Life After Death:

Techniques for the rapid prognostication of post-anoxic coma patients.

Mohammad M. Ghassemi
Massachusetts Institute of Technology

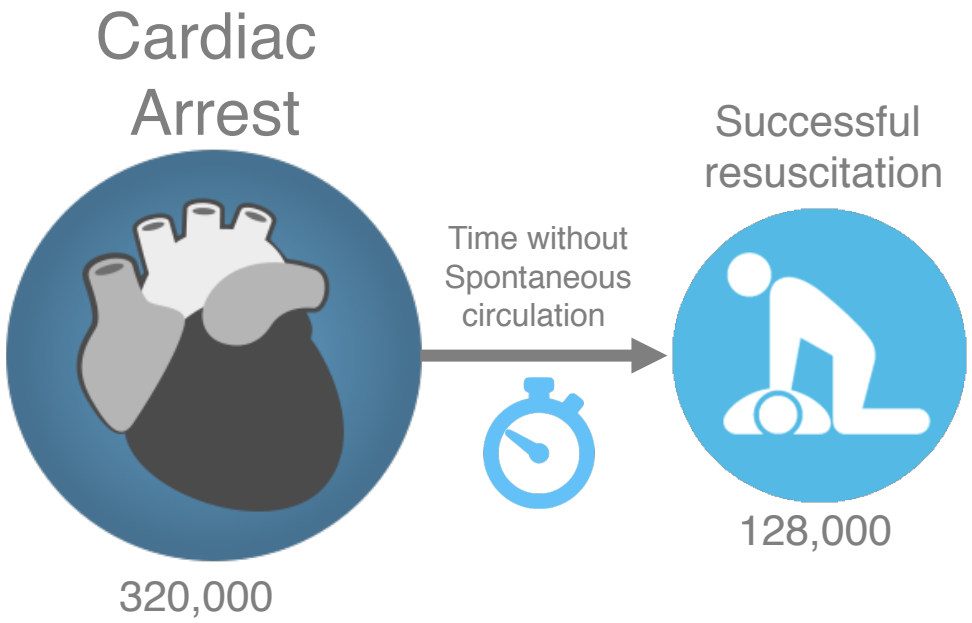
Cardiac arrest affects 320K people in the US annually

Cardiac Arrest

320,000

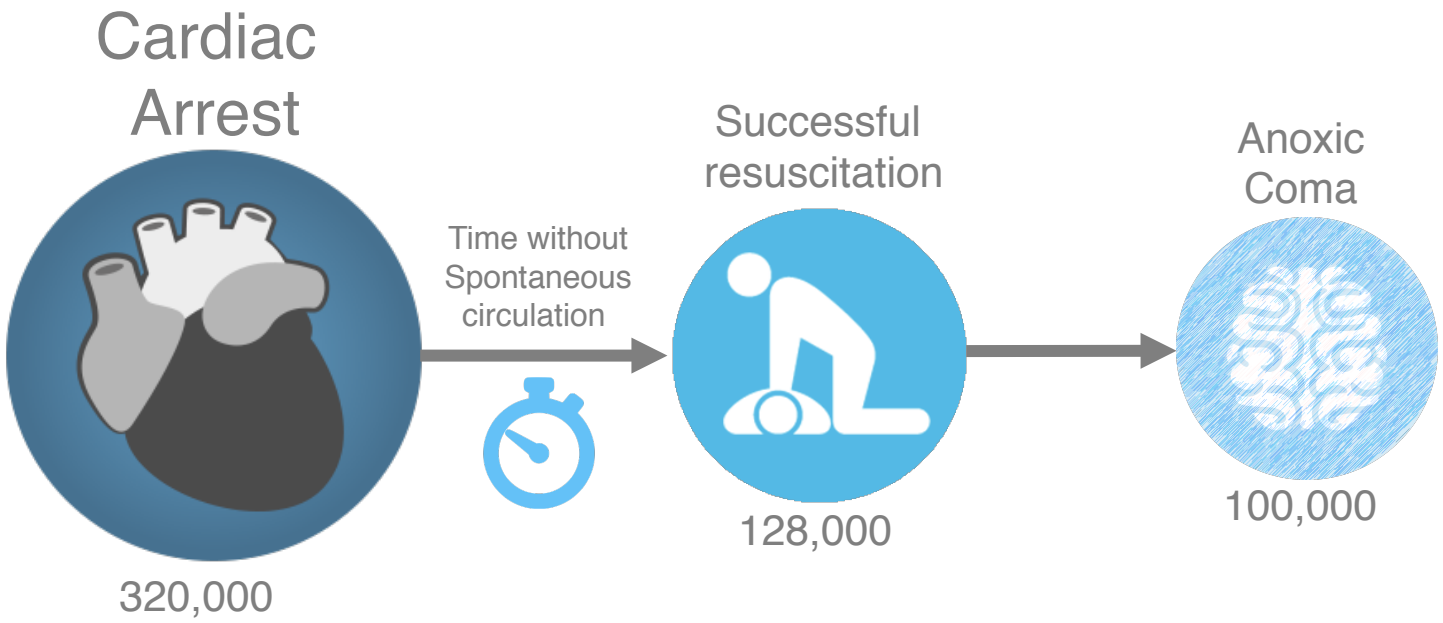
[4] G. Citerio, J. Bakker, M. Bassetti, D. Benoit, M. Cecconi, J. Curtis, G. Doig, M. Herridge, S. Jaber, and M. Joannidis, "Year in review in intensive care medicine 2014: I. cardiac dysfunction and cardiac arrest, ultrasound, neurocritical care, icu-acquired weakness, nutrition, acute kidney injury, and miscellaneous," *Intensive Care Medicine*, vol. 41, no. 2, pp. 179–191, 2015.

128K patients are successfully resuscitated



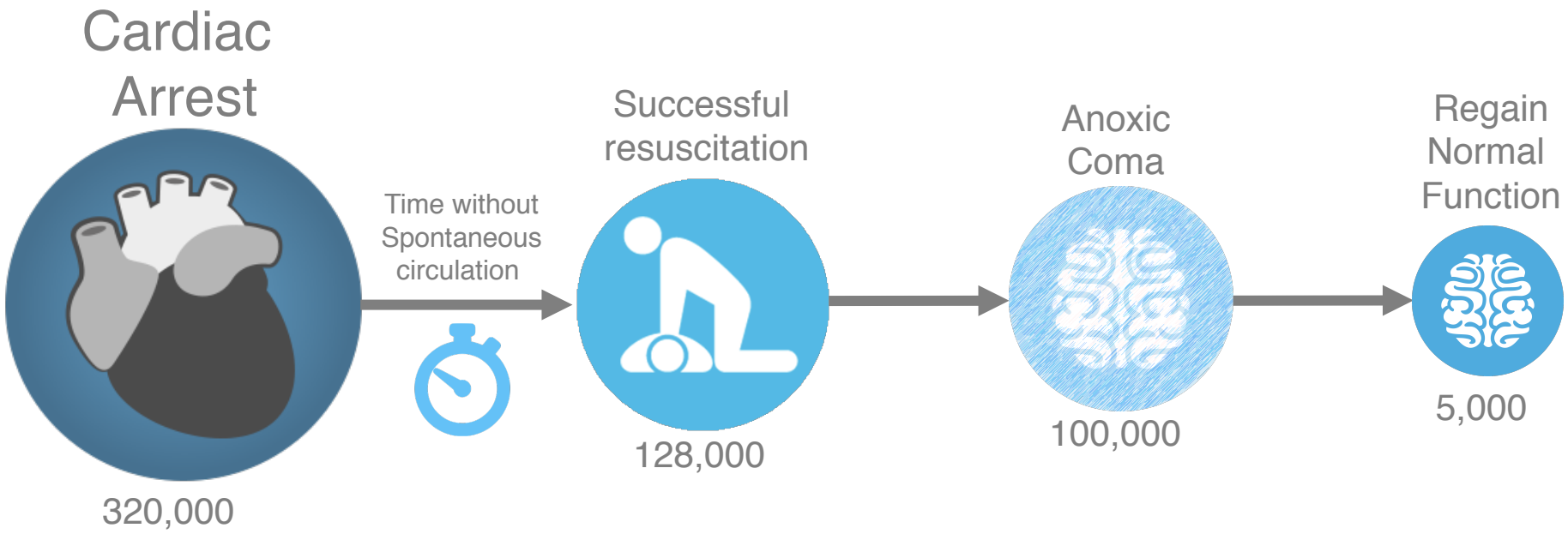
[5] K. Kern, "Optimal treatment of patients surviving out-of-hospital cardiac arrest," *JACC: Cardiovascular Interventions*, vol. 5, no. 6, pp. 597–605, 2012.

100K enter an indefinite, anoxia-induced coma



[6] R. Deng, M. Koenig, L. Young, and X. Jia, "Early quantitative gamma-band eeg marker is associated with outcomes after cardiac arrest and targeted temperature management," *Neurocritical Care*, vol. 23, no. 2, pp. 262–273, 2015.

10K will survive, but only 5K will regain normal function



[7] P. R. L. I. for Health Policy, "Icu outcomes (mortality and length of stay) methods, data collection tool and data." http://healthpolicy.ucsf.edu/content/icu-outcomes, 2016.

Outcome metric

Cerebral Performance Category	Conscious (Y/N)	Cerebral Disability	Consequences
1	Y	Mild	May resume independent activities with minimal complications
2	Y	Moderate	May resume independent activities while requiring some assistance
3	Y	Severe	Will require assistance to perform activities and may involve paralysis, or dementia.
4	N	Vegetative	Requires assistance to survive and will be minimally aware/responsive
5	N	Brain-Death	Requires assistance to survive and will be totally unaware/response

Outcome metric

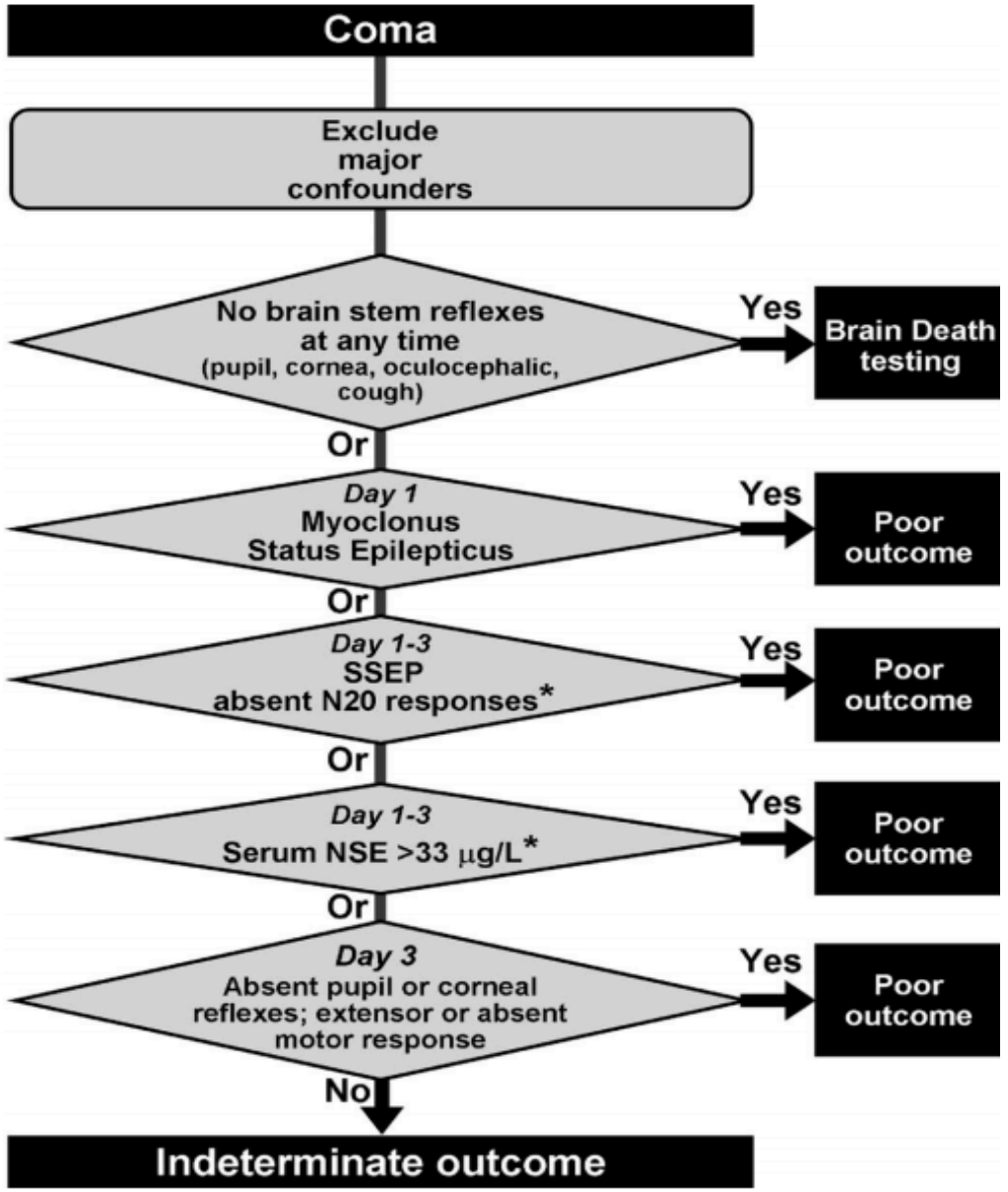
Cerebral Performance Category	Conscious (Y/N)	Cerebral Disability	Consequences	
1	Y	Mild	May resume independent activities with minimal complications	Good
2	Y	Moderate	May resume independent activities while requiring some assistance	
3	Y	Severe	Will require assistance to perform activities and may involve paralysis, or dementia.	Bad
4	N	Vegetative	Requires assistance to survive and will be minimally aware/responsive	
5	N	Brain-Death	Requires assistance to survive and will be totally unaware/response	

Goals of prognostication

- **Primary:** Prevent premature withdrawal care
- **Secondary:** Prevent unnecessary care
  (up to $20,000 per day in ICU)

Current Prognostic Guidelines

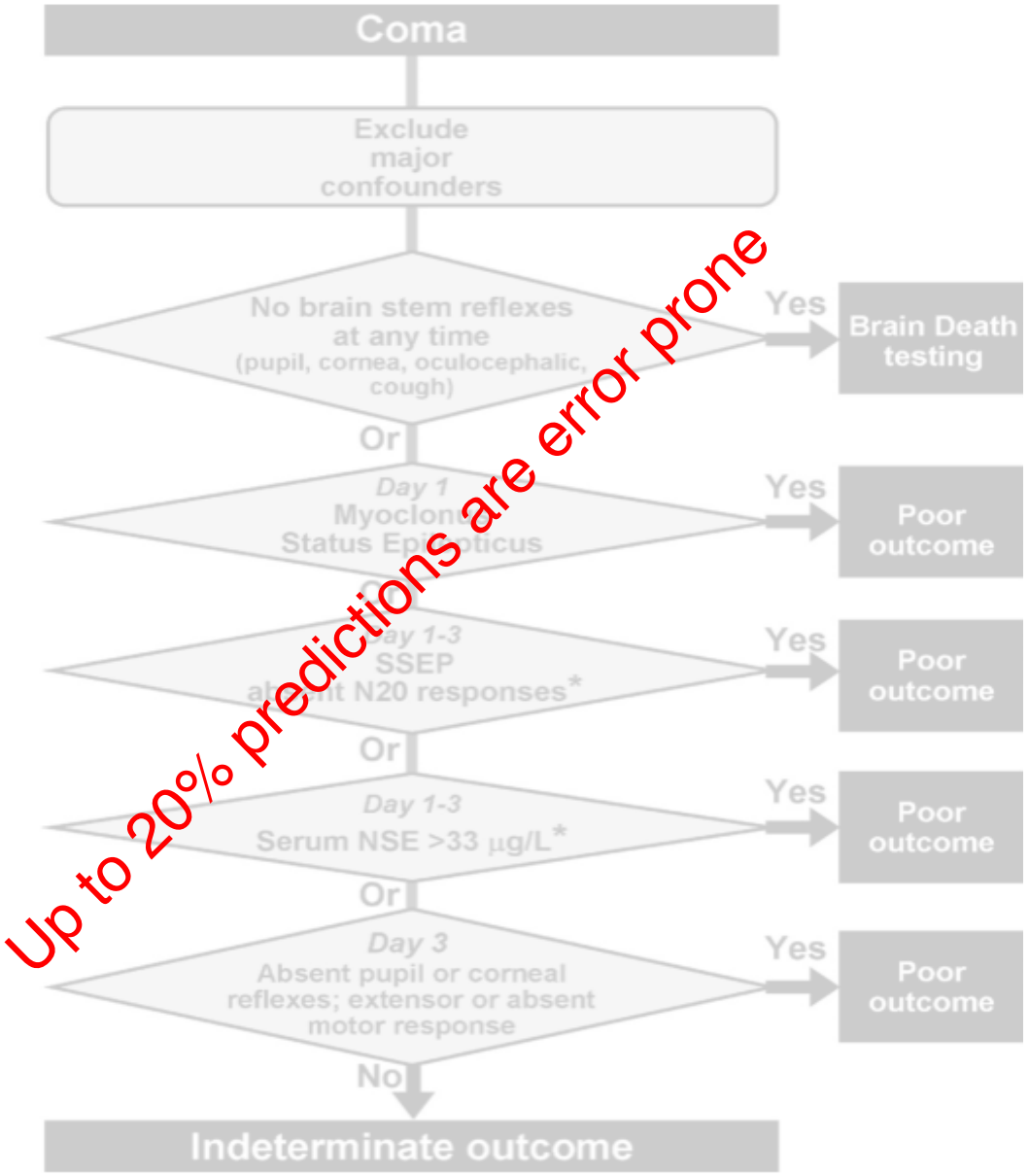
- A sequence of clinical observations and tests performed following cardiac arrest
- Accurate in predicting poor neurological outcomes when severe deficits are present (FPR < 1%)
- Less guidance in cases where clear-cut signs are lacking, or under varying protocols (e.g. therapeutic hypothermia)



The American Academy of Neurology

Current Prognostic Guidelines

- A sequence of clinical observations and tests performed following cardiac arrest
- Accurate in predicting poor neurological outcomes when severe deficits are present (FPR < 1%)
- Less guidance in cases where clear-cut signs are lacking, or under varying protocols (e.g. therapeutic hypothermia)



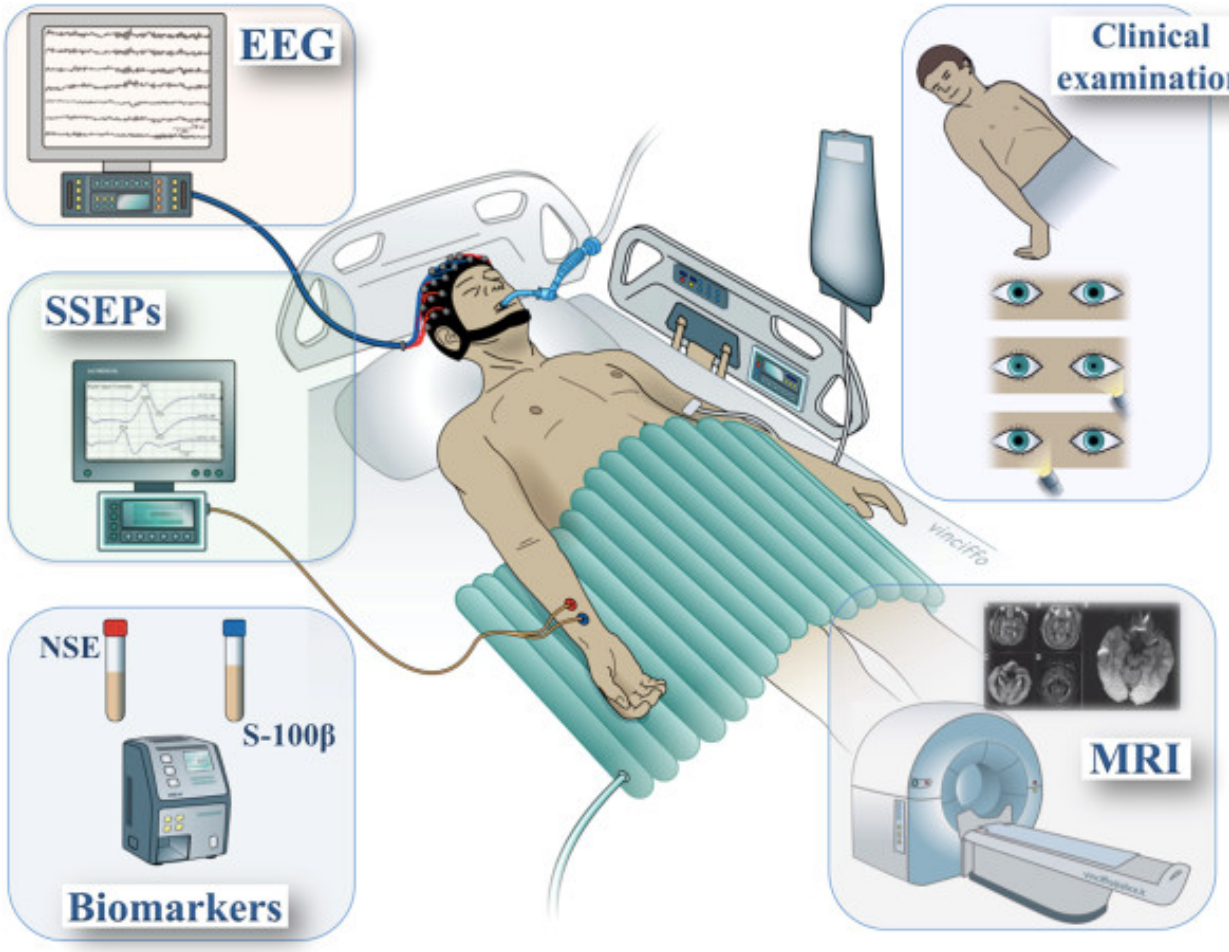
The American Academy of Neurology

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

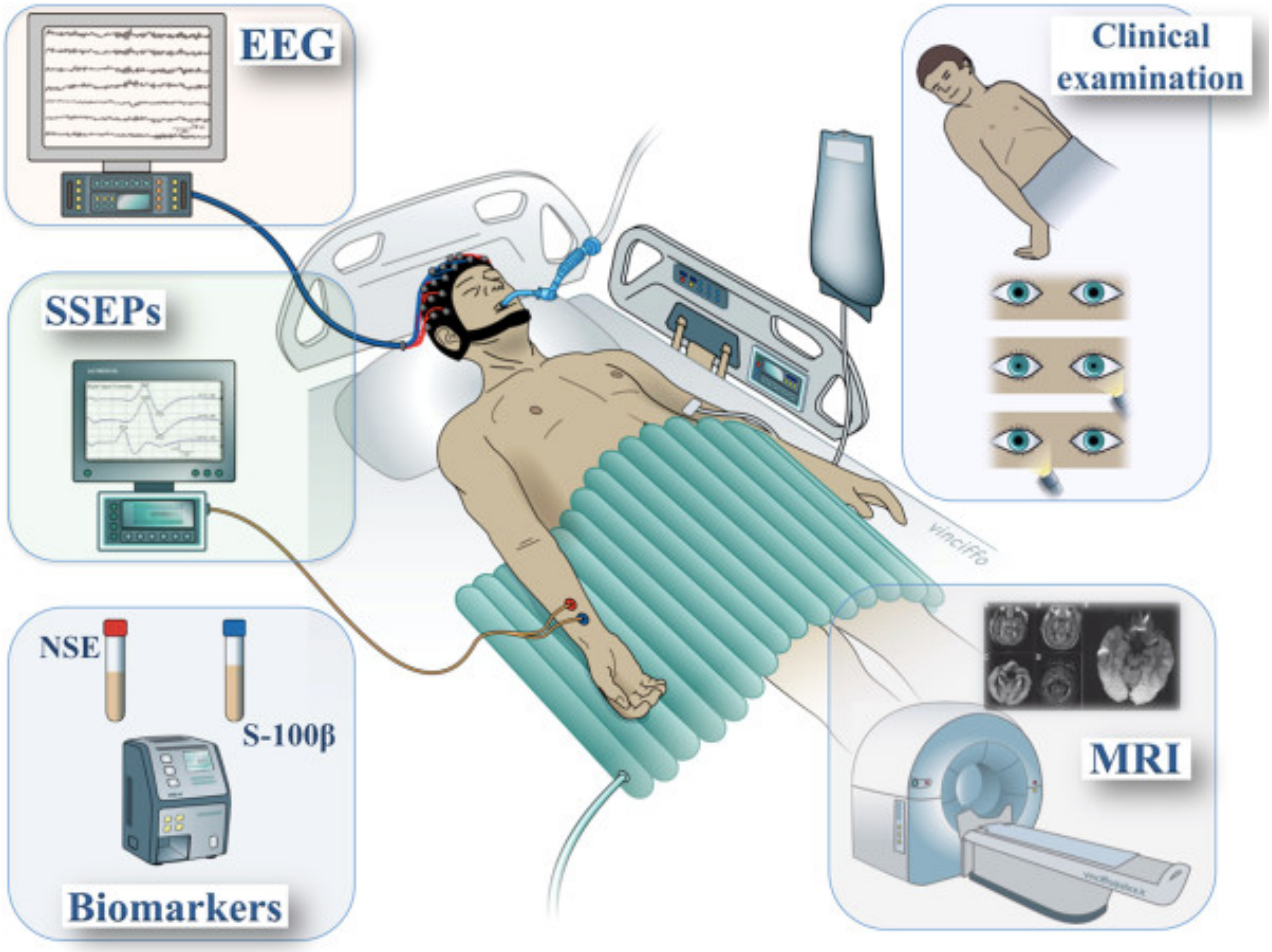
How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Conclusions

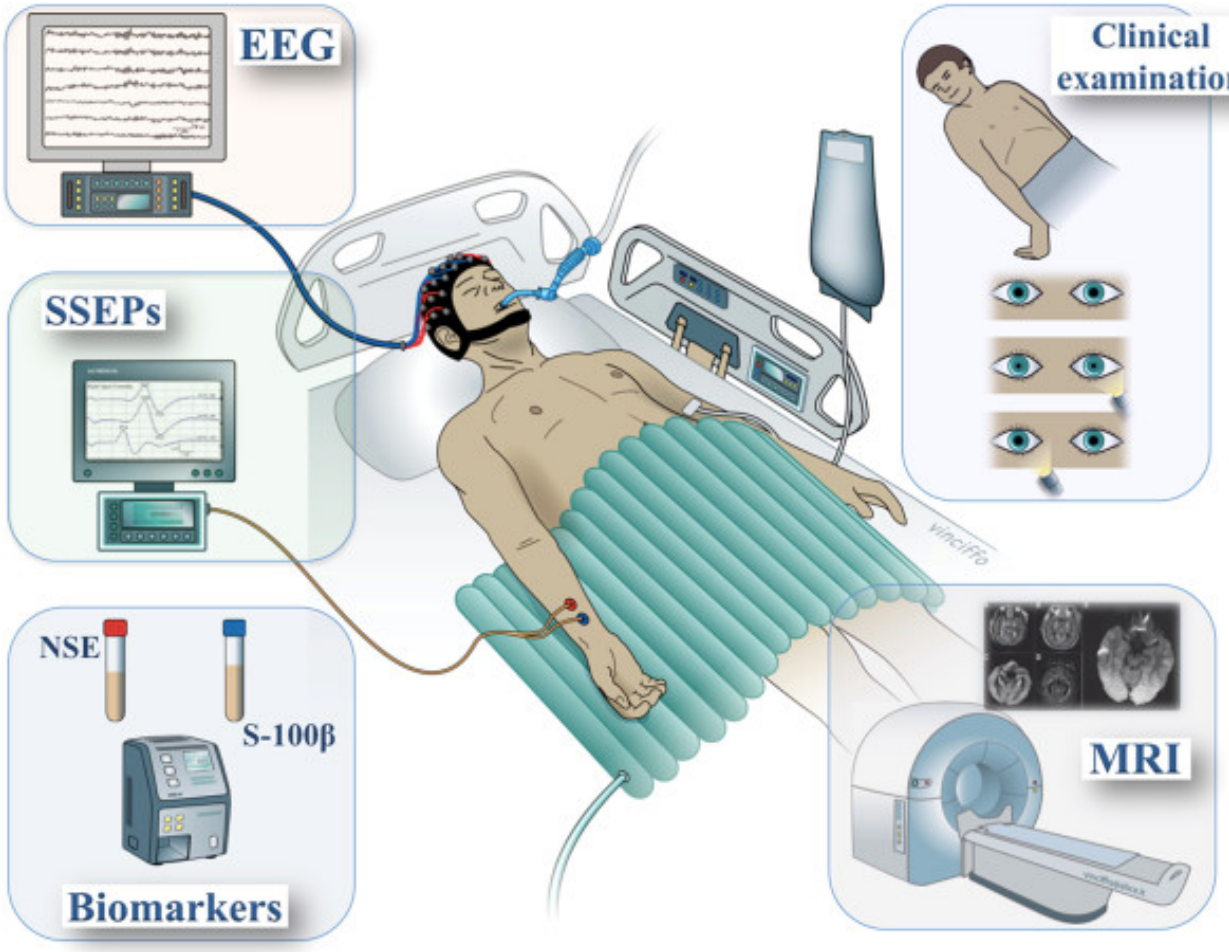
- **Limited Sample Sizes:** larger samples are needed

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

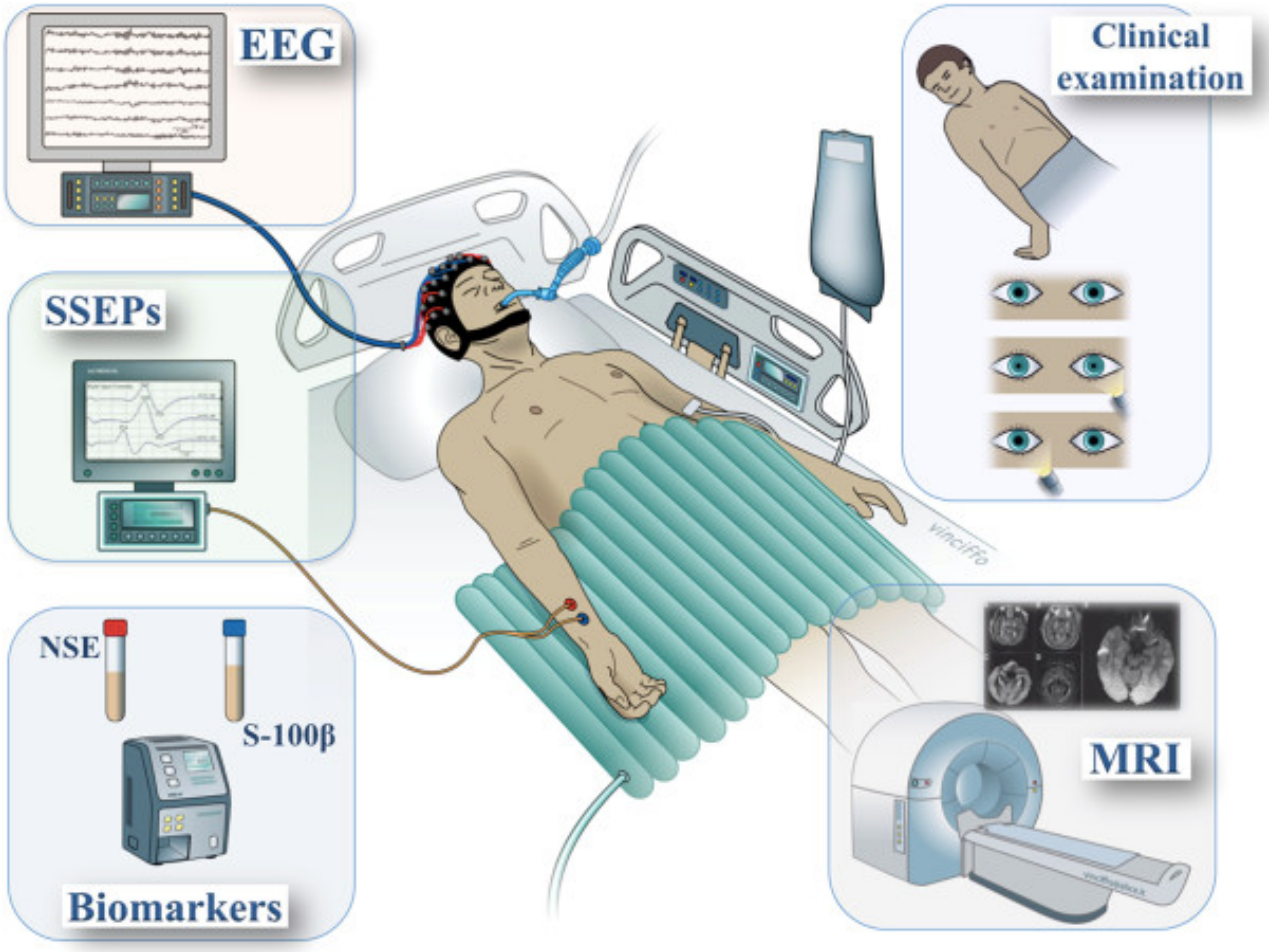
How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Conclusions

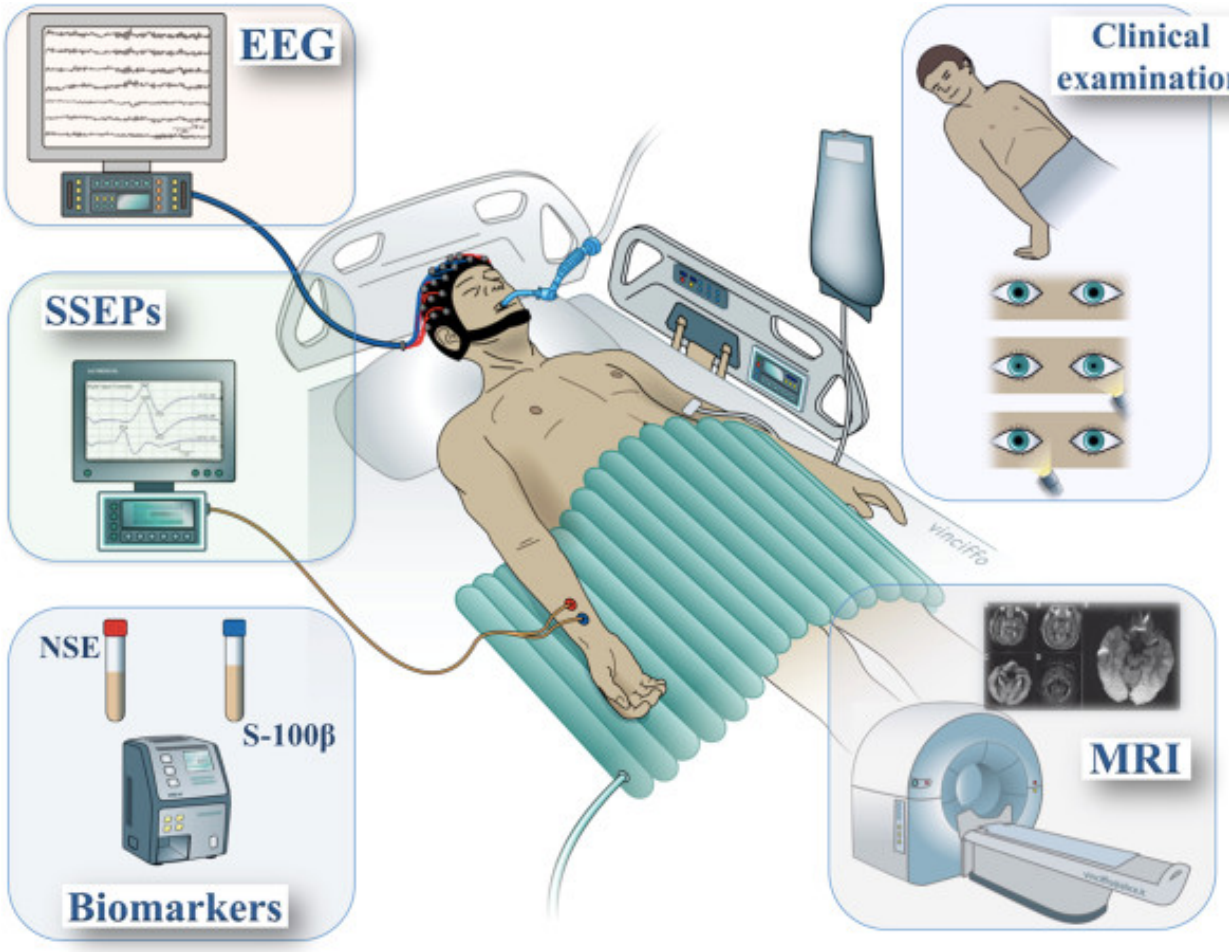
- **Limited Sample Sizes:** larger samples are needed
- **Time Specific:** prediction should be possible at all points in time

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

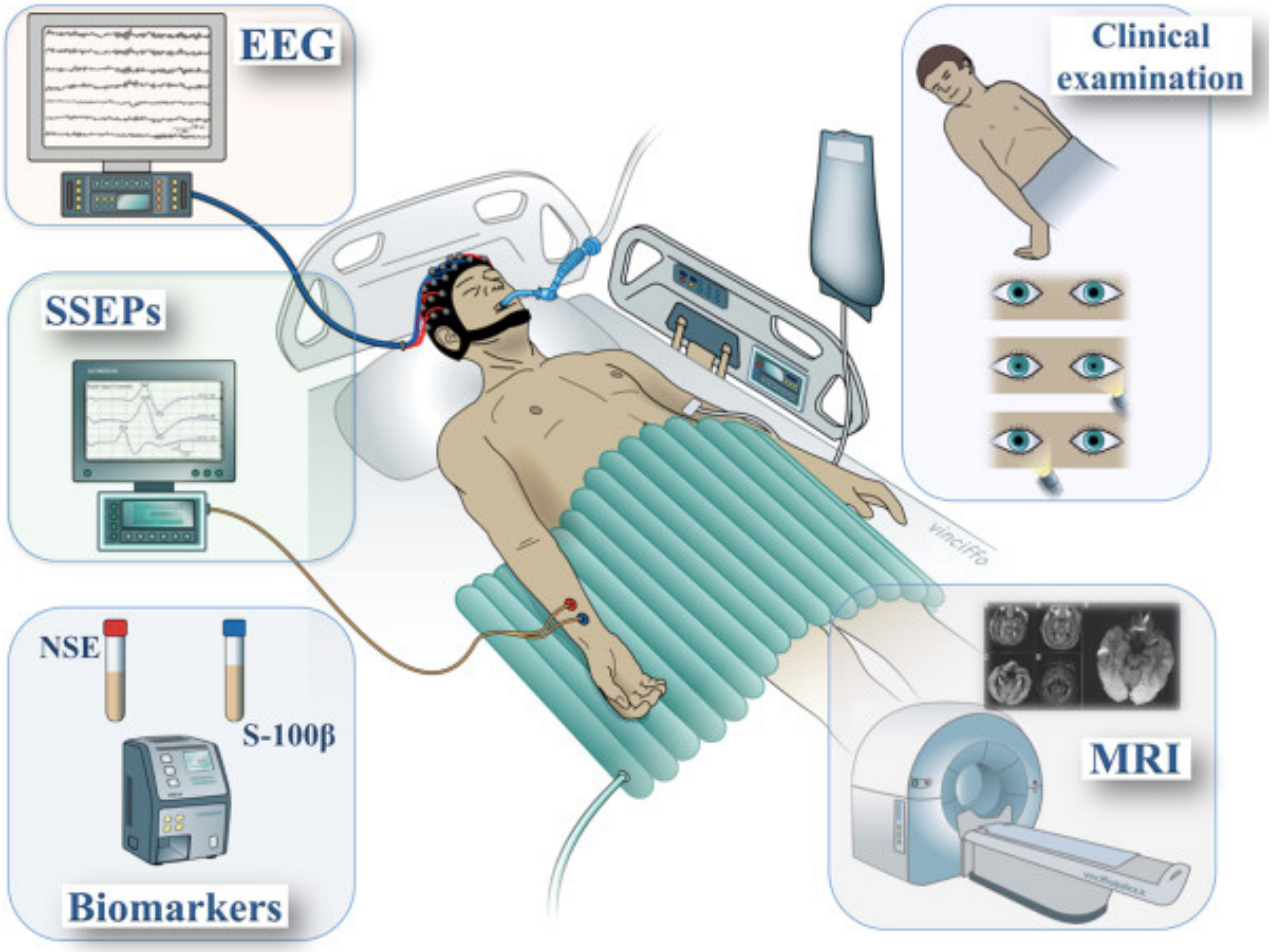
How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Conclusions

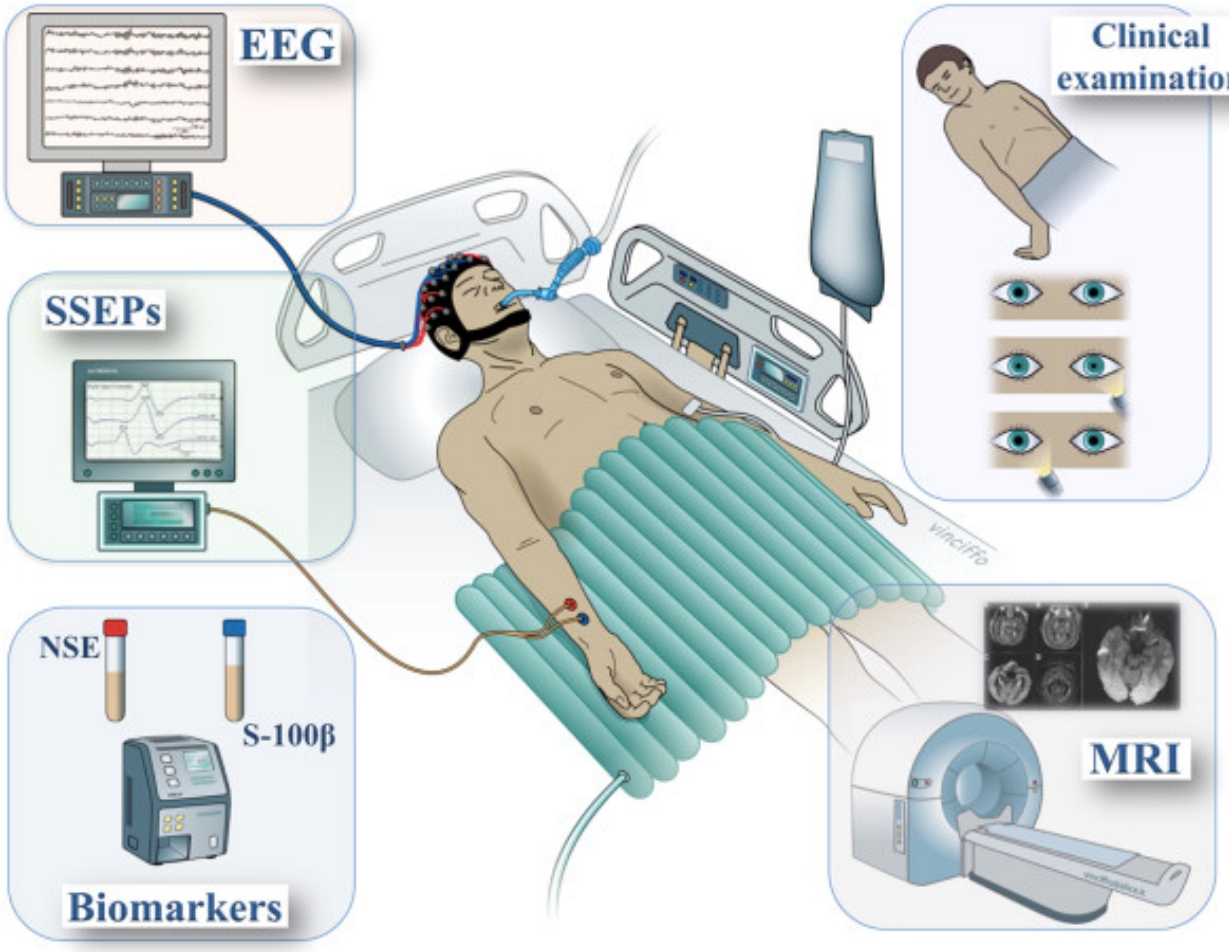
- **Limited Sample Sizes:** larger samples are needed
- **Time Specific:** prediction should be possible at all points in time
- **Classification focused:** risk scoring may be better

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

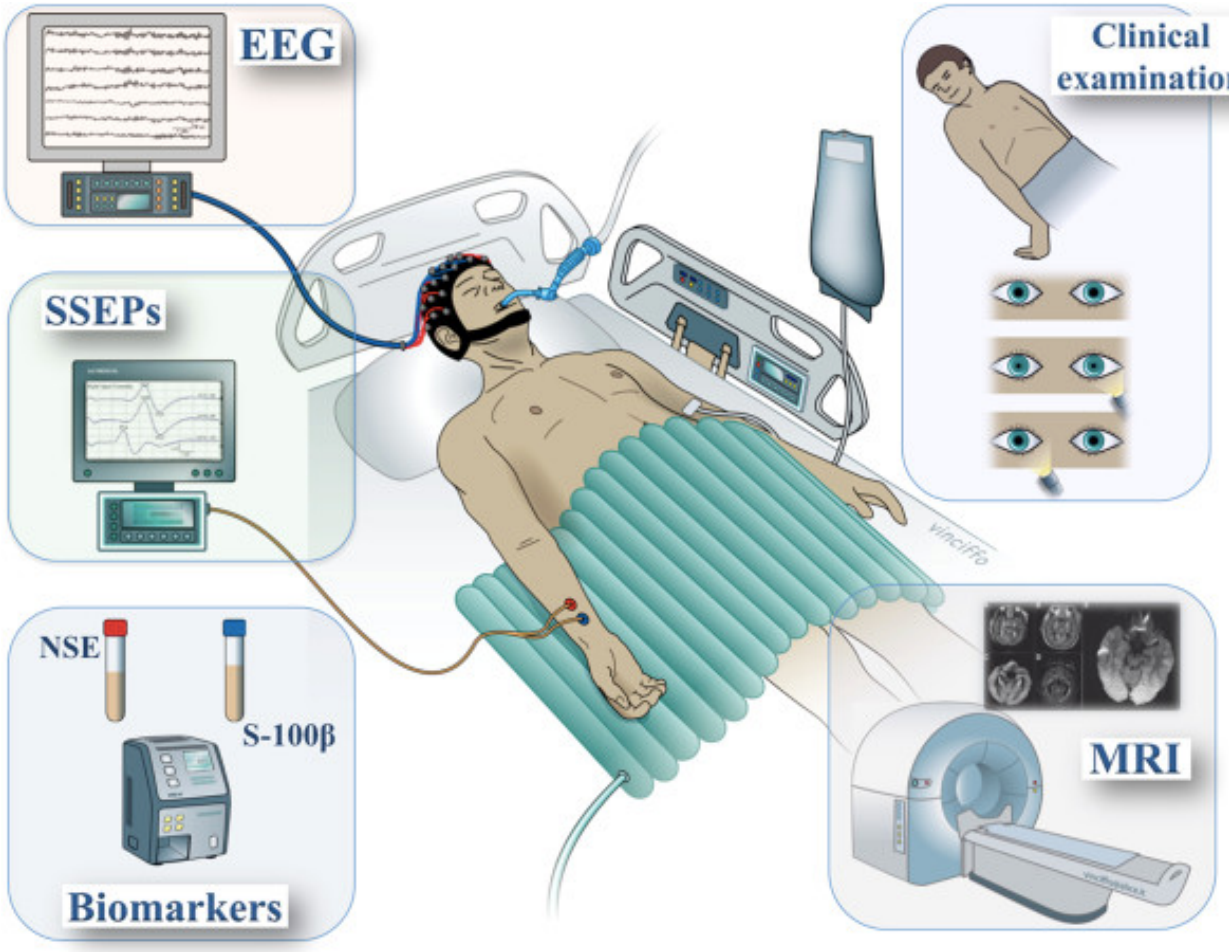
How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

Review of the Literature

EEG Features
Tjepkema-Cloostermans, 2017
AUC = 0.94 at 24 hours
N = 283

Pain-related SSEP
Zanatta, 2015
AUC = 0.84 at 72 hours
N = 167

NSE
Dagmar, 2017
AUC = 0.77 at 24 hours
N = 153



Quantitative Pupillometry
Suys, 2015
AUC = 0.81 at 24 hours
N = 50

MRI
Jeon, 2017
AUC = 0.89 at 24 hours
N = 47

How to assess prognosis after cardiac arrest and therapeutic hypothermia, Taccone et al, 2014

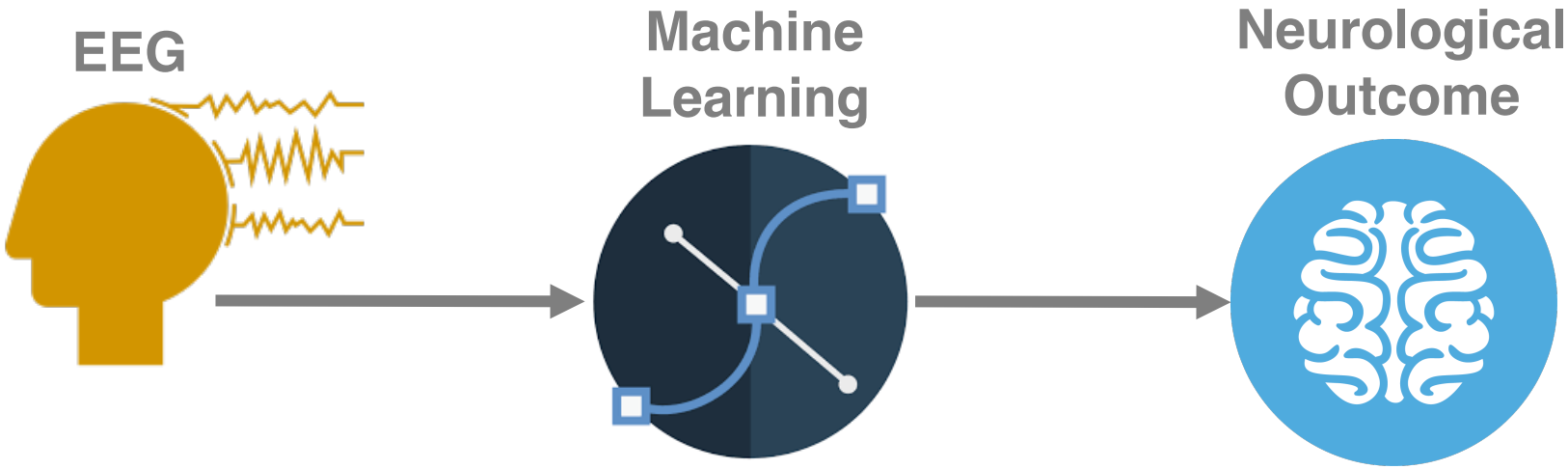
Conclusions

- **Limited Sample Sizes:** larger samples are needed
- **Time Specific:** prediction should be possible at all points in time
- **Classification focused:** risk scoring may be better
- **EEG alone can provide state-of-the-art performance**

Conclusions

- **Limited Sample Sizes:** larger samples are needed
- **Time Specific:** prediction should be possible at all points in time
- **Classification focused:** risk scoring may be better
- **EEG alone can provide state-of-the-art performance**
  - **deep learning could eliminate the need for feature engineering**

Our Goal: Build a better prognostication algorithm

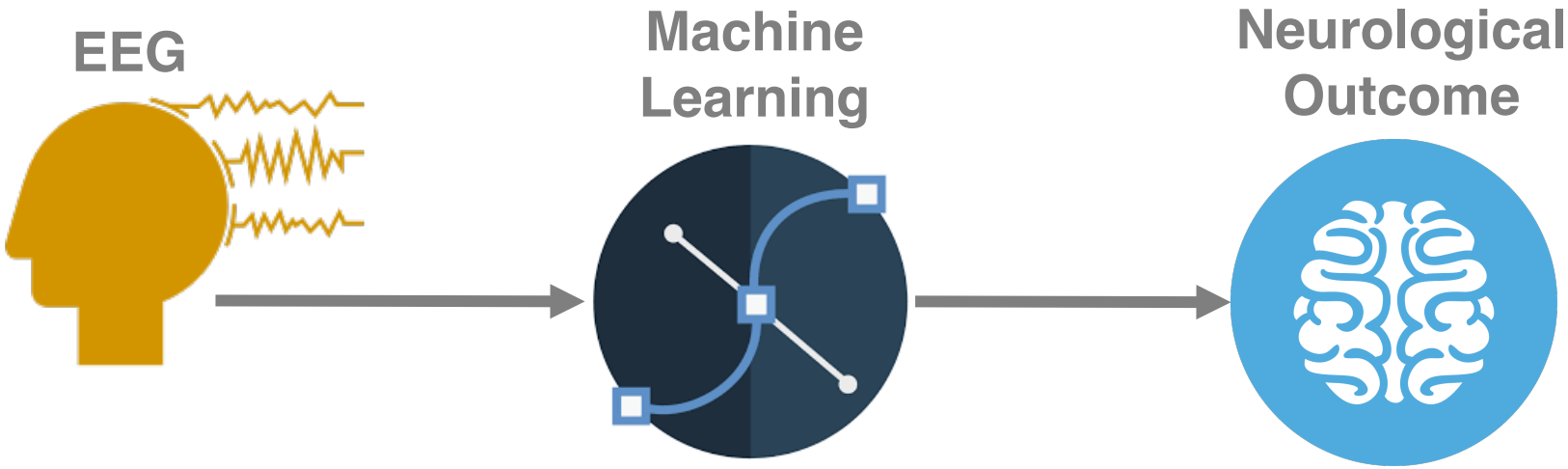


1. ‘Bigger’ data sets

2. Time-sensitive

3. Deployable

Our Goal: Build a better prognostication algorithm

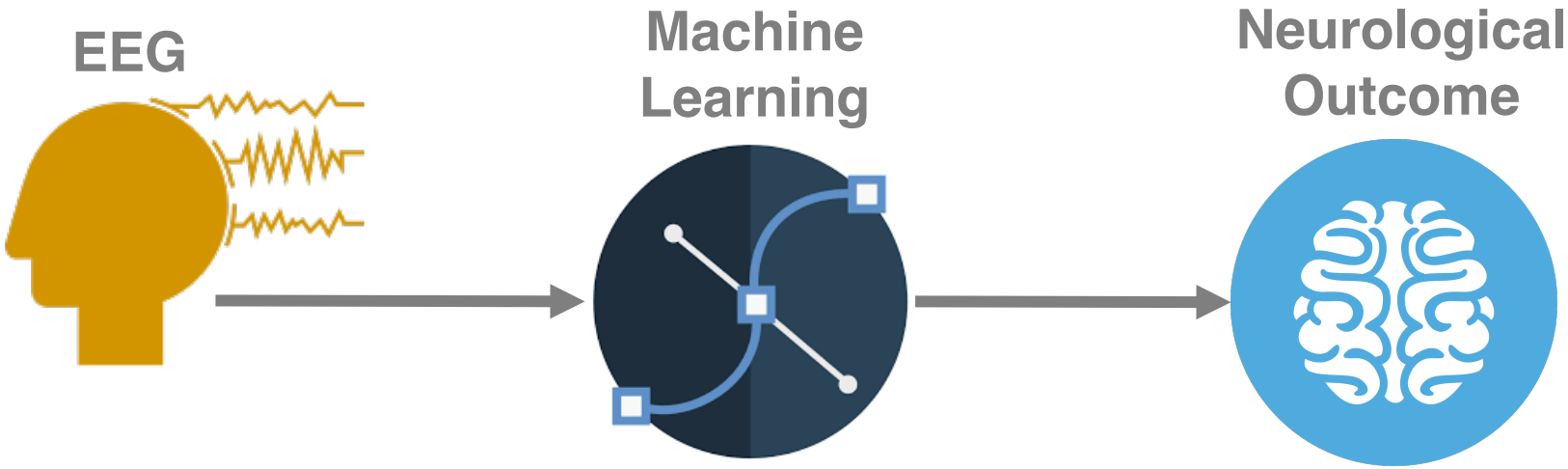


1. ‘Bigger’ data sets

2. Time-sensitive

3. Deployable

Our Goal: Build a better prognostication algorithm

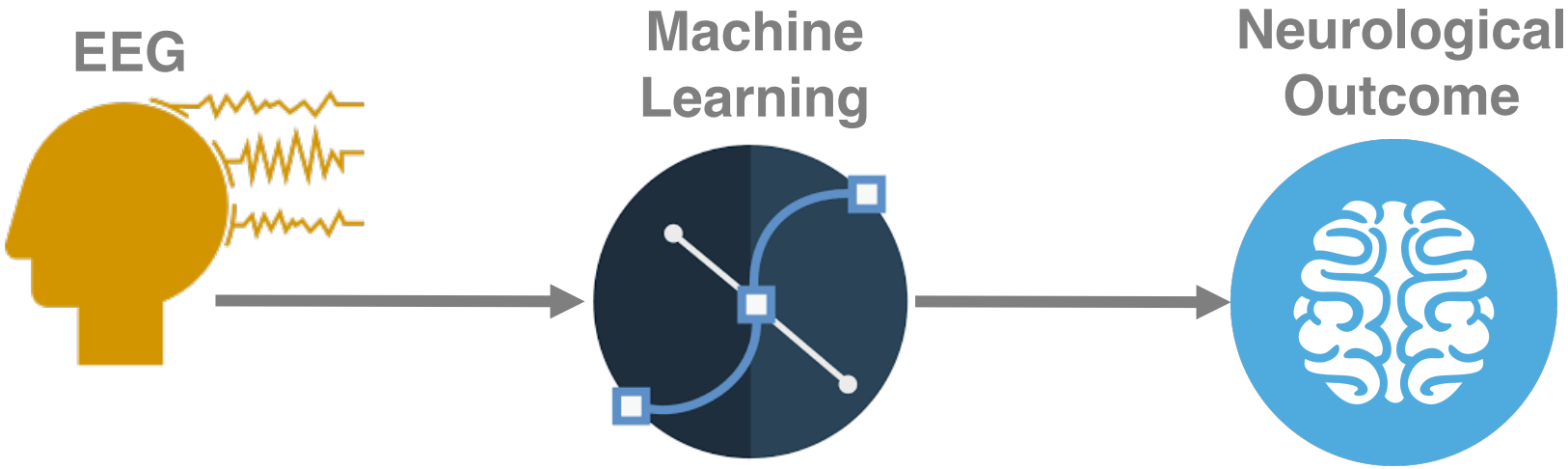


1. ‘Bigger’ data sets

2. Time-sensitive

3. Deployable

Our Goal: Build a better prognostication algorithm

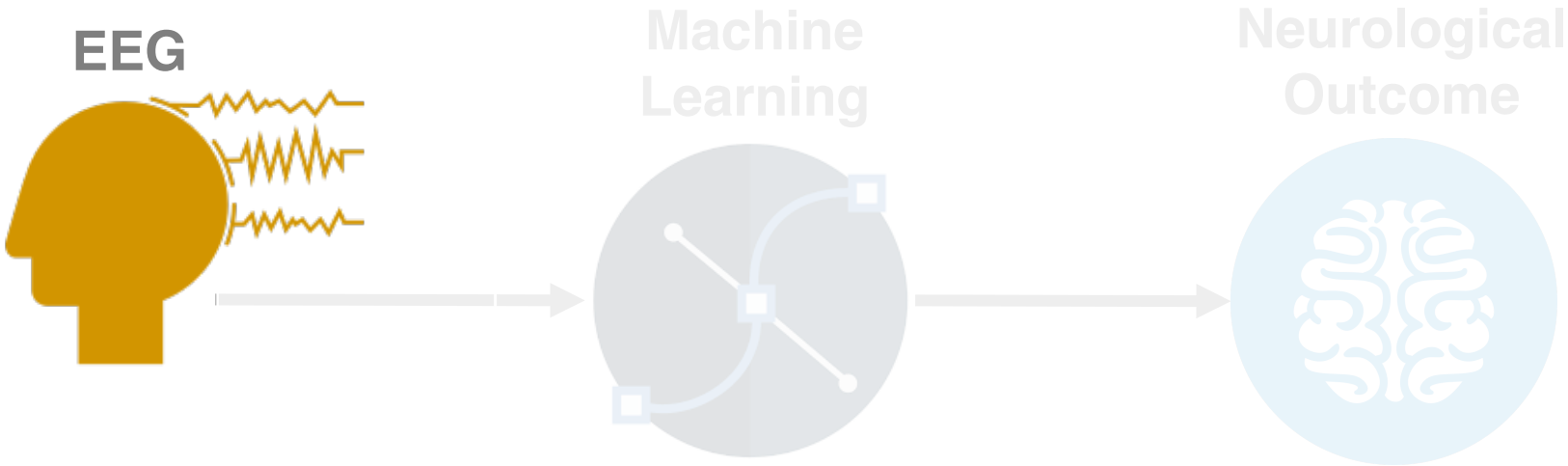


1. ‘Bigger’ data sets

2. Time-sensitive

3. Deployable

Data collection



1. ‘Bigger’ data sets

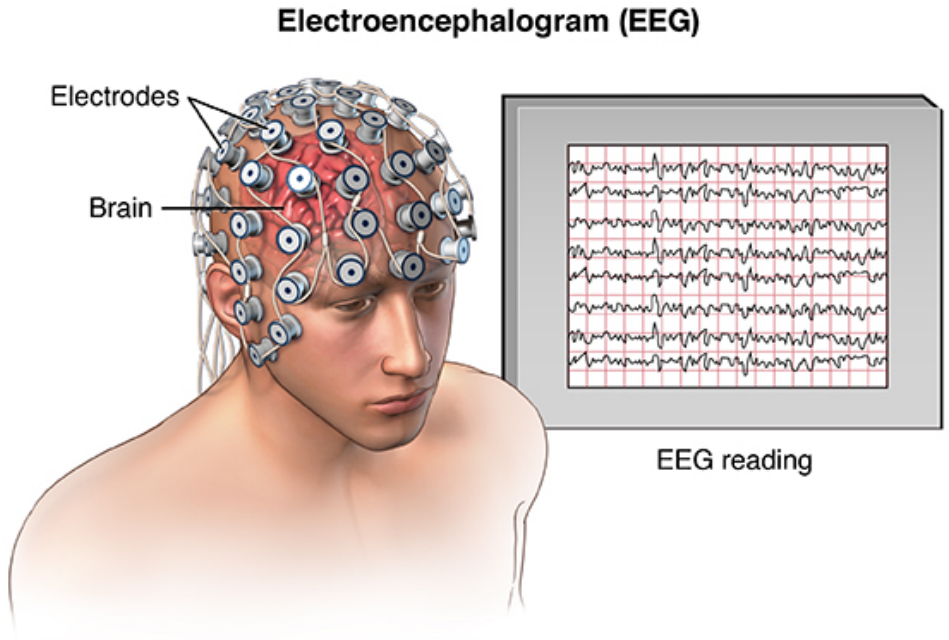
2. Time-sensitive

3. Deployable

Electroencephalography (EEG)

- Each EEG electrode records an **ensemble** of cellular activity near the location of the electrode

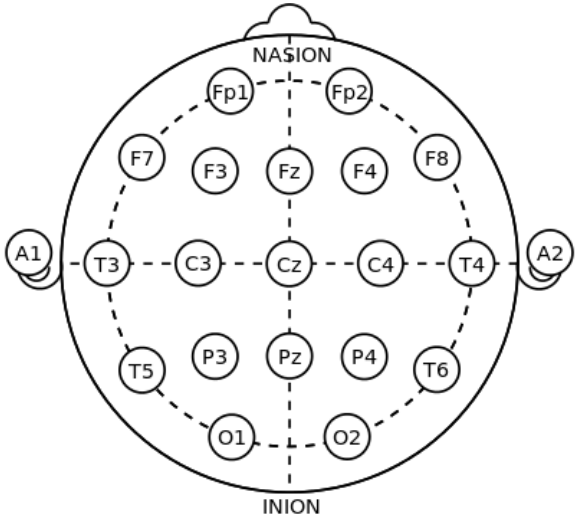
- Electrode Placement was in accordance with the International 10-20 system



Electroencephalography (EEG)

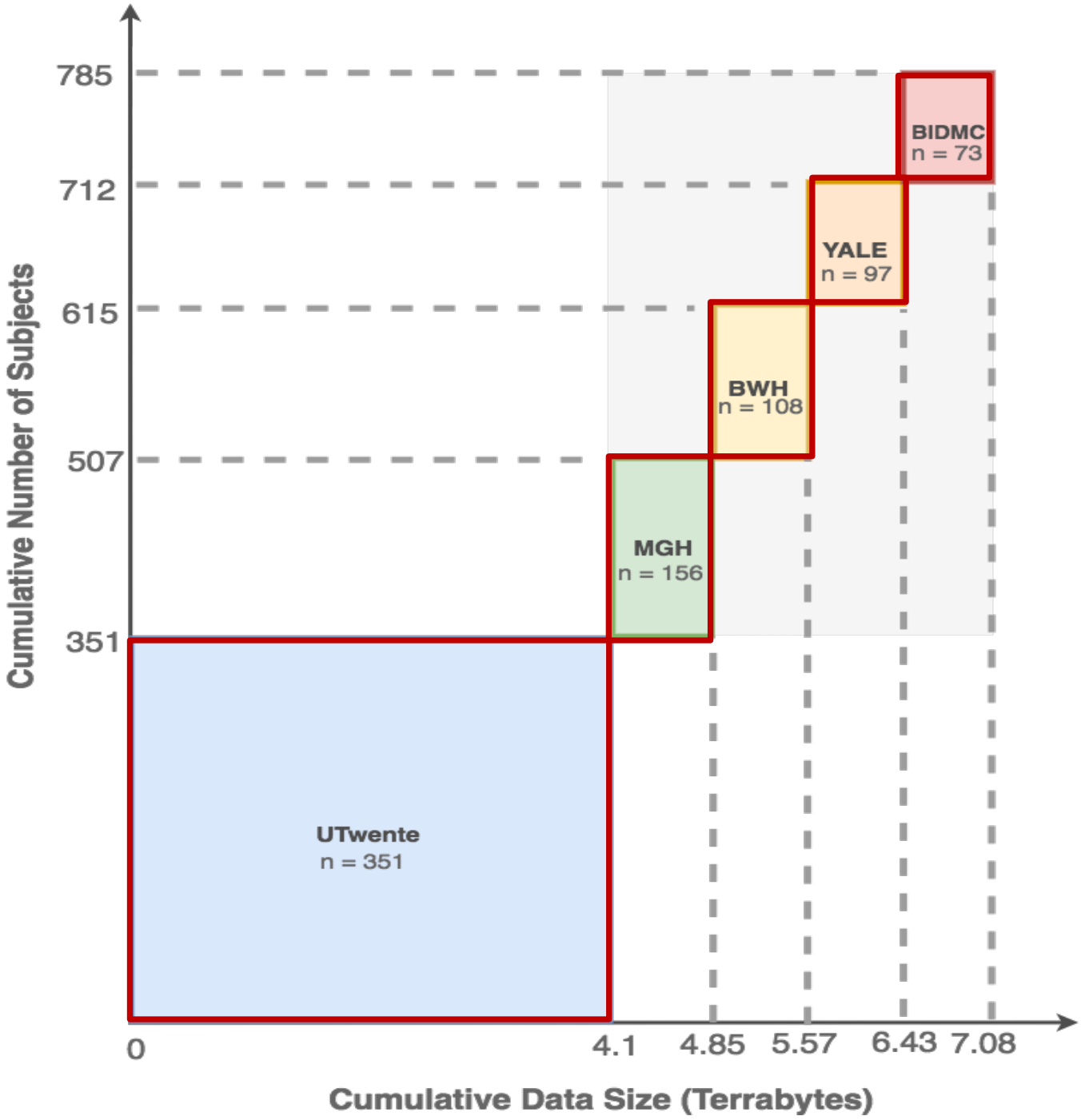
- Each EEG electrode records an **ensemble** of cellular activity near the location of the electrode

- Electrode Placement was in accordance with the International 10-20 system



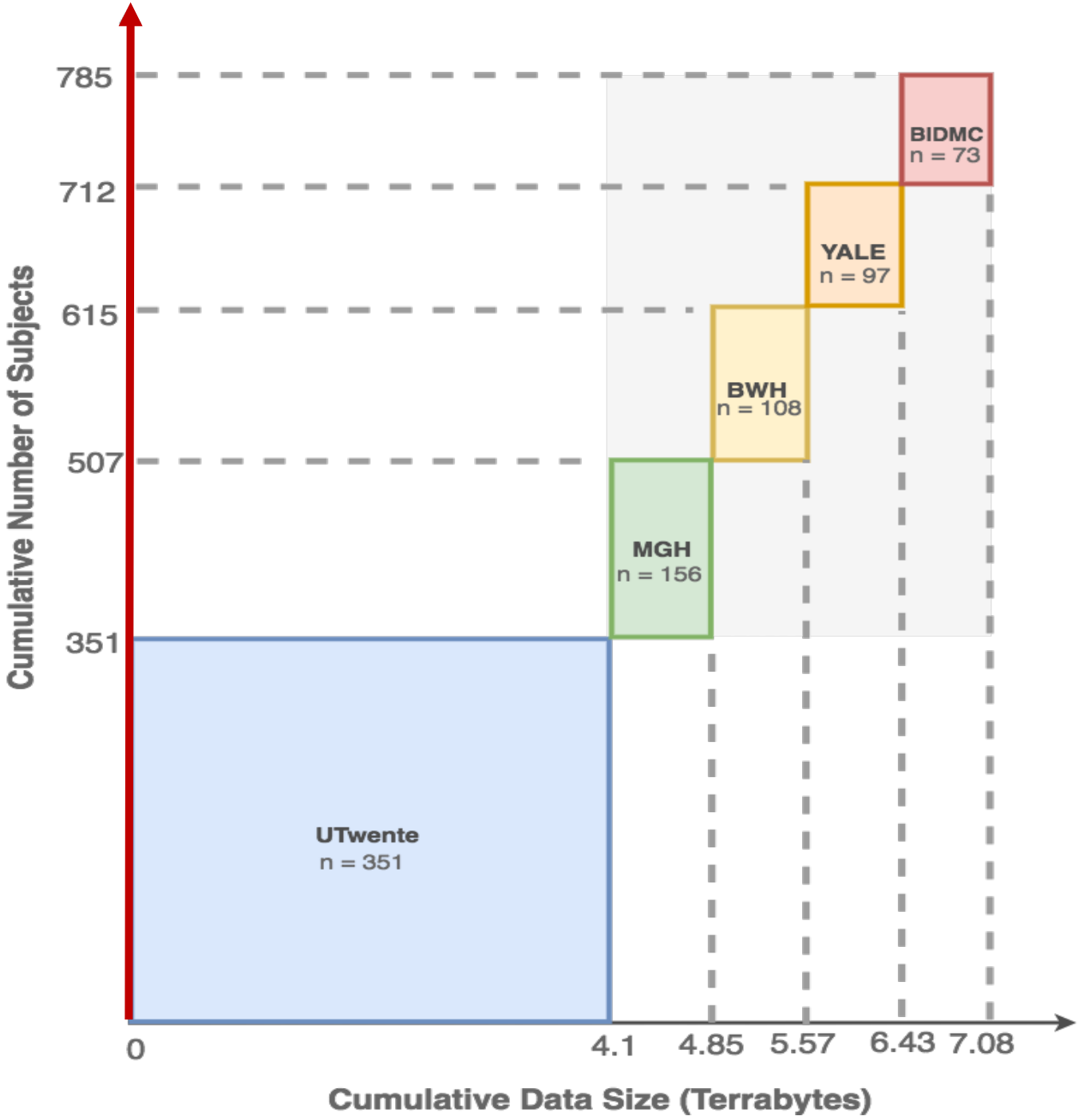
EEG data

- 5 contributing institutions
- 785 unique patients
- 7 terabytes of data



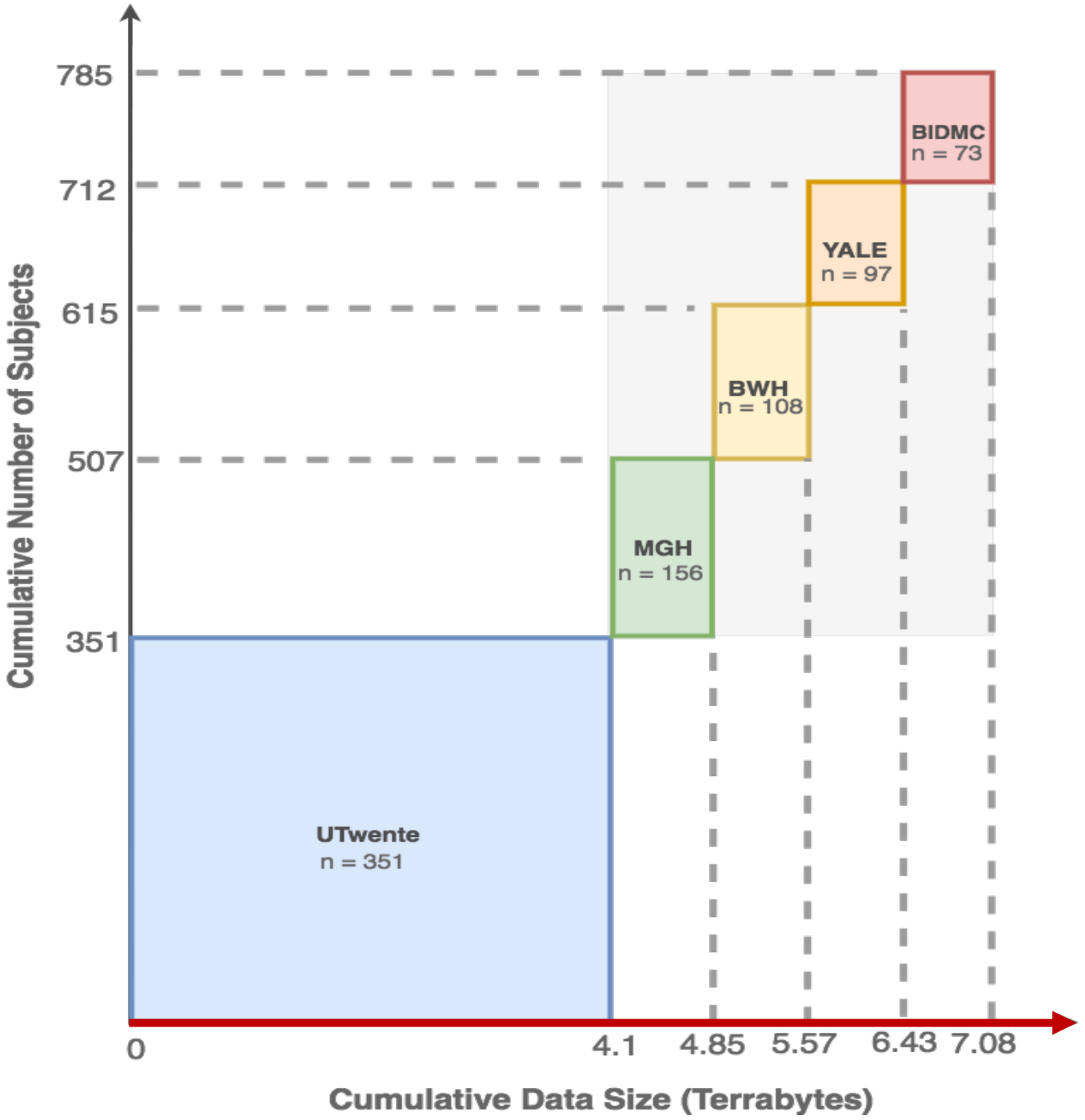
EEG data

- 5 contributing institutions
- 785 unique patients
- 7 terabytes of data



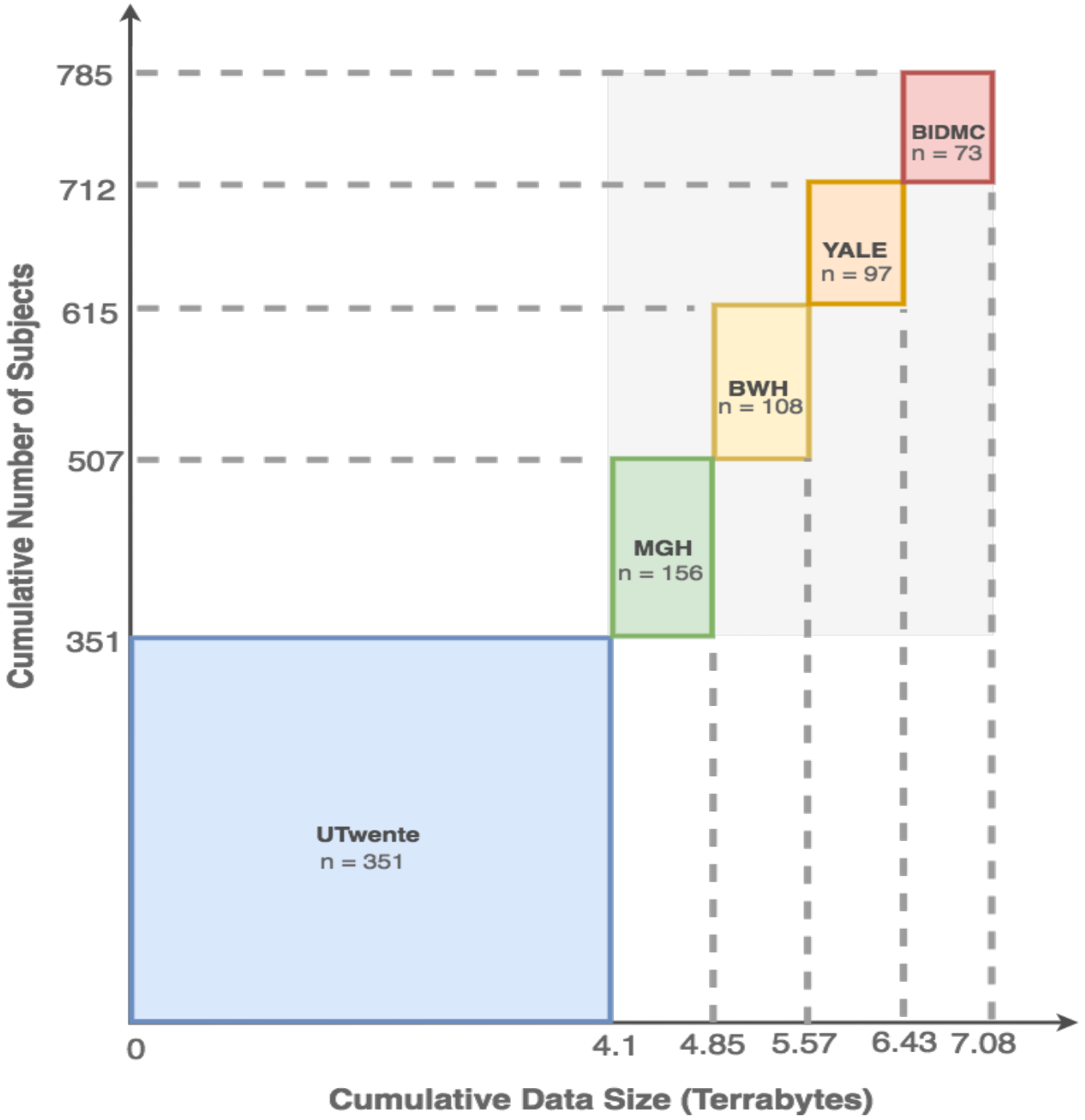
EEG data

- 5 contributing institutions
- 785 unique patients
- 7 terabytes of data



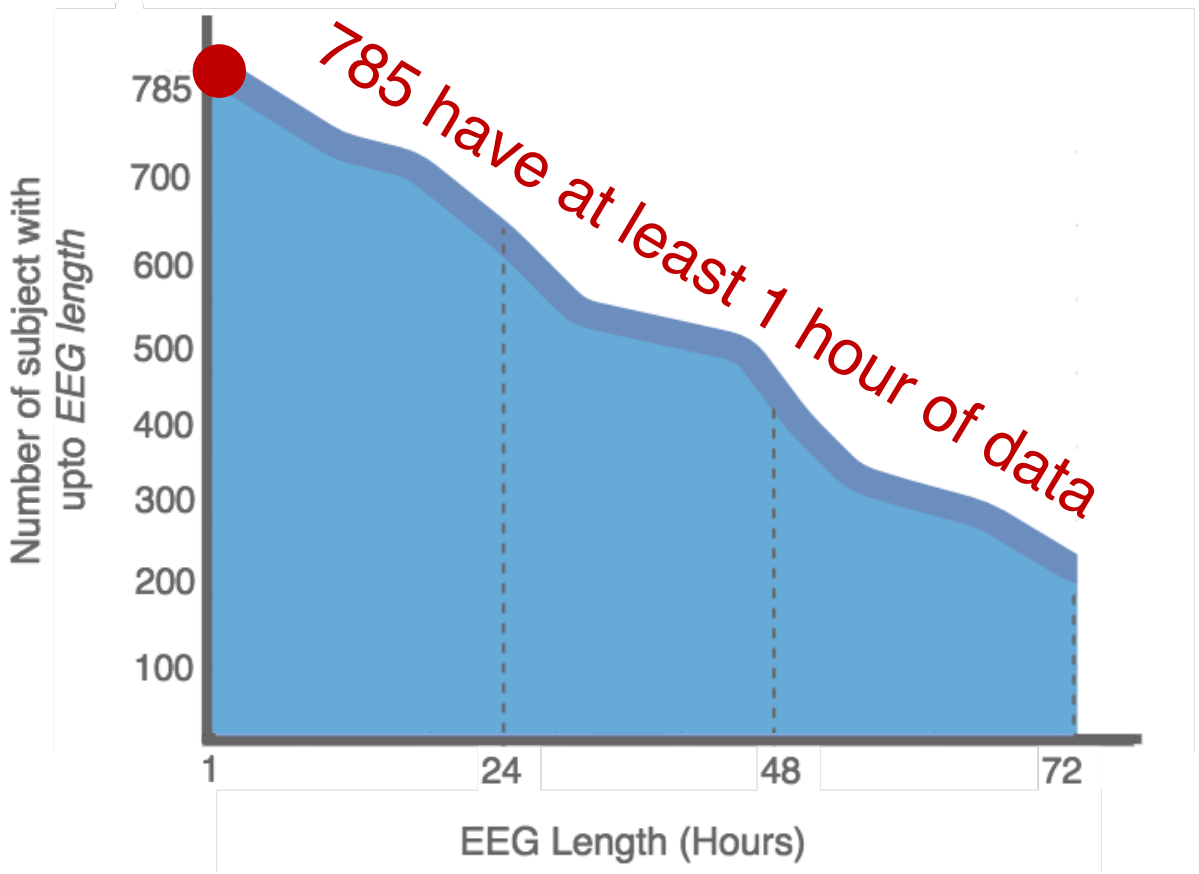
EEG data

- 5 contributing institutions
- 785 unique patients
- 7 terabytes of data
- Over 2x larger than existing archives described in the literature



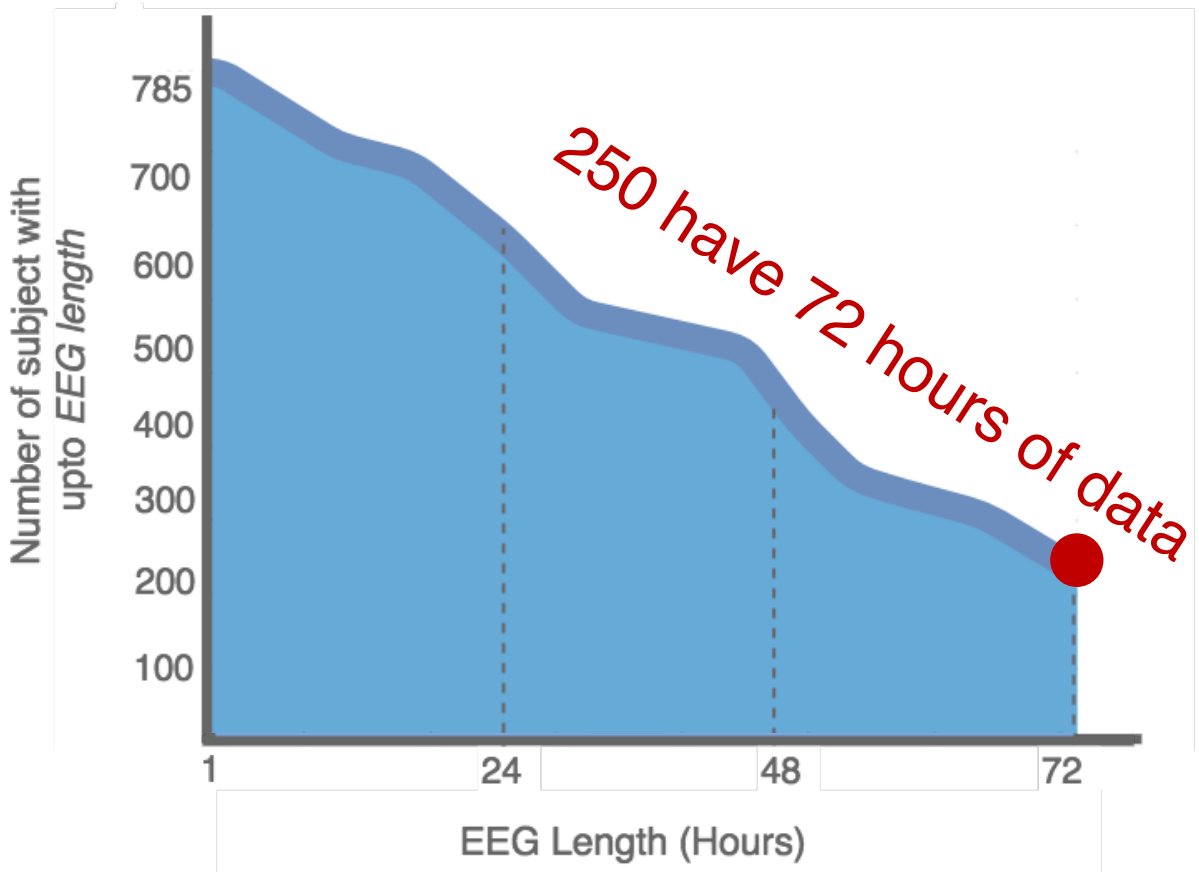
EEG temporal properties

- Data densities linearly decrease over time
- EEG withdrawal is assessed approximately once every 24 hours



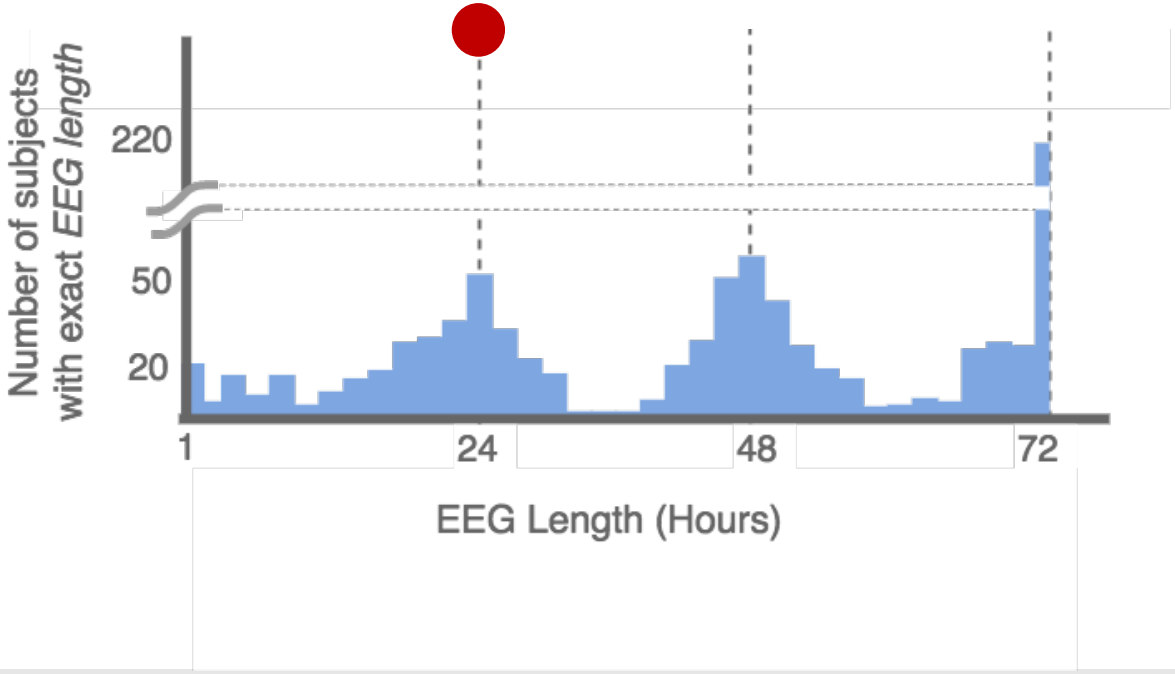
EEG temporal properties

- Data densities linearly decrease over time
- EEG withdrawal is assessed approximately once every 24 hours



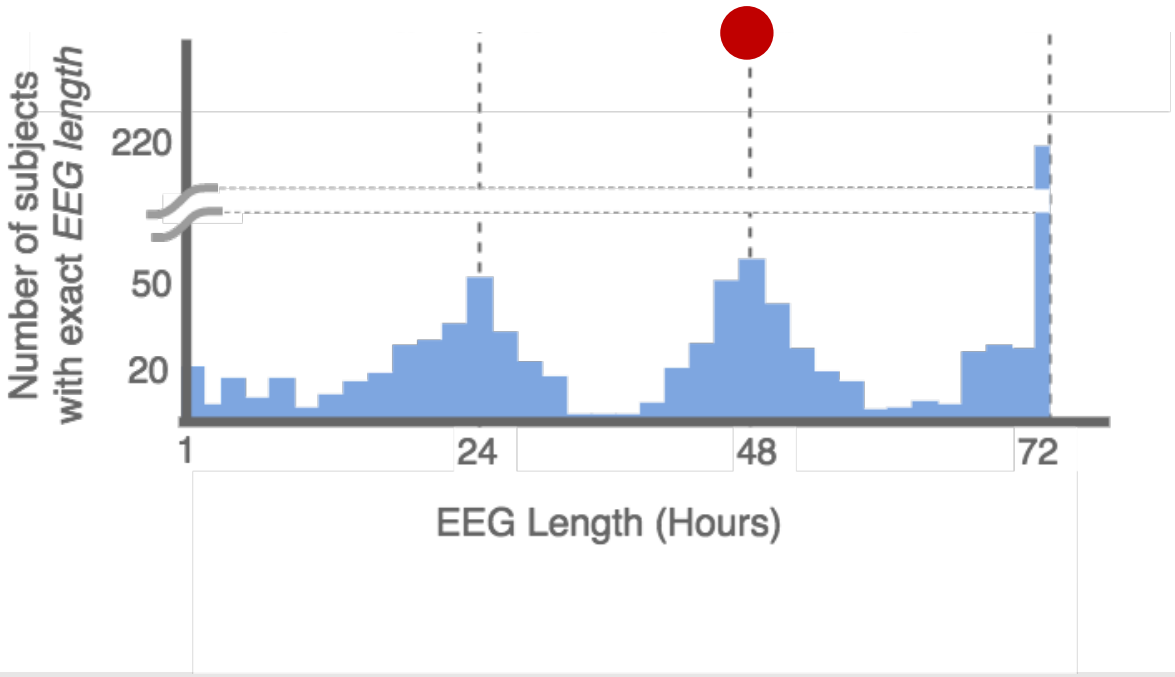
EEG temporal properties

- Data densities linearly decrease over time
- EEG withdrawal is assessed approximately once every 24 hours



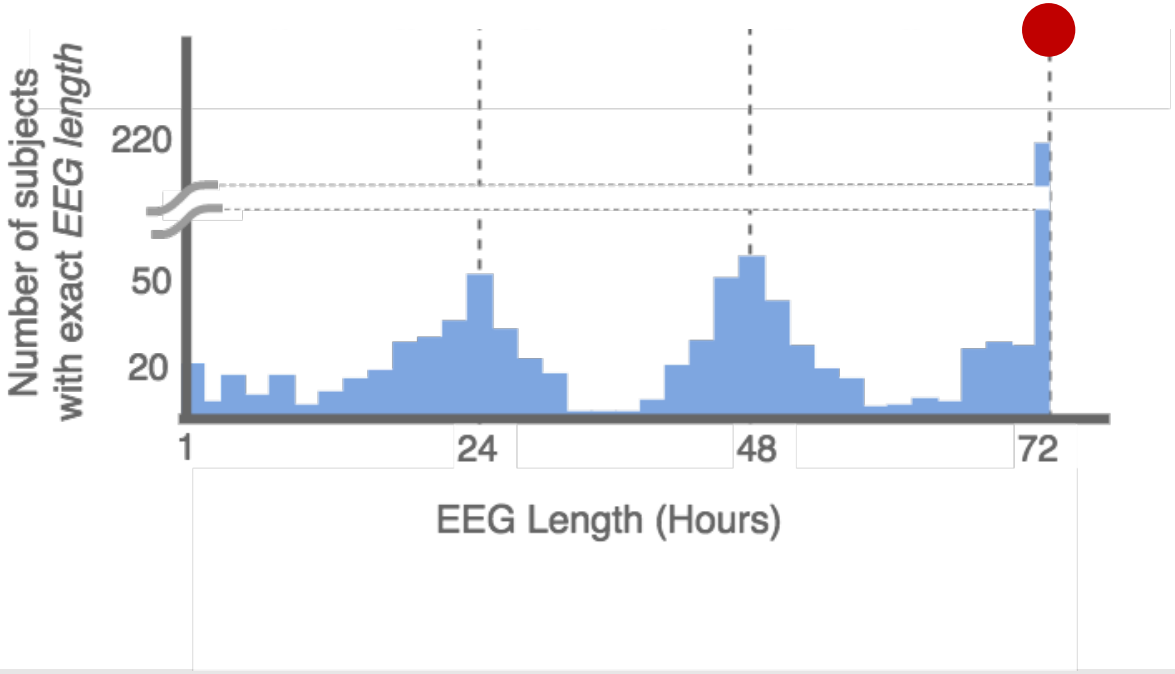
EEG temporal properties

- Data densities linearly decrease over time
- EEG withdrawal is assessed approximately once every 24 hours



EEG temporal properties

- Data densities linearly decrease over time
- EEG withdrawal is assessed approximately once every 24 hours

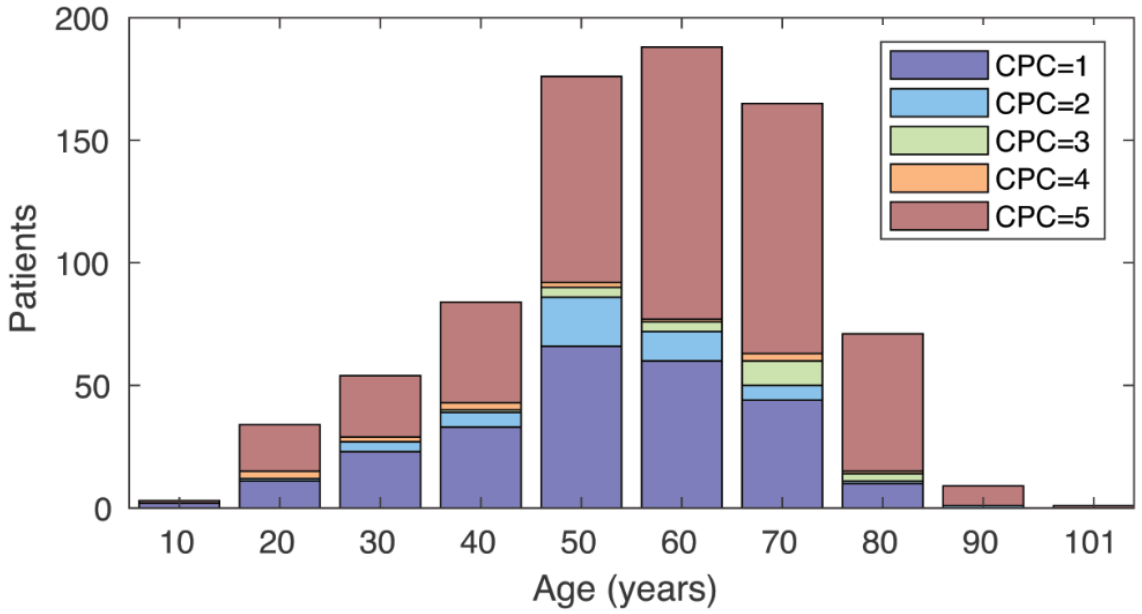


Clinical Characteristics

Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other (PEA / Asystole)	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

Clinical Characteristics

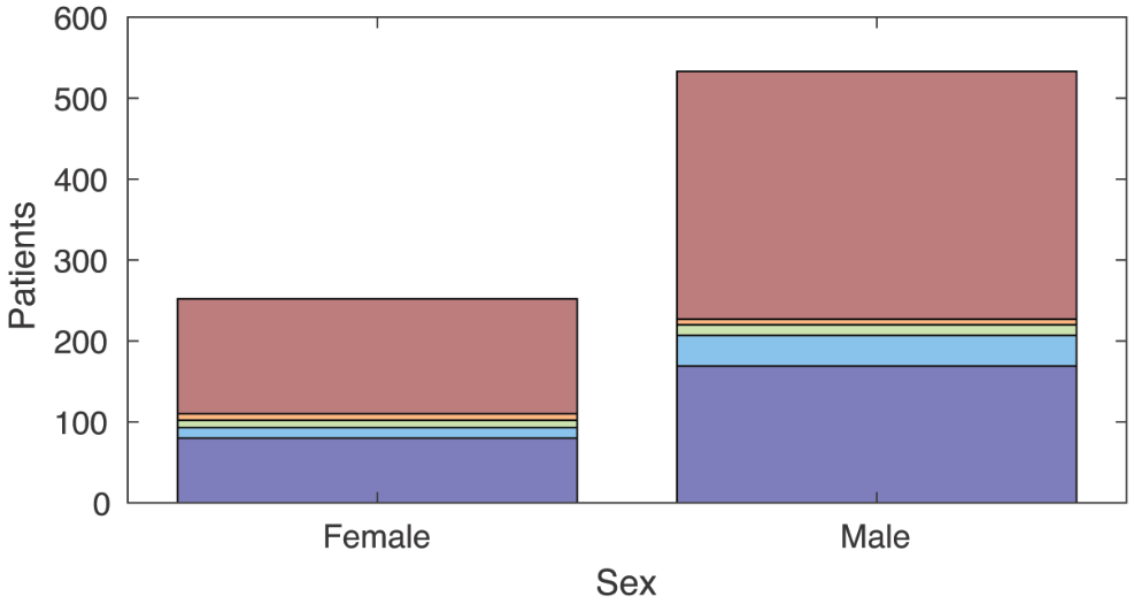
- Average Age: 60
- Older patients do worse



Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
Time to ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

Clinical Characteristics

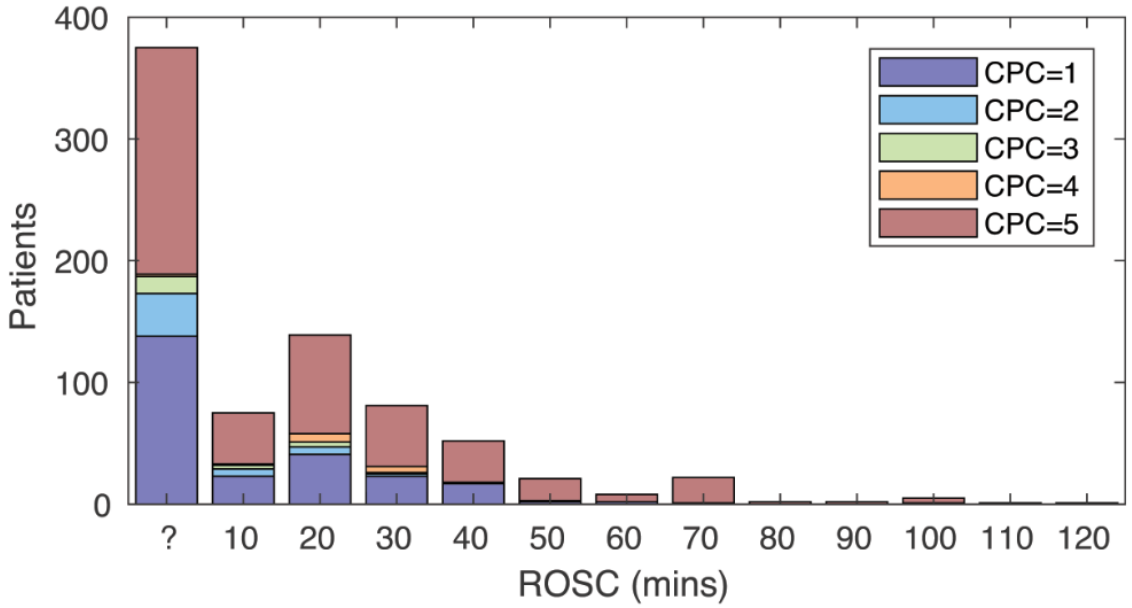
- 68% Male
- No difference w.r.t. outcome



Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
Time to ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

Clinical Characteristics

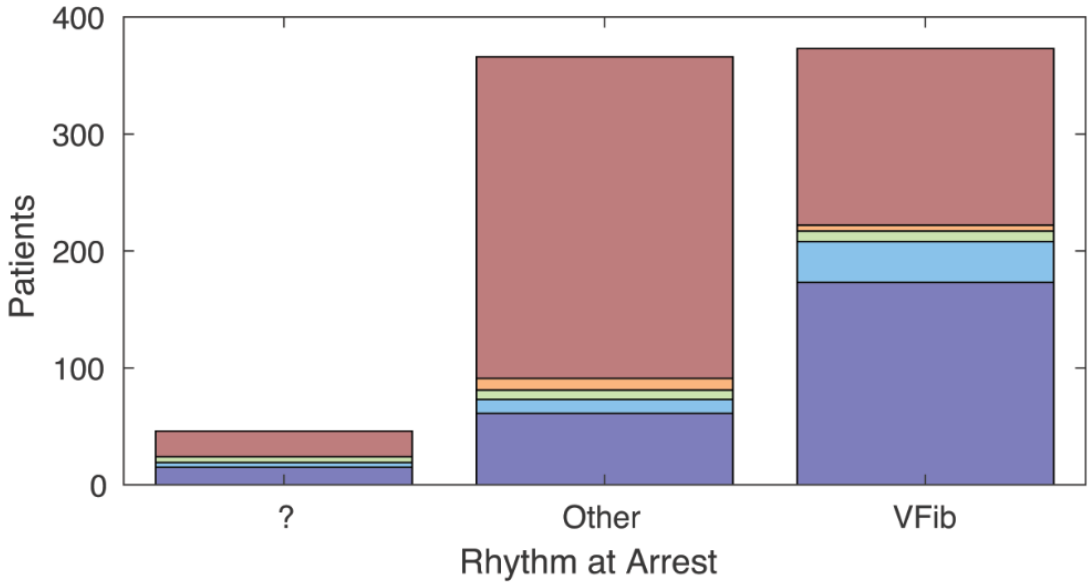
- Average ROSC time: 23 mins
- Longer time to ROSC is bad



Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
Time to ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

Clinical Characteristics

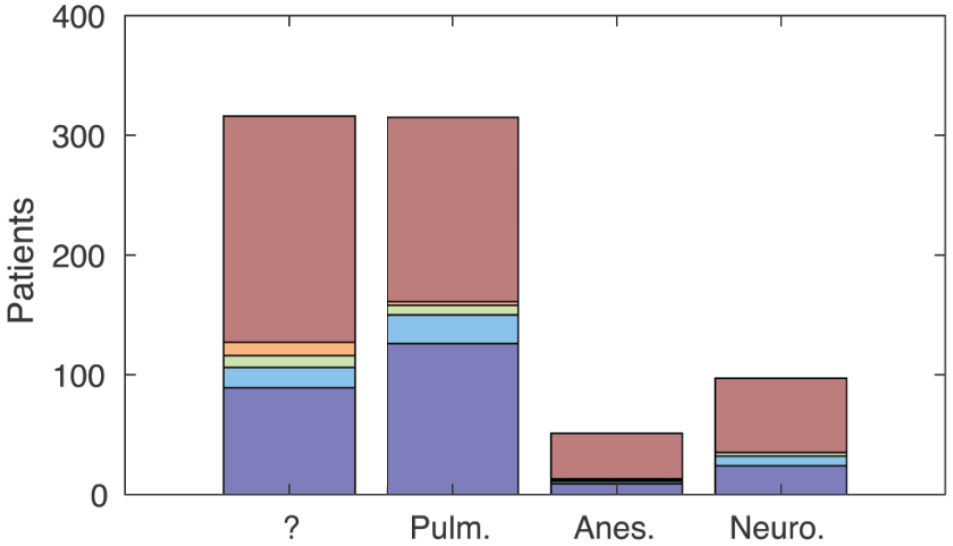
- 50% have Vfib arrest rhythm
- VFib patients do better



Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
Time to ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other (PEA / Asystole)	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

Clinical Characteristics

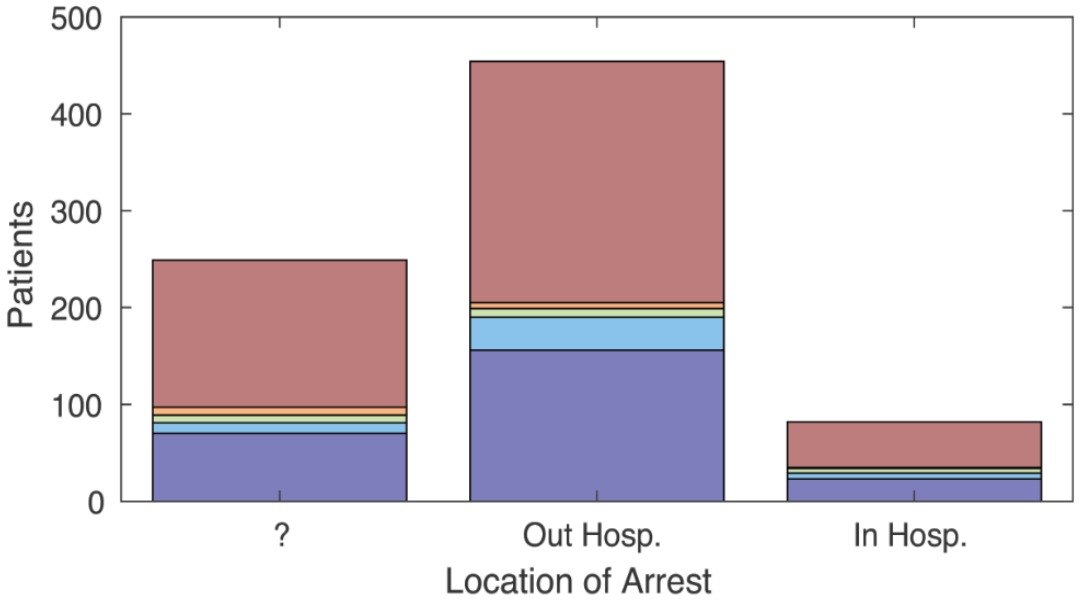
- Causes are often unknown



Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
Time to ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

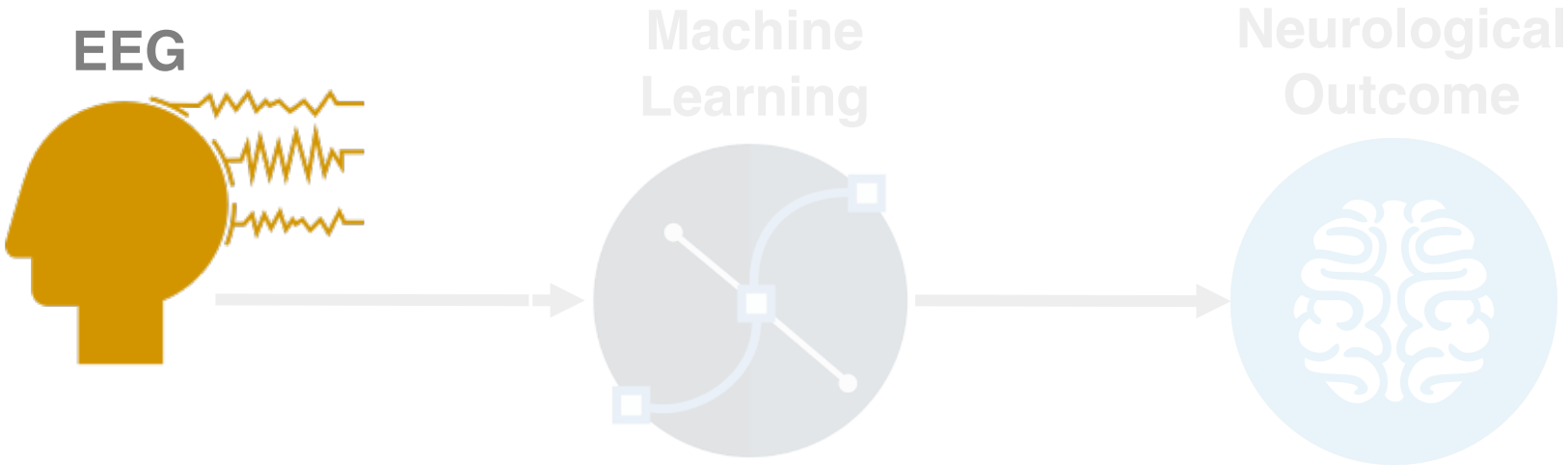
Clinical Characteristics

- Location often unknown
- Most are outside hops.



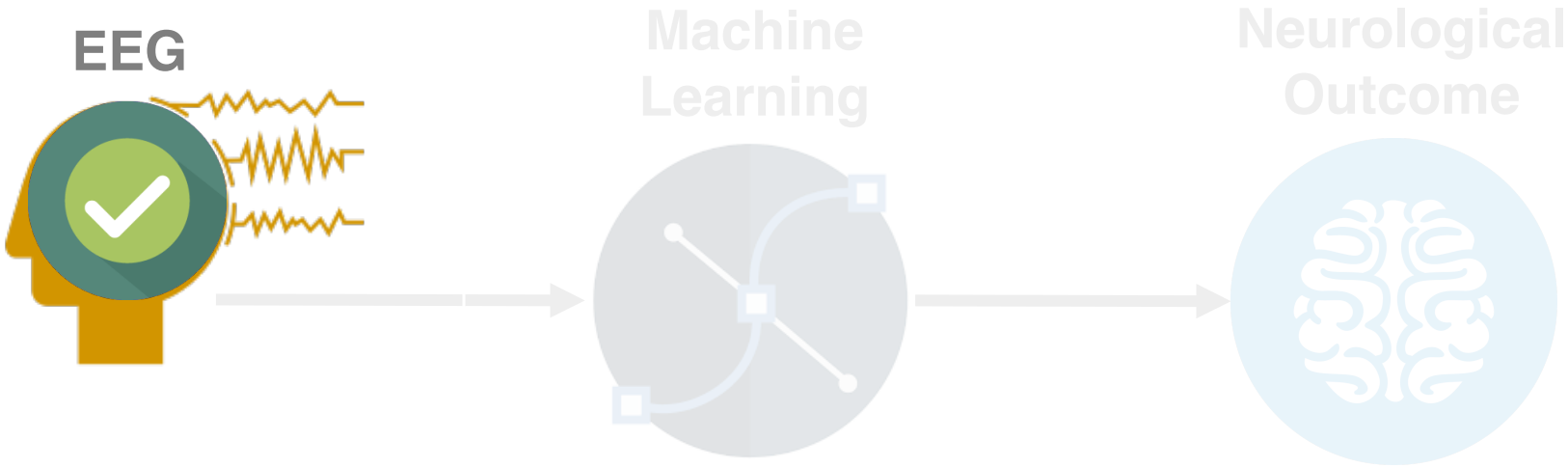
Clinical Feature (Mean)	Good Outcomes	Bad Outcomes
Age (Years)	57	63
Gender (% Male)	68	68
Time to ROSC (Mins)	19.47	25.7
Rhythm at Arrest (%)		
VFib	69	34
Other	25	61
Unknown	*6*	*5*
Cause of Arrest (%)		
Pulmonary	50	34
Anesthesia	3.6	8.5
Neurologic	9.6	13.8
Other/Unknown (%)	*36.8*	*43.7*
Arrest Location		
In Hospital (%)	9	11
Out of Hospital (%)	63	56
Unknown (%)	28	33

Data Collection



1. ‘Bigger’ data sets
2. Time-sensitive
3. Deployable

Data Collection

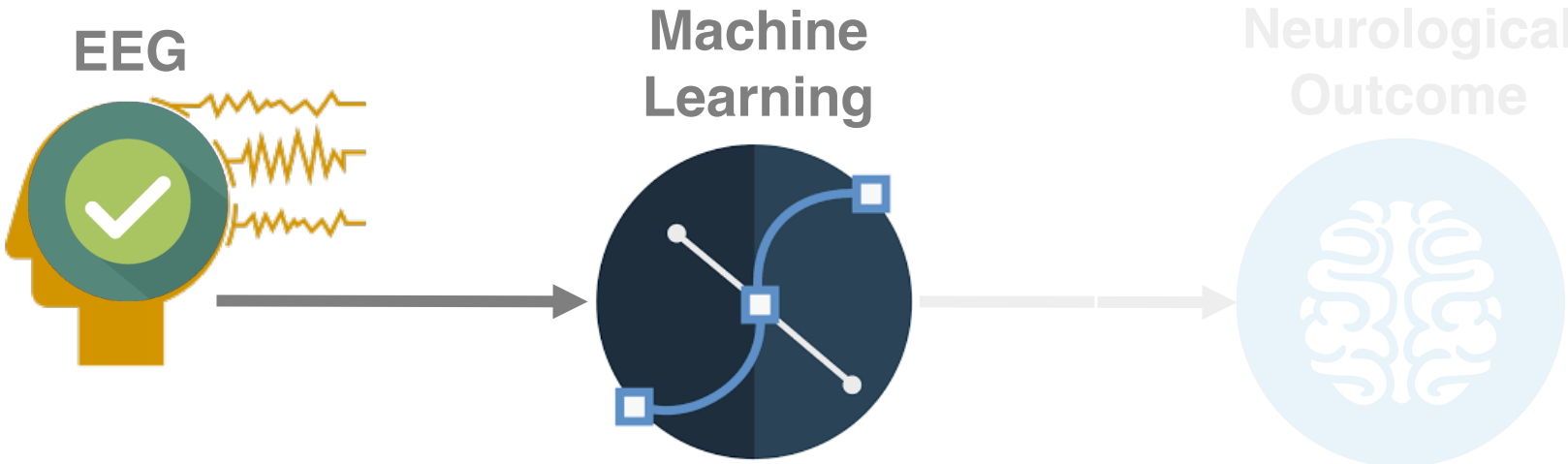


1. Collected an EEG archive 2x larger than largest set previously described in the literature

2. Time-sensitive

3. Deployable

Develop *time-sensitive* modeling approaches



1. Collected an EEG archive 2x larger than largest set previously described in the literature

2. Time-sensitive

3. Deployable

Three Approaches

Two Feature Based:

- 1. Sequential Logistic Regression with ‘Elastic’ memories.

- 2. Sequential Logistic Regression with Dynamic Bayesian network constraints

One Data Driven

- 3. Deep neural networks

Three Approaches

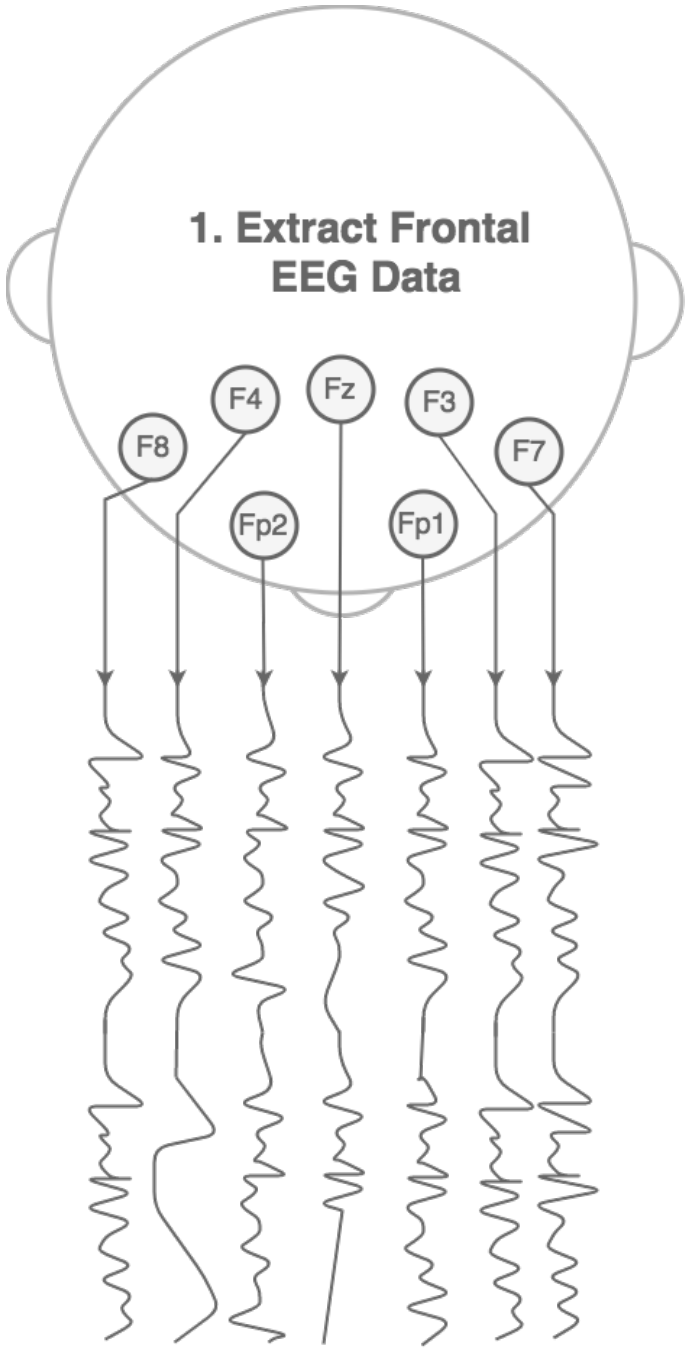
Two Feature Based:

- 1. Sequential Logistic Regression with ‘Elastic’ memories.

- 2. Sequential Logistic Regression with Dynamic Bayesian network constraints

One Data Driven

- 3. Deep neural networks



2. Identify artifacts in five second epochs

2.1 Disconnects and saturation

2.2 Eye and muscle artifact

2.3 Moment-based artifacts

3. Choose five minute epochs with minimal artifact

3.1 Generate artifact score

3.2 Identify cleanest epochs in each hour of data

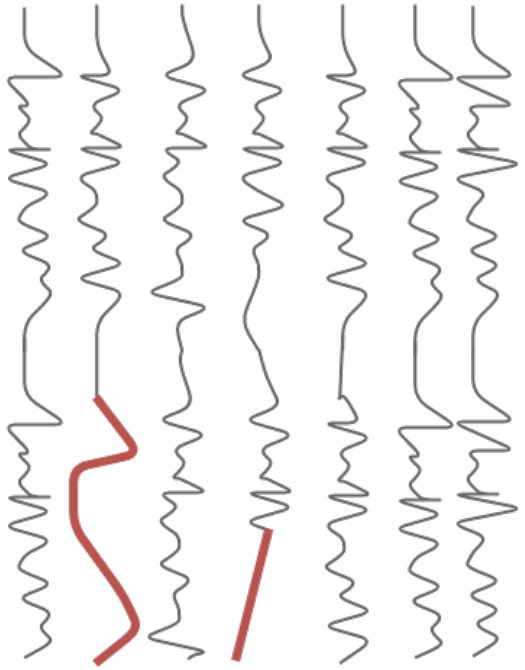
1. Extract Frontal EEG Data

2. Identify artifacts in five second epochs

2.1 Disconnects and saturation

2.2 Eye and muscle artifact

2.3 Moment-based artifacts



3. Choose five minute epochs with minimal artifact

3.1 Generate artifact score

3.2 Identify cleanest epochs in each hour of data

1. Extract Frontal EEG Data

2. Identify artifacts in five second epochs

2.1 Disconnects and saturation

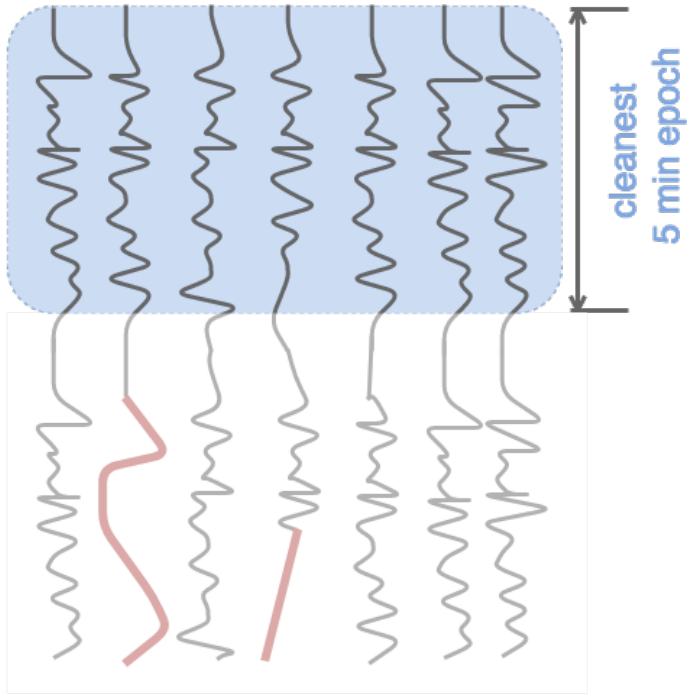
2.2 Eye and muscle artifact

2.3 Moment-based artifacts

3. Choose five minute epochs with minimal artifact

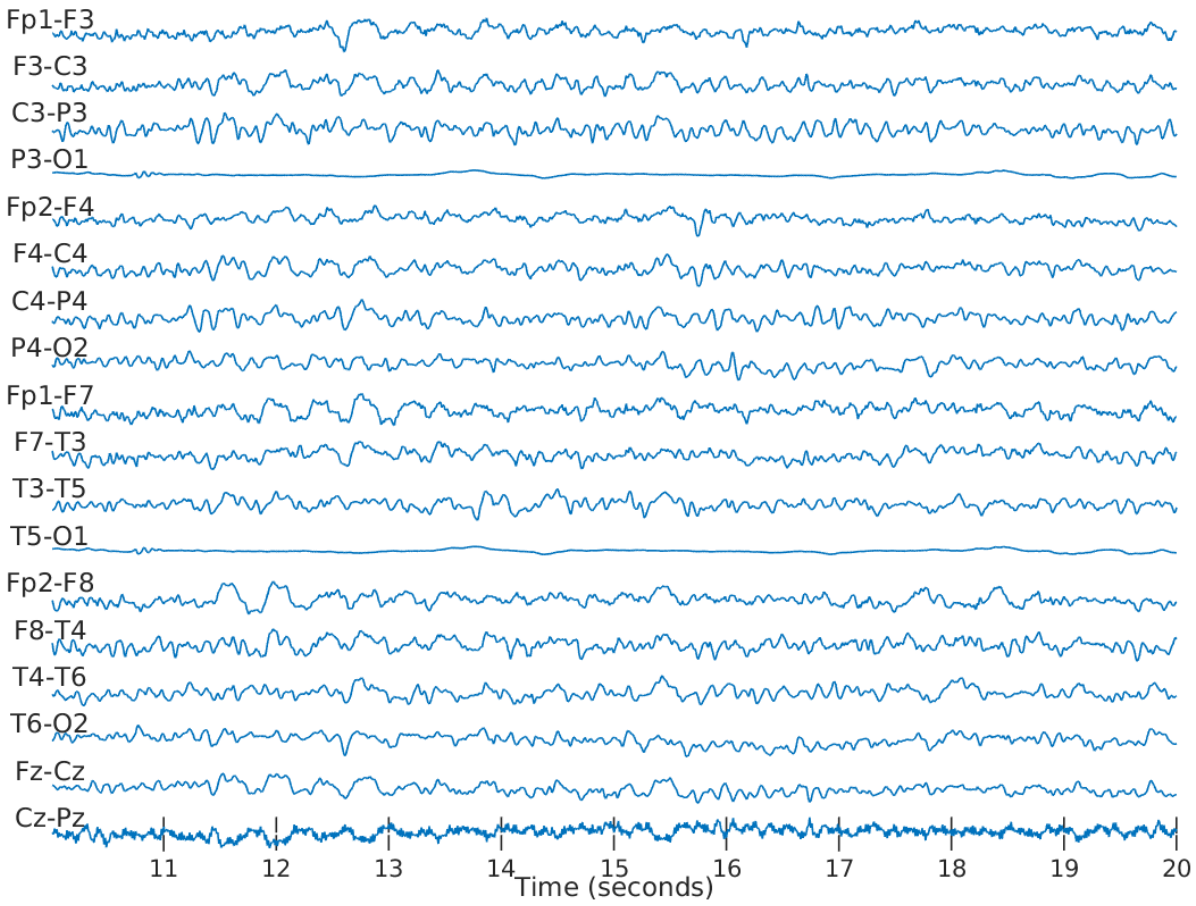
3.1 Generate artifact score

3.2 Identify cleanest epochs in each hour of data



57 features to describe three signal qualities

- Complexity (21 features)
- Category (31 features)
- Connectivity (7 features)



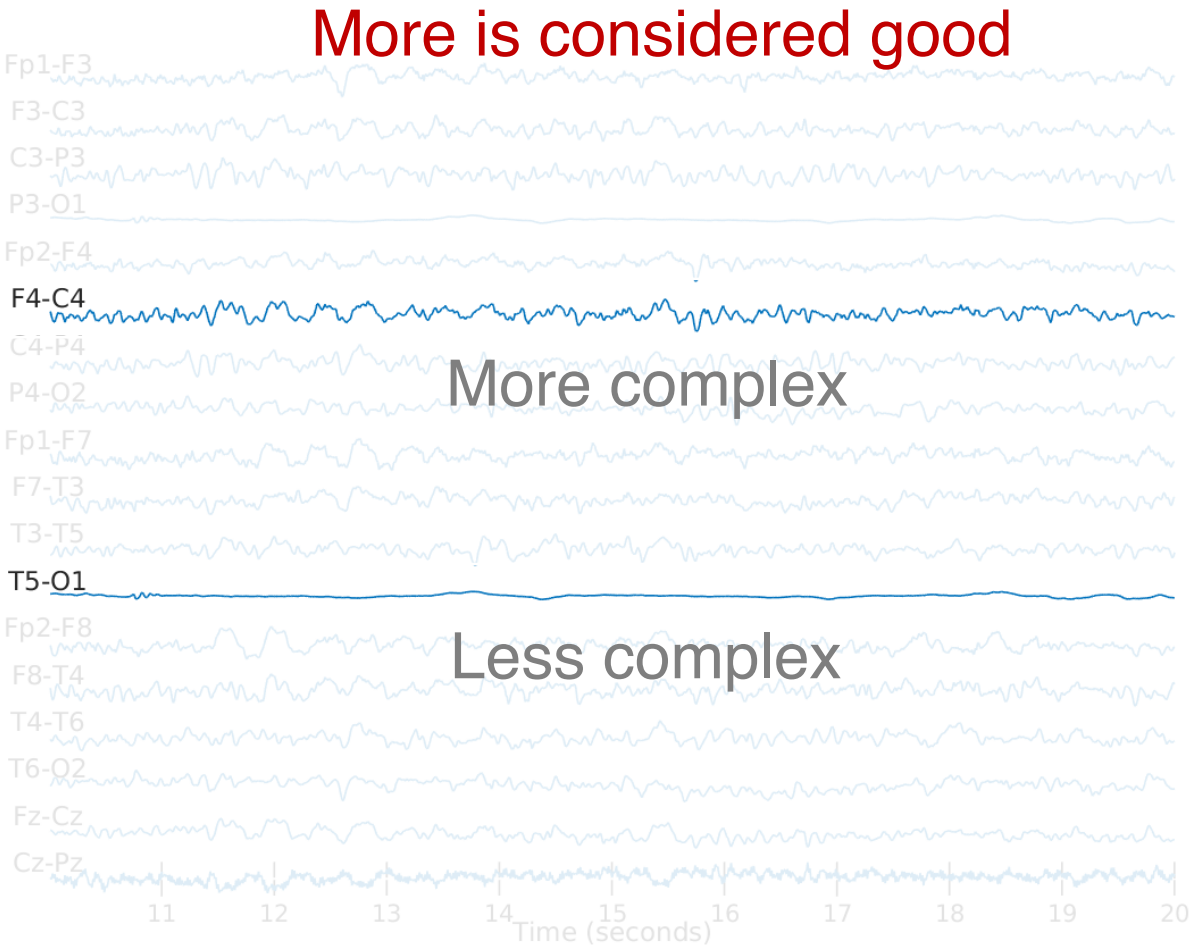
57 features to describe three signal qualities

- Complexity (21 features)

  e.g. Shannon Entropy

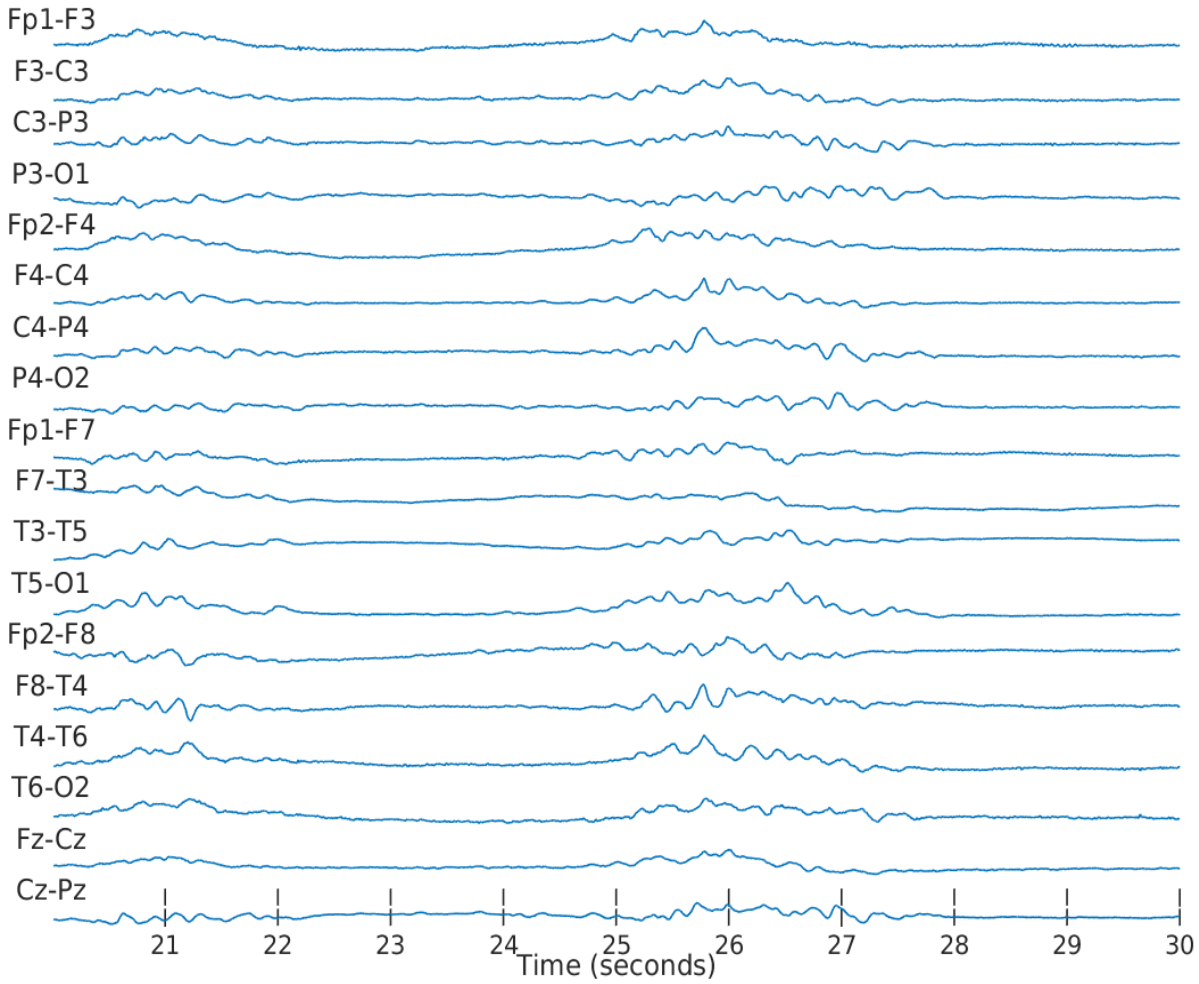
- Category (31 features)

- Connectivity (7 features)



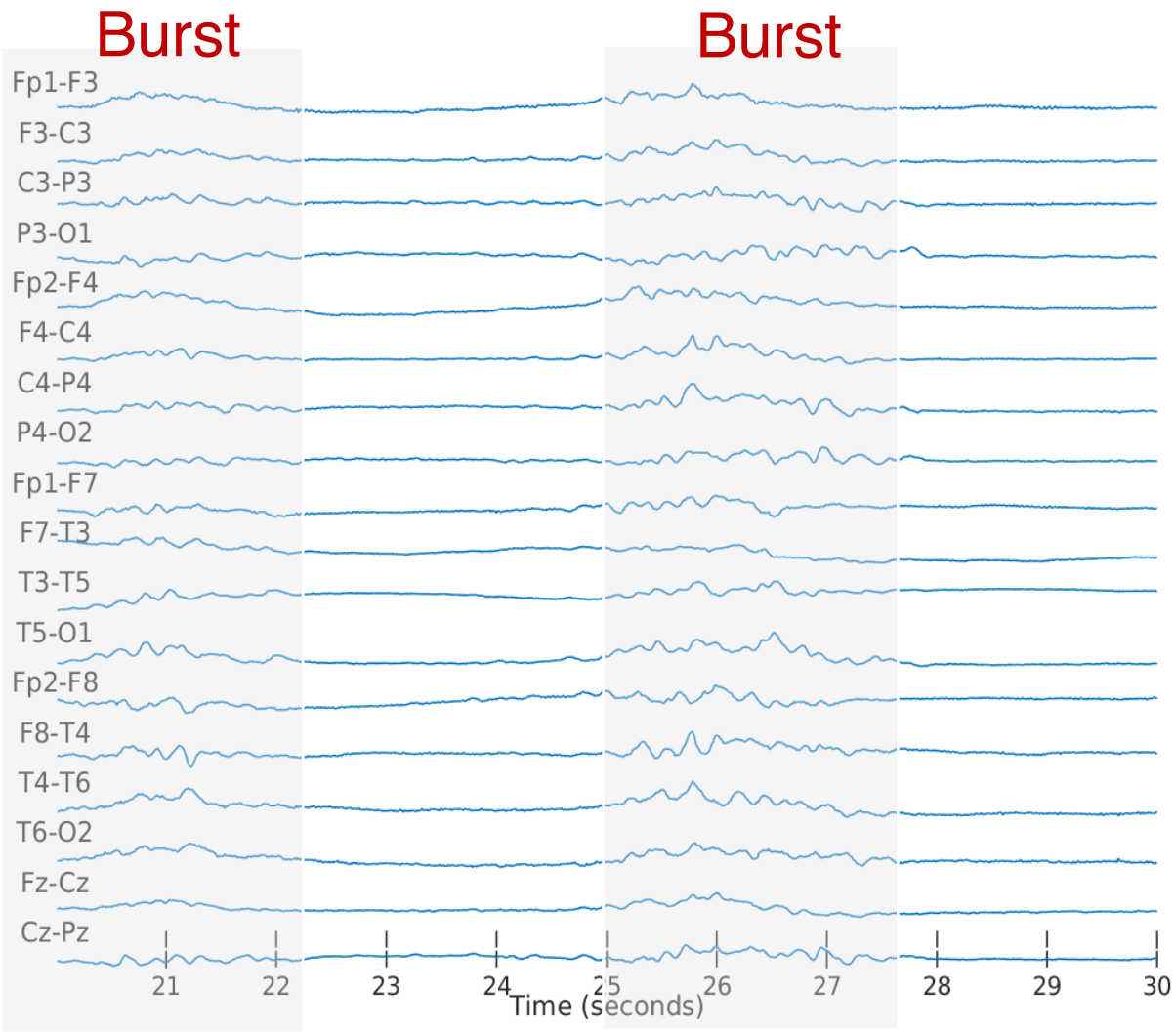
57 features to describe three signal qualities

- Complexity (21 features)
- Category (31 features)
- Connectivity (7 features)



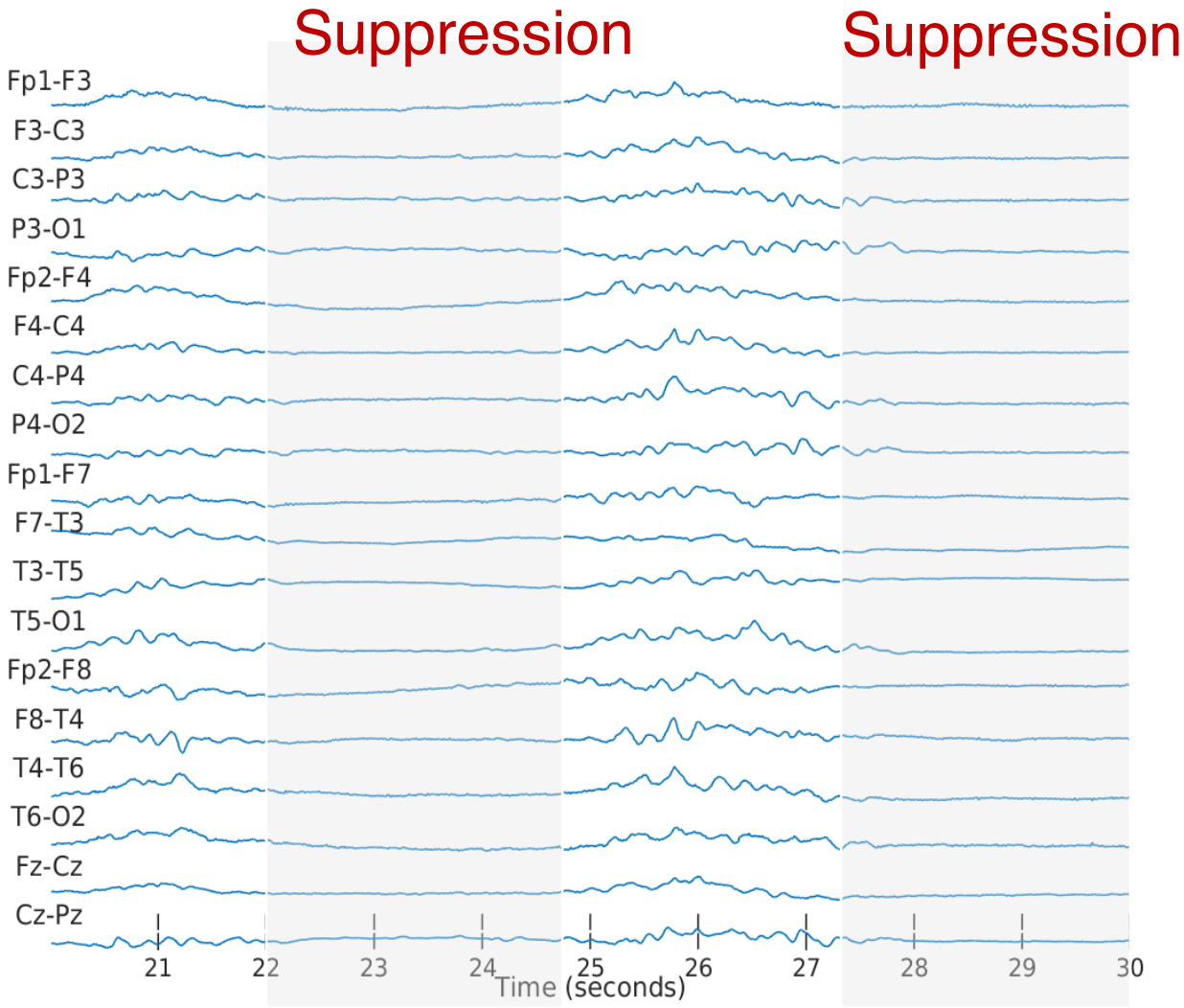
57 features to describe three signal qualities

- Complexity (21 features)
- Category (31 features)
  e.g. Burst Suppression
- Connectivity (7 features)



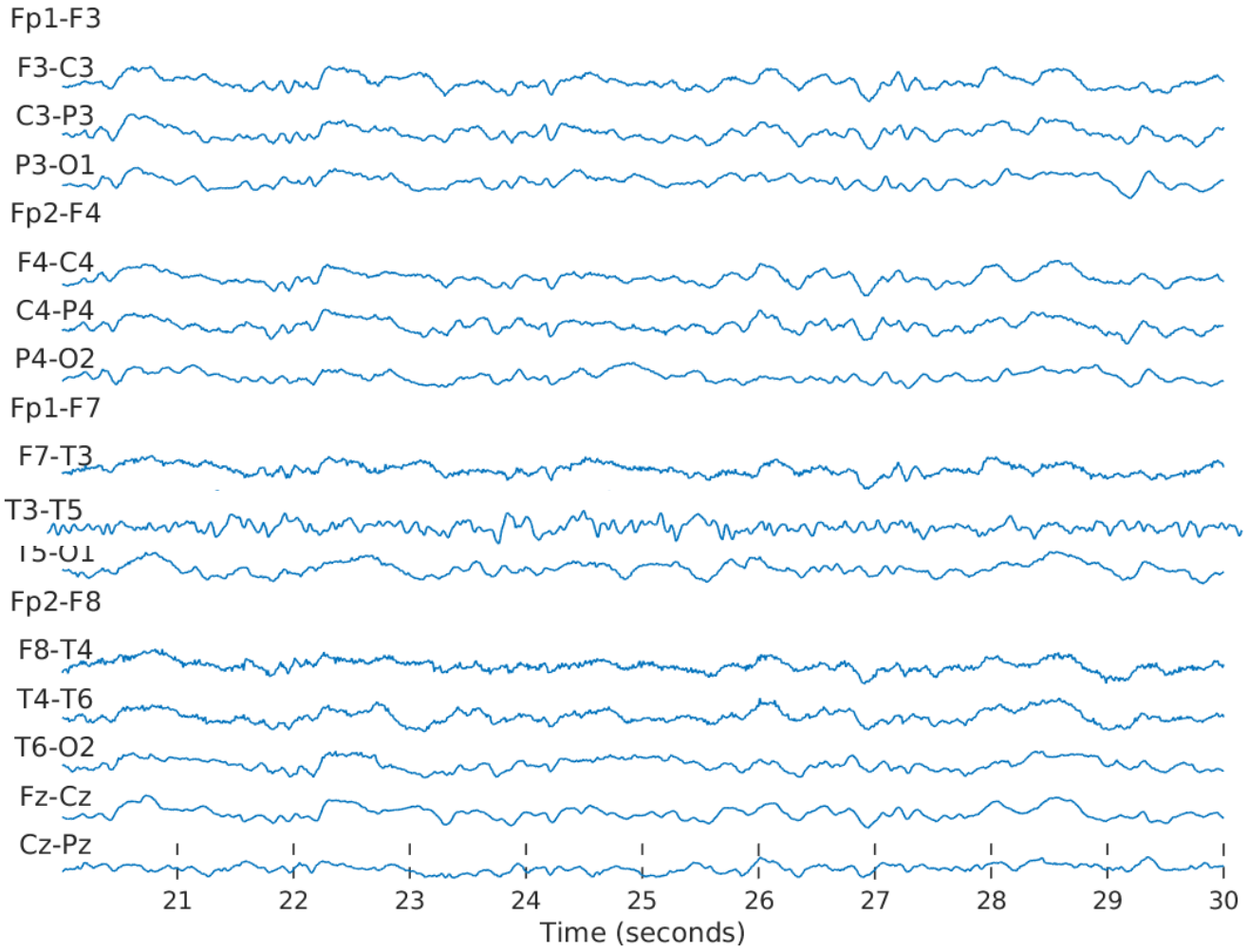
57 features to describe three signal qualities

- Complexity (21 features)
- Category (31 features)
  e.g. Burst Suppression
- Connectivity (7 features)



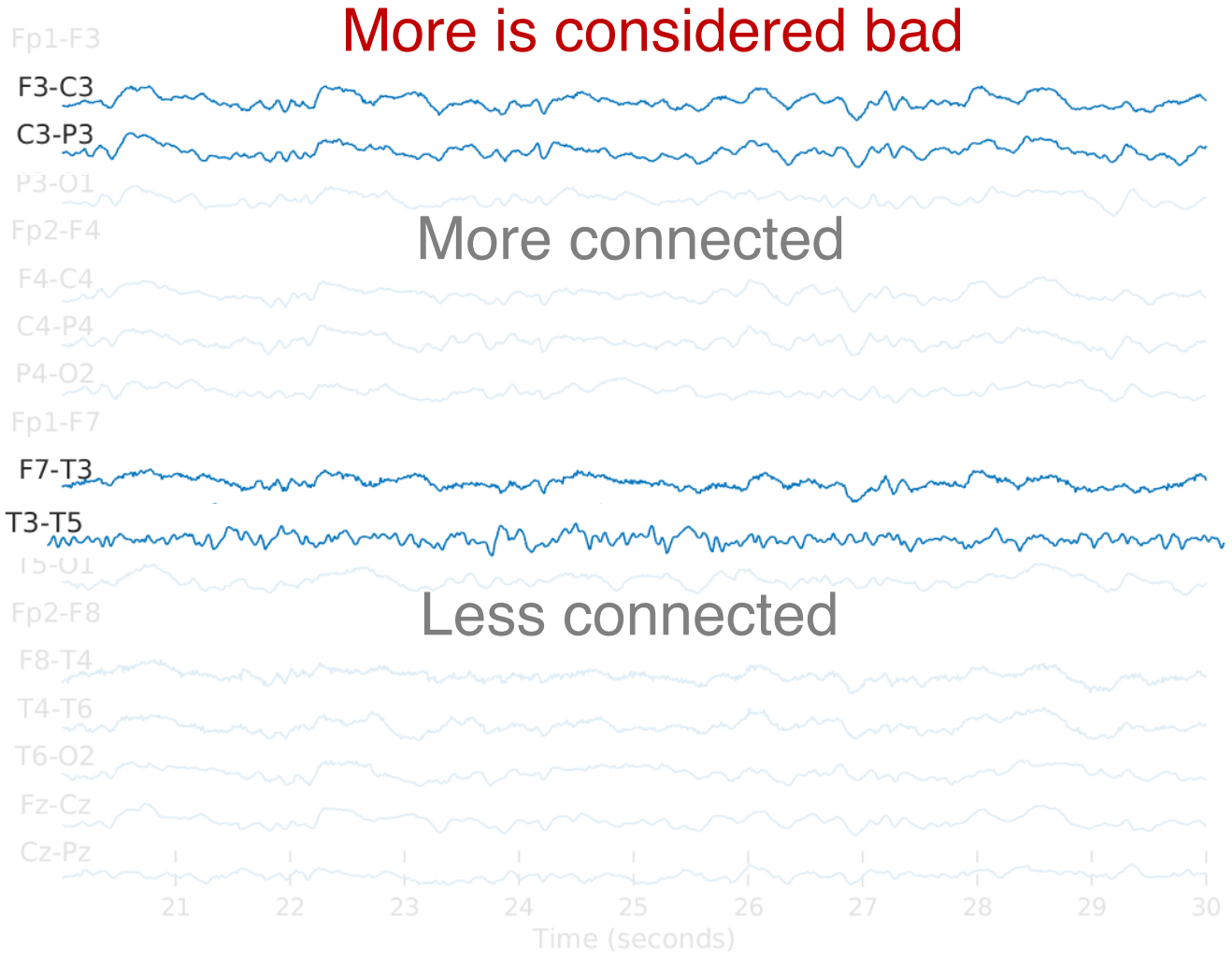
57 features to describe three signal qualities

- Complexity (21 features)
- Category (31 features)
- Connectivity (7 features)



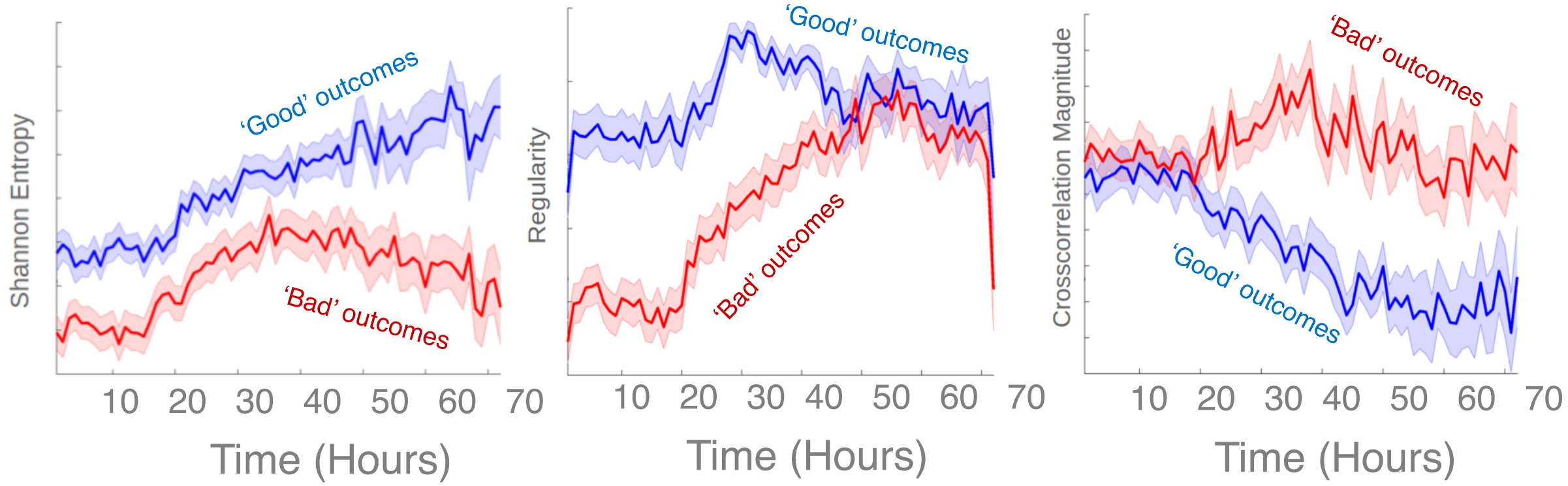
57 features to describe three signal qualities

- Complexity (21 features)
- Category (31 features)
- Connectivity (7 features)
  e.g. cross correlation



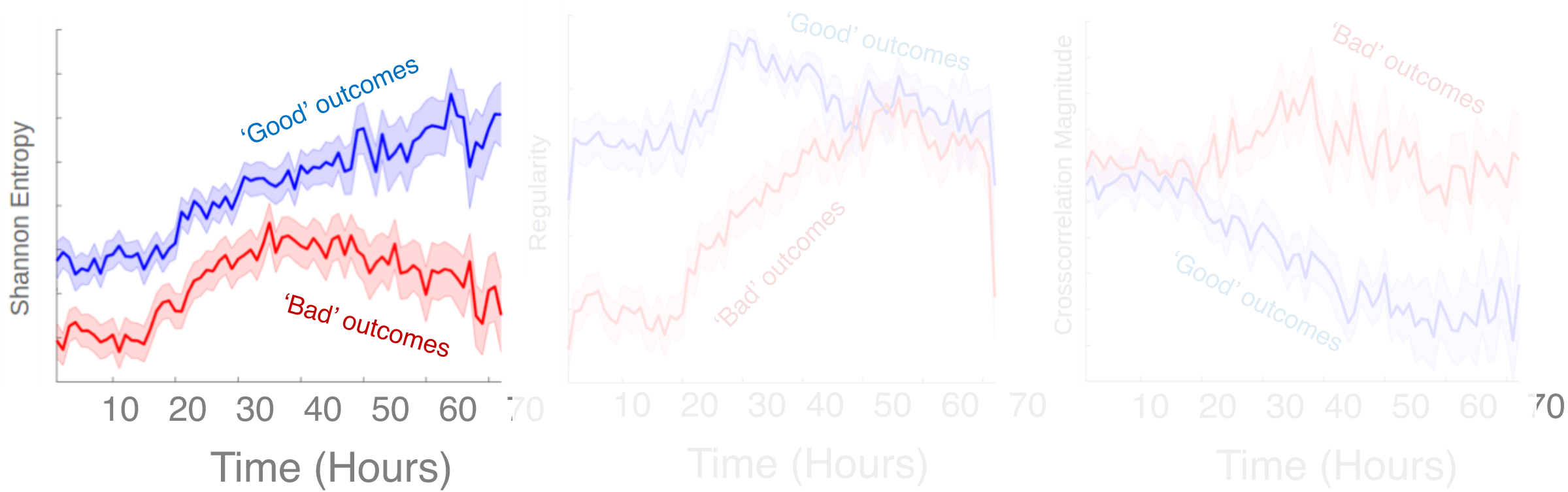
Feature-outcome relationship is time-dependent

The mean and standard error of three features for the study population, partitioned by outcome



Feature-outcome relationship is time-dependent

Shannon Entropy (a measure of complexity) distinguishes ‘Good’ and ‘Bad’ consistently



Feature-outcome relationship is time-dependent

But the value that optimally discriminates over time, changes.

Feature-outcome relationship is time-dependent

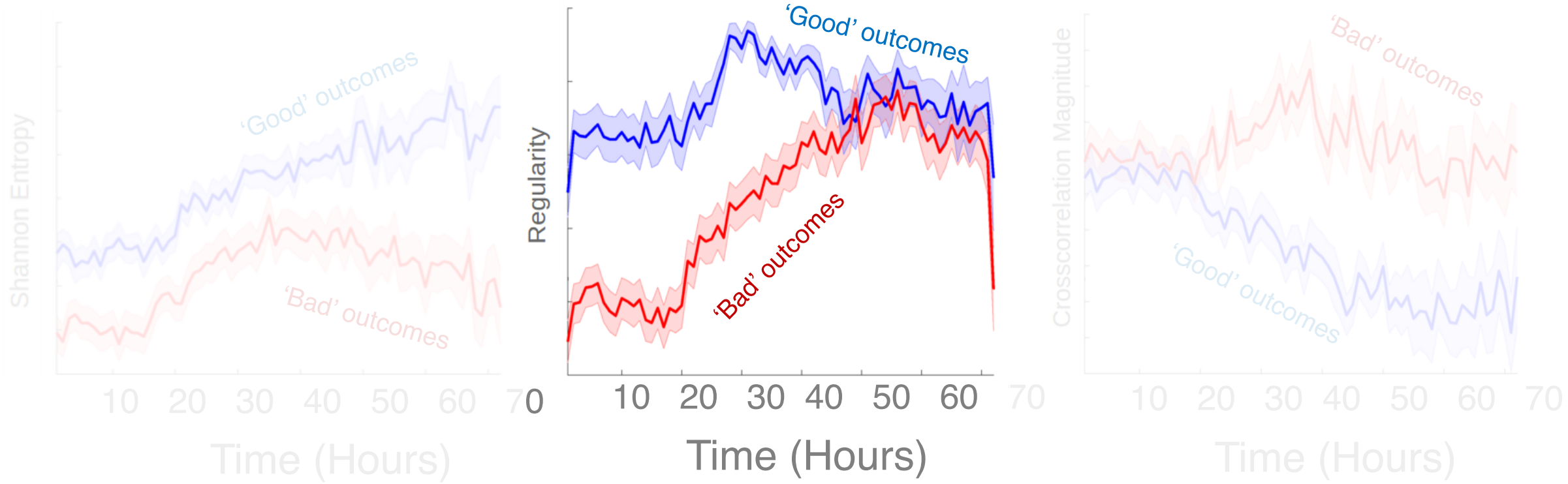
But the value that optimally discriminates over time, changes.

Feature-outcome relationship is time-dependent

But the value that optimally discriminates over time, changes.

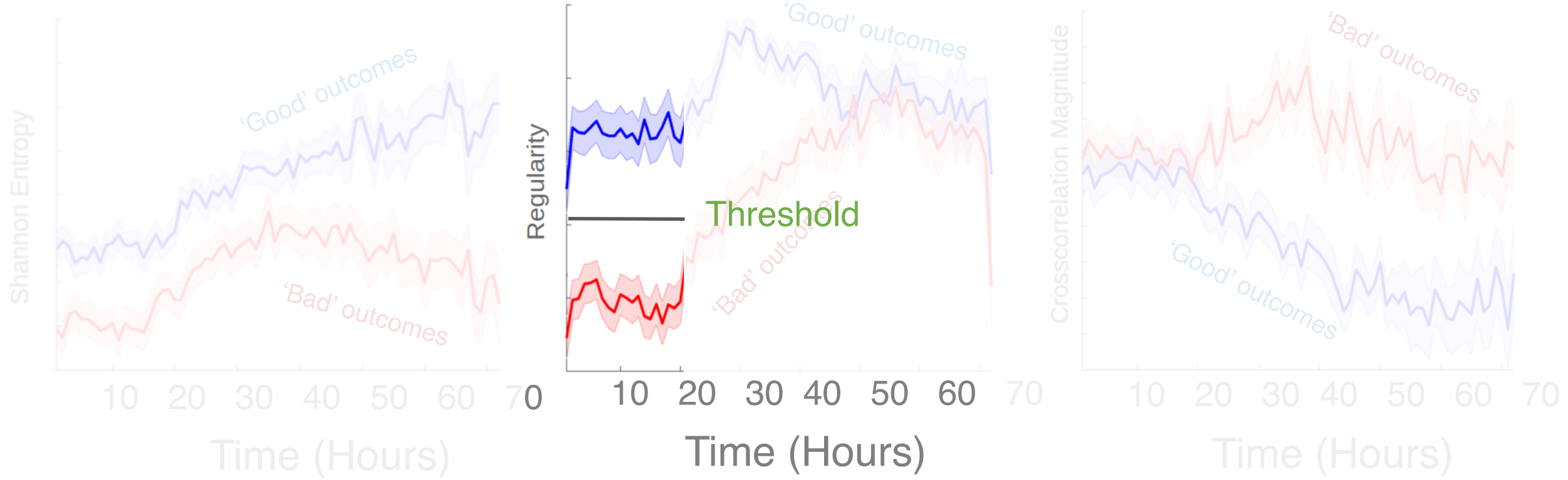
Feature-outcome relationship is time-dependent

Regularity (a measure of burst-suppression) distinguishes ‘Good’ and ‘Bad’ earlier



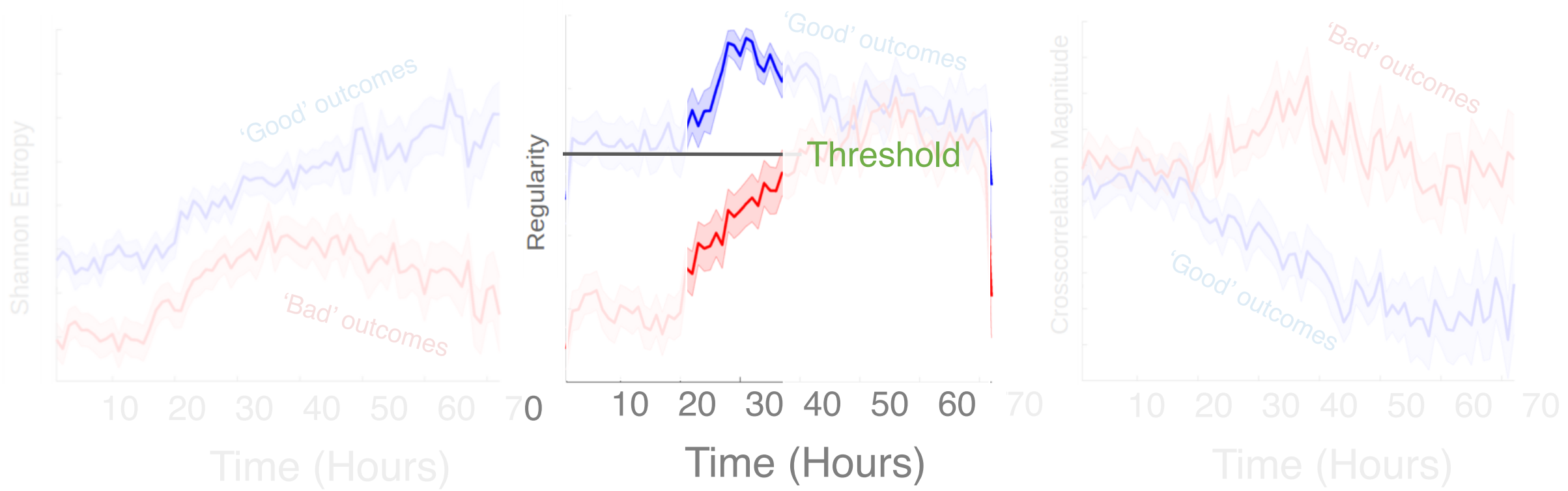
Feature-outcome relationship is time-dependent

But the value that optimally discriminates over time, changes.



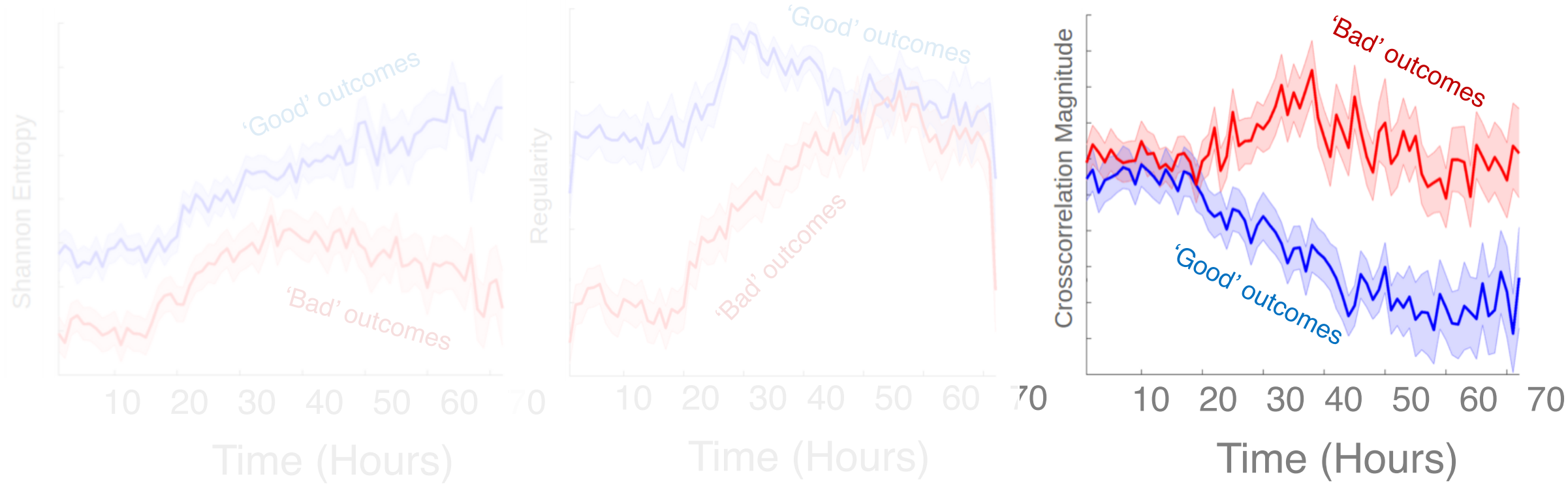
Feature-outcome relationship is time-dependent

But the value that optimally discriminates over time, changes.



Feature-outcome relationship is time-dependent

Cross correlation (a measure of complexity) distinguishes ‘Good’ and ‘Bad’ later



Conclusion

- The features-outcome relationship changes over time
- Needs modelling approach where coefficients evolve over time

Three Approaches

Two Feature Based:

- 1. Sequential Logistic Regression with ‘Elastic’ memories

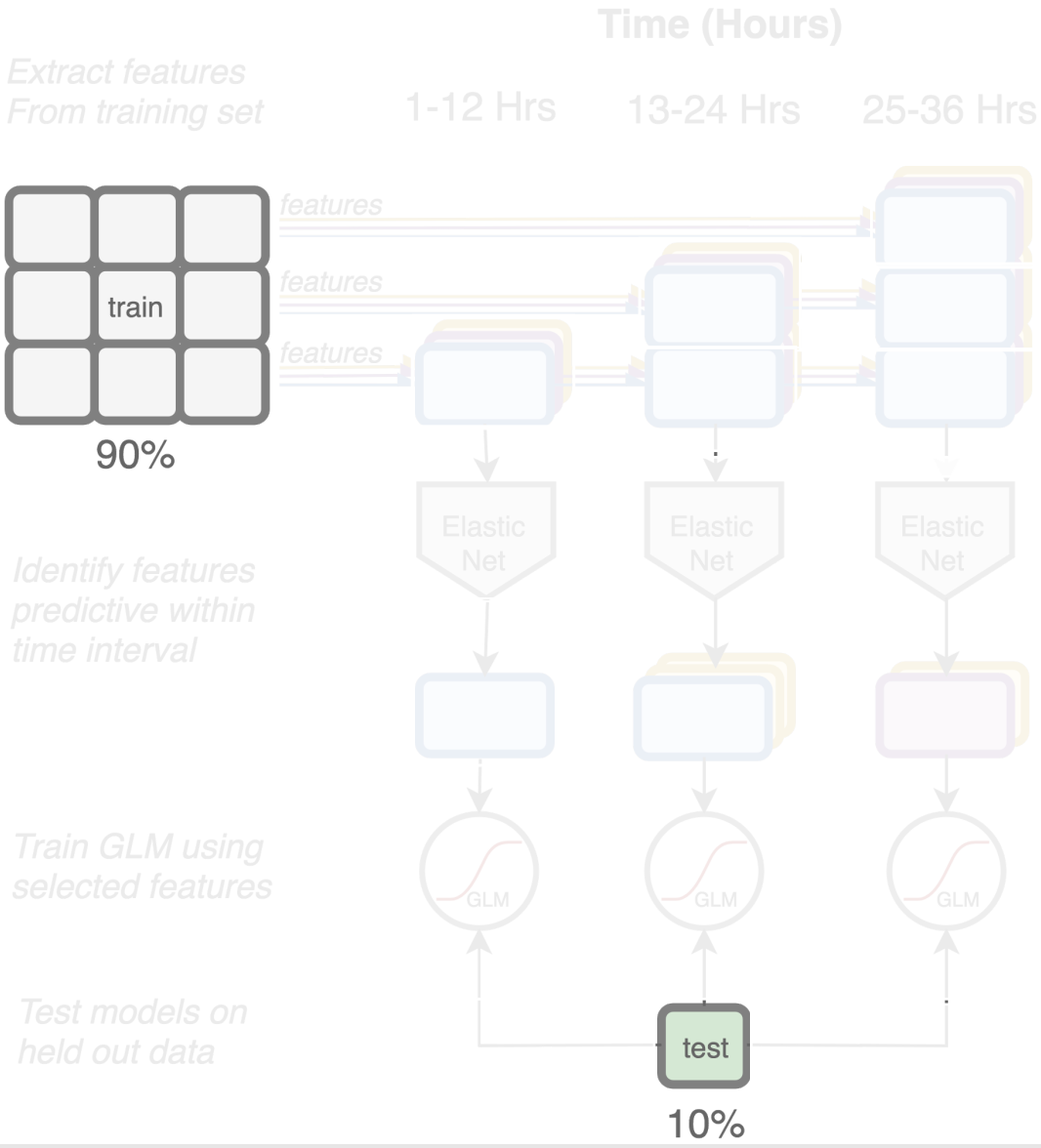
- 2. Sequential Logistic Regression with Dynamic Bayesian network constraints

One Data Driven

- 3. Deep neural networks

Logistic Regression with *Elastic Memories*

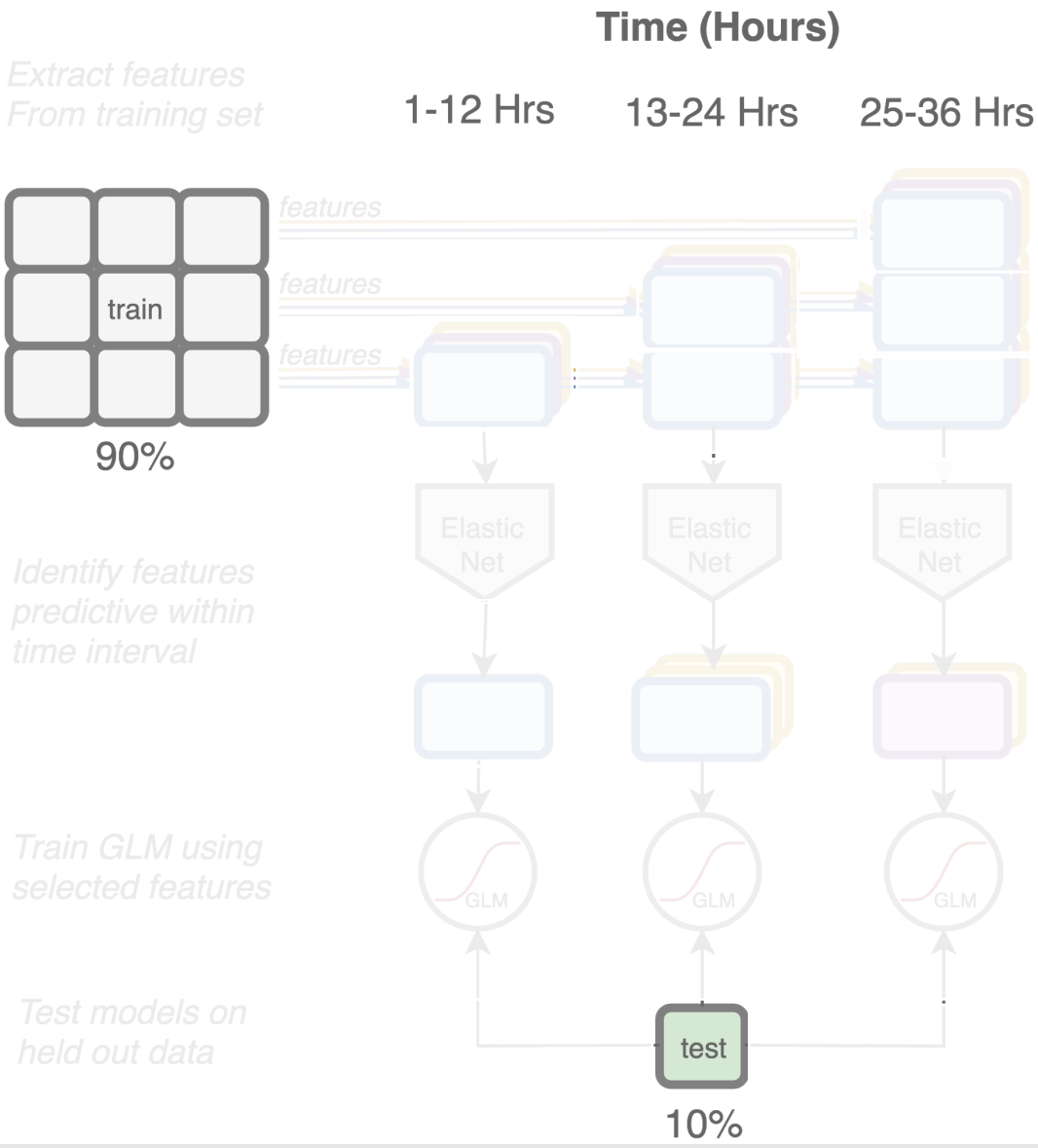
- n = 438 patients
- We used 52/57 features
- Validation: 10-fold



Logistic Regression with *Elastic Memories*

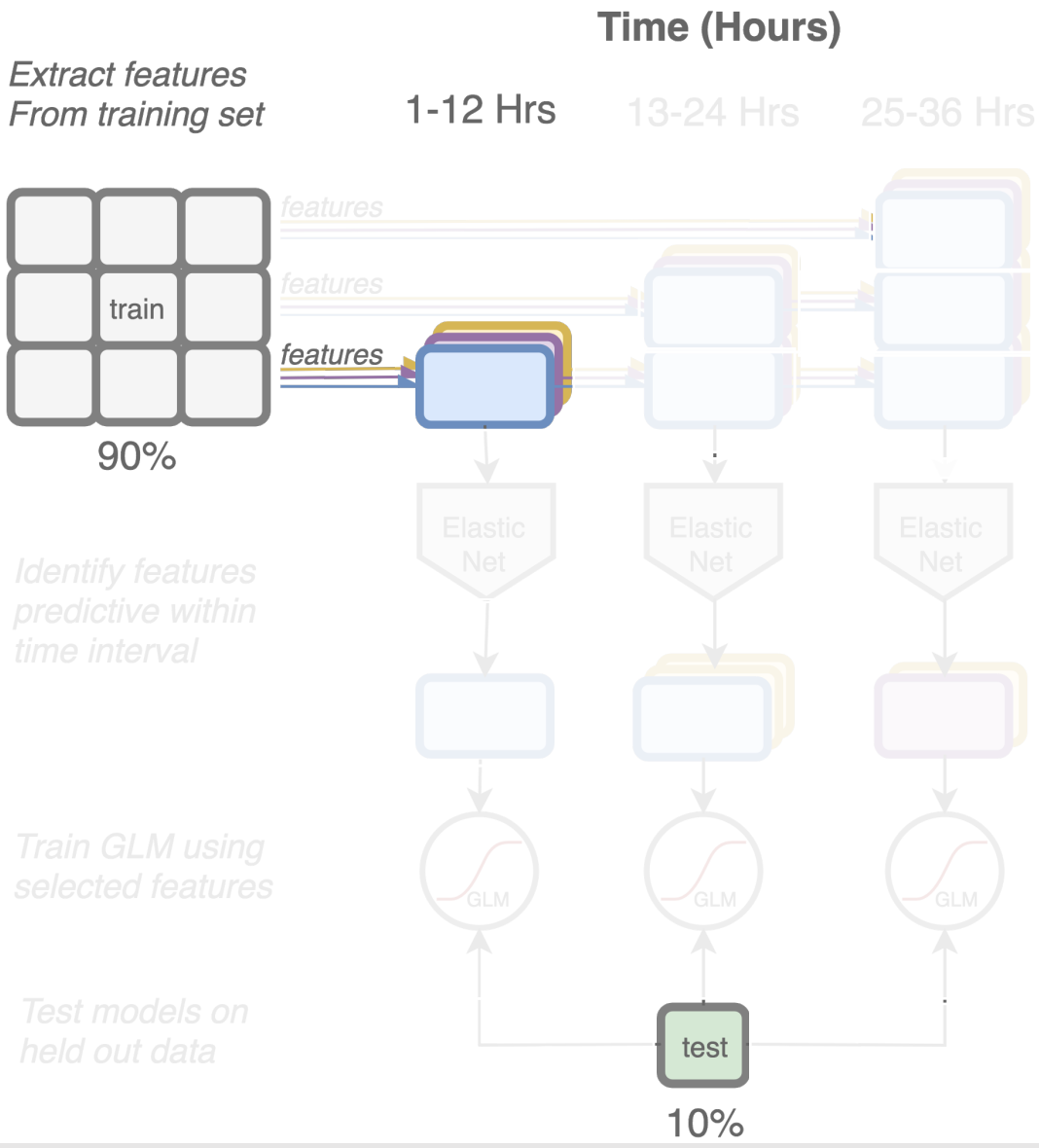
- We train a series of models that classify patient outcomes, in particular time intervals:

  1-12 hours
  13-24 hours
  25-36 hours etc.



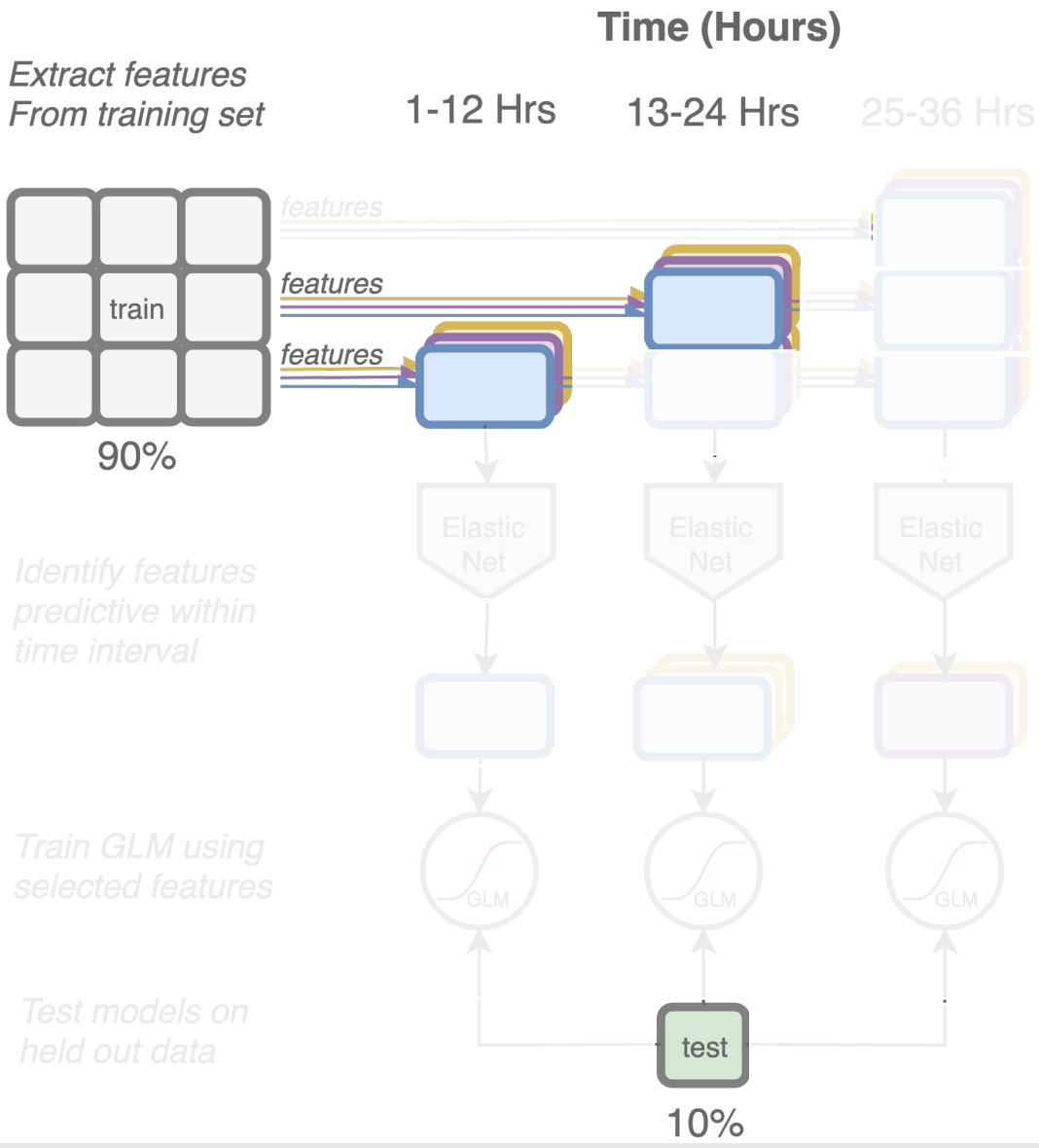
Logistic Regression with *Elastic Memories*

- Features are extracted at particular time intervals
- 1-12 hours



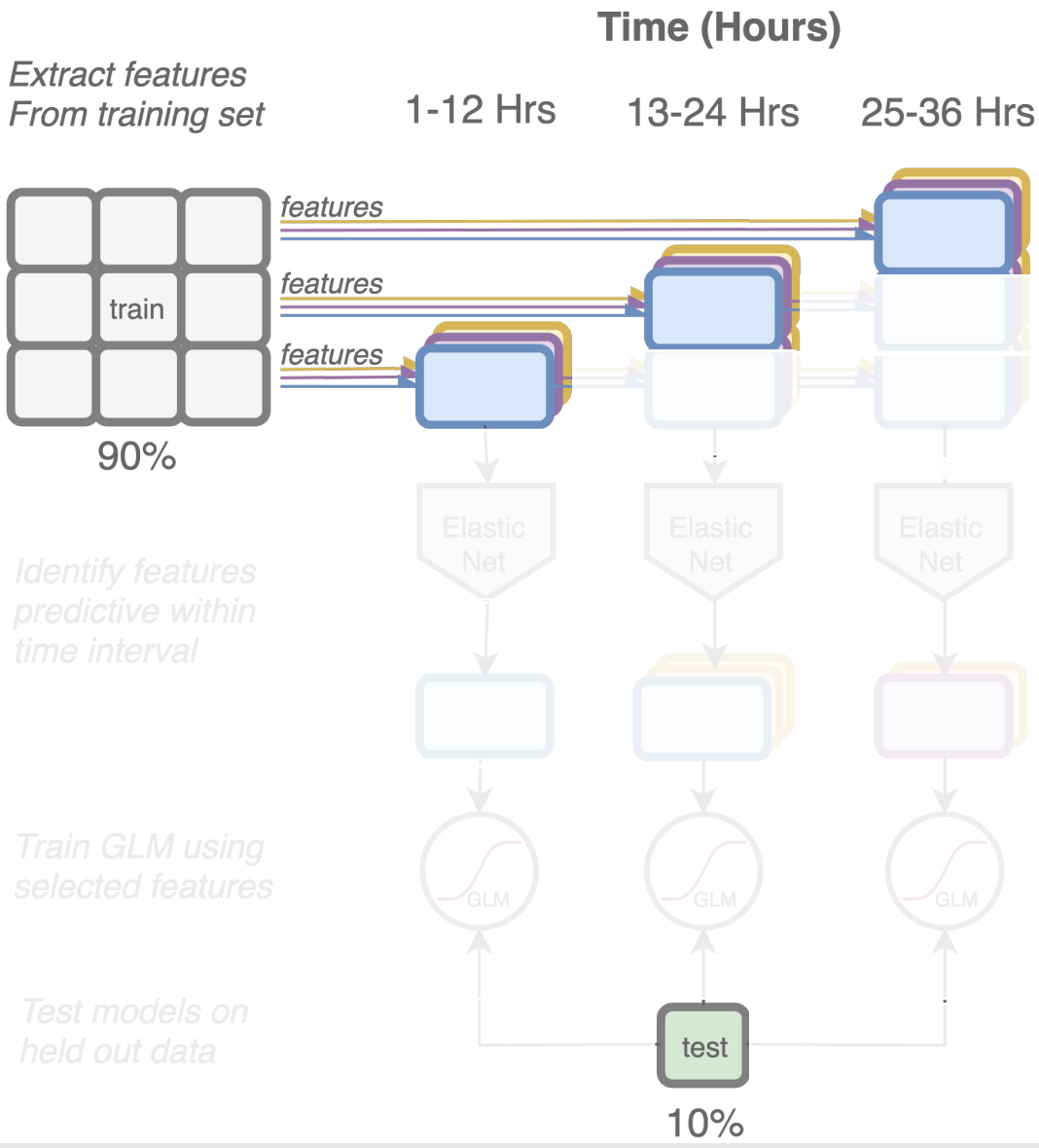
Logistic Regression with *Elastic Memories*

- Features are extracted at particular time intervals
- 13- 24 hours



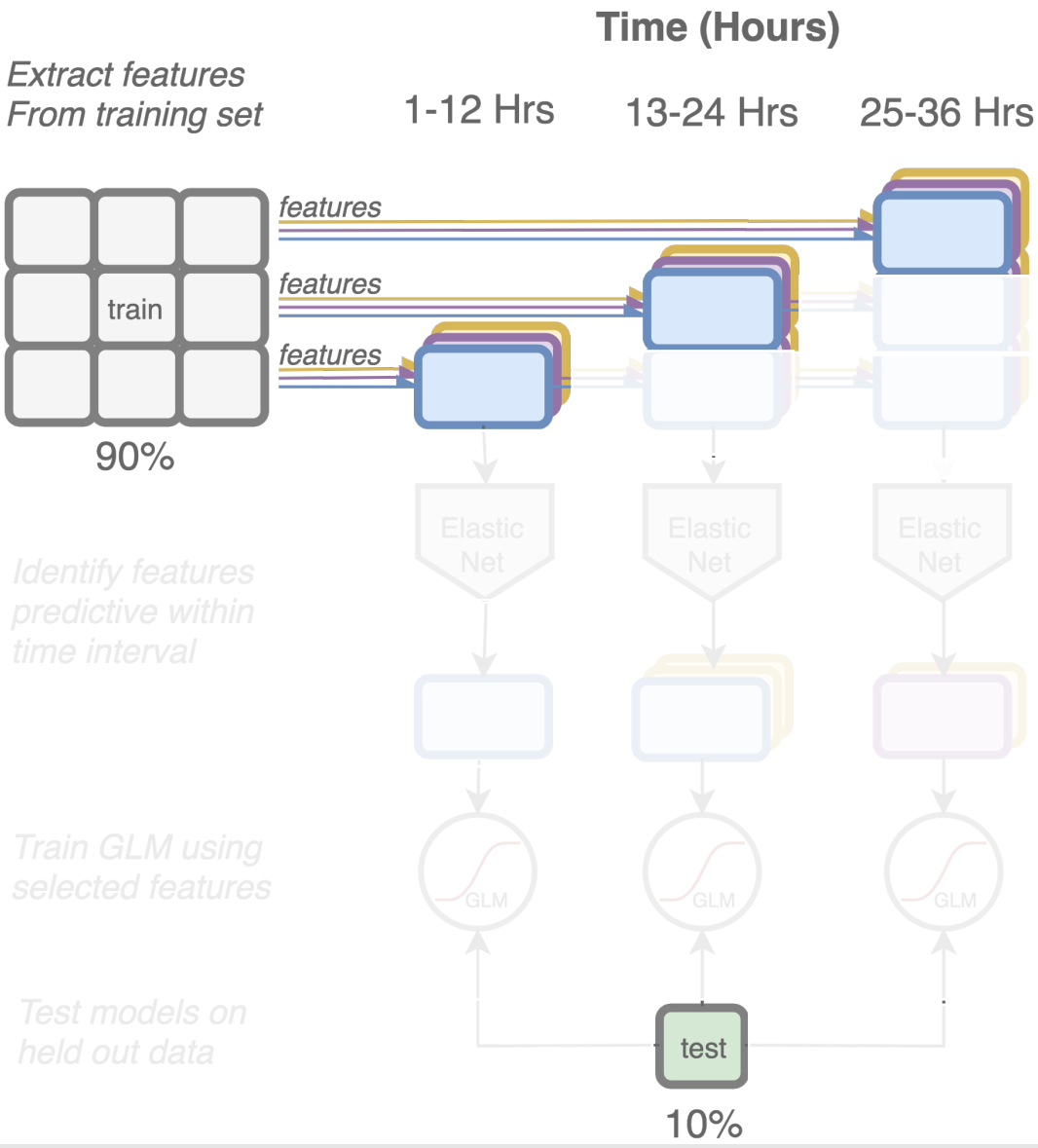
Logistic Regression with *Elastic Memories*

- Features are extracted at particular time intervals
- 25-36 hours, and so on…



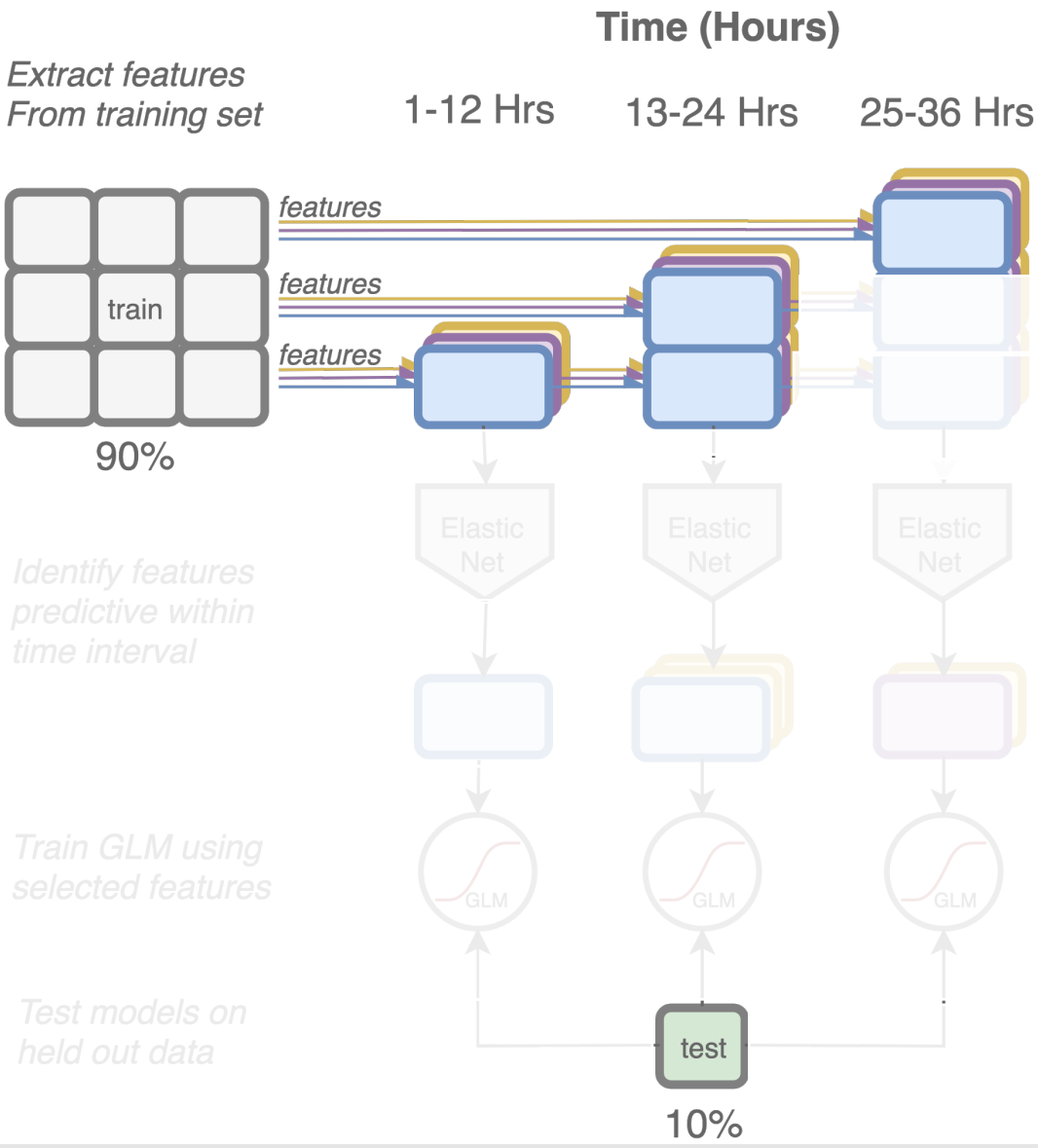
Logistic Regression with *Elastic Memories*

- Features used by models in earlier time intervals are passed forward as ‘memories’ for models in future time intervals



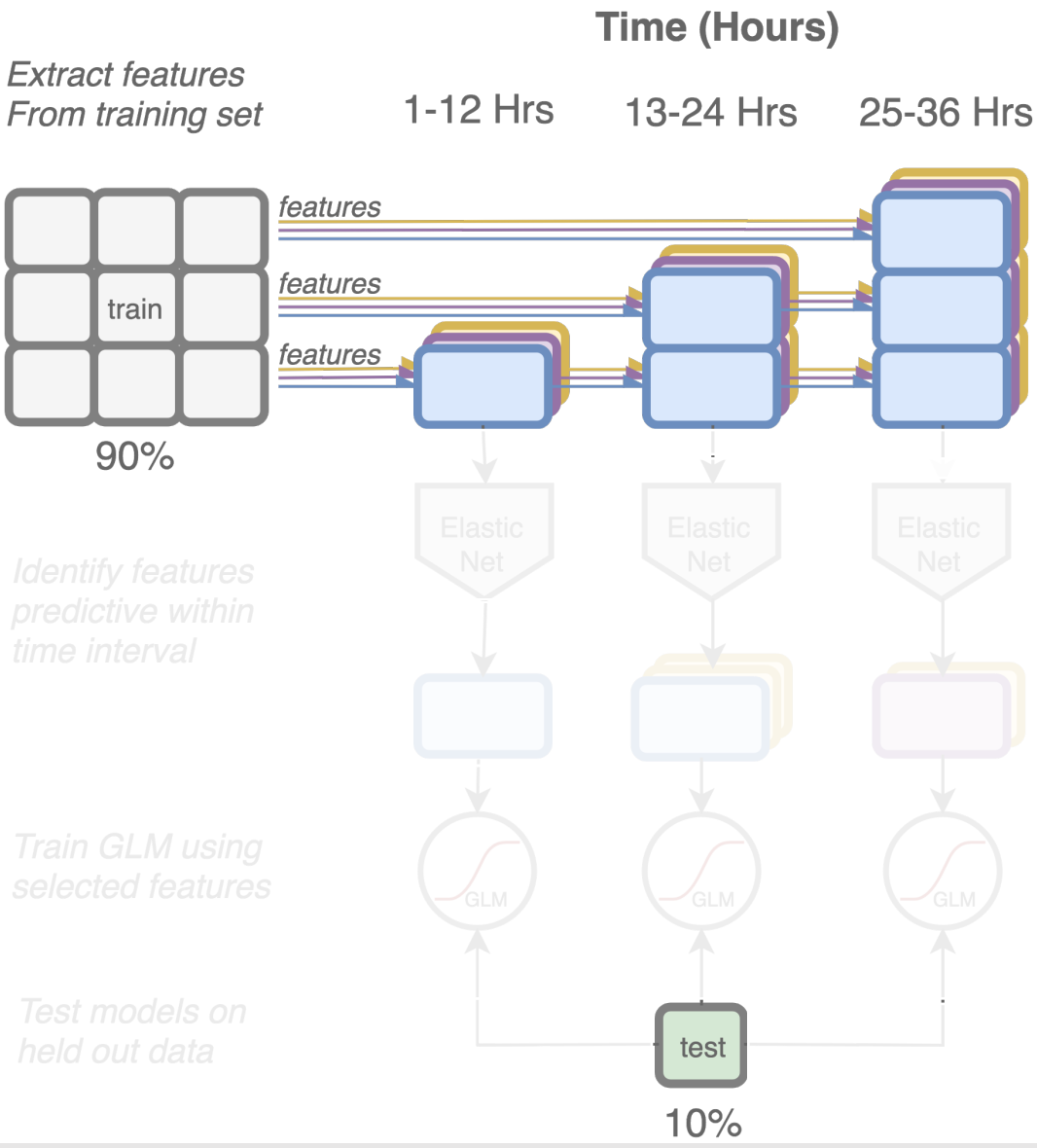
Logistic Regression with *Elastic Memories*

- Features used by models in earlier time intervals are passed forward as ‘memories’ for models in future time intervals



Logistic Regression with *Elastic Memories*

- Features used by models in earlier time intervals are passed forward as ‘memories’ for models in future time intervals

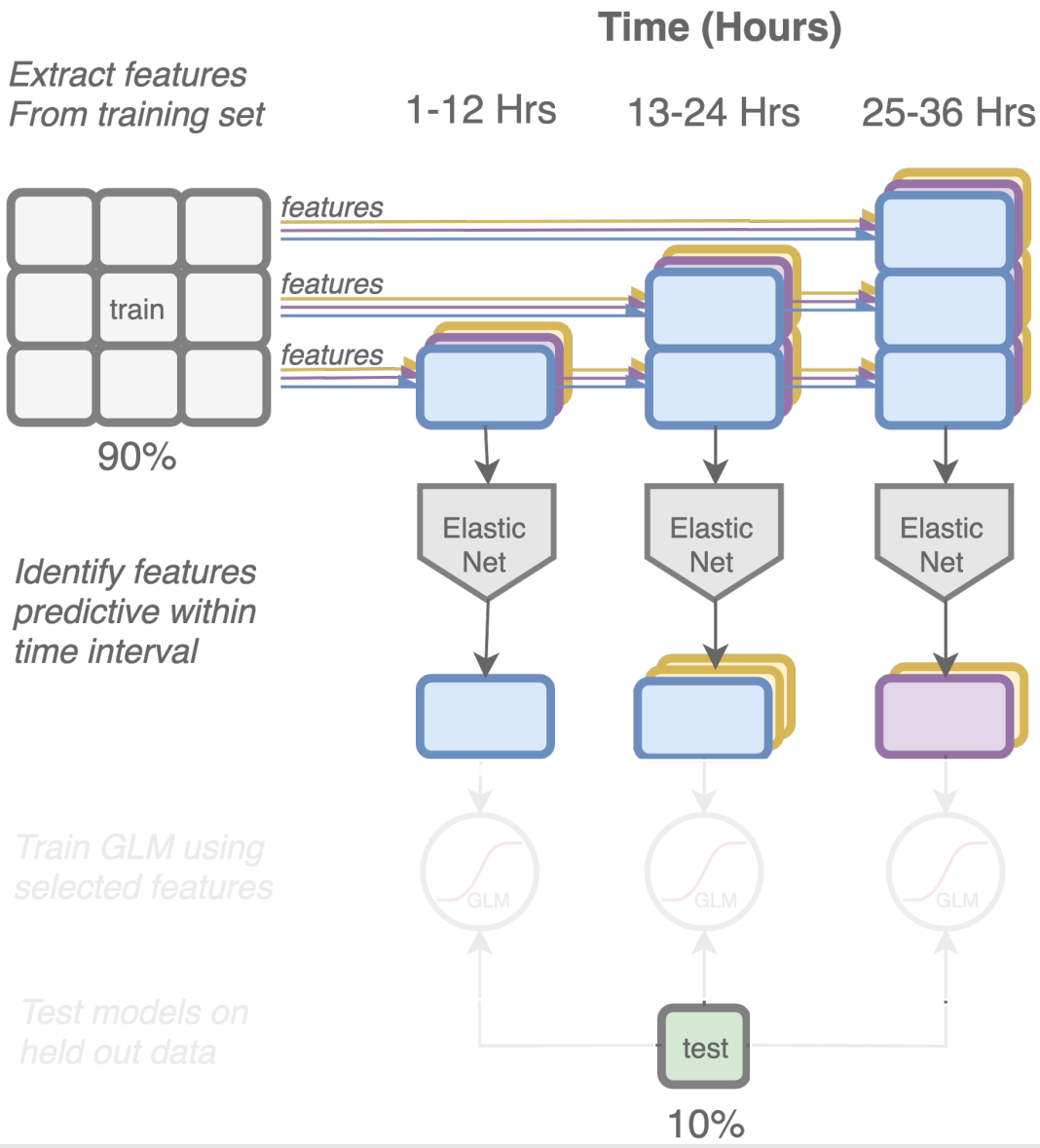


Logistic Regression with *Elastic Memories*

- We retain only the most important features using Elastic Net

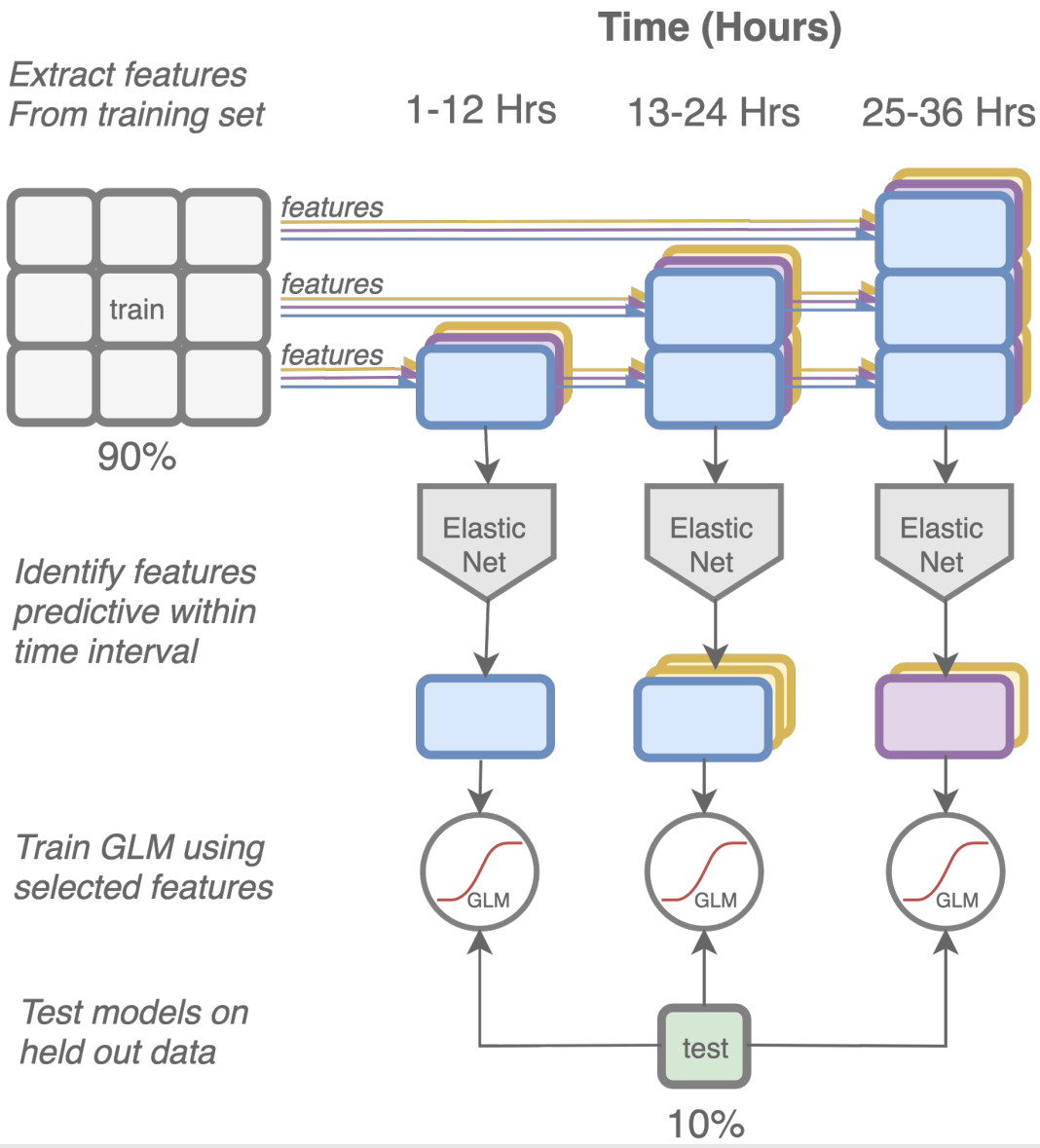
- Penalizes the size of the regression coefficients based on both their l^1 norm and their l^2 norm :

$$argmax_{\beta} \sum_{i} \log L(y_i, ; \beta, x_i) - \lambda[\alpha||\beta||_1 + \frac{1}{2}(1-\alpha)||\beta||_2^2]$$



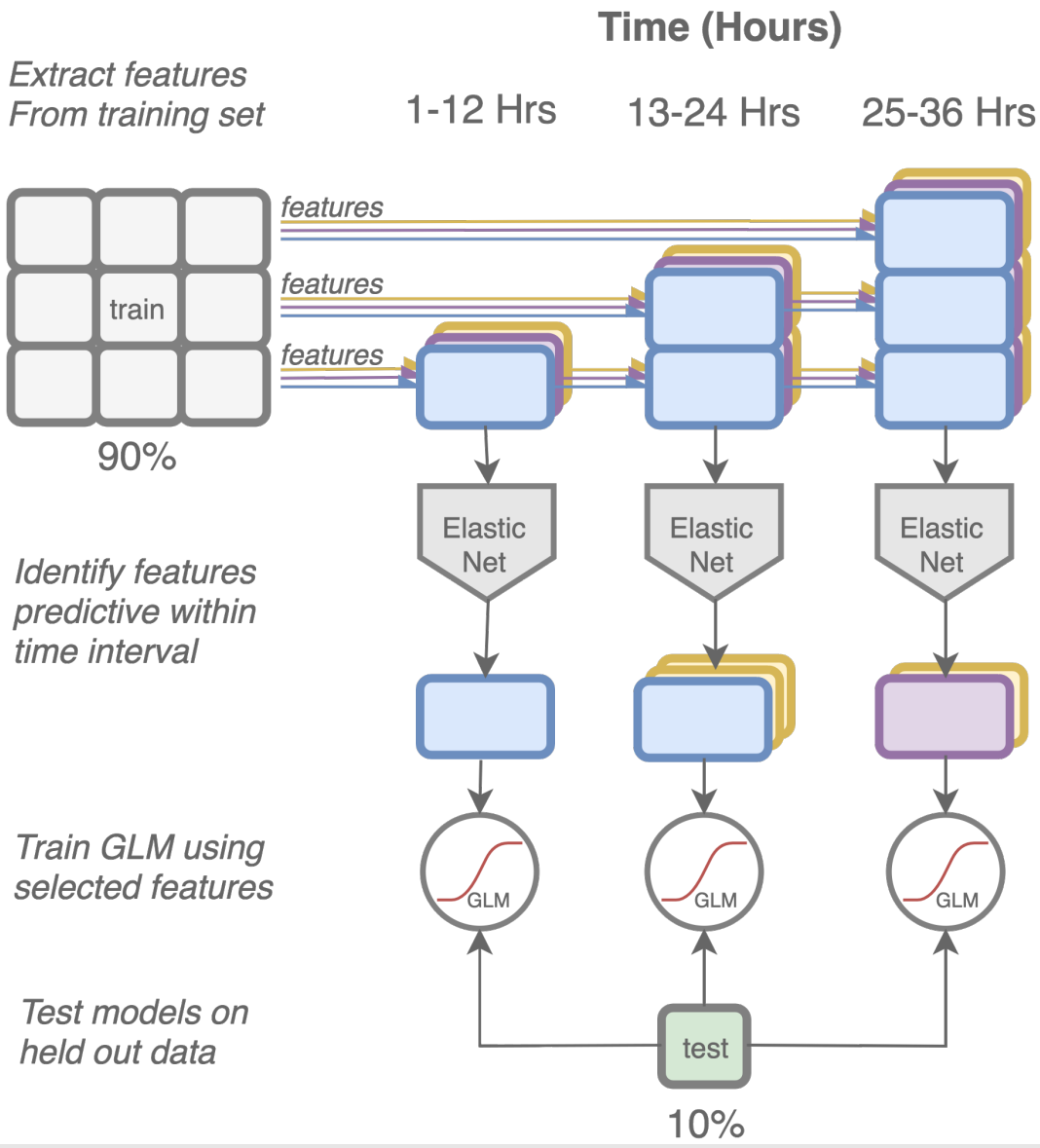
Logistic Regression with *Elastic Memories*

- A logistic regression model with the selected features is used to evaluate performance on the held out test-sets



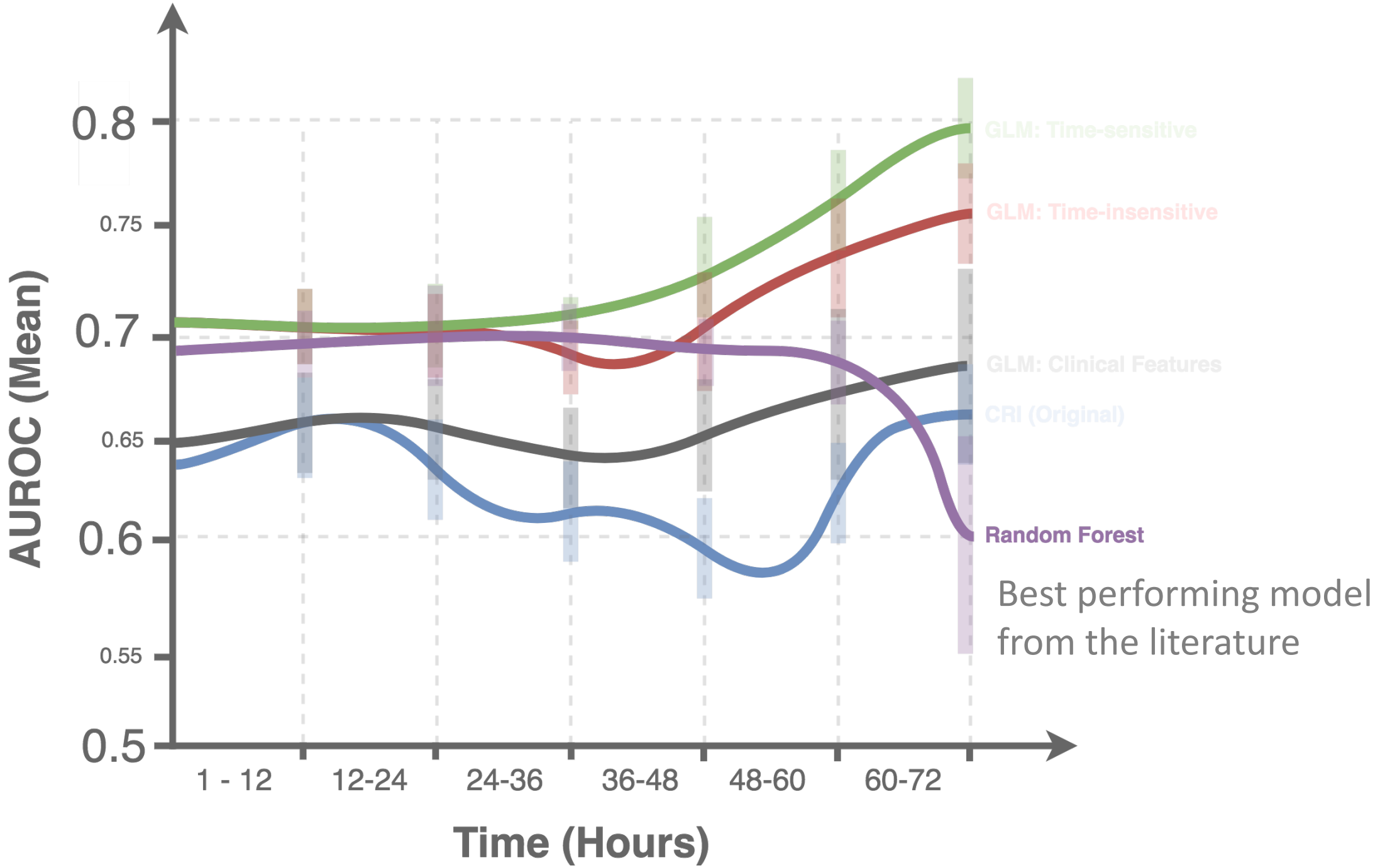
Logistic Regression with *Elastic Memories*

- A logistic regression model with the selected features is used to evaluate performance on the held out test-sets
- This process was designed to imitate how providers might perform prognosis

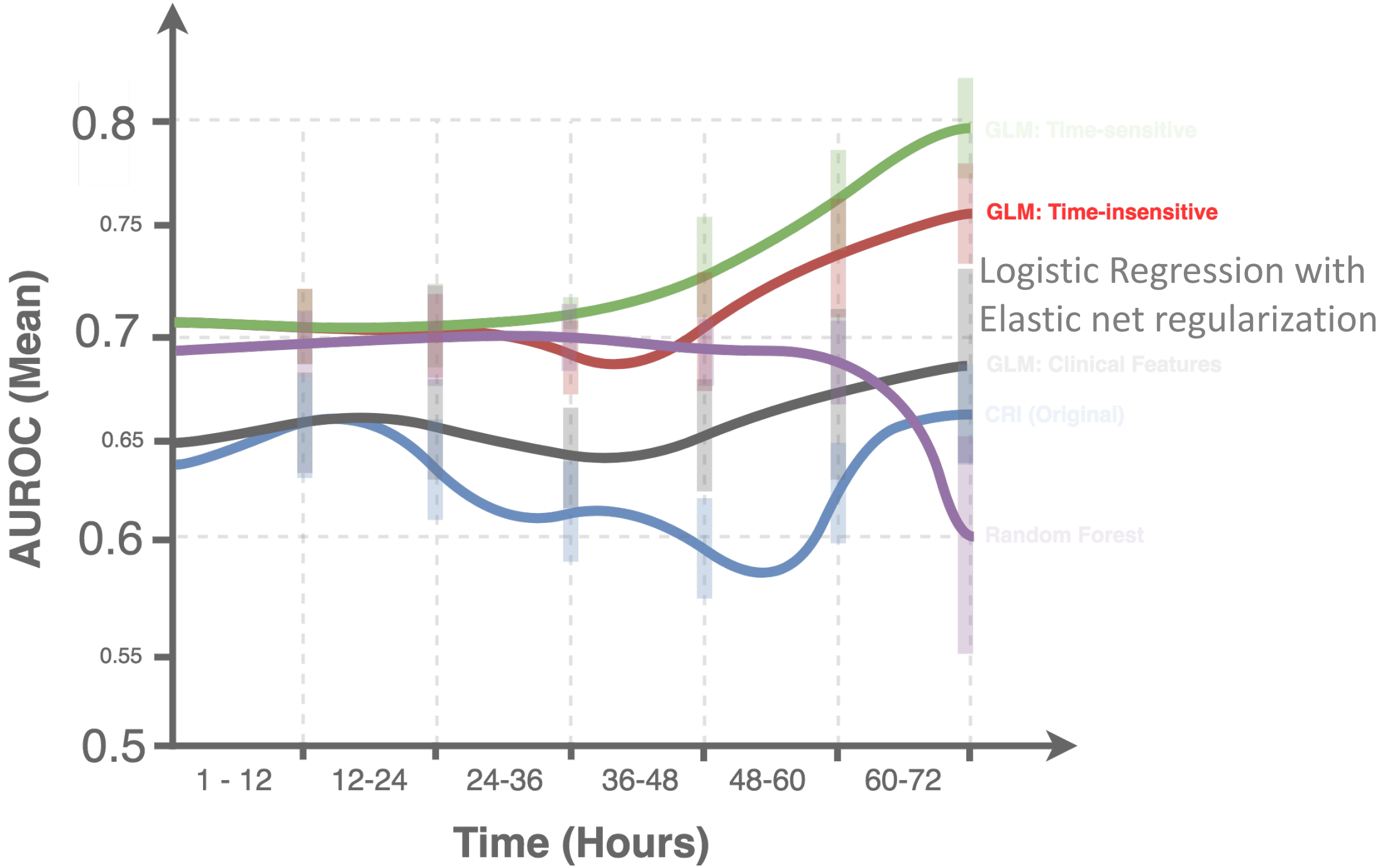


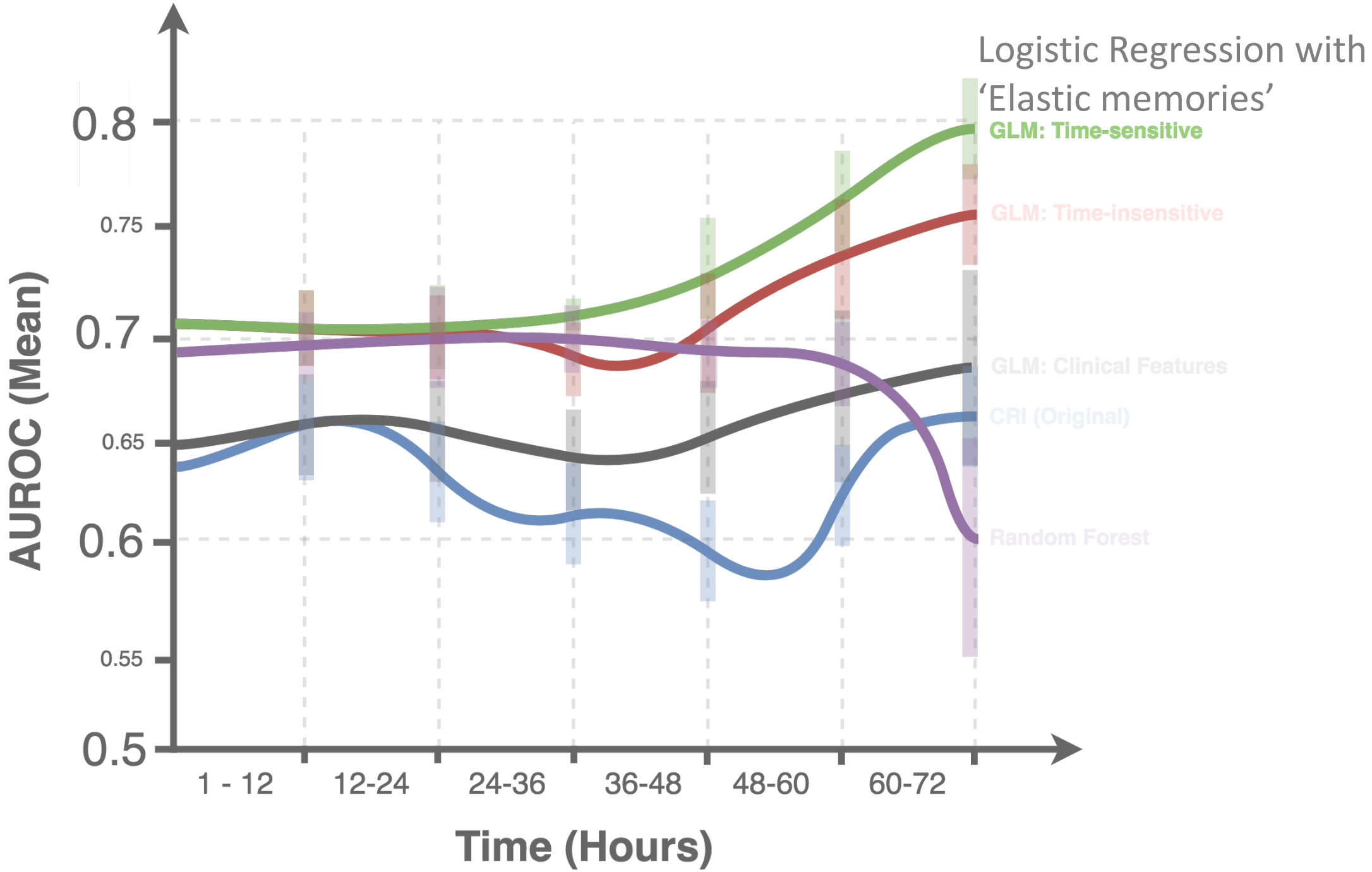
Coefficients Mat Reflect Three Kinds of Relationships

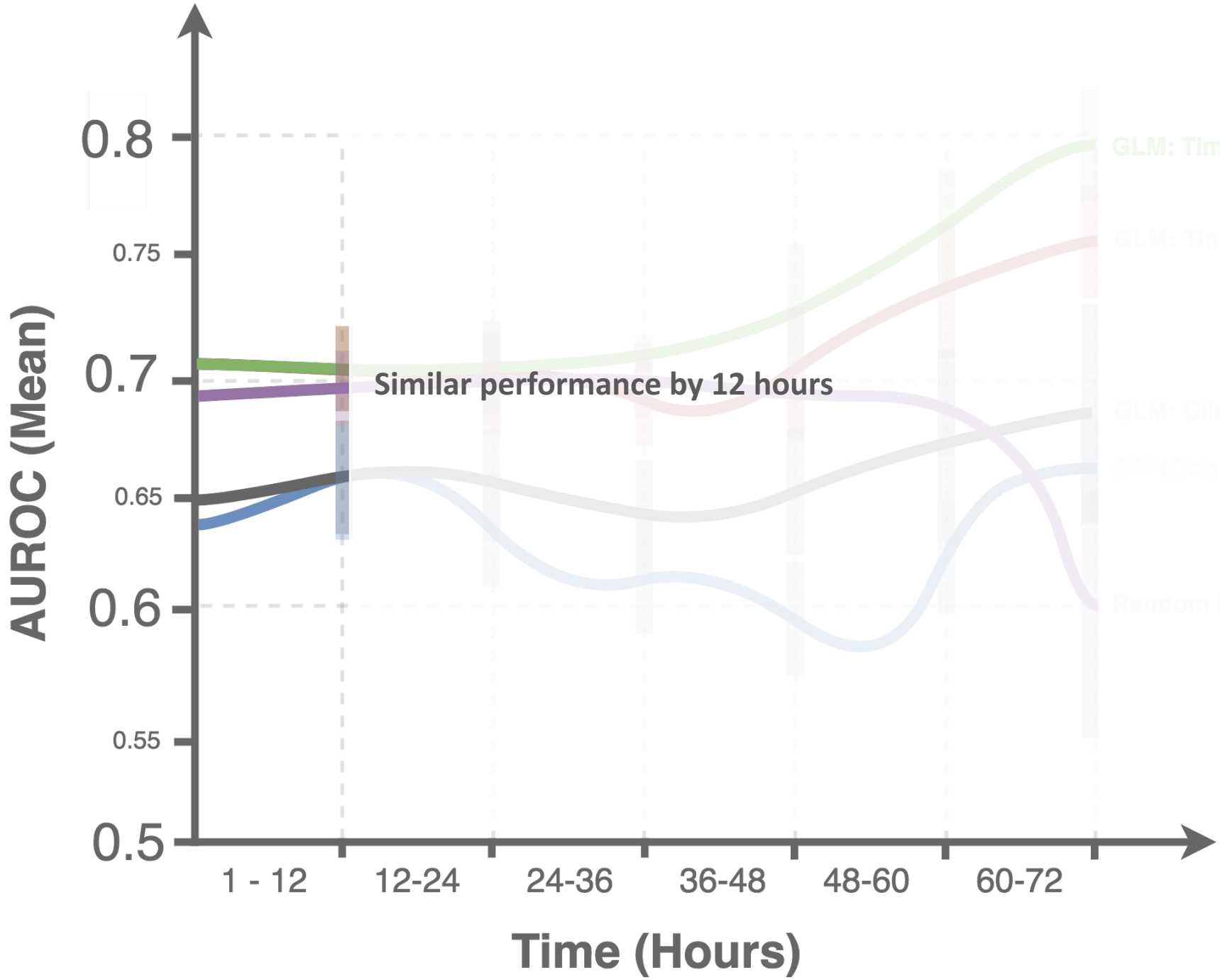
- An **immediate value** is most predictive of outcomes.
- A **prior value** is most predictive of outcomes.
- The **cumulative** value is most predictive of outcomes.

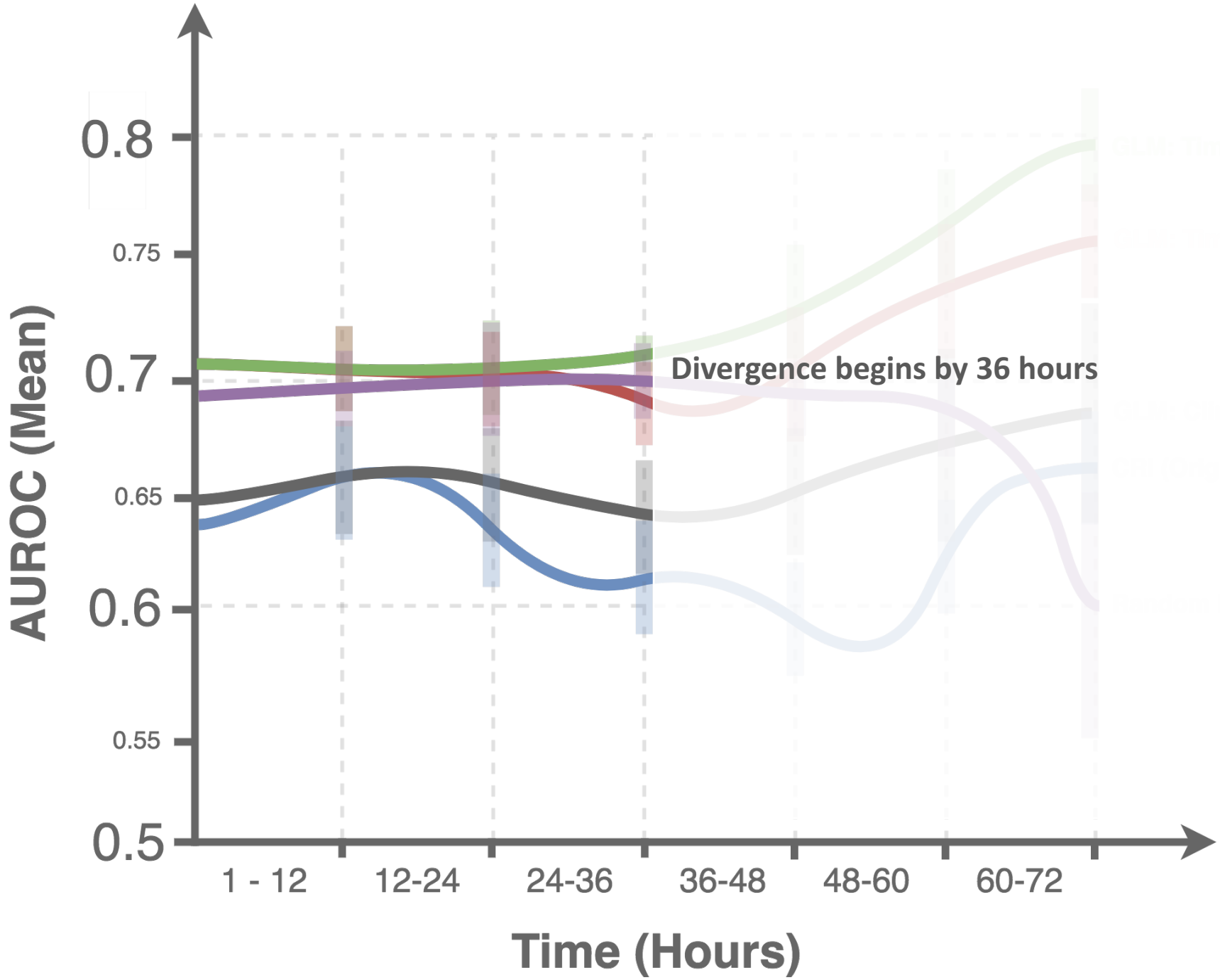


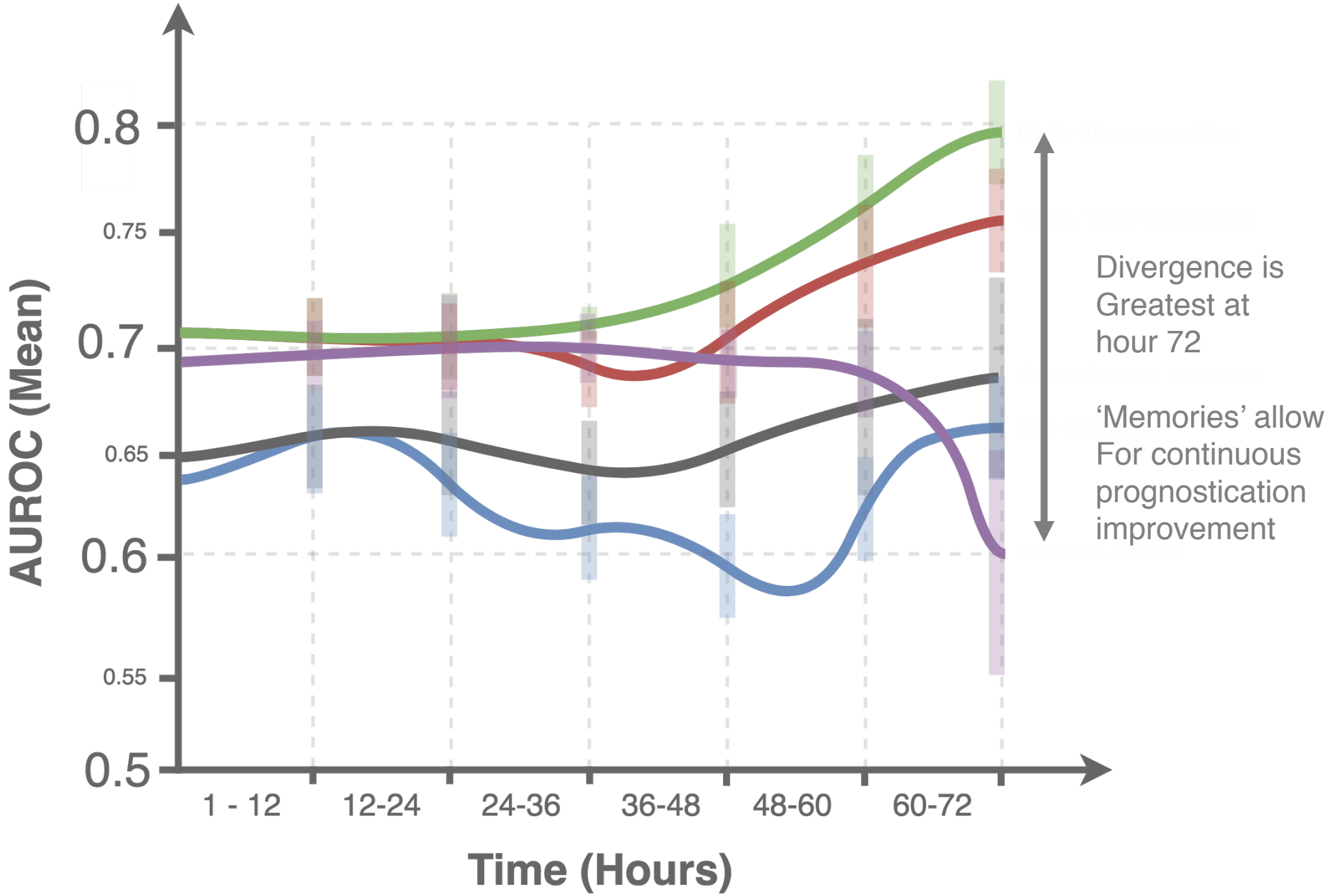
Best performing model from the literature


AUROC (Mean)
0.8
0.75
0.7
0.65
0.6
0.55
0.5
1 - 12
12-24
24-36
36-48
48-60
60-72
Time (Hours)
GLM: Time-sensitive
GLM: Time-insensitive
Logistic Regression with
Elastic net regularization
GLM: Clinical Features
CRI (Original)
Random Forest


Logistic Regression with
'Elastic memories'
GLM: Time-sensitive
GLM: Time-insensitive
GLM: Clinical Features
CRI (Original)
Random Forest
AUROC (Mean)
0.8
0.75
0.7
0.65
0.6
0.55
0.5
1 - 12
12-24
24-36
36-48
48-60
60-72
Time (Hours)

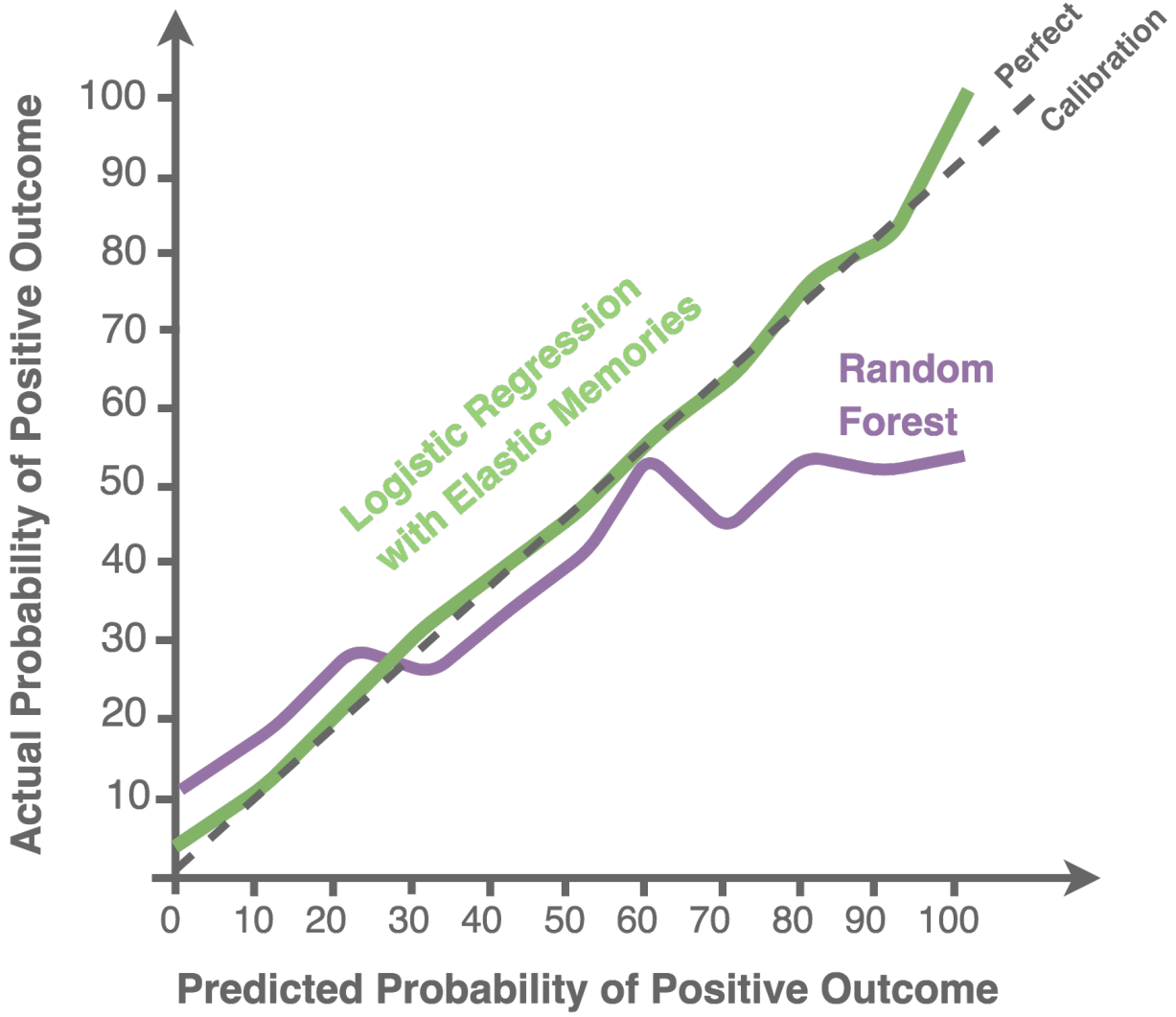

AUROC (Mean)
0.8
0.75
0.7
0.65
0.6
0.55
0.5
1 - 12
12-24
24-36
36-48
48-60
60-72
Time (Hours)
Similar performance by 12 hours
GLM: Time-sensitive
GLM: Time-insensitive
GLM: Clinical Features
CRI (Original)
Random Forest


AUROC (Mean)
0.8
0.75
0.7
0.65
0.6
0.55
0.5
Divergence begins by 36 hours
1 - 12
12-24
24-36
36-48
48-60
60-72
Time (Hours)


AUROC (Mean)
0.8
0.75
0.7
0.65
0.6
0.55
0.5
1 - 12
12-24
24-36
36-48
48-60
60-72
Time (Hours)
Divergence is Greatest at hour 72
‘Memories’ allow For continuous prognostication improvement

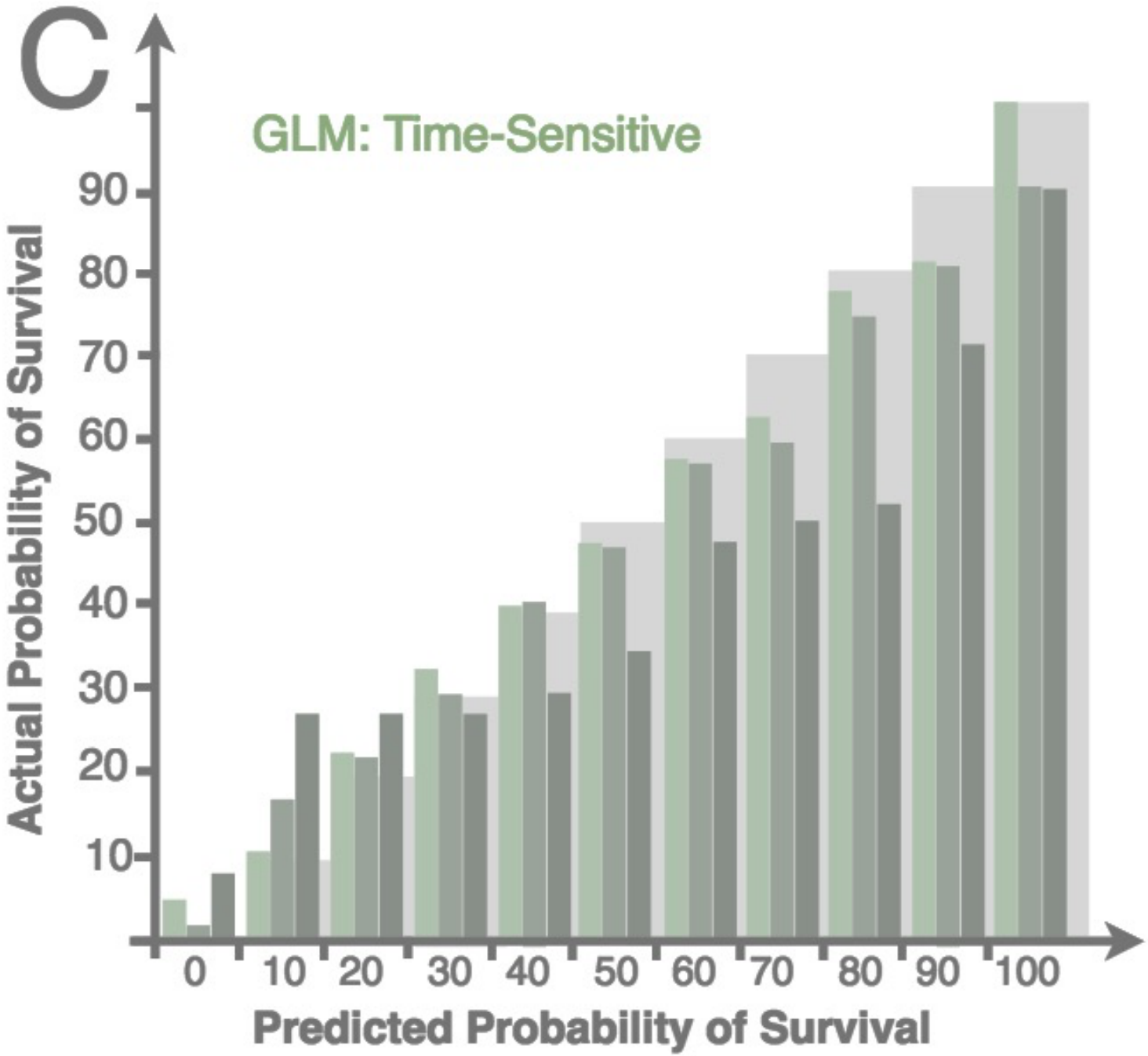
Calibration

- Our approach exhibited enhanced calibration compared to the literature baseline
- This allows for a more nuanced use of the model, compared to existing approaches



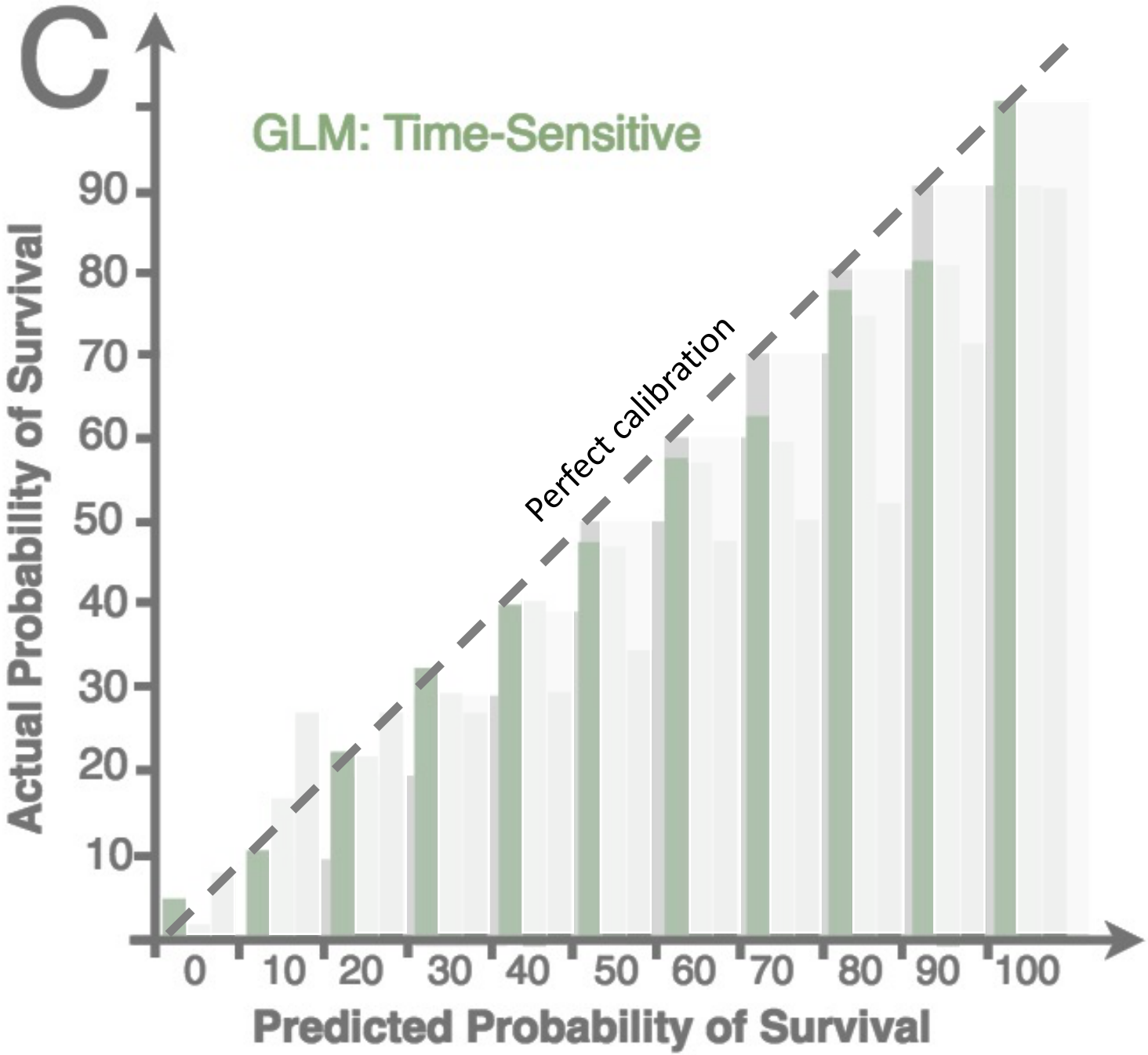
Calibration

- Our approach was also well calibrated over time.



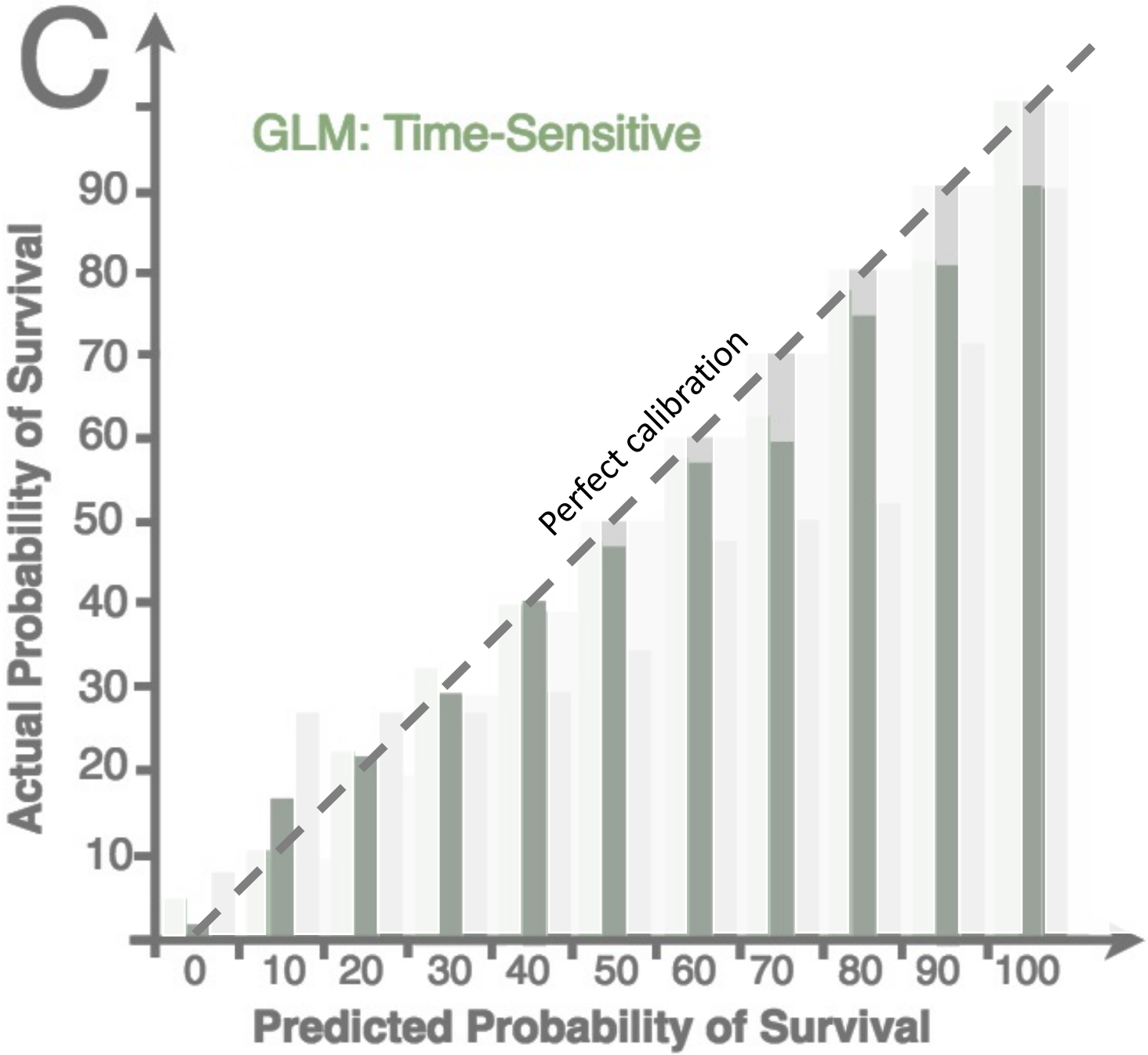
Calibration

- Our approach was also well calibrated over time.
- Day 1



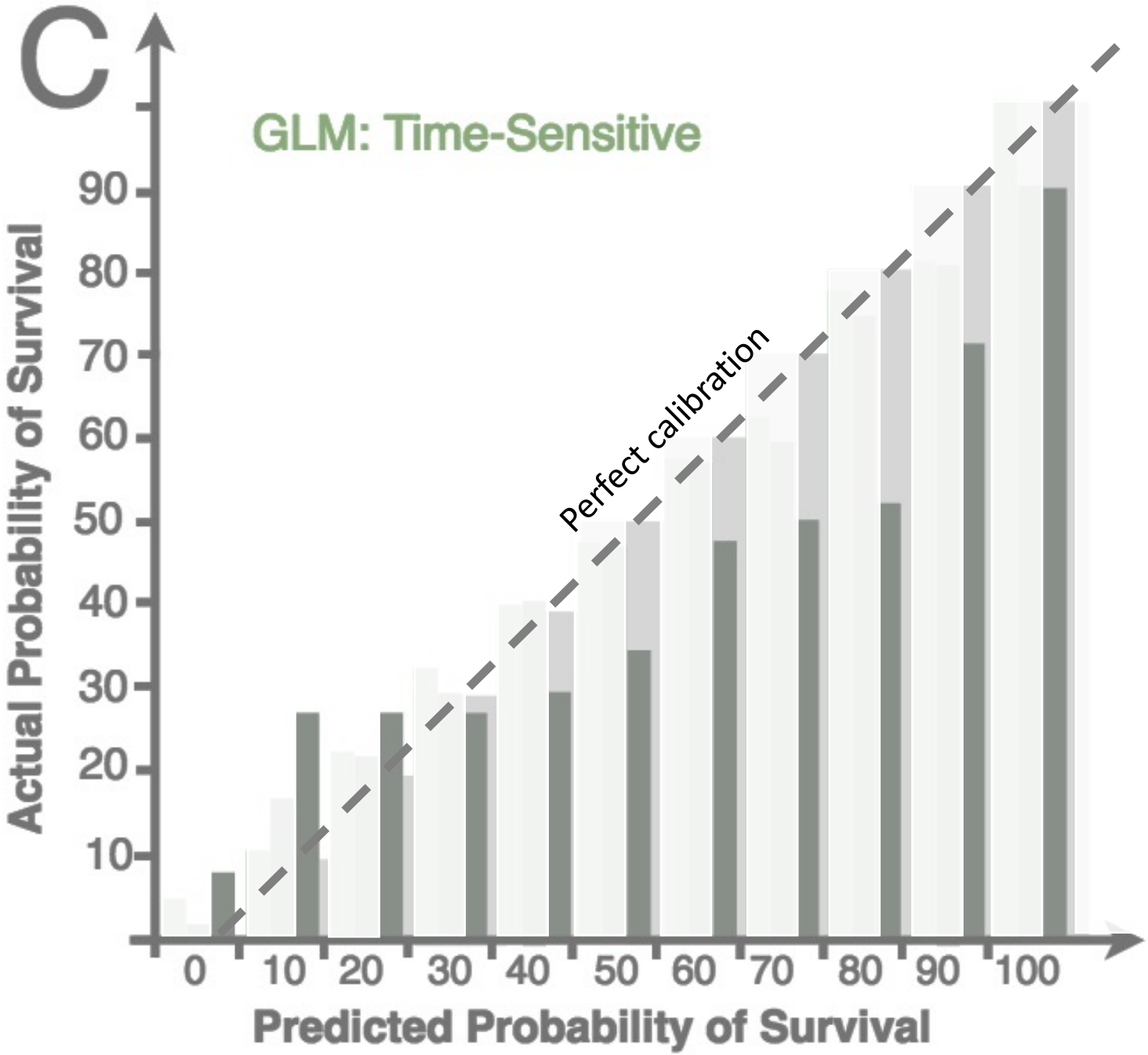
Calibration

- Our approach was also well calibrated over time.
- Day 2



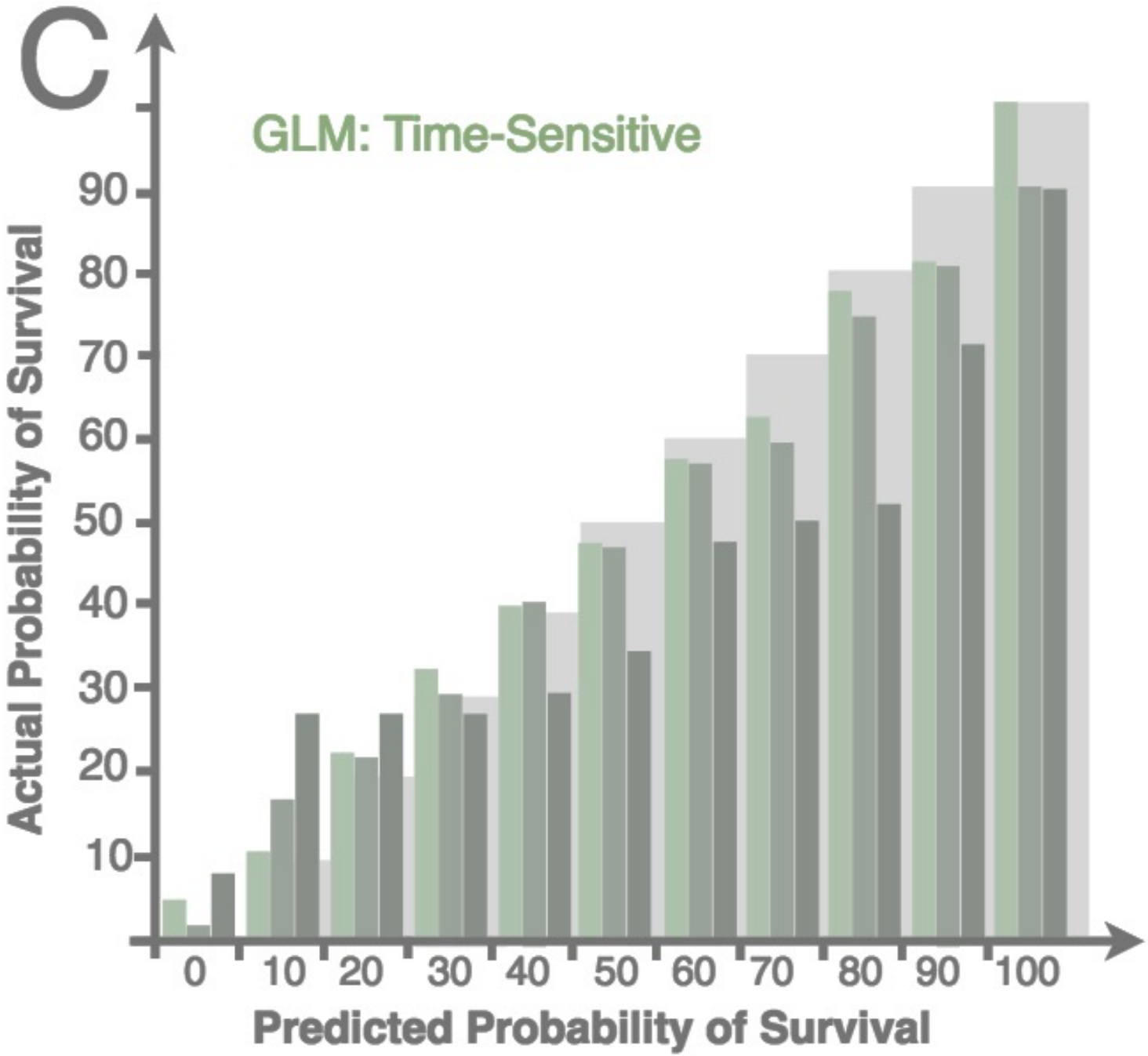
Calibration

- Our approach was also well calibrated over time.
- Day 3

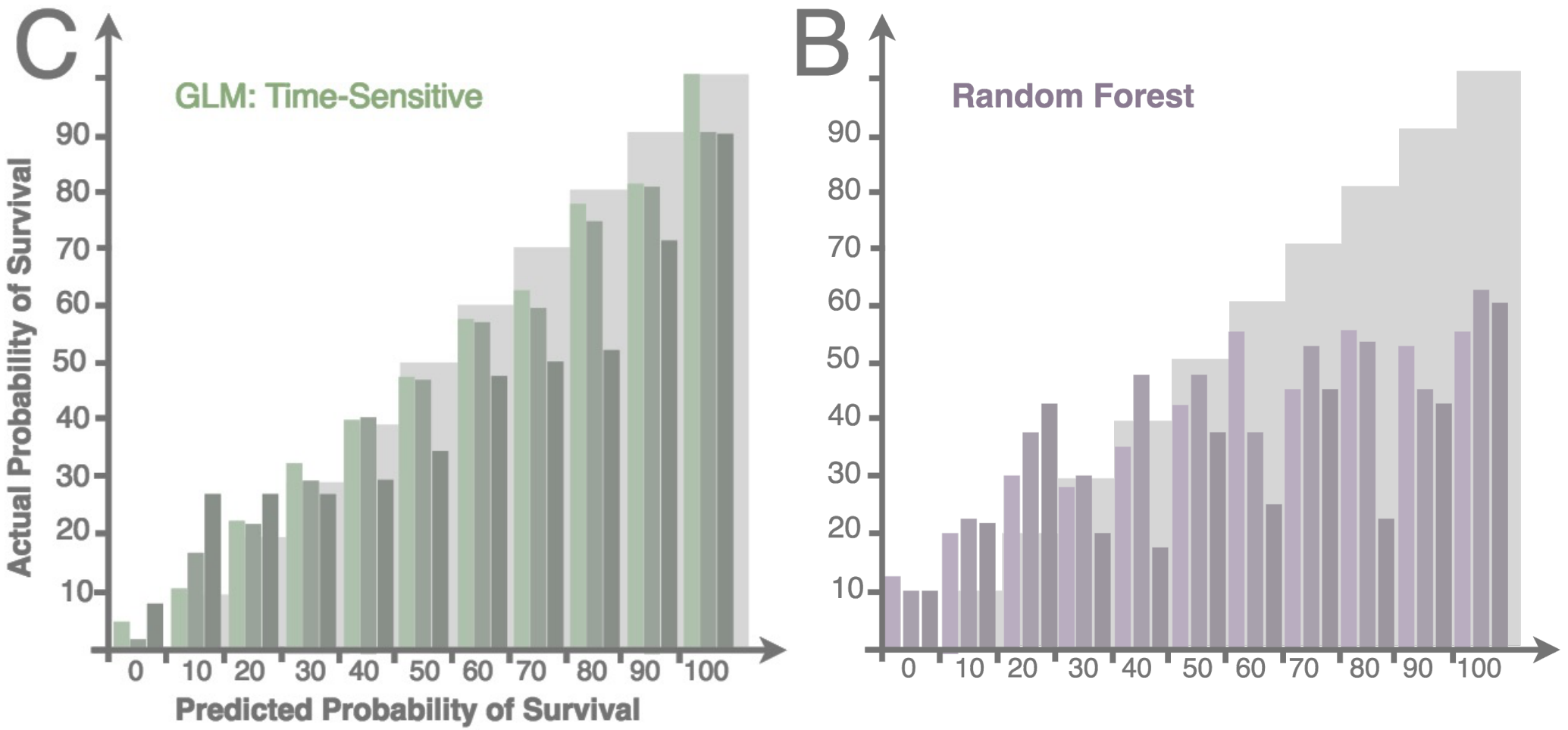


Calibration

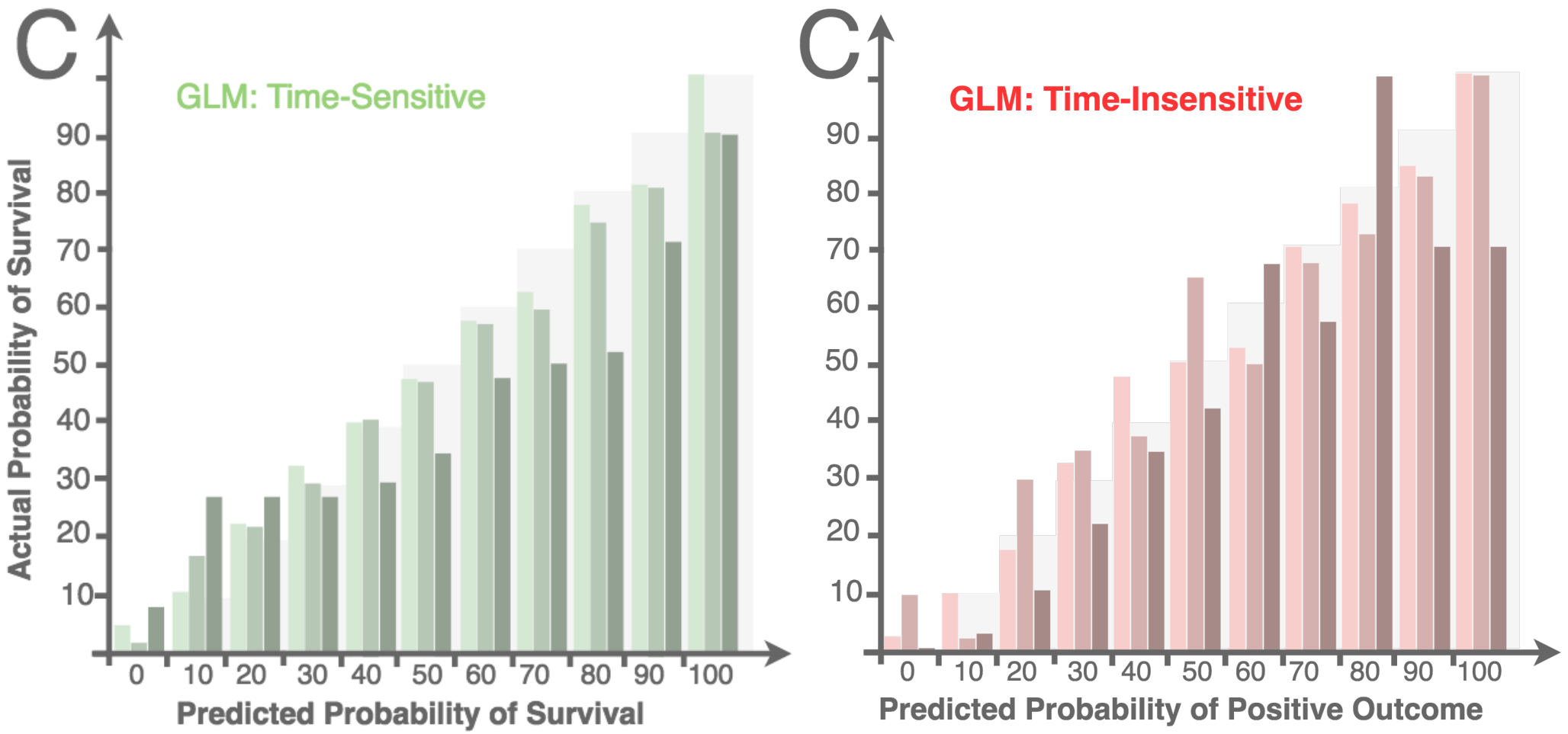
- Our approach was also well calibrated over time.



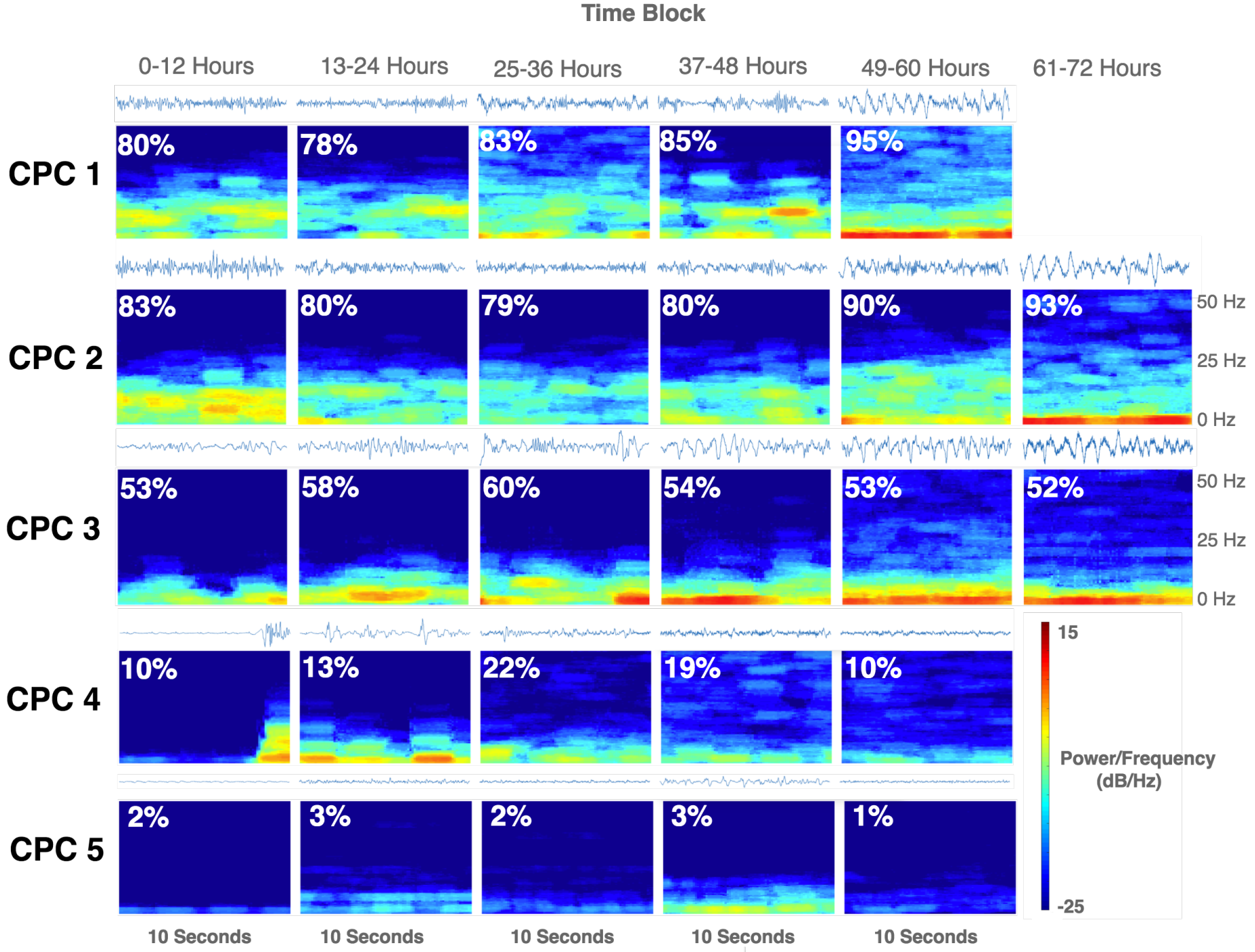
Calibration



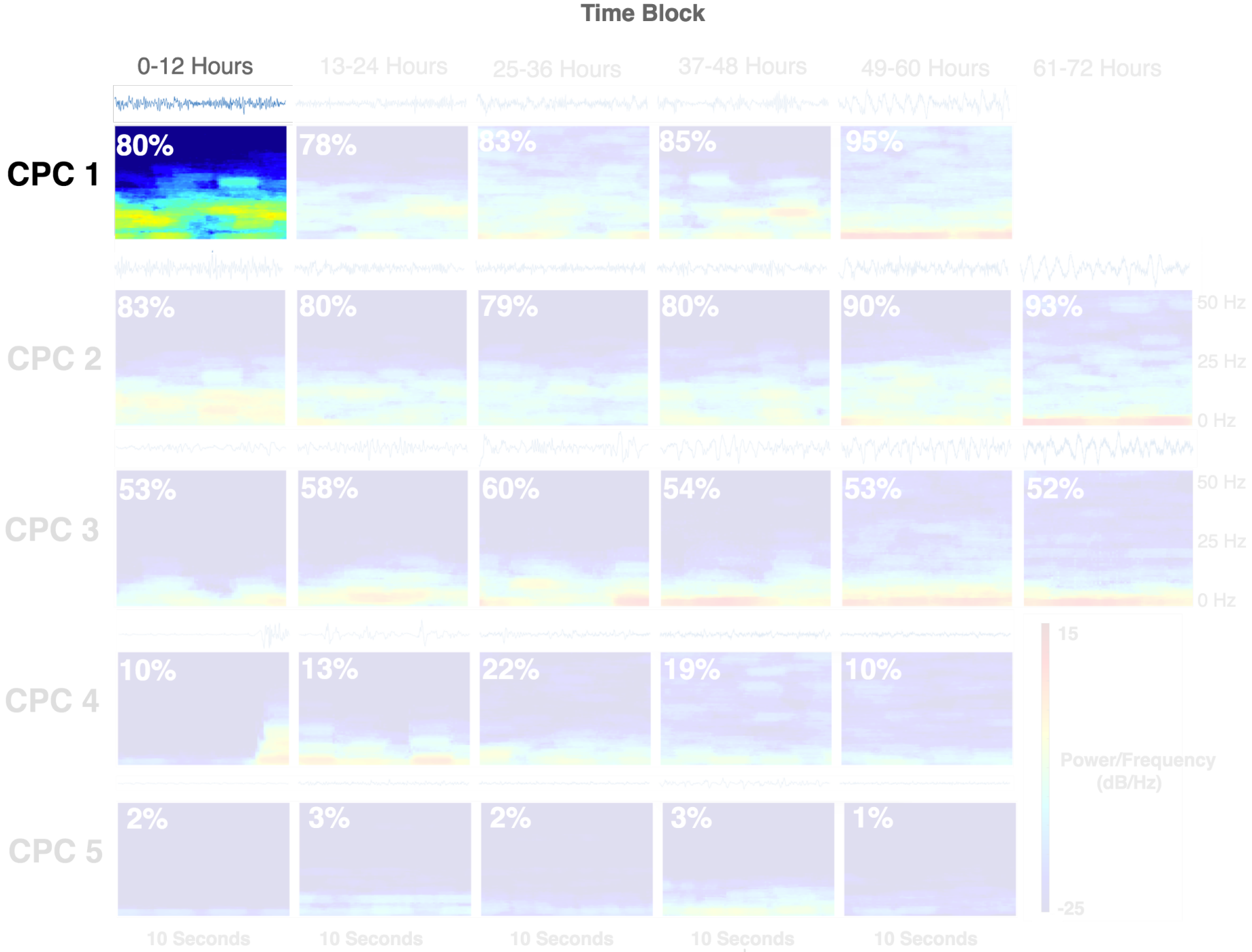
Calibration



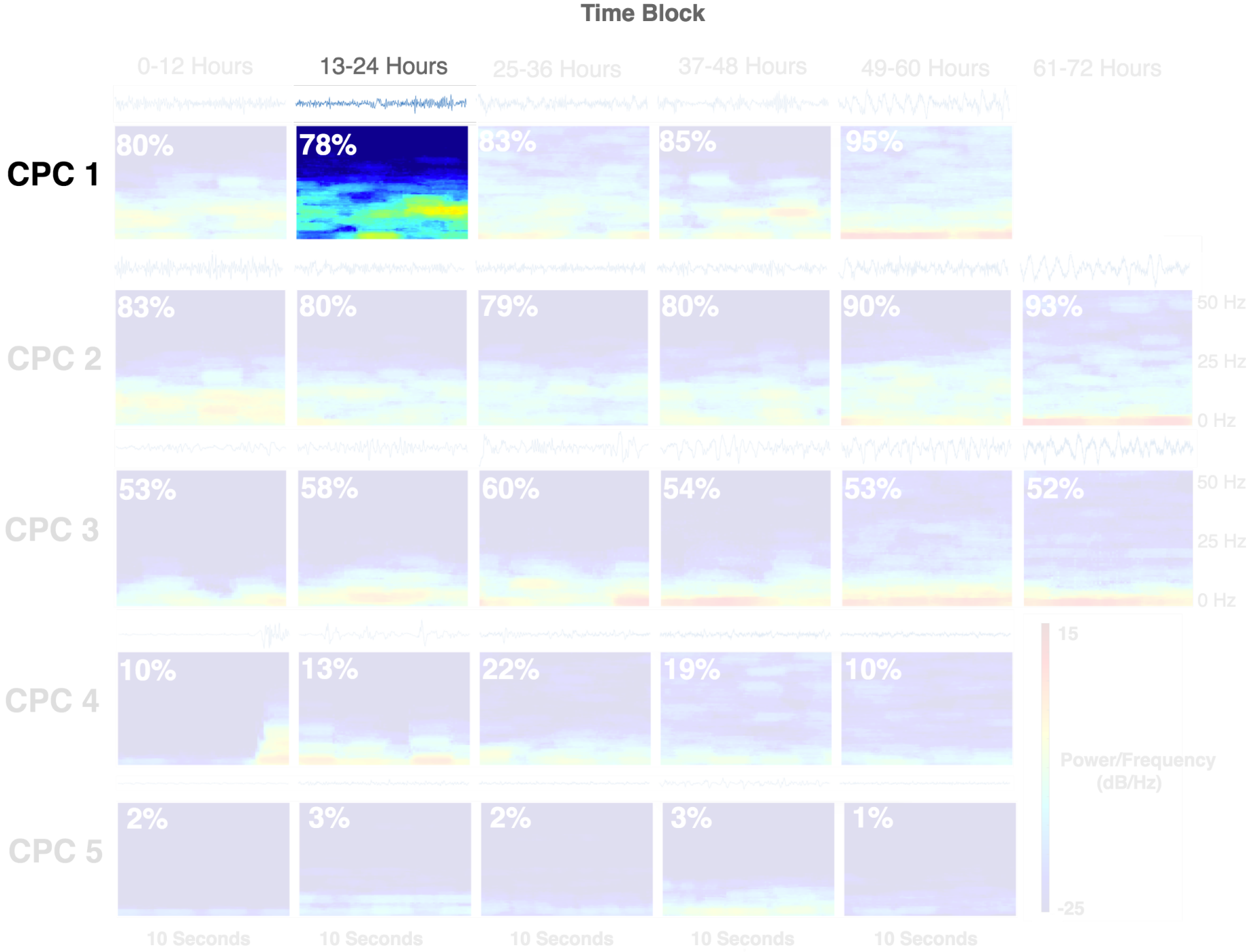
Exemplary Patients



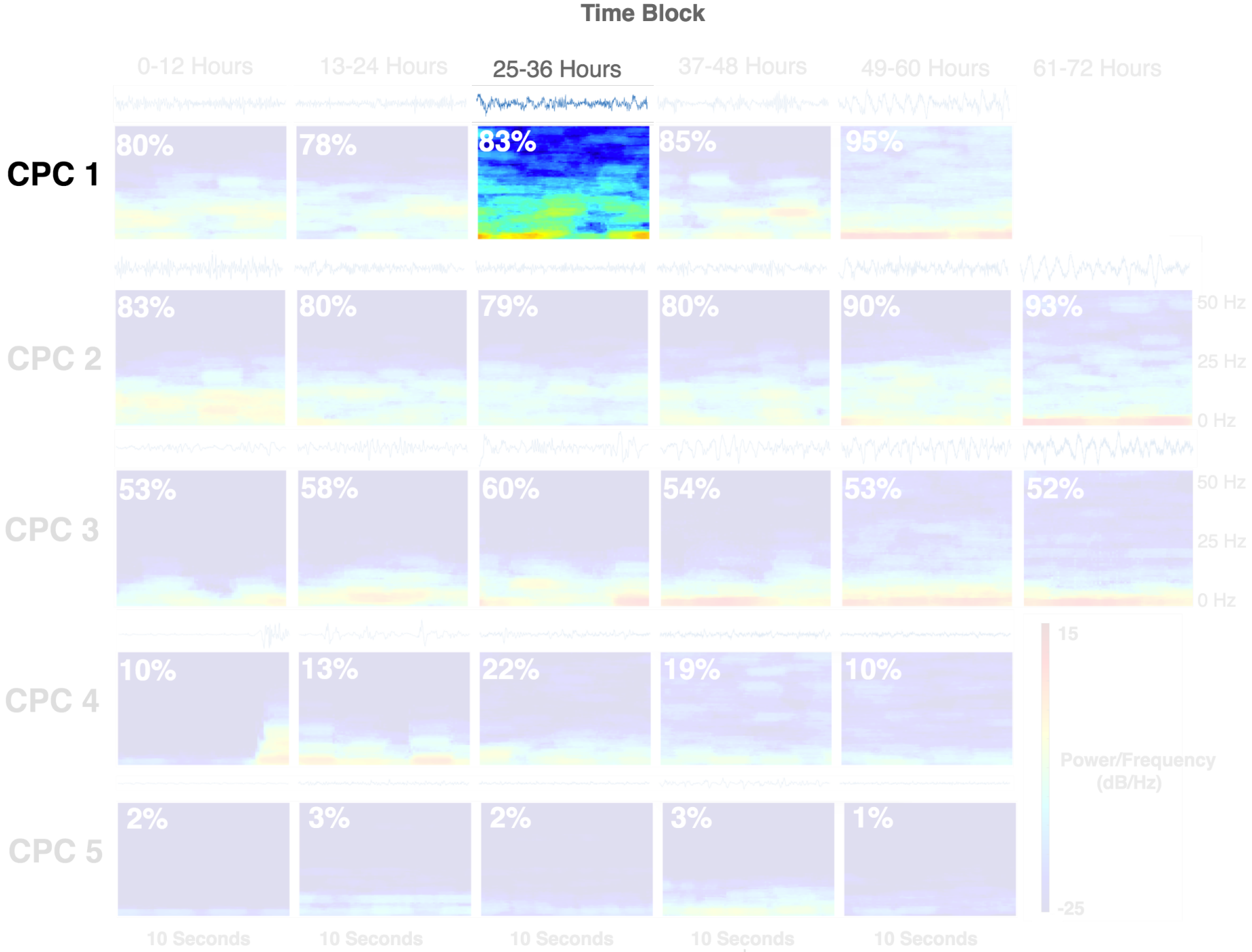
Exemplary Patients



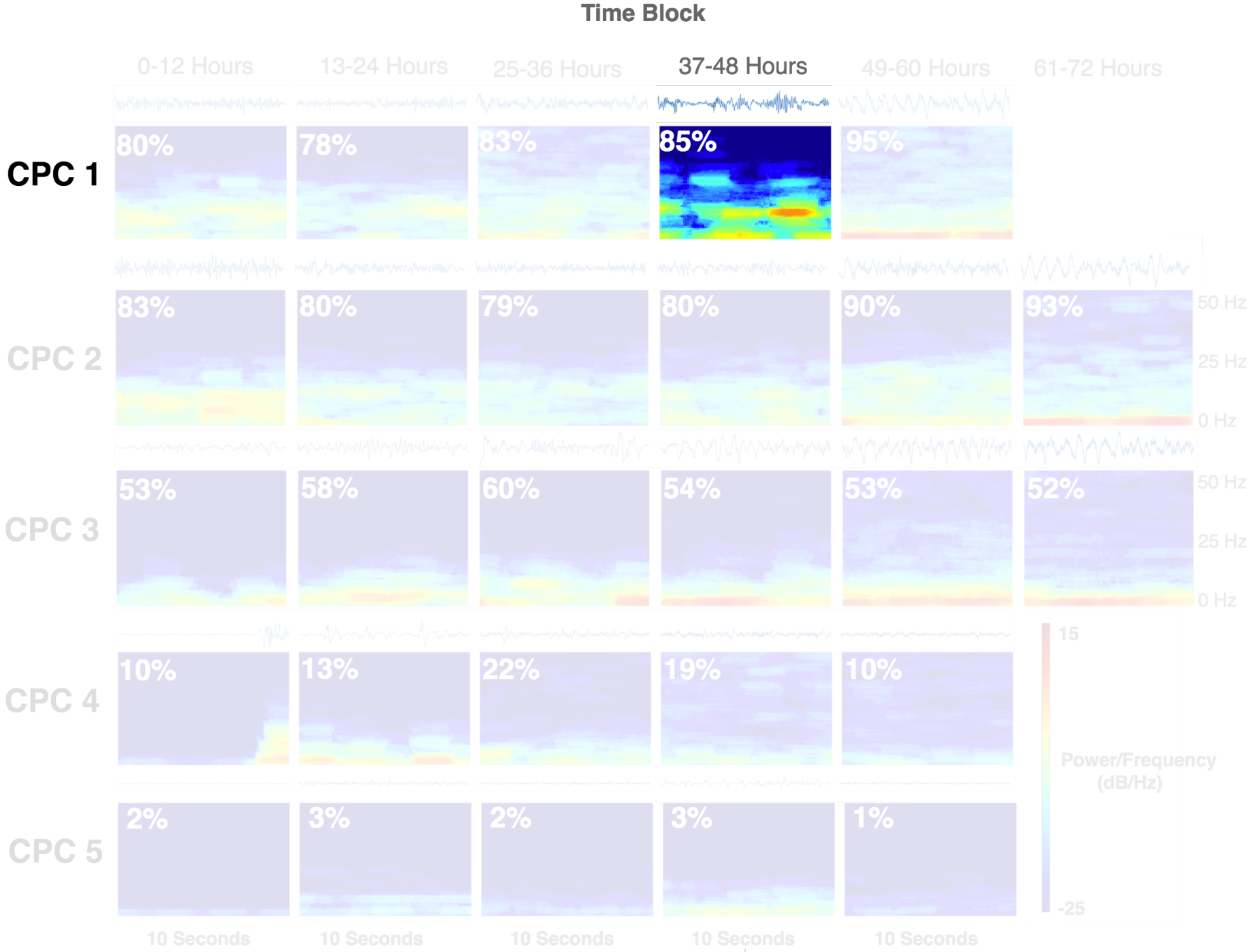
Exemplary Patients



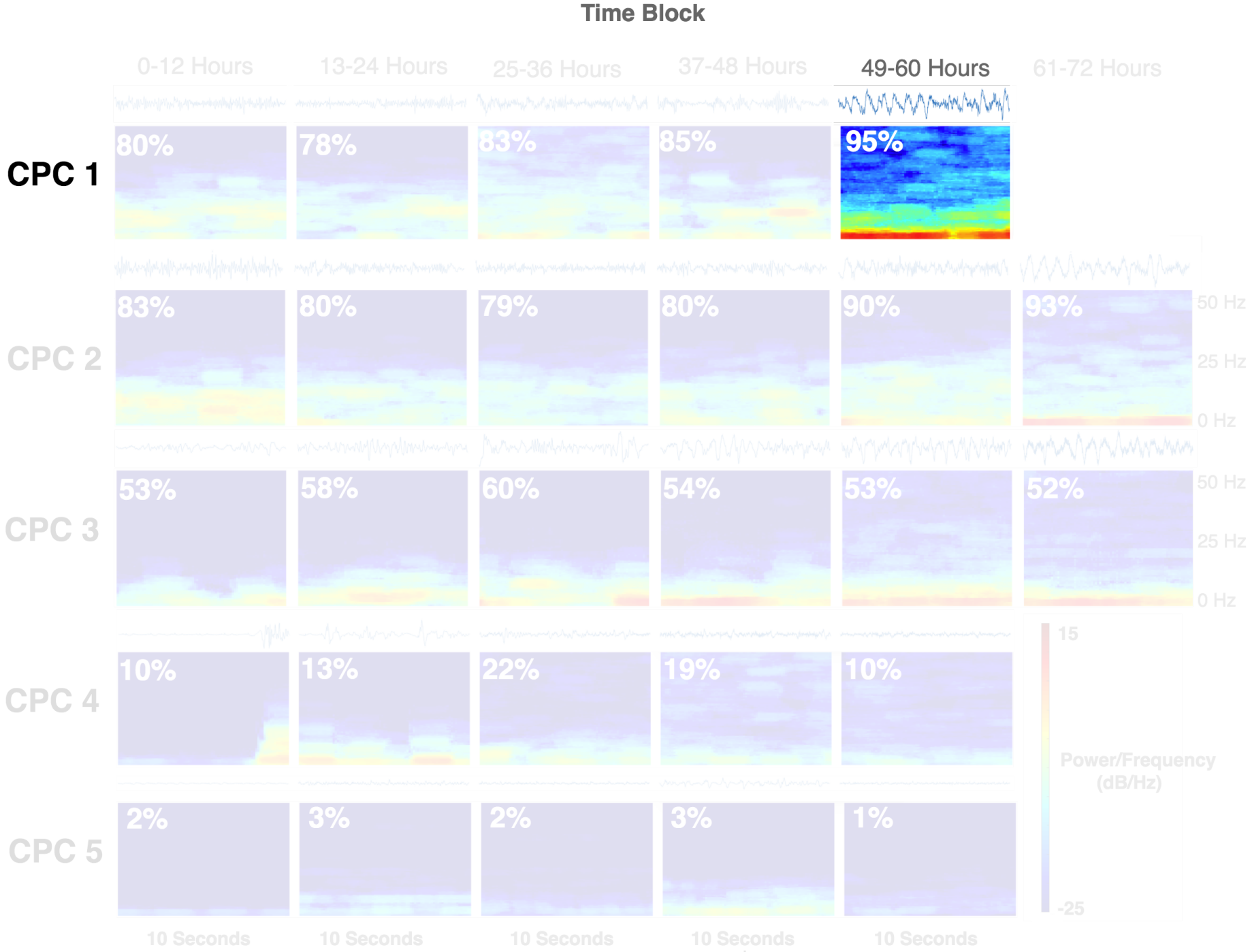
Exemplary Patients



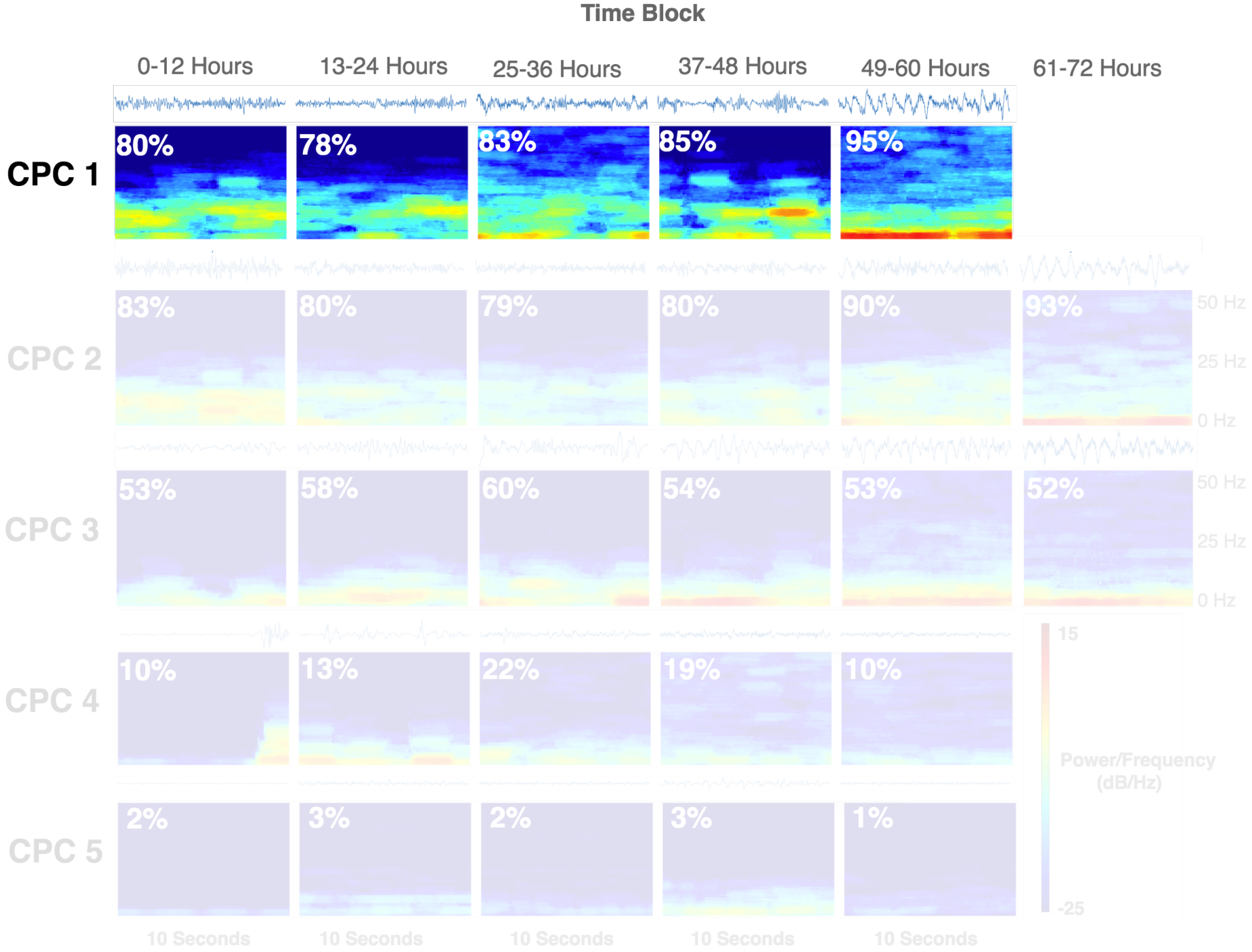
Exemplary Patients



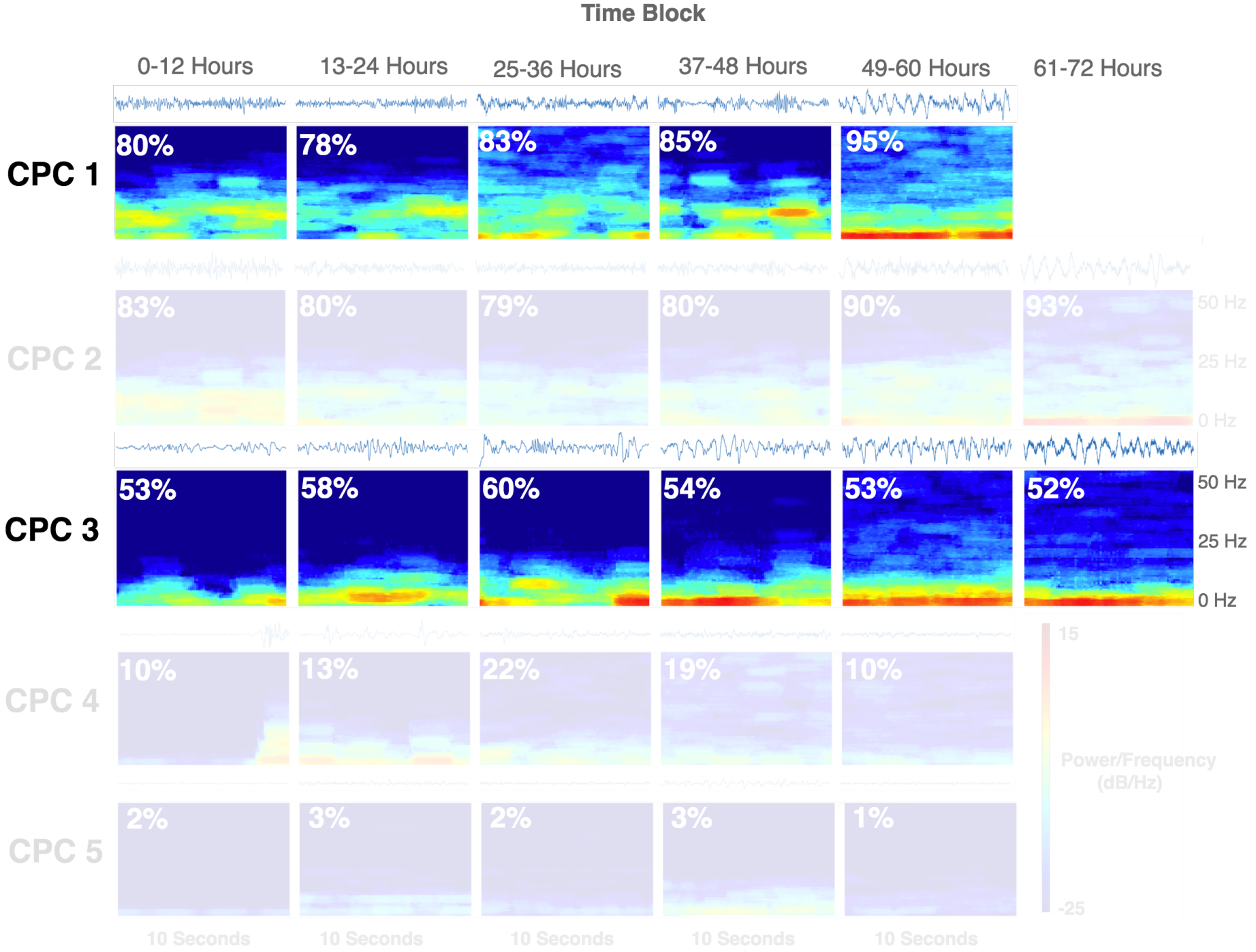
Exemplary Patients



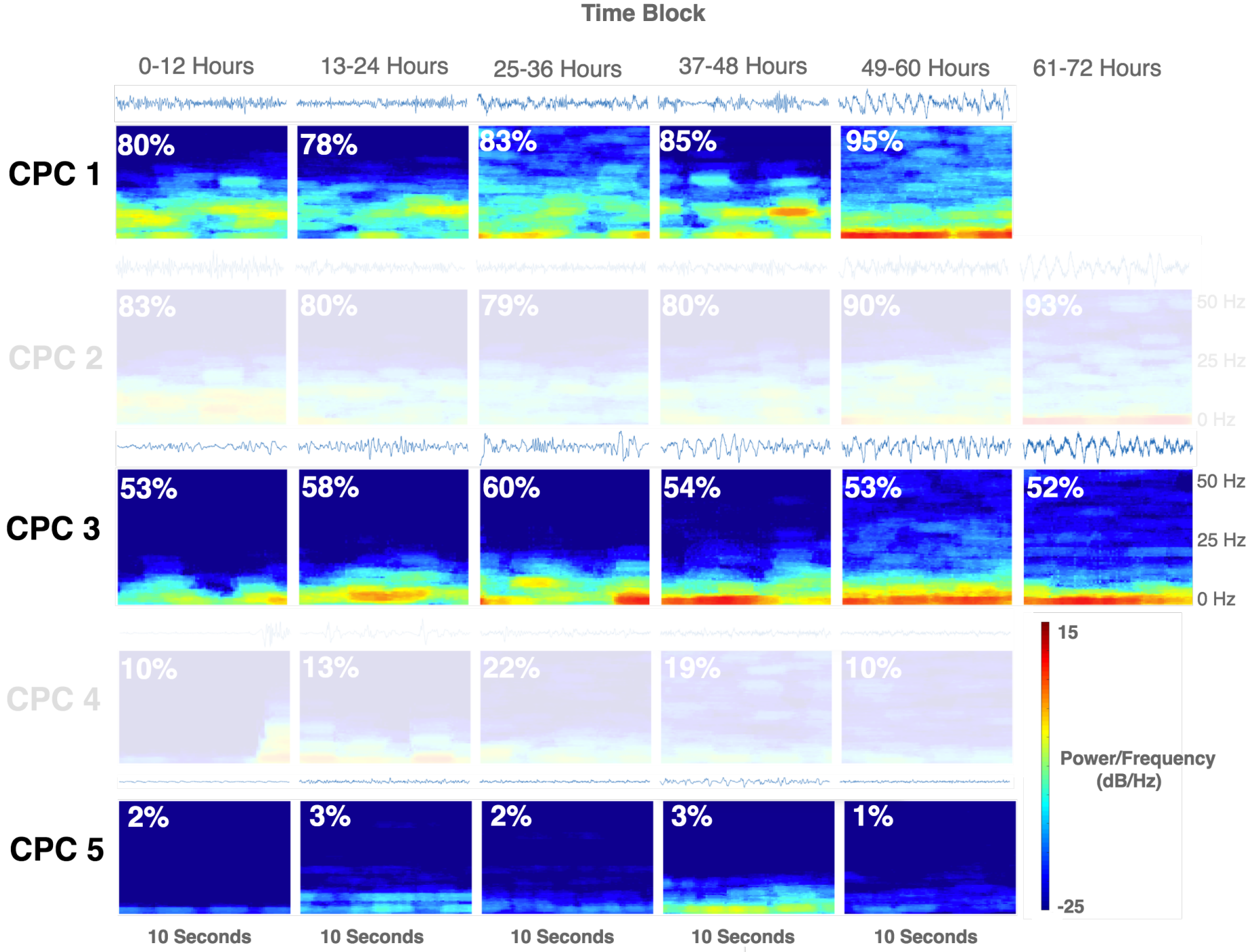
Exemplary Patients



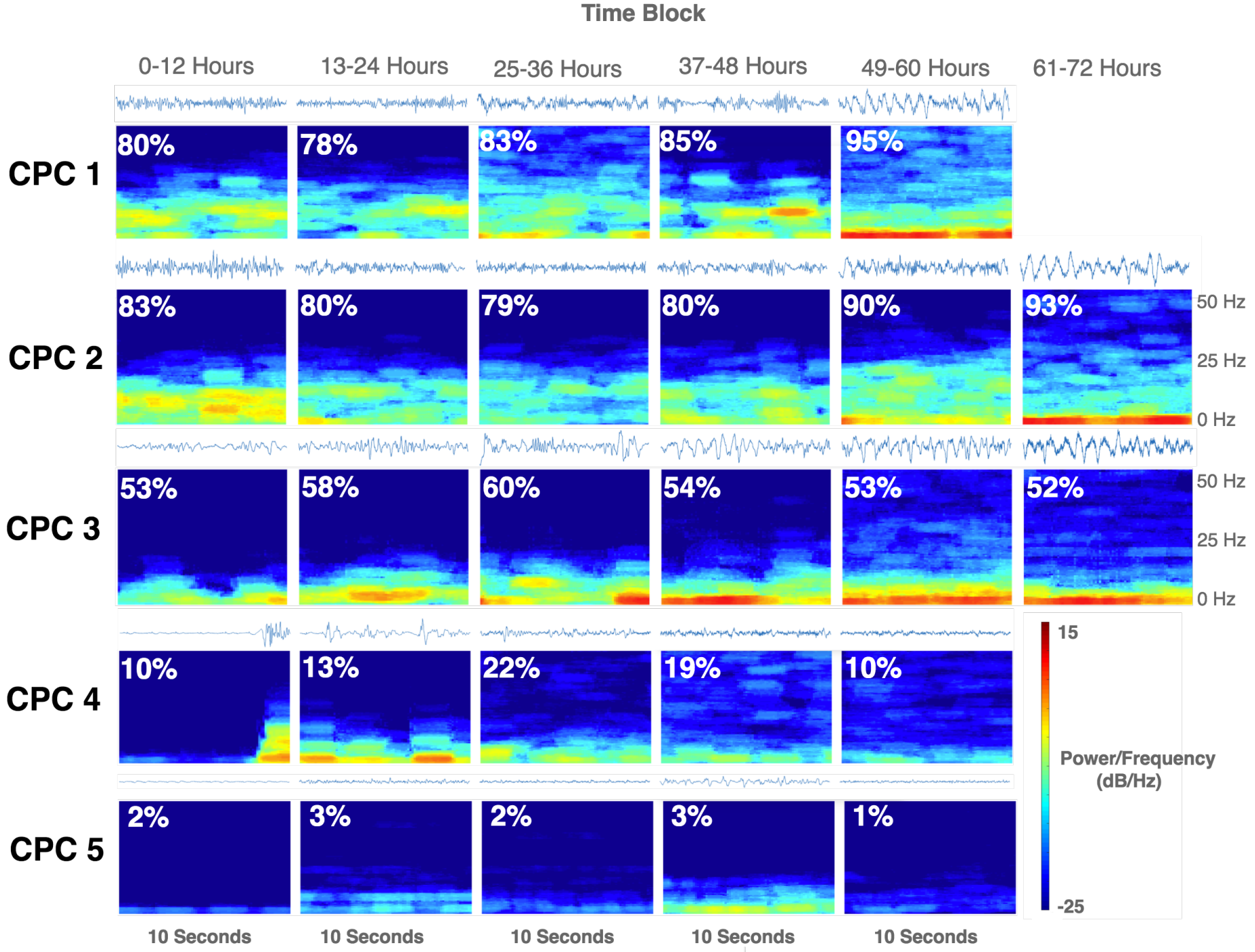
Exemplary Patients



Exemplary Patients



Exemplary Patients



Pros/Cons of the Approach

- Advantages
  - Does not require time of arrest
  - Prognostic performance improves over time using ‘memories’
  - Well calibrated
- Disadvantages
  - Alignment with respect to EEG initiation harm physiological interpretations
  - Assumes that feature coefficients in neighboring intervals are independent
  - Does not account for spatial information

Three Approaches

Two Feature Based:

- ~~1. Sequential Logistic Regression with ‘Elastic’ memories.~~
- 2. Sequential Logistic Regression with Dynamic Bayesian network constraints

One Data Driven

- 3. Deep neural networks

Logistic Regression with *DBN*

- n = 785 patients
- 10 EEG, 5 clinical features
- Validation:
- Leave-one-subject-out

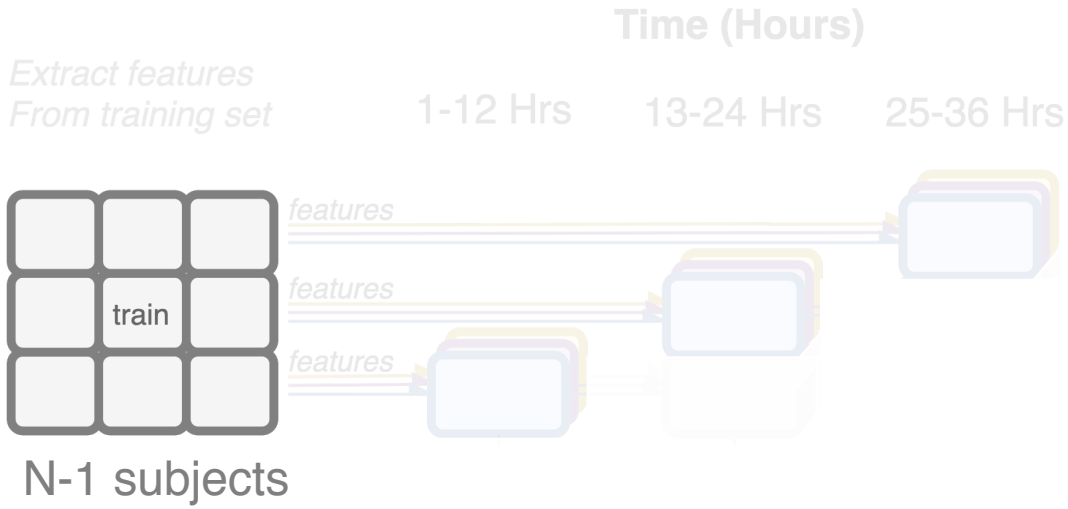
Logistic Regression with *DBN*

- n = 785 patients
- 10 EEG, 5 clinical features
- Validation:
- Leave-one-subject-out

Signal Feature
Complexity
Shannon Entropy
Hjorth Complexity
False Nearest Neighbor
Category
Standard Deviation
Regularity (burst-suppression)
Diffuse Slowing
Spikes
Connectivity
Coherence - delta
Phase Lag Index
Cross-correlation Magnitude

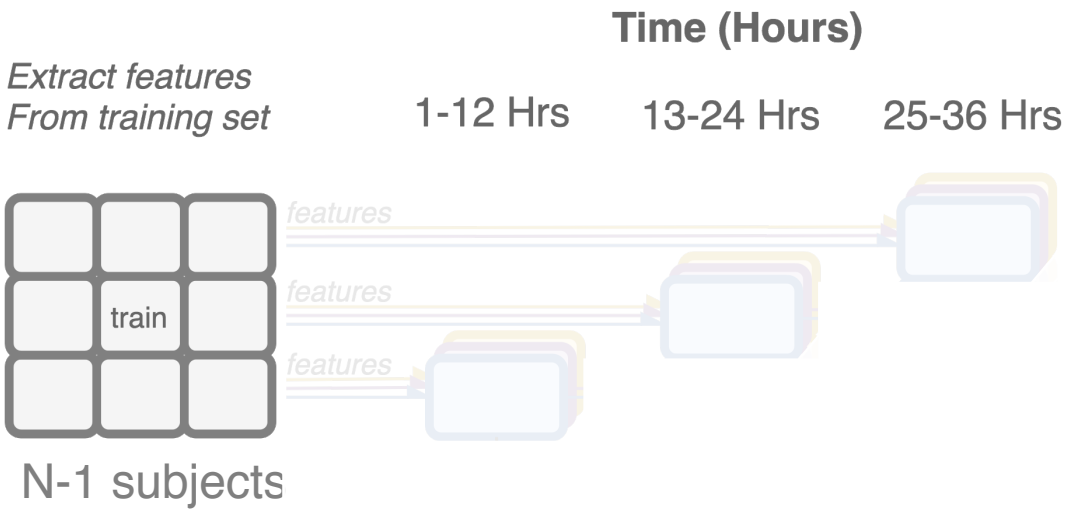
Logistic Regression with *DBN*

- n = 785 patients
- 10 EEG, 5 clinical features
- Validation:
- Leave-one-subject-out



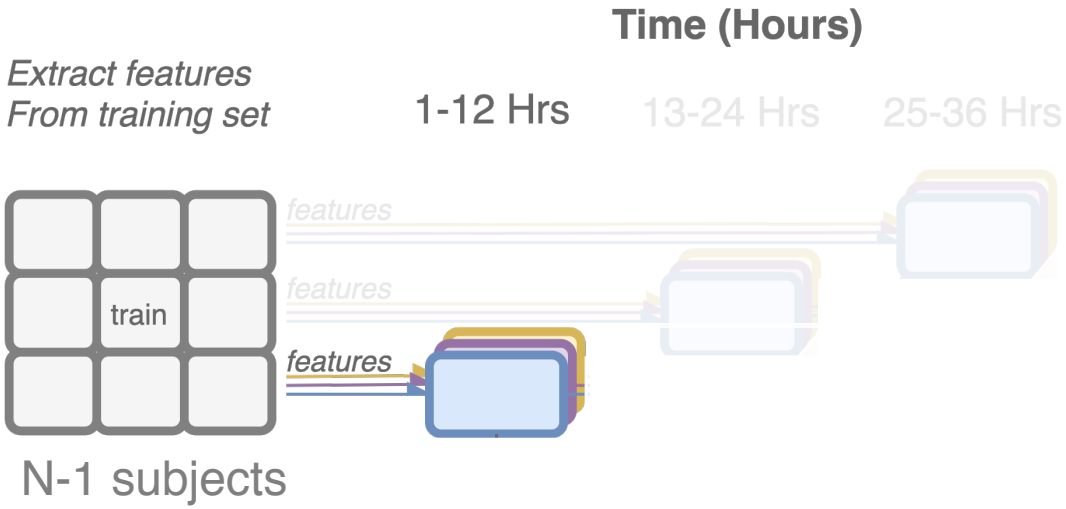
Logistic Regression with *DBN*

- We extracted features in 12 hour intervals:



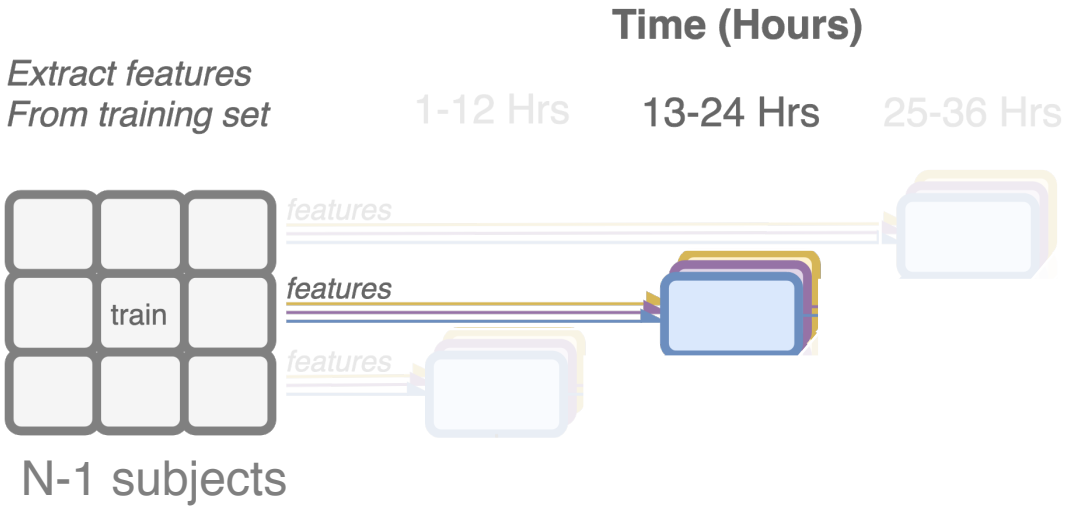
Logistic Regression with *DBN*

- We extracted features in 12 hour intervals:
- 1-12 hours



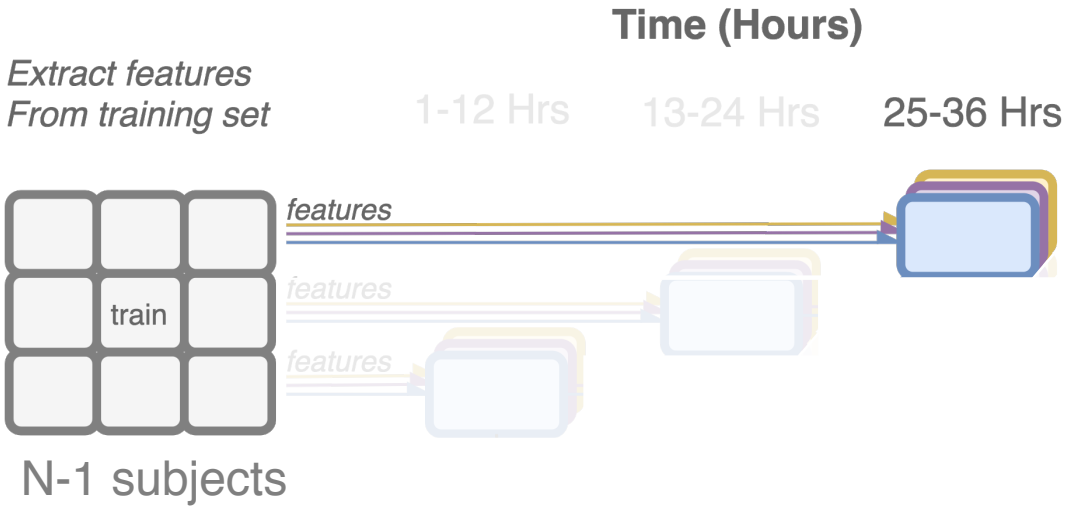
Logistic Regression with *DBN*

- We extracted features in 12 hour intervals:
- 13-24 hours



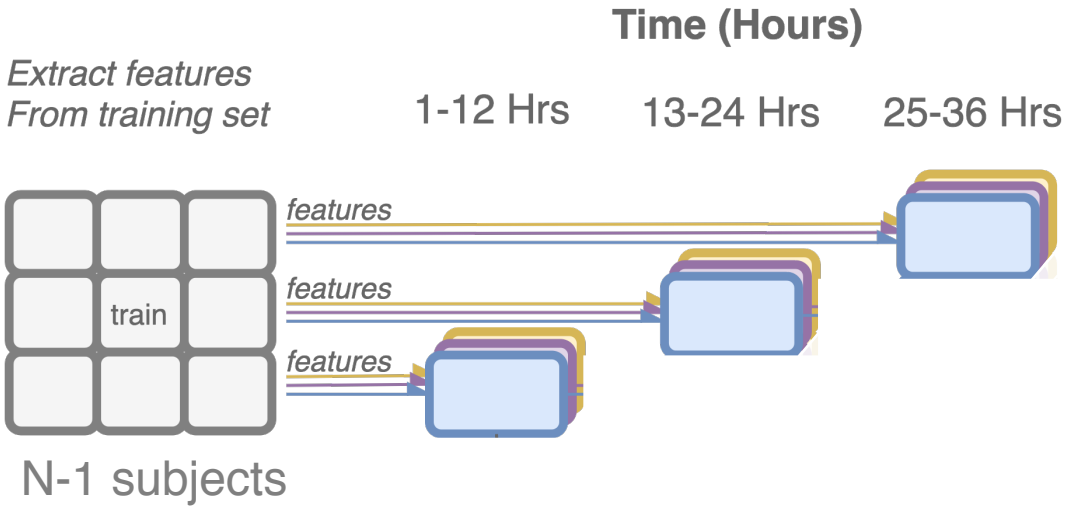
Logistic Regression with *DBN*

- We extracted features in 12 hour intervals:
- 25-36 hours



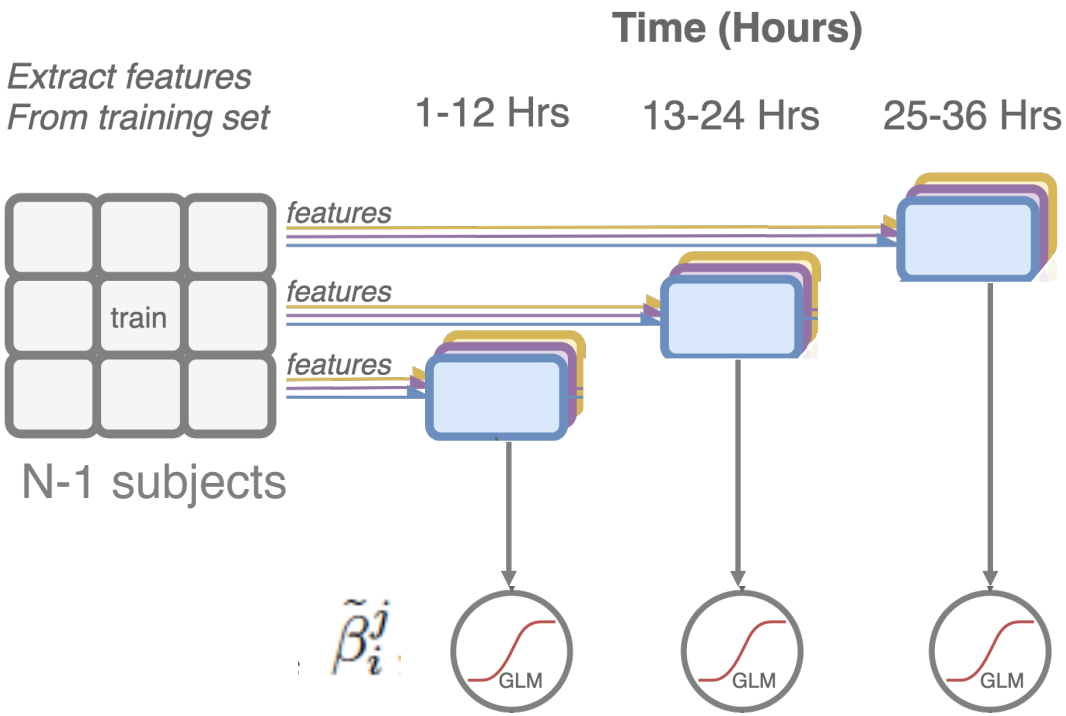
Logistic Regression with *DBN*

- We extracted features in 12 hour intervals:
- etc…



Logistic Regression with *DBN*

- We learn parameters of several logistic regressions: *coefficient measures* $\tilde{\beta}_i^j$
- With corresponding *coefficient measure errors* R_i^j ,which are simply the standard errors



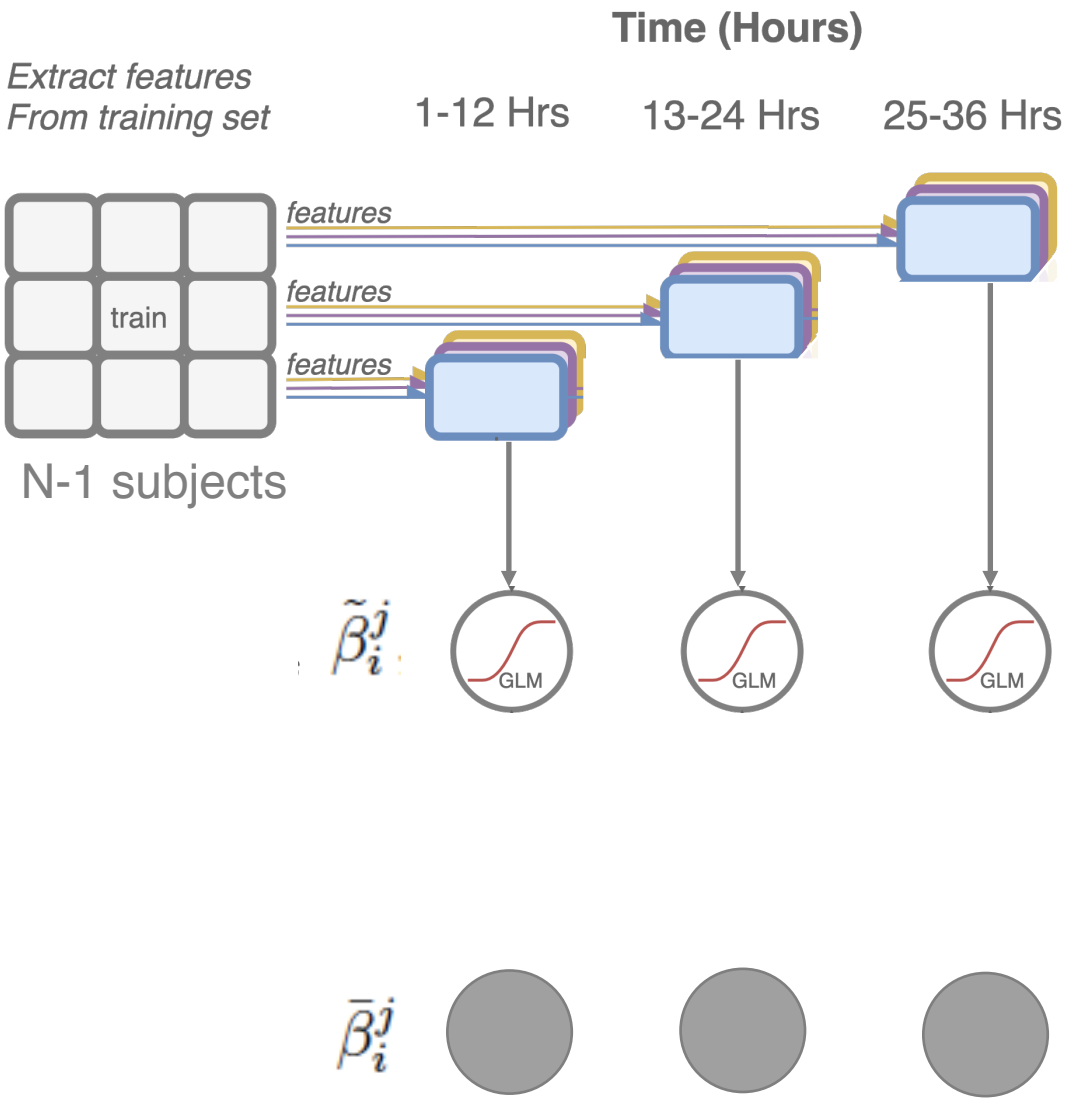
Logistic Regression with *DBN*

- To relate the coefficient measurements across time intervals, we specify a model of coefficient dynamics which provide *coefficient estimates* $\bar{\beta}_i^j$

$$f(\beta_i^{j-1}) = \beta_i^{j-1} + \Delta(\beta_i^{j-1}) = \bar{\beta}_i^j$$

- With *estimate error* denoted by $\bar{P}_i^j$

$$\bar{P}_i^j = P_i^{j-1} + Q$$



Logistic Regression with *DBN*

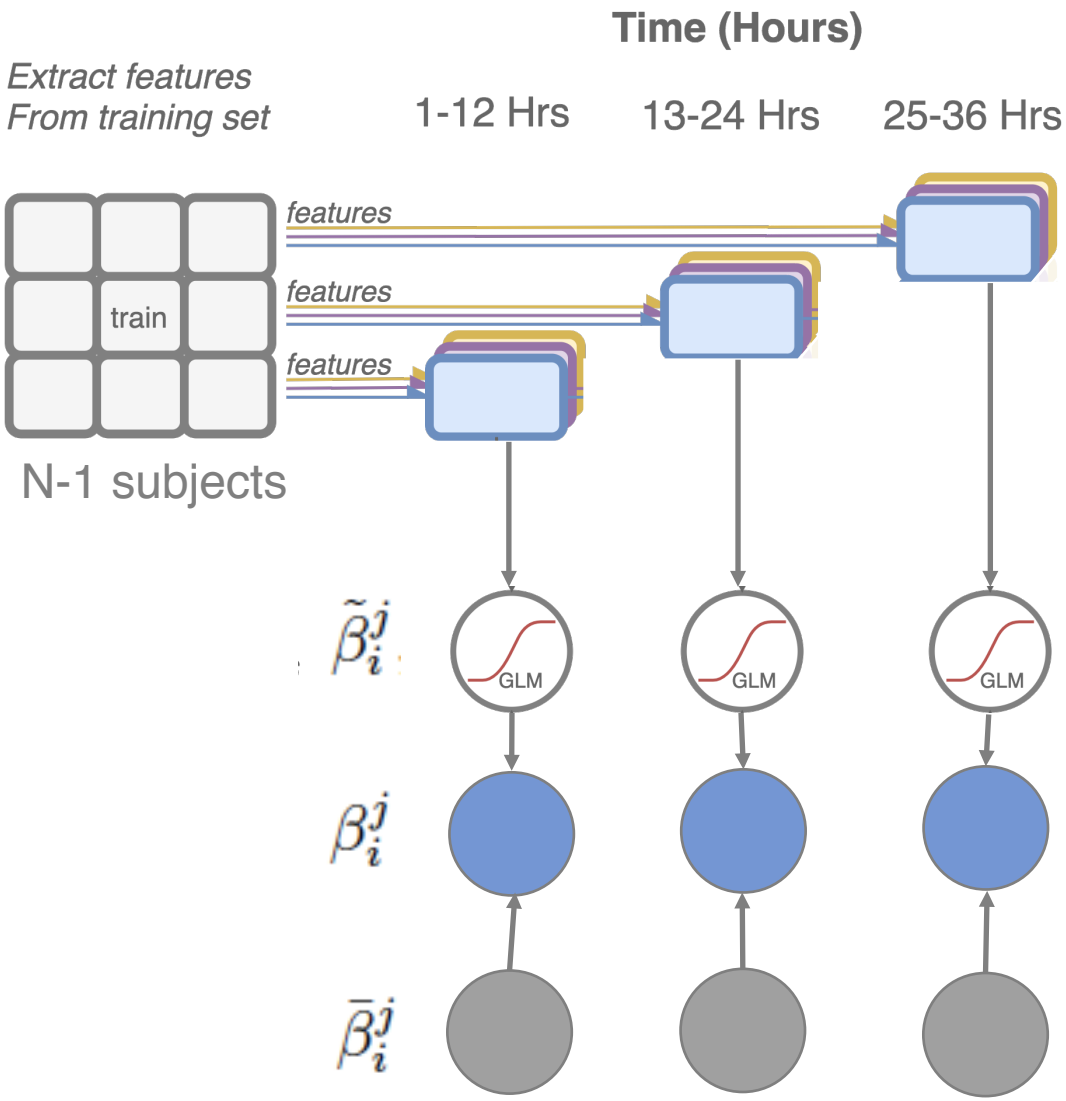
- A simple DBN (the kalman filter) recursively determine the optimal coefficient estimates

$$\beta_i^j = \bar{\beta}_i^j + K_i^j[\tilde{\beta}_i^j - \bar{\beta}_i^j]$$

$$K_{(i,j)} = \frac{\bar{P}_i^j}{\bar{P}_i^j + R_i^j}$$

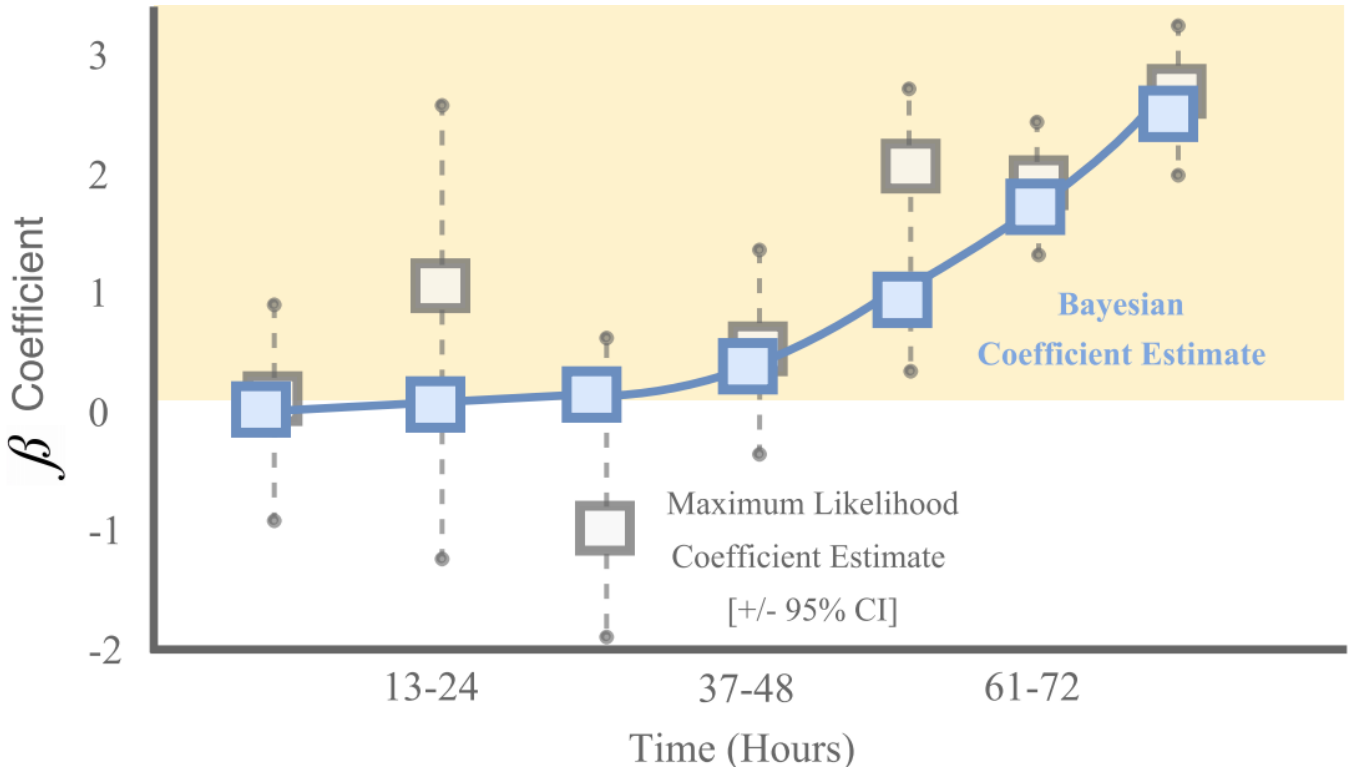
- With coefficient uncertainty

$$P_i^j = (1 - K_i^j)\bar{P}_i^{j-1}$$



Logistic Regression with *DBN*

- DBN estimates rely more strongly on prior estimates when present estimates have high standard error



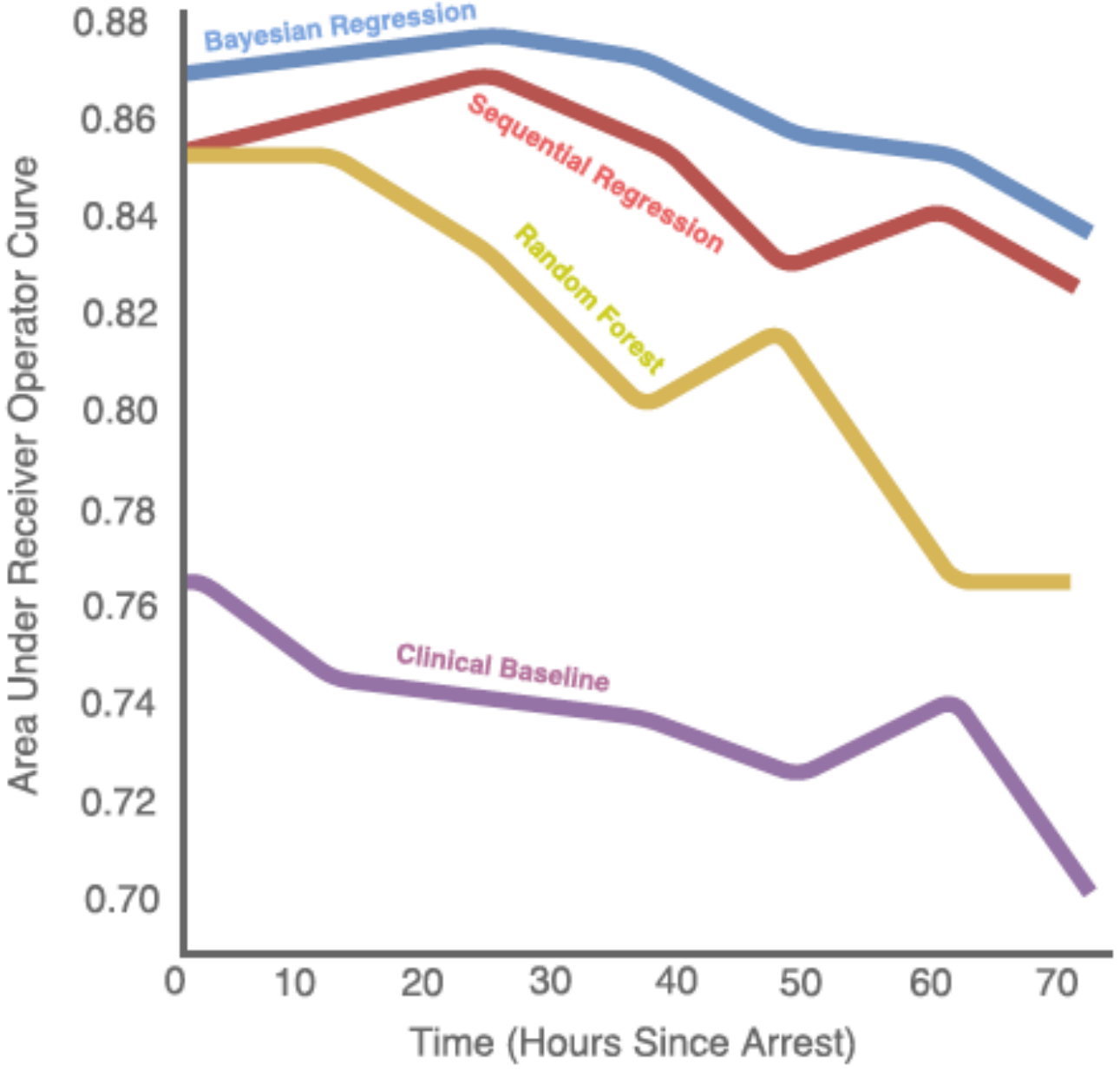
Results

The Sequential Logistic Regression with DBN had the best overall prognostic performance when compared to several baseline approaches.

Model	Features	Temporal Assumptions	AUC	TPR FPR=0.05	TPR FPR=0.01
Logistic Regression	5 Clinical, 10 EEG	Kalman Filter	0.86	0.49	0.29
Logistic Regression	5 Clinical, 10 EEG	Independent	0.84	0.46	0.26
Logistic Regression	10 EEG	Kalman Filter	0.83	0.47	0.25
Random Forest	5 EEG [2]	Independent	0.82	0.44	0.21
Logistic Regression	10 EEG	Independent	0.81	0.44	0.19
Logistic Regression	5 Clinical	None	0.73	0.26	0.11

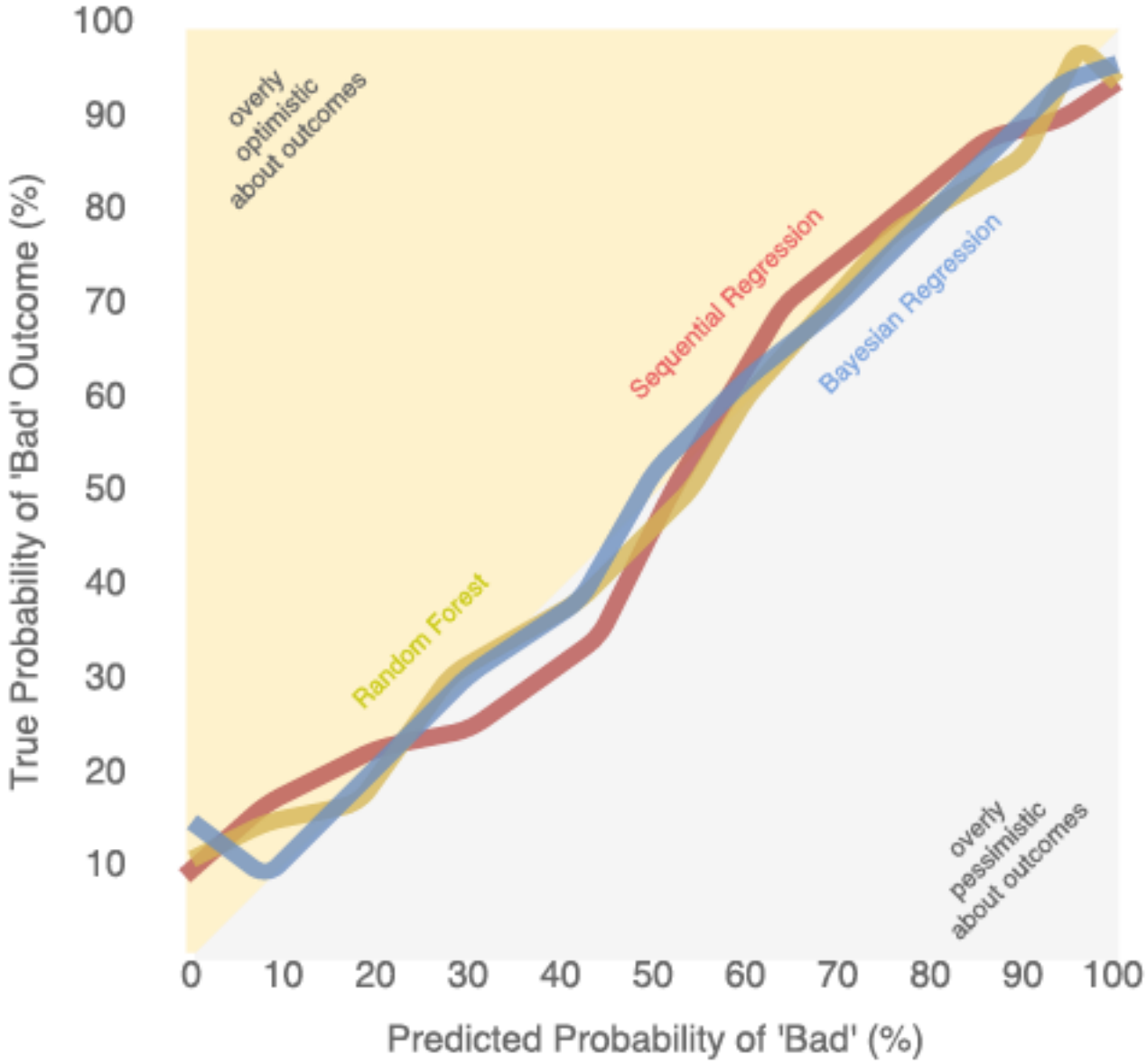
Results over time

- The DBN approach had the best performance over time (peak AUC: 0.87)
- performance decreases over time (min AUC: 0.84)



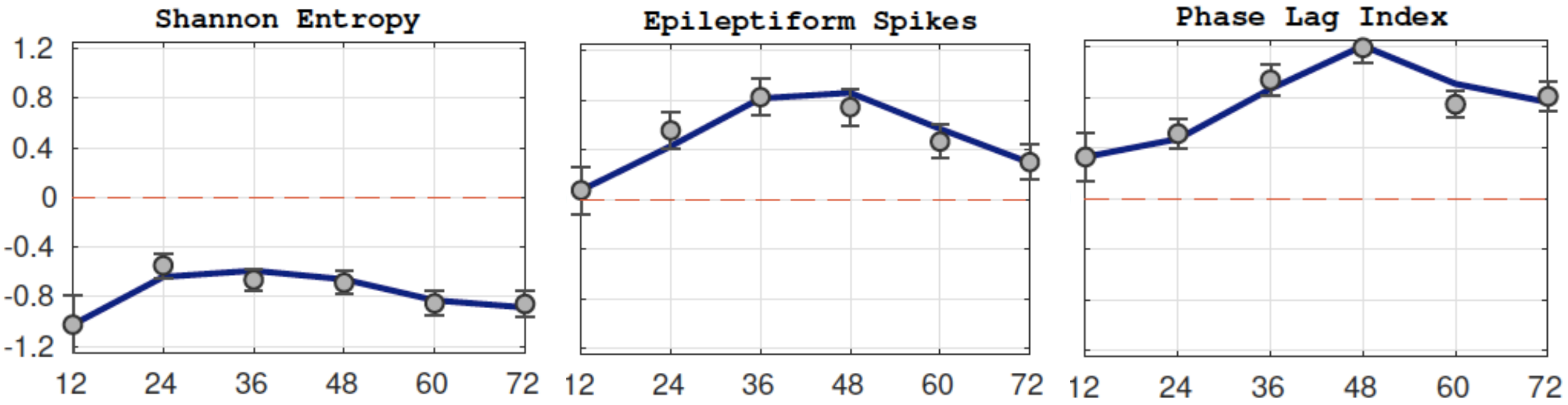
Calibration

- The DBN approach was the best calibrated of the tested approaches
- The direction of mis-calibration was '*optimistic*', tending to over-estimate '*good*' outcomes.



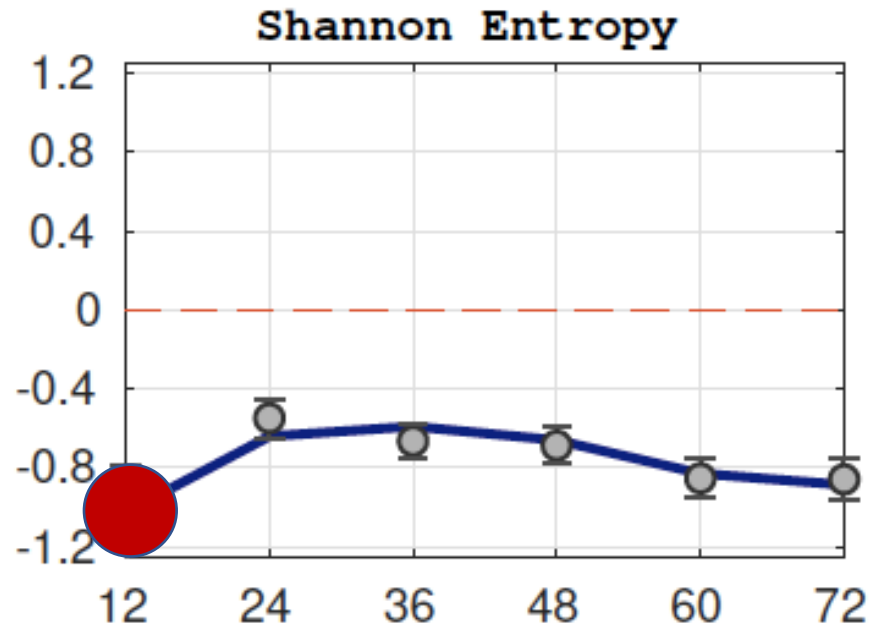
Feature coefficients over time

Model coefficients evolve over time in a sensible way, and are interpretable.



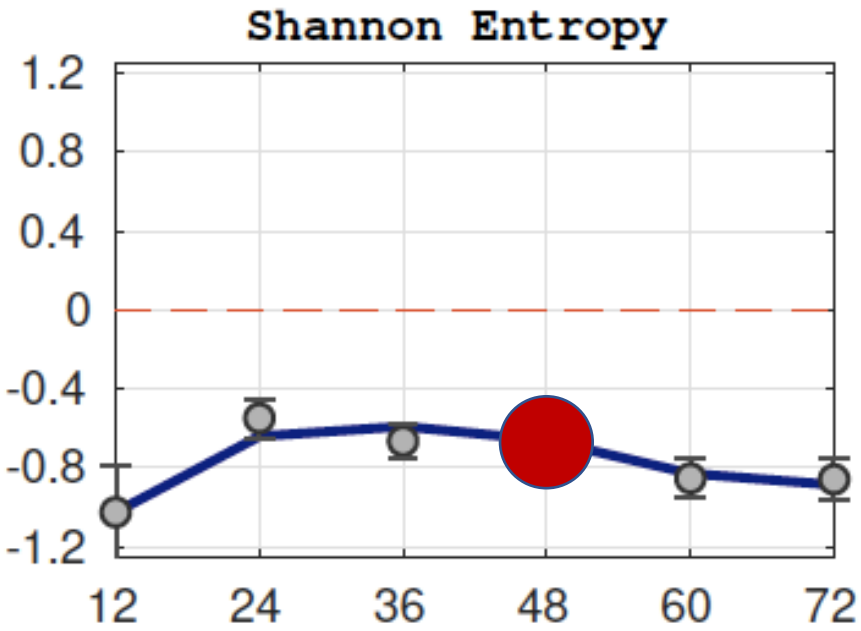
Feature coefficients over time

Entropy 1 std above the mean at 12 hours → 0.33 times as likely to have a bad outcome



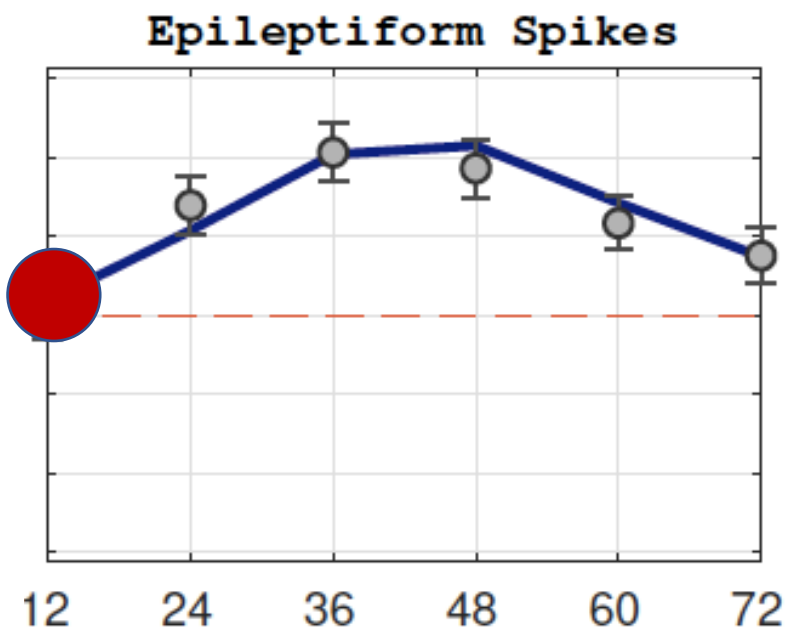
Feature coefficients over time

Entropy 1 std above the mean at 48 hours → 0.5 times as likely to have a bad outcome



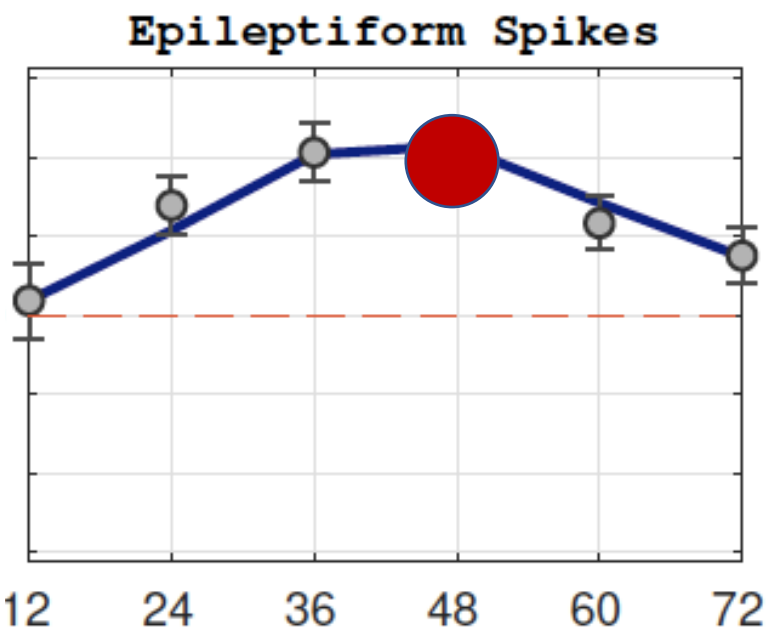
Feature coefficients over time

Epileptiforms 1 std below the mean at 12 hours → 1.1 times as likely to have a bad outcome



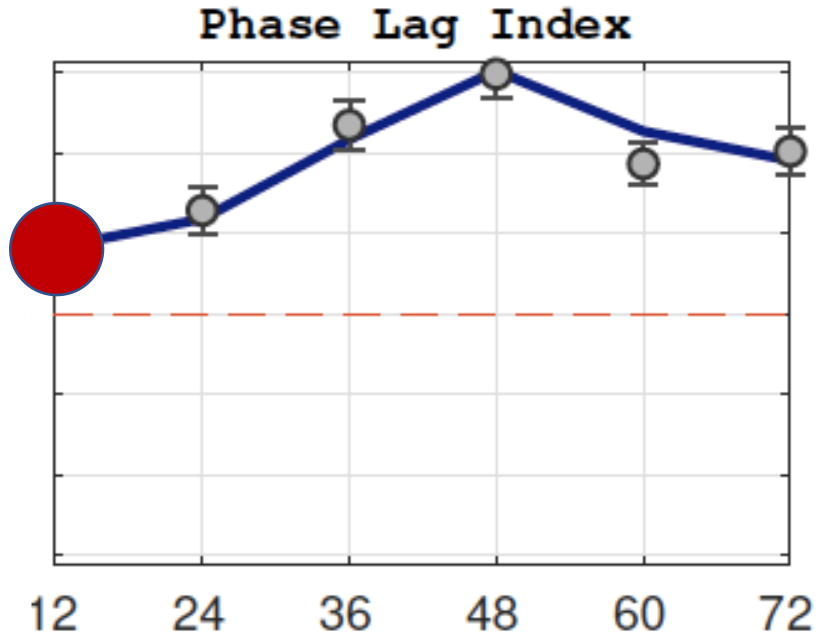
Feature coefficients over time

Epileptiforms 1 std below the mean at 48 hours → 2.2 times as likely to have a bad outcome



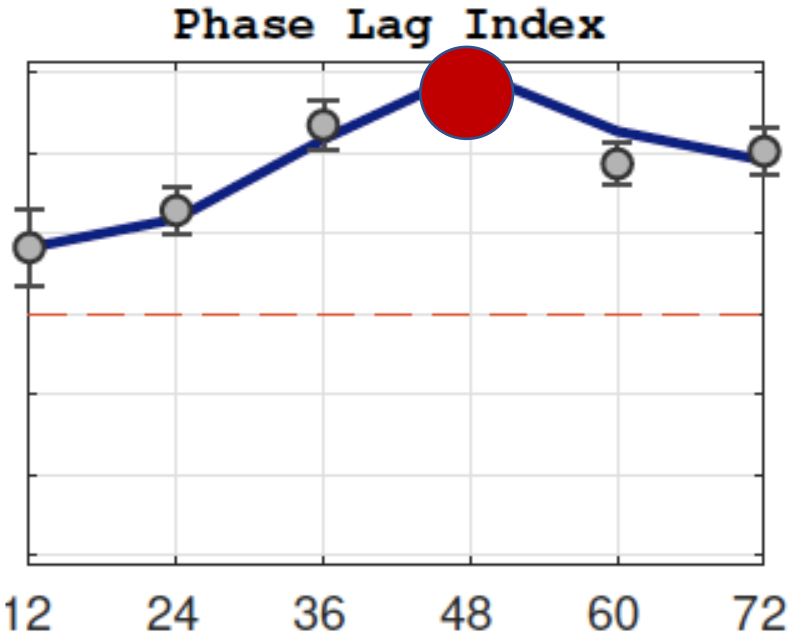
Feature coefficients over time

PLI 1 std below the mean at 12 hours → 1.5 times as likely to have a bad outcome



Feature coefficients over time

PLI 1 std below the mean at 48 hours → 3 times as likely to have a bad outcome



Pros/Cons of the Approach

- Advantages
  - Features are lightweight: 10 in total
  - Model is interpretable and may aid providers without formal deployment
  - Excellent LOOCV prognostication performance relative to baselines
- Disadvantages
  - Without ‘memories’, performance decreases over time
  - Possibility that important features were overlooked
  - Does not account for spatial information

Three Approaches

Feature Based:

- ~~1. Sequential Logistic Regression with ‘Elastic’ memories.~~

- ~~2. Sequential Logistic Regression with Dynamic Bayesian network constraints~~

Data Driven

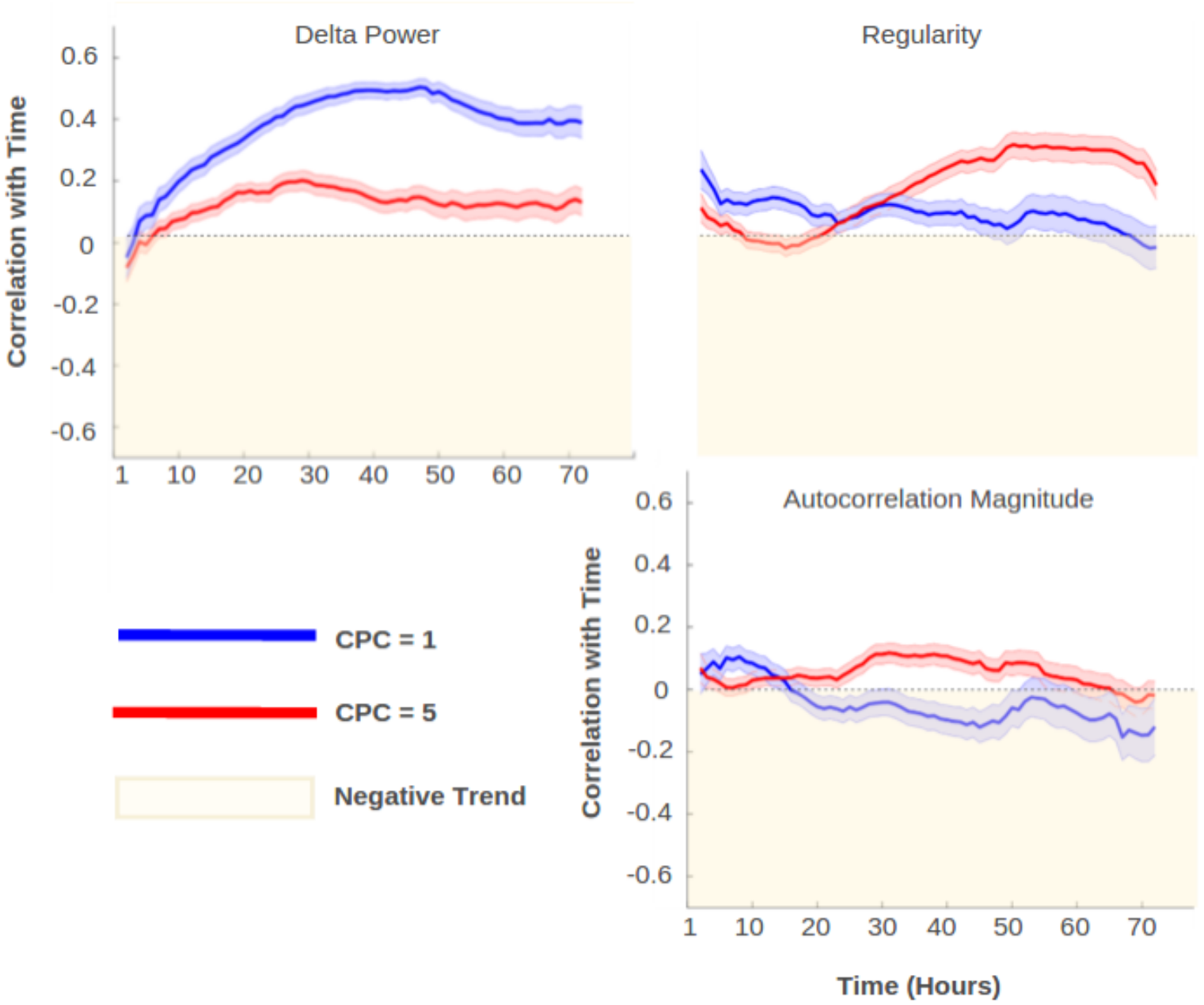
- 3. Deep neural networks

Limitations of the feature-based approaches

- Requires thoughtful pre-processing pipeline
- Requires computation of features
- By averaging over channels, features eliminate spatial information
- Possibility that certain features might have been missed

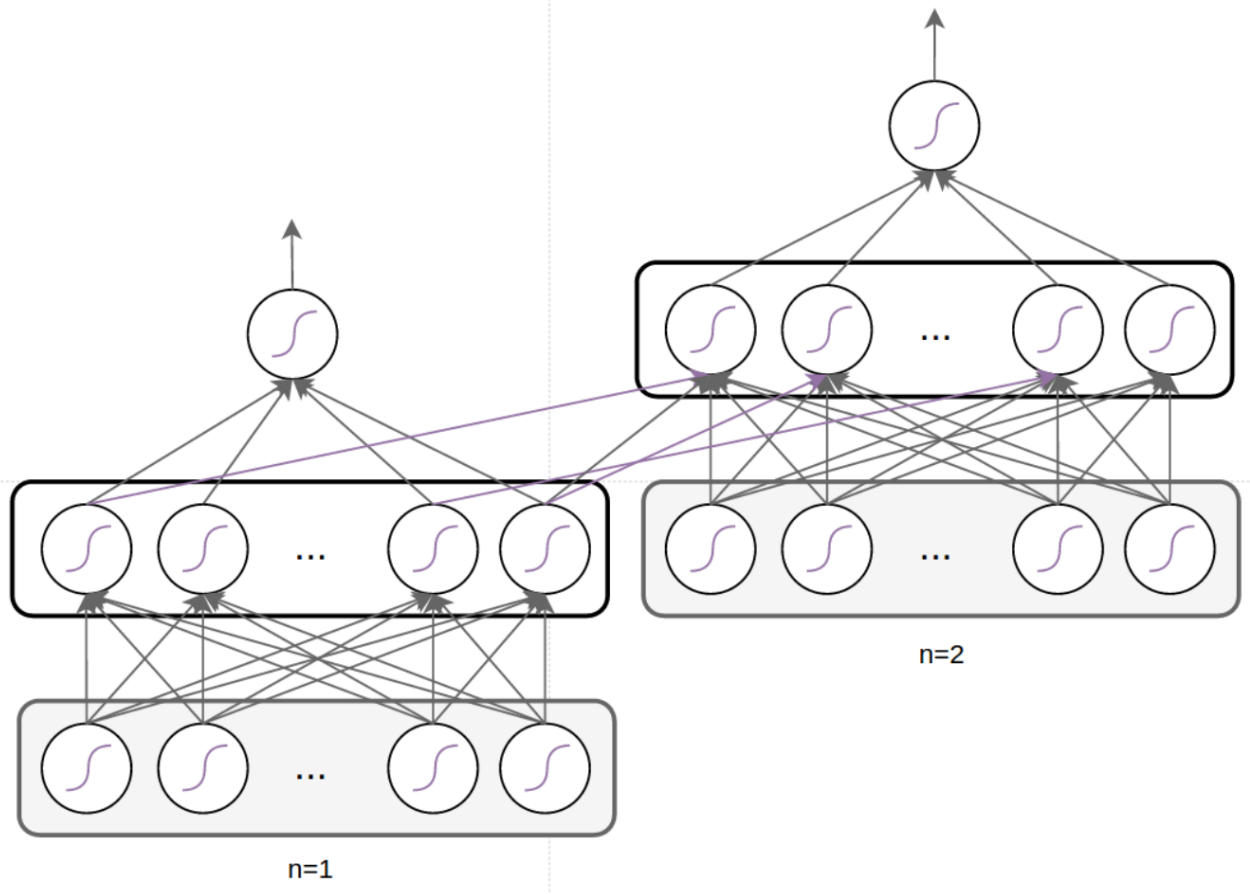
Motivation

- Even if all 57 features were used, **derivatives** of features may be useful features



Motivation

- Deep Neural Networks provide an potential solution to this problem because they derive features from the data automatically

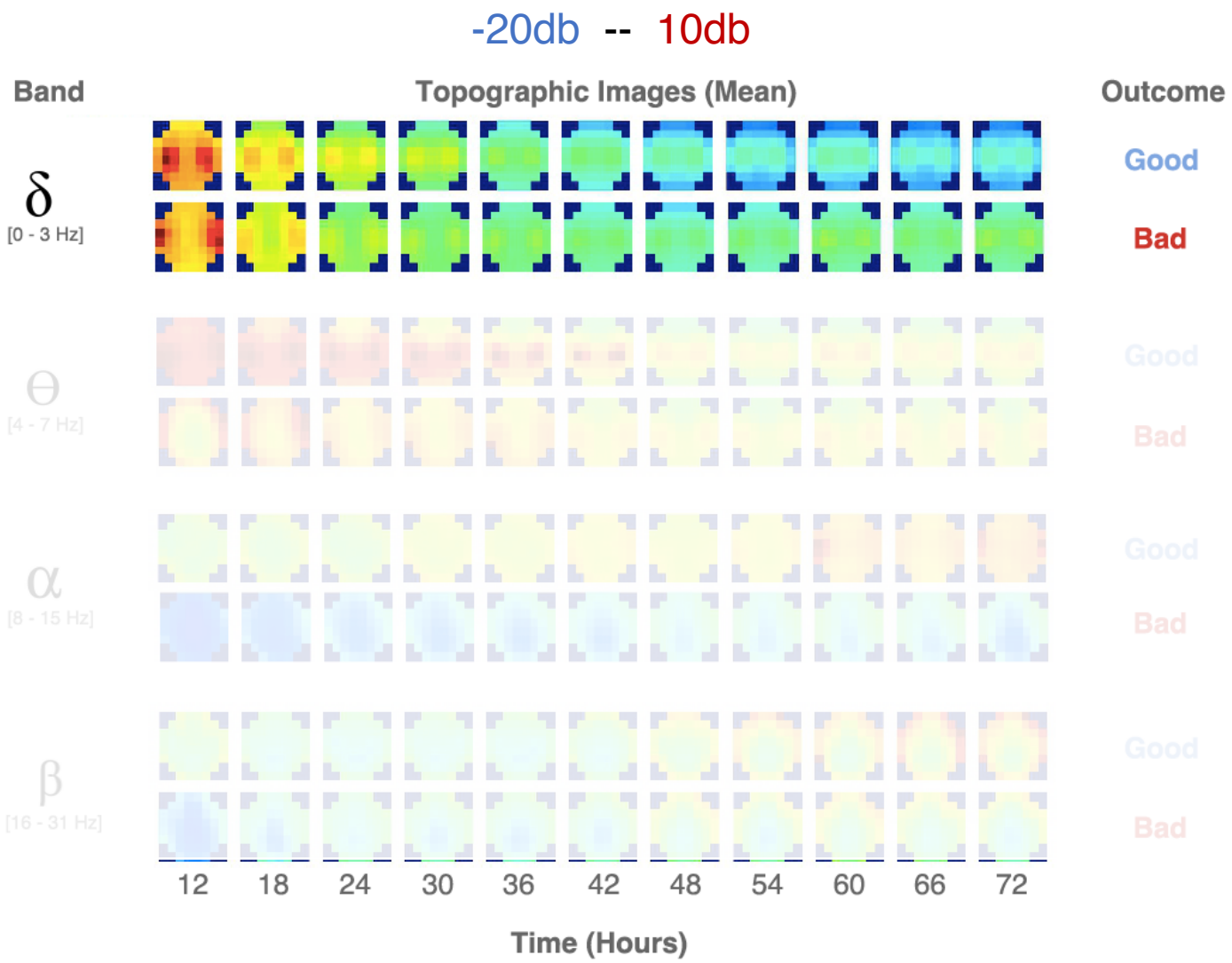


Deep Networks

- n = 724 patients
- 4 topographic energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out

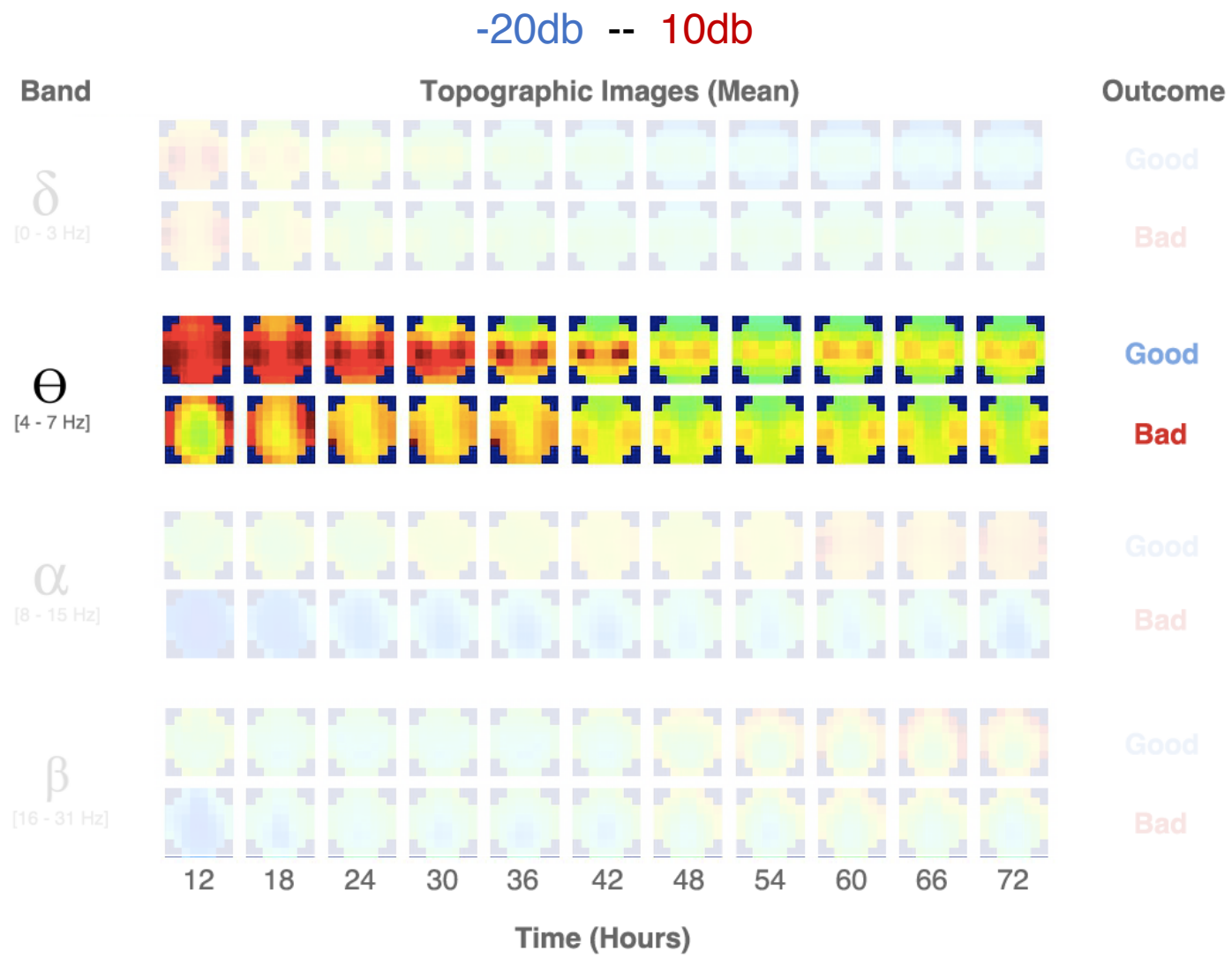
Deep Networks

- n = 724 patients
- 4 topographic energy plots: [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation: Leave-one-institution-out



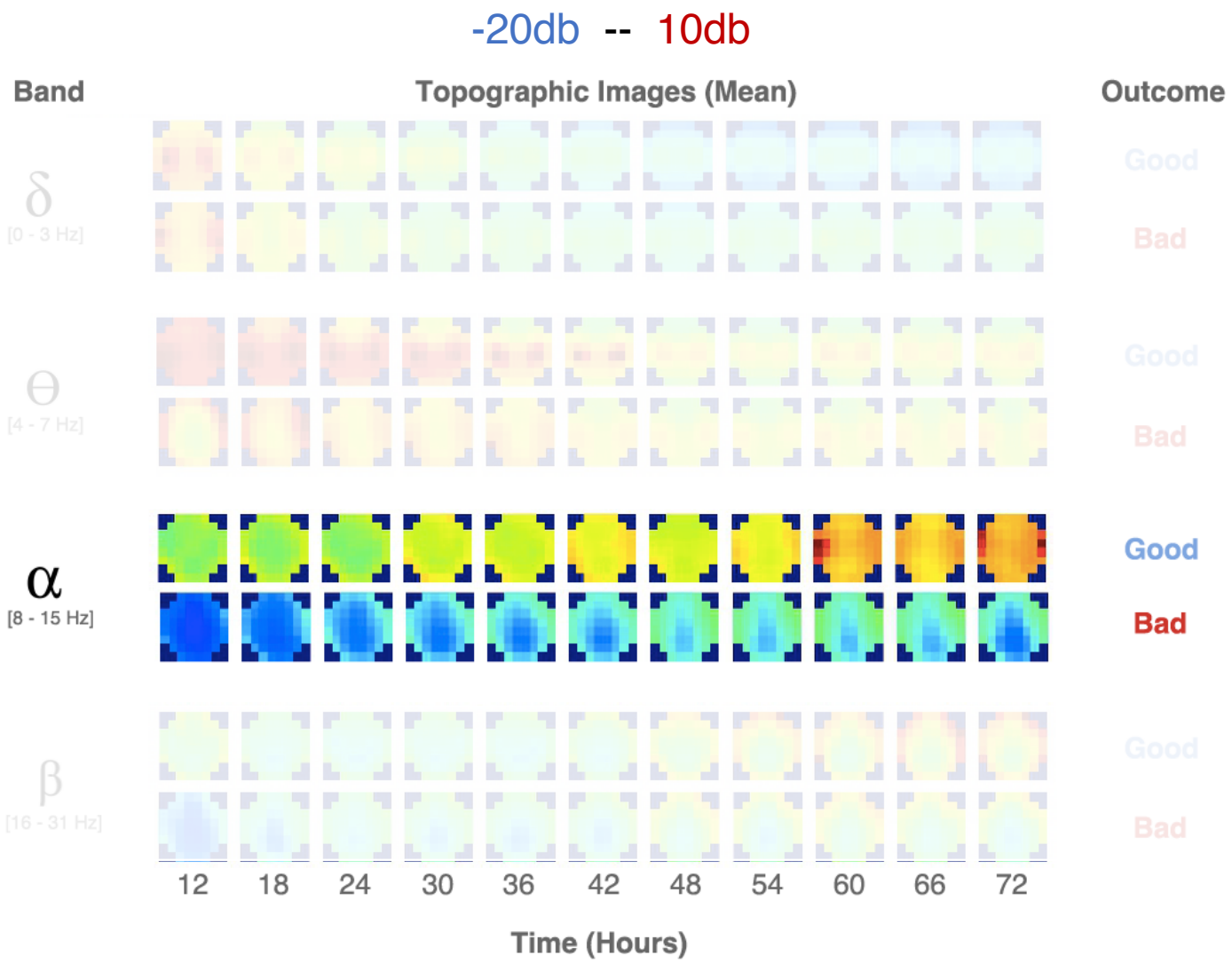
Deep Networks

- n = 724 patients
- 4 topographic energy plots: [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation: Leave-one-institution-out



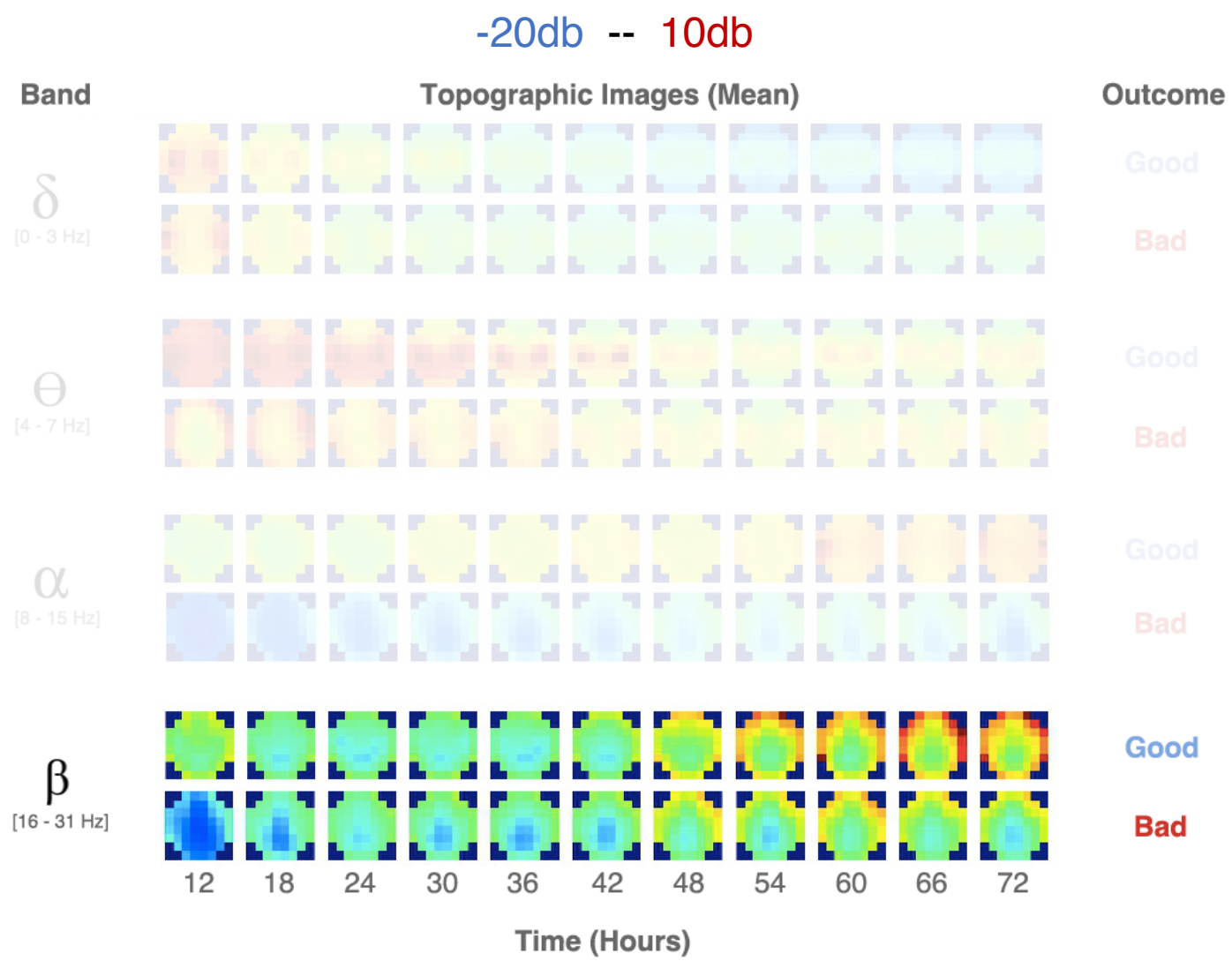
Deep Networks

- n = 724 patients
- 4 topographic energy plots: [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation: Leave-one-institution-out



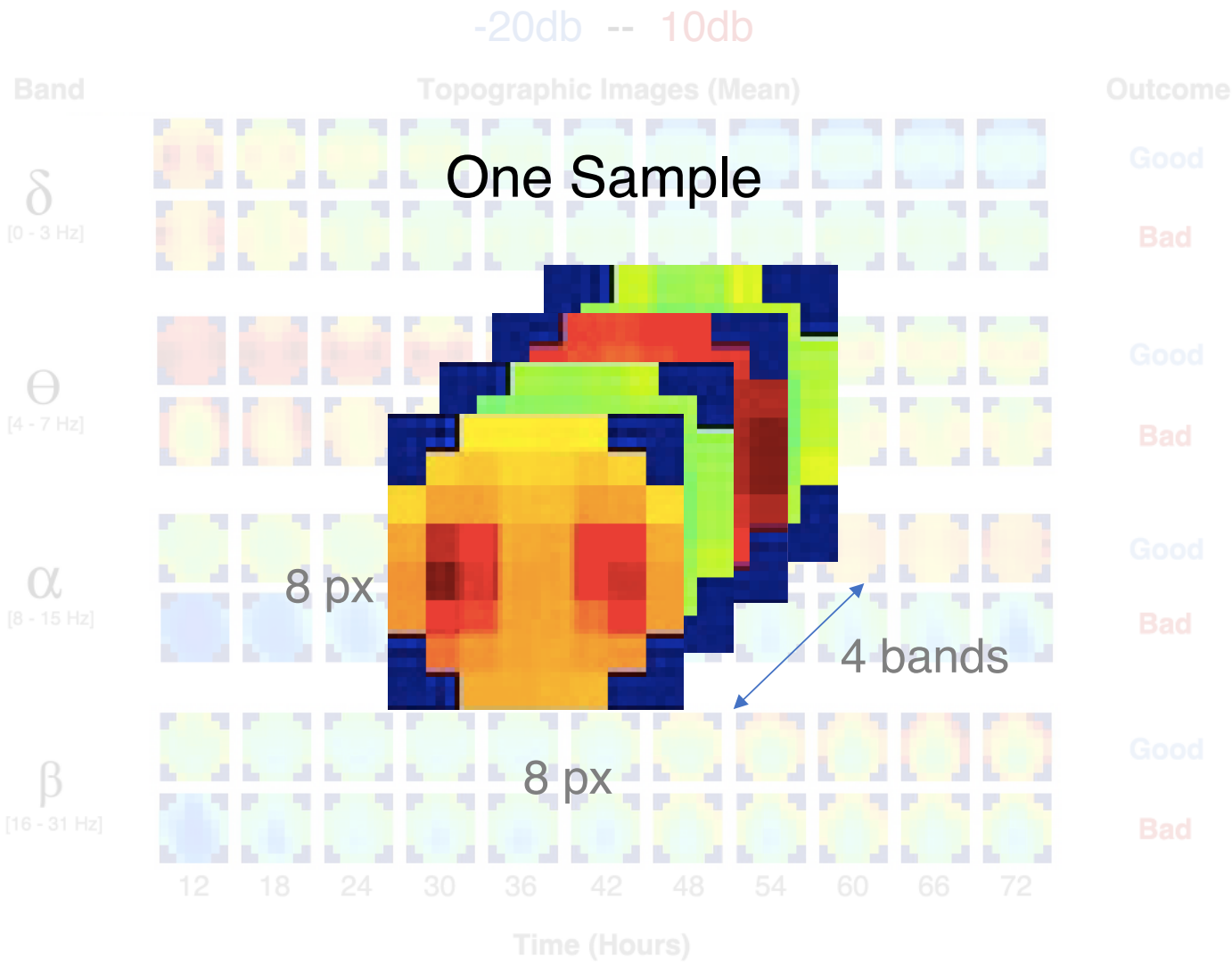
Deep Networks

- n = 724 patients
- 4 topographic energy plots: [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation: Leave-one-institution-out



Deep Networks

- n = 724 patients
- 4 topographic energy plots: [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation: Leave-one-institution-out



Deep Networks

- n = 724 patients
- 4 topographic energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out

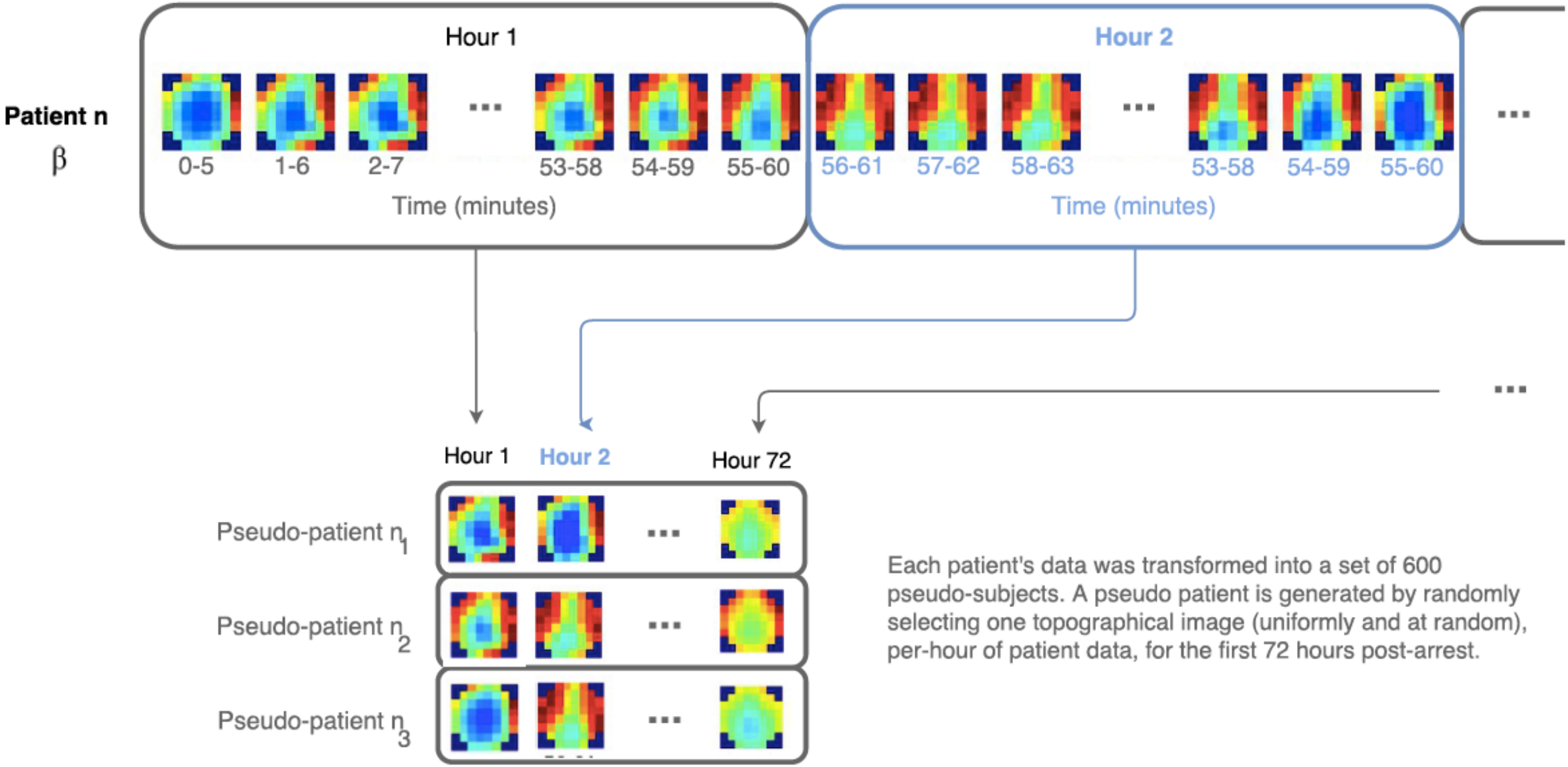


Figure 8.3: **Pseudo-patient Generation.**

Figure 8.3: **Pseudo-patient Generation.**

Figure 8.3: **Pseudo-patient Generation.**

Figure 8.3: **Pseudo-patient Generation.**

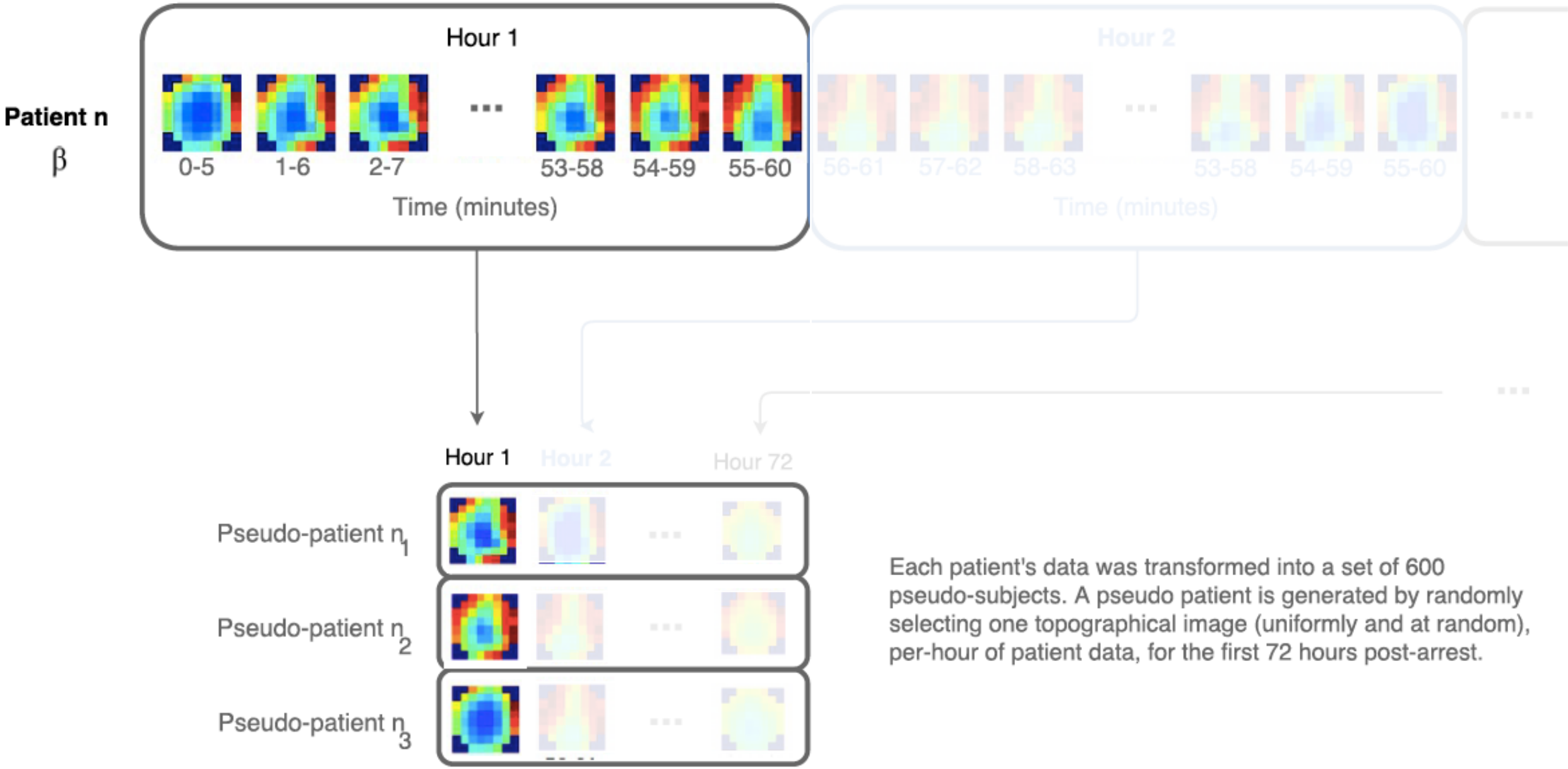


Figure 8.3: **Pseudo-patient Generation.**

Figure 8.3: **Pseudo-patient Generation.**

Figure 8.3: **Pseudo-patient Generation.**

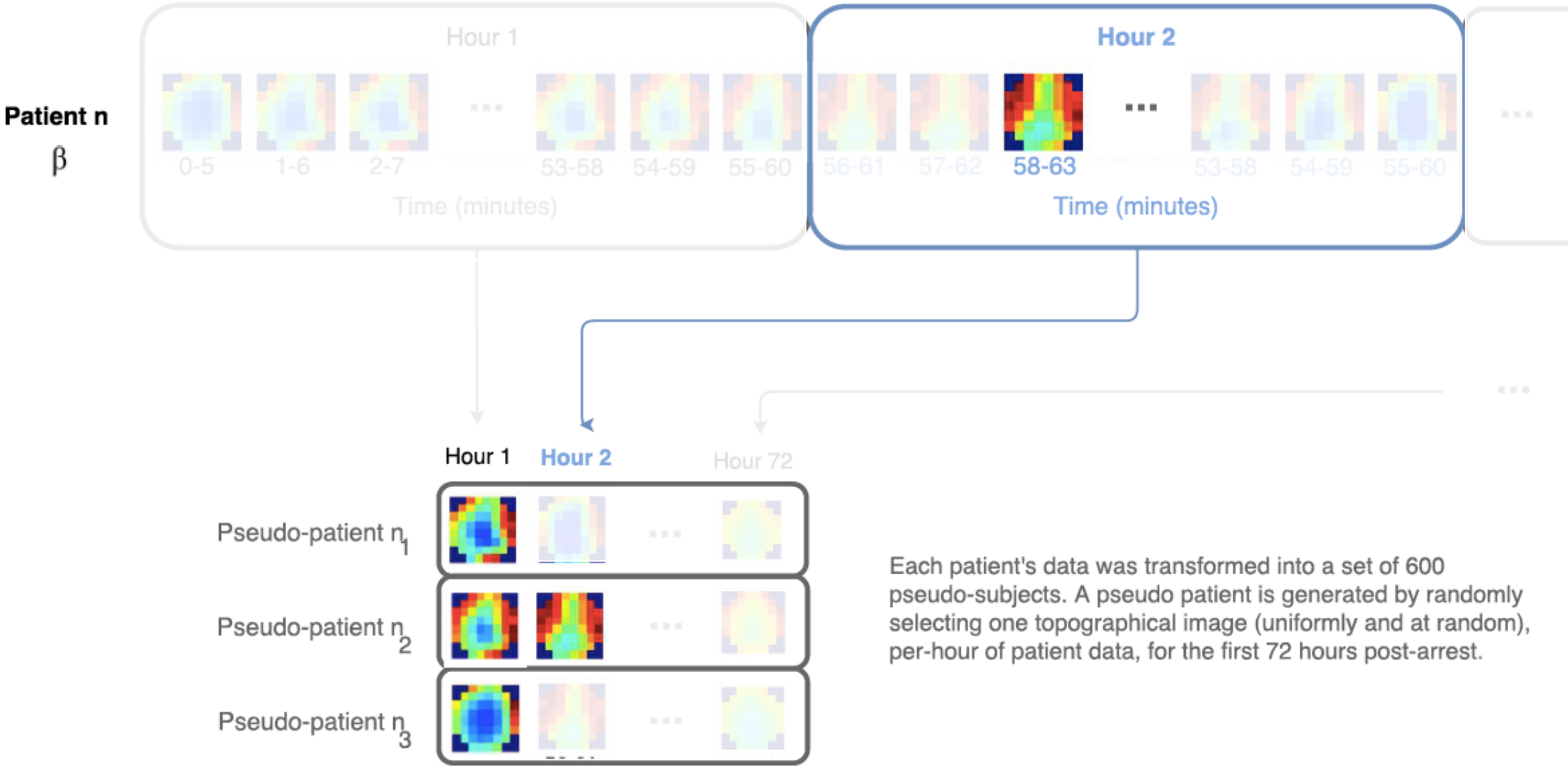


Figure 8.3: **Pseudo-patient Generation.**

Figure 8.3: **Pseudo-patient Generation.**

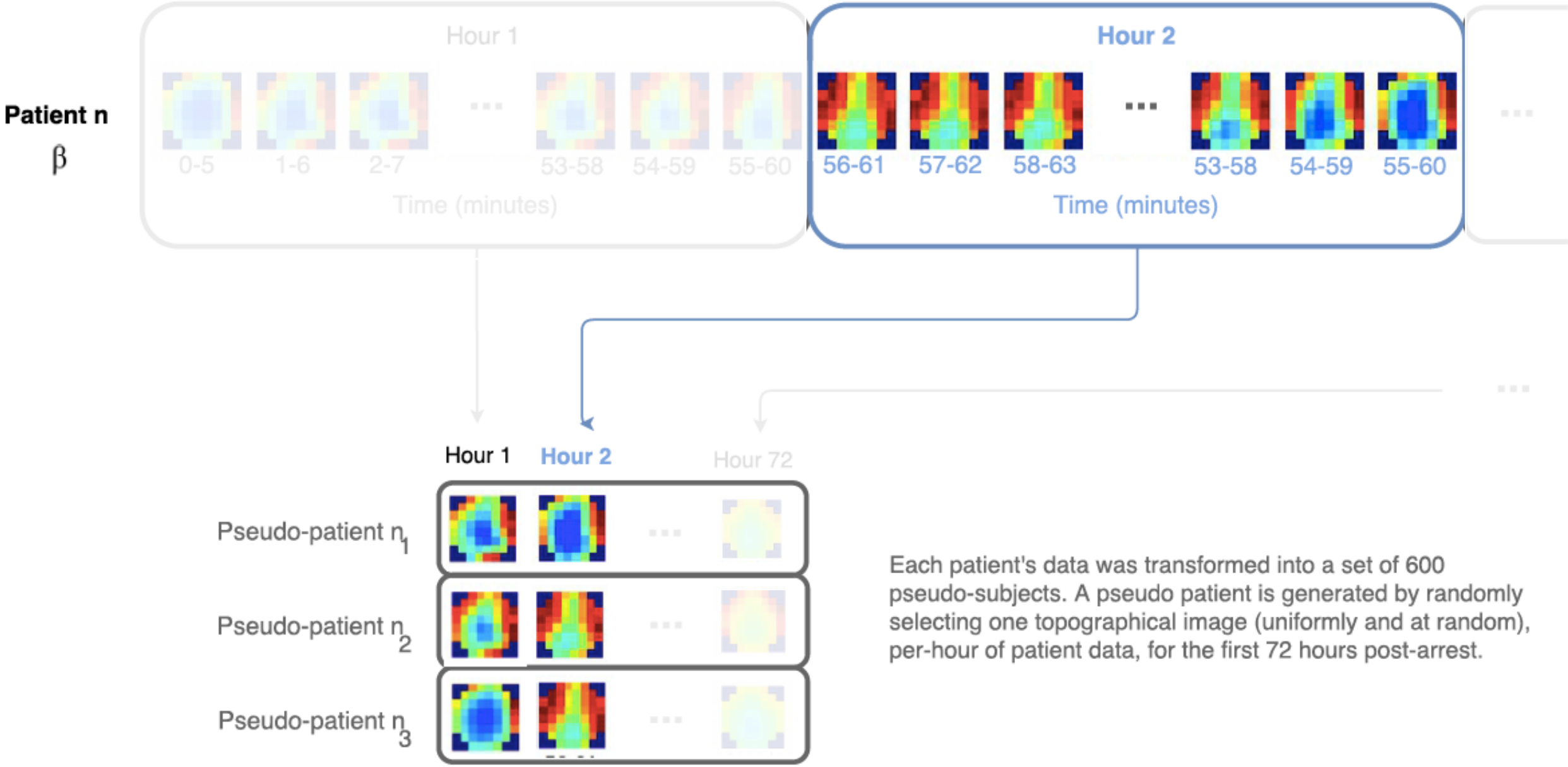


Figure 8.3: **Pseudo-patient Generation.**

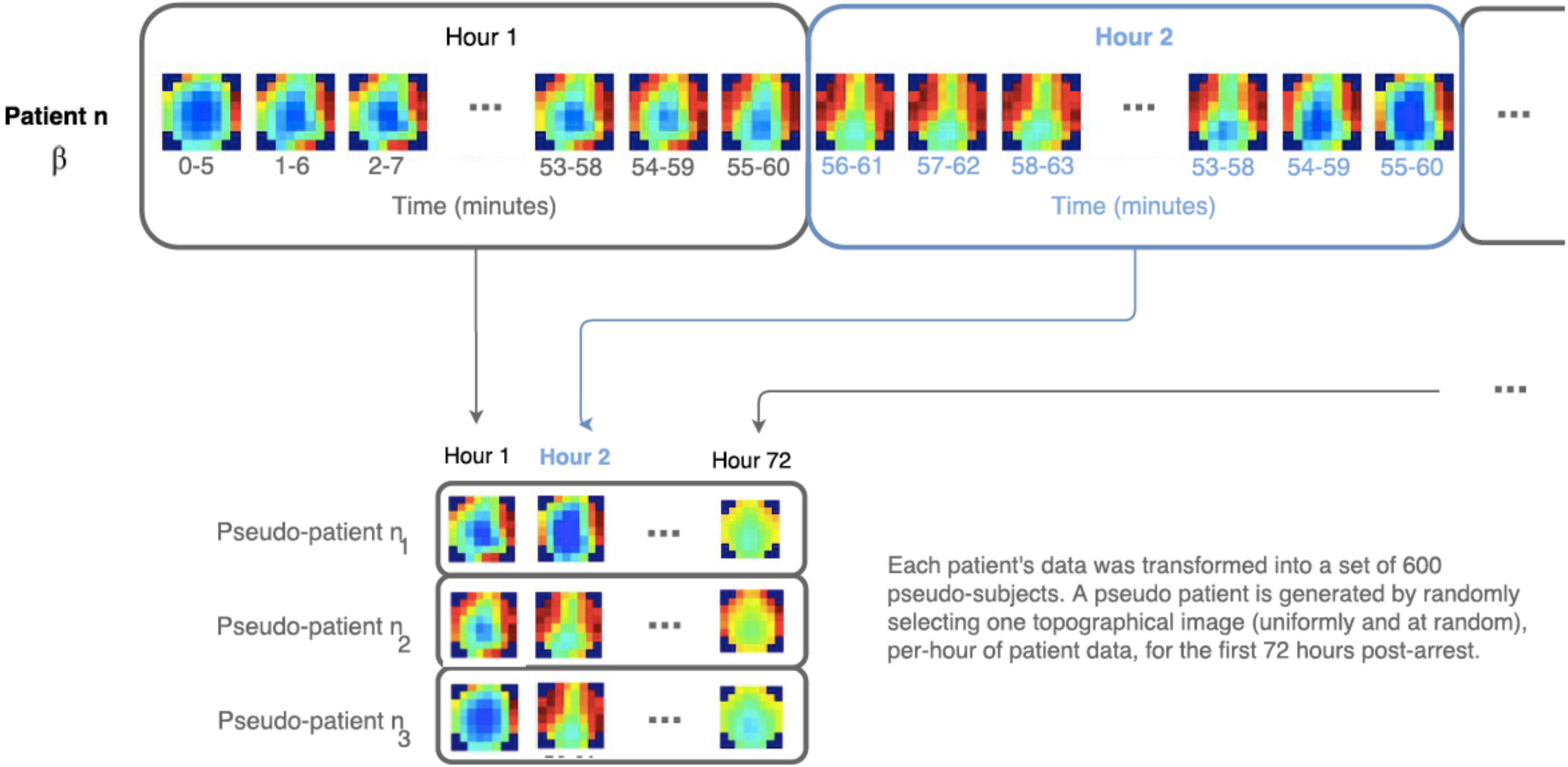
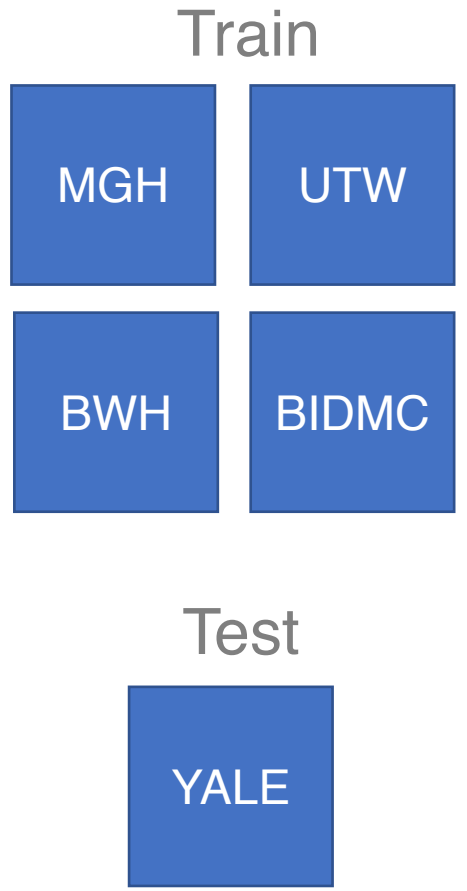


Figure 8.3: **Pseudo-patient Generation.**

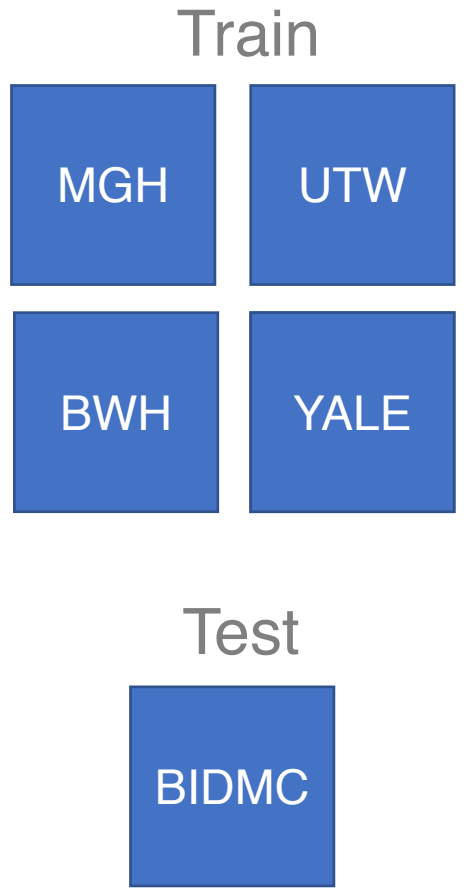
Deep Networks

- n = 724 patients
- 4 energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out



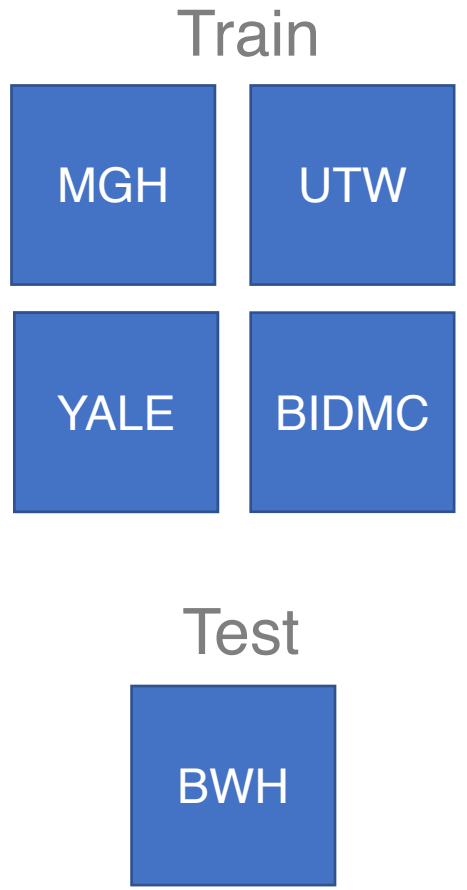
Deep Networks

- n = 724 patients
- 4 energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out



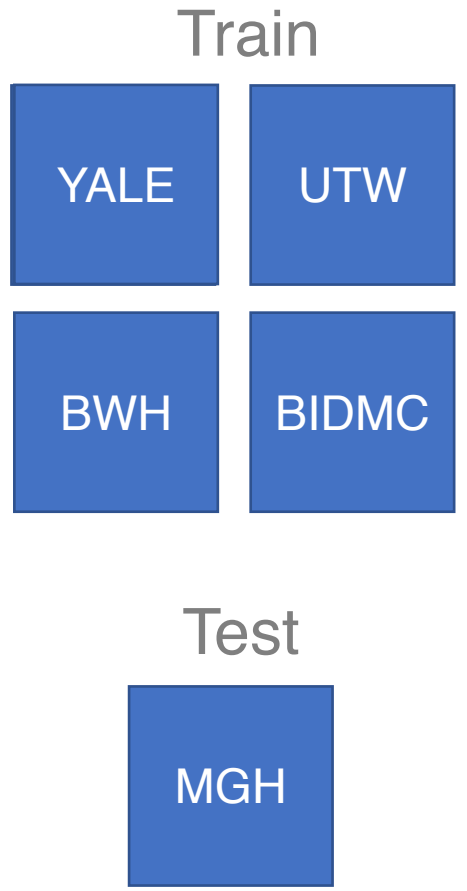
Deep Networks

- n = 724 patients
- 4 energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out



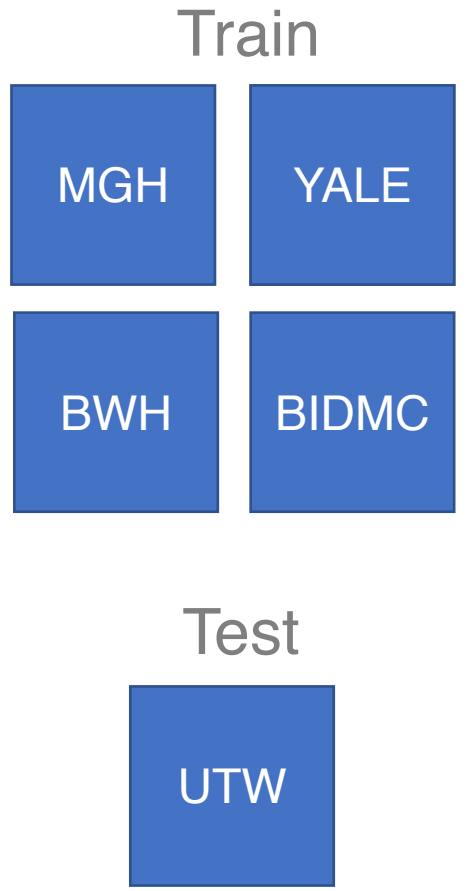
Deep Networks

- n = 724 patients
- 4 energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out



Deep Networks

- n = 724 patients
- 4 energy plots:
  [0-3, 4-7, 8-15, 16-31] Hz
- Pseudo-subject generation
- Validation:
  Leave-one-institution-out

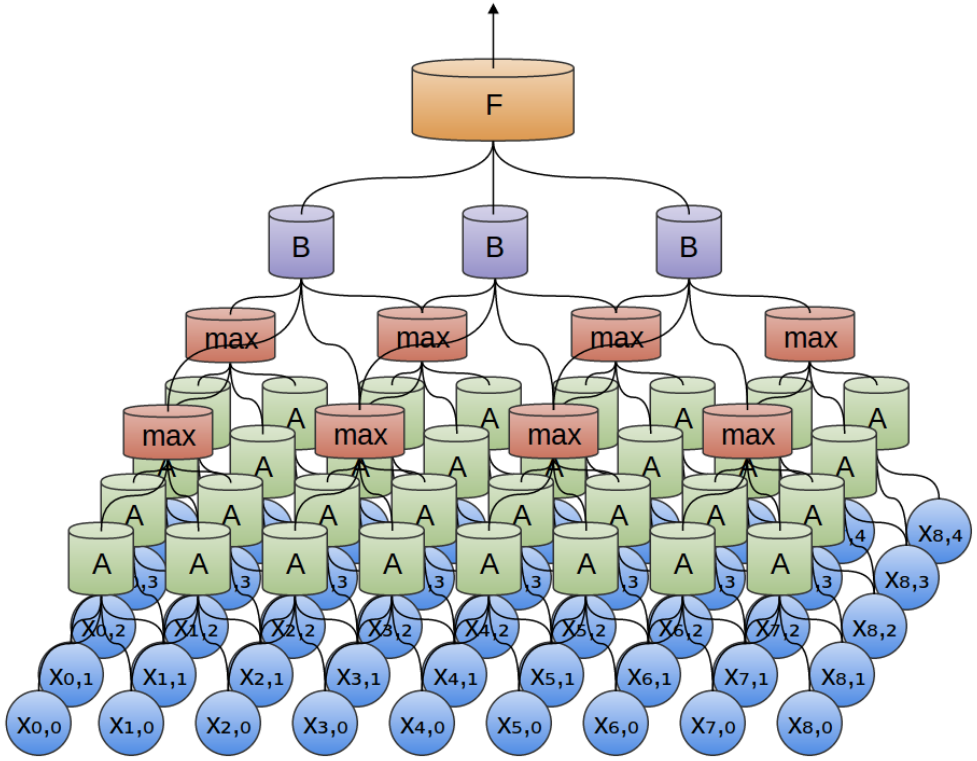


Deep Networks

- 300 topologies were tested
- 100, 2 Dimensional Convolutional Neural Networks
- 100, 3 Dimensional Convolutional Neural Networks
- 100, Long-term Recurrent Convolutional Networks

Deep Networks

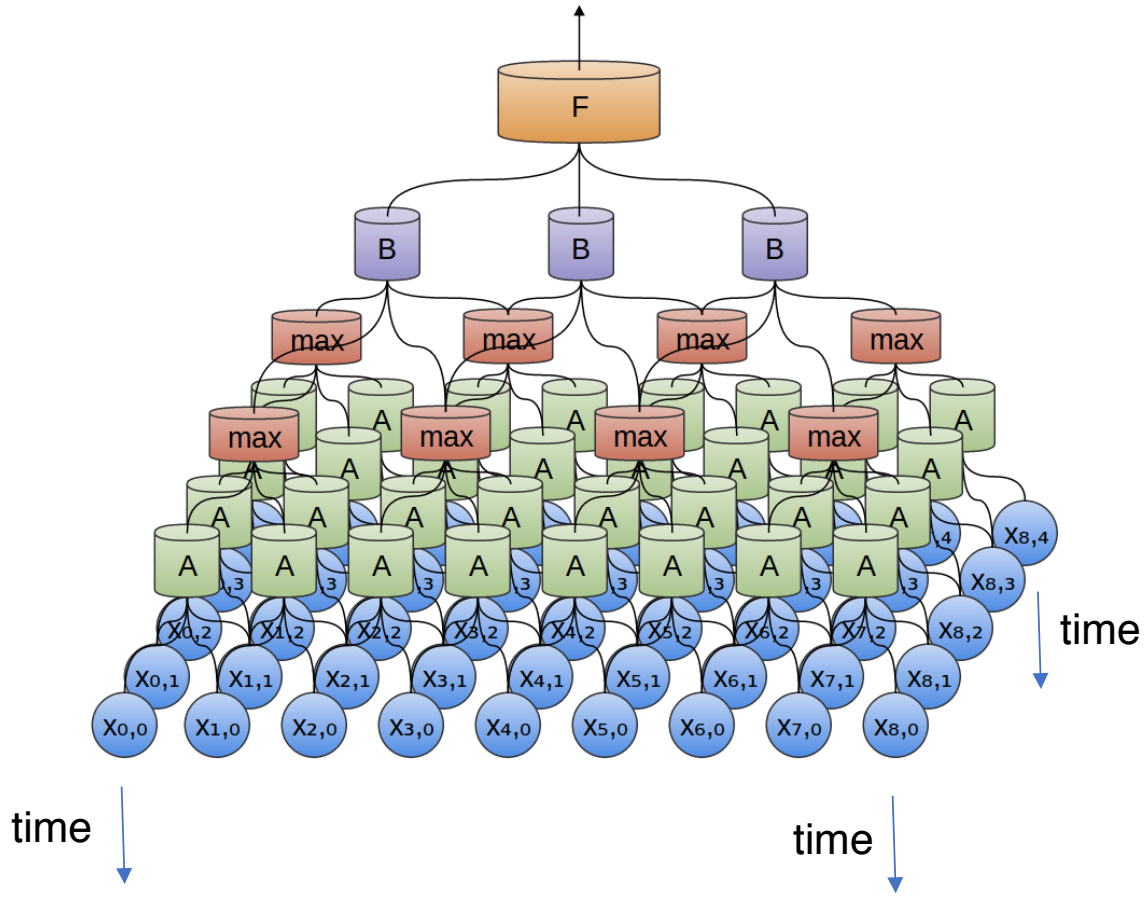
- 300 topologies were tested
- 100, 2 Dimensional Convolutional Neural Networks
- 100, 3 Dimensional Convolutional Neural Networks
- 100, Long-term Recurrent Convolutional Networks



Source: http://alifar76.github.io/cnn-microbiome/

Deep Networks

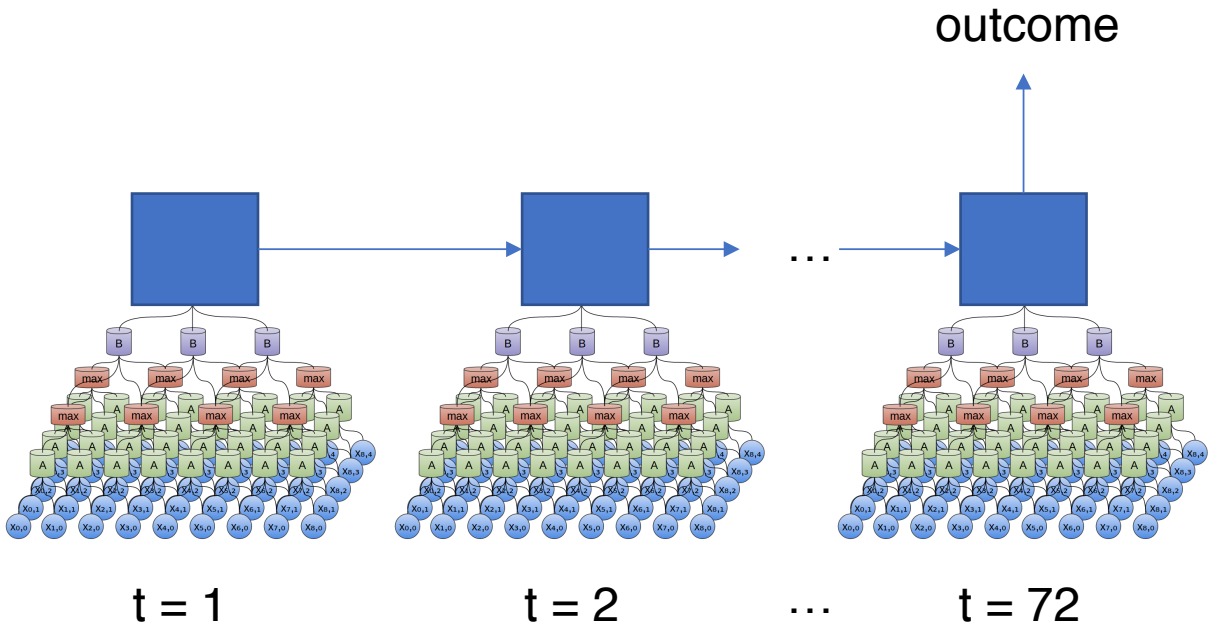
- 300 topologies were tested
- 100, 2 Dimensional Convolutional Neural Networks
- 100, 3 Dimensional Convolutional Neural Networks
- 100, Long-term Recurrent Convolutional Networks



Source: http://alifar76.github.io/cnn-microbiome/

Deep Networks

- 300 topologies were tested
- 100, 2 Dimensional Convolutional Neural Networks
- 100, 3 Dimensional Convolutional Neural Networks
- 100, Long-term Recurrent Convolutional Networks



Results

- 3D CNNs provide performance on par with the best-performing feature-based approaches.

	AUROC	TPR| FPR=0	TPR| FPR=1%	TPR| FPR=5%
Neural Networks				
2D CNN	0.71 (0.09)	0 (0.0)	0.09 (0.06)	0.23 (0.10)
**3D CNN	0.81 (0.04)	0.24 (0.05)	0.30 (0.05)	0.45 (0.09)
LRCN	0.79 (0.08)	0.19 (0.11)	0.34 (0.15)	0.48 (0.15)
SVM				
*Linear	0.76 (0.13)	0.17 (0.18)	0.21 (0.16)	0.32 (0.19)
Quadratic	0.72 (0.14)	0.14 (0.09)	0.16 (0.09)	0.26 (0.12)
Fine Gaussian	0.75 (0.14)	0.18 (0.14)	0.2 (0.12)	0.31 (0.19)
Medium Gaussian	0.56 (0.03)	0.01 (0.02)	0.01 (0.02)	0.06 (0.04)
Coarse Gaussian	0.71 (0.13)	0.07 (0.12)	0.08 (0.12)	0.15 (0.1)
Discriminant Analysis				
Linear	0.77 (0.15)	0.21 (0.13)	0.23 (0.14)	0.34 (0.25)
*Quadratic	0.78 (0.14)	0.22 (0.13)	0.25 (0.14)	0.4 (0.22)
Cubic	0.72 (0.13)	0 (0)	0 (0)	0.15 (0.24)
Decision Trees				
Simple	0.68 (0.11)	0 (0)	0 (0)	0 (0)
Medium	0.7 (0.09)	0.05 (0.11)	0.05 (0.11)	0.11 (0.16)
Complex	0.69 (0.07)	0 (0)	0 (0)	0 (0)
*RUS Boosted	0.8 (0.1)	0.25 (0.2)	0.25 (0.2)	0.39 (0.17)
Ensemble Boosted	0.74 (0.12)	0.12 (0.17)	0.12 (0.17)	0.14 (0.19)
Ensemble Bagged	0.77 (0.1)	0.09 (0.14)	0.11 (0.13)	0.22 (0.15)

Results

- 3D CNNs provide performance on par with the best-performing feature-based approaches.

	AUROC	TPR| FPR=0	TPR| FPR=1%	TPR| FPR=5%
Neural Networks				
2D CNN	0.71 (0.09)	0 (0.0)	0.09 (0.06)	0.23 (0.10)
**3D CNN	0.81 (0.04)	0.24 (0.05)	0.30 (0.05)	0.45 (0.09)
LRCN	0.79 (0.08)	0.19 (0.11)	0.34 (0.15)	0.48 (0.15)
SVM				
*Linear	0.76 (0.13)	0.17 (0.18)	0.21 (0.16)	0.32 (0.19)
Quadratic	0.72 (0.14)	0.14 (0.09)	0.16 (0.09)	0.26 (0.12)
Fine Gaussian	0.75 (0.14)	0.18 (0.14)	0.2 (0.12)	0.31 (0.19)
Medium Gaussian	0.56 (0.03)	0.01 (0.02)	0.01 (0.02)	0.06 (0.04)
Coarse Gaussian	0.71 (0.13)	0.07 (0.12)	0.08 (0.12)	0.15 (0.1)
Discriminant Analysis				
Linear	0.77 (0.15)	0.21 (0.13)	0.23 (0.14)	0.34 (0.25)
*Quadratic	0.78 (0.14)	0.22 (0.13)	0.25 (0.14)	0.4 (0.22)
Cubic	0.72 (0.13)	0 (0)	0 (0)	0.15 (0.24)
Decision Trees				
Simple	0.68 (0.11)	0 (0)	0 (0)	0 (0)
Medium	0.7 (0.09)	0.05 (0.11)	0.05 (0.11)	0.11 (0.16)
Complex	0.69 (0.07)	0 (0)	0 (0)	0 (0)
*RUS Boosted	0.8 (0.1)	0.25 (0.2)	0.25 (0.2)	0.39 (0.17)
Ensemble Boosted	0.74 (0.12)	0.12 (0.17)	0.12 (0.17)	0.14 (0.19)
Ensemble Bagged	0.77 (0.1)	0.09 (0.14)	0.11 (0.13)	0.22 (0.15)

Results

- 3D CNNs provide performance on par with the best-performing feature-based approaches.

	AUROC	TPR| FPR=0	TPR| FPR=1%	TPR| FPR=5%
Neural Networks				
2D CNN	0.71 (0.09)	0 (0.0)	0.09 (0.06)	0.23 (0.10)
**3D CNN	0.81 (0.04)	0.24 (0.05)	0.30 (0.05)	0.45 (0.09)
LRCN	0.79 (0.08)	0.19 (0.11)	0.34 (0.15)	0.48 (0.15)
SVM				
*Linear	0.76 (0.13)	0.17 (0.18)	0.21 (0.16)	0.32 (0.19)
Quadratic	0.72 (0.14)	0.14 (0.09)	0.16 (0.09)	0.26 (0.12)
Fine Gaussian	0.75 (0.14)	0.18 (0.14)	0.2 (0.12)	0.31 (0.19)
Medium Gaussian	0.56 (0.03)	0.01 (0.02)	0.01 (0.02)	0.06 (0.04)
Coarse Gaussian	0.71 (0.13)	0.07 (0.12)	0.08 (0.12)	0.15 (0.1)
Discriminant Analysis				
Linear	0.77 (0.15)	0.21 (0.13)	0.23 (0.14)	0.34 (0.25)
*Quadratic	0.78 (0.14)	0.22 (0.13)	0.25 (0.14)	0.4 (0.22)
Cubic	0.72 (0.13)	0 (0)	0 (0)	0.15 (0.24)
Decision Trees				
Simple	0.68 (0.11)	0 (0)	0 (0)	0 (0)
Medium	0.7 (0.09)	0.05 (0.11)	0.05 (0.11)	0.11 (0.16)
Complex	0.69 (0.07)	0 (0)	0 (0)	0 (0)
*RUS Boosted	0.8 (0.1)	0.25 (0.2)	0.25 (0.2)	0.39 (0.17)
Ensemble Boosted	0.74 (0.12)	0.12 (0.17)	0.12 (0.17)	0.14 (0.19)
Ensemble Bagged	0.77 (0.1)	0.09 (0.14)	0.11 (0.13)	0.22 (0.15)

Results

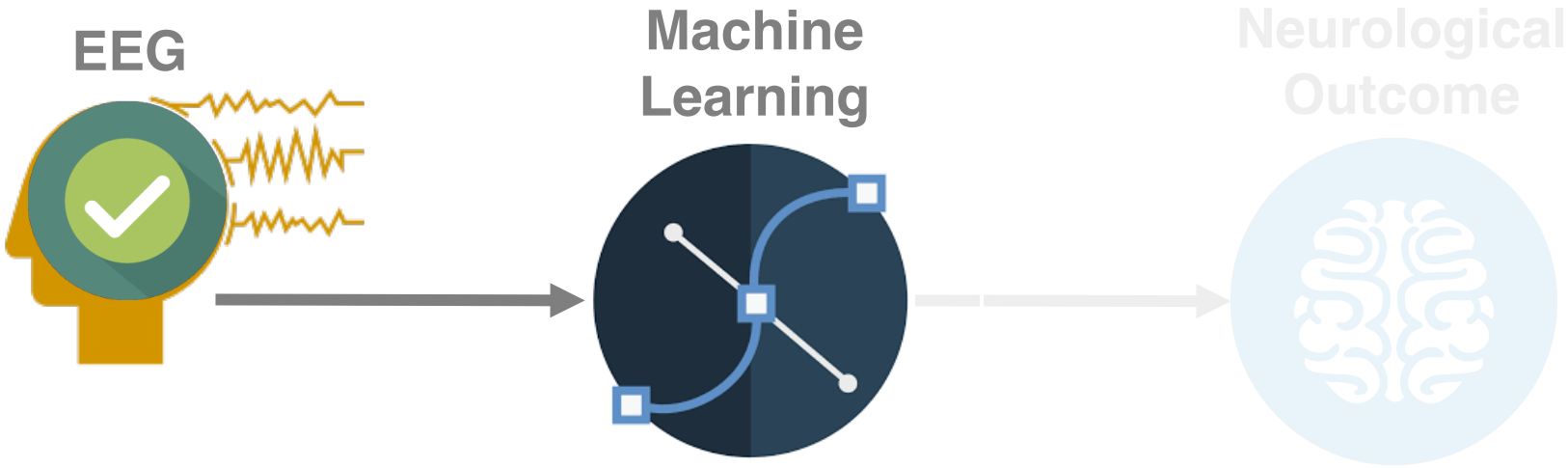
- 3D CNNs provide performance on par with the best-performing feature-based approaches.

	AUROC	TPR| FPR=0	TPR| FPR=1%	TPR| FPR=5%
Neural Networks				
2D CNN	0.71 (0.09)	0 (0.0)	0.09 (0.06)	0.23 (0.10)
**3D CNN	0.81 (0.04)	0.24 (0.05)	0.30 (0.05)	0.45 (0.09)
LRCN	0.79 (0.08)	0.19 (0.11)	0.34 (0.15)	0.48 (0.15)
SVM				
*Linear	0.76 (0.13)	0.17 (0.18)	0.21 (0.16)	0.32 (0.19)
Quadratic	0.72 (0.14)	0.14 (0.09)	0.16 (0.09)	0.26 (0.12)
Fine Gaussian	0.75 (0.14)	0.18 (0.14)	0.2 (0.12)	0.31 (0.19)
Medium Gaussian	0.56 (0.03)	0.01 (0.02)	0.01 (0.02)	0.06 (0.04)
Coarse Gaussian	0.71 (0.13)	0.07 (0.12)	0.08 (0.12)	0.15 (0.1)
Discriminant Analysis				
Linear	0.77 (0.15)	0.21 (0.13)	0.23 (0.14)	0.34 (0.25)
*Quadratic	0.78 (0.14)	0.22 (0.13)	0.25 (0.14)	0.4 (0.22)
Cubic	0.72 (0.13)	0 (0)	0 (0)	0.15 (0.24)
Decision Trees				
Simple	0.68 (0.11)	0 (0)	0 (0)	0 (0)
Medium	0.7 (0.09)	0.05 (0.11)	0.05 (0.11)	0.11 (0.16)
Complex	0.69 (0.07)	0 (0)	0 (0)	0 (0)
*RUS Boosted	0.8 (0.1)	0.25 (0.2)	0.25 (0.2)	0.39 (0.17)
Ensemble Boosted	0.74 (0.12)	0.12 (0.17)	0.12 (0.17)	0.14 (0.19)
Ensemble Bagged	0.77 (0.1)	0.09 (0.14)	0.11 (0.13)	0.22 (0.15)

Pros/Cons of the Approach

- Advantages
  - No feature engineering
  - No pre-processing required
  - Performance is on par with, or exceeds, feature-based approaches
- Disadvantages
  - Requires 72 hours of data
  - Requires tuning of 300K+ parameters
  - Models are difficult to interpret

Develop *time-sensitive* modeling approaches

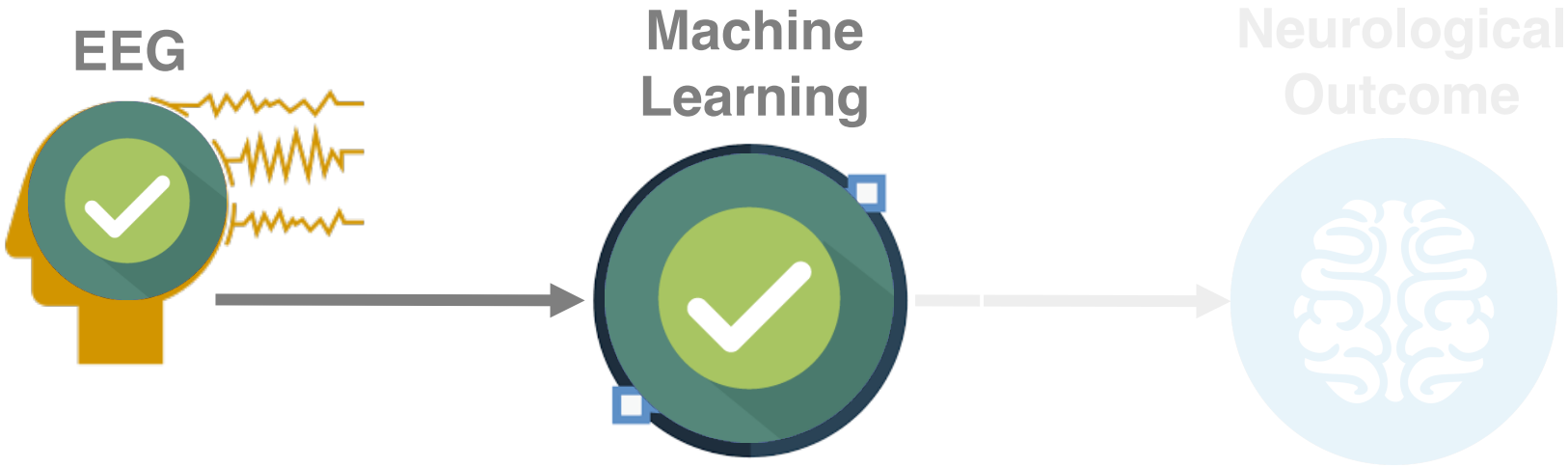


1. Collected an EEG archive 2x larger than largest set previously described in the literature

2. Time-sensitive

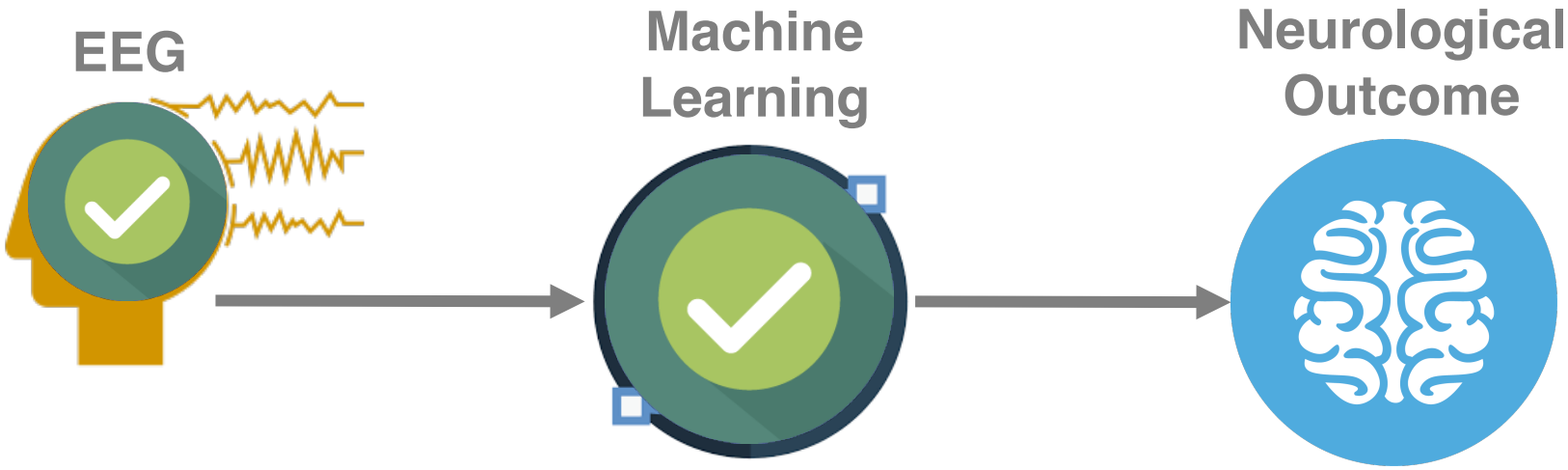
3. Deployable

Develop *time-sensitive* modeling approaches



1. Collected an EEG archive 2x larger than largest set previously described in the literature
2. Demonstrated 3 time-sensitive approaches that out-perform state-of-the-art
3. Deployable

Choosing a model for deployment

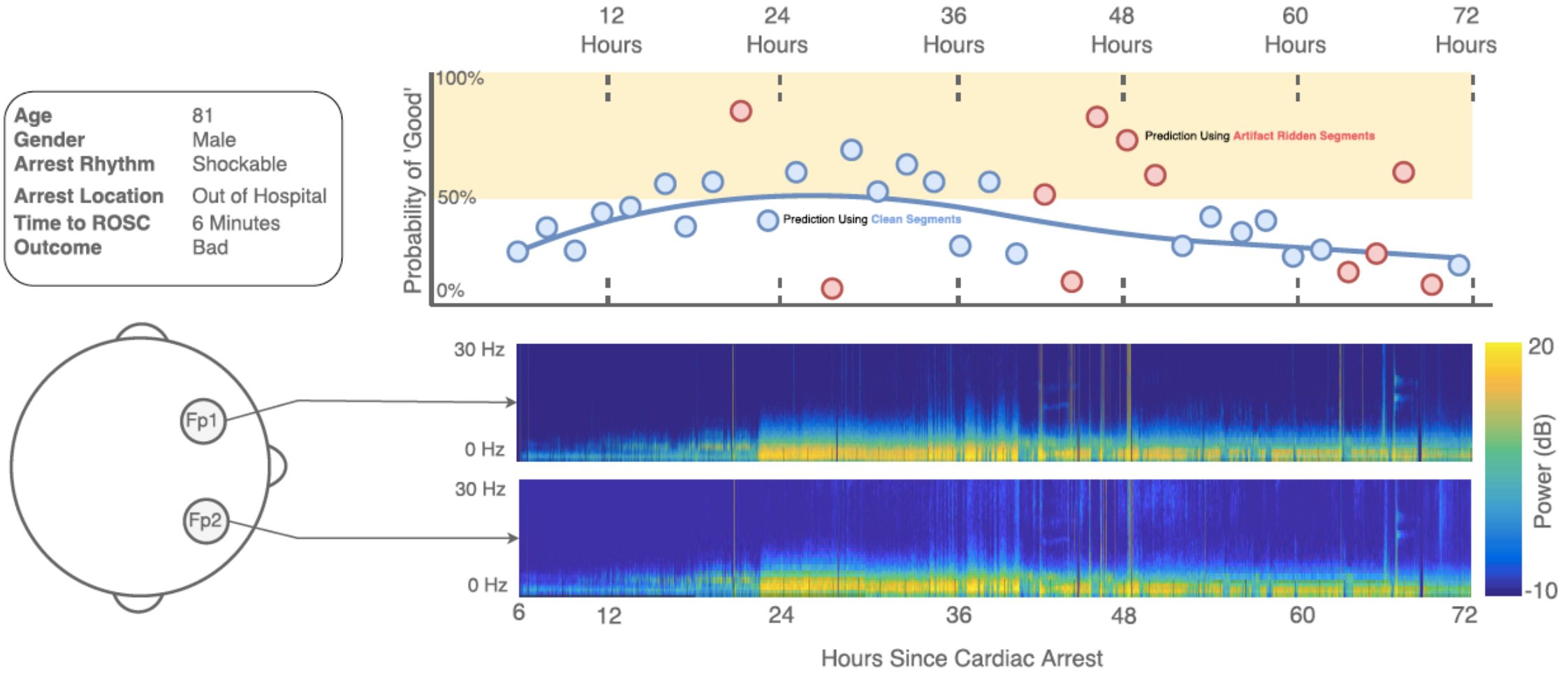


1. Collected an EEG archive 2x larger than largest set previously described in the literature
2. Demonstrated 3 time-sensitive approaches that out-perform state-of-the-art
3. Deployable

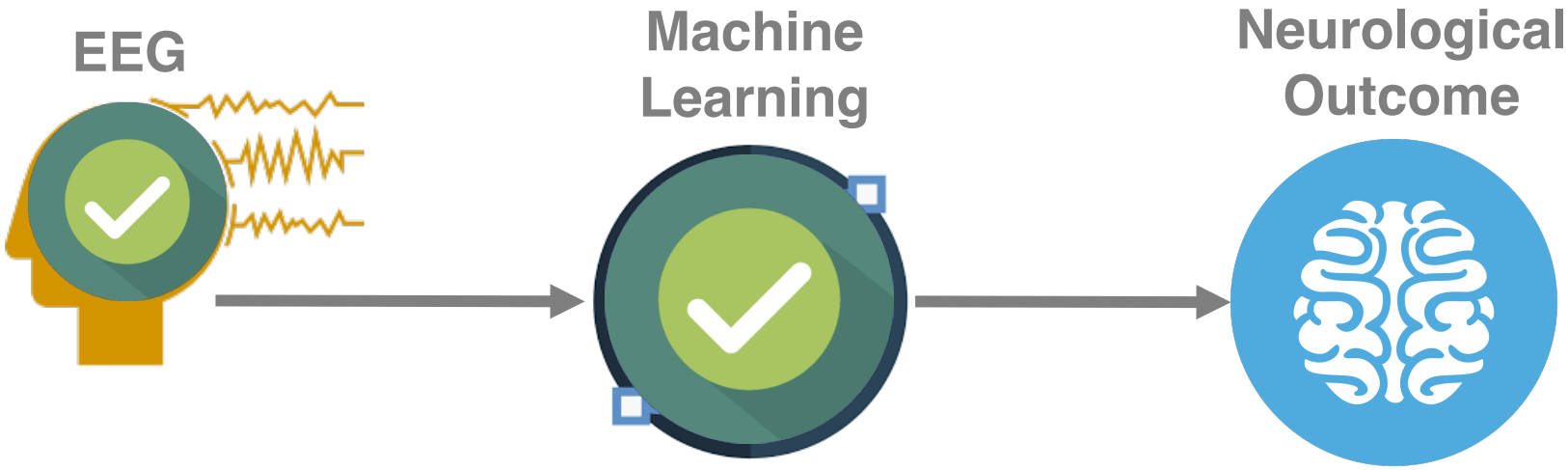
Criteria for deployment

- Interpretable
  - Clear why model has provided the prognostication at hand
- Robust
  - Robust to issues of missing data (e.g. lost channels)
- Well-calibrated
  - Accurate mapping between predicted and actual probabilities of outcome

Proposed model for deployment:
sequential logistic regression with DBN constraints

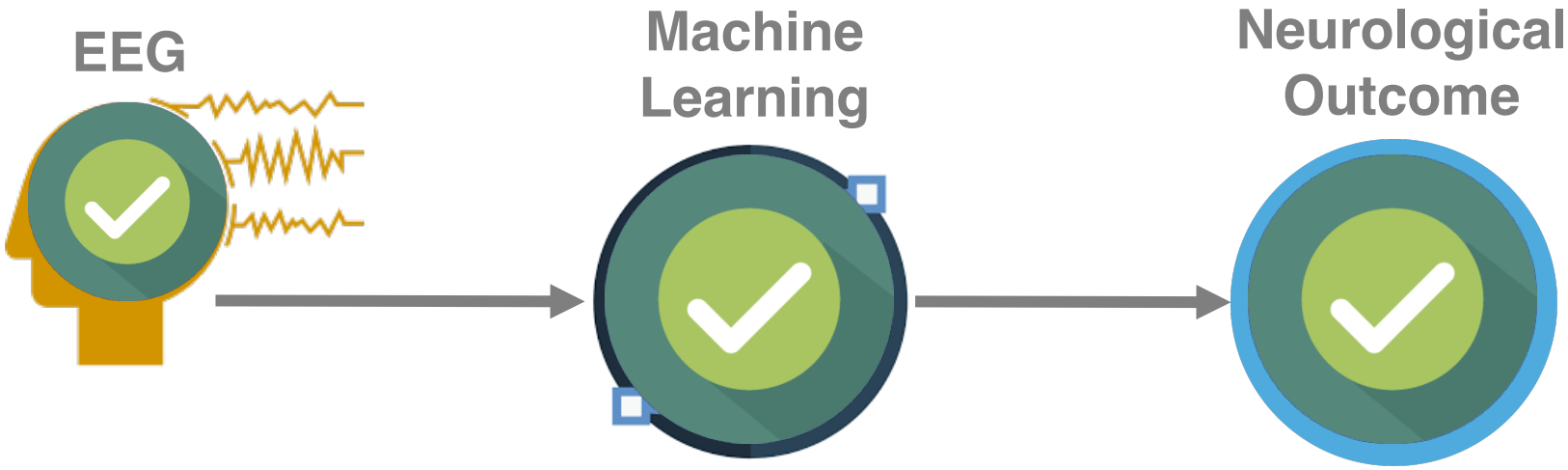


Choosing a model for deployment



1. Collected an EEG archive 2x larger than largest set previously described in the literature
2. Demonstrated 3 time-sensitive approaches that out-perform state-of-the-art
3. Deployable

Assessing performance: classification and calibration



1. Collected an EEG archive 2x larger than largest set previously described in the literature

2. Demonstrated 3 time-sensitive approaches that out-perform state-of-the-art

3. Our proposed approach provides excellent classification, is well-calibrated, and Interpretable

Summary

Collected 785 PAC patients from five university affiliated hospitals

- 35,000 hours of 21-channel continuous EEG recordings,
- A selection of clinical covariates and an ordinal measure of

Extracted 57 quantitative EEG features that capture three signal properties

- Complexity: the degree of randomness in the EEG signal,
- Category: qualitative descriptors of signal characteristics or behaviours and
- Connectivity: interactions between EEG electrodes.

Tested novel methods for time-sensitive classification of outcomes

- Penalized, sequential, logistic regression using 57 multi-scale features,
- Logistic regression using 10 qEEG features constrained by a DBN and
- A variety of deep neural network architectures

Acknowledgements and thanks

Contributors:
Edilberto Amorim
Tuka Alhanai
Jong Woo Lee
Michel van Putten
Jeannette Hofmeijer
Adithya Sivaraju
Nicolas Gaspard
Barry Ruijter
Siddharth Biswal
Valdery Moura Junior
Michael Donnino
Susan Herman

Committee:

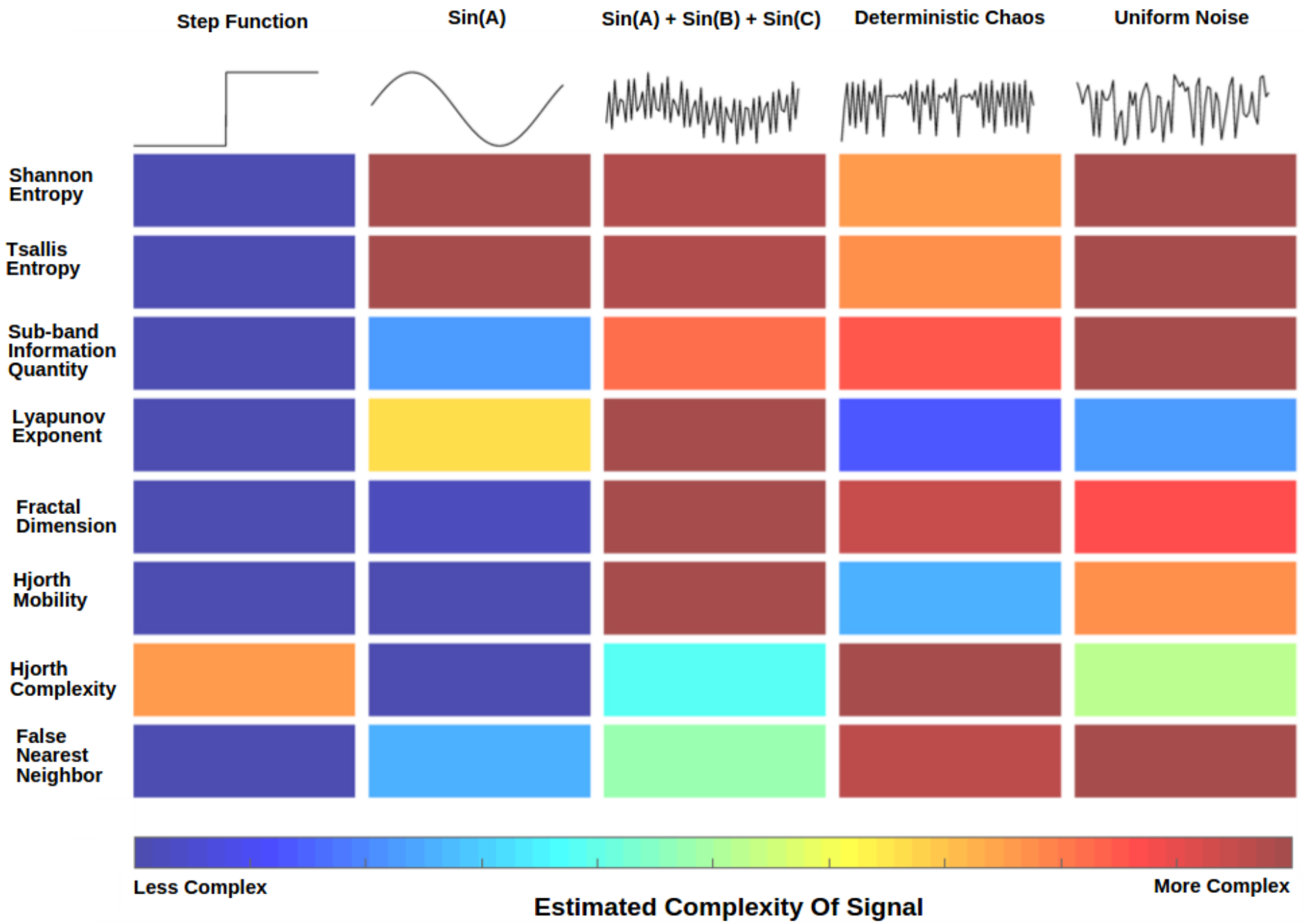
Prof. Roger G. Mark

Prof. Emery N. Brown

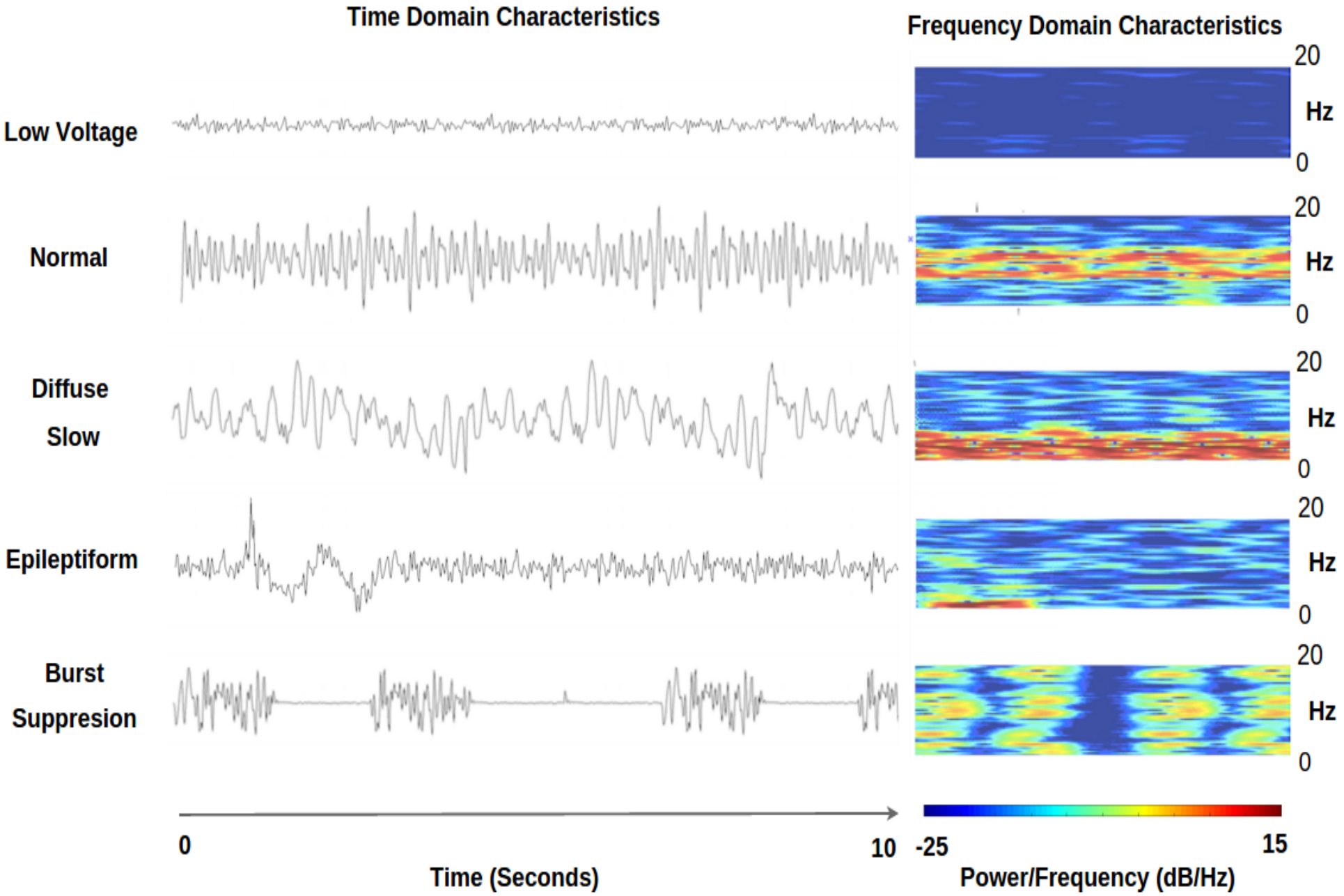
Dr. M. Brandon Westover

Prof. Thomas Heldt

Complexity


Step Function
Sin(A)
Sin(A) + Sin(B) + Sin(C)
Deterministic Chaos
Uniform Noise
Shannon Entropy
Tsallis Entropy
Sub-band Information Quantity
Lyapunov Exponent
Fractal Dimension
Hjorth Mobility
Hjorth Complexity
False Nearest Neighbor
Less Complex
More Complex
Estimated Complexity Of Signal

Category


Time Domain Characteristics
Frequency Domain Characteristics
Low Voltage
Normal
Diffuse
Slow
Epileptiform
Burst
Suppresion
20
Hz
0
0
10
Time (Seconds)
-25
15
Power/Frequency (dB/Hz)

Connectivity