Machine Learning for Healthcare: Disease progression modeling

David Sontag
Clinical Machine Learning Group
MIT

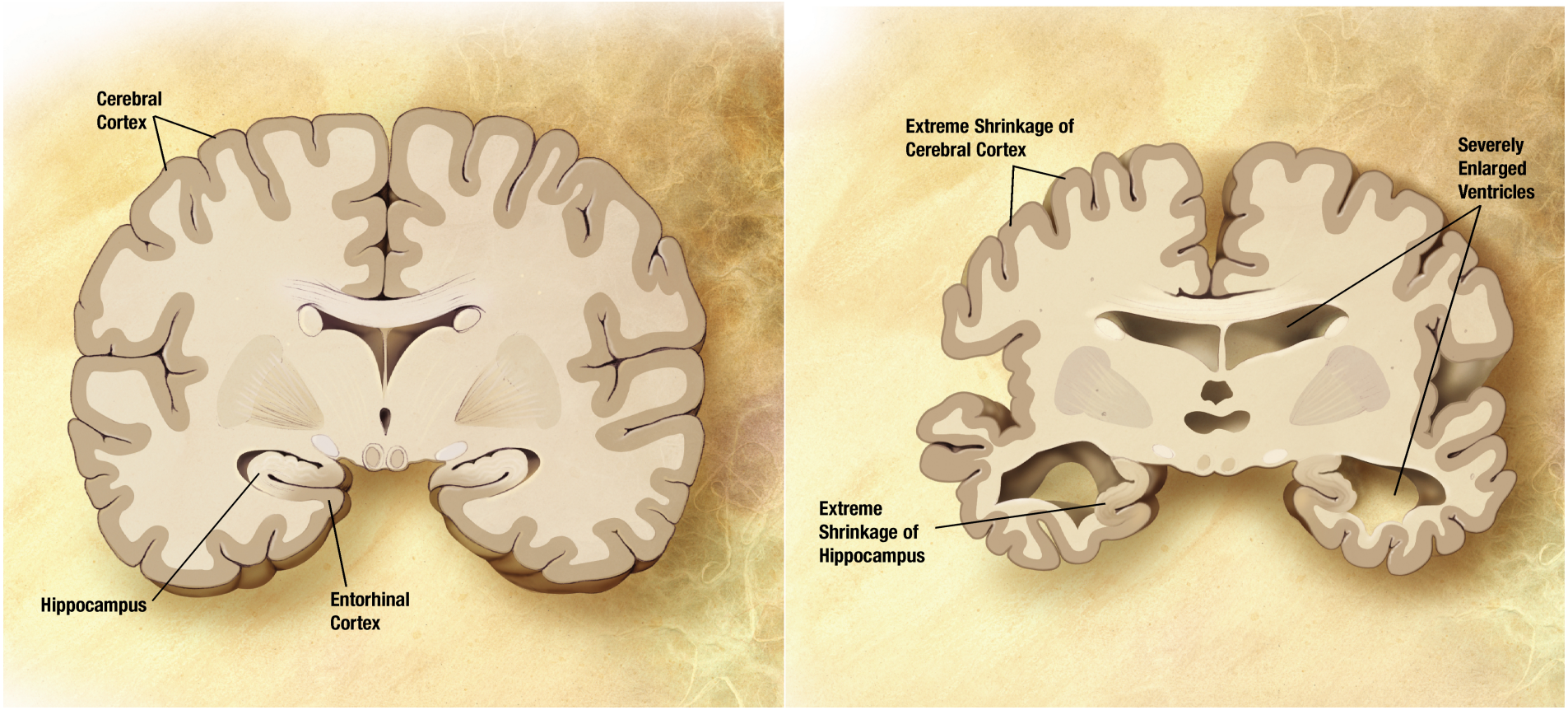
Goals of disease progression modeling

- Predictive:
  - *What will this patient's future trajectory look like?*
- Descriptive:
  - *Find markers of disease stage and progression, statistics of what to expect when*
  - *Discover new disease subtypes*
- Key challenges we will tackle:
  - Seldom directly observe disease stage, but rather only indirect observations (e.g. symptoms)
  - Data is censored – don't observe beginning to end

Goals of disease progression modeling

- **Predictive:**
  - ***What will this patient's future trajectory look like?***
- Descriptive:
  - *Find markers of disease stage and progression, statistics of what to expect when*
  - *Discover new disease subtypes*
- Key challenges we will tackle:
  - Seldom directly observe disease stage, but rather only indirect observations (e.g. symptoms)
  - Data is censored – don't observe beginning to end

Predicting disease progression in Alzheimer's disease



[Image credit: Wikipedia; "Alzheimer's Disease Education and Referral Center, a service of the National Institute on Aging."]

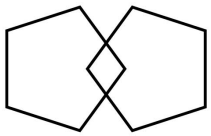
Disease status quantified by cognitive score (continuous valued)

MINI MENTAL STATE EXAMINATION (MMSE)

Name:

DOB:

Hospital Number:

One point for each answer	DATE:		
ORIENTATION Year Season Month Date Time	...../ 5	...../ 5	...../ 5
Country Town District Hospital Ward/Floor	...../ 5	...../ 5	...../ 5
REGISTRATION Examiner names three objects (e.g. apple, table, penny) and asks the patient to repeat (1 point for each correct. THEN the patient learns the 3 names repeating until correct).	...../ 3	...../ 3	...../ 3
ATTENTION AND CALCULATION Subtract 7 from 100, then repeat from result. Continue five times: 100, 93, 86, 79, 65. (Alternative: spell "WORLD" backwards: DLROW).	...../ 5	...../ 5	...../ 5
RECALL Ask for the names of the three objects learned earlier.	...../ 3	...../ 3	...../ 3
LANGUAGE Name two objects (e.g. pen, watch).	...../ 2	...../ 2	...../ 2
Repeat "No ifs, ands, or buts".	...../ 1	...../ 1	...../ 1
Give a three-stage command. Score 1 for each stage. (e.g. "Place index finger of right hand on your nose and then on your left ear").	...../ 3	...../ 3	...../ 3
Ask the patient to read and obey a written command on a piece of paper. The written instruction is: "Close your eyes".	...../ 1	...../ 1	...../ 1
Ask the patient to write a sentence. Score 1 if it is sensible and has a subject and a verb.	...../ 1	...../ 1	...../ 1
COPYING: Ask the patient to copy a pair of intersecting pentagons 	...../ 1	...../ 1	...../ 1
TOTAL:	...../ 30	...../ 30	...../ 30

MMSE scoring

24-30: no cognitive impairment
18-23: mild cognitive impairment
0-17: severe cognitive impairment

Predicting disease progression in Alzheimer's disease

- Goal: Predict disease status in *6*, *12*, *24*, *36, and 48 months*
- Five different regression tasks?
- Challenge: data sparsity
  - Total number of patients is small
  - Labels are noisy
  - Due to censoring, fewer patients at later time points

[Zhou et al., Modeling Disease Progression via Fused Sparse Group Lasso, KDD '12]

Predicting disease progression in Alzheimer's disease

- Goal: Predict disease status in *6*, *12*, *24*, *36, and 48 months*
- Five different regression tasks?
- Challenge: data sparsity

Number of patients M months after baseline
(Alzheimer's Disease Neuroimaging Initiative)

M06	M12	M24	M36	M48
648	642	569	389	87

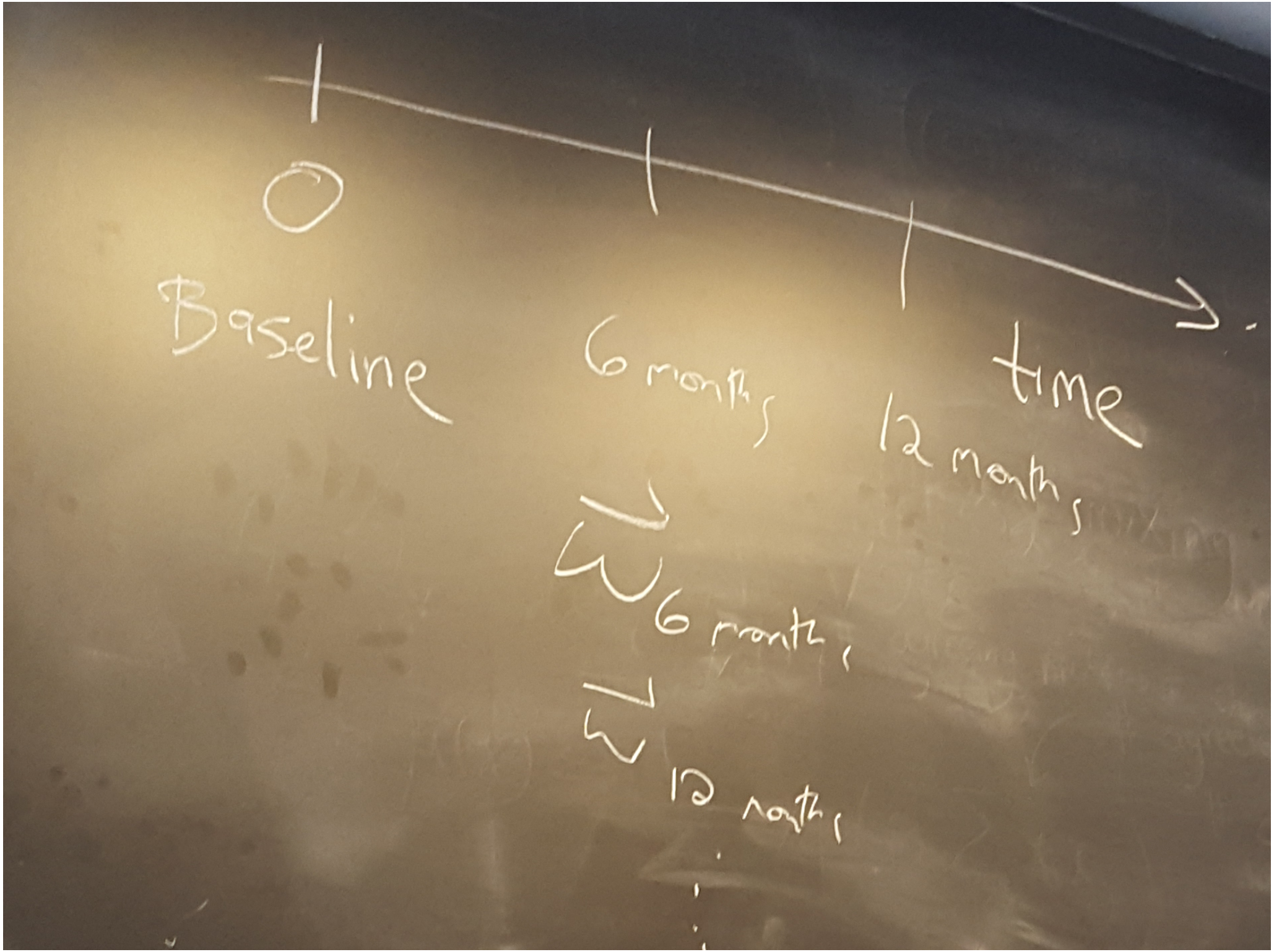
M06 = 6 months after baseline

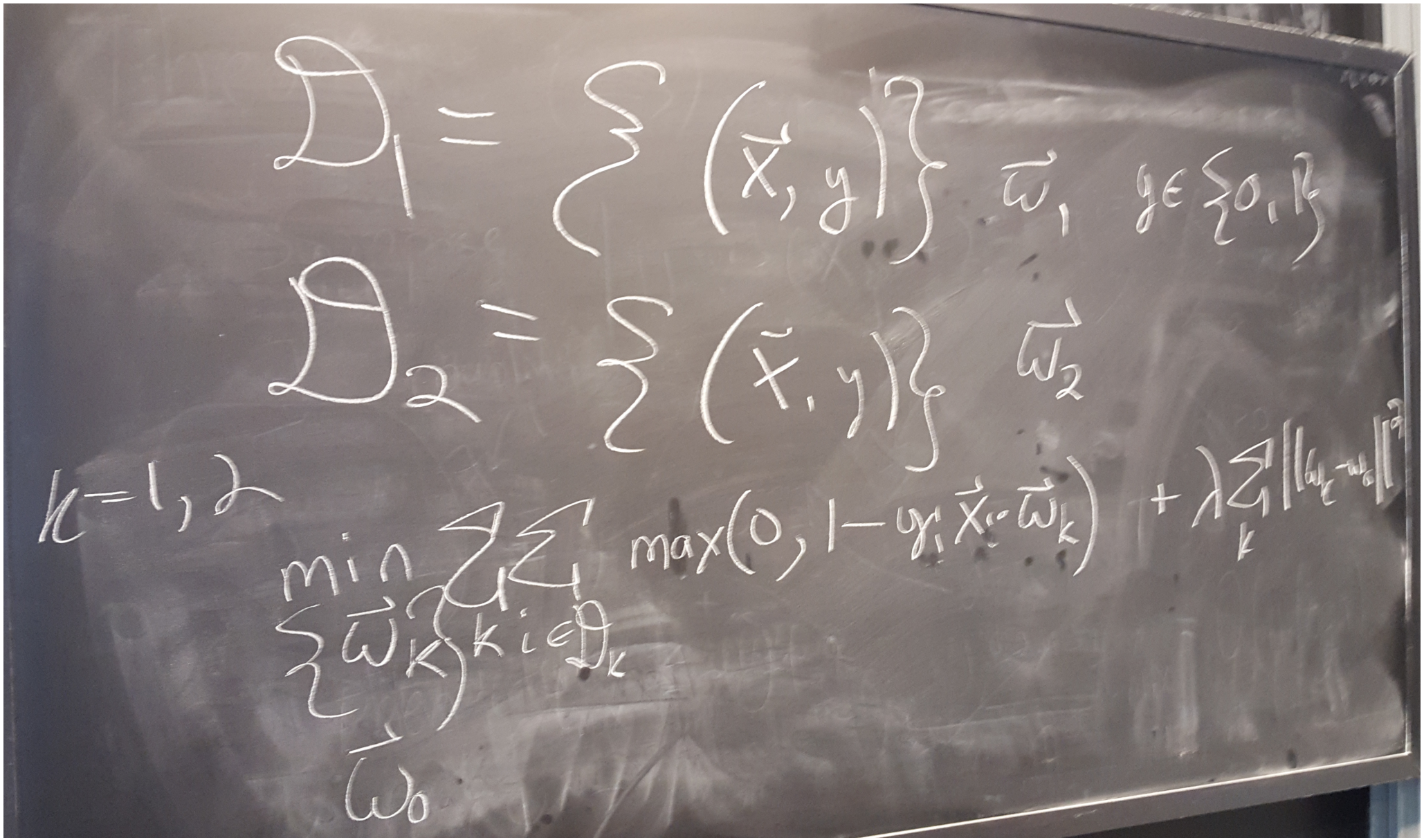
[Zhou et al., Modeling Disease Progression via Fused Sparse Group Lasso, KDD '12]

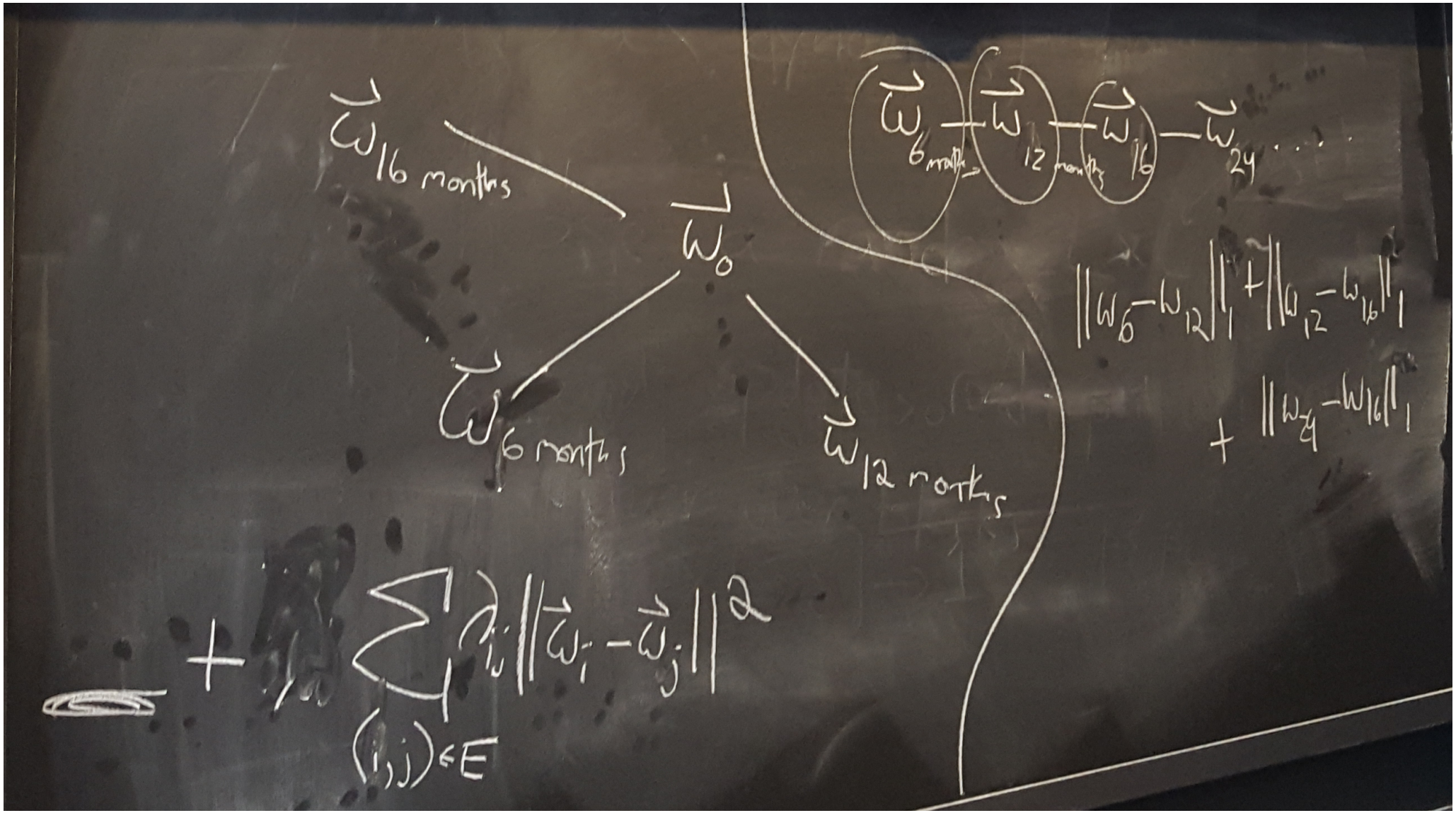
Multi-task learning

- Goal: Predict disease status in *6*, *12*, *24*, *36, and 48 months*
- Rather than learn several independent models, view as *multi-task* learning
  - Select common set of biomarkers for all time points
  - Also allow for specific set of biomarkers at different time points
  - Incorporate temporal smoothness in models

[Zhou et al., Modeling Disease Progression via Fused Sparse Group Lasso, KDD '12]


0
1
1
Baseline
6 months
12 months
time
6 months
12 months


$D_1 = \{(\vec{x}, y)\}$ $\vec{w}_1$ $y \in \{0, 1\}$
$D_2 = \{(\vec{x}, y)\}$ $\vec{w}_2$
$k = 1, 2$
$\min \sum_{k} \sum_{i} \max(0, 1 - y_i \vec{x}_i \cdot \vec{w}_k) + \lambda \sum_{k} \|w_k - w_0\|^2$
$\{\vec{w}_k\}^k \; i \in D_k$
$\vec{w}_0$


$\vec{w}_{16\ months}$
$\vec{w}_0$
$\vec{w}_{6\ months}$
$\vec{w}_{12\ months}$
$\vec{w}_{6\ months}$ — $\vec{w}_{12\ months}$ — $\vec{w}_{16}$ — $\vec{w}_{24}$ . . .
$\|w_6 - w_{12}\|_1 + \|w_{12} - w_{16}\|_1$
$+ \|w_{24} - w_{16}\|_1$
$+ \lambda \sum_{(i,j) \in E} \beta_{ij} \|\vec{w}_i - \vec{w}_j\|^2$

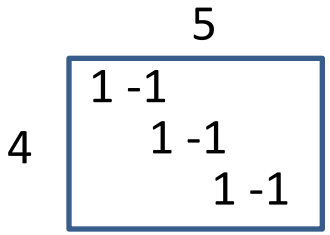
Convex fused sparse group lasso

- Simultaneously learn all 5 models by solving the following convex optimization problem:

$$\min_W L(W) + \lambda_1 \|W\|_1 + \lambda_2 \left\|RW^T\right\|_1 + \lambda_3 \|W\|_{2,1}$$

- Squared loss: $L(W) = \|S \odot (XW - Y)\|_F^2$
  (*S* is a mask to account for labels missing in subset of tasks)

- Group Lasso penalty $\|W\|_{2,1}$ given by $\sum_{i=1}^{d} \sqrt{\sum_{j=1}^{t} W_{ij}^2}$
- R =



[Zhou et al., KDD '12]

Features

MRI scans (white matter parcellation volume, etc.) +

Demographic	age, years of education, gender
Genetic	ApoE-$\varepsilon 4$ information
Baseline cognitive scores	MMSE, ADAS-Cog, ADAS-MOD, ADAS sub-scores, CDR, FAQ, GDS, Hachinski, Neuropsychological Battery, WMS-R Logical Memory
Lab tests	RCT1, RCT11, RCT12, RCT13, RCT14, RCT1407, RCT1408, RCT183, RCT19, RCT20, RCT29, RCT3, RCT392, RCT4, RCT5, RCT6, RCT8

371 in total

[Zhou et al., Modeling Disease Progression via Fused Sparse Group Lasso, KDD '12]

Results (averaged over 5 time points)

Baseline – independent regressors

Temporal smoothing helps!

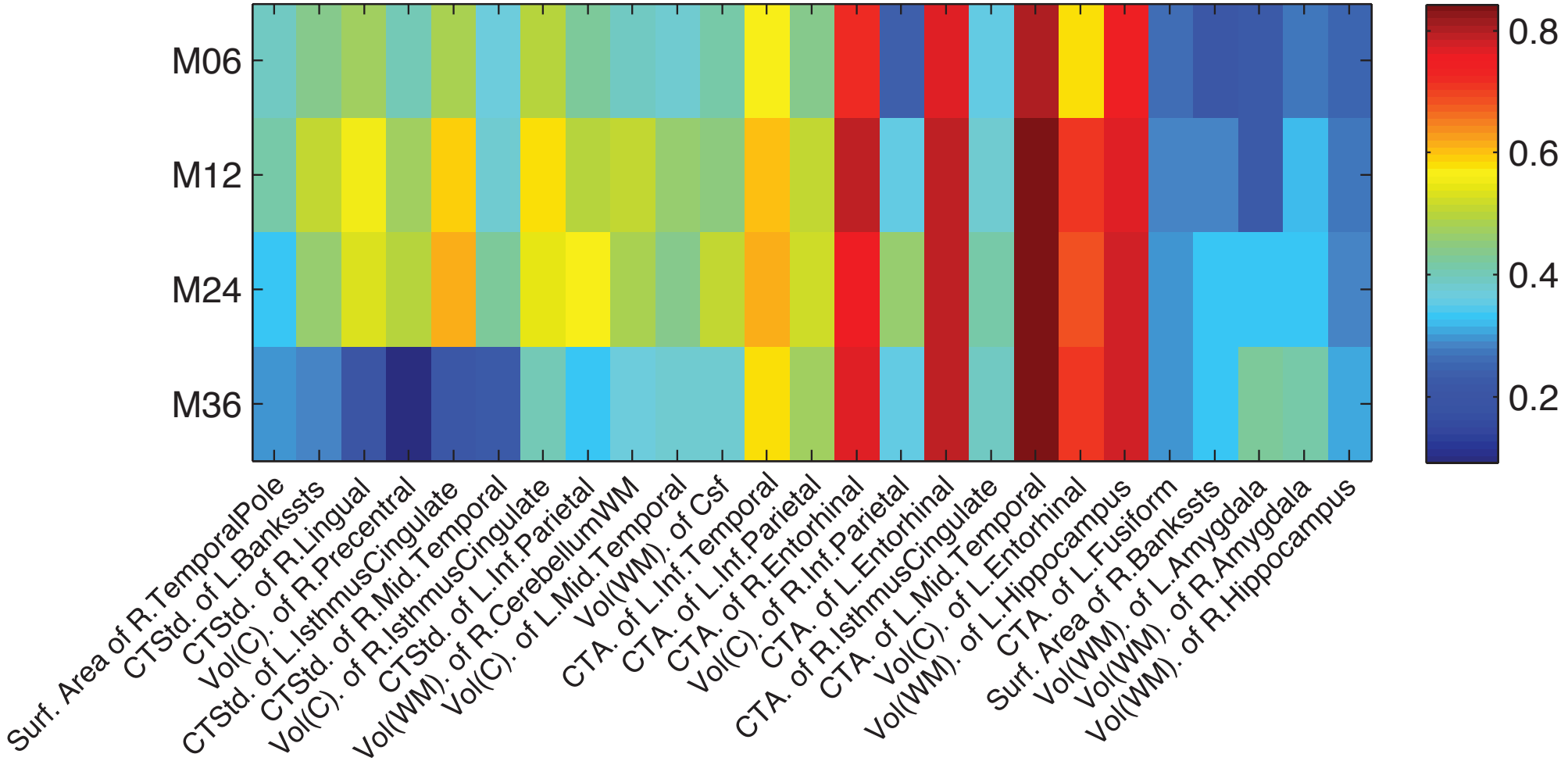
	Ridge	cFSGL1 $\lambda_2 = 20$	cFSGL2 $\lambda_2 = 50$	cFSGL3 $\lambda_2 = 100$
		Target: MMSE		
nMSE	0.548 ± 0.057	0.428 ± 0.052	0.400 ± 0.053	$\mathbf{0.395 \pm 0.052}$
R	0.689 ± 0.030	0.772 ± 0.030	0.790 ± 0.032	$\mathbf{0.796 \pm 0.031}$

nMSE – normalized mean squared error. Smaller is better

R – average R^2 (correlation coefficient). Larger is better

$$\min_W L(W) + \lambda_1 \|W\|_1 + \lambda_2 \left\|RW^T\right\|_1 + \lambda_3 \|W\|_{2,1}$$

Feature importance varies by time



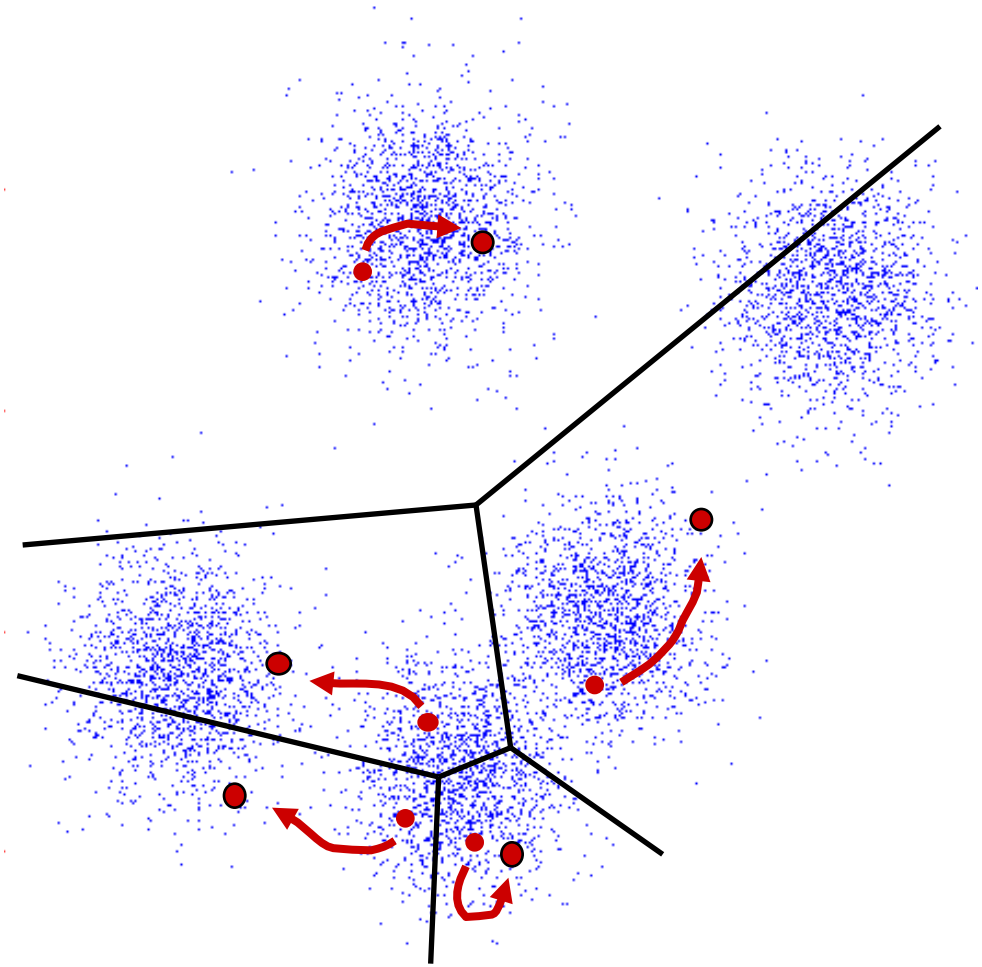
(a) Target: ADAS-Cog (25 stable features)

Goals of disease progression modeling

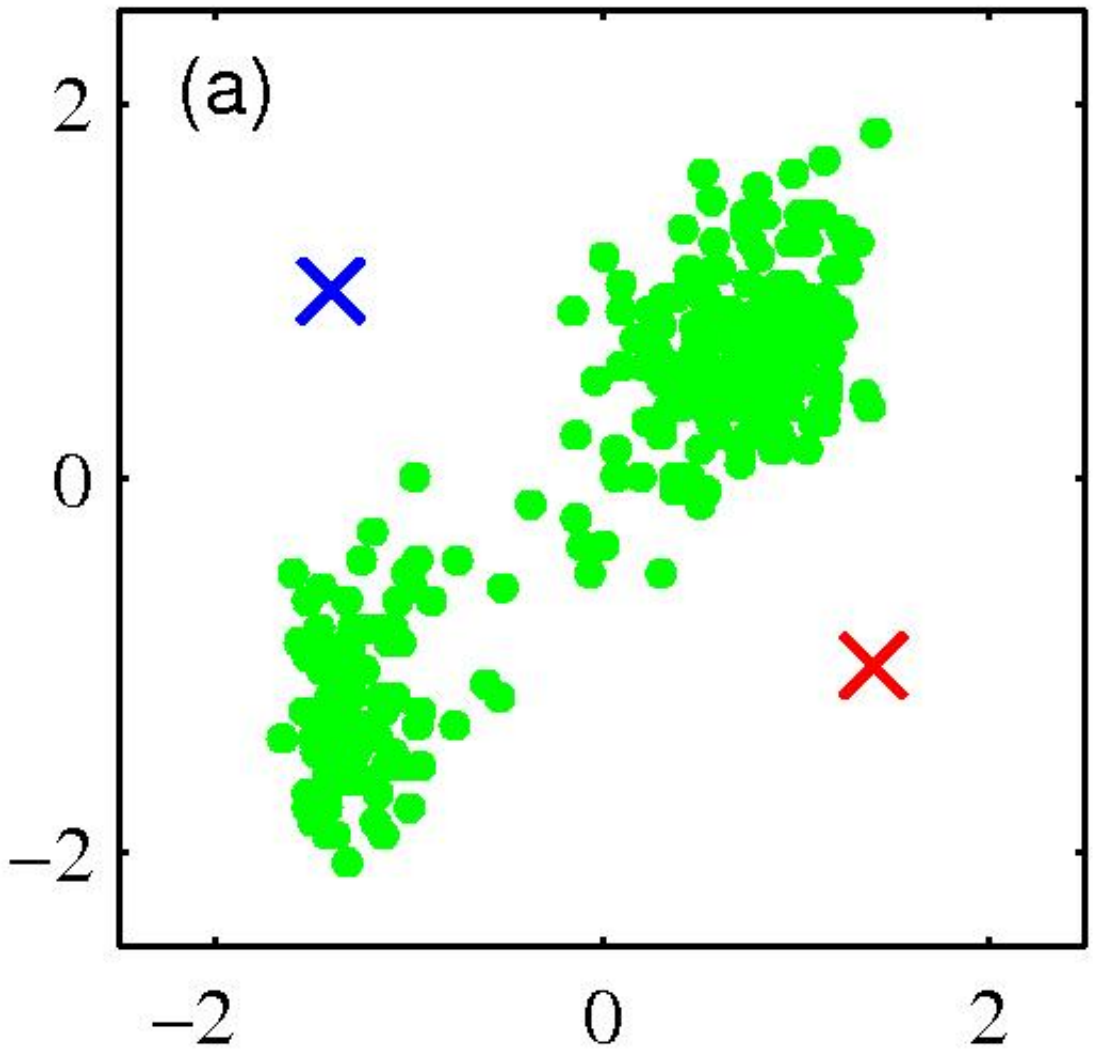
- Predictive:
  - *What will this patient's future trajectory look like?*
- **Descriptive:**
  - *Find markers of disease stage and progression, statistics of what to expect when*
  - ***Discover new disease subtypes***
- Key challenges we will tackle:
  - Seldom directly observe disease stage, but rather only indirect observations (e.g. symptoms)
  - Data is censored – don't observe beginning to end

K-Means

- An iterative clustering algorithm
  - Initialize: Pick *K* random points as cluster centers
  - Alternate:
    1. Assign data points to closest cluster center
    2. Change the cluster center to the average of its assigned points
  - Stop when no points' assignments change



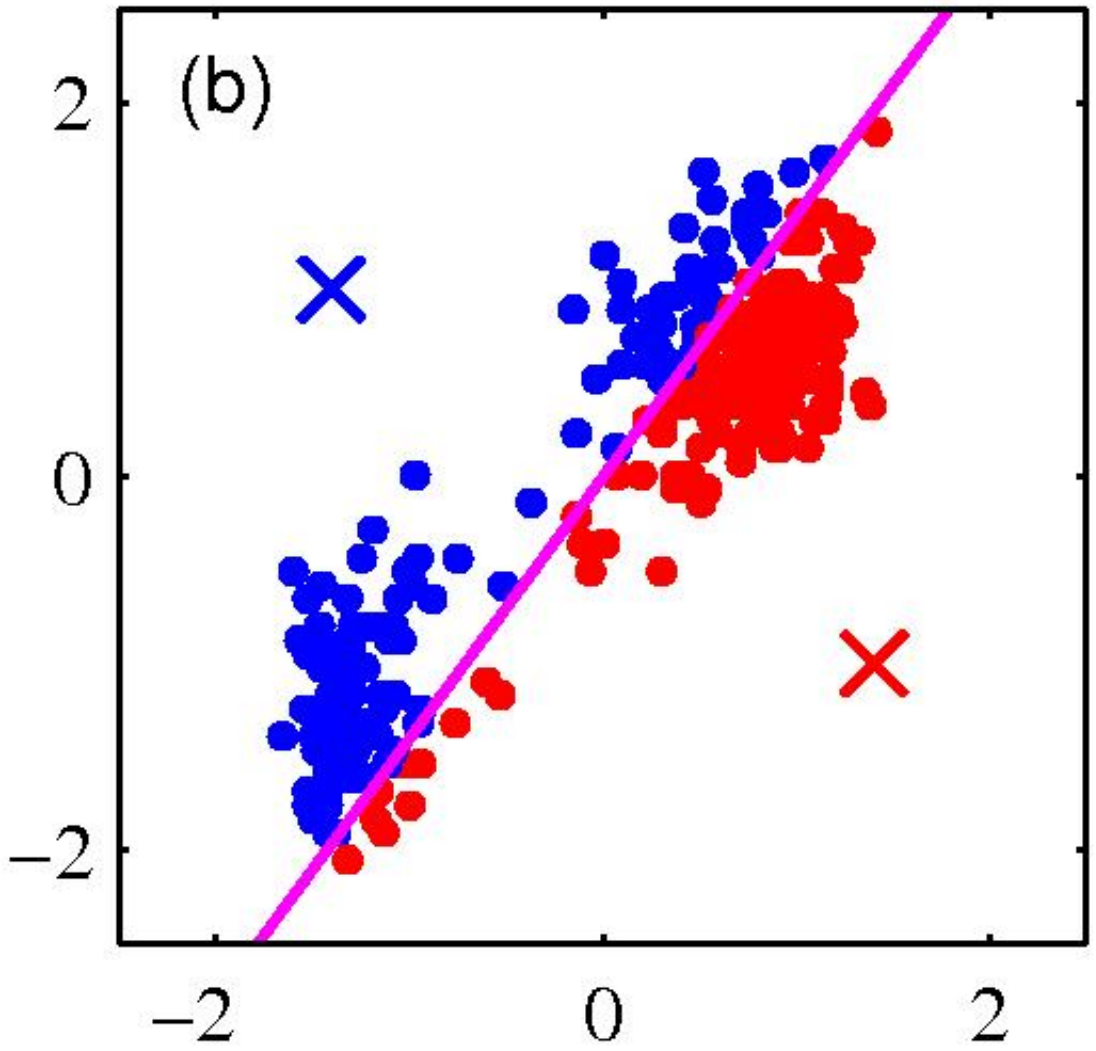
K-means clustering: Example



- Pick *K* random points as cluster centers (means)

Shown here for *K*=2

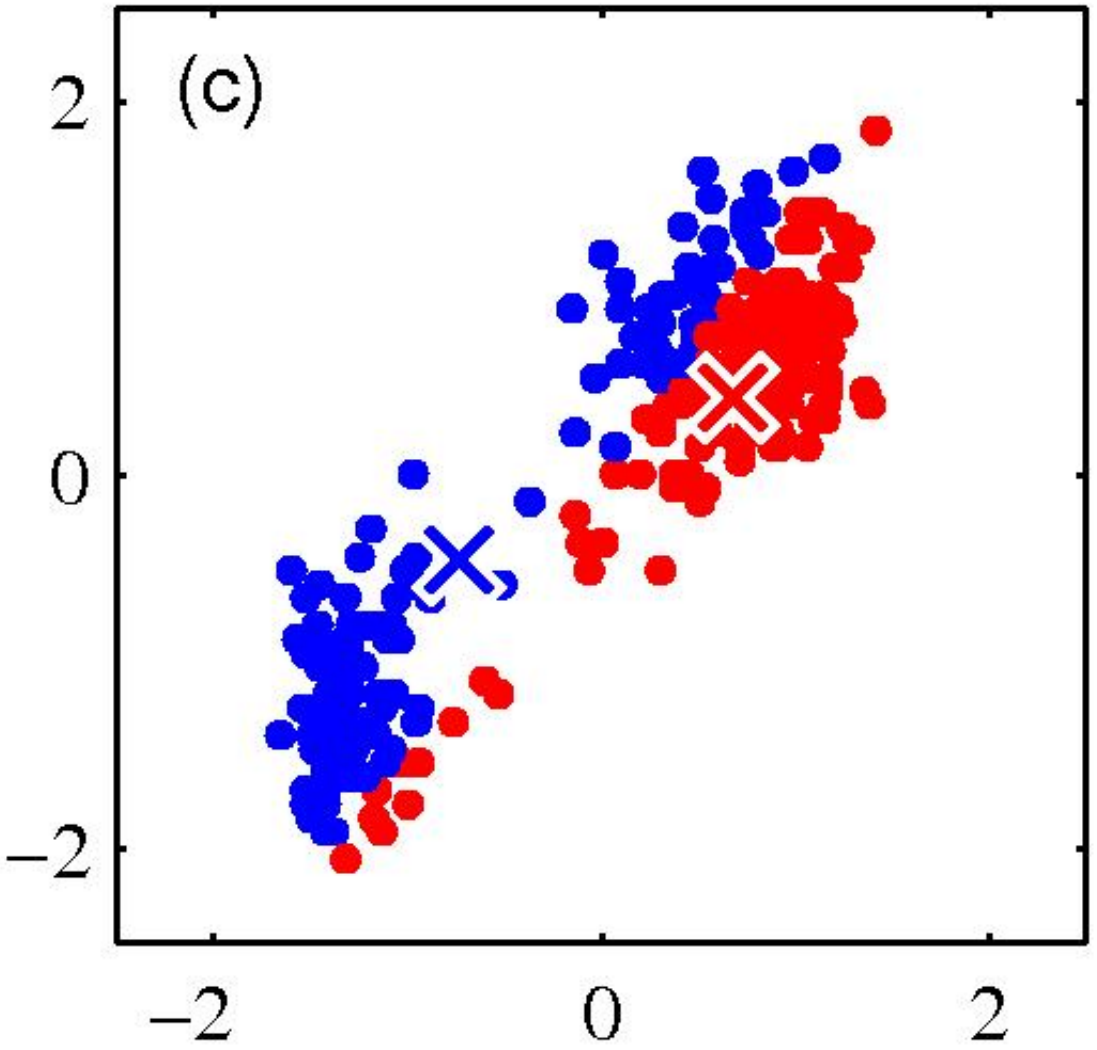
K-means clustering: Example



Iterative Step 1

- Assign data points to closest cluster center

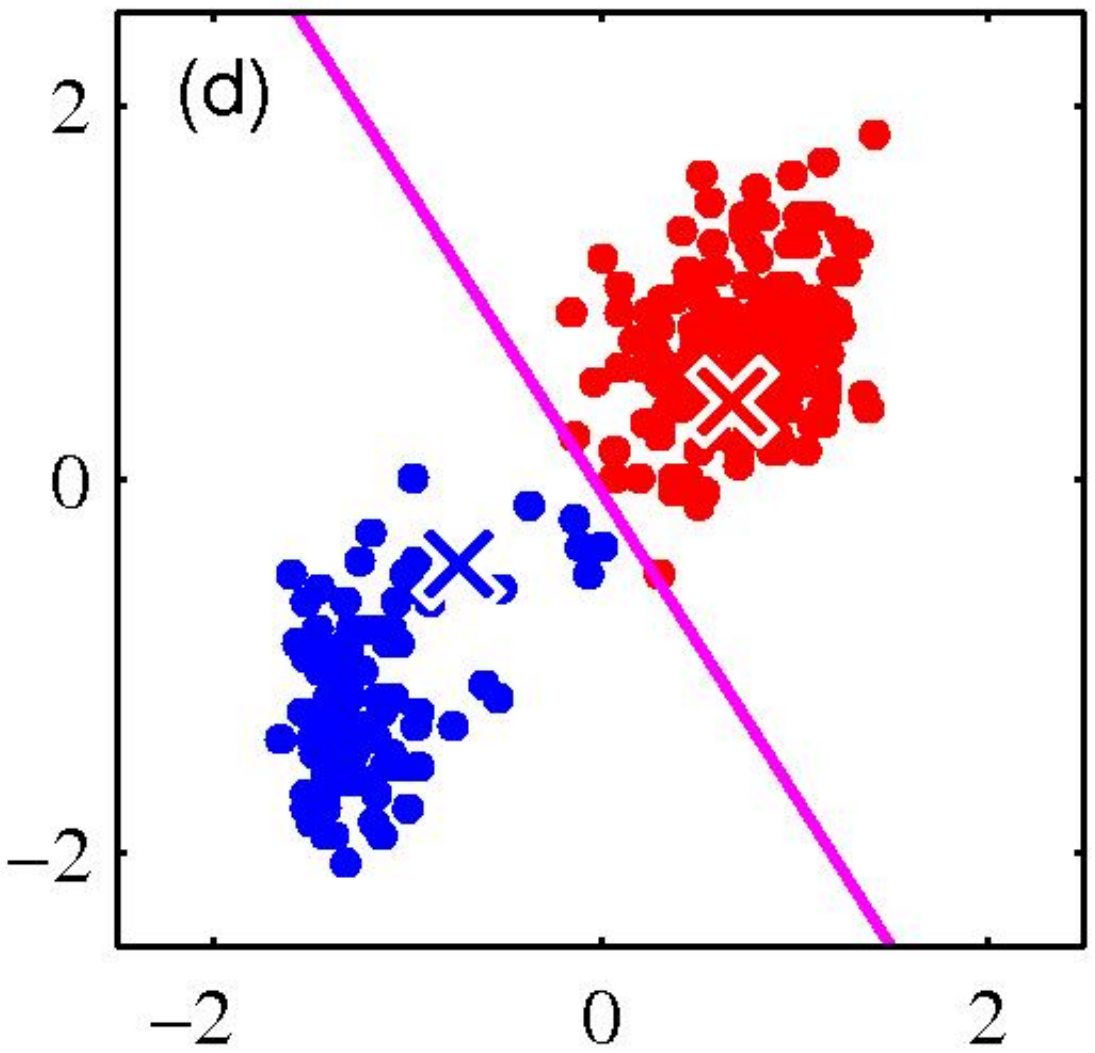
K-means clustering: Example



Iterative Step 2

- Change the cluster center to the average of the assigned points

K-means clustering: Example



- Repeat until convergence

Asthma: the problem

- 5 to 10% of people with severe asthma remain poorly controlled despite maximal inhaled therapy

[Holgate ST, Polosa R. The mechanisms, diagnosis, and management of severe asthma in adults. Lancet. 2006; 368:780–793]

[whatasthmais.com]

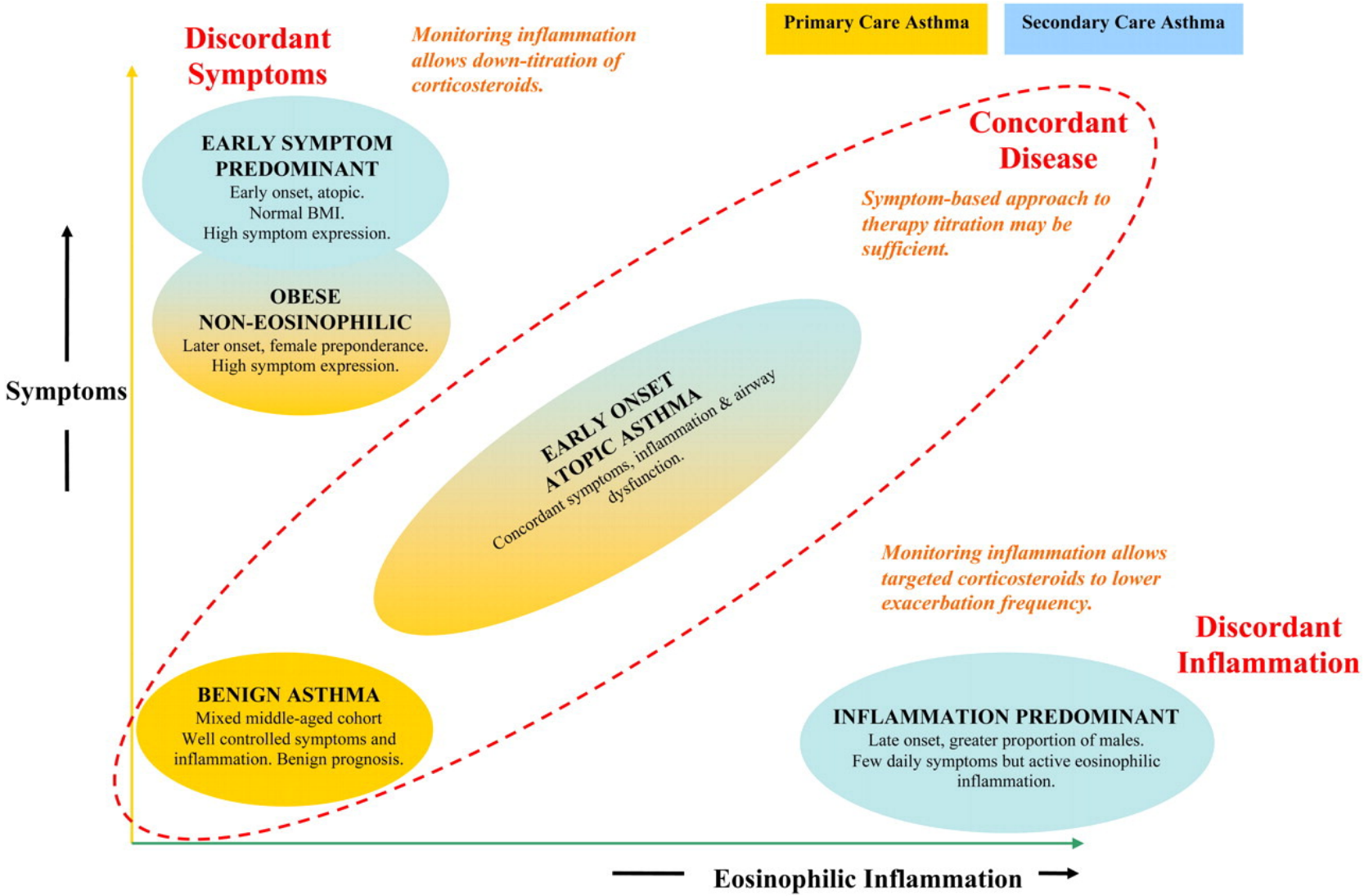
Asthma: the question

"It is now recognised that there are distinct asthma phenotypes and that distinct therapeutic approaches may only impinge on some aspects of the disease process within each subgroup"

- What are the processes (genetic or environmental) that underlie different subtypes of asthma?
- Which aspects of airway remodelling are important in disease subtypes?
- What are the best biomarkers of disease progression or treatment response?
- Why are some patients less responsive to conventional therapies than others?

[Adcock et al., "New targets for drug development in asthma". The Lancet, 2008]

Discovering subtypes from data



[Haldar et al., *Am J Respir Crit Care Med*, 2008]

The data

- All patients had physician diagnosis of asthma and at least one recent prescription for asthma therapy
- All were current nonsmokers
- *Data set #1*: 184 patients recruited from primary-care practices in the UK
- *Data set #2*: 187 patients from refractory asthma clinic in the UK
- *Data set #3*: 68 patients from 12 month clinical study
- Features: *z* scores for continuous variables, 0/1 for categorical
  - Some of the continuous variables log-transformed to approximate a normal distribution

[Haldar et al., *Am J Respir Crit Care Med*, 2008]

Comparison of Baseline Characteristics in the three Asthma Populations

Variable	Primary Care (*n = 184*)	Secondary Care (*n = 187*)	Longitudinal Cohort (*n = 68*)
Sex, % female	54.4	65.8	47.1
Age, yr (SD)	49.2 (13.9)	43.4 (15.9)	52.4 (14.6)
Age of onset, yr (SD)	24.7 (19)	20.3 (18.4)	31.1 (23.7)
Atopic status, % positive	72.8	73.8	57.4
Body mass index, kg/m^2 (SD)	27.5 (5.4)	28.5 (6.5)	28.0 (5.9)
PC_{20} methacholine†, mg/ml	1.04 (1.13)	†	0.67 (0.68)
Peak flow variability, amp % mean	17 (0.38)	32.2 (0.48)	13.8 (0.29)
FEV_1 change with bronchodilator, %	1.63 (1.16)	12.8 (0.41)	3.2 (1.04)
Post-bronchodilator FEV_1, % predicted	91.4 (21)	82.1 (21.1)	80.2 (20.6)
Sputum eosinophil count, %	1.32 (0.62)	2.9 (0.99)	2.4 (0.81)
F_{ENO}‡, ppb	31.6 (0.33)	43 (0.32)	4.32 (0.64)‡
Sputum neutrophil count, %	55.09 (0.31)	46.7 (0.32)	41.1 (0.35)
Modified JACS§ (SD)	1.36 (0.74)	2.02 (1.16)	1.42 (1.26)
Dose of inhaled corticosteroid, BDP equivalent/μg (SD)	632 (579)	1,018 (539)	1,821 (1,239)
Long-acting bronchodilator use, %	40.2	93	86.7

Definition of abbreviations: amp = amplitude; BDP = beclomethasone dipropionate; JACS = Juniper Asthma Control Score

[Haldar et al., *Am J Respir Crit Care Med*, 2008]

Clusters in primary care

(found by K-means)

Variable	Primary Care (*n* = 184)	Cluster 1 Early-Onset Atopic Asthma (*n* = 61)	Cluster 2 Obese Noneosinophilic (*n* = 27)	Cluster 3 Benign Asthma (*n* = 96)	Significance (*P* Value)*
Sex†, % female	**54.4**	45.9	81.5	52.1	0.006
Age, yr (SD)	**49.2 (13.9)**	44.5 (14.3)	53.9 (14)	50.8 (13)	0.003
Age of onset†, yr (SD)	**24.7 (19)**	14.6 (15.4)	35.3 (19.6)	28.2 (18.3)	<0.001
Atopic status†, % positive	**72.8**	95.1	51.9	64.6	<0.001
Body mass index†, kg/m^2 (SD)	**27.5 (5.4)**	26.1 (3.8)	36.2 (5.5)	26 (3.6)	<0.001
PC_{20} methacholine†‡, mg/ml	**1.04 (1.13)**	0.12 (0.86)	1.60 (0.93)	6.39 (0.75)	<0.001
PC_{20} >8 mg/ml, n (%)	**64 (34.7)**	2 (3.3)	6 (22.2)	56 (58.3)	<0.001
Peak flow variability†‡, amp % mean	**17 (0.38)**	20 (0.47)	21.9 (0.32)	14.8 (0.32)	0.039
FEV_1 change with bronchodilator‡, %	**1.63 (1.16)**	4.5 (0.91)	1.82 (1.16)	0.83 (1.22)	<0.001
Post-bronchodilator FEV_1, % predicted	**91.4 (21)**	86.9 (20.7)	91.5 (21.4)	94.2 (20.7)	0.107
Sputum eosinophil count†‡, %	**1.32 (0.62)**	3.75 (0.64)	1.55 (0.51)	0.65 (0.44)	<0.001
F_{ENO}‡§, ppb	**31.6 (0.33)**	57.5 (0.27)	25.8 (0.29)	22.8 (0.27)	<0.001
Sputum neutrophil count‡, %	**55.09 (0.31)**	45.87 (0.24)	72.71 (0.13)	57.56 (0.36)	0.038
Modified JACS† (SD)	**1.36 (0.74)**	1.54 (0.58)	2.06 (0.73)	1.04 (0.66)	<0.001
Dose of inhaled corticosteroid, BDP equivalent/μg (SD)	**632 (579)**	548 (559)	746 (611)	653 (581)	0.202
Long-acting bronchodilator use, %	**40.2**	34.4	48.2	41.7	0.442
Previous hospital admission or emergency attendance, no. per patient	**0.60 (1.57)**	1.04	0.26	0.20	0.037
Previous outpatient attendance, % attended	**15%**	22%	19%	6%	0.121
Severe asthma exacerbations (requiring oral corticosteroids) in past 12 mo, no. per patient	**1.25 (1.94)**	1.86 (0.32)	1.07 (0.32)	0.39 (0.18)	0.002

Clusters in secondary care

Variable	Secondary Care (*n = 187*)	Cluster 1 Early Onset, Atopic (*n = 74*)	Cluster 2 Obese, Noneosinophilic (*n = 23*)	Cluster 3 Early Symptom Predominant (*n = 22*)	Cluster 4 Inflammation Predominant (*n = 68*)	Significance (*P* Value)*
Sex†, % female	**65.8**			68.2	47.1	<0.001
Age, yr (SD)	**43.4 (15.9)**			35.5 (15.5)	50.6 (15.1)	<0.001
Age of onset†, yr (SD)	**20.3 (18.4)**			12.6 (15)	32.6 (19.1)	<0.001
Atopic status†, % positive	**73.8**			81.8	63.2	0.024
Body mass index†, kg/m^2 (SD)	**28.5 (6.5)**			23.6 (3.1)	27 (3.9)	<0.001
Peak flow variability‡, amp % mean	**32.2 (0.48)**			24.2 (0.65)	27.6 (0.36)	0.002
FEV_1 change with bronchodilator‡, %	**12.8 (0.41)**	24.5 (0.31)	9.3 (0.35)	4.5 (0.33)	9.8 (0.34)	<0.001
Post-bronchodilator FEV_1, % predicted (SD)	**82.1 (21.1)**	79.0 (21.9)	79.0 (18.5)	79.5 (26.1)	87.2 (18.5)	0.093
Sputum eosinophil count†‡, %	**2.9 (0.99)**	4.2 (0.76)	1.3 (1.01)	0.1 (0.9)	8.4 (0.64)	<0.001
F_{ENO}‡§, ppb	**43 (0.32)**	51.2 (0.36)	24.2 (0.27)	22.6 (0.30)	53.1 (0.32)	<0.001
Sputum neutrophil count, %‡	**46.7 (0.32)**	45.4 (0.39)	49.3 (0.22)	51.3 (0.23)	45.9 (0.29)	0.892
Modified JACS† (SD)	**2.02 (1.16)**	2.63 (0.93)	2.37 (1.09)	2.11 (1.11)	1.21 (0.95)	<0.001
Dose of inhaled corticosteroid, BDP equivalent/μg (SD)	**1,018 (539)**	1,168 (578)	1,045 (590)	809 (396)	914 (479)	0.008
Long-acting bronchodilator use, %	**93.0**	91.9	95.4	90.9	94.1	0.999

Resembled clusters from primary care – i.e., these are common across spectrum of severity

Objective measures of disease severity show more advanced disease

How should we treat asthma?

- Now we use 3rd dataset – 68 patients over 12 months
- Randomized control trial with two arms:
  - Standard clinical care ("clinical")
  - Regular monitoring of airway inflammation using induced sputum, to titrate steroid therapy to maintain normal eosinophil counts ("sputum")
- Original study found no difference in corticosteroid usage
  - But, this could have been explained by heterogeneity in treatment response!

[Haldar et al., *Am J Respir Crit Care Med*, 2008]

Patients in different clusters respond differently to treatment! (analysis using 3rd dataset from 12 month study)

Cluster (found using *baseline* data)	Outcomes	Treatment strategy: Clinical (*n* = *10*)	Treatment strategy: Sputum (*n* = *8*)	Significance
1: Obese female	Δ Inhaled corticosteroid dose*/μg per day (SEM)	−400 (328)	−462 (271)	0.89
	Severe exacerbation frequency over 12 mo (SEM)	1.40 (0.78)	1.50 (0.80)	0.93
	Number commenced on oral corticosteroids	2	1	0.59
		Clinical (*n* = *15*)	Sputum (*n* = *24*)	
2: Inflammation predominant	Δ Inhaled corticosteroid dose*/μg per day (SEM)	+753 (334)	+241 (233)	0.22
	Severe exacerbation frequency over 12 mo (SEM)	3.53 (1.18)	0.38 (0.13)	0.002
	Number commenced on oral corticosteroids	2	9	0.17
		Clinical (*n* = *7*)	Sputum (*n* = *4*)	
3: Early symptom predominant	Δ Inhaled corticosteroid dose*/μg per day (SEM)	+1,429 (429)	−400 (469)	0.022
	Severe exacerbation frequency over 12 mo (SEM)	5.43 (1.90)	2.50 (0.87)	0.198
	Number commenced on oral corticosteroids	6	0	Undefined

[Haldar et al., *Am J Respir Crit Care Med*, 2008]

Summary – two approaches

- **Supervised:**
  predict future disease status
- **Unsupervised:**
  which patients look similar / different? Do clusters have different outcomes?

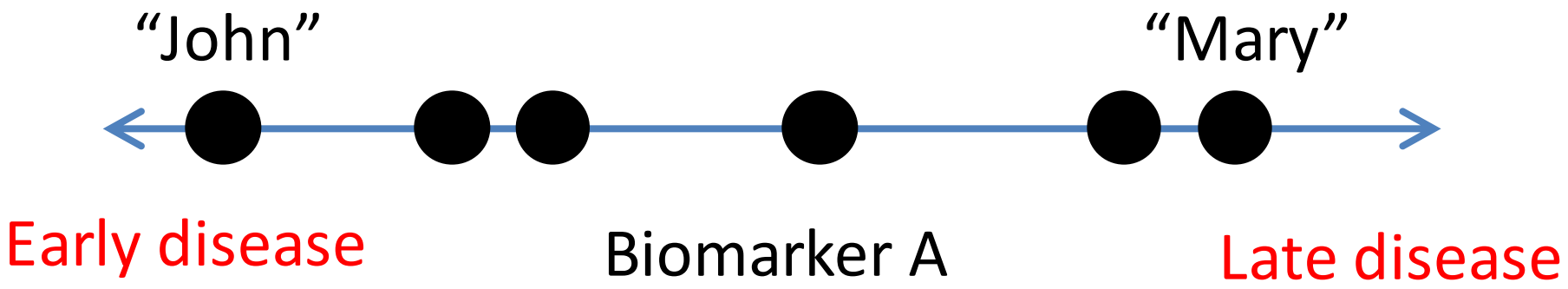
Limitations of what we've described thus far

- Can't differentiate between subtype and stage
  - Patients assumed to be aligned at baseline
- Only make use of one time point per patient
- Assumes single factor (cluster) explains all variation

How can we discover stages?

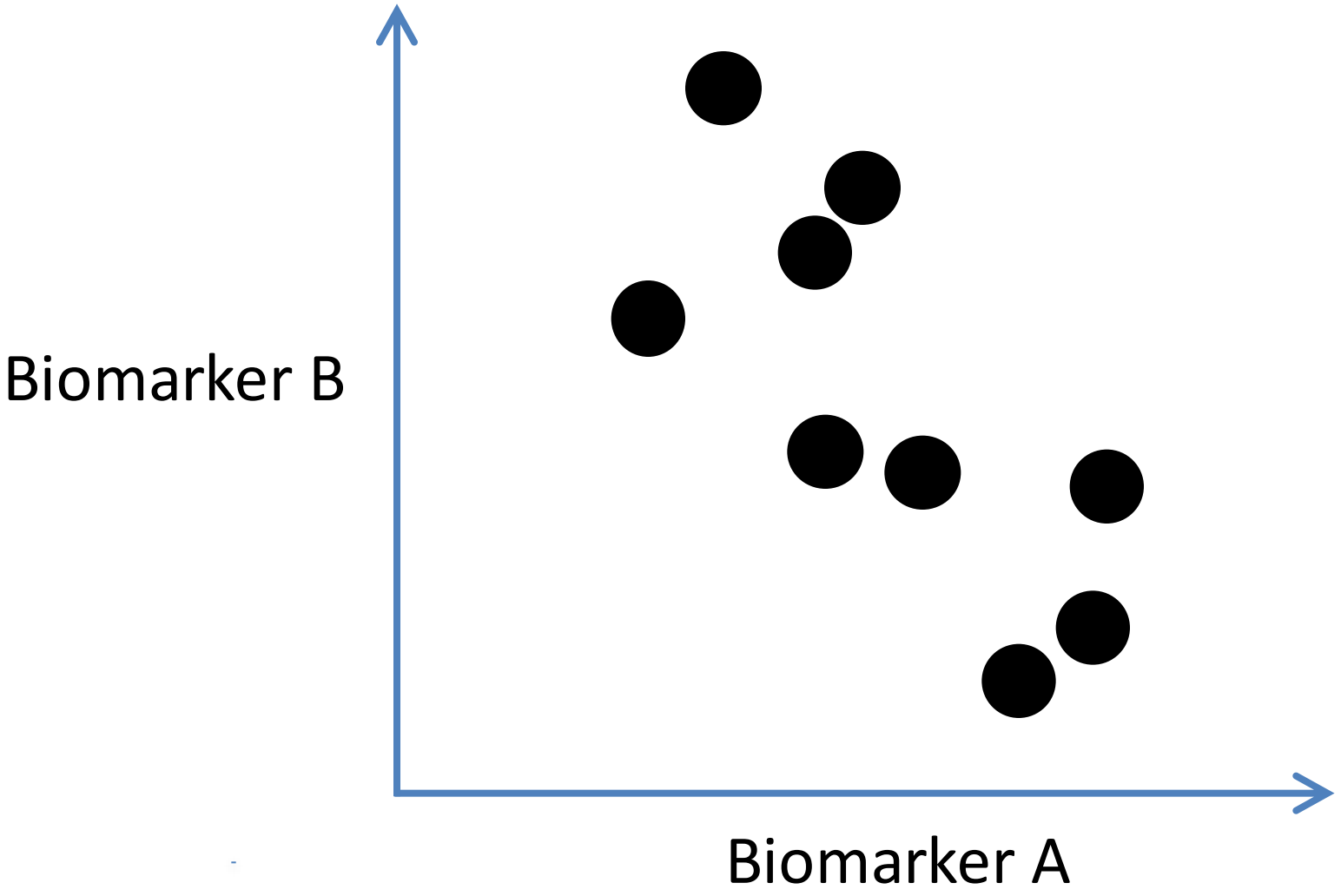
1. **Intuition on staging from cross-sectional data**
2. Staging with pseudo-time methods
3. Staging with probabilistic models: missing data & multiple time points
   - Case study: chronic obstructive pulmonary disease
   - Case study: Type 2 diabetes

In 1-D, might assume that low values correspond to an early disease stage (or vice-versa)

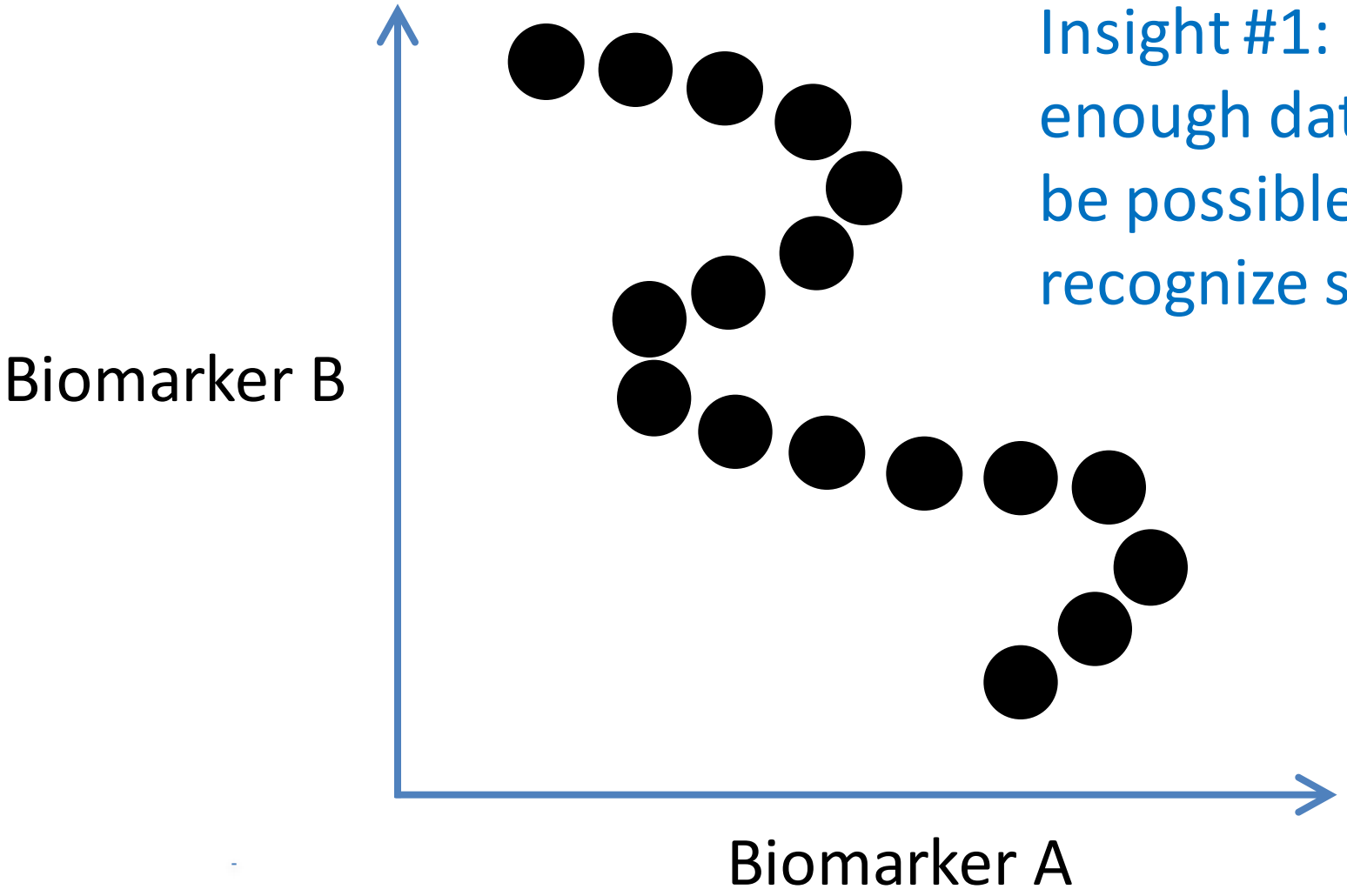


Assume samples were all taken today

What about in higher dimensions?



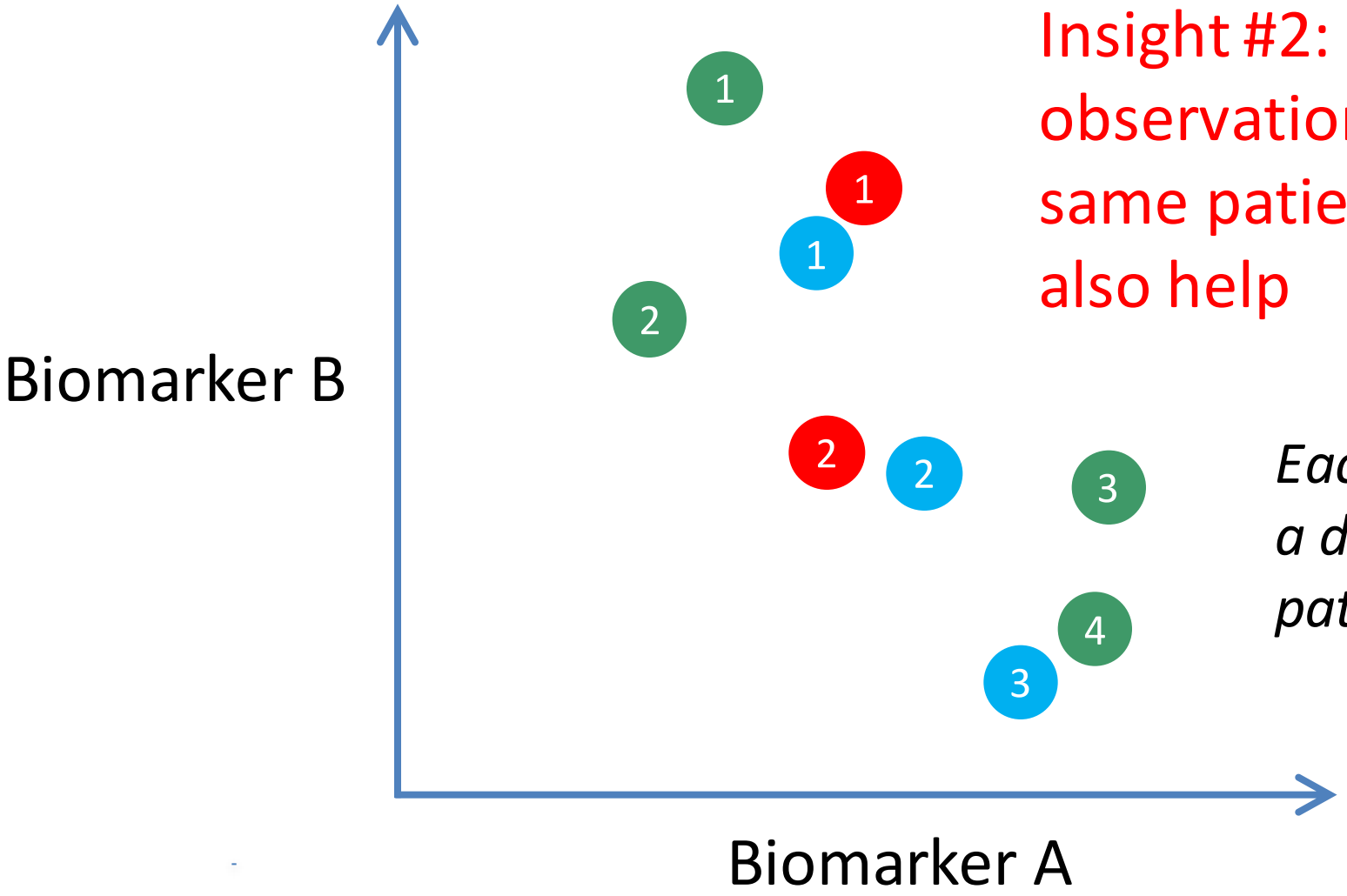
What about in higher dimensions?



Insight #1: with enough data, may be possible to recognize structure

[Bendall et al., Cell 2014 (human B cell development)]

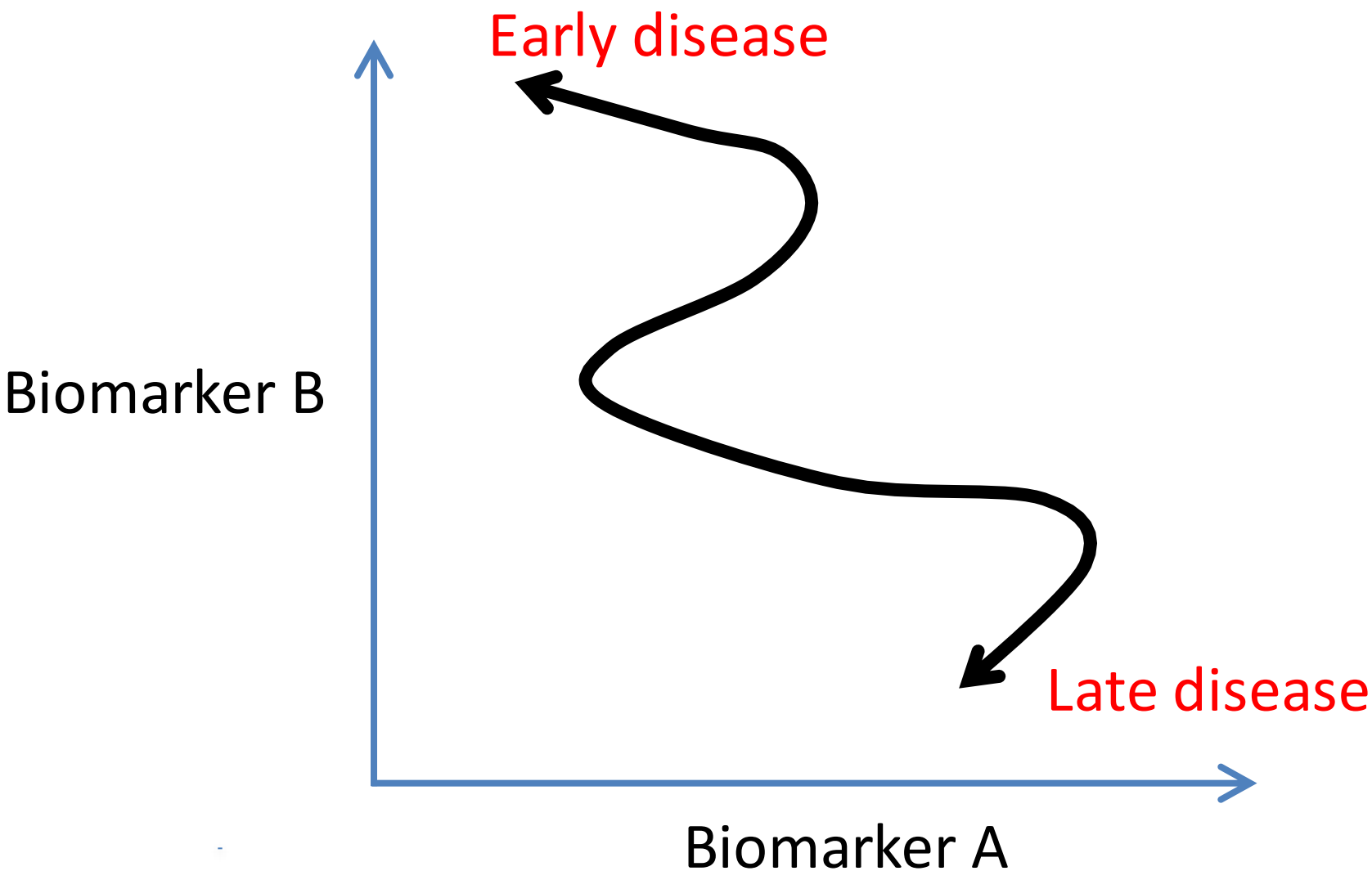
What about in higher dimensions?



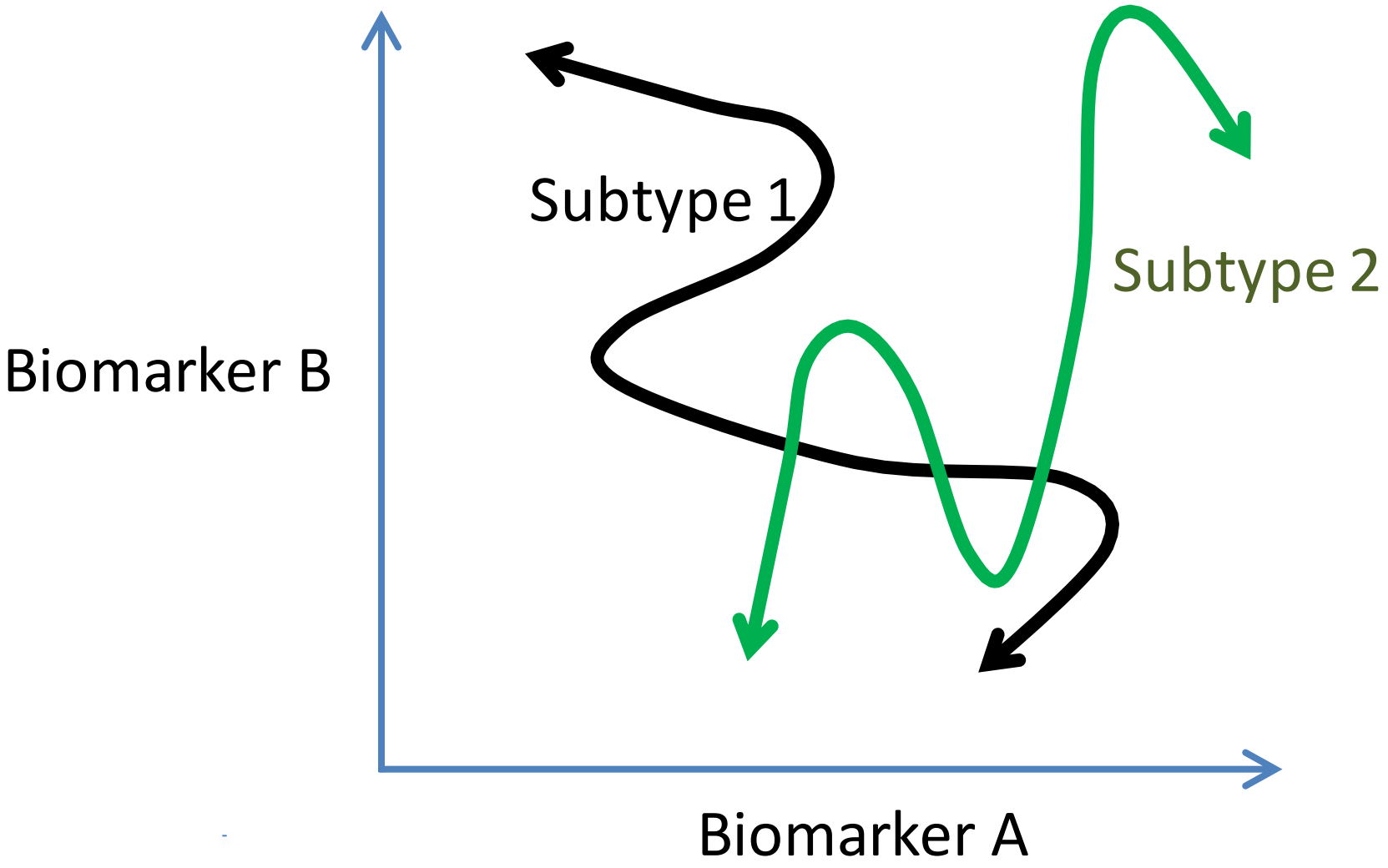
Insight #2: sequential observations from same patient can also help

Each color is a different patient

What about in higher dimensions?



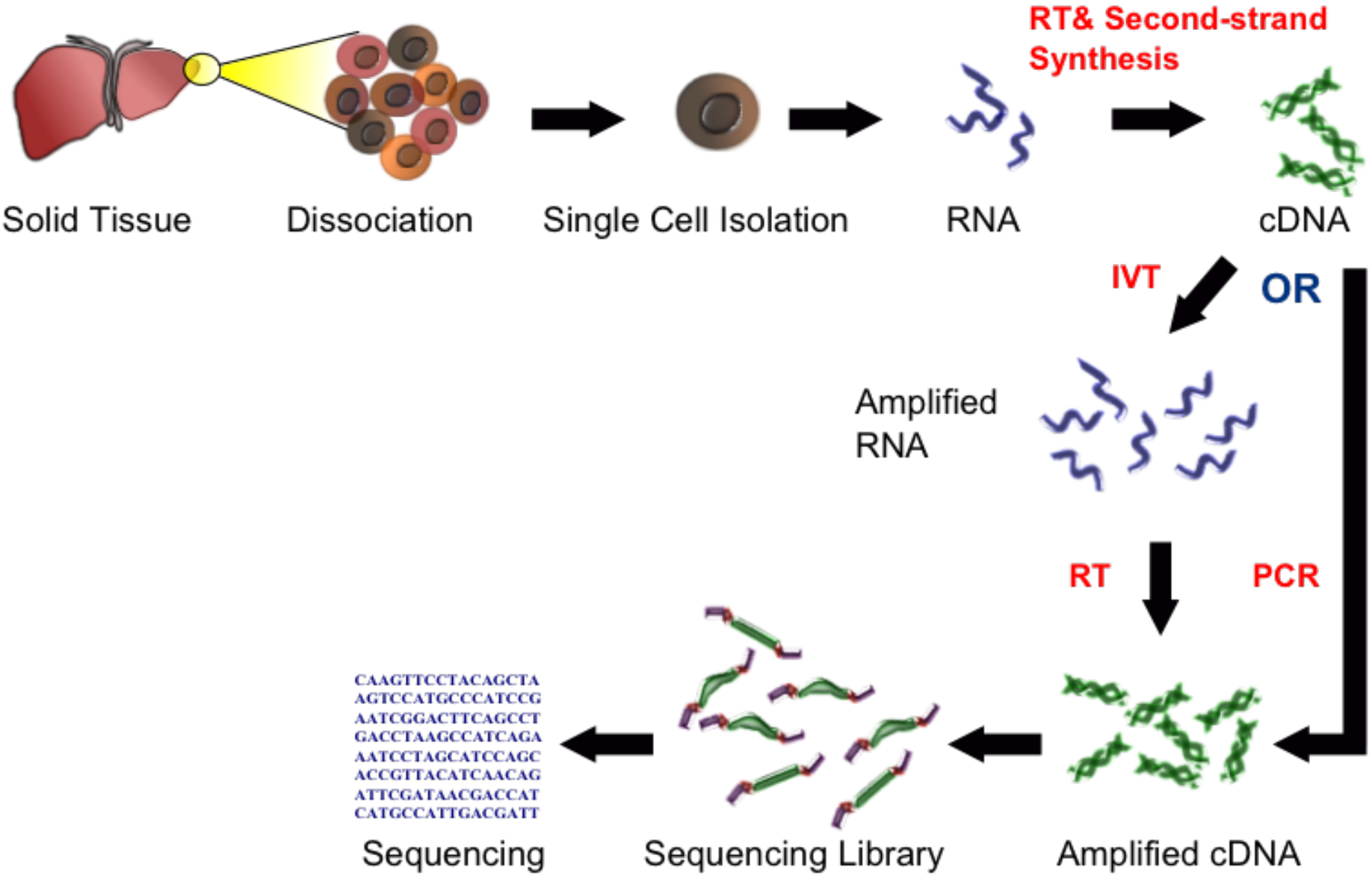
May also seek to discover disease subtypes



How can we discover stages?

1. Intuition on staging from cross-sectional data
2. **Staging with pseudo-time methods**
3. Staging with probabilistic models: missing data & multiple time points
   - Case study: chronic obstructive pulmonary disease
   - Case study: Type 2 diabetes

Single-cell sequencing



[Figure source: https://en.wikipedia.org/wiki/Single_cell_sequencing]

Inferring original trajectory from single-cell data

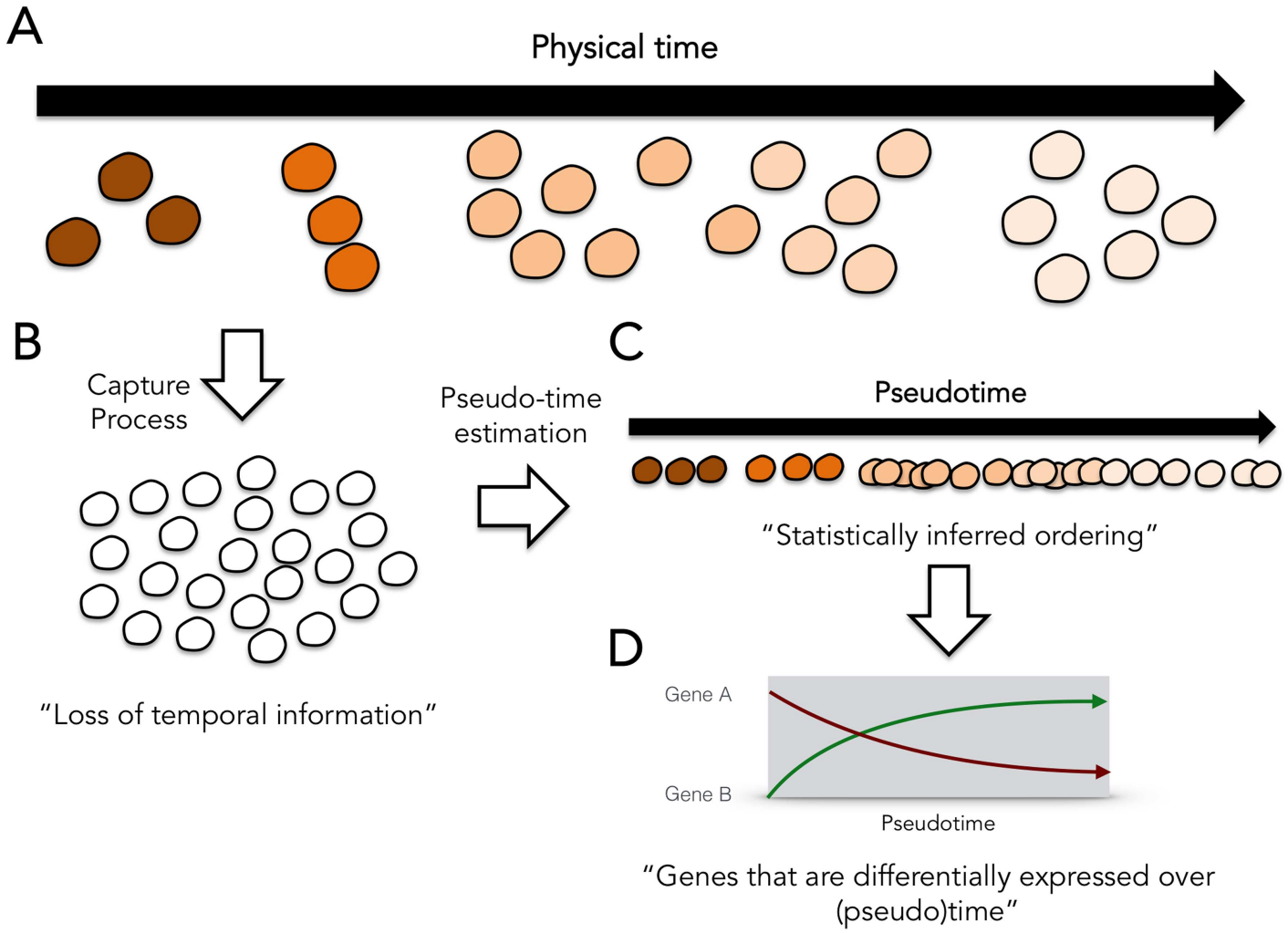
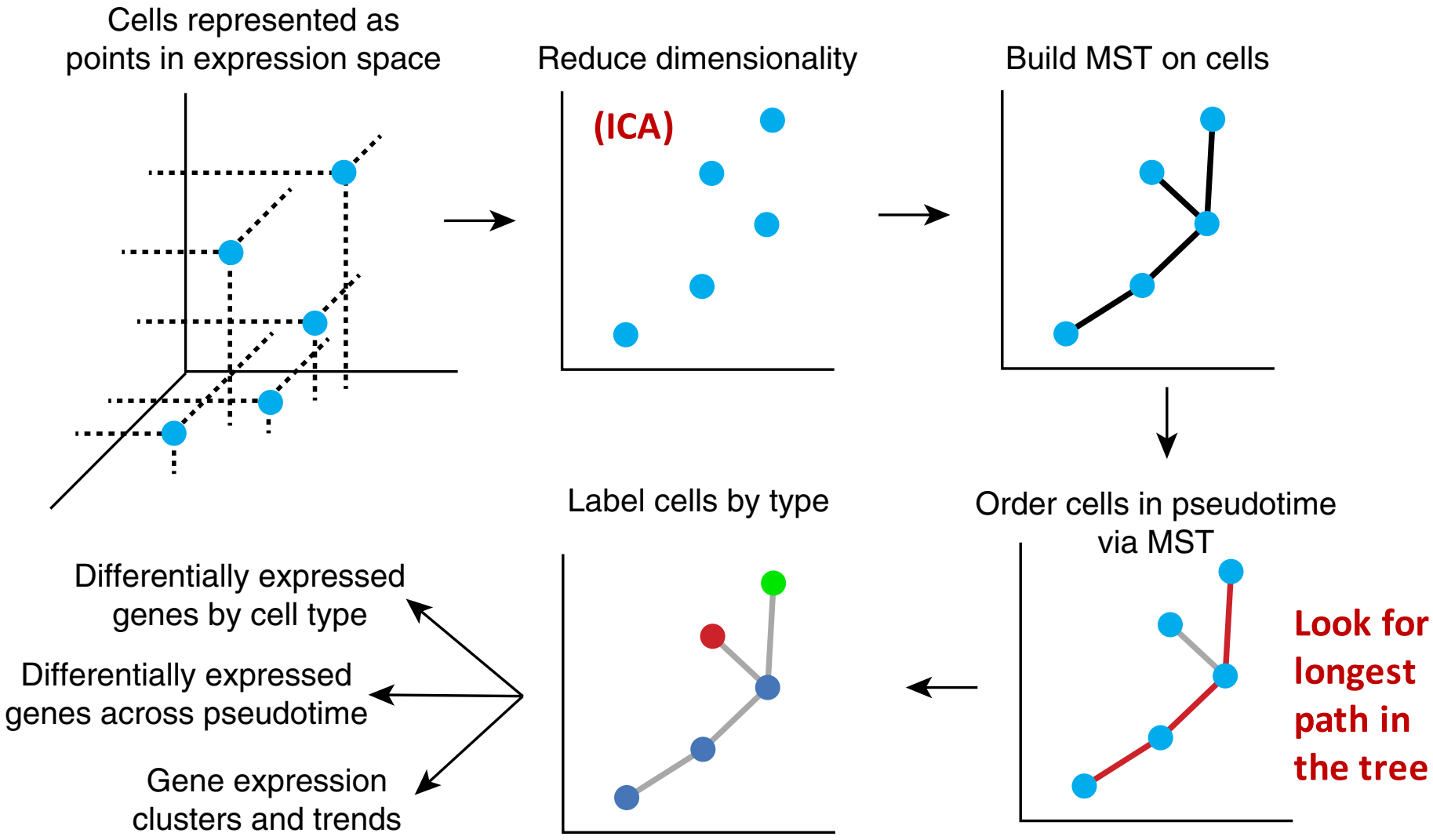


Fig 1. The single cell pseudotime estimation problem. (A) Single cells at different stages of a temporal process. (B) The temporal labelling information is lost during single cell capture. (C) Statistical pseudotime estimation algorithms attempt to reconstruct the relative temporal ordering of the cells but cannot fully reproduce physical time. (D) The pseudotime estimates can be used to identify genes that are differentially expressed over (pseudo)time.

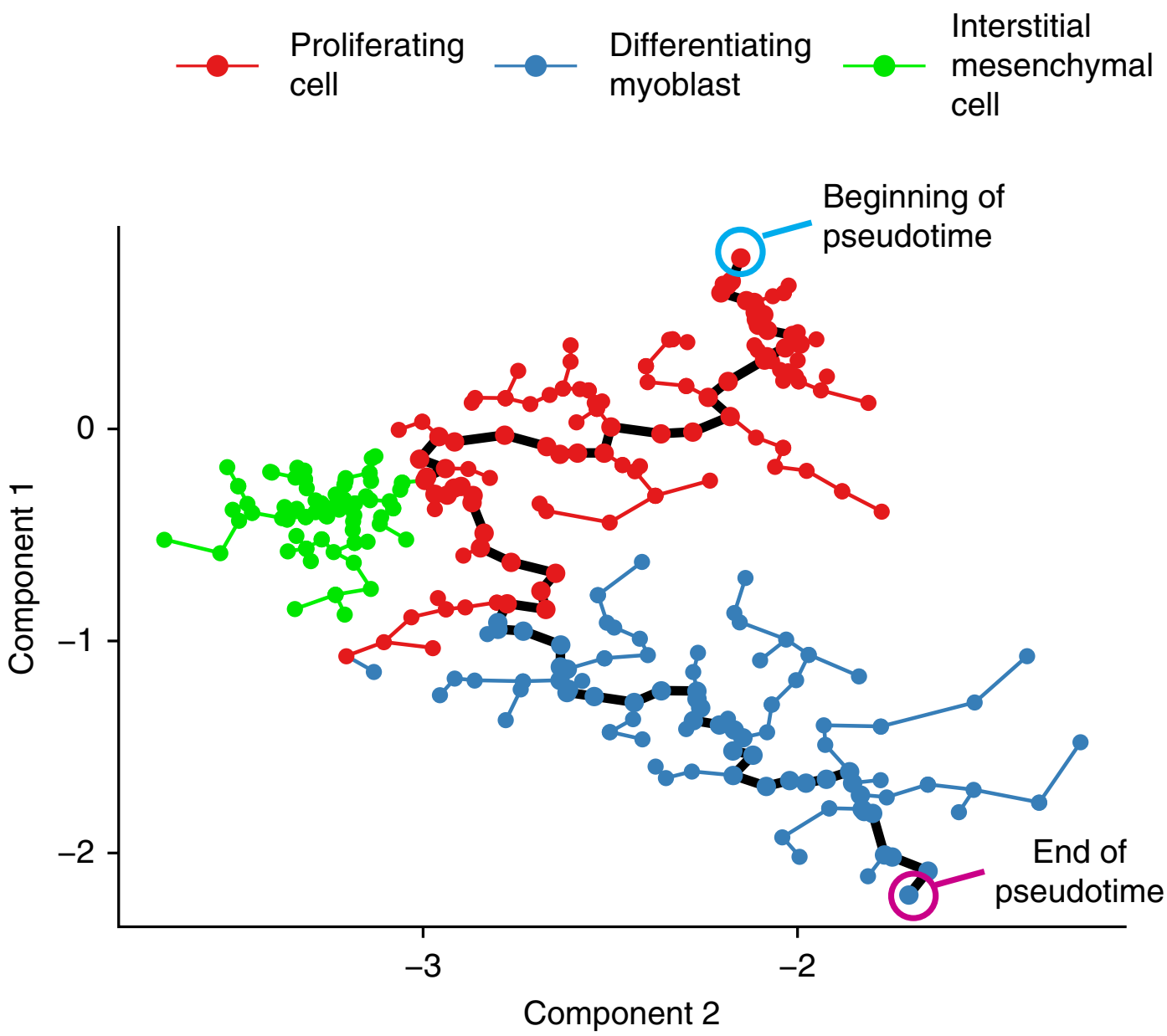
[Figure from: Campbell & Yau, PLOS Computational Biology, 2016]

MST-based approach (Monocle)



[Magwene et al., *Bioinformatics*, 2003; Trapnell et al., *Nature Biotechnology*, 2014]

MST-based approach (Monocle)



[Trapnell et al., *Nature Biotechnology*, 2014]

How can we discover stages?

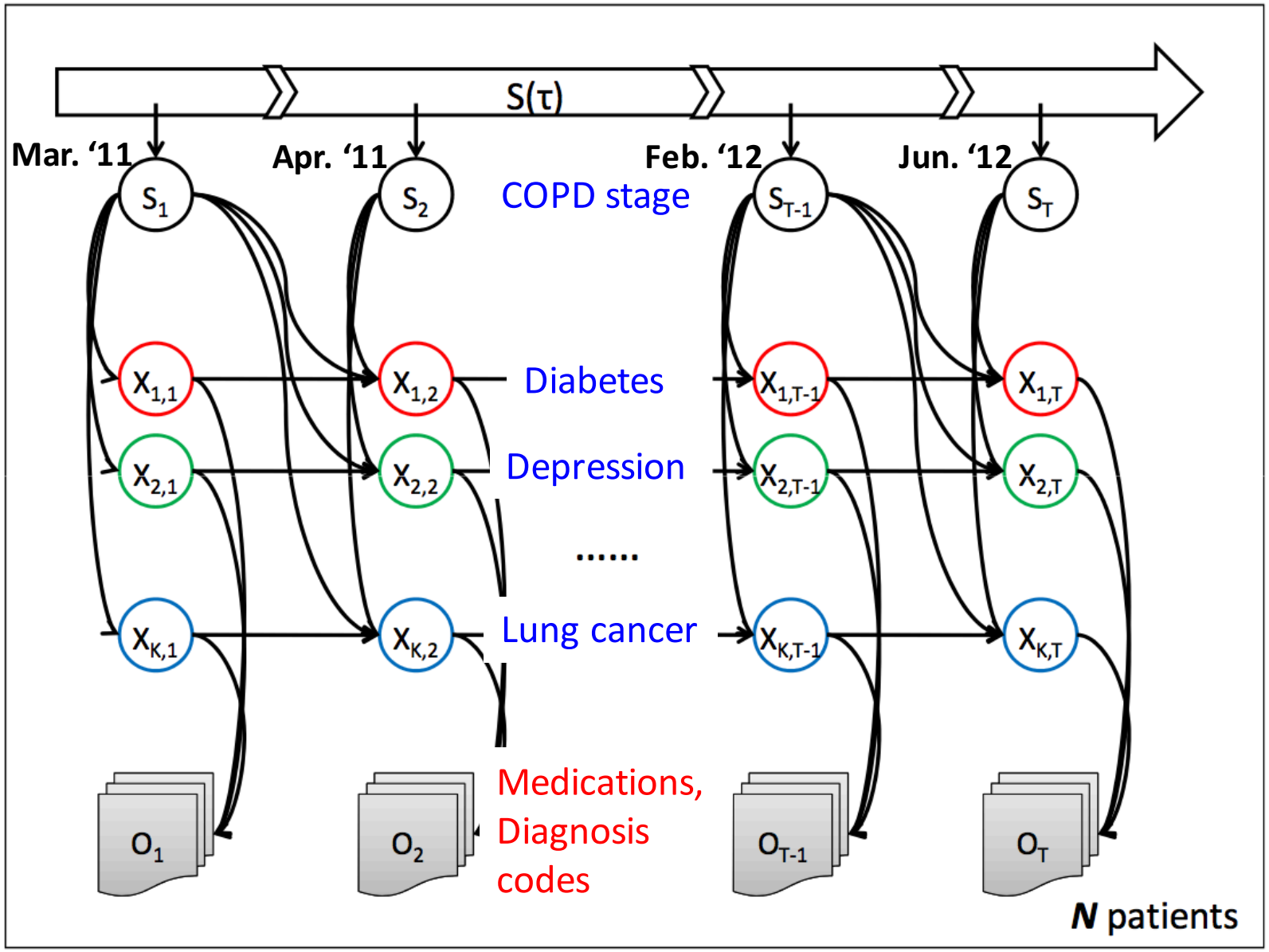
1. Intuition on staging from cross-sectional data
2. Staging with pseudo-time methods
3. **Staging with probabilistic models: missing data & multiple time points**
   - **Case study: chronic obstructive pulmonary disease**
   - Case study: Type 2 diabetes

Can we learn 10-year progression of COPD from EHR data?

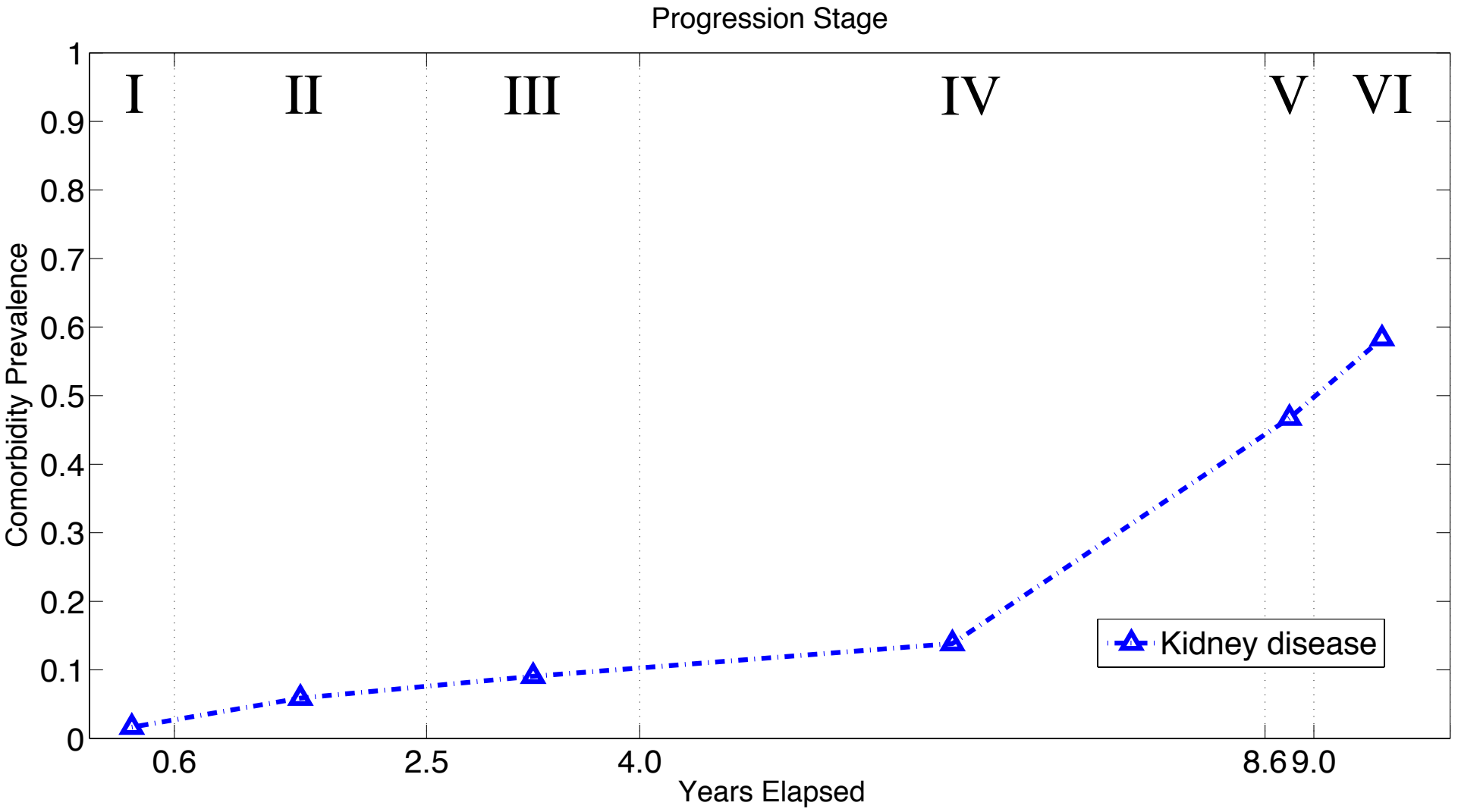
- Only 2-4 years of data for each patient
- High-dimensional, with **lots of missing data**
- No ground truth – not even spirometry

[Xiang, Sontag, Wang, "Unsupervised learning of Disease Progression Models", KDD 2014]

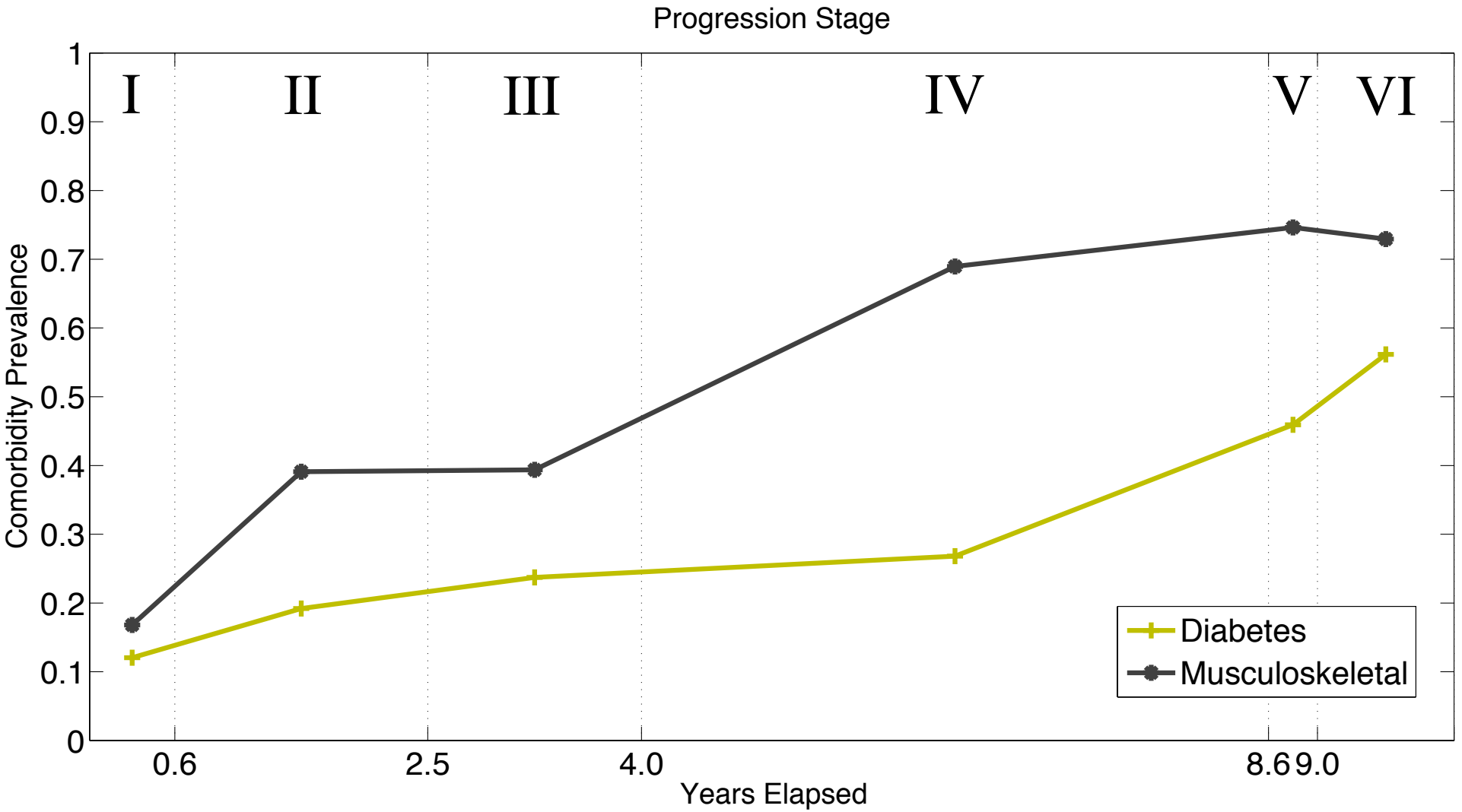
Probabilistic model of disease progression



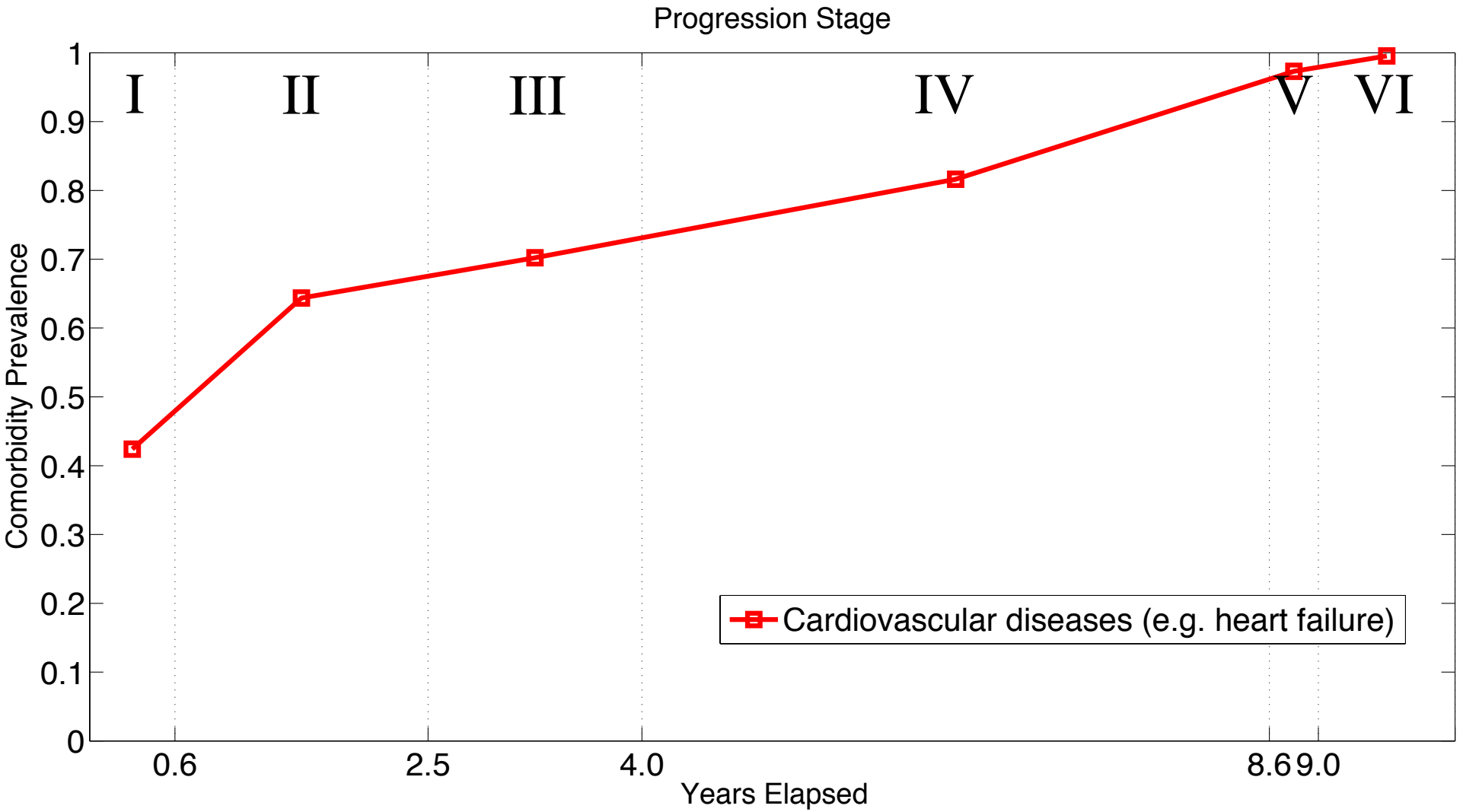
Inferred prevalence of comorbidities across stages (Kidney disease)



Inferred prevalence of comorbidities across stages (Diabetes & Musculoskeletal disorders)



Inferred prevalence of comorbidities across stages (Cardiovascular disease)



Inf

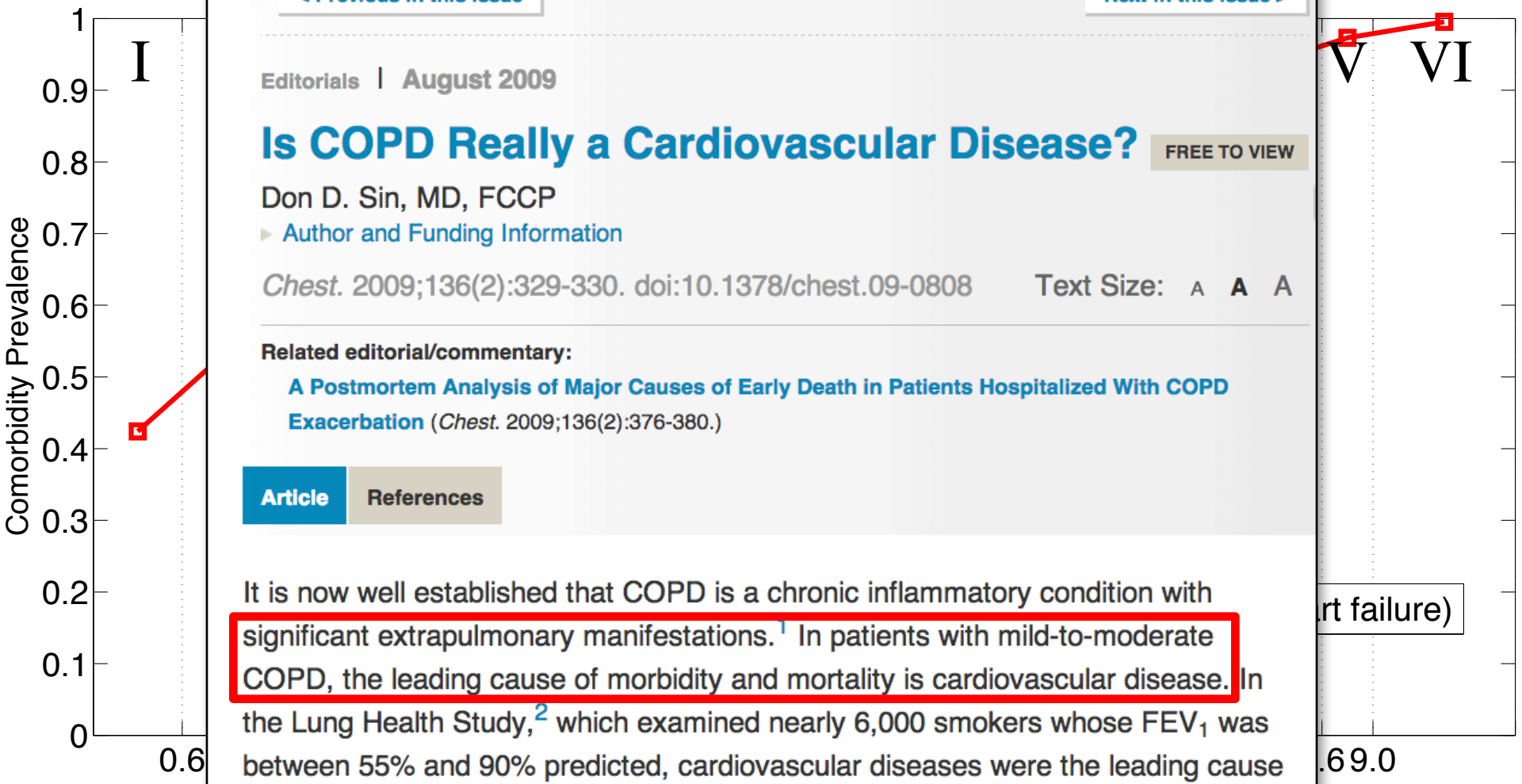
CHEST JOURNAL
Official Publication of the American College o

oss

August 2009, Vol 136, No. 2

< Previous in this issue

Next in this issue >



Editorials | August 2009

Is COPD Really a Cardiovascular Disease?

FREE TO VIEW

Don D. Sin, MD, FCCP

Author and Funding Information

Chest. 2009;136(2):329-330. doi:10.1378/chest.09-0808

Text Size: A A A

Related editorial/commentary:

A Postmortem Analysis of Major Causes of Early Death in Patients Hospitalized With COPD Exacerbation (*Chest*. 2009;136(2):376-380.)

Article | References

It is now well established that COPD is a chronic inflammatory condition with significant extrapulmonary manifestations.[1] In patients with mild-to-moderate COPD, the leading cause of morbidity and mortality is cardiovascular disease. In the Lung Health Study,[2] which examined nearly 6,000 smokers whose FEV_1 was between 55% and 90% predicted, cardiovascular diseases were the leading cause of hospitalization, accounting for nearly 50% of all hospital admissions, and the second leading cause of mortality, accounting for a quarter of all deaths.

How can we discover stages?

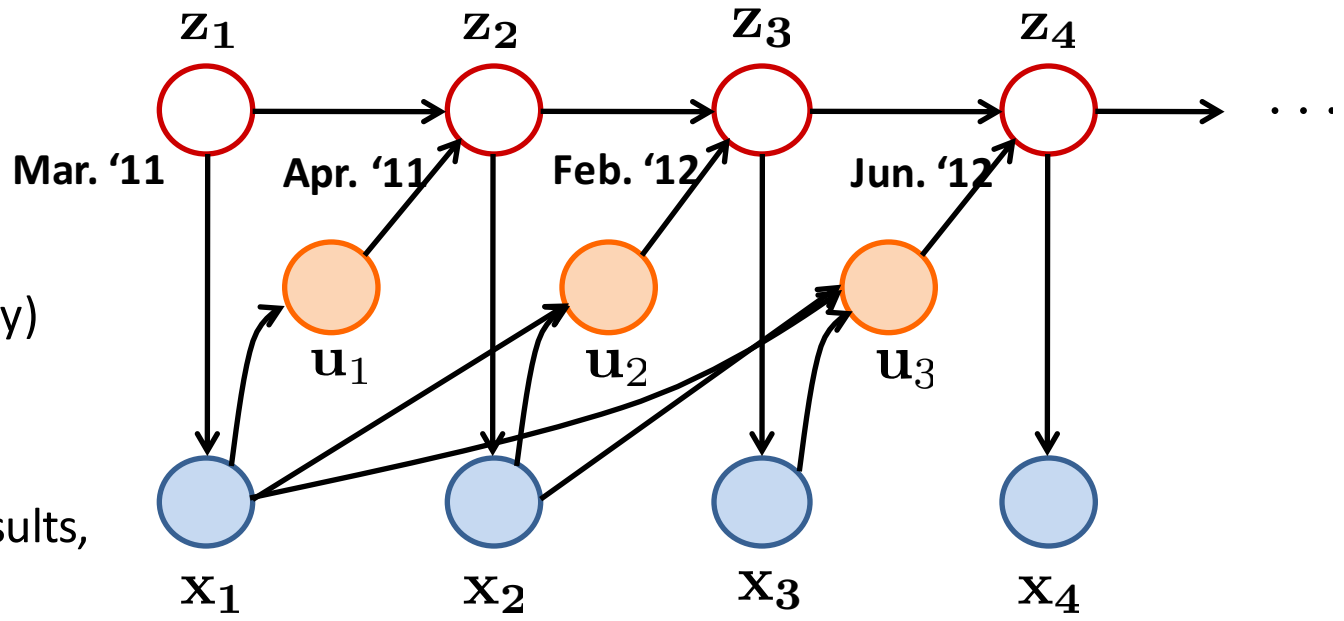
1. Intuition on staging from cross-sectional data
2. Staging with pseudo-time methods
3. **Staging with probabilistic models: missing data & multiple time points**
   - Case study: chronic obstructive pulmonary disease
   - **Case study: Type 2 diabetes**

Deep Markov models (DMMs)

Patient state $\mathbf{z}_t \in \mathbb{R}^{100}$

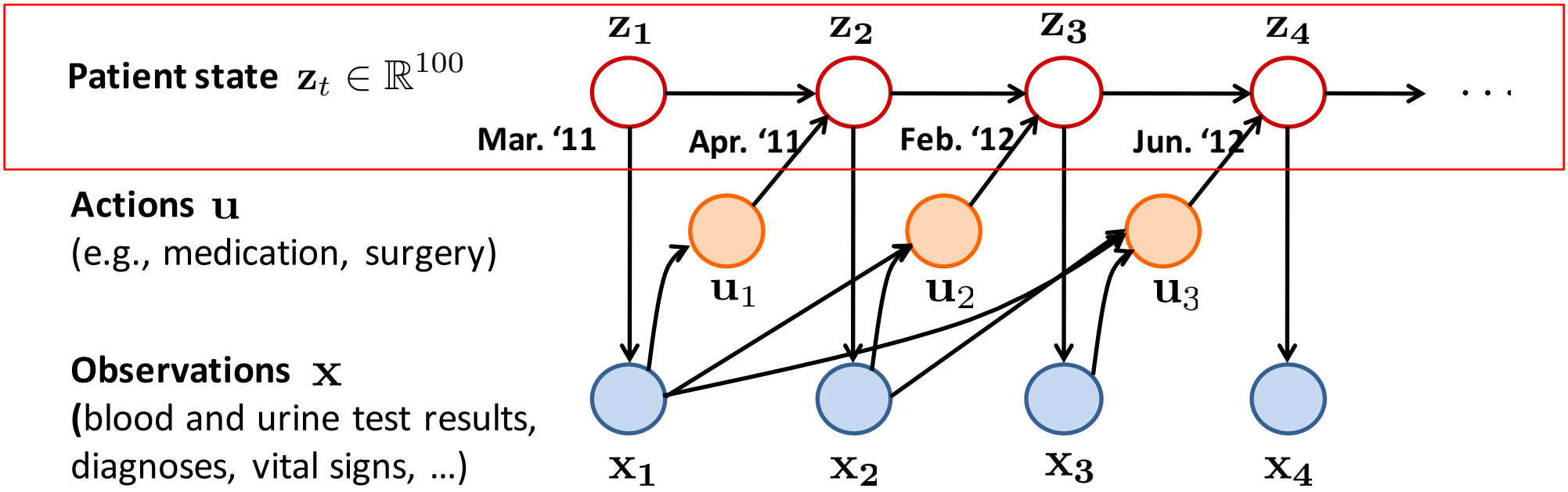
Actions $\mathbf{u}$
(e.g., medication, surgery)

Observations $\mathbf{x}$
(blood and urine test results, diagnoses, vital signs, …)



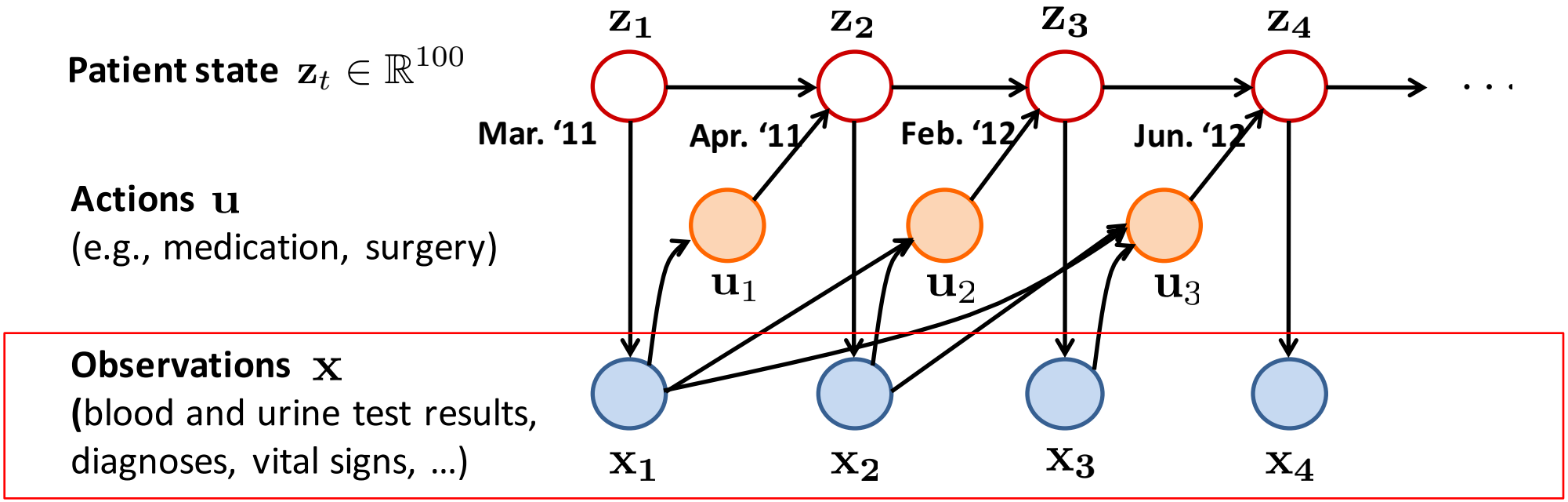
[Krishnan, Shalit, Sontag, AAAI '17]

Deep Markov models (DMMs)



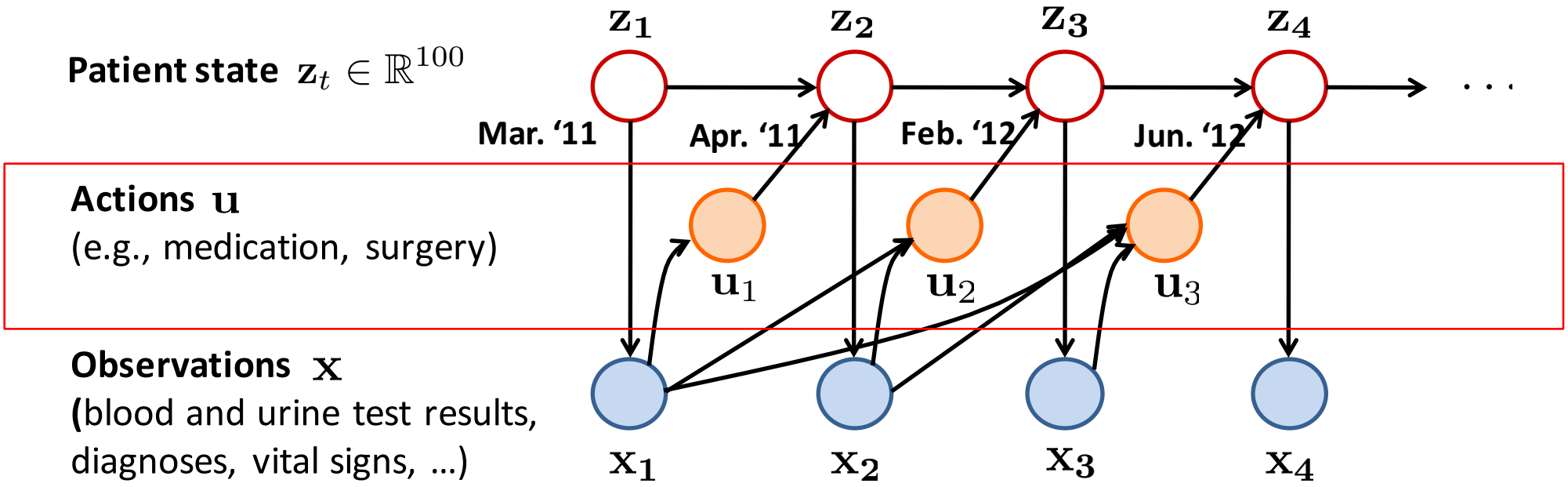
[Krishnan, Shalit, Sontag, AAAI '17]

Deep Markov models (DMMs)



[Krishnan, Shalit, Sontag, AAAI '17]

Deep Markov models (DMMs)



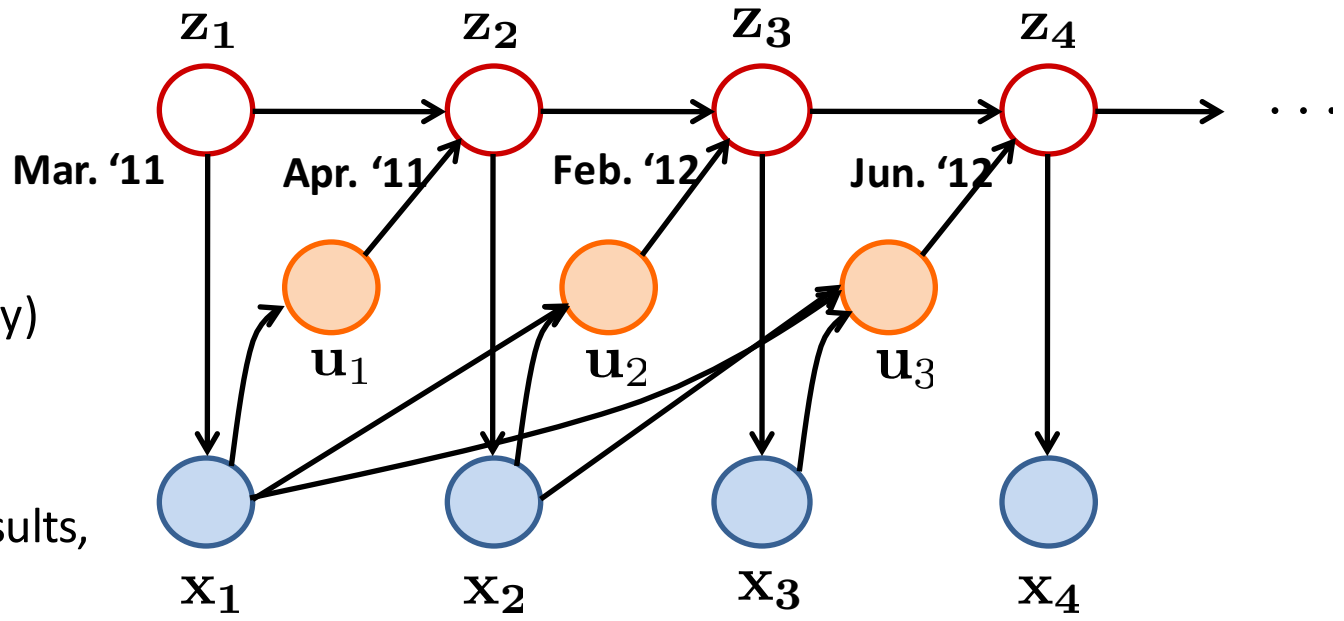
[Krishnan, Shalit, Sontag, AAAI '17]

Deep Markov models (DMMs)

Patient state $\mathbf{z}_t \in \mathbb{R}^{100}$

Actions $\mathbf{u}$
(e.g., medication, surgery)

Observations $\mathbf{x}$
(blood and urine test results, diagnoses, vital signs, …)



- Provides an in-silico model for assessing effect of interventions (actions), by forward sampling in model

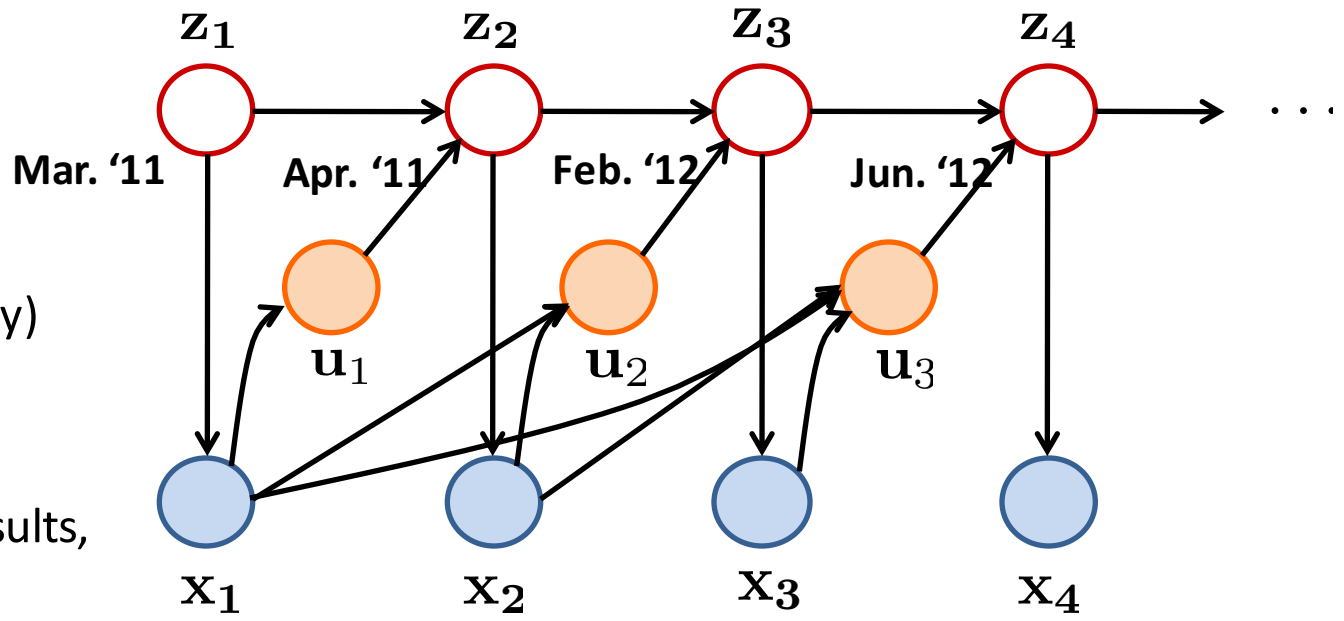
[Krishnan, Shalit, Sontag, AAAI '17]

Deep Markov models (DMMs)

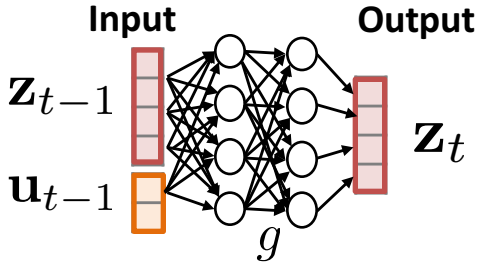
Patient state $\mathbf{z}_t \in \mathbb{R}^{100}$

Actions $\mathbf{u}$
(e.g., medication, surgery)

Observations $\mathbf{x}$
(blood and urine test results, diagnoses, vital signs, …)



- Provides an in-silico model for assessing effect of interventions (actions), by forward sampling in model
- Transition & emission distributions given by deep neural networks:

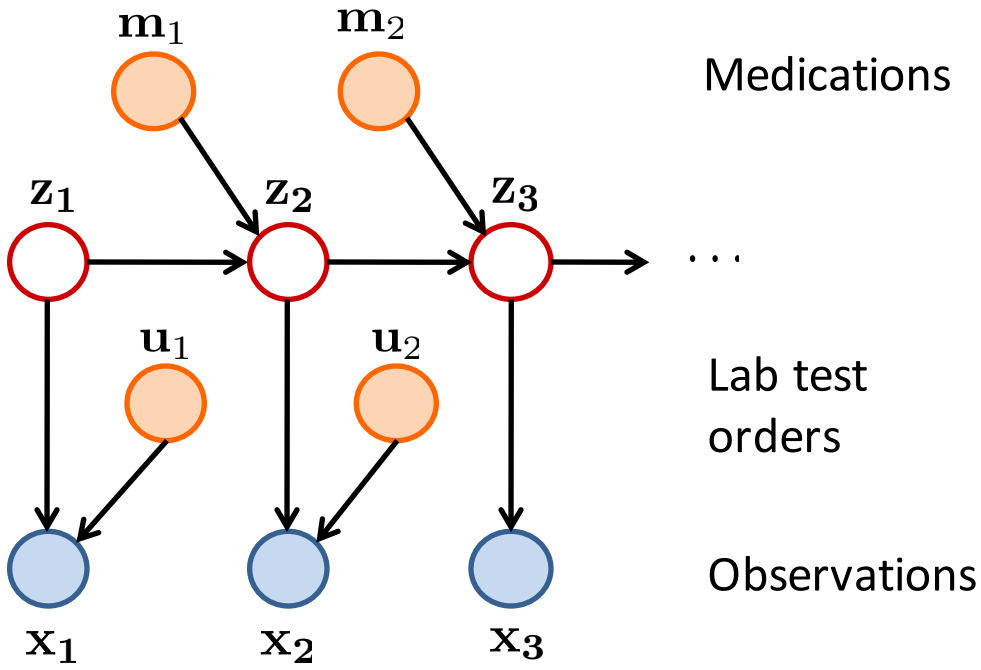


$$\mathbf{z}_t \sim \mathcal{N}(g(\mathbf{z}_{t-1}, \mathbf{u}_{t-1}), s(\mathbf{z}_{t-1}, \mathbf{u}_{t-1}))$$

[Krishnan, Shalit, Sontag, AAAI '17]

Learning the effect of diabetic treatments

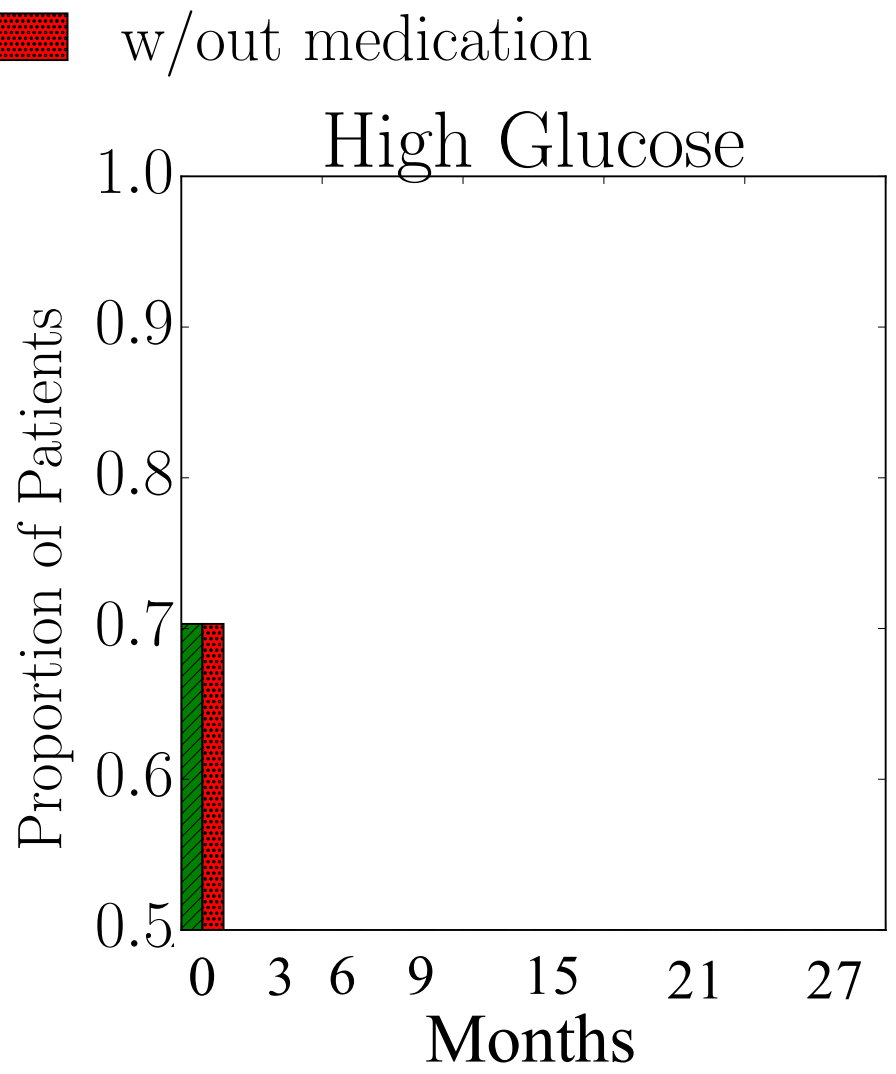
- Long-term: which diabetes medications work best for whom?
- **Actions:** 9 diabetic drugs including Metformin and Insulin (**m**), lab test orders (**u**)



- *Here we just do a sanity check. 8000 diabetic & pre-diabetic patients, 4 years of data.*

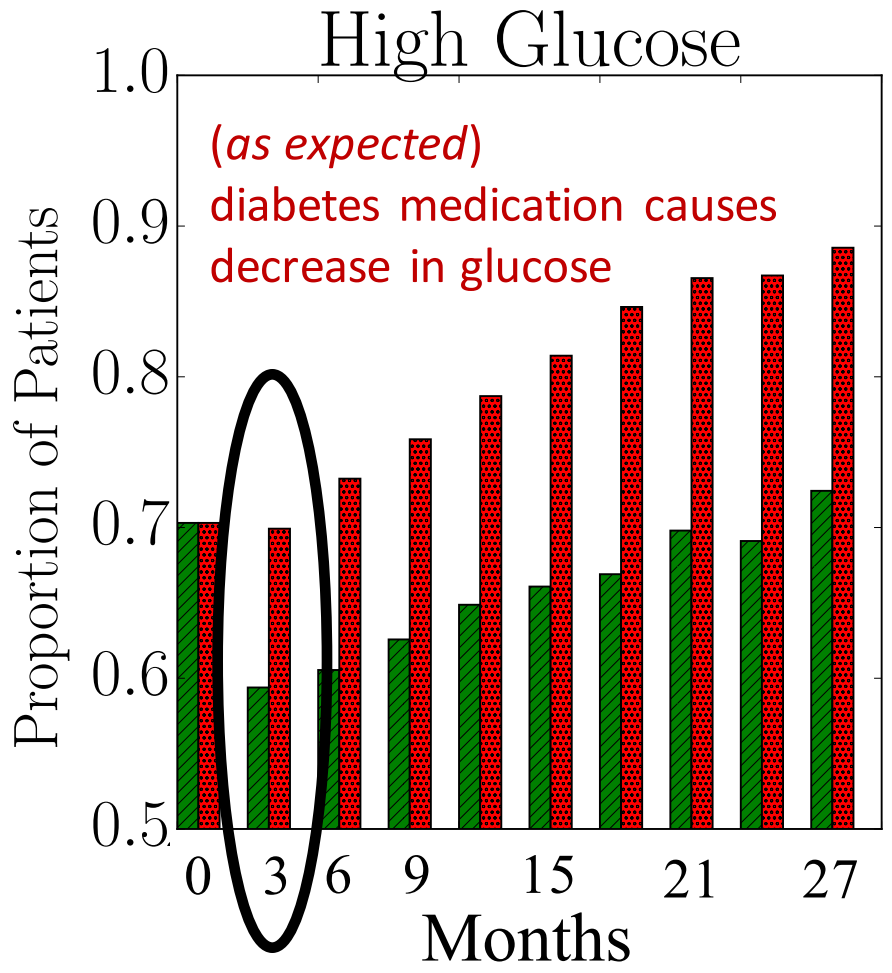
Effect of diabetes treatments on glucose

1. Align patients by when they were first prescribed Metformin
2. Sample future patient data ***using the medications they truly received***
3. Sample future patient data ***as if they never received medication***



Effect of diabetes treatments on glucose

1. Align patients by when they were first prescribed Metformin
2. Sample future patient data ***using the medications they truly received***
3. Sample future patient data ***as if they never received medication***



Summary

- Incredible potential for clinical data to be used for:
  - Population-level understanding of disease progression
  - Discovering new disease subtypes
  - Predicting future outcomes such as survival time and complications
  - Personalizing therapy by identifying who will respond best to treatment
- Key advance is to show how to do these from high-dimensional, noisy, incomplete patient trajectories