Offline Reinforcement Learning for Community-Acquired Pneumonia Management: A Feasibility Study

Alex Beeson[1,2] (✉), Keith Couper[2,3], and Giovanni Montana[1,4,5]

[1] Warwick Manufacturing Group, University of Warwick, Coventry, UK
[2] Warwick Medical School, University of Warwick, Coventry, UK
[3] Critical Care Unit, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
[4] Department of Statistics, University of Warwick, Coventry, UK
[5] Alan Turing Institute, London, UK

Appendix

PIONEER data acquisition

The process for acquiring data first involves submitting an application outlining the scope of the project and the required data fields based on a data dictionary. This application is reviewed by a Data Trust Committee (DTC) which approves the release of data subject to satisfactory discussions about the project and use of requested data items with the applicant. This is then followed by a precise detailing of data items required for the project and inclusion/exclusion criteria for the patient cohort. Once specifics have been agreed, a Data Licence Agreement is drawn-up and following sign-off the data is provided to the applicant via secure file transfer. Data is pseudonymised for the purpose of curation and then anonymised prior to transfer.

Initially, applicants are required to apply for a synthetic data set in order to provide assurances to the DTC that the requested data is suitable for the proposed project. This synthetic data replicates the format of real data, but the items themselves are randomly sampled. Once assurances have been provided, applicants are then able to submit an application for real patient data.

For the purpose of our proposed study, the scope of the project was investigating the feasibility of using offline reinforcement reinforcement learning methods for optimising treatment pathways for patients with CAP. We requested data items relating to the monitoring and treatment of patients with CAP which included observations, tests, assessments, radiography, drug administration, oxygen therapy, demographics, admissions data and outcomes such as 30-day mortality. For our inclusion criteria we specified that patients were of adult age (18 years or older) with a diagnosis of CAP based on an entry of "CURB-65" in medical notes or ICD-10/SNOMED codes. CURB-65 is a diagnostic metric used to determine the severity of CAP upon presenting in hospital [6] while ICD-10 and SNOMED are medical classification systems used by the NHS. The list of codes

used is provided in Table A. We also required that patients received antibiotics within 48 hours of admission, either in hospital or at discharge.

Table A. Diagnosis criteria for community-acquired pneumonia (CAP). NEC = not elsewhere classified.

Source	Value
ICD-10 code	X-J13 (Pneumonia due to Streptococcus pneumoniae)
ICD-10 code	X-J14 (Pneumonia due to Haemophilus influenzae)
ICD-10 code	X-J15 (Bacterial pneumonia, NEC)
ICD-10 code	X-J16 (Pneumonia due to other infectious organisms, NEC)
ICD-10 code	X-J17 (Pneumonia in diseases classified elsewhere)
ICD-10 code	X-J18 (Pneumonia, organism unspecified)
SNOMED	385093006 (Community acquired pneumonia *(disorder)*)
Medical notes	Entry of "CURB65", "CURB-65", "CRB65", "CRB-65"

We first applied for synthetic data and following acquisition we conducted a high-level assessment and concluded that the format was such that RL methods could in principle be applied. We were however limited to any formative analysis due to the underlying randomness of individual items. After providing assurances to the DTC, we applied for and received approval for real patient data. This application yielded data on 36,885 patients who attended hospital between April 2018 and September 2022.

Data structure and composition

A summary of data held within each table is provided in Table B. Figure A provides a visualisation of this data structure for a patient with a single care spell.

Data pre-processing

Here we detail the procedures used to convert the PIONEER source data into the states, actions and rewards required for RL. These procedures centred around tasks relating to cohort refinement, variable selection, variable derivation, state-action pair construction and finally reward function and terminal state definition.

Software

All data pre-processing was carried out using the KNIME data analytics platform [1]. This platform makes use of visual workflows to facilitate data wrangling, allowing users to perform complex data transformations in a clear and structured manner.

Table B. Overview of the main tables in the PIONEER dataset, summarizing core data items and row counts.

Table name	Data items	Rows
Admissions	Demographics, arrival/discharge times, outcomes	47,972
Assessments	Waterlow, falls, ACT10 scores	5,813,787
Diagnosis	Diagnosis name and code	2,005,178
DnarTeal	Do-not-resuscitate preferences	209,736
DrugAdmin	Drug name, dosage, administration route	14,168,655
Imaging	Image type (not the image itself)	98,590
LabAnalysis	Analysis type and measurement	15,563,490
Microbiology	Specimen type, organism, and measurement	313,314
Observation	Observations and measurements	22,070,964
Prescription	Prescribed drug name, dosage, administration route	1,699,509
Therapy	Physiotherapy treatment	237,844
Ventilation	Oxygen delivery device, FiO_2 measurement	2,607,307

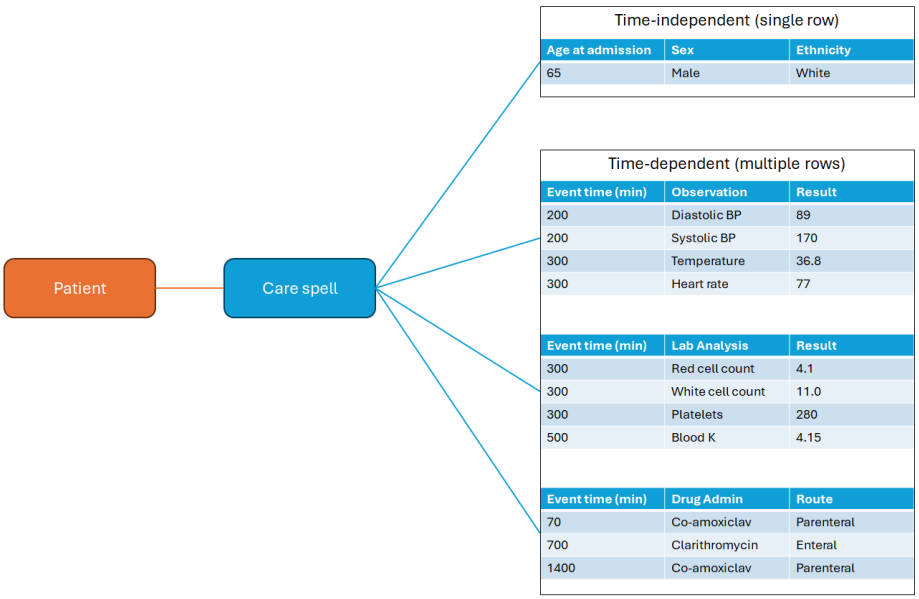

Fig. A. PIONEER data structure. Data is fictitious but representative

Cohort refinement

The inclusion/exclusion criteria provided as part of our application yielded data on 47,972 care spells for 36,885 patients. An initial inspection of the data revealed a significant number of patients with a diagnosis relating to Covid-19. Owing to complications and inconsistencies surrounding the recording of data for patients with Covid-19 (particularly during the early months of the global pandemic) coupled with the potential for overlap between Covid-19 and CAP, we took the decision to exclude these patients. Specifically, we excluded care spells with at least one *diagnosis name* containing variants of the string "COVID-19" and "Coronavirus".

Part of our inclusion criteria was a diagnosis of pneumonia using medical notes and ICD-10. These alone were insufficient for differentiating between patients with CAP and HAP. In order to refine our population to patients with CAP we set up rules based on when the diagnosis of pneumonia was made, the antibiotics received and time they were first administered. Specifically, we filtered for care spells that within the first 24 hours since admission received (a) a diagnosis of pneumonia and (b) an antibiotic taken from the CAP section of the Trust's Adult Guidelines for Antimicrobial Prescribing [10]. These antibiotics were: Amoxicillin, Co-amoxiclav, Clarithromycin, Doxycycline and Levofloxacin. If a patient received a pneumonia diagnosis or was administered one or more of these antibiotics more than 24 hours since admission, we classed them as HAP and excluded them from our patient population.

This refinement process, which we summarise in Figure B, resulted in a final cohort size of 10,707 care spells for 9,147 patients. 88% of patients had a single care spell, 9% had two care spells, 2% had three care spells and <1% had 4-9 care spells.

Variable selection

For our variable selection process, we conducted a table by table review of data items and determined their suitability for the proposed study based on four criteria: clinical relevance, usability, coverage and the ability to impute values if they were missing. For clinical relevance, we assessed items based on their ability to represent patient characteristics and health (states and rewards) as well as administered treatments (actions) in the context of managing CAP. For usability, items were assessed based on the feasibility of representing them as inputs/outputs in deep learning models (i.e. neural networks). In terms of coverage, we looked at the percentage of care spells for each item, as well as categories of items and the entire table itself. Finally, for missing values, we determined whether such data could reasonably be imputed with nominal or derived values. This would allow us to use rarely recorded but potentially informative data items by ensuring consistency in model inputs (i.e. ensuring all features have complete data). We summarise the outcomes of this selection process in Table F and provide further details in the following text.

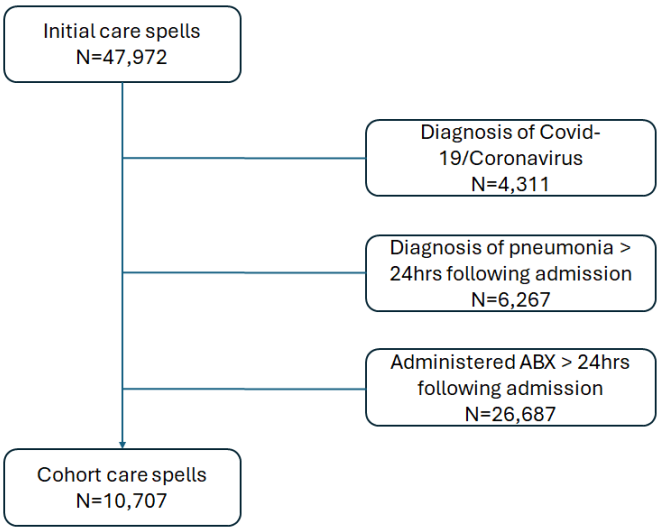

Fig. B. Cohort refinement. ABX = antibiotic.

Admissions By virtue of the Trust's procedures for recording and linking data, every patient and care spell had an Admissions record. Within the Admissions table, all care spells had associated demographics data (age, sex, ethnicity) which we selected to be part of the state and a binary indicator for 30-day mortality which we selected to be part of the reward. ICU admission time was used to filter out data recorded after a patient entered intensive care due to differences in monitoring and treatment compared to non-ICU patients. Other potential state variables such as social and economic indicators (deciles) and health indicators like smoking and alcohol consumption were excluded due to poor coverage/quality.

Assessments Nearly all care spells (99.1%) had data in the Assessments table. Out of all assessments, only the Abbreviated Mental Test (AMT-10) was deemed clinically relevant and usable. However, only 4.5% of care spells undertook this assessment and recorded a final value (score). The extremely low coverage of this assessment meant it was not selected as a variable for the state.

Diagnosis Since our inclusion criteria was based on a diagnosis of pneumonia using ICD-10 and SNOMED coding, all care spells had data in the Diagnosis table. The diagnosis itself was identified as a clinical relevant variable, however the level of granularity these were recorded at meant there were thousands of diagnosis names/codes, many of which were only used a handful of times. To make the data more usable we performed a filtering and grouping procedure. First, we filtered out names/code that had been used less than 20 times. For the remaining names/codes we grouped them into 22 categories to create a Diagnosis

Type variable, which went on to be used as part of the state. The full list of categories is provided in Table C. In addition, we back-dated these comorbidities to the start of a care spell since they were all long-standing conditions that would have been present upon admission.

Table C. Comorbidities used after applying filtering and grouping procedures. The table shows the number of care spells in which each condition was documented.

Comorbidity	**Care spell count**
Aneurysm	245
Anti-coagulant use	2161
Cancer	2664
Chronic kidney disease	2485
Chronic lung disease	4940
Coronary artery disease	3309
Cystic Fibrosis	4
Dementia	1863
Diabetes mellitus	3047
Epilepsy	467
Heart failure	2815
Hypercholesterolaemia	766
Hyperlipidaemia	207
Hypertension	5907
Liver disease	624
Mental health	4929
Neurological (e.g. stroke, Parkinson's)	1020
Obesity	1302
Osteoarthritis	1133
Other cardiac	4431
Peripheral vascular disease	394
Pulmonary hypertension	200

DnarTeal Less than half (48%) of care spells had data relating to do not resuscitate preferences. Of those that did this was not recorded in such a way that could be easily transformed into state variables. Furthermore, such preferences were not identified as clinical relevant. As such, data on resuscitation preferences was not included in the study.

DrugAdmin Since both our inclusion criteria and cohort refinement procedure required that patients were administered at least one CAP specific antibiotic, all care spells had data in the DrugAdmin table. The administration of a drug was identified as clinically relevant and the associated data the basis for the action space. However, due to factors such as co-morbidities, the list of drugs ran into the thousands and grew further still when route and dosage were taken

into consideration. As such, we based our action space on the Trust's Guidelines for Antimicrobial Prescribing, using the list of recommended drugs as one component and the administration route as another. We concatenated these components the create the variable *drug-route* but did not incorporate the dosage into the action space as this was fixed by the guidelines.

Imaging Around 80% of care spells had imaging data. Out of all image types, those relating to chest X-rays were identified as clinically relevant and these were recorded for 78.0% of care spells. The image itself was not available, but whether or not a patient had a chest X-ray was deemed relevant as in practice this is often used to confirm a diagnosis and can thus impact treatment decisions. As such, a time-dependent binary variable was created to indicate whether a care spell had received results from a chest X-ray and this was used as part of the state. It was assumed that if no chest X-ray data was available the patient did not have one.

LabAnalysis Practically all care spells (99.9%) had at least one type of laboratory test, however some types were much more common than others, with certain tests only performed for a small minority of patients. This reflects the individualised nature of managing CAP in practice, with clinicians requesting different types of tests to assist them in their decision-making. Out of the 52 types of tests conducted, 24 were selected to become part of the state based on having coverage of 80% or higher, and a further 3 selected due to their clinical relevance and the ability to derive a representative value when missing (see Variable Derivation). The remaining tests and their respective coverages are presented in Table D.

Microbiology Less than one fifth (18%) of care spells had data relating to microbiology. This low coverage and the fact the data could not be easily used or imputed led to its exclusion from the study.

Observation Practically all care spells (99.9%) had at least one type of observation. Out of the 12 types of observation, only 2 were excluded due to low coverage and lack of ability to impute missing values. The remaining observations were selected to become part of the state based on having very high coverage and the ability to derive representative values for the missing cases (see Variable Derivation).

Prescription The Prescription table contained prescription data for drugs in the DrugAdmin table. This was supplementary data used by pharmacy departments to monitor drug usage, as opposed to data on the administration of the drugs themselves. As such, it did not provide any further relevant information over that already held in the DrugAdmin table and was therefore discarded.

Table D. Laboratory tests excluded from this study due to having coverage below 80%.

Variable	**Coverage (%)**
AKI Stage	36.2
Anion Gap	60.2
APTT (sec)	55.0
APTT Ratio	34.6
Asp.Aminotransferase	1.4
Bicarbonate	0.8
Bilirubin	77.6
Blood Cl	79.9
Blood Glucose	79.8
CO Hb	78.7
ESR	5.6
Estimated GFR	77.6
Fluid Creatinine	0.0
Fluid LDH	0.6
Gamma G.T.	2.5
INR	70.2
Lactate Dehyd'genase	7.8
Met Hb	78.5
Nucleated RBCs	78.9
pCO_2 (temp corrected)	68.8
pH (temp corrected)	68.8
Plasma Lactate	0.0
pO_2 (temp corrected)	68.8
PoC SO_2	78.1
Urine Sodium	0.3

Therapy Around 72% of care spells had physiotherapy data. The format of the data was such that it couldn't easily be represented as a state variable (beyond a binary indicator that a patient had/hadn't received physiotherapy which wasn't considered informative) nor incorporated into an action space with drug administration in a straightforward manner. As a consequence, no physiotherapy data was included in the study

Ventilation Just under two-thirds (65%) of care spells had data on mechanical ventilation. This mostly related to the oxygen delivery device which was available for practically all care spells that received ventilation. Although not necessarily informative by itself, it was considered useful for deriving FiO_2 values in conjunction with variables from other tables (see Variable Derivation), which would compensate for the extremely low coverage of FiO_2 within the ventilation table itself (0.2%).

Variable derivation

As part of our variable selection process we noted several instances of missing data items that could be reasonable derived using other variables. We summarise these derivations as follows

NEWS2 The National Early Warning Score (NEWS) is a metric used to detect clinical deterioration in adults [8]. A score is given to patients based on measurements of respiratory rate, oxygen saturation, blood pressure, pulse, consciousness, temperature and whether a patient is receiving oxygen therapy. Higher scores indicate more severe deterioration. NEWS2 was a variable in the Observation table that had high but incomplete coverage (86%). However, there was near complete coverage of variables that constituted NEWS2 and hence we were able to derive missing values using the NEWS2 scoring system. In Table E we provide a mapping of PIONEER variables to those in the scoring system.

Table E. Mapping the National Early Warning Score (NEWS2) components to the corresponding PIONEER observation variables.

NEWS2	**PIONEER**
Respiration rate	Respiratory rate
SpO_2	O_2 saturation
Air or oxygen	Oxygen flow rate (0 = air, >0 = oxygen)
Systolic blood pressure	Systolic blood pressure
Pulse	Heart rate
Consciousness	AVPU scale (4 = alert; <4 = CVPU)
Temperature	Temperature

Table F: Summary of variable selection. IC: Insufficient coverage; NR: Not clinically relevant; NU: Not easily usable; NI: Not easily imputable. MAP: Mean arterial pressure.

Table	Variable	Coverage (%)	Used	Reason not used
Admissions	Age	100	Y	
	Sex	100	Y	
	Ethnicity	100	Y	
	30-day mortality	100	Y	
Assessments	All variables		N	IC, NR, NU, NI
Diagnosis	Co-morbidity	100	Y	
DnarTeal	All variables		N	IC, NR, NU, NI
DrugAdmin	Drug	100	Y	
	Route	100	Y	
Imaging	Chest X-ray taken	78.0	Y	
Lab Analysis	Alanine Transferase	94.0	Y	
	Albumin	95.9	Y	
	Alkaline Phosphatase	95.5	Y	
	Base Excess	81.3	Y	
	Basophils	96.9	Y	
	Blood K	98.4	Y	
	Blood Na	98.8	Y	
	C-reactive protein	94.3	Y	
	Calcium	87.9	Y	
	Eosinophils	96.9	Y	
	Haematocrit	97.2	Y	
	Haemoglobin	98.6	Y	
	Lymphocytes	97.0	Y	
	Mean cell Hb	97.2	Y	
	Mean cell volume	97.2	Y	
	Monocytes	97.0	Y	
	Neutrophils	97.0	Y	
	pCO_2	81.4	Y	
	Platelets	97.1	Y	
	PO_2	81.4	Y	
	Red cell count	97.2	Y	
	Total Protein	89.5	Y	
	Urea	97.2	Y	
	White cell count	97.2	Y	
	25 other types	<80	N	IC, NR, NU, NI
Microbiology	All variables		N	IC, NU, NI
Observation	AVPU scale	100	Y	
	Diastolic BP	100	Y	
	Heart rate	100	Y	

Continued on next page

Table F (continued)

Table	Variable	Coverage (%)	Used	Reason not used
	NEWS2 score	86	Y	
	O_2 flow rate	97.4	Y	
	O_2 saturation	100	Y	
	Respiratory rate	100	Y	
	Systolic BP	100	Y	
	Temperature	99.9	Y	
	MAP	2.5	N	IC
	Urine	42.1	N	IC
Prescription	All variables		N	NR
Therapy	All variables		N	NR, NU, NI
Ventilation	Device	99.9	Y	
	FiO_2	0.2	N	IC

D-Dimer, Troponin-I, Troponin-T D-Dimer is a protein related to blood clotting and Troponin-I/Troponin-T are proteins relating to heart muscle contraction [9,7]. Only patients with suspected blood clotting and/or a heart condition will have these protein levels measured and thus only a subset of patients in the Lab Analysis table had associated values (21%, 20% and 6% coverage, respectively). Unlike NEWS2 we could not derive values using other variables, nor was it appropriate to impute based on available values due to low coverage. Therefore, to include such test results in our state we converted them to categorical variables, mapping existing values to "normal" or "high" based on clinical criteria and all others as "not recorded".

To derive missing D-dimer and Troponin-I we used the Trust's lab ranges [11] and for Troponin-T we used the Wythenshawe Hospital Laboratory Medicine Handbook [4]. Each derivation is summarised in Table G.

Table G. Categorization of D-dimer and Troponin (I/T) test results based on clinical thresholds.

Protein	Normal	High
D-dimer (all)	<250 ng/ml	≥250 ng/ml
Troponin-I (men)	<34 ng/L	≥34 ng/L
Troponin-I (women)	<16 ng/L	≥16 ng/L
Troponin-T (all)	<14 ng/L	≥14 ng/L

FiO_2 Fraction of Inspired Oxygen (FiO_2) is the concentration of oxygen in a gas mixture and is used as an estimate for the amount of oxygen a person inhales [3]. It is a particularly important measure in conditions that affect breathing, and is thus clinically relevant in the management of CAP. FiO_2 was a variable in the Ventilation table, but the coverage was extremely low. However, we were able to

derive an FiO_2 estimate using the oxygen flow rate from the Observation table and the oxygen delivery device from the Ventilation table. In the event there was an oxygen flow rate and no device, we used the median across flow rates. In the event there was a device and no oxygen flow rate, we used the median across devices. For patients not receiving oxygen therapy we set their FiO_2 to 0.21, which is the value for air (i.e normal breathing).

To derive missing FiO_2 we used the ventilation device from the Ventilation table and oxygen flow rate from the Observation table. For O_2FACE, O_2HFF, O_2NASAL, O_2NONINV and O_2TRACH, the FiO_2 value is based on the Shapiro formula [2]. O_2VENTURI and O_2RESPIFLO are based on systems at the local hospital and manufacturer stated FiO_2. We summarise the mapping for each ventilation device in Table H.

Table H: Mapping from oxygen device and flow rate to estimated FiO_2. For O_2FACE, O_2HFF, O_2NASAL, O_2NONINV and O_2TRACH, the FiO_2 value is based on the Shapiro formula [2]. O_2VENTURI and O_2RESPIFLO are based on systems at the local hospital and manufacturer stated FiO_2.

Device	**Oxygen flow rate**	**FiO_2**
O_2FACE	0.5	0.22
	1	0.24
	1.5	0.26
	2	0.28
	2.5	0.30
	3	0.32
	3.5	0.34
	4	0.36
	4.5	0.38
	5	0.40
	5.5	0.42
	6	0.44
	6.5	0.46
	7	0.48
	7.5	0.50
	8	0.52
	8.5	0.54
	9	0.56
	10	0.60
	10.5	0.62
	11	0.64
	12	0.68
	13	0.72
	14	0.76
	15	0.80
O_2HFF	2	0.28
	3	0.32
	4	0.36
	4.1	0.364
	5	0.40
	6	0.44
	7	0.48
	7.5	0.50
	8	0.52
	9	0.56
	10	0.60
	10.5	0.62
	11	0.64

Continued on next page

Table H (continued)

Device	Oxygen flow rate	FiO_2
	12	0.68
	12.5	0.70
	13	0.72
	14	0.76
	15	0.80
O_2NASAL	0.5	0.22
	1	0.24
	1.5	0.26
	2	0.28
	2.5	0.30
	3	0.32
	3.5	0.34
	4	0.36
	4.5	0.38
	5	0.40
	5.5	0.42
	6	0.44
	7	0.48
	8	0.52
O_2NONINV	0.5	0.22
	1	0.24
	1.5	0.26
	2	0.28
	2.5	0.30
	3	0.32
	3.5	0.34
	4	0.36
	4.5	0.38
	5	0.40
	5.5	0.42
	6	0.44
	7	0.48
	8	0.52
	8.5	0.54
	9	0.56
	9.8	0.592
	10	0.60
	11	0.64
	12	0.68
	13	0.72
	14	0.76
	15	0.80

Continued on next page

Table H (continued)

Device	Oxygen flow rate	FiO_2
O_2TRACH	0.5	0.22
	1	0.24
	2	0.28
	3	0.32
	3.5	0.34
	4	0.36
	4.5	0.38
	5	0.40
	6	0.44
	7	0.48
	8	0.52
	8.5	0.54
	9	0.56
	10	0.60
	12	0.68
	13	0.72
	15	0.80
O_2VENTURI	2	0.24
	4	0.28
	6	0.31
	8	0.35
	10	0.40
	15	0.60
O_2RESPIFLO	5	0.28
	6	0.35
	8	0.60
	9	0.40
	11	0.60
	15	0.80
O_2NOREMASK	10	0.60
	11	0.60
	12	0.80
	13	0.80
	14	0.98
	15	0.98

State-action pair construction

After determining the variables for our study, we turned our attention to the construction of state-action pairs. In general, this is one of the more challenging aspect of using RL in real-world settings. Unlike simulated environments, where the agent receives a completely updated state and a reward signal at regular intervals each time an action is taken, in real-world settings this is seldom the

case and the agent will instead receive information on different parts of the state at irregular intervals as time progresses.

Within the PIONEER data set individual items (observations, tests, measurements, treatments etc.) were recorded at varying frequencies both within and between tables. Furthermore, due to the individualised nature of patient care, there was considerable variation between care spells, with some patients being monitored more closely and/or treated more frequently than others in response to their changing health status. We provide examples of this in Figure C in which we plot histograms for time intervals between measurements for four state variables, capping the interval at 36 hours for illustrative purposes.

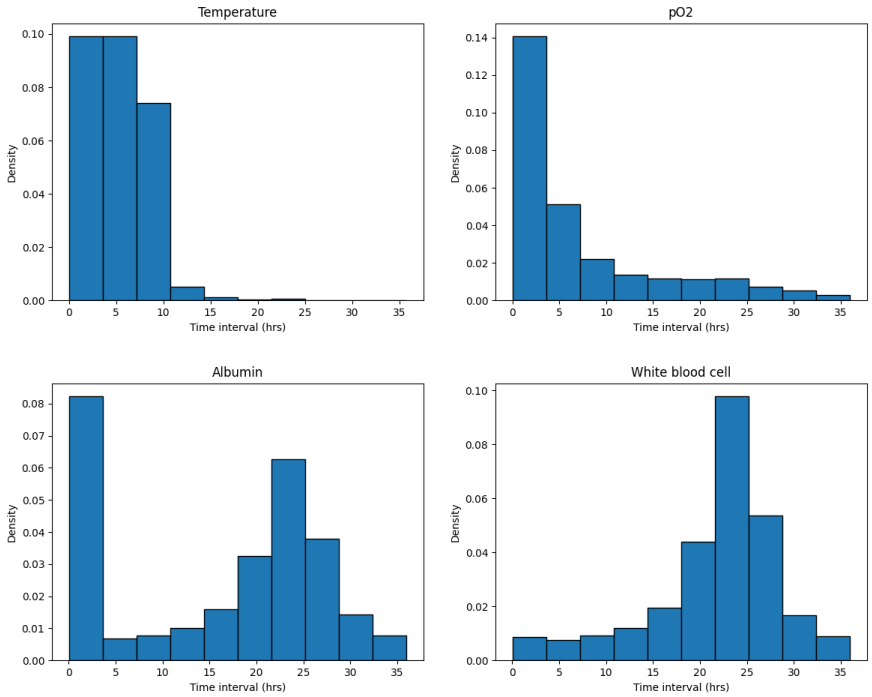

Fig. C. Time intervals between successive measurements of four state variables. Variables were measured at irregular intervals reflecting how individual patients were monitored in the clinical setting.

The fact that the PIONEER data was recorded in this way meant that when we pivoted and joined together data items from individual tables the resulting table contained a lot of apparent "missing" values, as illustrated with an example in Figure D. In order to accommodate data of this type, we adopted a sample-and-hold strategy in which the latest value was brought forward at each time step until a new value was recorded [5]. Such a strategy has been regularly employed

in related literature as it reasonably approximates the hospital setting, whereby clinicians use the latest information available to inform their treatment decisions. In addition, it provides a practical form of interpolation in settings where the nature of missingness is non-random or unknown.

Source data (fictitious example)

Table	Care spell id	Event time	Item / Value
Observation	101	500	Temperature: 36.8
Observation	101	500	Heart rate: 65
Observation	101	700	Temperature: 36.7
Lab Analysis	101	600	Albumin: 22
Lab Analysis	101	600	pO_2: 3.98
Lab Analysis	101	700	Albumin: 21
DrugAdmin	101	900	Drug-Route: Amoxicillin-Enteral

↓

Pivoted and joined data

Care spell id	Event time	Temperature	Heart rate	Albumin	pO_2
101	500	36.8	65	–	–
101	600	–	–	22	3.98
101	700	36.7	–	21	–
101	900	–	–	–	–

(Drug-Route: Amoxicillin-Enteral at event time = 900)

Fig. D. Pivot and join. Top table is source data; bottom table is pivoted and joined. For each table, rows are grouped by *care spell id* and *event time*, and each column header is an item. Next, a full outer join on these keys yields a single table with potential missing values due to irregular sampling. Data is fictitious but representative.

Before we imputed missing values with previous values however, we first had to consider what constituted a time step for our study and how we aggregated data so that it approximated a MDP. In general, such a time step should reflect the frequency with which actions are taken. For our particular study this translates to the frequency with which antibiotics are administered. However, in practice there is considerable variation when it comes to antibiotic administration due to several factors. These include, for example, the fact that certain drugs may only be administered once a day while others are administered multiple times, patients can move from one course of antibiotics to another, the route of administration can change, a patient's health can suddenly deteriorate and require immediate intervention, etc. Furthermore, even though the Trust has established guidelines for antibiotics administration, this only forms one part of the decision-making process and clinicians will use other sources of information and their own judgement in deciding which drug to eventually administer. The end result is a large amount of variability in treatment across the PIONEER

data set as evidenced in Figure E where we plot histograms for time intervals between drug administration for four antibiotics and administrations routes.

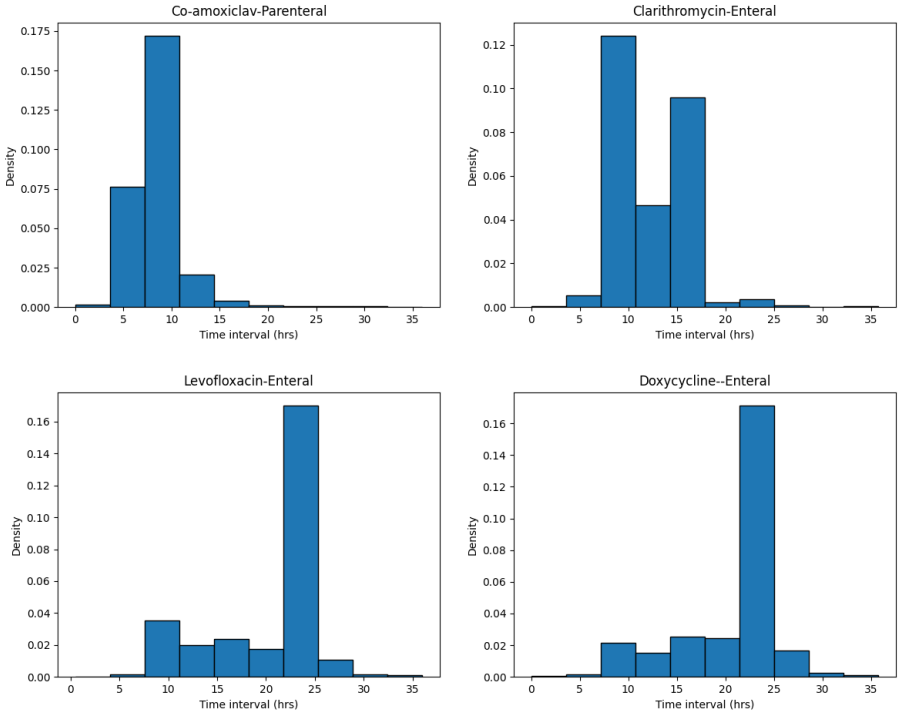

Fig. E. Time intervals between successive administration of four antibiotics. Antibiotics were administered at different intervals reflecting how drugs were prescribed and how individual patients were treated in the clinical setting.

In light of this variability in antibiotic administration, we settled on two approaches for defining our time step and aggregating data items to create our state-action pairs. The first was based on a fixed time step, with both state and action variables aggregated into windows of fixed length throughout a care spell trajectory. The second was based on a variable time step according to when actions were taken, with state variables being aggregated in line with action variables. We detail each of these in the following sub-sections, commenting on their respective strengths and limitations. In each case the starting table resembled that of the bottom of Figure D.

Approach 1: Fixed time step For this approach we first created windows of fixed length and then aggregated data within each window according to the

data type. For the state, we used the median for numerical variables and mode for categorical variables, while for the actions we concatenated values for *drug-route*, which extended our action set to include drug combinations. To avoid a combinatorial explosion in the number of drug combinations, we limited the number of drugs to two, using the latest two drugs in cases where there were three or more drugs administered in a window. Furthermore, for cases where a drug switched administration route during a window, we used the latest route since in practice the same drug would not have been administered via two different routes at the same time.

We noted in some instances that patients did not receive their first antibiotic until after the first time step. For these cases we set the *drug-route* variable to "no treatment" for the preceding windows. We also noted instances where patients did not receive antibiotics for extended periods of time. In these cases we set the *drug-route* variable to "no treatment" if no antibiotics were administered for 36 hours, reflecting the upper limit of time intervals for drug administration as per Figure E as well as allowing for a buffer of 12 hours for drugs usually administered once-a-day.

At this stage we then applied our sample-and-hold strategy and populated missing values for the current window with the latest available from previous windows. In cases where no previous values were available we first imputed nominal values where this was appropriate and for the remaining missing values we used the median across all care spells. The full list of median values is provided in Table I and the rationale for nominal values as follows.

- FiO_2: If there is no evidence a patient is receiving oxygen therapy, we set FiO_2=0.21 which is the value for regular breathing
- AVPU: If there is no assessment of a patient alertness/consciousness, we set AVPU=4 which is the value for full alertness/consciousness
- NEWS2: If there is no assessment of deterioration, we set NEWS2=0 which is the value for low risk
- Chest X-ray: If there is no evidence a chest X-ray has been taken, we set the Chest X-ray=0 which is the value for not taken.

Finally, we added our time independent state variables, namely back-dated comorbidities and demographics. The resulting table was a set of state-action pairs that was feature complete for each fixed-length window. For the length of the window itself, we used 8 hours, 12 hours and 24 hours to simulate the administration of antibiotics three times a day, twice a day and once a day, respectively.

The primary strength of this approach is that it maintains the fixed time step assumption of MDPs. It also allows for drug combinations, which may better capture the underlying treatment regime. The main limitation is that we are no longer accurately reflecting the treating a patient actually received. We are in effect stating that a patient received a particular treatment at fixed intervals when in reality the patient could have received the treatment anytime in the window, or not at all if brining forward previous treatments values. This

Table I: Median values used to impute missing measurements in the PIONEER dataset.

Category	Feature	Median
Observations	Diastolic BP	70
	Heart rate	84
	O_2 sats %	95
	Respiratory rate	18
	Systolic BP	122
	Temperature	36.4
Lab Analysis	Alanine transferase	18
	Albumin	29
	Alkaline phosphatase	100
	Base excess	0.1
	Basophils	0.04
	Blood K	4.2
	Blood Na	138
	Calcium	2.2
	C-reactive protein	74
	Eosinophils	0.1
	Haematocrit	0.34
	Haemoglobin	113
	Lymphocytes	1.1
	Mean cell Hb	29.6
	Mean cell volume	89.8
	Monocytes	0.73
	Neutrophils	7.5
	pCO_2	5.7
	Platelets	267
	PO_2	7.1
	Red cell count	3.84
	Total protein	62
	Urea	7
	White cell count	9.9

potentially creates a disconnect between what really happened in the clinical setting and what the agent learns from.

Approach 2: Variable time step For this approach we first created windows based on when antibiotics were administered and then used the latest value in each window to populate the components of the state. If a patient did not receive antibiotics for 36 hours we created additional windows for "no treatment", allowing us to account for extended periods of no drug administration (similar to the fixed time step approach). Since our cohort refinement procedure necessitated the administration of at least one antibiotic within the first 24 hours of admission,

we did not need to create similar "no treatment" windows at that start of a care spell trajectory.

At this stage we then followed the same procedure as the fixed time step approach, populating missing values with those in previous windows where available or with nominal or median values where not, before adding comorbidities and demographics variables to complete the state. The resulting table was a set of state-action pairs that was feature complete for each variable-length window.

The primary strength of this approach is that it accurately reflects the treatment a patient received, with the agent learning from actions that were actually taken in the clinical setting. The main limitation is that the fixed time step assumption of MDPs is violated. In addition, the action space no longer accommodates drug combinations and thus may fail to capture the patient's underlying treatment regime.

Illustrative example To help show how the resulting state-action pairs differ between the two approaches we provide an example based on the same set of starting data. In Figure F the source data is presented in the left-hand table with the resulting state-action pairs using the fixed time step and variable time step approaches presented in the top-right and bottom-right tables, respectively

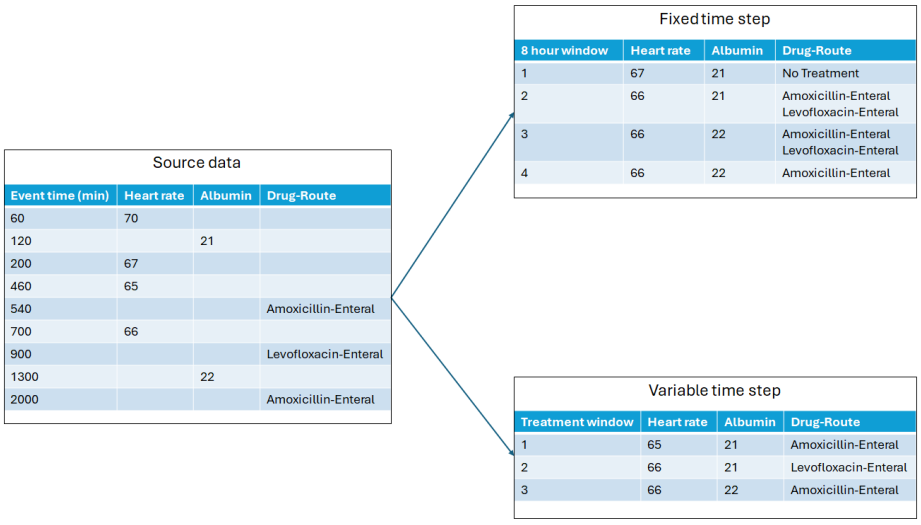

Fig. F. State-action pair construction. The left-hand table is the source starting data. The top-right table is the state-action pairs following the fixed time step approach with an 8hr window. The bottom-right table is the state-action pairs following the variable time step approach. Data is fictitious but representative

Reward function and terminal state definition

We elected to use a sparse reward based on 30-day mortality, a commonly used outcome measure for CAP [12] . If a care spell's 30-day mortality since admission indicator was recorded as 0 (survived) we set the reward to $+1$ while if it was recorded as 1 (did not survive) we set the reward to -1. Using a reward of this form effectively tasks the agent with reducing 30-day mortality. The terminal state was set to be the final state prior to 30-days since admission. This ensured all data available to the agent was recorded within the 30-day period following admission.

Processed data sets

The result of our pre-processing procedures was four sets of state, action and reward transitions, one for each of the fixed length windows (8hr, 12hr, 24hr) and one for the variable length window. Our resulting state space consisted of 74 features based on observations, laboratory tests, radiography, co-morbidities and demographics, our action space constituted antibiotic drugs with administration routes and our rewards were based on a sparse signal defined by 30-day mortality.

The composition of the data sets is summarised in Table J. A list of final actions is provided in Table K. For the fixed time step approach there were an additional 32 actions resulting from drug-route combinations, a full list of which is provided in Table L. A summary of cohort characteristics is presented in Table N. For completeness, in Table M we provide details of missingness for each set of transitions.

Table J. Overview of the processed datasets using fixed 8hr, 12hr, 24hr windows and variable-length windows based on antibiotic administration. "Trajectory length" refers to the number of transitions in a care spell.

Item	**Value**
Total number transitions (8hr)	217,988
Total number transitions (12hr)	151,922
Total number transitions (24hr)	78,731
Total number transitions (Variable)	180,949
Median trajectory length (8hr)	17
Median trajectory length (12hr)	12
Median trajectory length (24hr)	6
Median trajectory length (Variable)	15

Table K. List of actions under the variable time step approach (combining a drug and its administration route).

Drug	**Route**	**Action**
Amoxicillin	Enteral	Amoxicillin-Enteral
Amoxicillin	Parenteral	Amoxicillin-Parenteral
Clarithromycin	Enteral	Clarithromycin-Enteral
Clarithromycin	Parenteral	Clarithromycin-Parenteral
Co-amoxiclav	Enteral	Co-amoxiclav-Enteral
Co-amoxiclav	Parenteral	Co-amoxiclav-Parenteral
Doxycycline	Enteral	Doxycycline-Enteral
Levofloxacin	Enteral	Levofloxacin-Enteral
Levofloxacin	Parenteral	Levofloxacin-Parenteral
None	–	No Treatment

Table L. The set of possible actions under the fixed time step approach, accommodating up to two drug-route pairs.

Action
Amoxicillin-Enteral
Amoxicillin-Enteral, Clarithromycin-Enteral
Amoxicillin-Enteral, Clarithromycin-Parenteral
Amoxicillin-Enteral, Co-amoxiclav-Enteral
Amoxicillin-Enteral, Co-amoxiclav-Parenteral
Amoxicillin-Enteral, Doxycycline-Enteral
Amoxicillin-Enteral, Levofloxacin-Enteral
Amoxicillin-Enteral, Levofloxacin-Parenteral
Amoxicillin-Parenteral
Amoxicillin-Parenteral, Clarithromycin-Enteral
Amoxicillin-Parenteral, Clarithromycin-Parenteral
Amoxicillin-Parenteral, Co-amoxiclav-Enteral
Amoxicillin-Parenteral, Co-amoxiclav-Parenteral
Amoxicillin-Parenteral, Doxycycline-Enteral
Amoxicillin-Parenteral, Levofloxacin-Enteral
Amoxicillin-Parenteral, Levofloxacin-Parenteral
Clarithromycin-Enteral
Clarithromycin-Enteral, Co-amoxiclav-Enteral
Clarithromycin-Enteral, Co-amoxiclav-Parenteral
Clarithromycin-Enteral, Doxycycline-Enteral
Clarithromycin-Enteral, Levofloxacin-Enteral
Clarithromycin-Enteral, Levofloxacin-Parenteral
Clarithromycin-Parenteral
Clarithromycin-Parenteral, Co-amoxiclav-Enteral
Clarithromycin-Parenteral, Co-amoxiclav-Parenteral
Clarithromycin-Parenteral, Levofloxacin-Parenteral
Co-amoxiclav-Enteral
Co-amoxiclav-Enteral, Doxycycline-Enteral
Co-amoxiclav-Enteral, Levofloxacin-Enteral
Co-amoxiclav-Enteral, Levofloxacin-Parenteral
Co-amoxiclav-Parenteral
Co-amoxiclav-Parenteral, Doxycycline-Enteral
Co-amoxiclav-Parenteral, Levofloxacin-Enteral
Co-amoxiclav-Parenteral, Levofloxacin-Parenteral
Doxycycline-Enteral
Doxycycline-Enteral, Levofloxacin-Enteral
Doxycycline-Enteral, Levofloxacin-Parenteral
Levofloxacin-Enteral
Levofloxacin-Parenteral
NoTreatment

Table M. Percent of transitions missing a given variable in the four processed datasets (fixed 8hr, fixed 12hr, fixed 24hr, variable). Values reflect cases where median imputation was necessary after sample-and-hold or nominal substitution could not be applied.

Variable	**8hr**	**12hr**	**24hr**	**Variable**
Alanine Transferase	7.4	7.1	6.8	10.7
Albumin	5.1	4.9	4.7	8.3
Alkaline Phosphatase	5.4	5.2	4.9	8.7
Base Excess	26.2	26.2	25.7	23.6
Basophils	4.1	4.0	3.8	6.9
Blood K	3.9	3.7	3.5	5.1
Blood Na	3.3	3.2	3.0	4.4
C-reactive protein	5.7	5.5	5.1	9.0
Calcium	13.7	13.3	12.6	18.2
Diastolic BP	0.4	0.2	0.1	1.5
Eosinophils	4.1	3.9	3.7	6.9
Haematocrit	3.8	3.6	3.5	6.5
Haemoglobin (g/L)	3.4	3.3	3.1	4.6
Heart rate	0.4	0.2	0.1	1.5
Lymphocytes	3.9	3.8	3.6	6.7
Mean cell Hb	3.8	3.6	3.5	6.5
Mean cell volume	3.8	3.6	3.5	6.5
Monocytes	3.9	3.8	3.6	6.7
Neutrophils	3.9	3.8	3.5	6.7
O_2 sats (%)	0.4	0.2	0.1	1.6
Platelets	3.9	3.7	3.5	6.6
Red cell count	3.8	3.6	3.5	6.5
Respiratory rate	0.4	0.2	0.1	1.5
Systolic BP	0.4	0.2	0.1	1.5
Temperature	0.4	0.3	0.1	1.6
Total Protein	11.9	11.5	10.7	16.3
Urea	3.6	3.5	3.3	6.8
White cell count	3.8	3.6	3.4	6.5
pCO_2	26.1	26.1	25.7	23.5
pO_2	26.1	26.1	25.7	23.6

Table N: Characteristics of the patient cohort used in this study. Ethnicity types: White, Black or Black British, Asian or Asian British, Chinese, Mixed, Any other ethnic group, Not known. Comorbidity types: Aneurysm, Anti-coagulant use, Cystic fibrosis, Cancer, Chronic lung disease, Coronary artery disease, Dementia, Diabetes mellitus, Epilepsy, Heart failure, Hypercholesterolaemia, Hyperlipidaemia, Hypertension, Liver disease, Mental health, Neurological, Obesity, Osteoarthritis, Other-cardiac, Peripheral vascular disease, Pulmonary hypertension, Chronic kidney disease. FiO_2 (OT) = derived values for patients on oxygen therapy. D-dimer, Troponin-I, Troponin-T: N = Normal, H = High, NR = Not Recorded.

Category	**Feature**	**Mean (SD)**
Demographics	Age (years)	73.9 (15.7)
	Male (N, %)	5390 (50.3%)
	Non-survivors (N, %)	1727 (16.1%)
	Ethnicity	7 types
	Comorbidity	22 types
Observations	AVPU scale	3.96 (0.25)
	Diastolic BP	70.7 (12.9)
	Heart rate	85.2 (17.4)
	NEWS2	3.17 (2.6)
	O_2 sats (%)	94.8 (3.3)
	Respiratory rate	18.6 (3.3)
	Systolic BP	125.1 (22.6)
	Temperature	36.4 (0.7)
Lab Analysis	Alanine transferase	35.7 (97.3)
	Albumin	29.1 (6.8)
	Alkaline phosphatase	134 (144)
	Base excess	0.15 (5.1)
	Basophils	0.05 (0.06)
	Blood K	4.21 (0.74)
	Blood Na	138 (6.1)
	Calcium	2.2 (0.2)
	C-reactive protein	101 (91.7)
	Eosinophils	0.17 (0.38)
	Haematocrit	0.34 (0.07)
	Haemoglobin	113 (22.5)
	Lymphocytes	1.72 (8.9)
	Mean cell Hb	29.4 (3.0)
	Mean cell volume	89.6 (7.7)
	Monocytes	0.87 (1.44)
	Neutrophils	9.01 (7.9)
	pCO_2	6.01 (1.82)
	Platelets	290 (143)

Continued on next page

Table N (continued)

Category	Feature	Mean (SD)
	pO_2	8.04 (5.46)
	Red cell count	3.83 (0.75)
	Total protein	62.0 (8.8)
	Urea	9.12 (6.94)
	White cell count	11.4 (11.2)
Imaging	Chest X-ray (N, %)	8344 (77.9%)
Other	FiO_2 (OT)	0.28 (0.23)
	D-dimer	N, H, NR
	Troponin-I	N, H, NR
	Troponin-T	N, H, NR

Additional cohort details

Table O provides values used for ethnicity, comorbidities and antibiotics.

Table O. Additional cohort details

Item	Value
Ethnicity	White, Black or Black British, Asian or Asian British, Chinese, Mixed, Any other ethnic group, Not known.
Comorbidities	Aneurysm, Anti-coagulant use, Cystic fibrosis, Cancer, Chronic lung disease, Coronary artery disease, Dementia, Diabetes mellitus, Epilepsy, Heart failure, Hypercholesterolaemia, Hyperlipidaemia, Hypertension, Liver disease, Mental health, Neurological, Obesity, Osteoarthritis, Other-cardiac, Peripheral vascular disease, Pulmonary hypertension, Chronic kidney disease.
Antibiotic+Route	Amoxicillin-Enteral, Amoxicillin-Parenteral, Clarithromycin-Enteral, Clarithromycin-Parenteral, Co-amoxiclav-Enteral, Co-amoxiclav-Parenteral, Doxycycline-Enteral, Levofloxacin-Enteral, Levofloxacin-Parenteral.

References

1. Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., Wiswedel, B.: KNIME: The Konstanz Information Miner. In: Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007). Springer (2007)
2. Duprez, F., Mashayekhi, S., Cuvelier, G., Legrand, A., Reychler, G.: A new formula for predicting the fraction of delivered oxygen during low-flow oxygen therapy. Respiratory Care **63**(12), 1528–1534 (2018)

3. Fuentes, S., Chowdhury, Y.S.: Fraction of inspired oxygen. In: StatPearls [Internet]. StatPearls Publishing (2022)
4. Hospital, W.: Laboratory medicine handbook. https://mft.nhs.uk/app/uploads/2023/09/Pathology-Handbook.pdf, 2023
5. Komorowski, M., Celi, L.A., Badawi, O., Gordon, A.C., Faisal, A.A.: The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nature medicine **24**(11), 1716–1720 (2018)
6. Lim, W.S., Van der Eerden, M.M., Laing, R., Boersma, W.G., Karalus, N., Town, G.I., Lewis, S., Macfarlane, J.: Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax **58**(5), 377–382 (2003)
7. Maynard, S., Menown, I., Adgey, A.: Troponin t or troponin i as cardiac markers in ischaemic heart disease (2000)
8. of Physicians, R.C.: National early warning score (NEWS) 2: Standardising the assessment of acute-illness severity in the nhs. https://www.rcp.ac.uk/media/ctulqqbn/news2-executive-summary_0.pdf, 2017
9. Tripodi, A.: D-dimer testing in laboratory practice. Clinical chemistry **57**(9), 1256–1262 (2011)
10. Trust, U.H.B.N.F.: Adult guidelines for antimicrobial prescribing. https://www.uhb.nhs.uk/Downloads/pdf/controlled-documents/AntimicrobialPrescribingGuidelines.pdf, 2023-09-01
11. Trust, U.H.B.N.F.: Clinical chemistry tests database. https://qehbpathology.uk/test-database, 2023-04-15
12. Wiemken, T.L., Furmanek, S.P., Mattingly, W.A., Guinn, B.E., Cavallazzi, R., Fernandez-Botran, R., Wolf, L.A., English, C.L., Ramirez, J.A.: Predicting 30-day mortality in hospitalized patients with community-acquired pneumonia using statistical and machine learning approaches. The University of Louisville Journal of Respiratory Infections **1**(3), 10 (2017)