Predicting Deterioration

Improving prediction by the introduction of a novel ‘seriously deteriorated’ outcome

Presented by Victoria Campbell & Anton van der Vegt

Thanks to our team: Imogen Mitchell, Arthas Flabouris, Naitik Mehta, Ian Scott, James Malycha, Daryl Jones

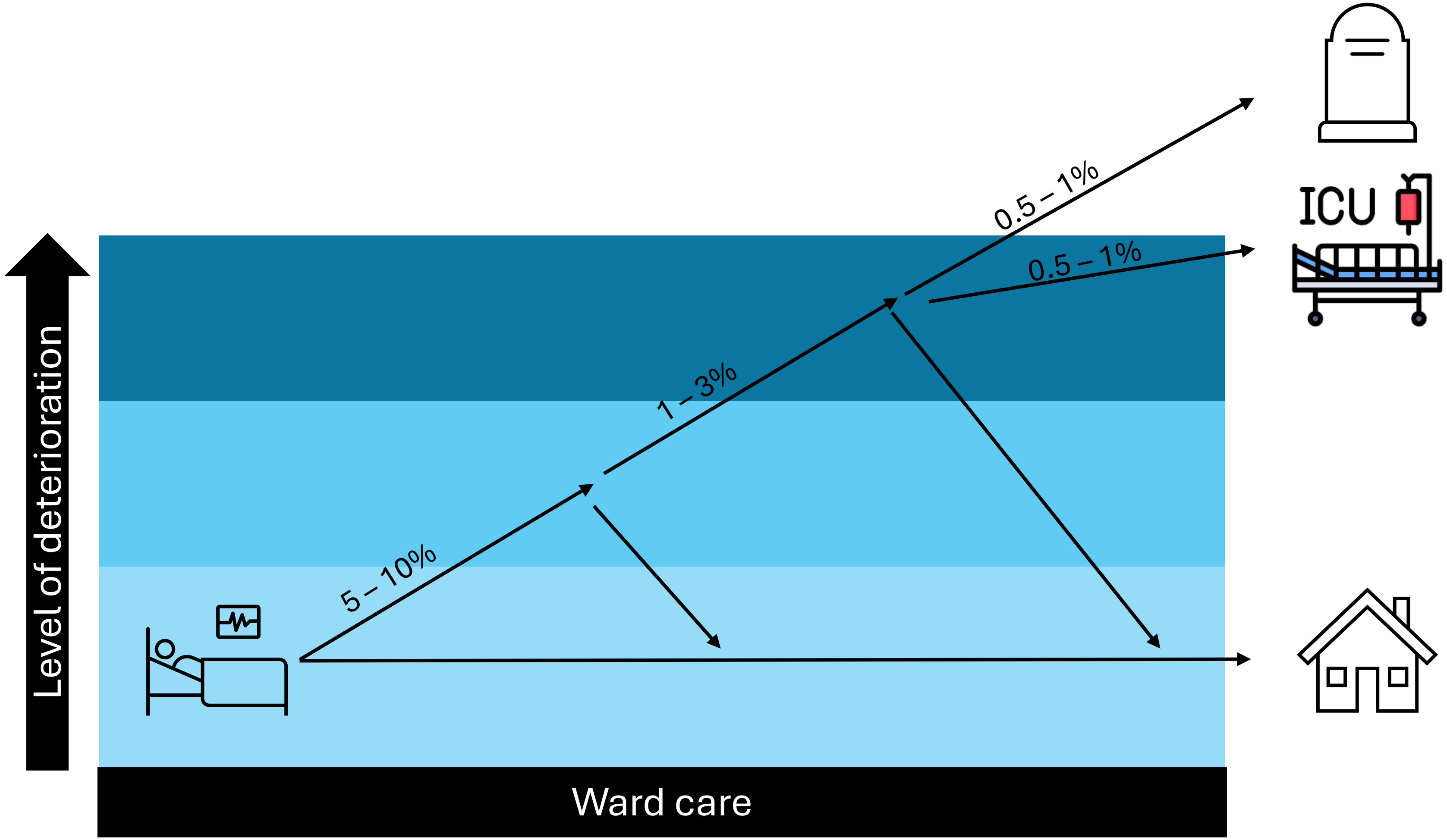
Conflict of interest

None reported

Background

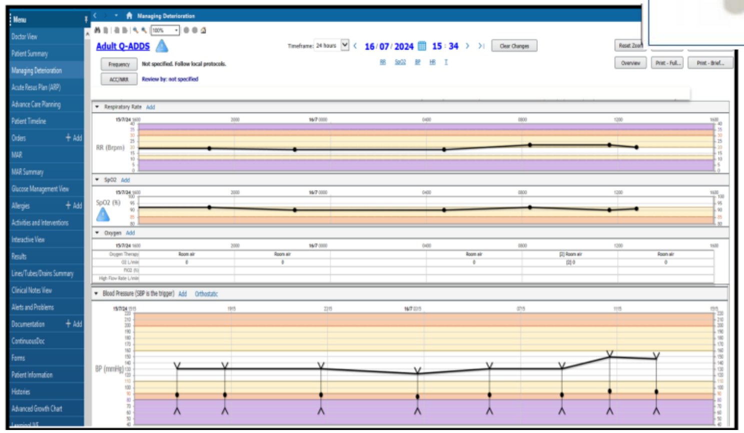
What is patient deterioration?

'A deteriorating patient is one who moves from one clinical state to a worse clinical state which increases their individual risk of morbidity, including organ dysfunction, protracted hospital stay, disability, or death' - Jones et al, 2013

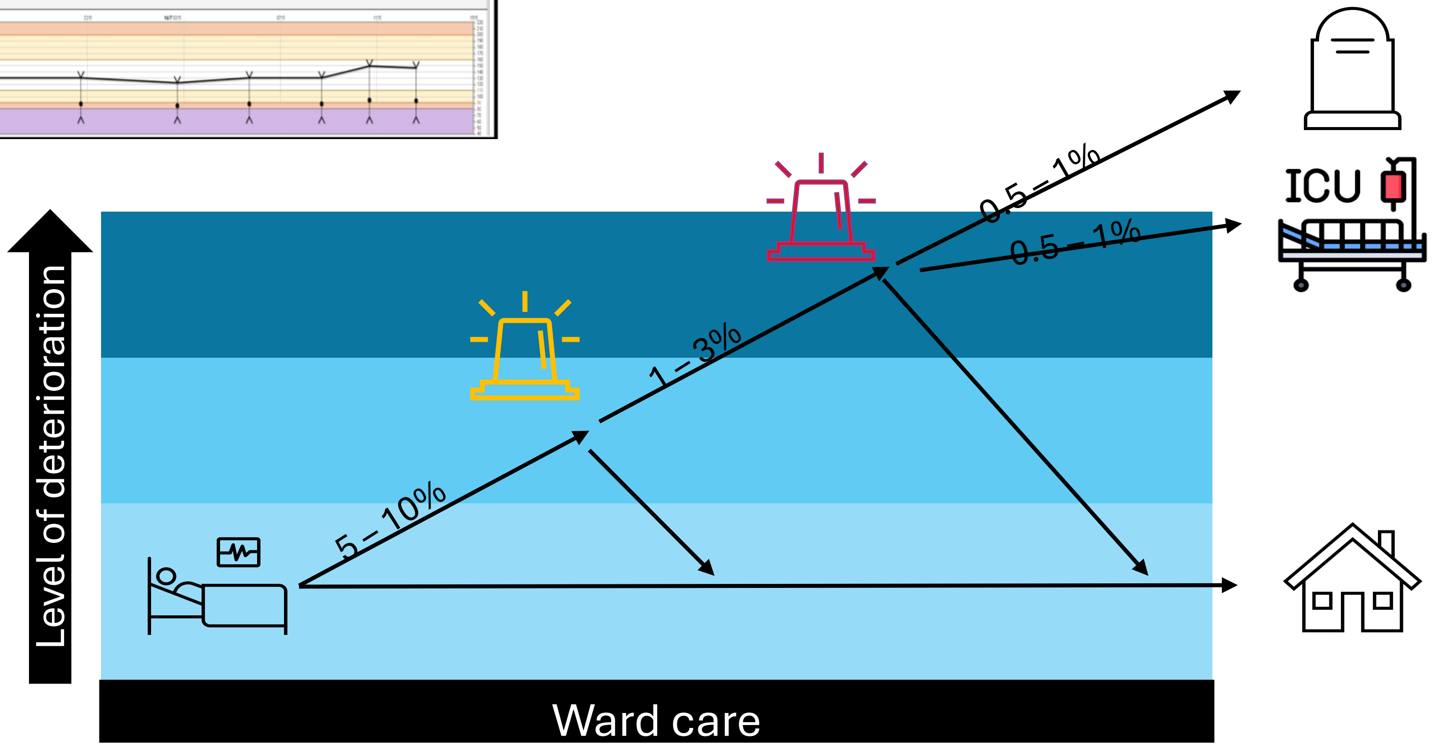


1. Jones D, Mitchell I, Hillman K, Story D. Defining clinical deterioration. Resuscitation. 2013;84(8):1029-1034. doi:10.1016/j.resuscitation.2013.01.013

Early Warning Systems



1. Recognition: Early Warning Tools (EWTs)
2. Notification: Alert staff to patient deterioration
3. Escalation: guidance
4. Response: appropriate expertise



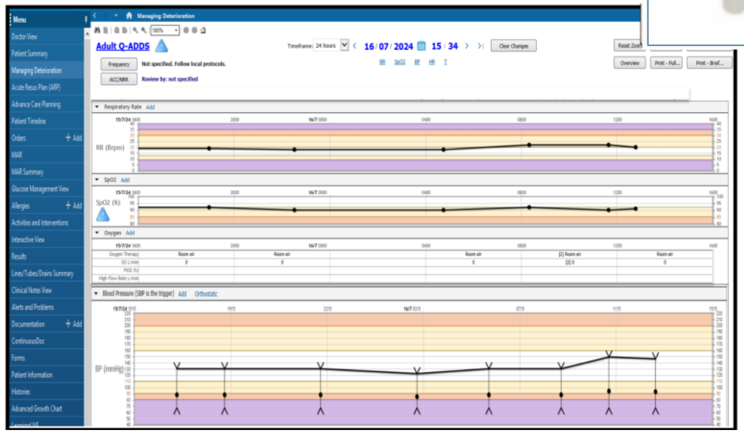
1. Jones D, Mitchell I, Hillman K, Story D. Defining clinical deterioration. Resuscitation. 2013;84(8):1029-1034. doi:10.1016/j.resuscitation.2013.01.013

Early Warning Systems

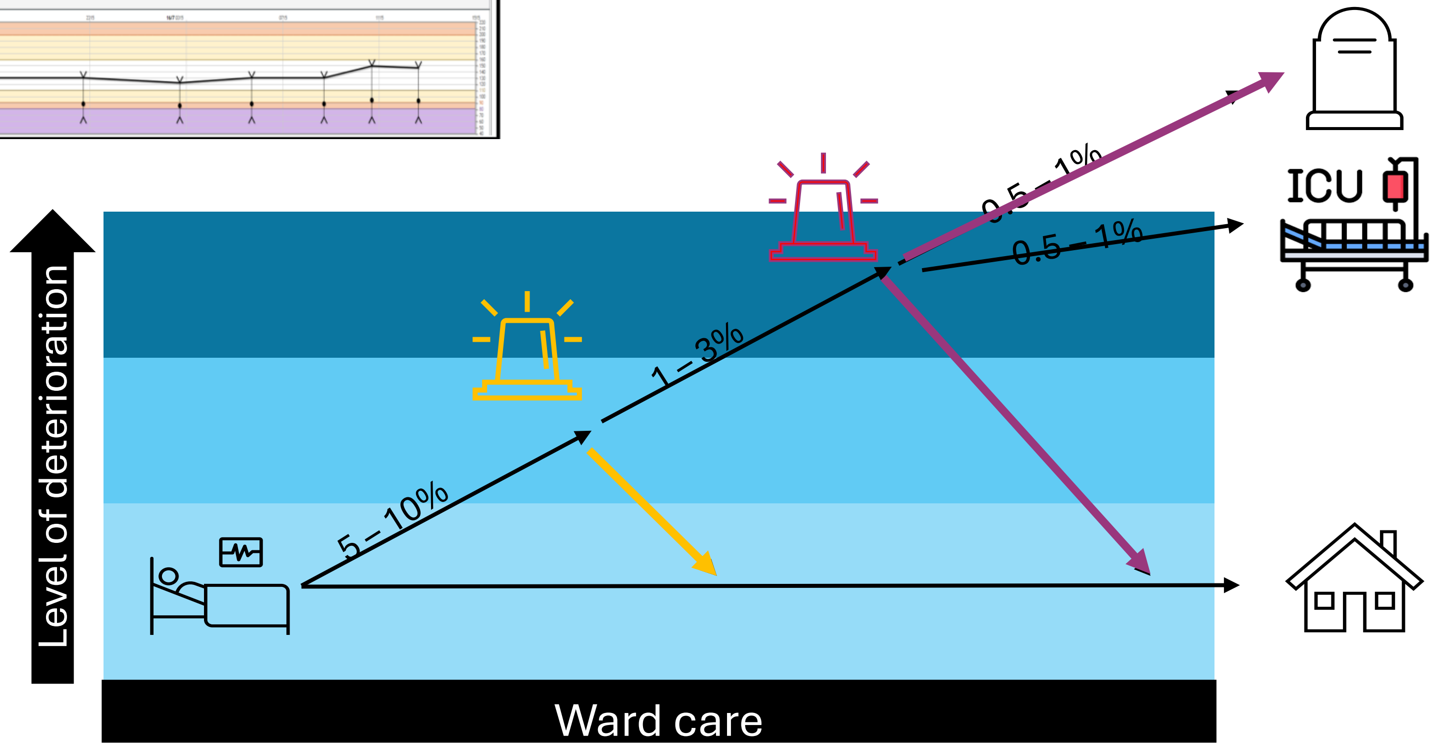
AUSTRALIAN COMMISSION ON SAFETY AND QUALITY IN HEALTH CARE

NSQHS STANDARDS

Standard 8: Recognising and Responding to Acute Deterioration



1. Recognition: Early Warning Tools (EWTs)
2. Notification: Alert staff to patient deterioration
3. Escalation: guidance
4. Response: appropriate expertise



1. Jones D, Mitchell I, Hillman K, Story D. Defining clinical deterioration. Resuscitation. 2013;84(8):1029-1034. doi:10.1016/j.resuscitation.2013.01.013

Current Early Warning Tools (EWTs)

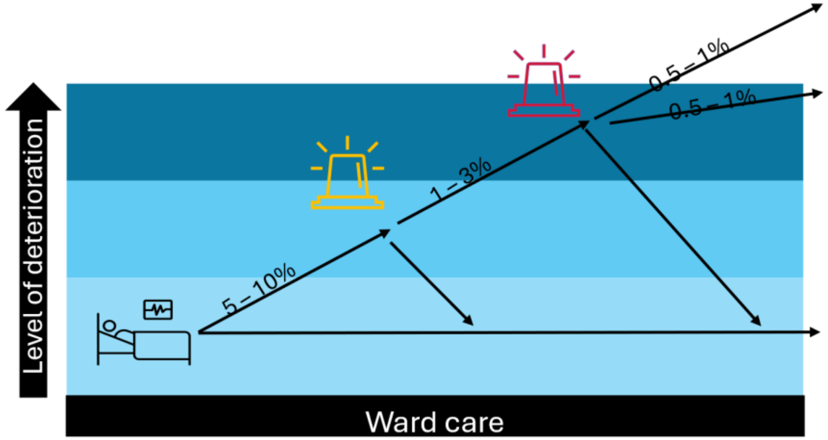
Rule-based
e.g., NEWS, Q-ADDS, BTF

Purpose: identify deterioration when it happens and respond: acts as a safety net.

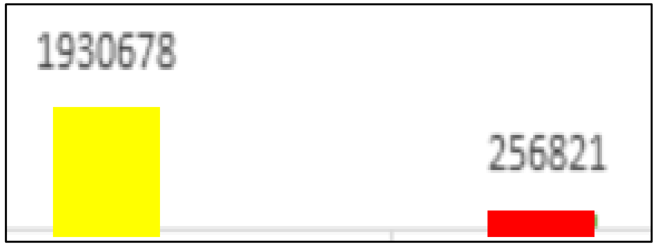
> Resuscitation. 2020 Aug:153:28-34. doi: 10.1016/j.resuscitation.2020.05.027. Epub 2020 Jun 3.

Predicting clinical deterioration with Q-ADDS compared to NEWS, Between the Flags, and eCART track and trigger tools

Victoria Campbell [1], Roger Conway [2], Kyle Carey [3], Khoa Tran [4], Adam Visser [5], Shaune Gifford [6], Mia McLanders [7], Dana Edelson [8], Matthew Churpek [9]



	% patients	Alerting vital sign sets	Calls per month	Critical care input
Q-ADDS MET	10%	1%	~200	~5%
Q-ADDS MO review	47%	14%	~2,500	~0.3%



Current Early Warning Tools (EWTs)

Rule-based
e.g., NEWS, Q-ADDS, BTF

Purpose: identify deterioration when it happens and respond: acts as a safety net.

AI-based
e.g., eCART, HAVEN, AAM

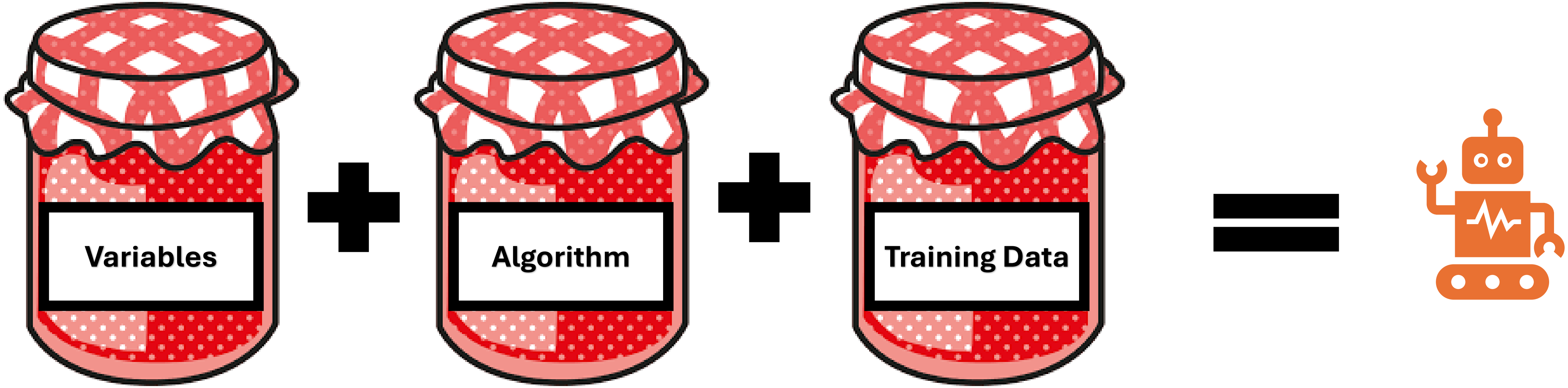
Purpose: predict deterioration

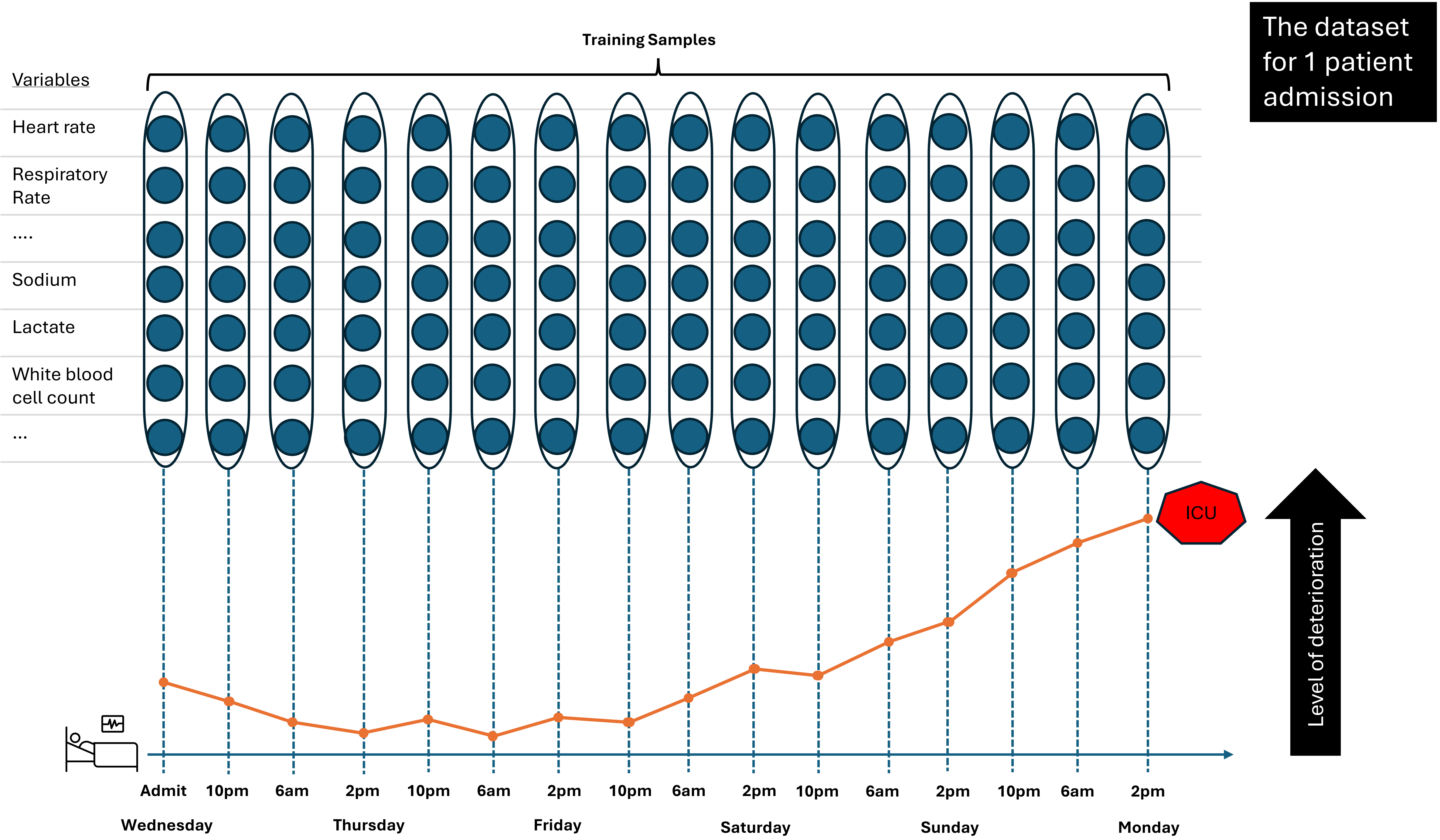
> before it occurs: prevent rather than respond

> with better accuracy: less false alerts

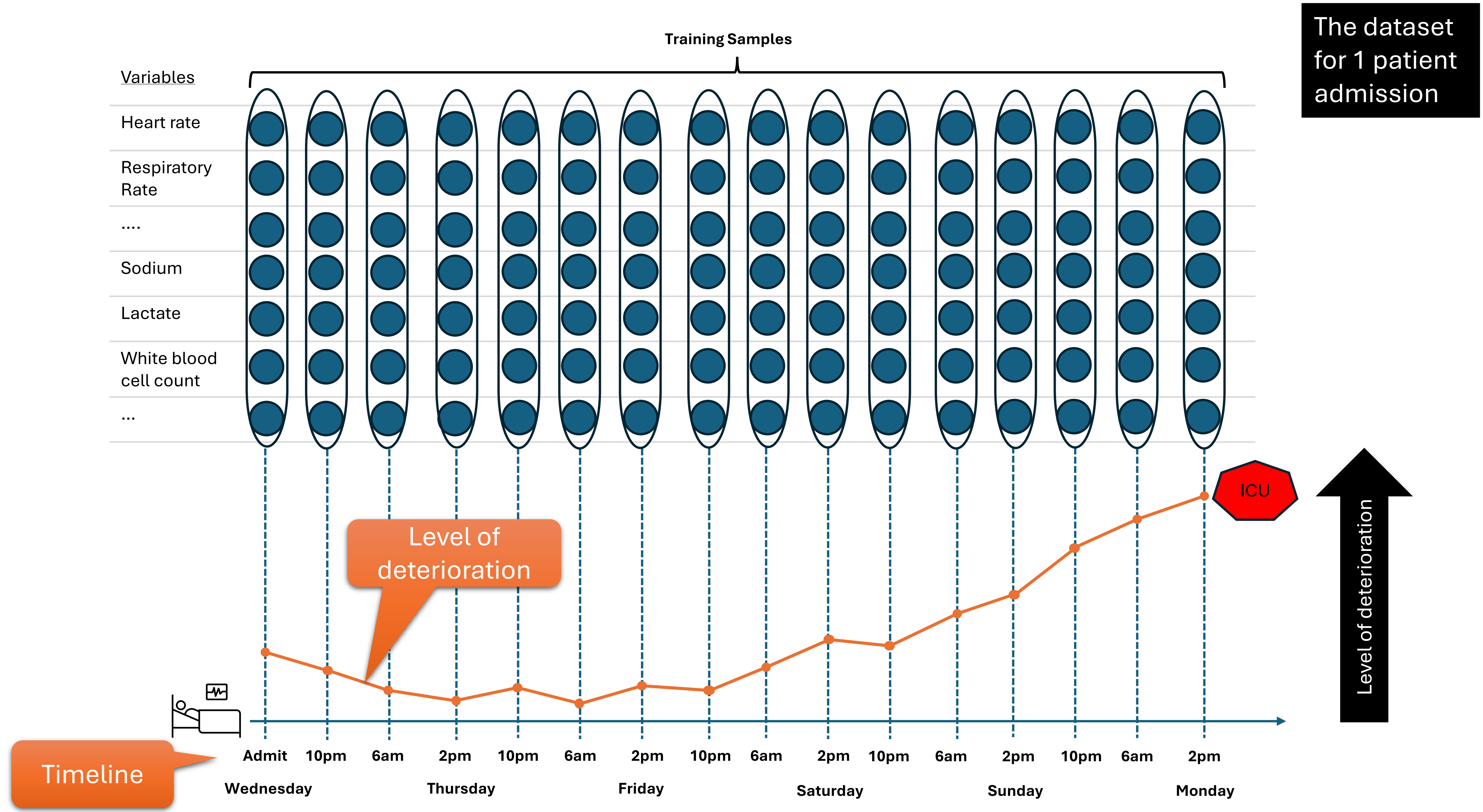
AI = ***Augmented*** Intelligence

AI patient risk prediction model ingredients

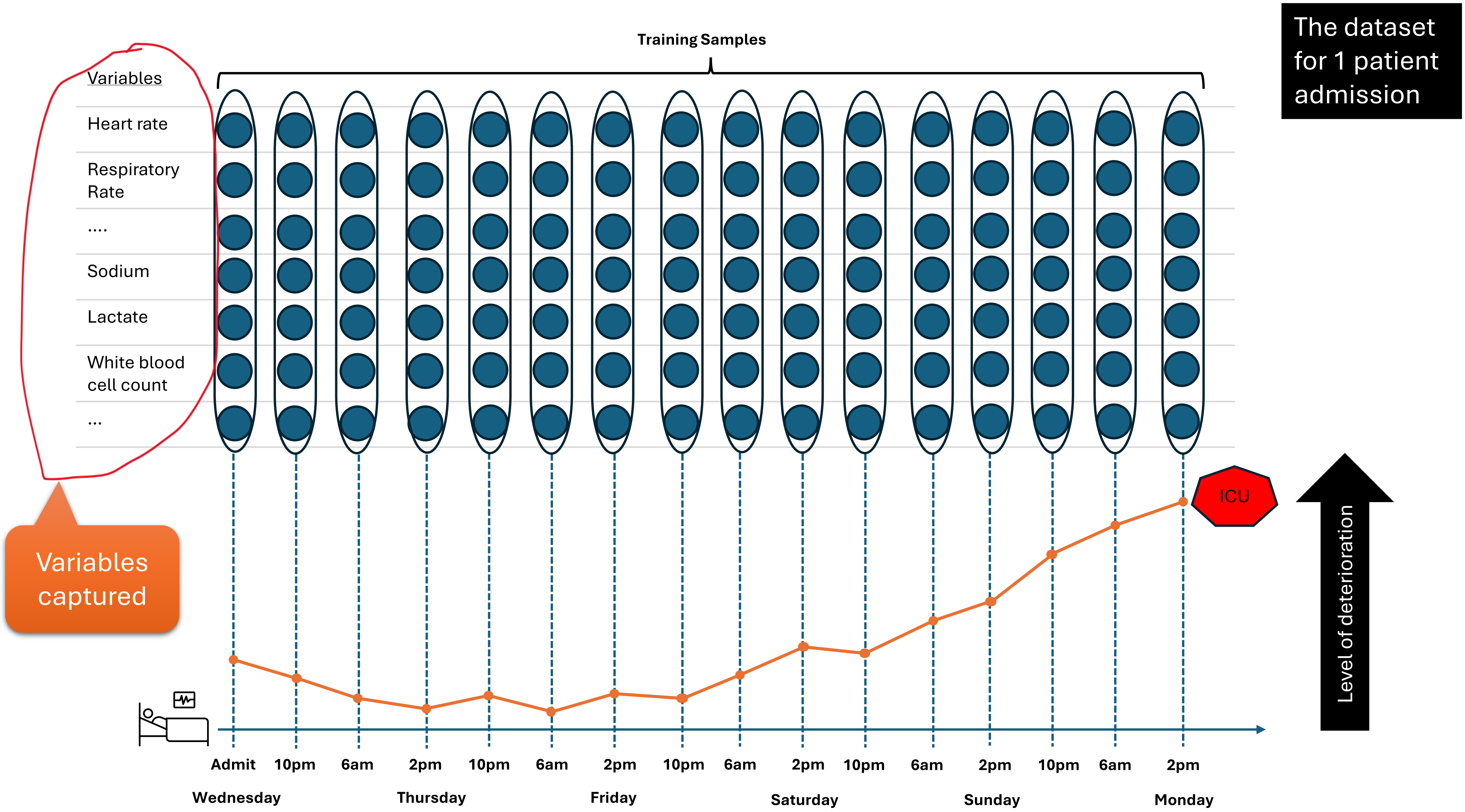




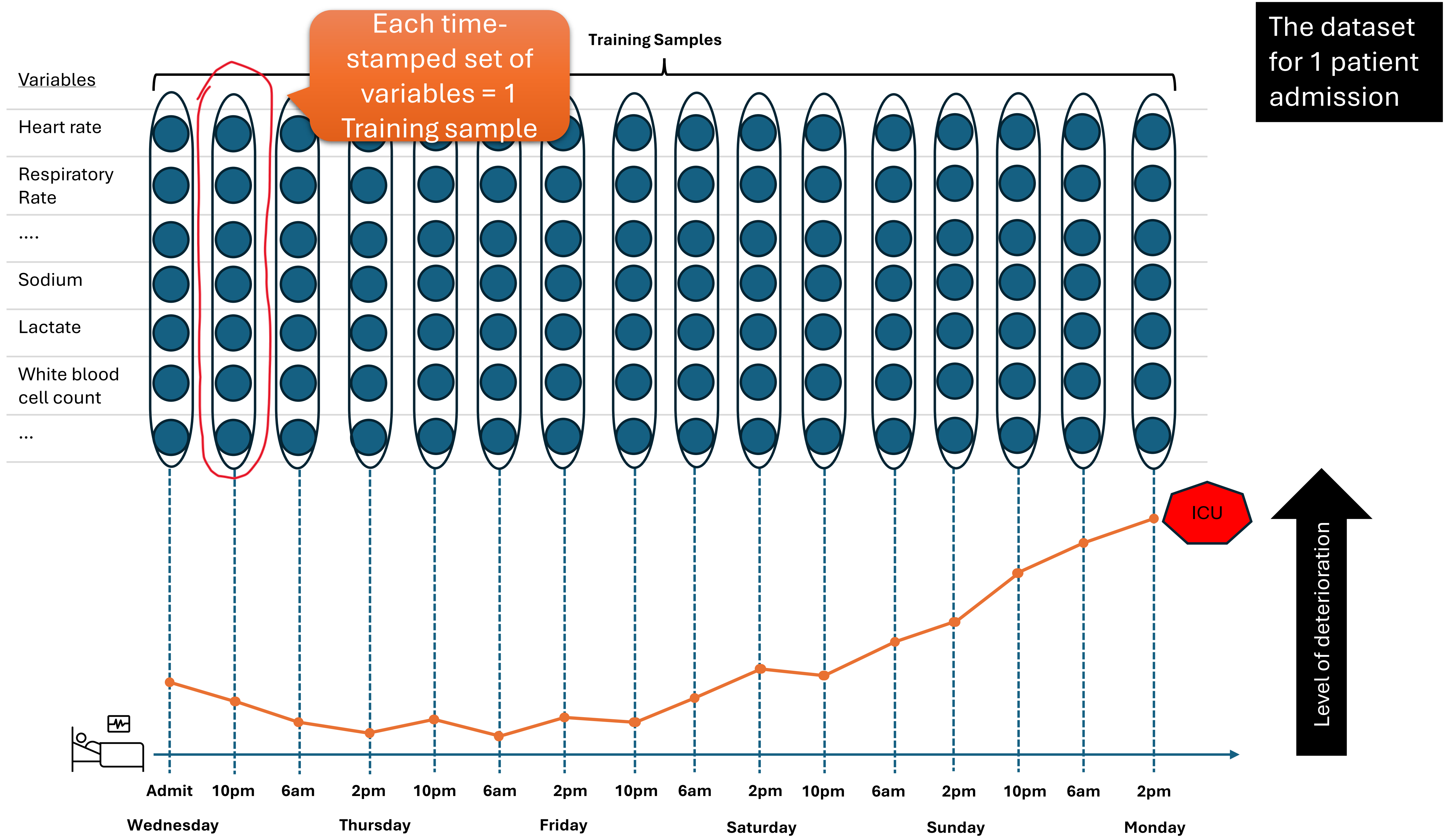
Training Samples
The dataset for 1 patient admission
Variables
Heart rate
Respiratory Rate
....
Sodium
Lactate
White blood cell count
...
ICU
Level of deterioration
Admit
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
Wednesday
Thursday
Friday
Saturday
Sunday
Monday



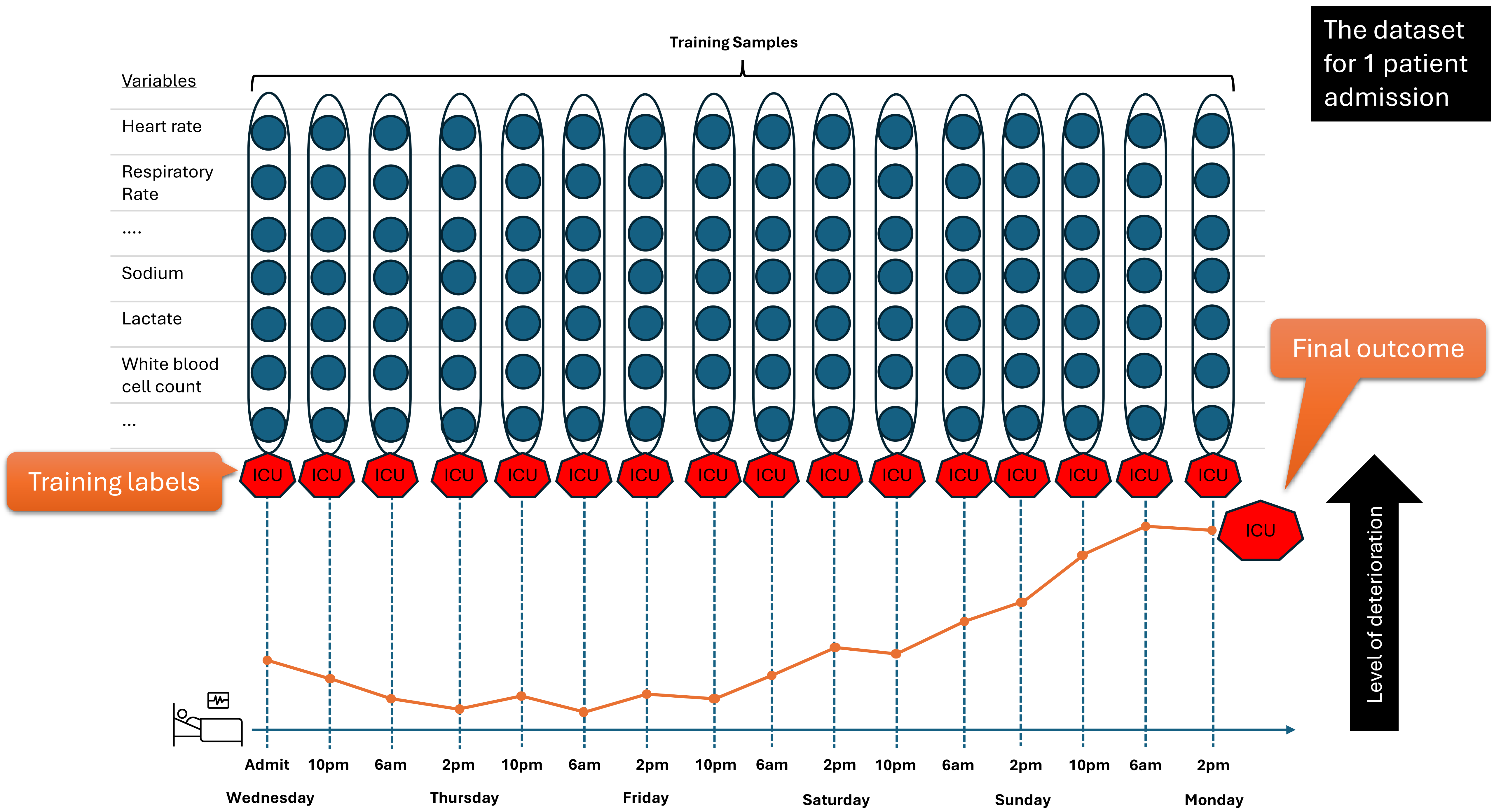
Training Samples
Variables
Heart rate
Respiratory Rate
....
Sodium
Lactate
White blood cell count
...
The dataset for 1 patient admission
Level of deterioration
ICU
Level of deterioration
Timeline
Admit
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
Wednesday
Thursday
Friday
Saturday
Sunday
Monday



The dataset for 1 patient admission
Training Samples
Variables
Heart rate
Respiratory Rate
....
Sodium
Lactate
White blood cell count
...
Variables captured
ICU
Level of deterioration
Admit
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
Wednesday
Thursday
Friday
Saturday
Sunday
Monday

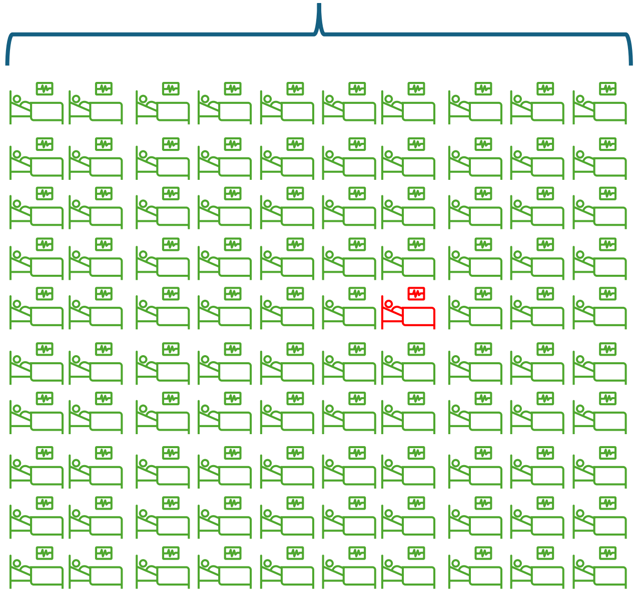


The dataset for 1 patient admission
Training Samples
Each time-stamped set of variables = 1 Training sample
Variables
Heart rate
Respiratory Rate
....
Sodium
Lactate
White blood cell count
...
ICU
Level of deterioration
Admit
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
Wednesday
Thursday
Friday
Saturday
Sunday
Monday


The dataset for 1 patient admission
Training Samples
Variables
Heart rate
Respiratory Rate
....
Sodium
Lactate
White blood cell count
...
Training labels
ICU
Final outcome
Level of deterioration
Admit
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
10pm
6am
2pm
Wednesday
Thursday
Friday
Saturday
Sunday
Monday

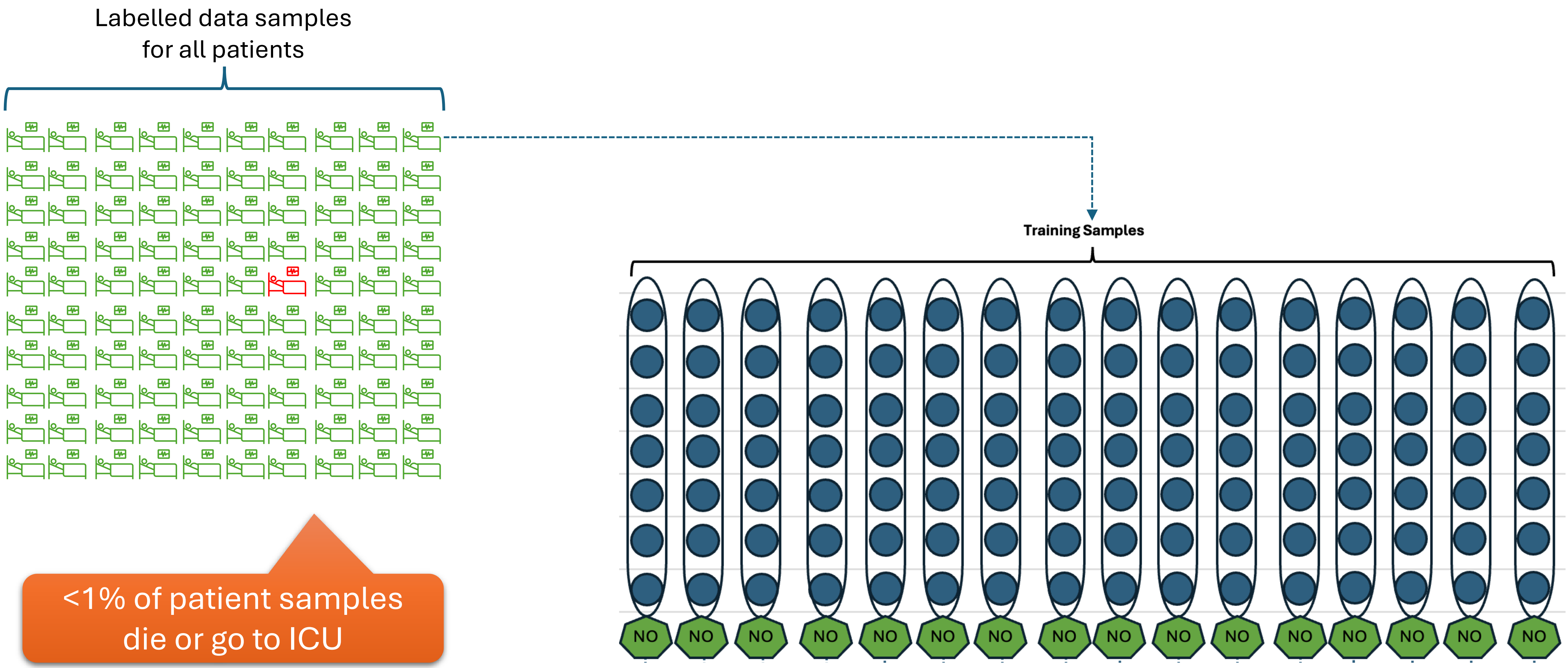
How machine learning systems ‘predict’ deterioration

Labelled data samples
for all patients



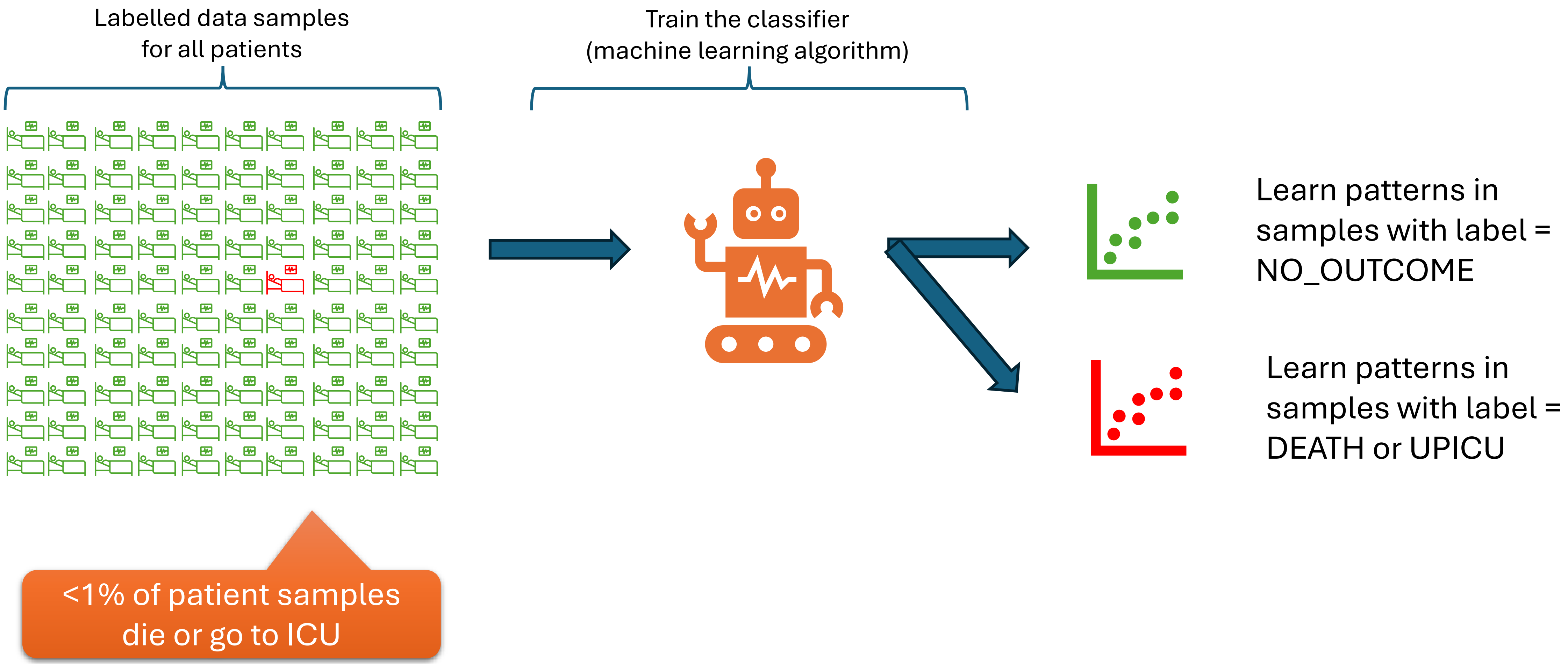
<1% of patient samples
die or go to ICU

How machine learning systems ‘predict’ deterioration

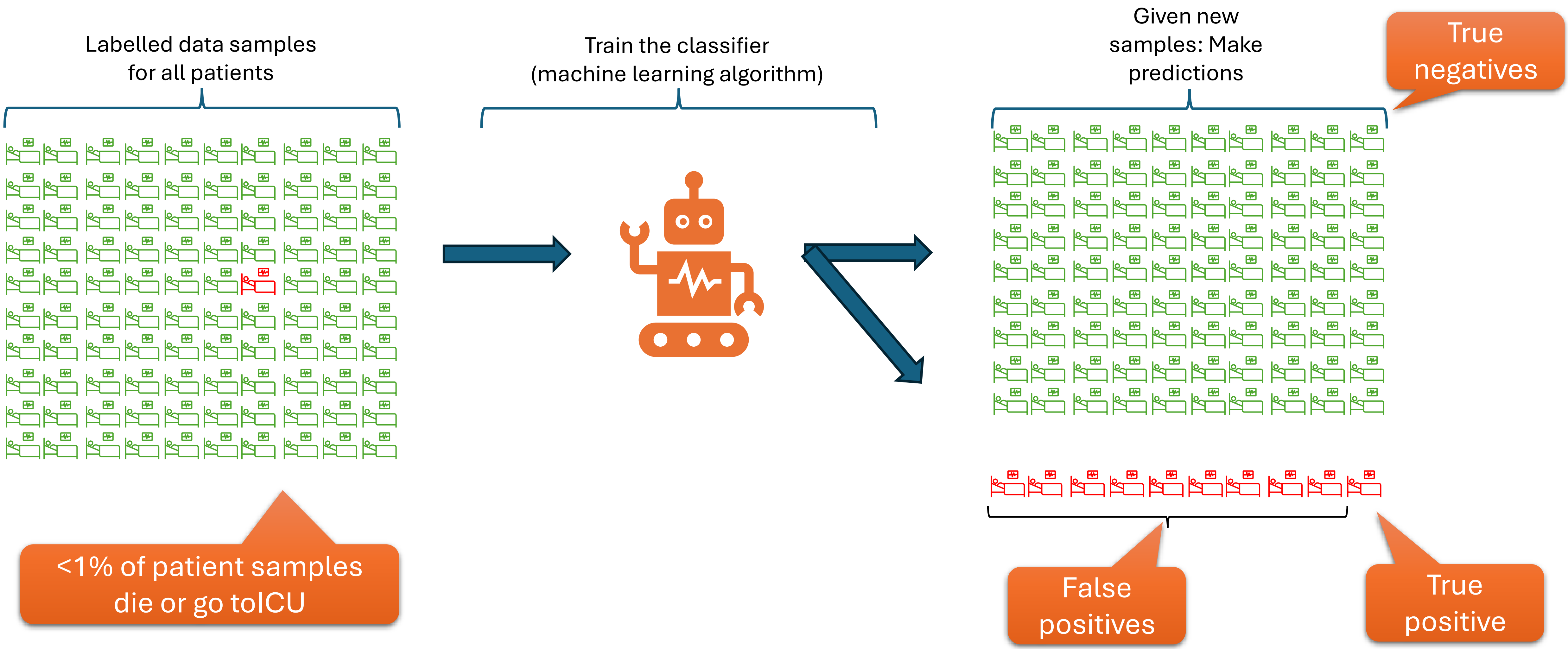


<1% of patient samples die or go to ICU

How machine learning systems ‘predict’ deterioration



How machine learning systems ‘predict’ deterioration



Problems with current outcomes

Background

Journal of the American Medical Informatics Association, 2023, 1–16
https://doi.org/10.1093/jamia/ocad220
Review

AMIA OXFORD

Review

Systematic review and longitudinal analysis of implementing Artificial Intelligence to predict clinical deterioration in adult hospitals: what is known and what remains uncertain

Anton H. van der Vegt, PhD,[*,1] Victoria Campbell , MBBS,[2,3] Imogen Mitchell, PhD,[4] James Malycha, PhD,[5] Joanna Simpson, MBChB,[6] Tracy Flenady, PhD,[7] Arthas Flabouris, MD,[8,9] Paul J. Lane, MBBS,[10] Naitik Mehta, MHServMgt,[11] Vikrant R. Kalke, MHServMgt,[11] Jovie A. Decoyna, MD,[3] Nicholas Es'haghi, BMedSc,[3] Chun-Huei Liu, BMedSc,[3] Ian A. Scott, MHA[1,12]

TABLE G1: CRITERIA UTILISED BY PAPERS

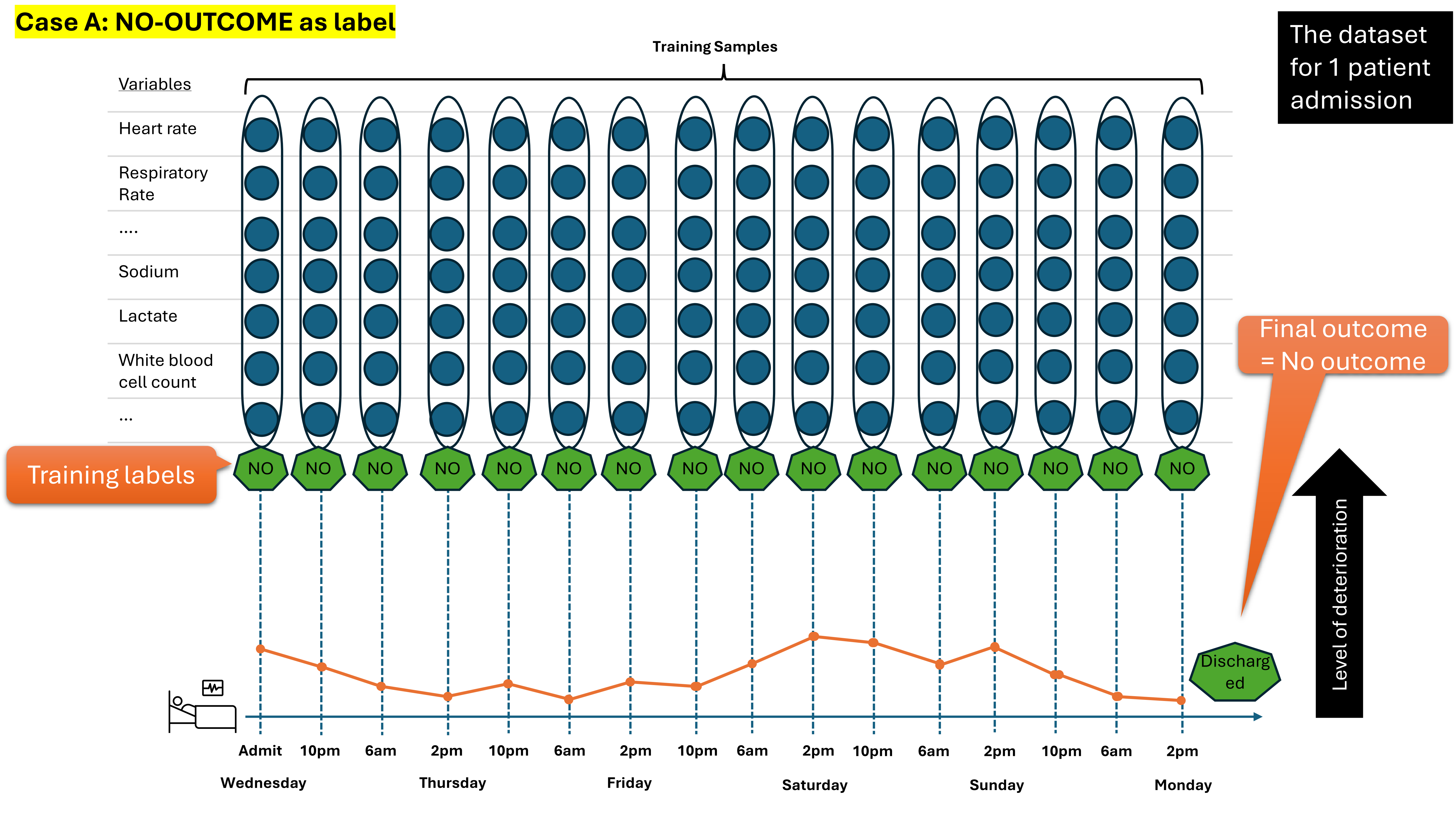
Criteria ID	Criteria specified (with references where applicable)
A	Transfer to ICU
B	In-hospital death
C	Cardiac arrest
D	Intubation and mechanical ventilation
E	Unplanned return to operating theatre
F	Medical emergency team (MET) call
G	Use of Oxygen therapy
H	Transfer to step-down unit
I	Transfer to palliative care
J	Issuance of palliative order
K	National Early Warning Score 2 (NEWS2) trigger/change

Typical training labels for predicting deterioration

- Transfer to ICU & death are used by most algorithm developers
- Other outcomes are:
  - Far less common – usually by 1 or 2 algorithms only
  - Rely upon less commonly collected data

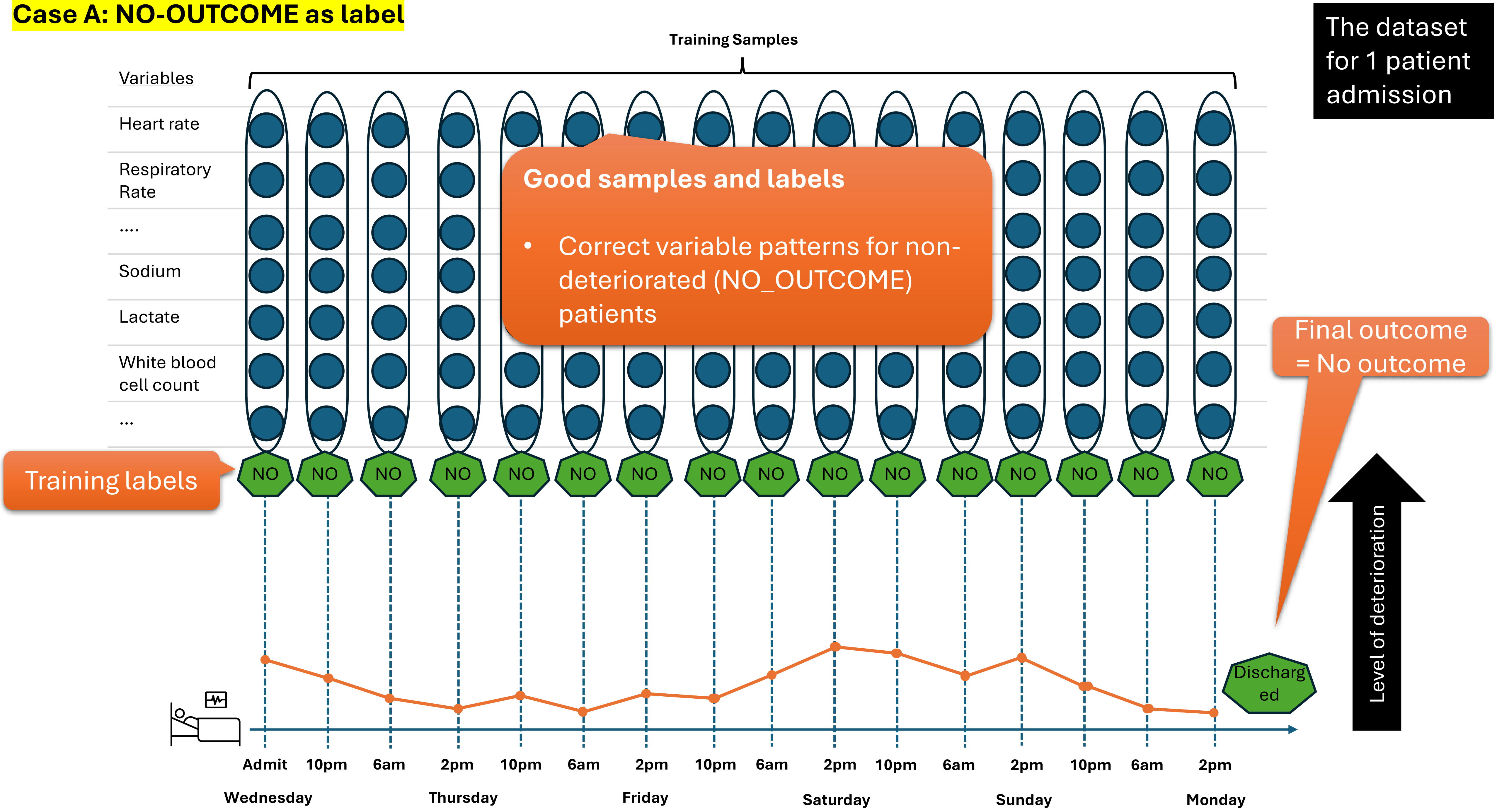
Let’s consider 4 scenarios

1. No or low deterioration patient (NO_OUTCOME)
2. Unplanned ICU patient (UPICU)
3. Patient who dies (DEATH)
4. Patient who seriously deteriorates (SDP)

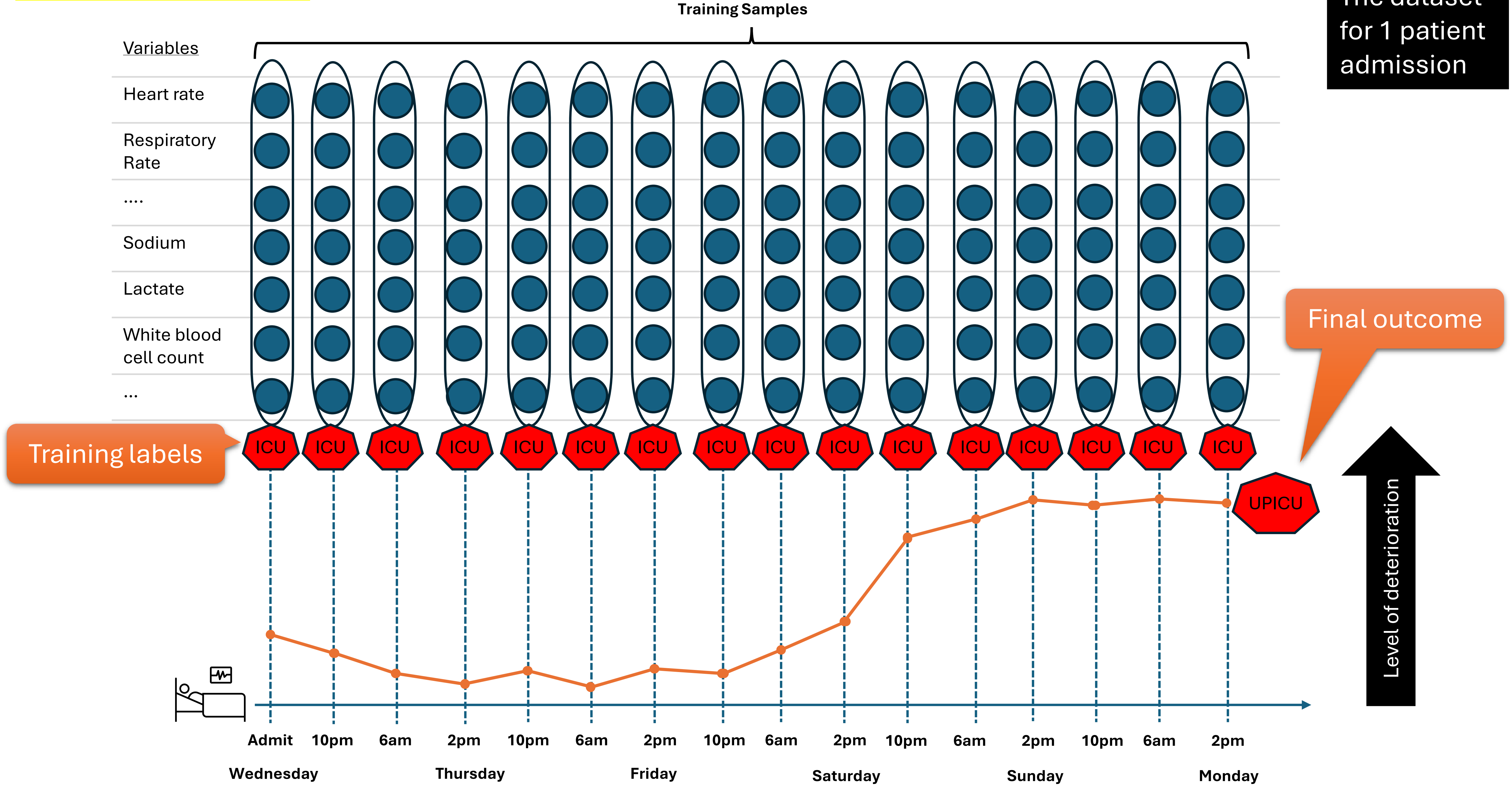


Case A: NO-OUTCOME as label
Training Samples
The dataset for 1 patient admission
Variables
Heart rate
Respiratory Rate
....
Sodium
Lactate
White blood cell count
...
Training labels
NO
Final outcome = No outcome
Level of deterioration
Discharged
Admit
10pm
6am
2pm
Wednesday
Thursday
Friday
Saturday
Sunday
Monday

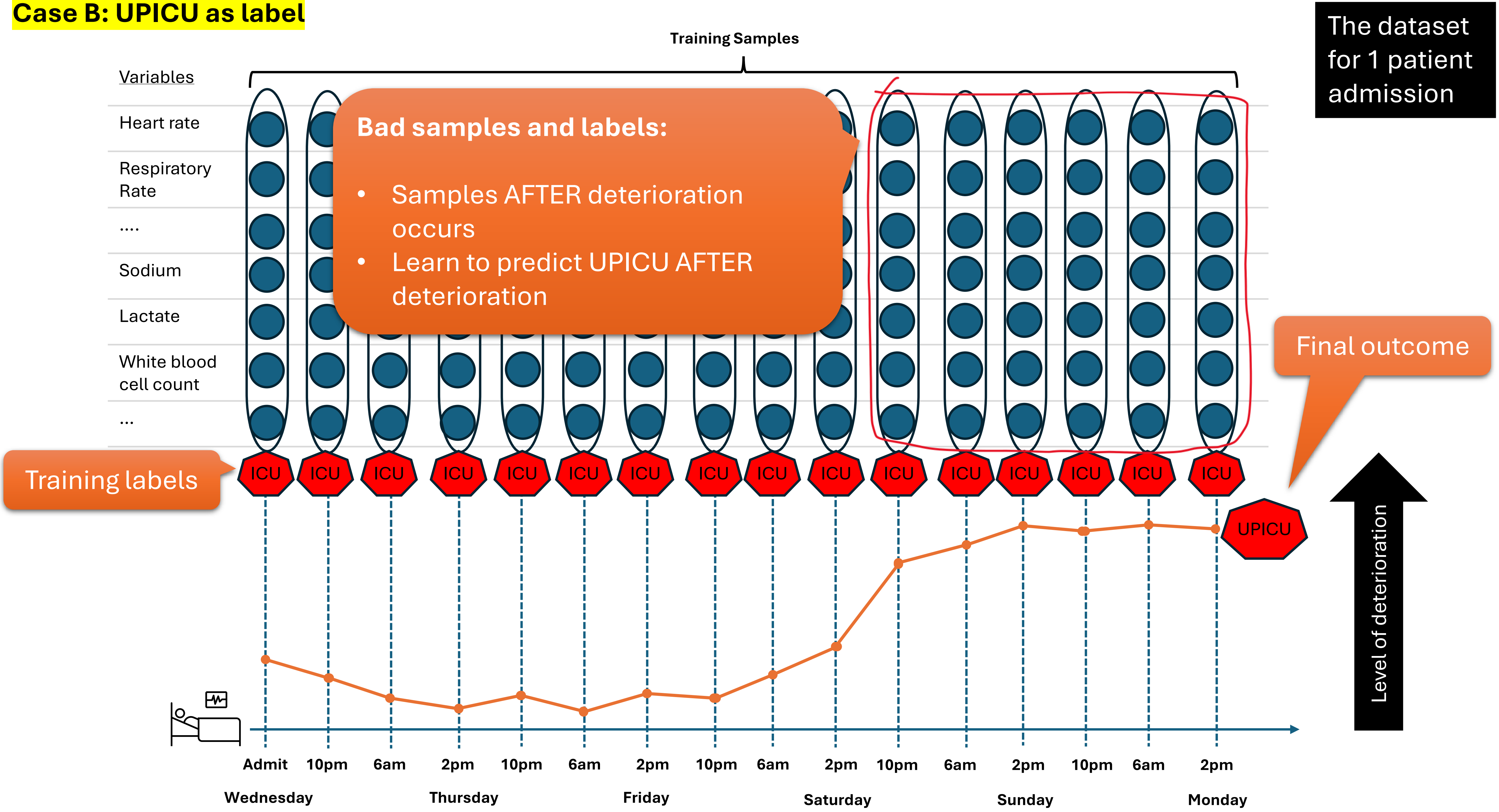
Case A: NO-OUTCOME as label



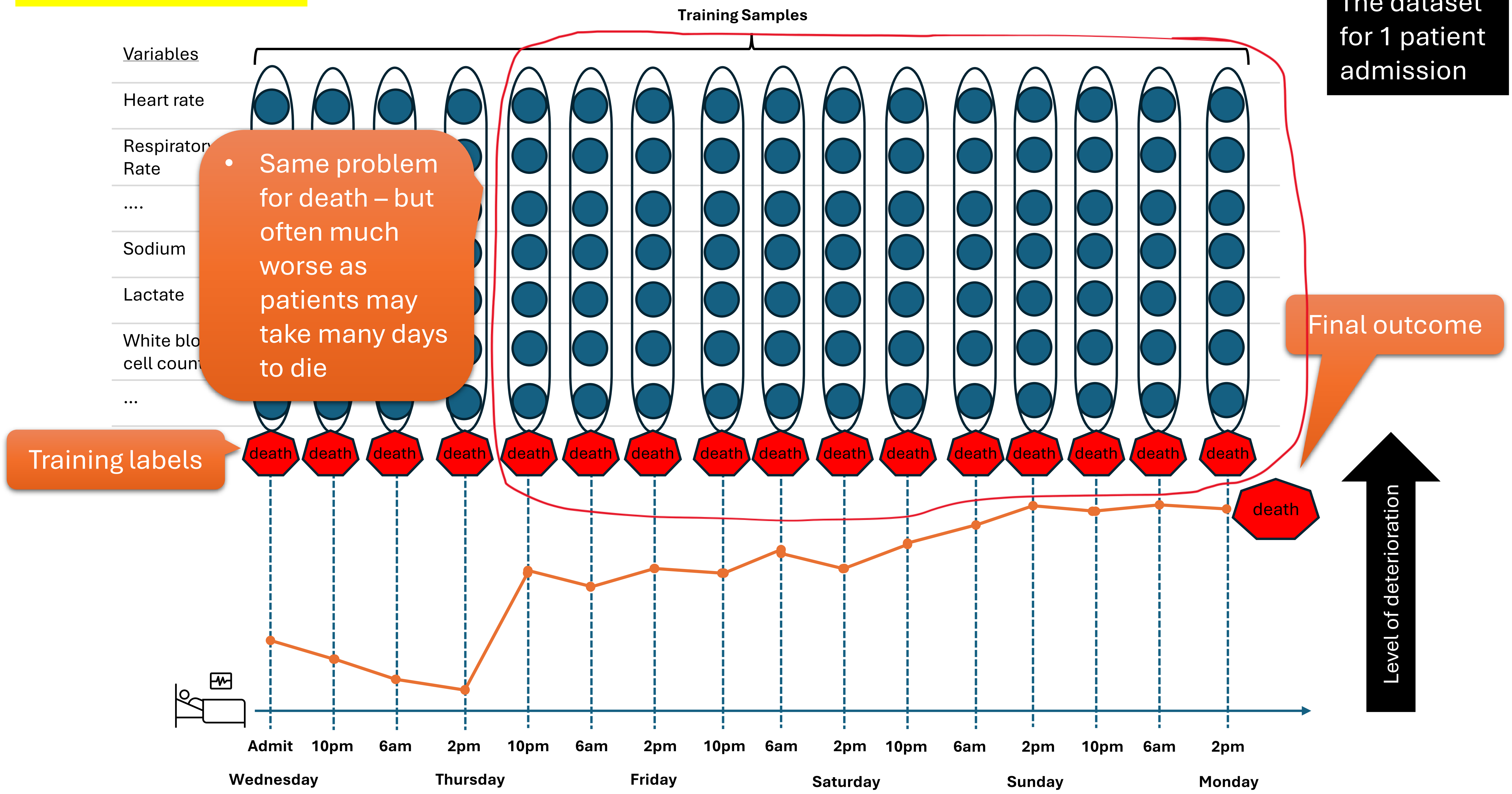
Case B: UPICU as label



Case B: UPICU as label

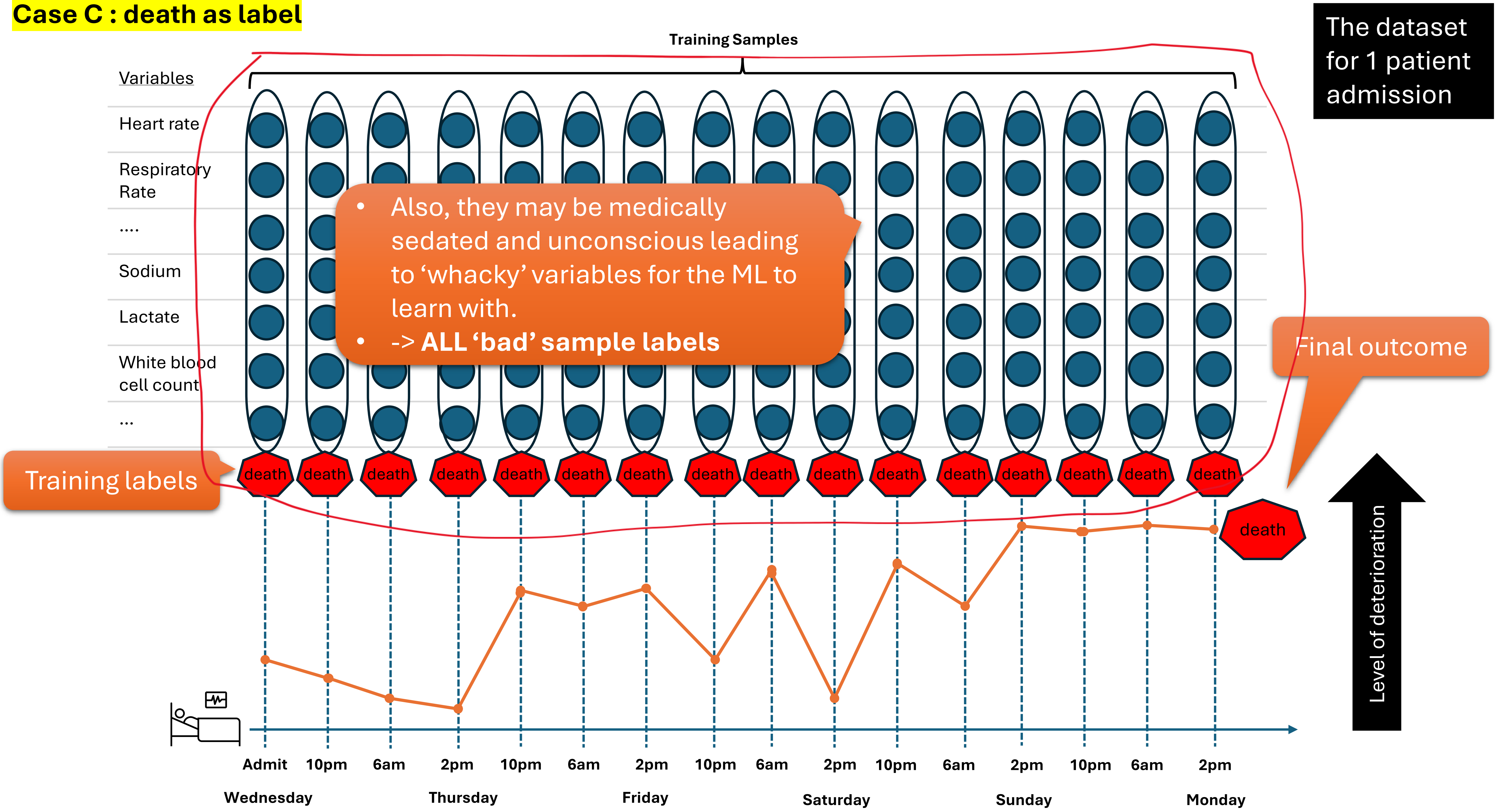


Case C : death as label

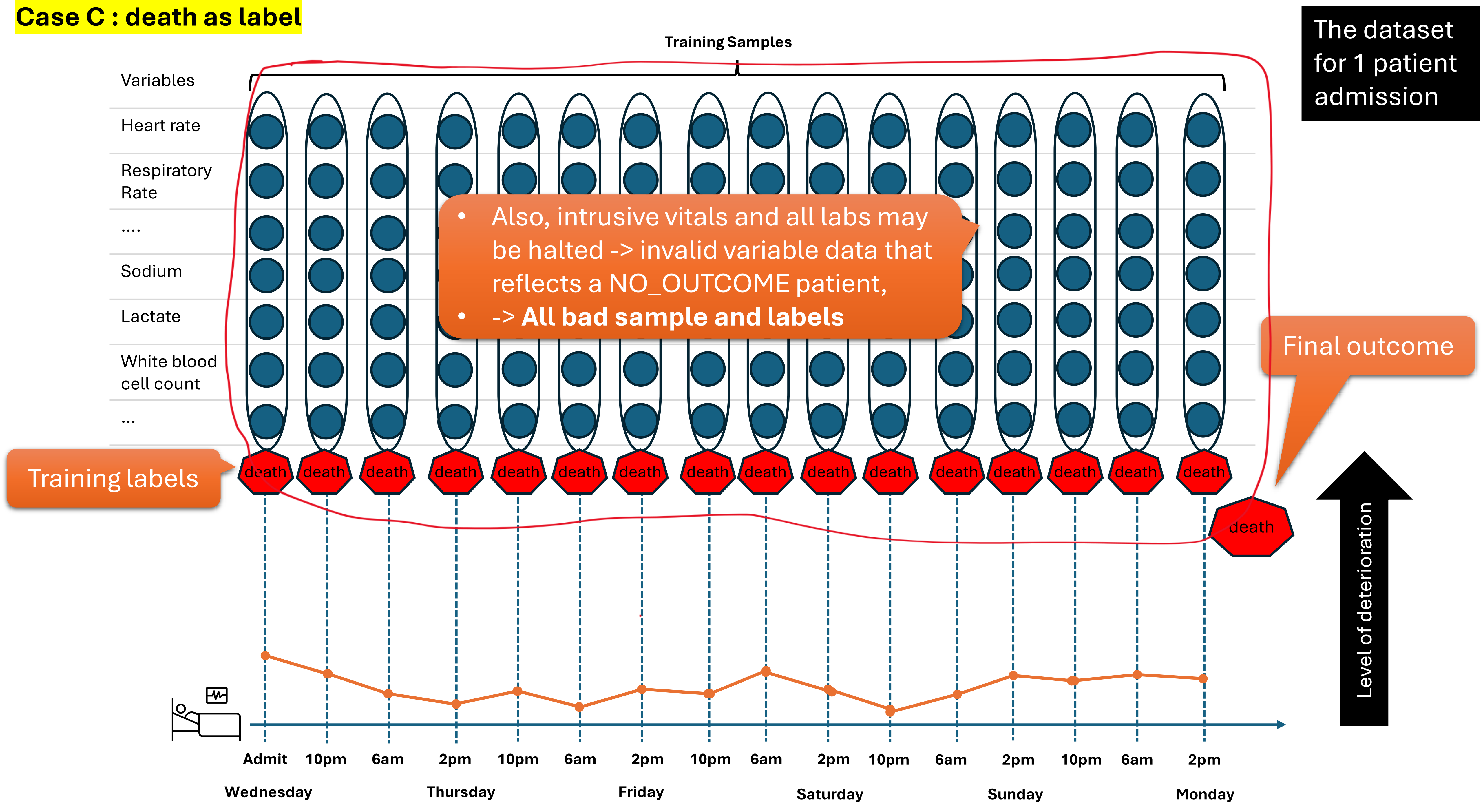


The dataset for 1 patient admission

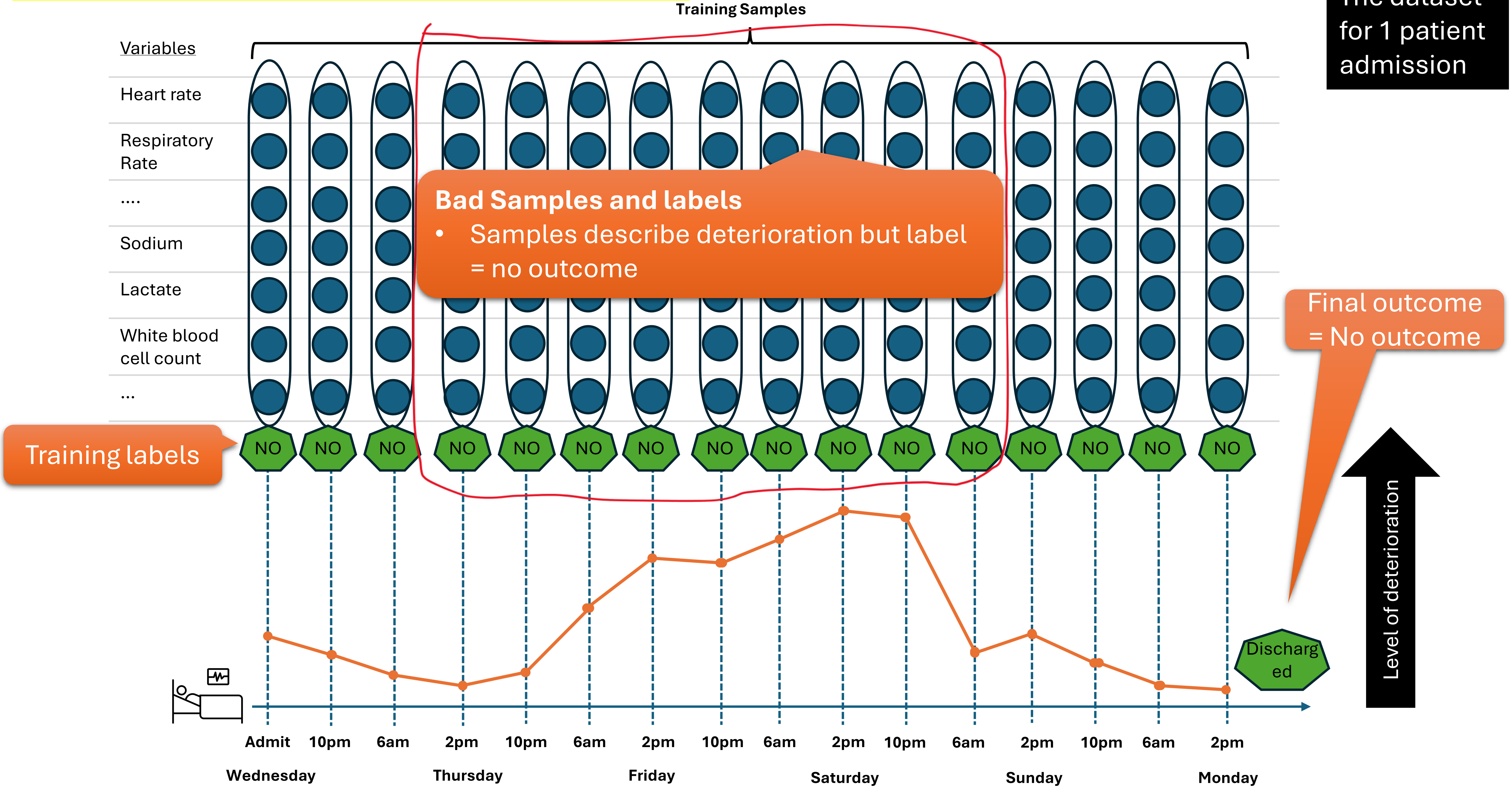
Case C : death as label



Case C : death as label



Case D: Seriously deteriorated – NO_OUTCOME label



In Summary

Unplanned ICU transfer (UPICU)

- The best outcome label we have, except when it occurs hours after serious deterioration, it is too late

Death

- The worst outcome label in hospitals where patient deaths are rarely ‘unexpected’
- Death often occurs long after the serious deterioration event occurred, with or without sedating medication with no labs/vitals taken, resulting in ‘confusing’ or wrong training labels

Serious deterioration

- Currently missing altogether – patients who seriously deteriorate and then recover on the ward

Our Research

Objectives

To develop a ‘seriously deteriorated patient’ (SDP) label for machine learners to use

To replace the ‘death’ label with the SDP and provide a composite SDP + UPICU label, whichever occurs first

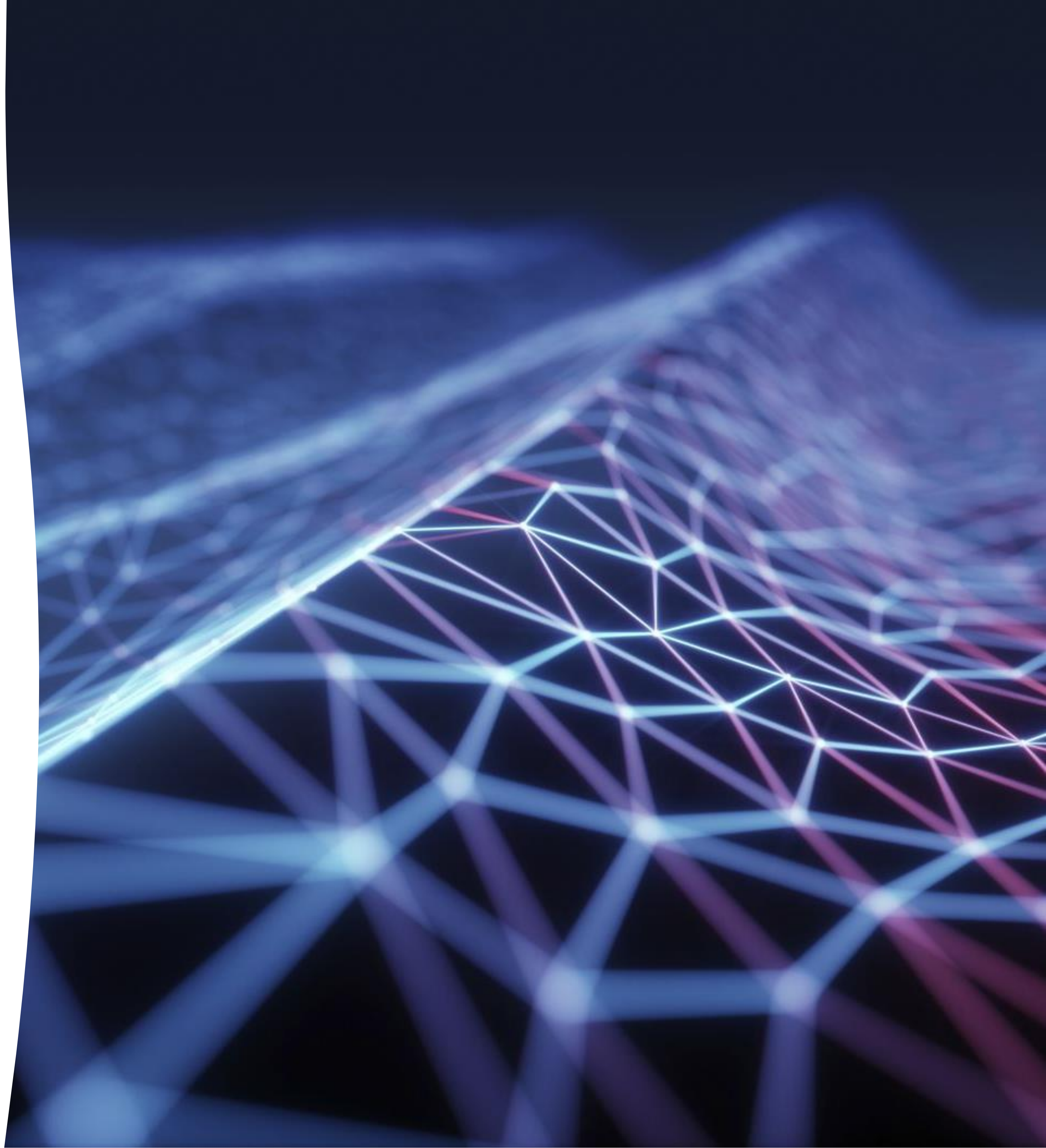
Criteria for a ‘good’ Seriously Deteriorated patient (SDP) label

It utilises widely captured variables that any research team around the world could construct from their Electronic Medical Record (EMR)

It is clinically valid: *any* clinician would agree that patients within the SDP cohort are very sick (seriously deteriorated)

It is more accurate than MET level alerts (i.e., 1 in 20 patients go to ICU)

It is a better composite label with UPICU than [Death + UPICU]



Method

Step 1: Identify a rule-based SDP candidate to identify seriously deteriorated patients

Step 2: Improve the SDP candidate using prior work and evidence

Step 3: Validate the novel SDP label against traditional labels

Method

Step 1: Identify a rule-based SDP candidate to identify seriously deteriorated patients

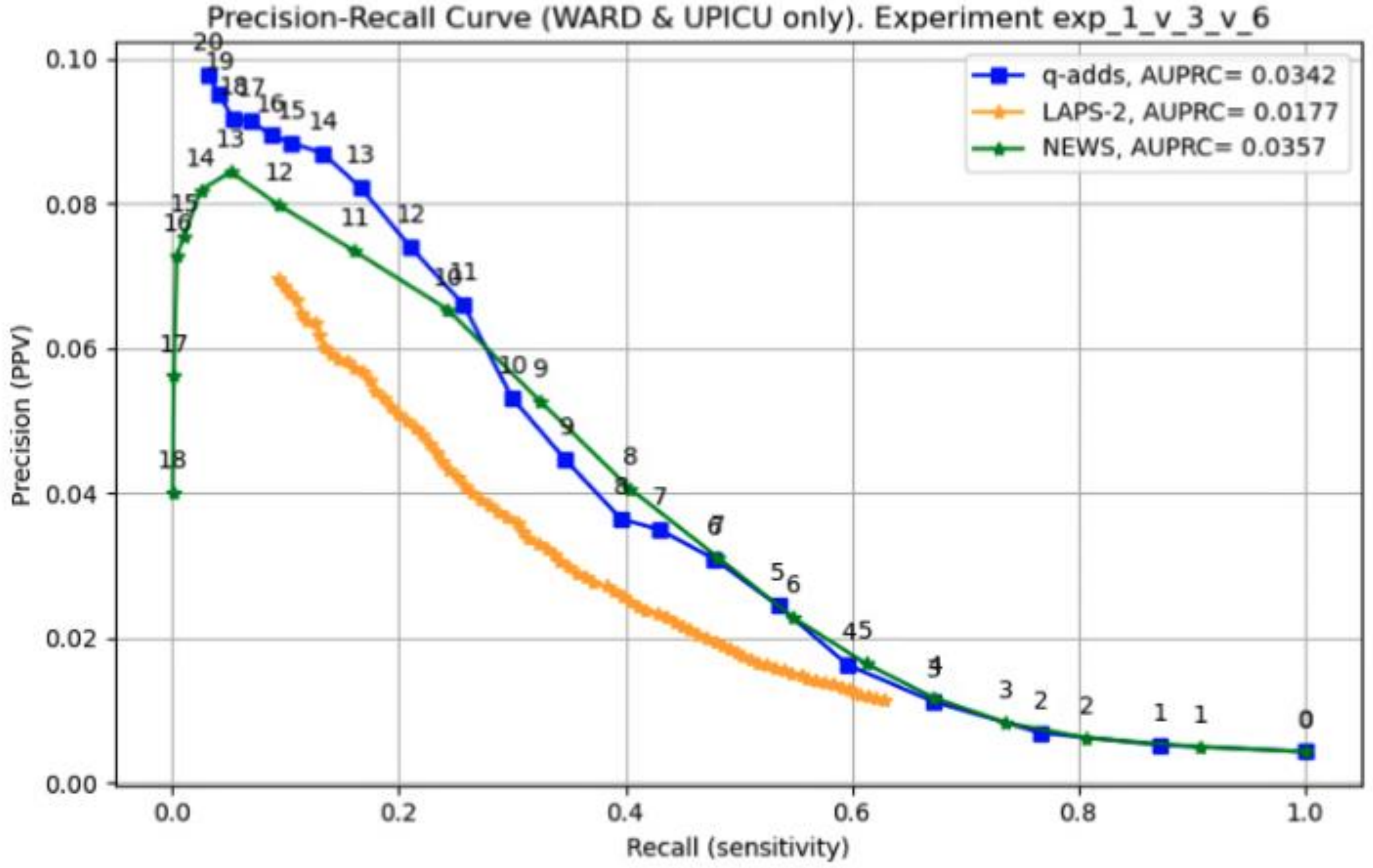
Step 2: Improve the SDP candidate using prior work and evidence

Step 3: Validate the novel SDP label against traditional labels

Step 1 Results

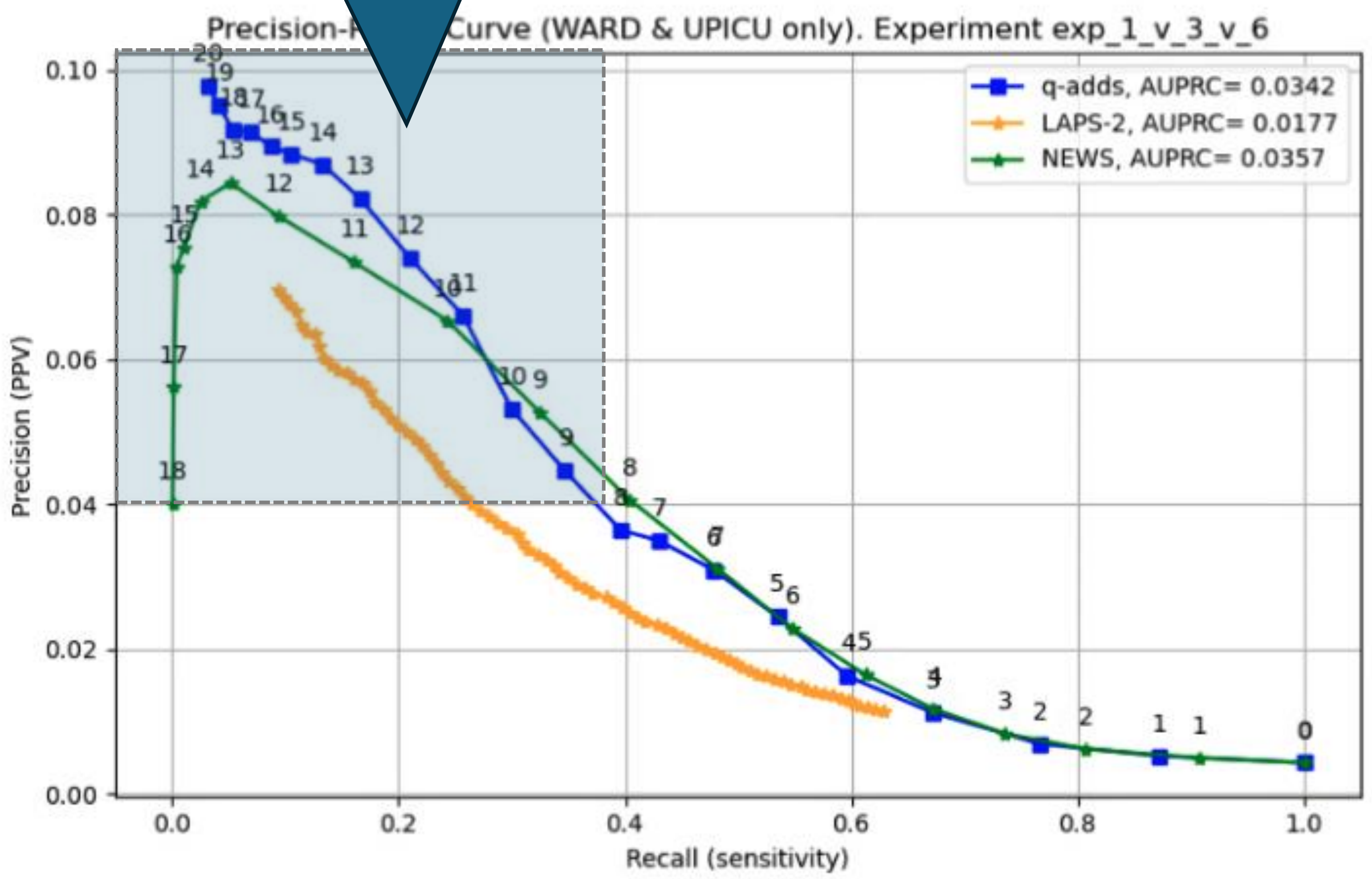
Start with rule-based early warning tools that are currently in use:

1. National Early Warning Score (NEWS) – in UK
2. Laboratory-based Acute Physiology Score (LAPS2)
3. Queensland Adult Deterioration Detection Score (Q-ADDS)

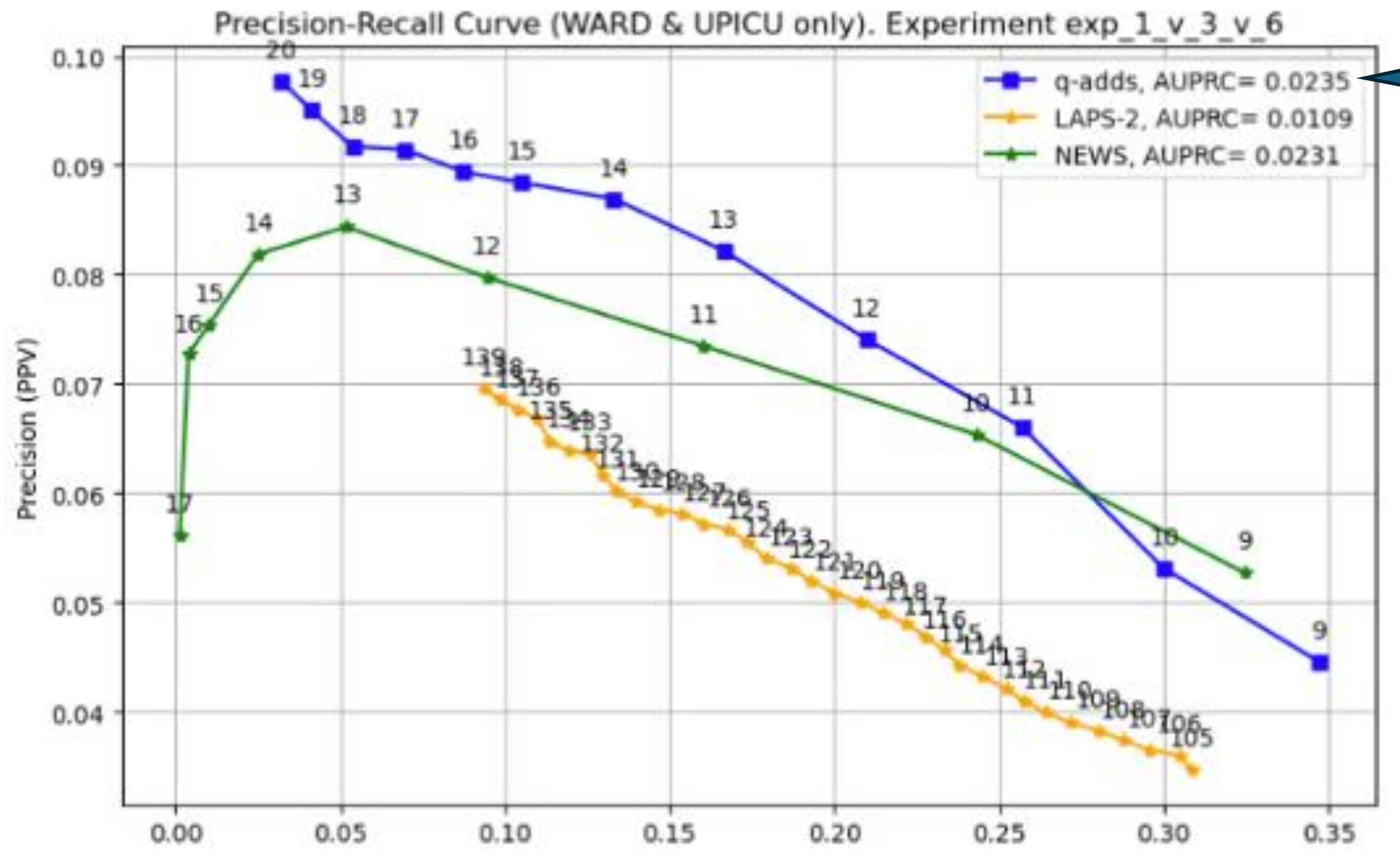


Candidate SDP ability to identify UPICU patients

Specifically interested in serious deterioration beyond MET level



Ability to identify UPICU patients



Select Q-adds: most effective candidate at identifying seriously deteriorated patients

Ability to identify UPICU patients

Method

Step 1: Identify a rule-based SDP candidate to identify seriously deteriorated patients

Step 2: Improve the SDP candidate using prior work and evidence

Step 3: Validate the novel SDP label against traditional labels

	8	4	3	2	1	0	1	2	3	4	8
Respiratory Rate	≤8				9-12	13-20	21-24	25-30		31-35	≥36
SpO2		≤84		85-89	90-91	≥92					
Oxygen Flow Rate or FiO2						≤1	2-5	6-11		12-14	≥15
						≤27	28-40	41-50		51-59	≥60
SBP	≤79	80-89		90-99	100-109	110-159	160-169	170-199		≥200	
Heart Rate	≤39			40-49		50-99	100-109	110-129	130-139	140-159	≥160
Temp		≤34		34.1-35	35.1-36	36.1-37.9	38-38.4	≥38.5			
AVPU	P/U				V	A					

Method to optimise

- Using Q-adds as the baseline SDP candidate
- Identify evidence based variables to test on the baseline
- Select positive contributors and finalise SDP score index
- Select cut-off score to define SDP cohort

Update candidates & rationale

change	rationale
high temp removed	• High temp not an indicator of physiological failure • influenced by paracetamol
T2RF	• Indicator of ventilation failure • blood gas performance indicator of clinician concern for patient
Neurological indicators (AVPU/GCS) removed	• Low conscious level common, and while warrants a MET call, most don't go to ICU • Supported by data/evidence (ref ADI) • common and often documented 'unresponsive' at end of life or transient (eg secondary to a seizure) and remain on the ward as long as no other concerning VS indicators
Take either SBP or MAP, whichever scoring worse	• ward charts trigger on SBP <90 or <80 • critical care inotrope decisions based on MAP ~<65 • tried adding a score 4 for MAP 55, 60, 65 cutoffs and all were slightly better but 55 best of the 3
more extreme HR scores higher	• to account for asymptomatic AF, a score of 5 was given to a HR >160
lactate	• novel, perfusion, sepsis • increasingly performed with sepsis awareness campaigns
K+	• severe Hyperkalemia life threatening and requires monitored bed and aggressive treatment
HCT	• used in LAPS2, relative LAPS2 ranges applied
albumin	• used in LAPS2, relative LAPS2 ranges applied • these were very low values • very low albumin is common in critically ill patients, esp sepsis
Ur:Cr ratio	• sign of hypovolemia • used in LAPS2, relative LAPS2 ranges applied
bicarbonate	• Acidosis in critical care considered bad

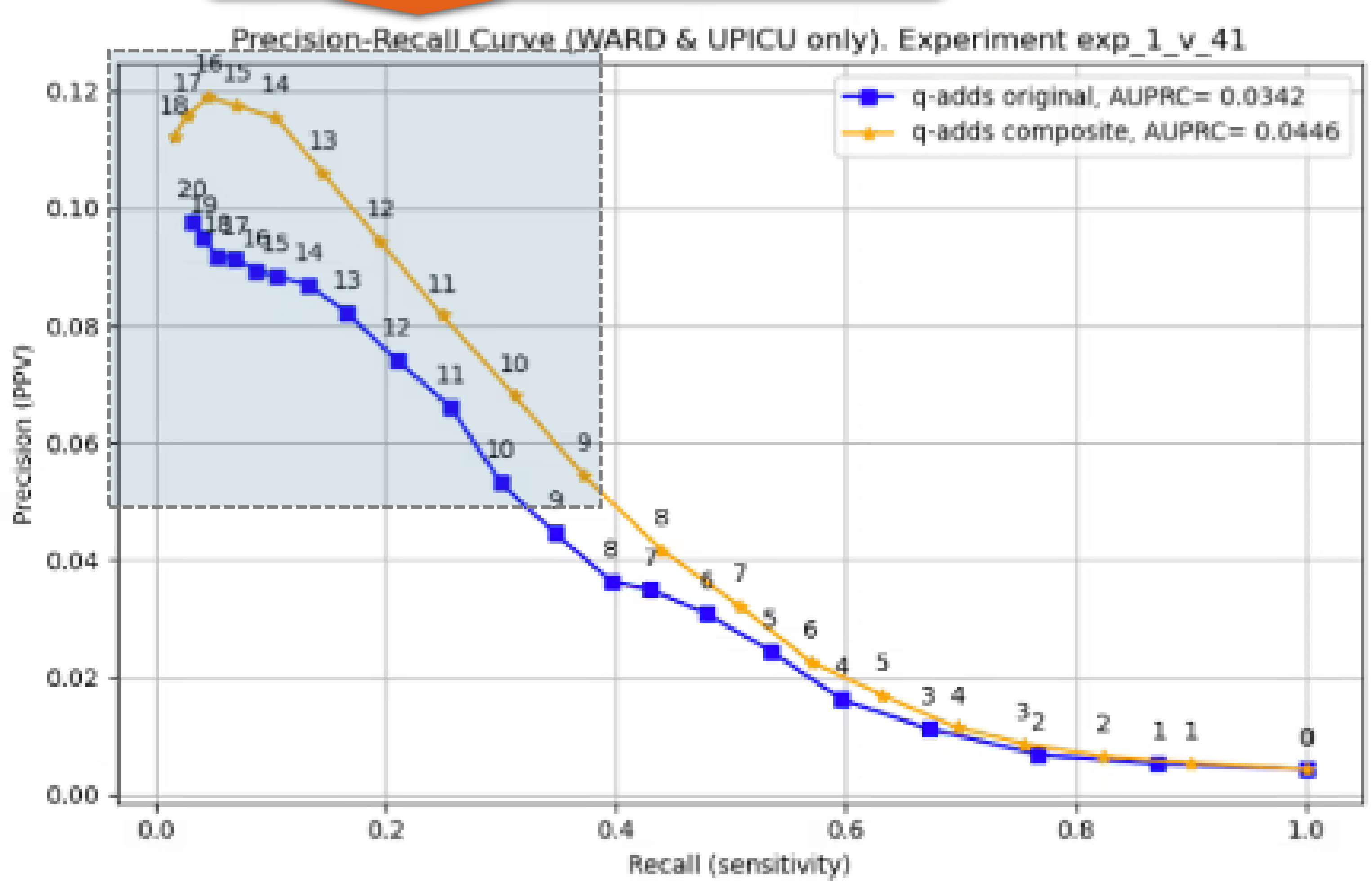
change	rationale
high temp removed	• High temp not an indicator of physiological failure • influenced by paracetamol
T2RF	• Indicator of ventilation failure • blood gas performance indicator of clinician concern for patient
Neurological indicators (AVPU/GCS) removed	• Low conscious level common, and while warrants a MET call, most don't go to ICU • Supported by data/evidence (ref ADI) • common and often documented 'unresponsive' at end of life or transient (eg secondary to a seizure) and remain on the ward as long as no other concerning VS indicators
Take either SBP or MAP, whichever scoring worse	• ward charts trigger on SBP <90 or <80 • critical care inotrope decisions based on MAP ~<65 • tried adding a score 4 for MAP 55, 60, 65 cutoffs and all were slightly better but 55 best of the 3
more extreme HR scores higher	• to account for asymptomatic AF, a score of 5 was given to a HR >160
lactate	• novel, perfusion, sepsis • increasingly performed with sepsis awareness campaigns
K+	• severe Hyperkalemia life threatening and requires monitored bed and aggressive treatment
HCT	• used in LAPS2, relative LAPS2 ranges applied
albumin	• used in LAPS2, relative LAPS2 ranges applied • these were very low values • very low albumin is common in critically ill patients, esp sepsis
Ur:Cr ratio	• sign of hypovolemia • used in LAPS2, relative LAPS2 ranges applied
bicarbonate	• Acidosis in critical care considered bad

Update candidates & rationale

Update candidates & rationale

change	rationale	Improvement result	
high temp removed	• High temp not an indicator of physiological failure • influenced by paracetamol	no	-
T2RF	• Indicator of ventilation failure • blood gas performance indicator of clinician concern for patient	yes	+
Neurological indicators (AVPU/GCS) removed	• Low conscious level common, and while warrants a MET call, most don't go to ICU • Supported by data/evidence (ref ADI) • common and often documented 'unresponsive' at end of life or transient (eg secondary to a seizure) and remain on the ward as long as no other concerning VS indicators	yes	++
Take either SBP or MAP, whichever scoring worse	• ward charts trigger on SBP <90 or <80 • critical care inotrope decisions based on MAP ~<65 • tried adding a score 4 for MAP 55, 60, 65 cutoffs and all were slightly better but 55 best of the 3	yes	+
more extreme HR scores higher	• to account for asymptomatic AF, a score of 5 was given to a HR >160	no	=
lactate	• novel, perfusion, sepsis • increasingly performed with sepsis awareness campaigns	yes	++
K+	• severe Hyperkalemia life threatening and requires monitored bed and aggressive treatment	no	-
HCT	• used in LAPS2, relative LAPS2 ranges applied	yes	+
albumin	• used in LAPS2, relative LAPS2 ranges applied • these were very low values • very low albumin is common in critically ill patients, esp sepsis	yes	+++
Ur:Cr ratio	• sign of hypovolemia • used in LAPS2, relative LAPS2 ranges applied	yes	+
bicarbonate	• Acidosis in critical care considered bad	no	-

30% improvement in AUPRC with final selection of upgrade candidates



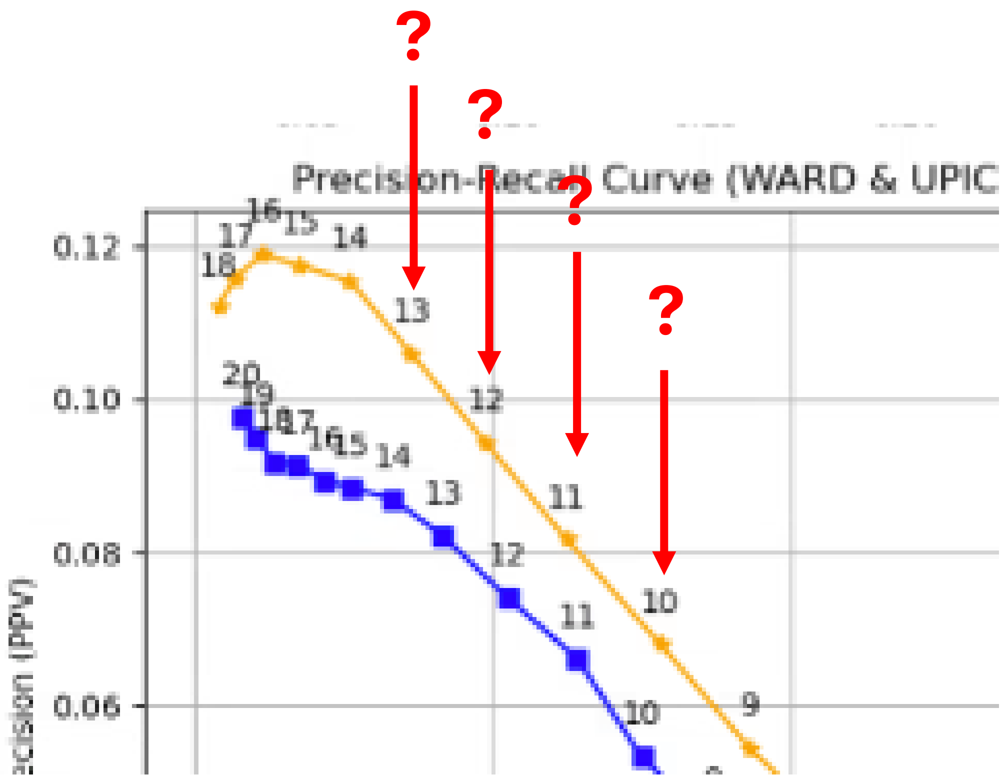
Before – After comparison of SDP

Selecting the SDP score cut-off: Trade-off

- Setting it higher ensures we capture only seriously deteriorated patients, but we capture fewer patients in the cohort

- Setting it lower captures more patients that deteriorate and recover, but if set too low may include less-than-seriously deteriorated patients

‘the SDP label’



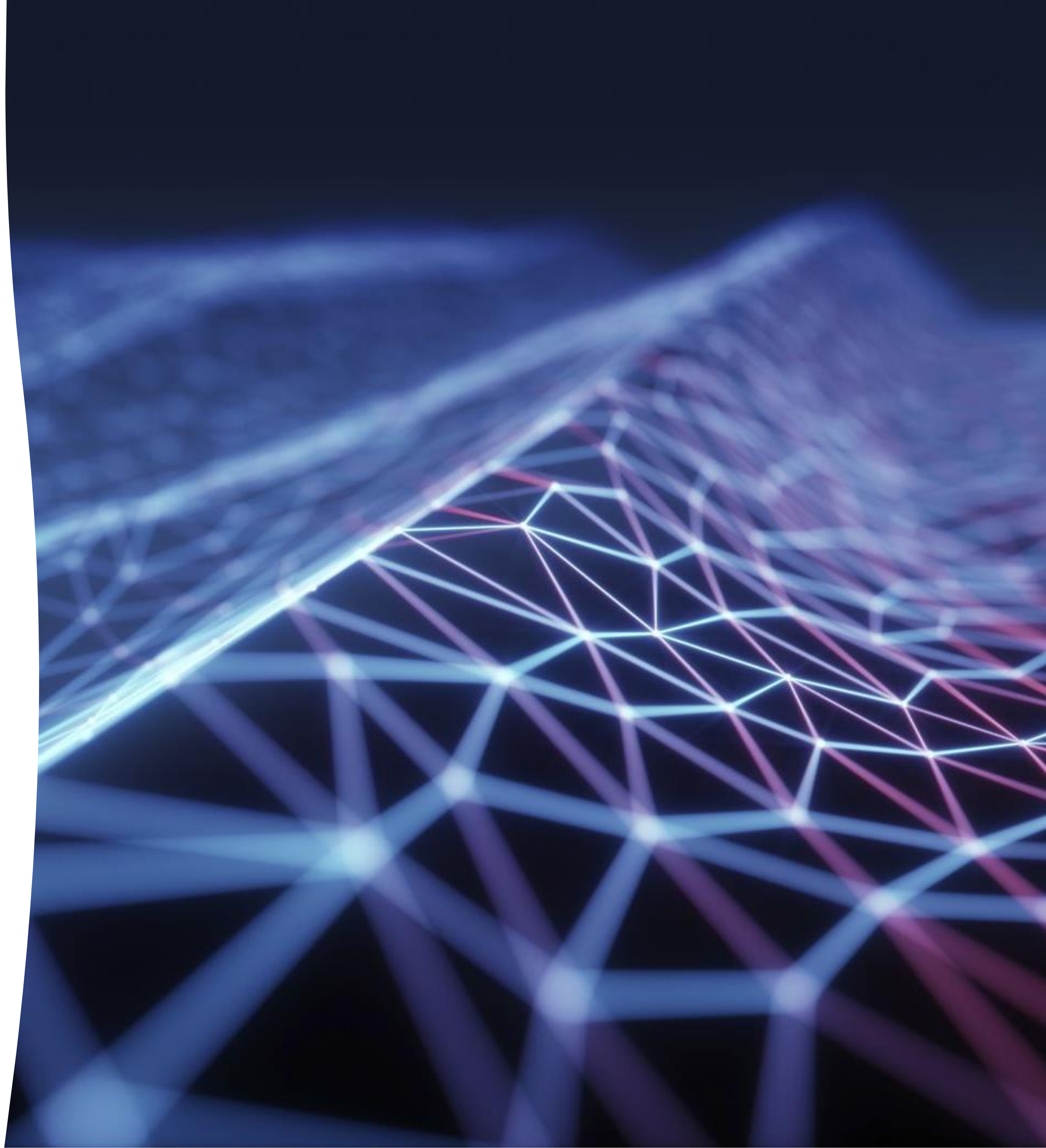
Criteria for a ‘good’ Seriously Deteriorated patient (SDP) label

It utilises widely captured variables that any research team around the world could construct from their Electronic Medical Record (EMR)

It is clinically valid: *any* clinician would agree that patients within the SDP cohort are very sick (seriously deteriorated)

It is more accurate than MET level alerts (i.e., 1 in 20 patients go to ICU)

It is a better composite label with UPICU than [Death + UPICU]



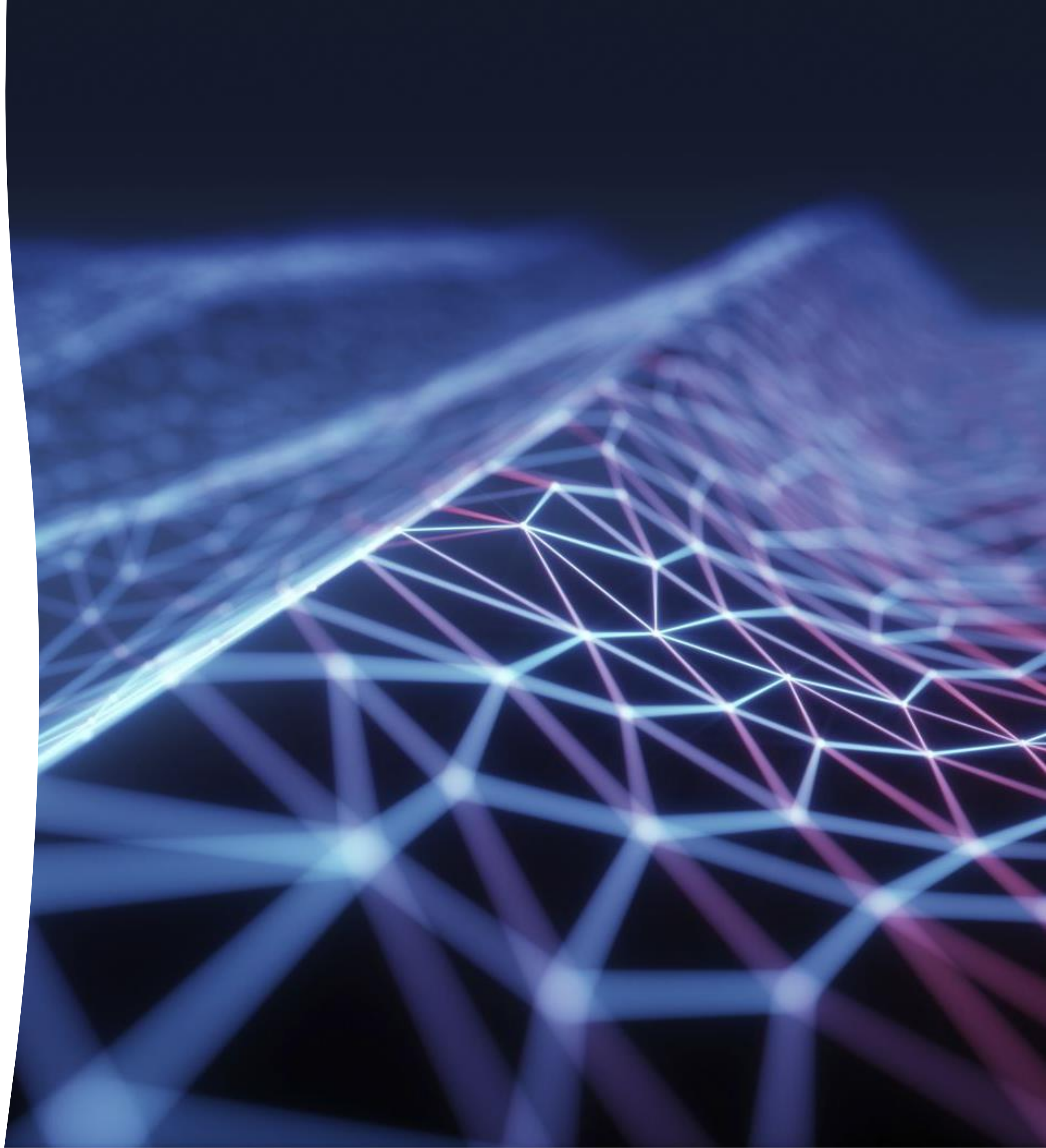
Criteria for a ‘good’ Seriously Deteriorated patient (SDP) label

It utilises widely captured variables that any research team around the world could construct from their Electronic Medical Record (EMR)

It is clinically valid: *any* clinician would agree that patients within the SDP cohort are very sick (seriously deteriorated)

It is more accurate than MET level alerts (i.e., 1 in 20 patients go to ICU)

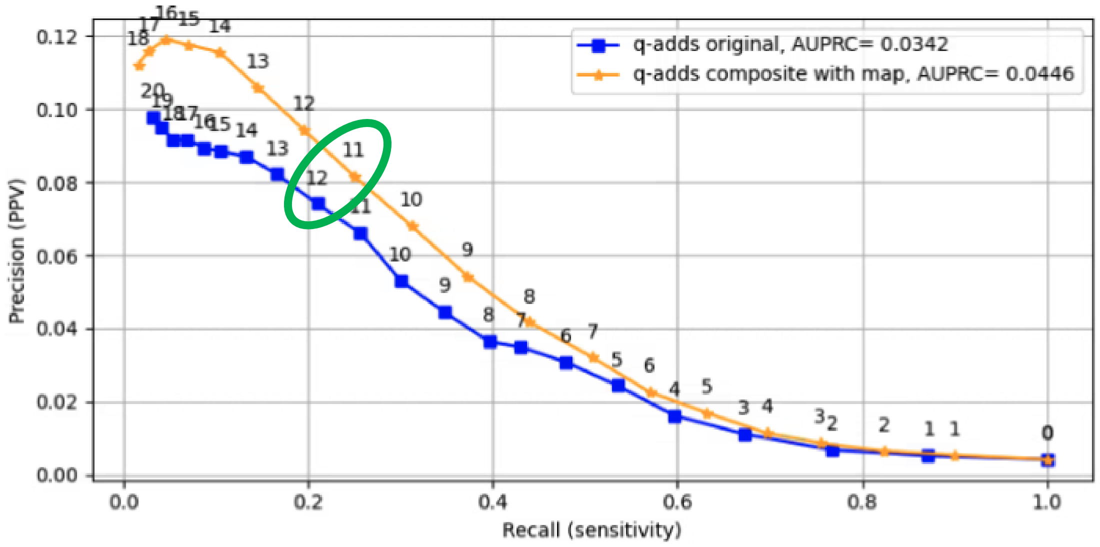
It is a better composite label with UPICU than Death



Clinically valid?

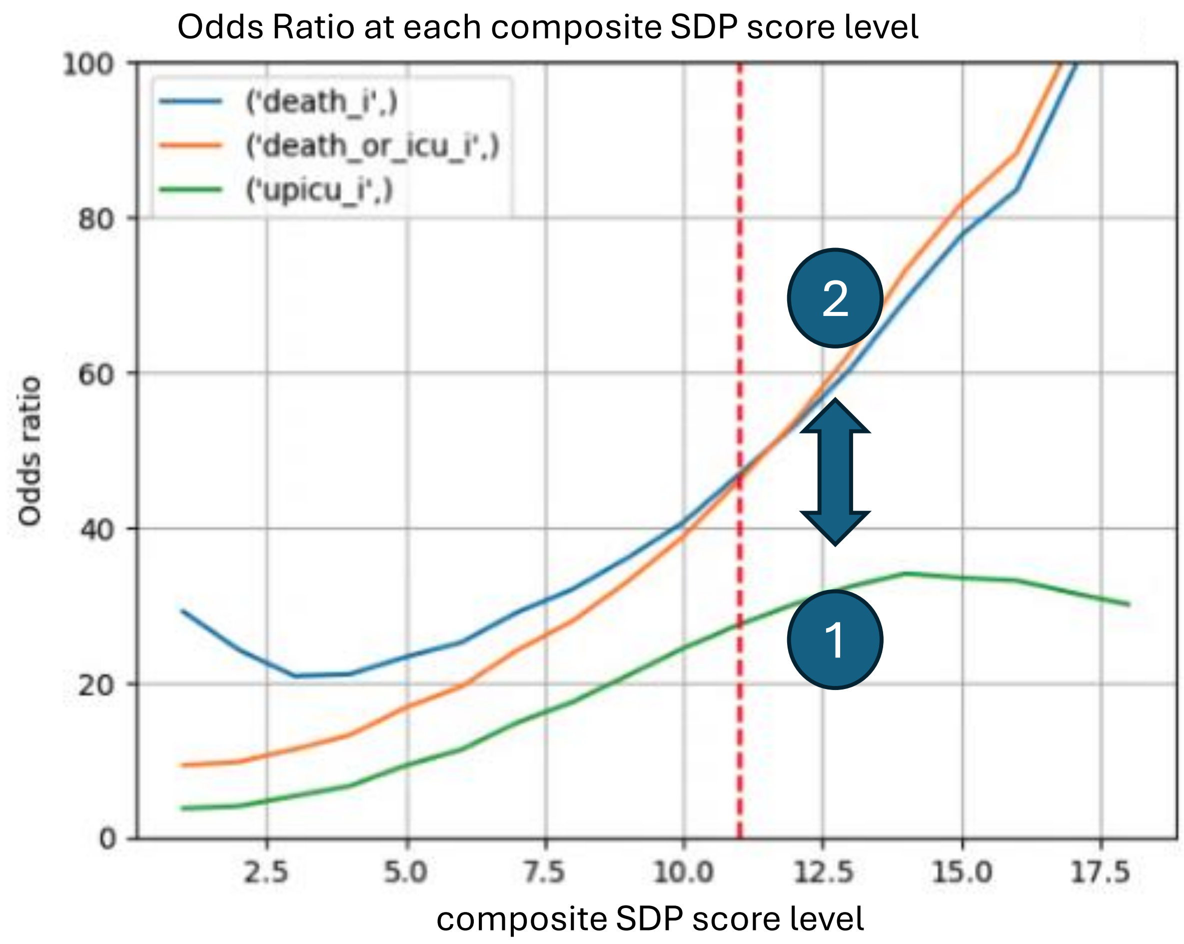
- True MET call data suggests UPICU transfers resulting from MET calls have a median score of ≥12

- SDP@11 is closest to this but with more precision and sensitivity



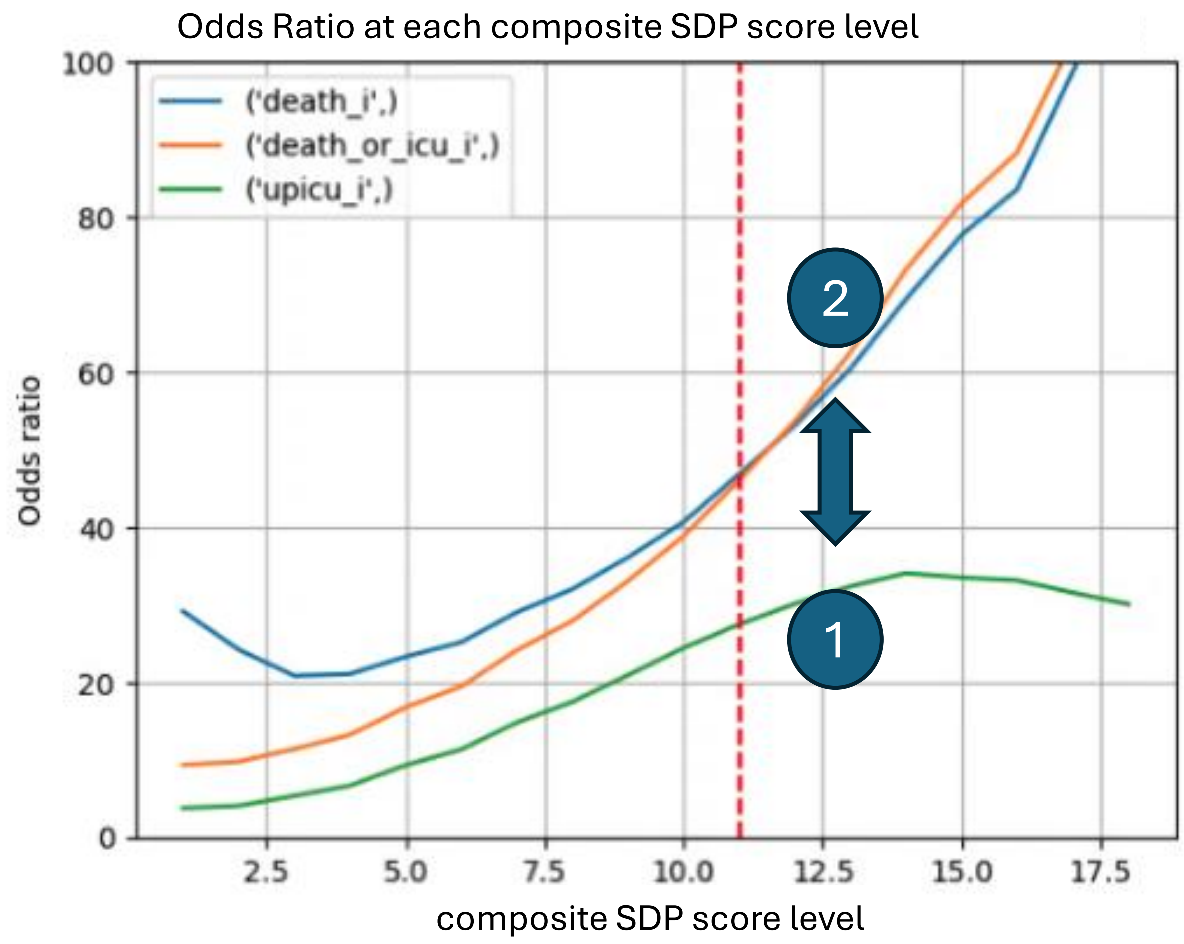
Clinically valid?

- At SDP score ~12-13 there is a plateau and then decline in the likelihood of a patient being moved to ICU
- This is associated with an increase in likelihood a patient will die
  - Hypothesis: this is the score point at which ICU-eligible patients have been removed from the cohort, and death is likely AND accepted in those who remain



Clinically valid?

- At SDP@11 – it is very clear the patient is seriously deteriorated



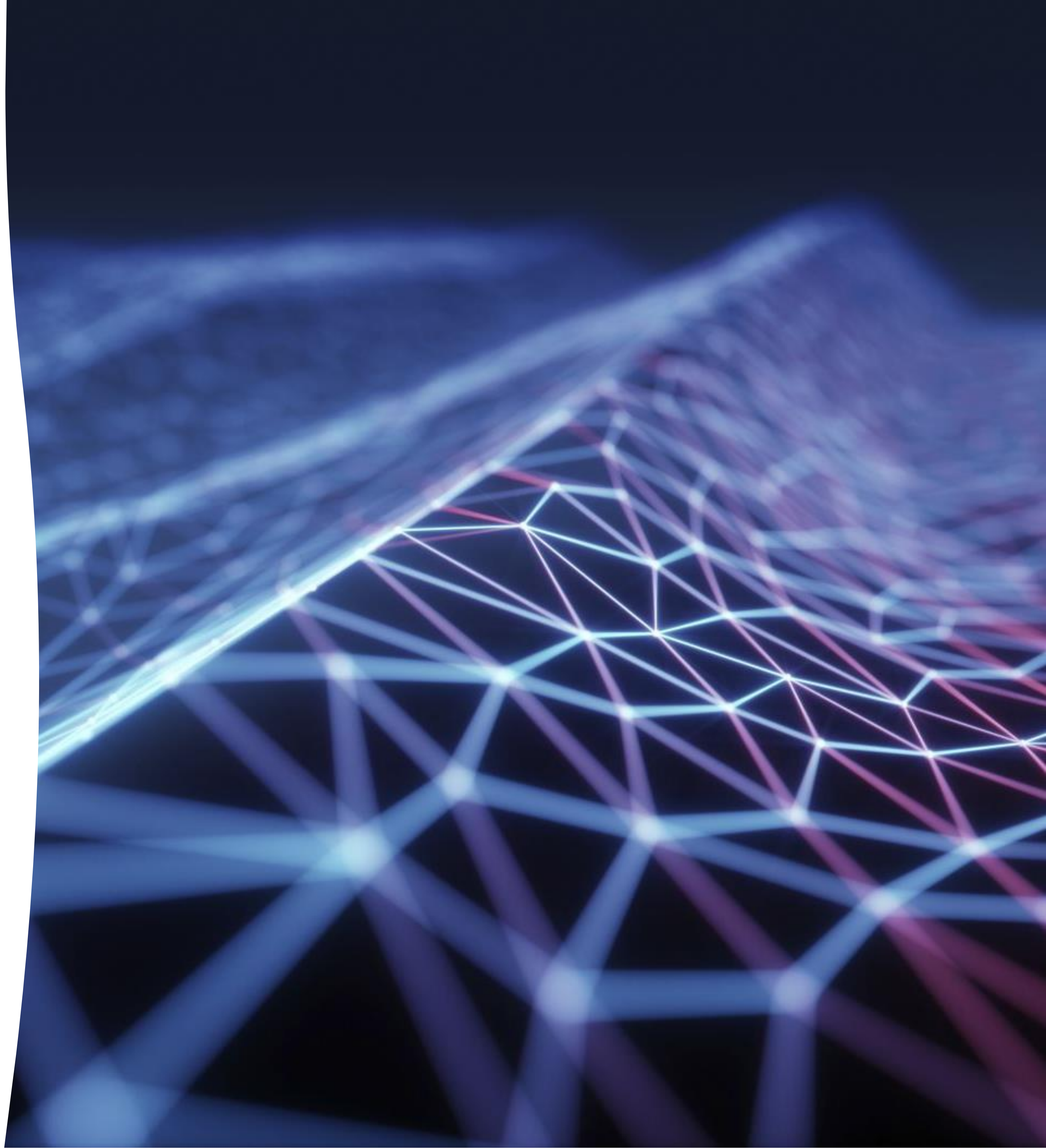
Criteria for a ‘good’ Seriously Deteriorated patient (SDP) label

It utilises widely captured variables that any research team around the world could construct from their Electronic Medical Record (EMR)

It is clinically valid: *any* clinician would agree that patients within the SDP cohort are very sick (seriously deteriorated)

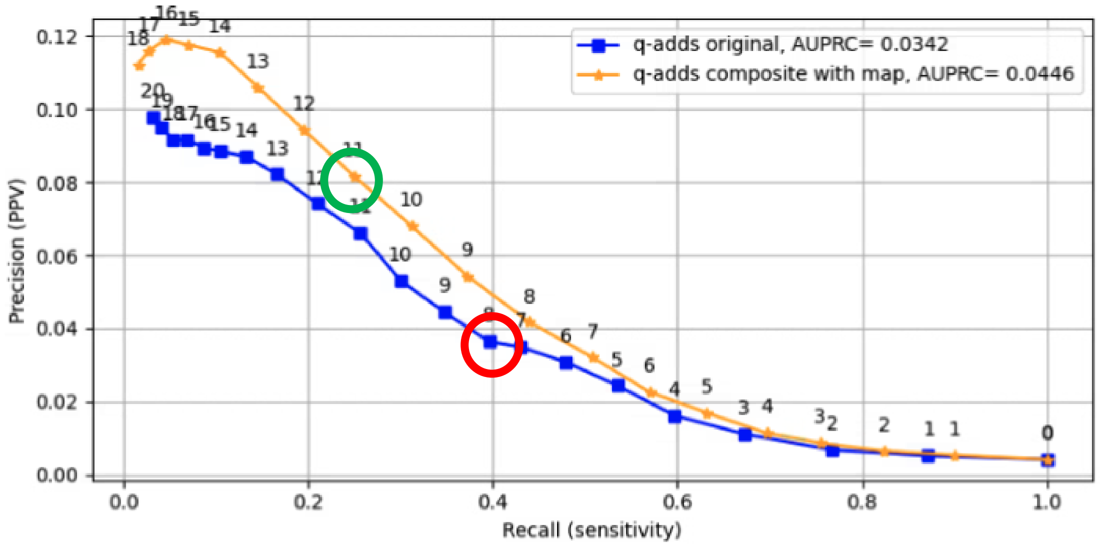
It is more accurate than current MET level alerts
(i.e., 1 in 20 patients go to ICU)

It is a better composite label with UPICU than Death



Is it more accurate than MET alert level?

- SDP@11 has twice the precision than Q-ADDS MET level score of 8



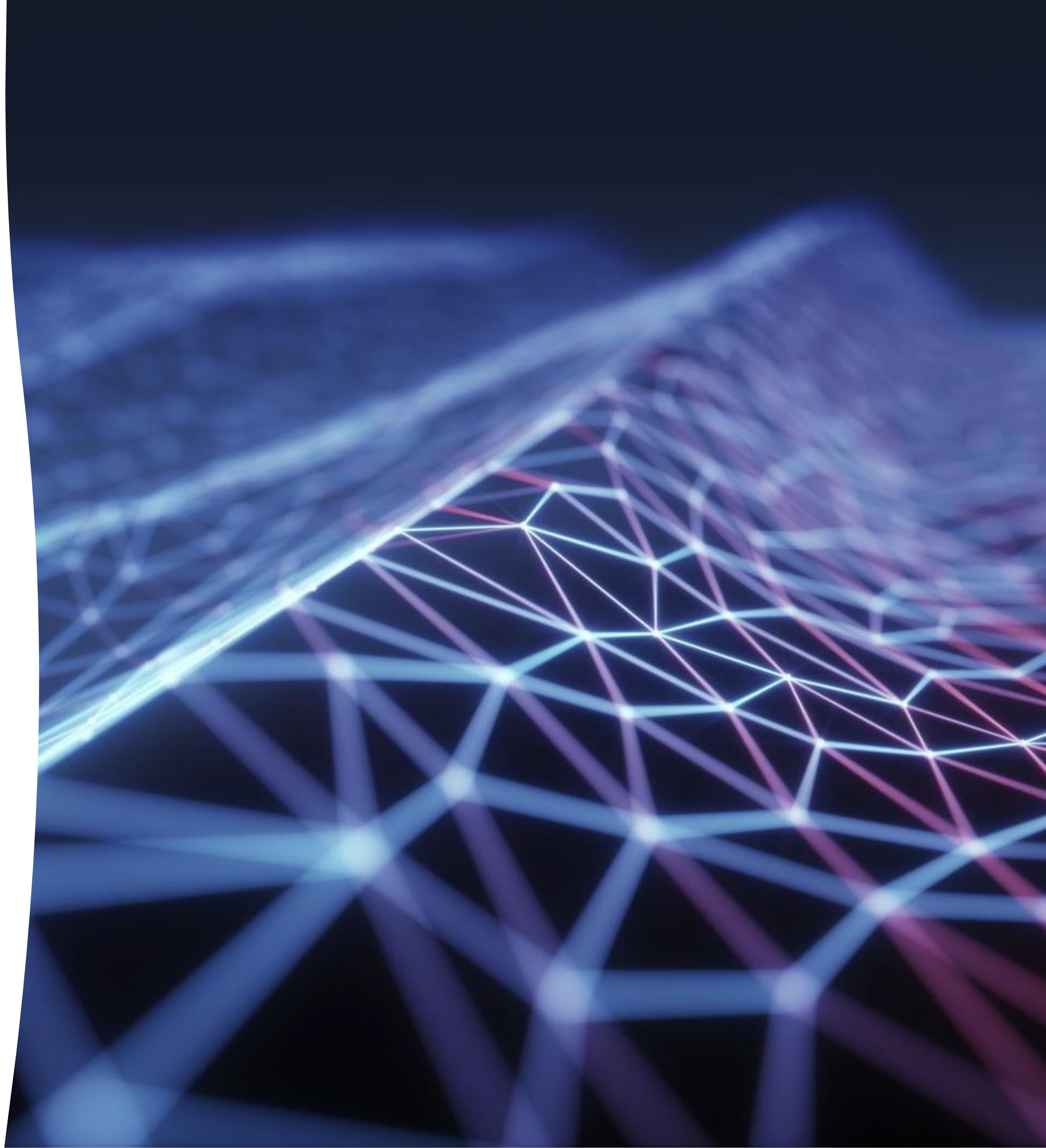
Criteria for a ‘good’ Seriously Deteriorated patient (SDP) label

It utilises widely captured variables that any research team around the world could construct from their Electronic Medical Record (EMR)

It is clinically valid: *any* clinician would agree that patients within the SDP cohort are very sick (seriously deteriorated)

It is more accurate than MET level alerts (i.e., 1 in 20 patients go to ICU)

It is a better composite label with UPICU than Death+UPICU



What does ***a better composite label*** with UPICU than Death, mean?

Enable the training of an algorithm that:

- Doesn’t damage the prediction of UPICU patients
- Improves the prediction of seriously deteriorated patients

Method

Identify a rule-based SDP candidate label to identify seriously deteriorated patients

Improve the SDP candidate using prior work and evidence

Validate the candidate SDP label (SDP@11) against traditional labels

Method

Using the e-CART machine learning algorithm:

- Train the algorithm on 36hrs data leading up to the outcome event
- Using 5-fold cross validation across complete dataset

Train eCART on:

- Traditional: UPICU & Death
- Novel: UPICU & SDP@11

Evaluate for:

- Effectiveness in predicting UPICU patients
- Effectiveness in predicting SDP patients

Dataset characteristics

- An episode is a continuous ward care period, excluding time in surgery and PACU which ends in discharge, death or transfer to ICU.

Number of episodes: 908,047

Episode labels:

UPICU (%)	4021	(0.44%)
Death (%)	6959	(0.77%)
SDP@11 (%)	10,186	(1.12%)
- upicu	- 886	(0.1%)
- death	- 1973	(0.22%)
- recovered	- 7327	(0.81%)

Median hrs between:

- SDP@11 & ICU	7.2 hours
- SDP@11 & death	82 hours

Dataset characteristics

- An episode is a continuous ward care period, excluding time in surgery and PACU which ends in discharge, death or transfer to ICU.

What does this mean?

Number of episodes: 908,047

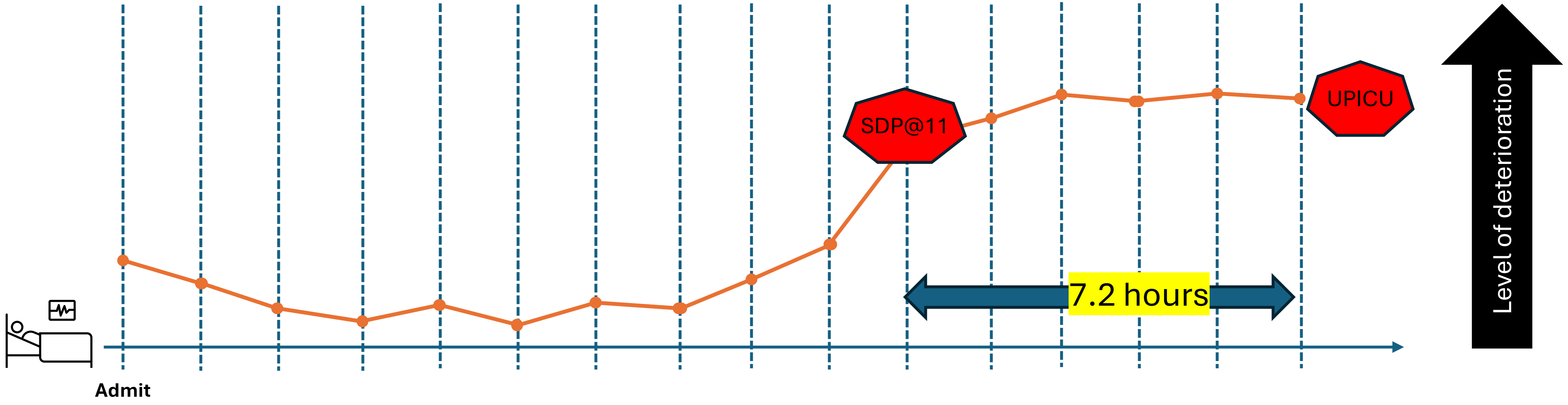
Episode labels:

UPICU (%)	4021	(0.44%)
Death (%)	6959	(0.77%)
SDP@11 (%)	10,186	(1.12%)
- upicu	- 886	(0.1%)
- death	- 1973	(0.22%)
- recovered	- 7327	(0.81%)

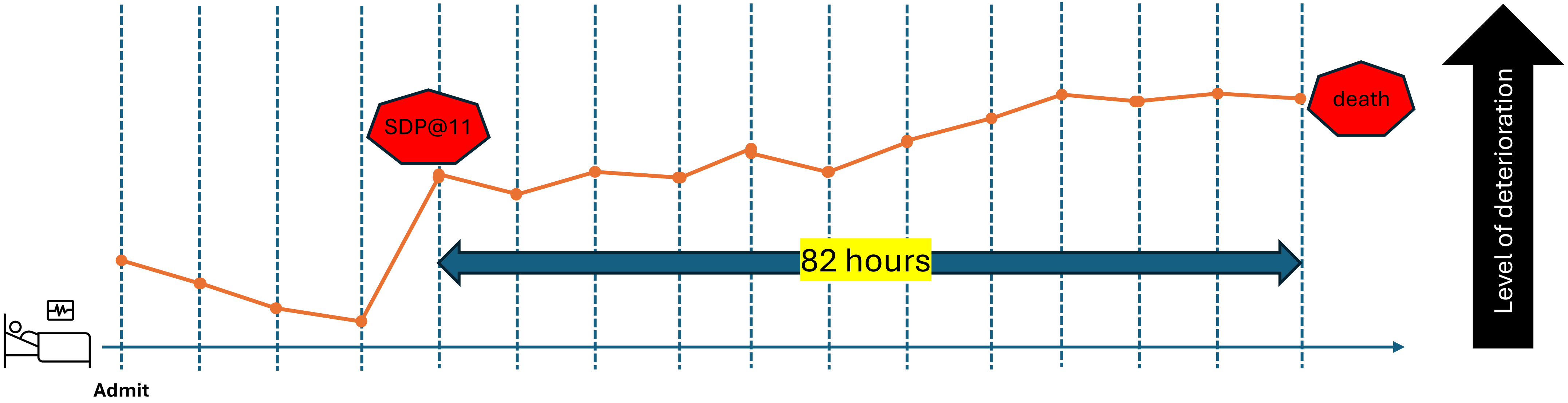
Median hrs between:

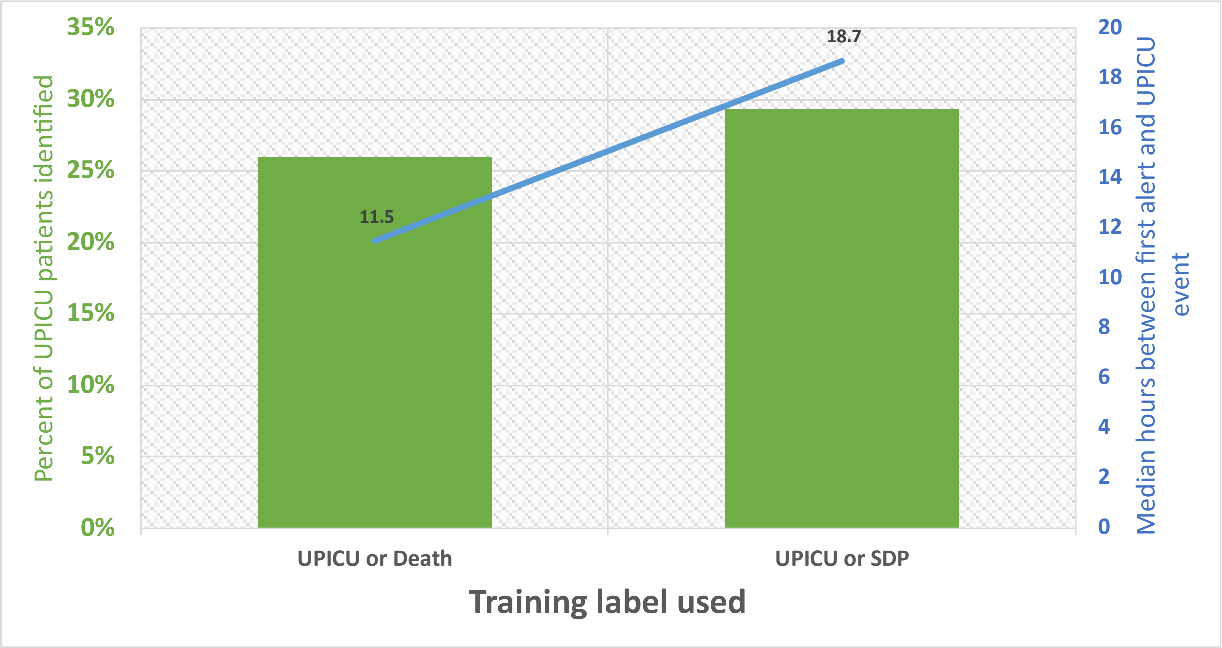
- SDP@11 & ICU	7.2 hours
- SDP@11 & death	82 hours

Case B: UPICU as label



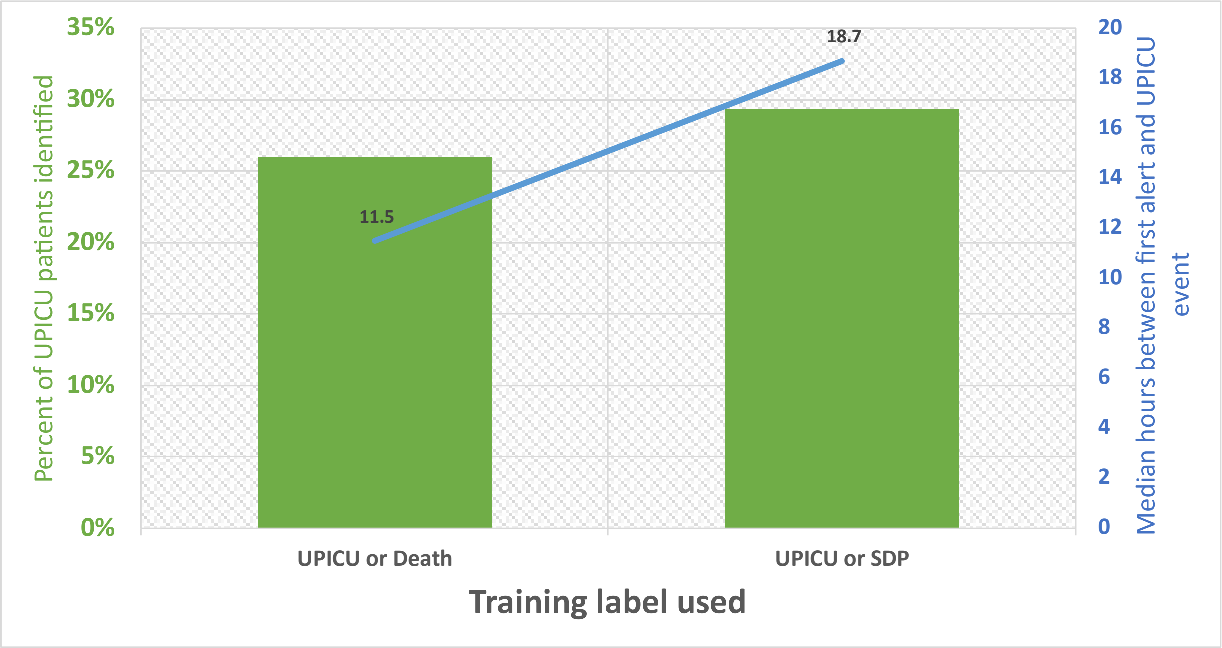
Case C : death as label





Results: Impact of using SDP@11 label on finding UPICU patients

- finds 13% more UPICU patients
- ‘predicts’ UPICU 7 hours earlier

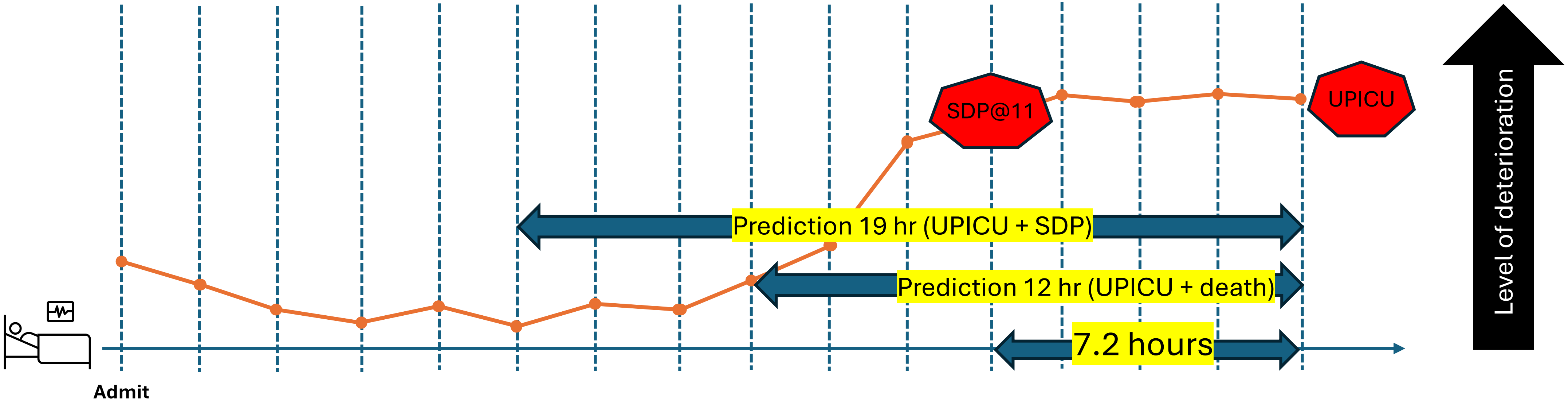


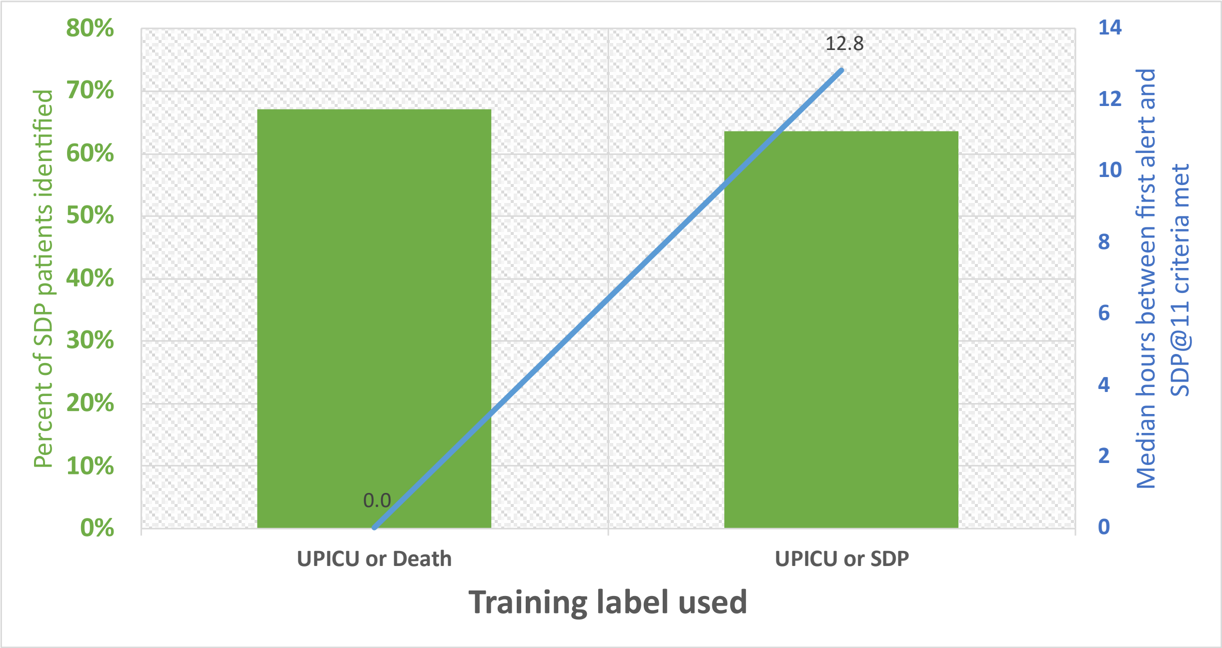
Results: Impact of using SDP@11 label on finding UPICU patients

- finds 13% more UPICU patients
- ‘predicts’ UPICU 7 hours earlier

What does this mean?

Case B: UPICU as label

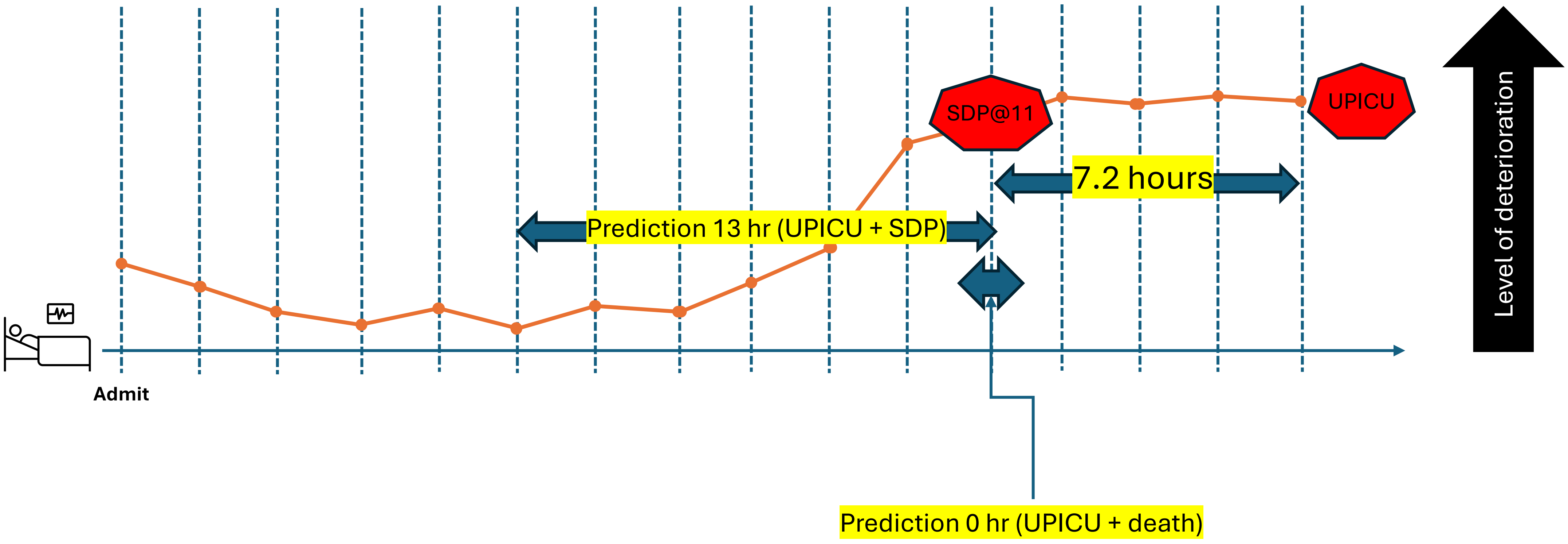




Results: Impact of using SDP@11 label on finding seriously deteriorated patients

- finds 5% **less** seriously deteriorated patients
- ‘predicts’ SDP patients 13 hours earlier

Case B: UPICU as label



Conclusions

New label (SDP@11) for seriously deteriorated patients

with commonly collected EMR variables

Training eCART on SDP@11:

Predict more UPICU patients earlier

Can now predict serious deterioration

Thank you

Questions