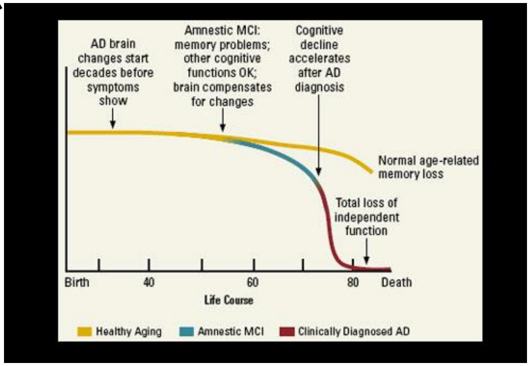
- Boundaries between disciplines are vanishing …
- *"The history of modern knowledge is concerned in no small degree with man's attempt to escape from his previous concepts" – Harold Himsworth*

Outline

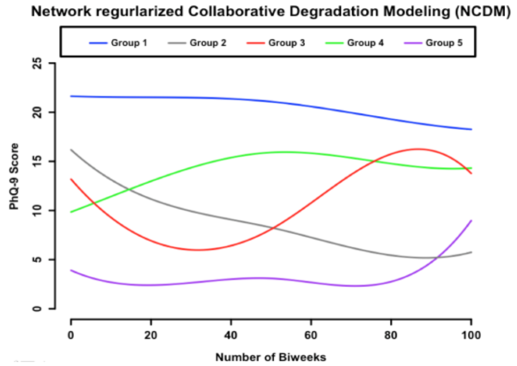
- Overview
- Data Analytics for Disease Management
  - *Topic I: New methods for personalization in disease modeling and monitoring*
  - *Topic II: Detection of depression from communication*
- Highlights of other Works
- Conclusion

From Reactive Care to Preventative Care

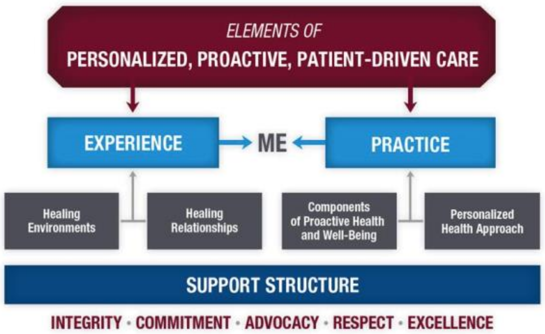
Massive individuals having heterogeneous disease trajectories



Management of the Aging population



Management of mental health such as Depression



VA cohort with complex chronic conditions

Process, limited resource, preventative care

Prognostics challenges

- Heterogeneity of the degradation processes
- Time-varying nature of the degradation

Monitoring challenges

- Need smart schedules for proactive health monitoring
- Able to update individuals' dynamic condition

Intervention challenges

- Difficult to identify the top high risk individuals
- Smart intervention resource allocation

https://www.va.gov/patientcenteredcare/about

Some Basics about Modeling of Disease Trajectory

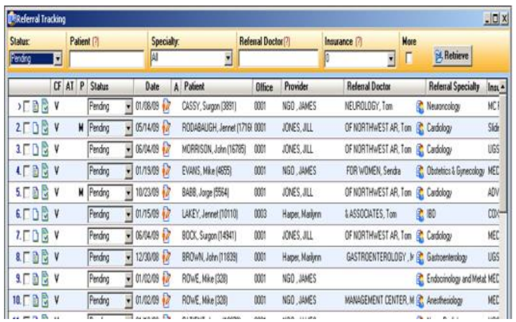
Individual Measurements

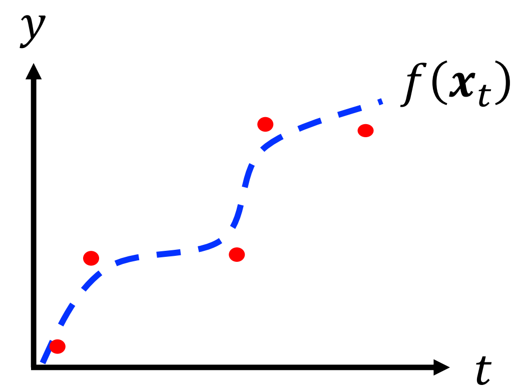
$$\mathbf{X} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_T]$$
$$\boldsymbol{y} = [y_1, \dots, y_T]$$

Computational Model

Disease Trajectory

$$y_t = f(\boldsymbol{x}_t) + \varepsilon_t, t = 1, \dots, T$$





Generalized Regression Model is widely used: $f(\boldsymbol{x}_t) = \boldsymbol{\Phi}(\boldsymbol{x}_t)^T \boldsymbol{\beta}$

Basis function,
e.g. Polynomial basis; Spline basis

Data Fusion to Create Contemporaneous Health Index

292 THE PATIENT WILL SEE YOU NOW

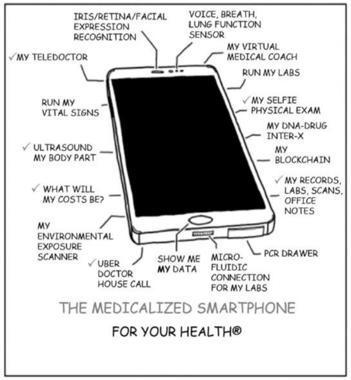
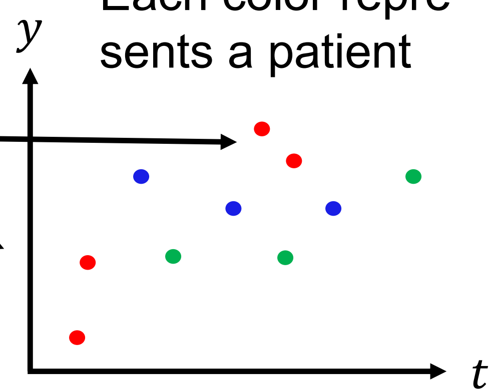


FIGURE E.1: The medicalized smartphone of the future. Check marks indicate functions that are now operational, at least in part. Adapted from xkcd.com.



Each color represents a patient

However, the abundance of individual data is collected at irregular time points (i.e., uneven distribution of measurement frequency), at different stages of disease progression, subject to enormous individual heterogeneity

Data fusion models convert individual data into health index

ELSEVIER

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

CHI: A contemporaneous health index for degenerative disease monitoring using longitudinal measurements

Yijun Huang[a], Qiang Meng[b], Heather Evans[c], William Lober[d], Yu Cheng[e], Xiaoning Qian[f], Ji Liu[a], Shuai Huang[b,*]

[a] Department of Computer Science, University of Rochester, United States
[b] Department of Industrial & Systems Engineering, University of Washington, United States
[c] Department of Surgery, University of Washington, United States
[d] Department of Biomedical Informatics and Medical Education, University of Washington, United States
[e] Healthcare Analytic Research, IBM T.J. Watson Research Center, United States
[f] Department of Electrical & Computer Engineering, Texas A&M University, United States

RESEARCH Open Access

DL-CHI: a dictionary learning-based contemporaneous health index for degenerative disease monitoring

Aven Samareh* and Shuai Huang

The Basic Framework for Collaborative Learning

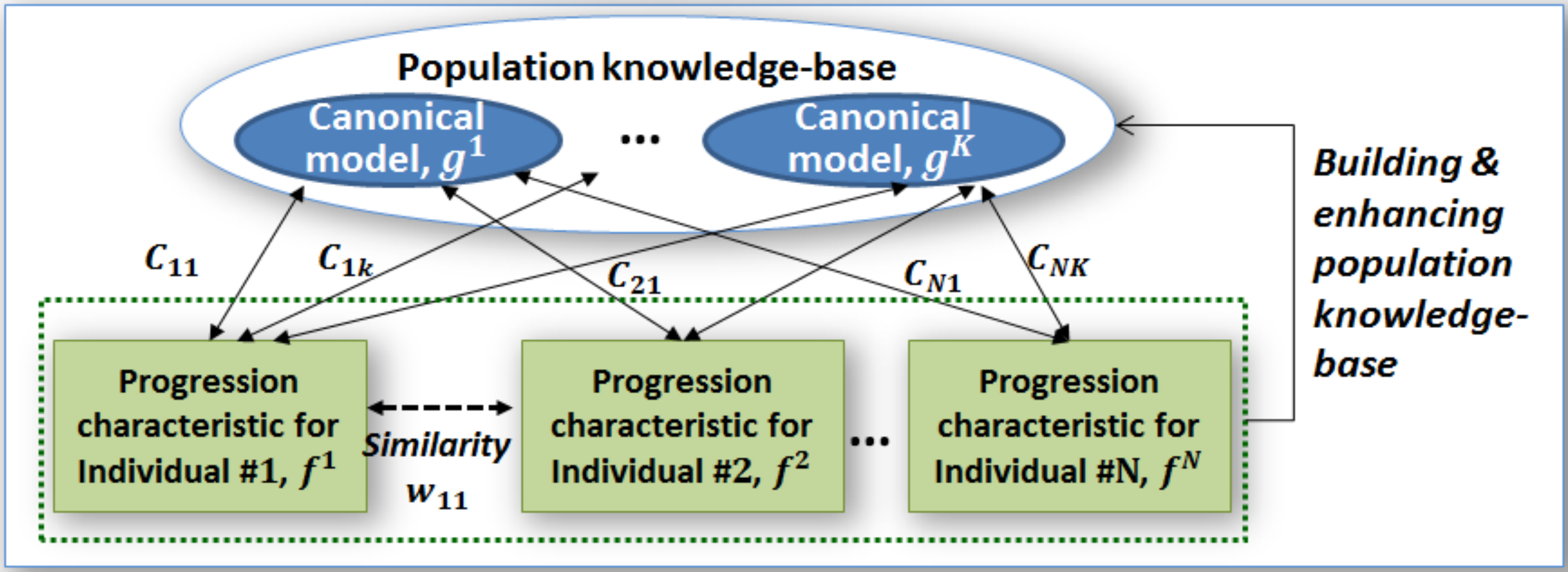
One-size-fit-all: builds one prediction model for all the subjects

Model Complexity

Fully individualized: builds a distinct model for each subject

Too simple

Too complex



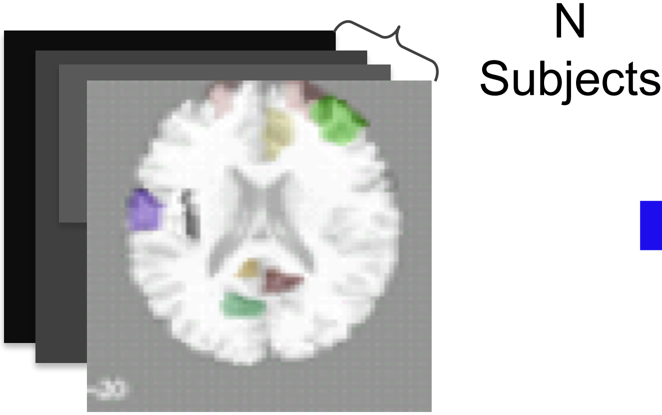
Our Formulation of the Collaborative Learning Framework

- Cluster structure is described by a set of latent models $g_k(\boldsymbol{x}),\ k = 1, \cdots, K$
- Use membership vector $\boldsymbol{c}$ for each subject

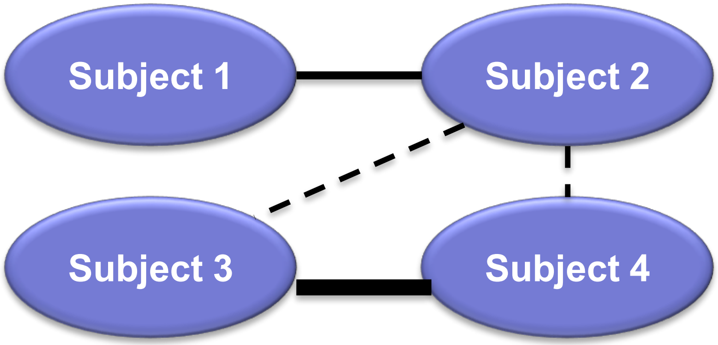
Goodness-of-fit | Regularization

$$\min_{\boldsymbol{c}_i, \mathbf{Q}} \sum_i \left\| \boldsymbol{y}_i - \sum_k c_{ik} g_k(\mathbf{X}_i) \right\|^2 + \lambda \sum_{j,l} ||\boldsymbol{c}_j - \boldsymbol{c}_l||^2 w_{jl}$$

subject to $c_{ik} \geq 0, \quad \sum_k c_{ik} = 1,\ g_k(\mathbf{X}_i) \geq \mathbf{0}$,
$\forall\ i = 1, \ldots, N$ and $k = 1, \ldots, K$.



Profiles on risk factors $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N$

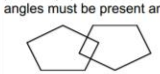


Similarity Matrix $\mathbf{W}$

SDM 2015; Math Bio 2016 IEEE Reliability, 2018

Application on Alzheimer's Disease

- The data was collected from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and processed by collaborators in Banner Alzheimer's Institute
- 478 subjects including 104 cognitively normal aging individuals (NC), 261 patients with mild cognitive impairment (MCI), and 133 AD patients (AD).
- ApoE genotypes, baseline MMSE, features extracted from MRI are used in the calculation of similarity.
- Quadratic model is used for modeling the disease trajectory of MMSE

Maximum Score	Patient's Score	Questions
5		"What is the year? Season? Date? Day? Month?"
5		"Where are we now? State? County? Town/city? Hospital? Floor?"
3		The examiner names three unrelated objects clearly and slowly, then the instructor asks the patient to name all three of them. The patient's response is used for scoring. The examiner repeats them until patient learns all of them, if possible.
5		"I would like you to count backward from 100 by sevens." (93, 86, 79, 72, 65, …) Alternative: "Spell WORLD backwards." (D-L-R-O-W)
3		"Earlier I told you the names of three things. Can you tell me what those were?"
2		Show the patient two simple objects, such as a wristwatch and a pencil, and ask the patient to name them.
1		"Repeat the phrase: 'No ifs, ands, or buts.'"
3		"Take the paper in your right hand, fold it in half, and put it on the floor." (The examiner gives the patient a piece of blank paper.)
1		"Please read this and do what it says." (Written instruction is "Close your eyes.")
1		"Make up and write a sentence about anything." (This sentence must contain a noun and a verb.)
1		"Please copy this picture." (The examiner gives the patient a blank piece of paper and asks him/her to draw the symbol below. All 10 angles must be present and two must intersect.) 
30		TOTAL

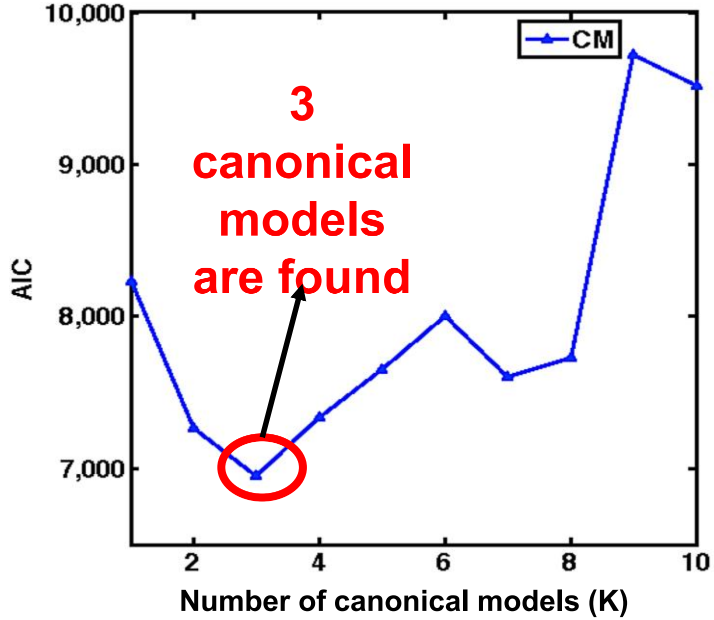
Cognitive degradation model [Sliwinski et al., 2003]:

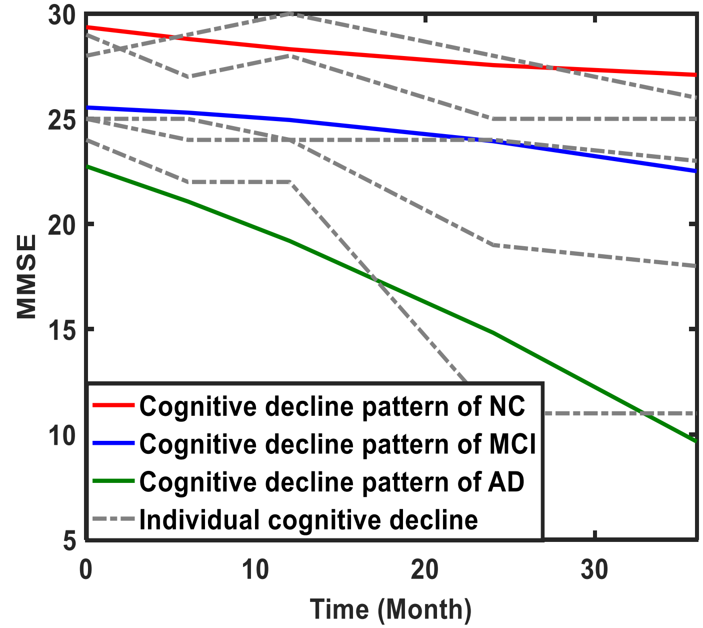
$$f_i(t) = \beta_{i0} + \beta_{i0} t + \beta_{i0} t^2 + \varepsilon_{it}$$

1. Reference: M. J. Sliwinski, S. M. Hofer, et al., "Modeling memory decline in older adults: The importance of preclinical dementia", *Psychology and Aging*, Vol. 18, pp.658–671, 2003.

Application on Alzheimer's Disease

Choose optimal number of canonical models





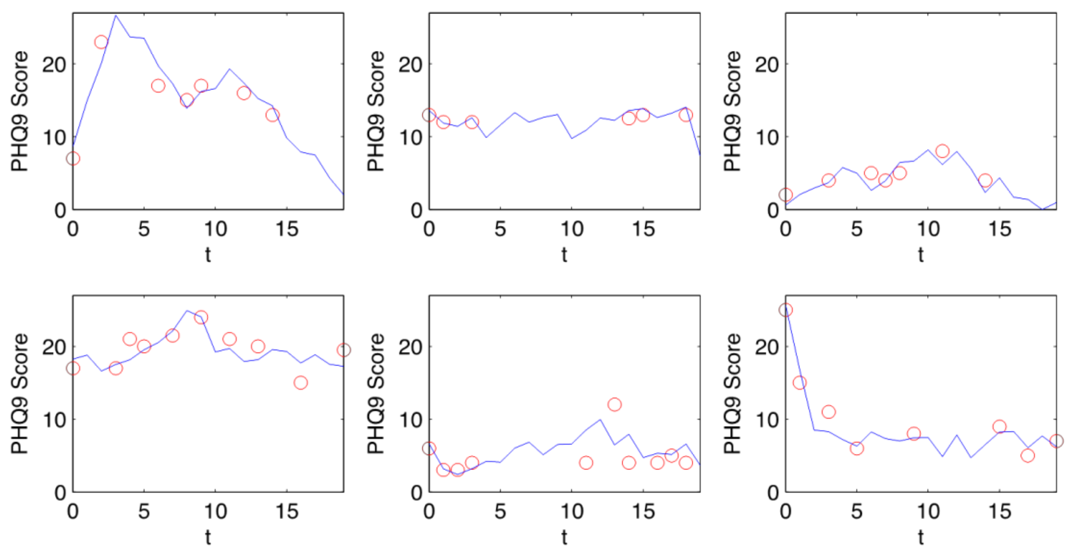
	IGM	CM	MEM	SCM
Target:MMSE				
nMSE	1.799	0.936	0.755	**0.531**
wR	0.580	0.618	0.660	**0.716**
M48 rMSE	4.874	4.330	3.705	**3.651**
M60 rMSE	8.326	5.458	5.040	**3.777**

Application on Depression

- Data comes from NHRN (Mental Health Research Network), largest depression dataset in U.S.
  - **3,159** subjects, each subject has
    - **more than 5** depression assessments (PHQ-9 scores).
    - Demographic features, treatment status, Charlson Comorbidity Score, 9th question score

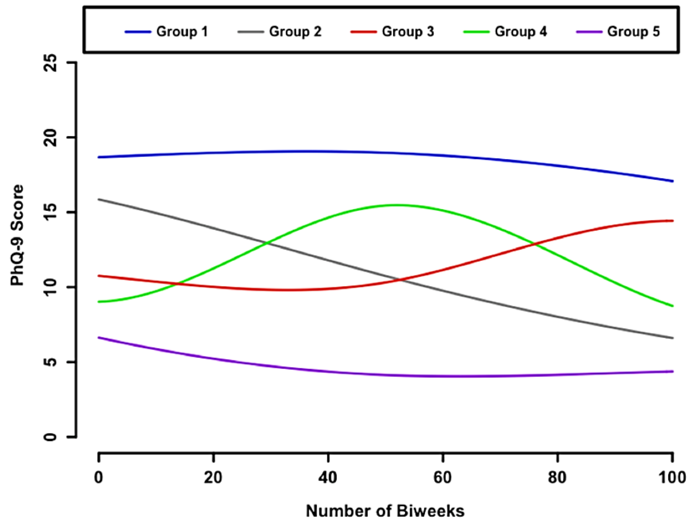
Over the last *2 weeks*, how often have you been bothered by any of the following problems? *(use "✓" to indicate your answer)*	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite — being so figety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3

Exemplary individual depression trajectories



15

Application on Depression

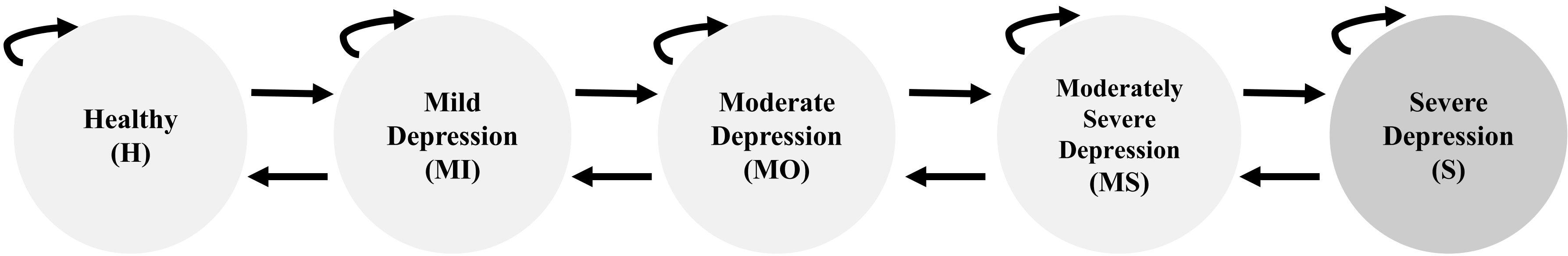


Five depression trajectory patterns are discovered

Our model leads to more accurate prognostics of depression trajectories.

Method	IGM	MEM	CM	SCM
Target: PHQ-9				
rMSE	12.534	5.913	5.178	**3.210**

Extension to Markov Disease Models



PHQ-9 Score	Depression Severity
1-4	Minimal
5-9	Mild
10-14	Moderate
15-19	Moderately Severe
20-27	Severe

	H	Mi	Mo	MS	S
H	0.762	0.228	0.010	0	0
Mi	0.097	0.665	0.215	0.023	0
Mo	0.002	0.129	0.691	0.165	0.014
MS	0	0.007	0.201	0.598	0.194
S	0	0	0.011	0.230	0.759

Markov Based Collaborative Learning

K Canonical Markov Models

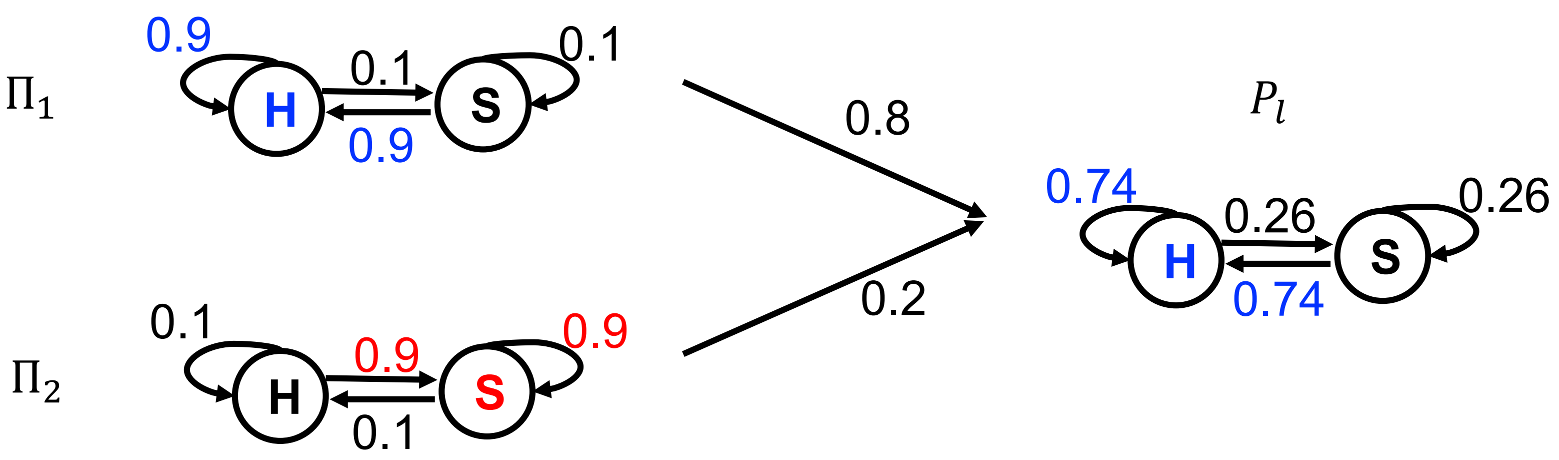
$$(\Pi_1, \theta_1), \dots, (\Pi_K, \theta_K) \xrightarrow{\boldsymbol{c}_i}$$

Individual Markov Models

$$P_l = \sum_k c_{ik} \Pi_k \qquad \pi_l = \sum_k c_{ik} \theta_k$$

Transition Matrix

Initial Distribution



Markov Based Collaborative Learning

Log-likelihood Function

$$\max_{c_i,\boldsymbol{\theta}_k,\boldsymbol{\Pi}_k} \sum_{i=1}^{N}\{\sum_s e_{is}\log[\sum_k c_{ik}\theta_{ks}] + \sum_{s_1}\sum_{s_2} N_i(s_1,s_2)\log[\sum_k c_{ik}\boldsymbol{\Pi}_k(s_1,s_2)]\}$$

$$-\frac{\lambda}{2}\sum_{j,l} w_{jl}\left\|\boldsymbol{c}_j - \boldsymbol{c}_l\right\|^2,$$

Regularizer

s.t. $\sum_{s_2}\boldsymbol{\Pi}_k(s_1,s_2) = 1, \sum_s \theta_{ks} = 1, \sum_k c_{ik} = 1$,

$\forall s_1 = 1,\dots,S, k = 1,\dots,K,\ \forall i = 1,\dots,N,$

all parameters are nonnegative. (2)

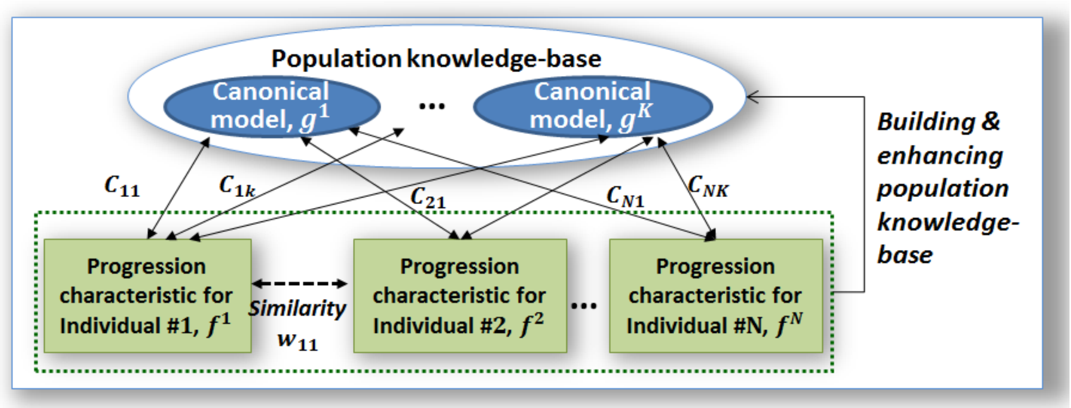
- e_{is}(indicator of initial state): $e_{is} = 1$ if $x_{i0} = s; e_{is} = 0$ otherwise
- $N_i(s_1,s_2)$: number of transitions from s_1 to s_2 on individual i.
- MLE of $P(i)$ can be obtained by solving:

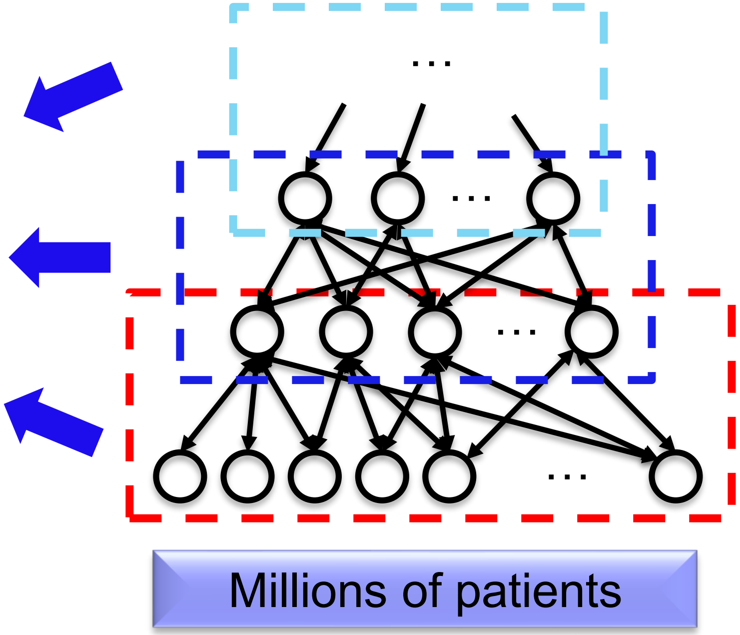
$$\max\{\log(\Pr(X_{it+1} = x_{i1})) + \sum_{s_1,s_2} N_i(s_1,s_2)\log(\mathrm{P}_i(s_1,s_2))\},$$

Extension: Hierarchical Collaborative Learning

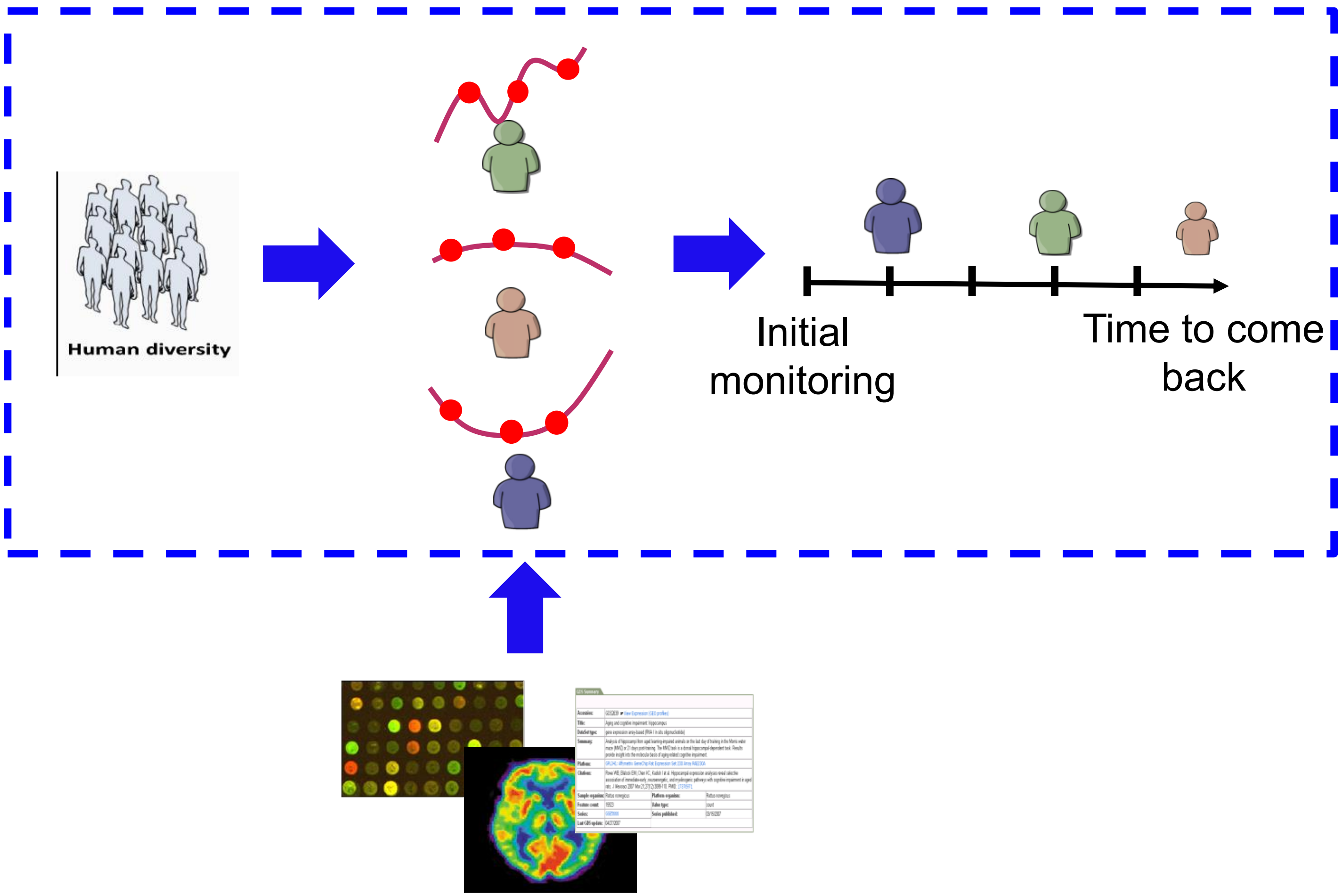
Collaborative learning is a concept that could be iterated

- A hierarchical collaborative learning framework could mitigate the problem if we aim to learn millions of personal models
- Canonical models that "span" the space for the personal models in a lower level, become the personal model for the canonical models in the next level



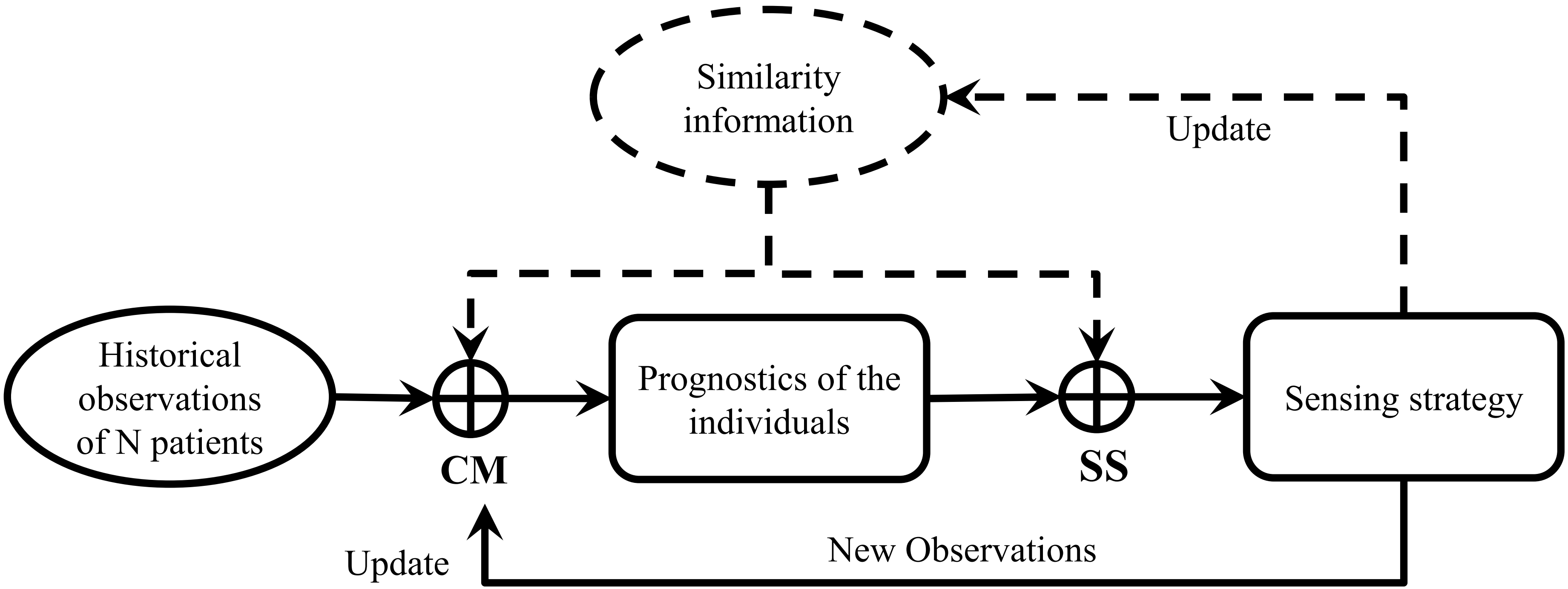


From Trajectory Modeling to Disease Monitoring



Sensing data

Collaborative Learning + Selective Sensing = Adaptive Patient Monitoring

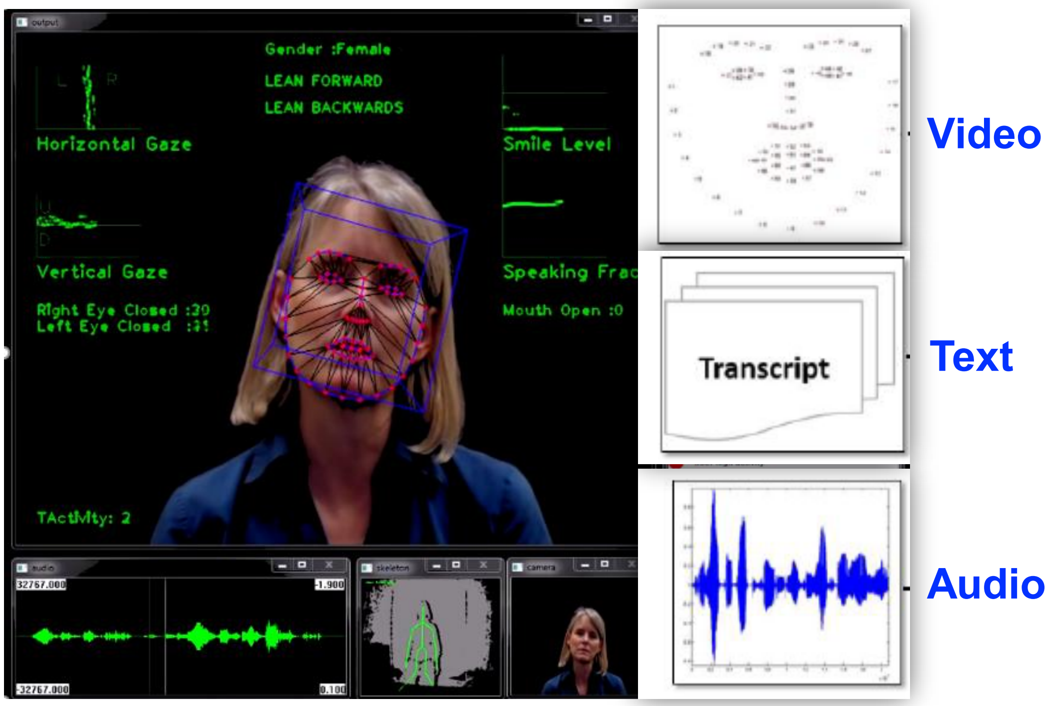


- ❖ A prediction model for each individual to predict the risk of disease onset
- ❖ Collaborative prognostics and selective sensing: adjust the risk scores based on the similarity of the individuals, and re-arrange the individuals from high-risk to low-risk
- ❖ Modeling updating: update the prediction model for each individual based on the new measurements of the selected individuals

Outline

- Overview
- Data Analytics for Disease Management
  - *Topic I: New methods for personalization in disease modeling and monitoring*
  - *Topic II: Detection of depression from communication*
- Highlights of other Works
- Conclusion

Detect Depression from Communication



IISE TRANSACTIONS ON HEALTHCARE SYSTEMS ENGINEERING
https://doi.org/10.1080/24725579.2018.1496494

Taylor & Francis
Taylor & Francis Group

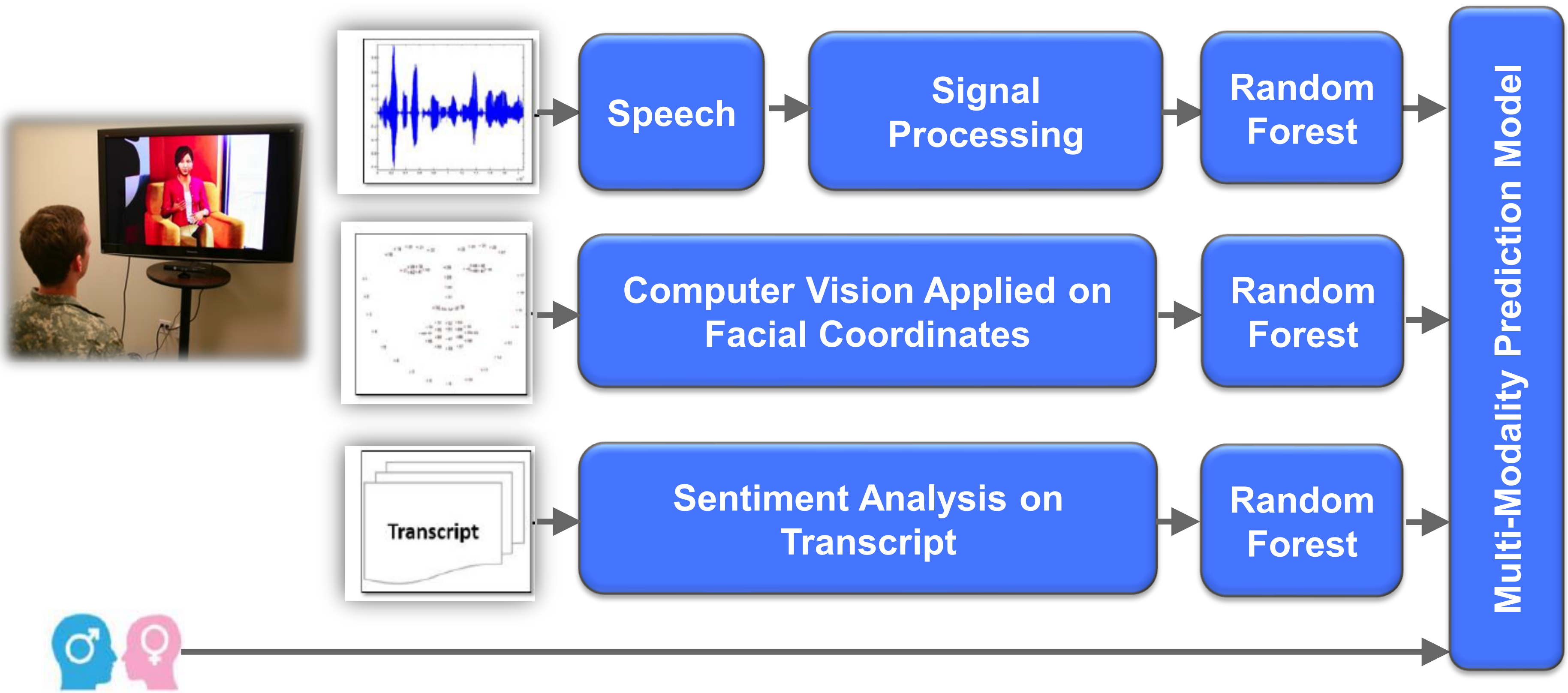
Detect depression from communication: how computer vision, signal processing, and sentiment analysis join forces

Aven Samareh[a], Yan Jin[b], Zhangyang Wang[c], Xiangyu Chang[d], and Shuai Huang[a]

[a]Industrial & Systems Engineering Department, University of Washington, Seattle, Washington, USA; [b]Research Engineer, JD.com, Inc., San francisco, California, USA; [c]Department of Computer Science and Engineering, Texas A&M University, College Station, Texas, USA; [d]School of Management, Xi'an Jiaotong University Shaanxi, P.R. China

https://www.theguardian.com/sustainable-business/2015/sep/17/ellie-machine-that-can-detect-depression

The Computational Pipeline



Characterization of the Condition by Biomarkers

Description of audio biomarkers used in a time domain

Audio Biomarkers	Description	No. of Biomarkers
Modulation of amplitude	It is used to find the amplitude of two signals that are multiplied by the superimposed signals.	1
Envelope	It represents the varying level of an audio signal over time.	1
Autocorrelation	It shows the repeating patterns between observations as a function of the time lag between them.	1
Onset detector	It is used to detect, a sudden change in the energy or any changes in the statistical properties of a signal.	1
Entropy of energy	It is a measure of abrupt changes in the energy level of an audio signal	1
Tonal power ratio	It is obtained by taking the ratio of the tonal power of the spectrum components to the overall power.	1
RMS power	Root mean square (RMS) approximates the volume of an audio frame.	1
ZCR	Zero Crossing Rate (ZCR) is the number of times the signal changes sign in a given period of time.	1

Characterization of the Condition by Biomarkers – cont'd

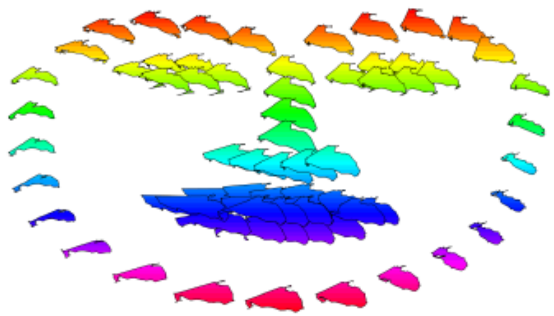
Description of audio biomarkers used in a frequency domain

Audio Biomarkers	Description	No. of Biomarkers
PLP	It is a technique to minimize the differences between speakers.	9
MFCC	It is a representation of the short-term power spectrum of an audio signal.	12
Spectral decrease	It computes the steepness of the decrease of the spectral envelope.	1
Spectral rolloff	It can be treated as a spectral shape descriptor of an audio signal.	1
Spectral flux	It is a measure of spectral change between two successive frames.	1
Spectral centroid	It is a measure to characterize the center mass of the spectrum.	1
Spectral slope	It is the gradient of the linear regression of a spectrum.	1
Spectral autocorrelation	It is a function that measures the regular harmonic spacing in the spectrum of the speech signal.	1

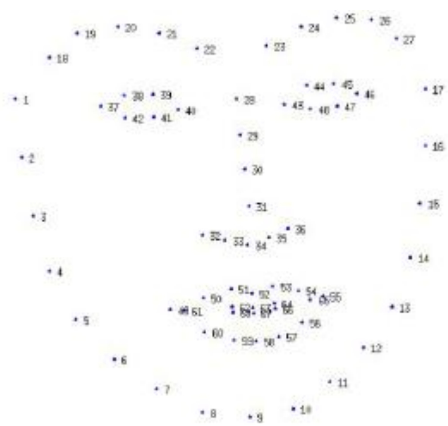
Overall 35 audio biomarkers

Characterization of the Condition by Biomarkers – cont'd

Head Biomarkers
41 biomarkers



Distance Biomarkers
92 biomarkers



Overall 133 video biomarkers

Basic Statistics of Words or Sentences

↓

Depression Related Words

↓

AFINN Sentiment Analysis

Overall 8 text biomarkers

Prediction Performance

root-mean-square error (**RMSE**)

mean absolute error (**MAE**)

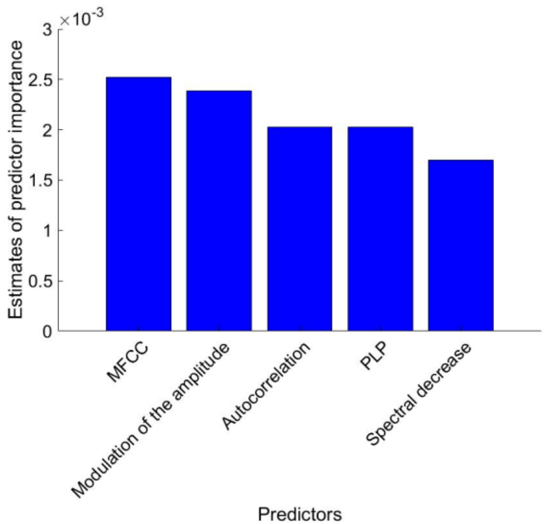
Biomarkers used	‘development’		‘train’	
	RMSE	MAE	RMSE	MAE
The baseline provided by the AVEC organizer				
Visual only	7.13	5.88	5.42	5.29
Audio only	6.74	5.36	5.89	4.78
Audio & Video	6.62	5.52	6.01	5.09
The model that doesn’t include gender variable				
Visual only	6.67	5.64	6.13	5.08
Audio only	6.00	5.25	5.62	4.89
Text only	5.95	5.21	5.68	5.17
Multi-modality prediction model	5.12	4.12	4.25	4.54
The model that includes the gender variable				
Visual only	5.65	4.87	4.99	4.46
Audio only	5.89	5.18	5.66	5.06
Text only	5.86	4.88	5.67	4.96
Multi-modality prediction model	4.78	4.05	4.35	3.69

Audio Biomarkers Relationship With Non-Linguistic Speech Patterns

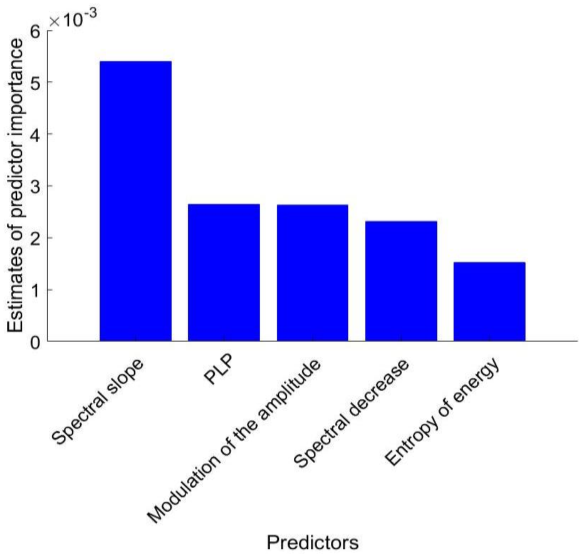
Audio Biomarkers	Loudness ↓	Pitch ↓	Silence ↑	Interruption ↑	Pauses ↑	Anger ↑	Laughter ↓
Modulation of Amplitude	[12,13]	[14,15]					
Envelope		[16]	[17]				
Autocorrelation	[18]	[19]					
Onset Detector		[20]	[21]				
Entropy of Energy	[22]	[23]					
Zero Crossing		[24]	[17]	[25]			
PLP	[26]				[29]	[30]	[31]
MFCC		[32,33]	[34]		[35,36]	[37,38]	[31]
Spectral Decrease	[39]	[16]	[40]				[31]
Spectral Roll off	[39]						[31]
Spectral Flux	[39]						[31]
Spectral Centroid	[39]						[31]
Spectral Slope	[39]						[31]
Spectral Autocorrelation	[39]						[31]

Validation and Interpretation

Audio biomarkers for females



Audio biomarkers for males

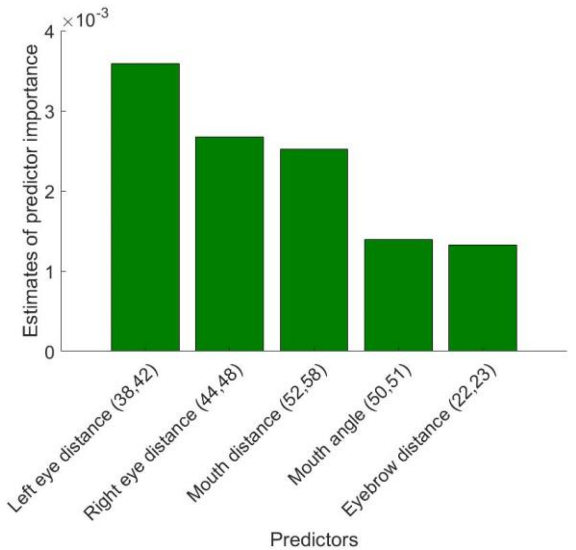


Selected biomarkers	Females	p-value	Males	p-value
Audio biomarkers	MFCC	$\boxed{0.0012}$	Spectral slope	<0.0001
	Modulation of amplitude	0.0008	PLP	<0.0001
	Autocorrelation	<0.0001	Modulation of amplitude	0.022
	PLP	<0.0001	Spectral decrease	0.012
	Spectral decrease	0.002	Entropy of energy	0.011
	Spectral slope	0.0080	MFCC	$\boxed{0.1140}$
	Entropy of energy	0.2100	Autocorrelation	0.0210

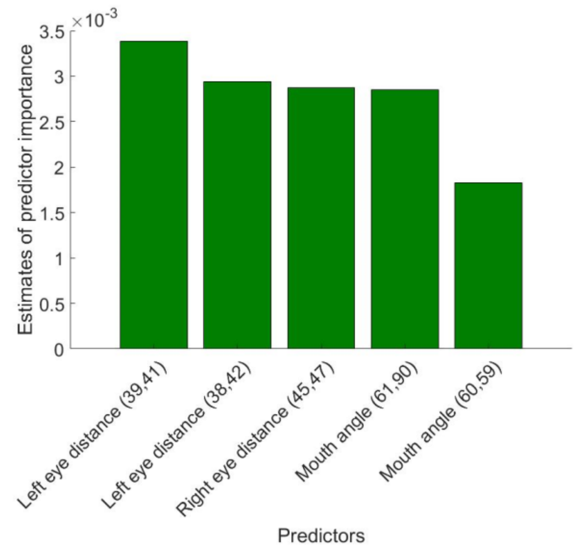
p-value of the selected top 5 significant biomarkers for females and males

Video Biomarkers

Video biomarkers for females



Video biomarkers for males

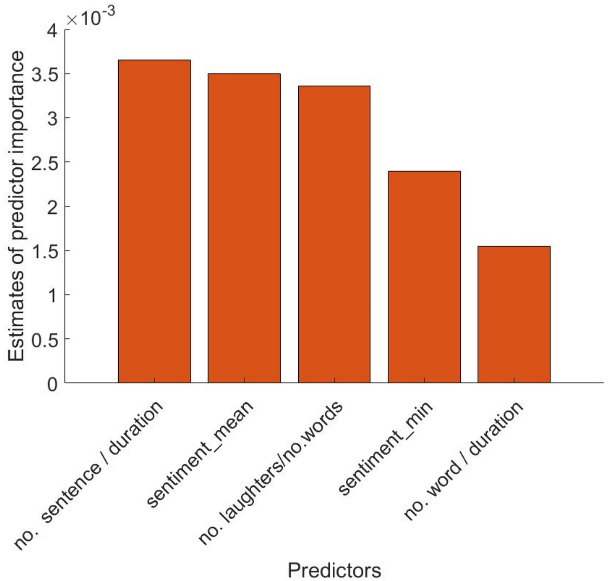


Selected biomarkers	Females	p-value	Males	p-value
Video biomarkers	Left eye distance (38,42)	<0.0001	Left eye distance (39,41)	<0.0001
	Right eye distance (44,48)	<0.0001	Left eye distance (38,42)	0.0028
	Mouth distance (52,58)	0.0014	Right eye distance (45,47)	0.0012
	Mouth angle (50,51)	0.001	Mouth angle (61,90)	<0.0001
	Eyebrow distance (22,23)	<0.0001	Mouth angle (60,59)	0.0311
	Left eye distance (39,41)	<0.0001	Left eye distance (38,42)	<0.0001
	Left eye distance (38,42)	<0.0001	Right eye distance (44,48)	<0.0001
	Right eye distance (45,47)	0.0063	Mouth distance (52,58)	0.6510
	Mouth angle (61,90)	<0.0001	Mouth angle (50,51)	0.0052
	Mouth angle (60,59)	<0.0001	Eyebrow distance (22,23)	0.1000

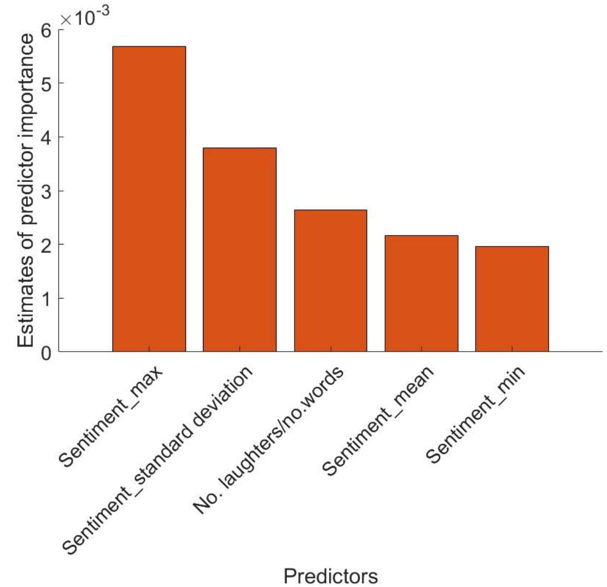
p-value of the selected top 5 significant biomarkers for females and males

Text Biomarkers

Text biomarkers for females



Text biomarkers for males



Selected biomarkers	Females	p-value	Males	p-value
Text biomarkers	No.sentence/duration	<0.0001	Sentiment_max	0.003
	Sentiment_mean	<0.0001	Sentiment_standard deviation	<0.0001
	No.laughters/no.words	0.032	No.laughters/no.words	0.126
	Sentiment_min	0.887	Sentiment_mean	0.366
	No. word/duration	0.311	Sentiment_min	0.369
	Sentiment_standard deviation	0.0036	No.sentence/duration	0.322

p-value of the selected top 5 significant biomarkers for females and males

Outline

- Overview
- Data Analytics for Disease Management
  - *Topic I: New methods for personalization in disease modeling and monitoring*
  - *Topic II: Detection of depression from communication*
- Highlights of other Works
- Conclusion

Surgical Site Infection (SSI)


Survey Case 1

Resize font:

Case 1

A 34 year old female had surgery to remove her spleen. She is fit and otherwise a healthy, non-smoker, without diabetes.

The day after her surgery she reports a low-grade fever and no fluid draining from the wound.

A photograph of her wound from this day is displayed below:

Based on the information given, does this patient have a surgical site infection?
* must provide value

- Yes
- No

reset

On a scale from 0 to 10 with 0 being least confident and 10 being most confident, how confident are you in your answer above?
* must provide value

0 1 2 3 4 5 6 7 8 9 10

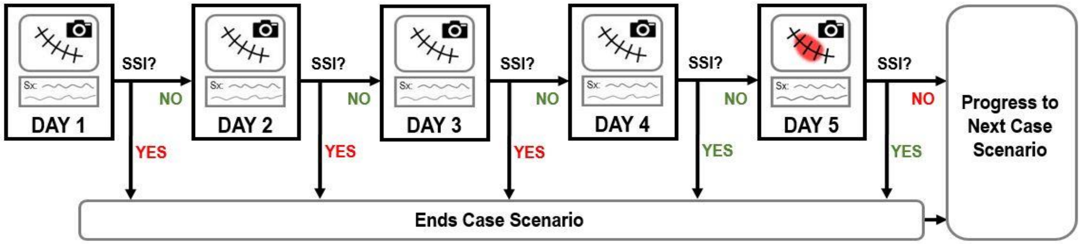
reset

What would you suggest they do next?
* must provide value

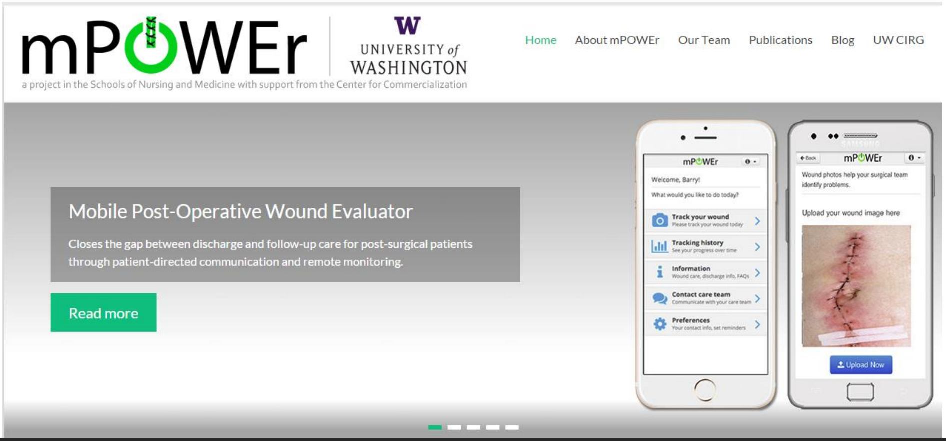
- Look at the wound again tomorrow
- Call their primary care provider for advice
- Call their surgeon to be seen in clinic in the next 2-3 days
- Call their surgeon for a clinic visit today
- Go to the nearest Emergency Department

reset

Next Page >>

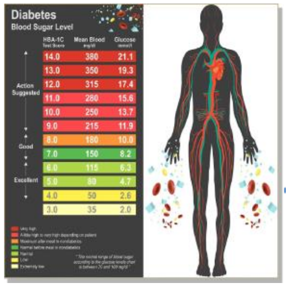


Wound photography could be used to monitor patients remotely after discharge, but timely review of the large volume of photos generated by an outpatient wound monitoring system may not be feasible by surgeons. **Crowdsourcing** this task may provide a viable alternative method of SSI detection



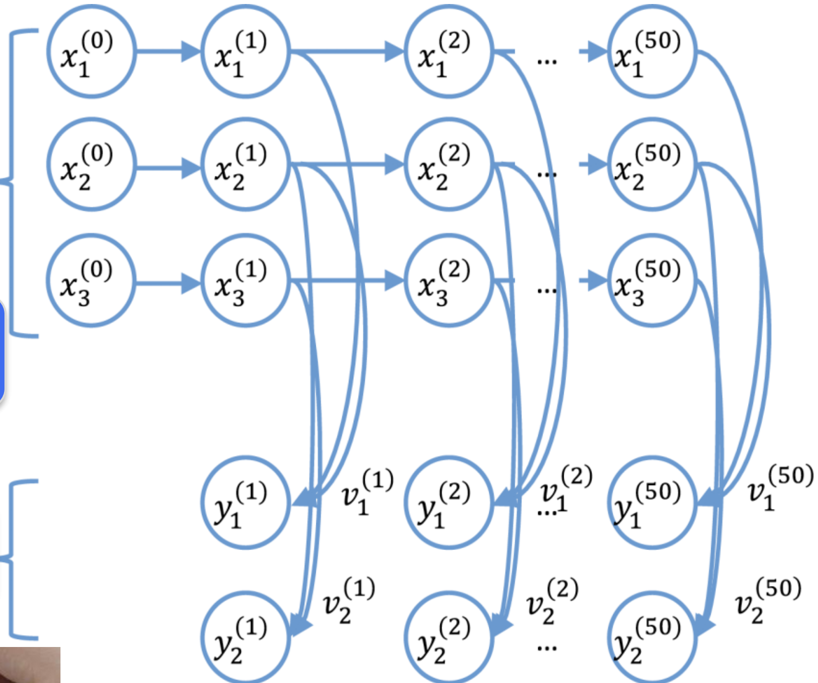
Surgery 2014, JACS 2016, JBI 2017, SI 2019

Dynamic Inspection of Latent Variables in State-Space Systems



Latent

Obs



Individual's health (latent) evolves over time

$$\mathbf{x}^{(t+1)} = \mathbf{B}\mathbf{x}^{(t)} + \boldsymbol{\epsilon}^{(t)}$$
$$\boldsymbol{\epsilon}^{(t)} \sim N(0, \mathbf{Q}).$$

"Cheap" measurements (observed) using wearable sensors

$$\mathbf{y}^{(t+1)} = \mathbf{Z}\mathbf{x}^{(t+1)} + \mathbf{v}^{(t)}$$
$$\mathbf{v}^{(t)} \sim N(0, \mathbf{R}).$$

1232 IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, VOL. 16, NO. 3, JULY 2019

Dynamic Inspection of Latent Variables in State-Space Systems

Tianshu Feng, Xiaoning Qian, *Senior Member, IEEE*, Kaibo Liu, *Member, IEEE*, and Shuai Huang, *Member, IEEE*

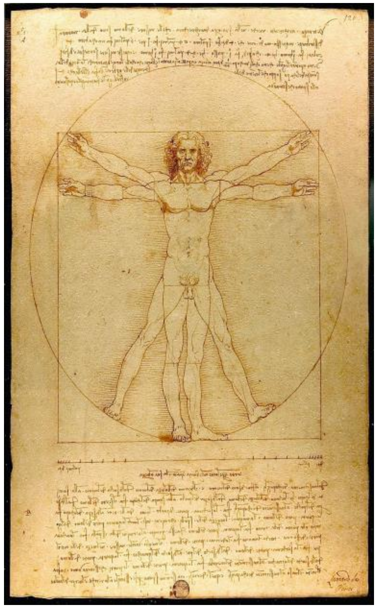
https://time.com/4703099/continuous-glucose-monitor-blood-sugar-diabetes/

JBI 2017; JHIR 2018; AISTAT 2019; SI 2019

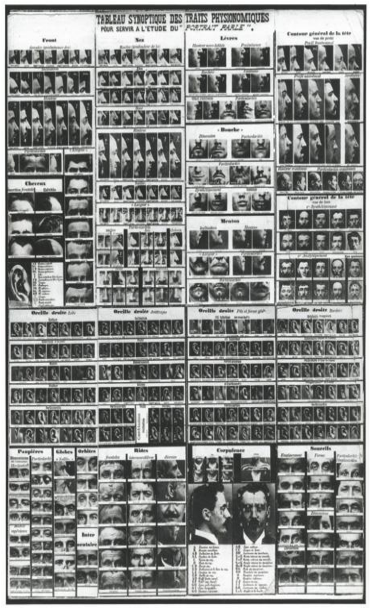
Outline

- Overview
- Data Analytics for Disease Management
  - *Topic I: New methods for personalization in disease modeling and monitoring*
  - *Topic II: Detection of depression from communication*
- Highlights of other Works
- Conclusion

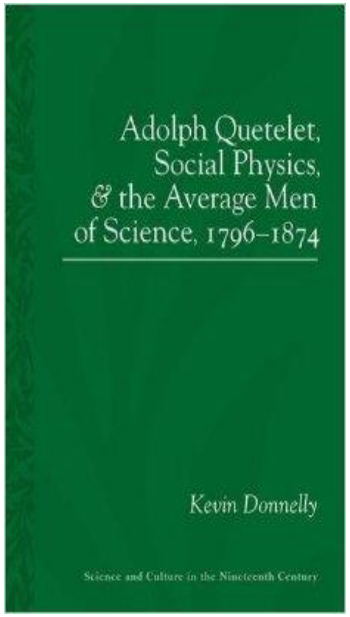
Norm and Derivation – An Old Song Sung to a New Tune of Data Science



Leonardo da Vinci's "***Vitruvian Man***"



Alphonse Bertillon's synoptic table of physiognomic traits



Quetelet's "***Average men***"

- ❖ Boundaries between disciplines are vanishing
- ❖ Drawing boundary is an important skill for engineers!

292 THE PATIENT WILL SEE YOU NOW

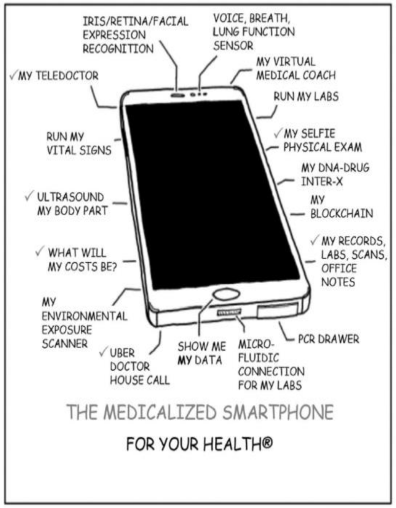
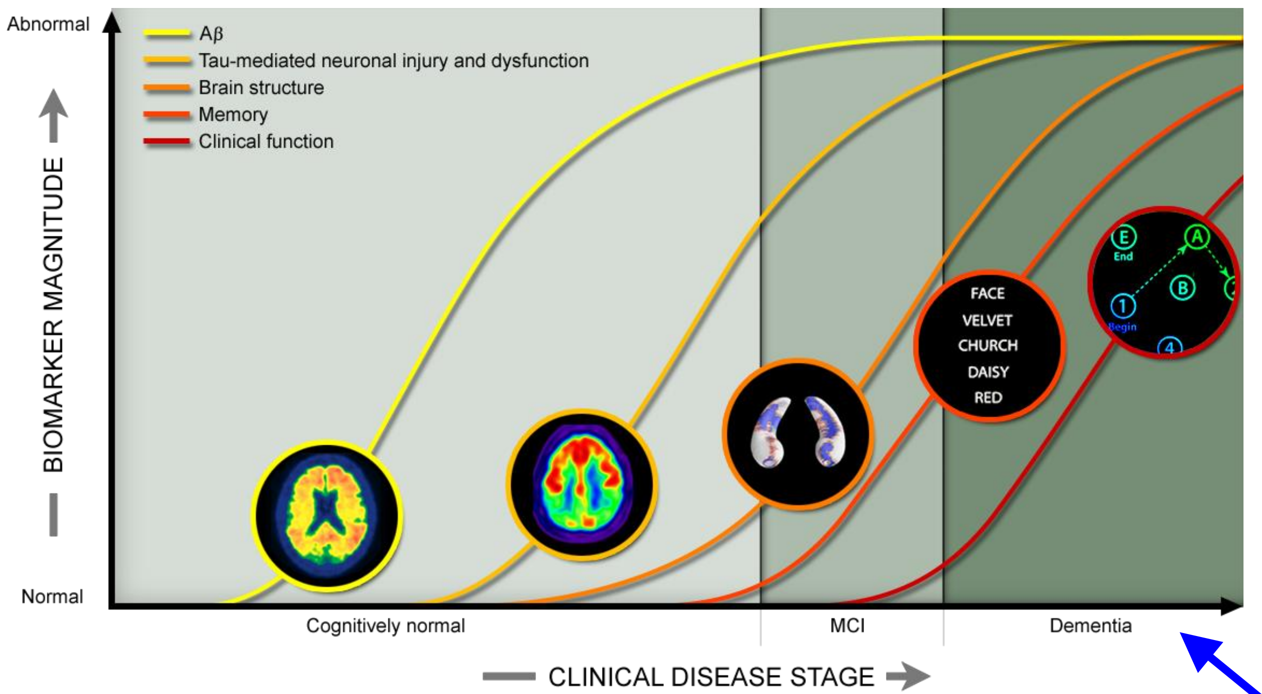
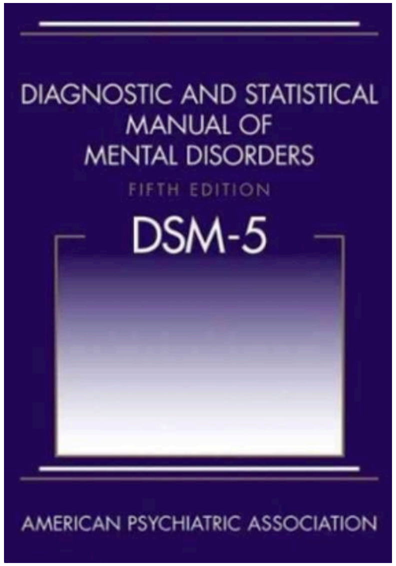


FIGURE E.1: The medicalized smartphone of the future. Check marks indicate functions that are now operational, at least in part. Adapted from xkcd.com.

https://en.wikipedia.org/wiki/Vitruvian_Man

What is Alzheimer's Disease?

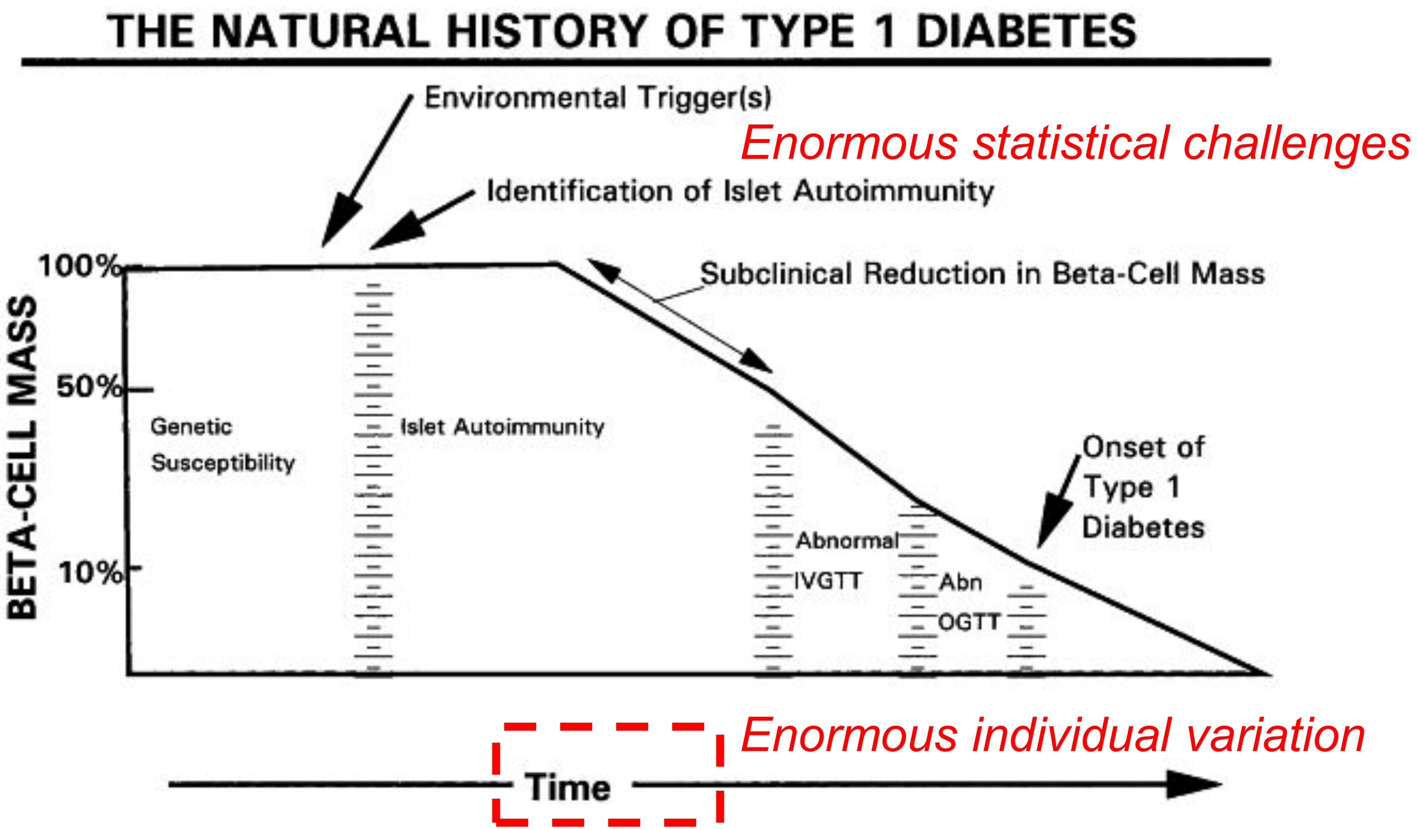




- A diseased condition has a definition that is usually in the later stage of the progression
- *Or, shall we take the disease as a process*

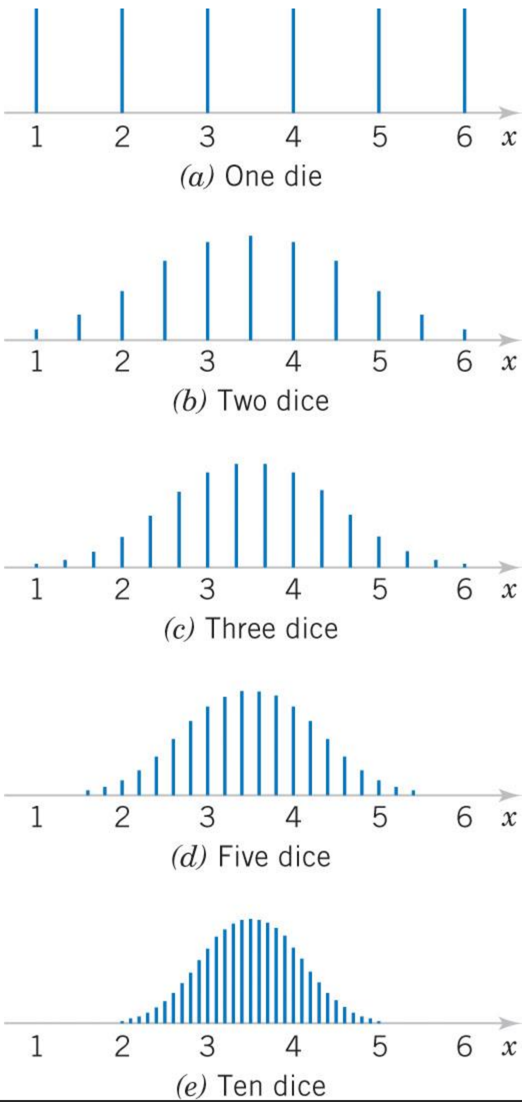
http://adni.loni.usc.edu/

What is Type 1 Diabetes (T1D)?



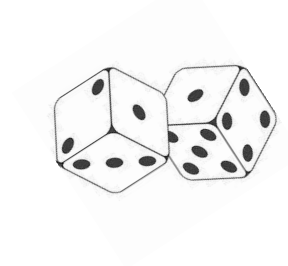
- Risk prediction and monitoring using complex biomarkers
- Seek of surrogate endpoints
- Answer questions regarding progression rate, time to onset, etc.
- Mechanistic understanding: identify environmental triggers, regulators

When does a System Emerge?



Throwing 1 die

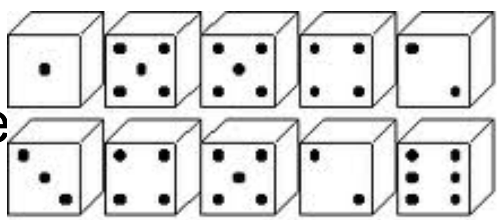
Throwing 2 dice



Throwing 3 dice

Throwing 5 dice

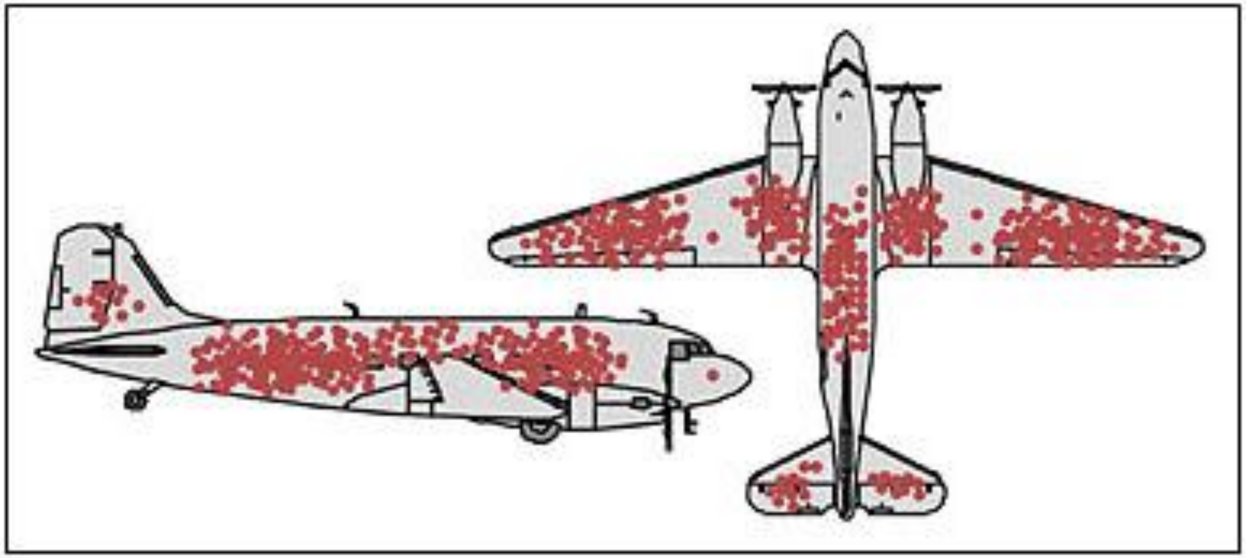
Throwing 10 dice



Models are Important, same as Domain Insight

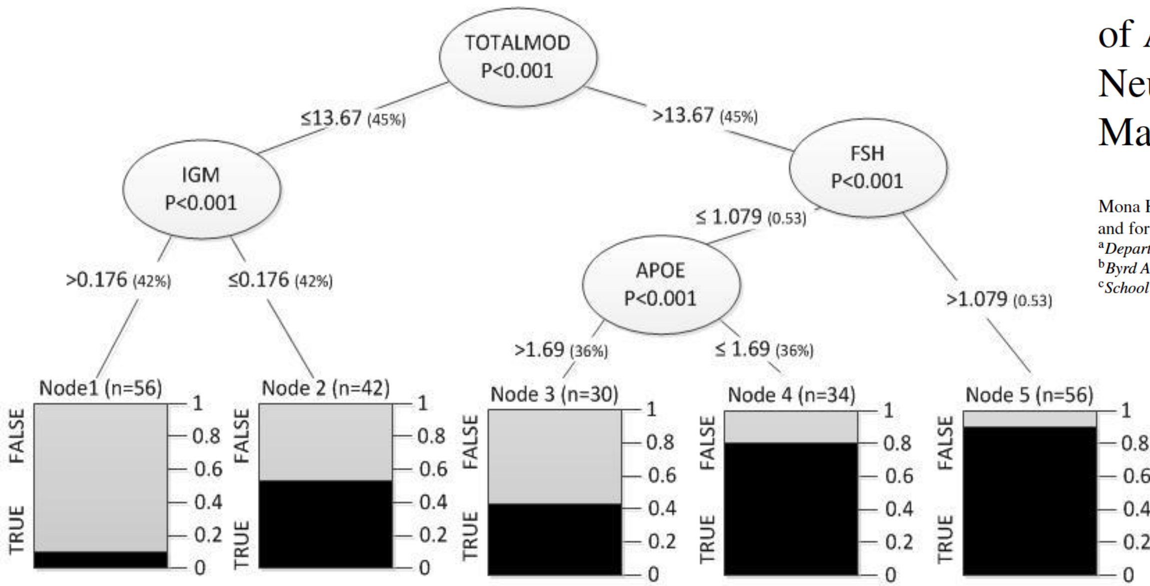
The story of the statistician Abraham Wald in World War II

- The Allied AF lost many aircrafts, so they decided to armor their aircrafts up
- However, limited resources are available – which parts of the aircrafts should be armored up?
- Abraham Wald stayed in the runaway, to catalog the bullet holes on the returning aircrafts



Credit: Cameron Moll

Communication with Multidisciplinary Experts Gives You New Perspectives



Identifying Cost-Effective Predictive Rules of Amyloid-β Level by Integrating Neuropsychological Tests and Plasma-Based Markers

Mona Haghighi[a], Amanda Smith[b], Dave Morgan[b], Brent Small[c], Shuai Huang[a,b,*]
and for the Alzheimer's Disease Neuroimaging Initiative (ADNI)[1]
[a]*Department of Industrial and Management Systems Engineering, University of South Florida, Tampa, FL, USA*
[b]*Byrd Alzheimer's Institute, University of South Florida, Tampa, FL, USA*
[c]*School of Aging Studies, University of South Florida, Tampa, FL, USA*

Why 60% accuracy is still very valuable

- ❖ **Anti-amyloid clinical trials need large-scale screening: $3,000 per PET scan**
- ❖ **If the PET scan shows negative result, $3,000 is a waste**
- ❖ **Blood measurements cost $200 per visit**
- ❖ **Question: can we use blood measurements to predict the amyloid?**
- ❖ **Benefit: enrich the cohort pool with more amyloid positive cases**

Thank you. Questions?