

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Scuola di Economia e Statistica
Corso di laurea Magistrale in
BIOSTATISTICA



**PREDIZIONE DELLA PRESSIONE PARZIALE ARTERIOSA DI
OSSIGENO IN PAZIENTI AMMESSI IN TERAPIA INTENSIVA:
UN MODELLO AD EFFETTI MISTI CON PARTICOLARE
ATTENZIONE AI SOGGETTI INTUBATI**

Relatore: Prof. Vincenzo Bagnardi

Correlatore: Dott. Stefano Finazzi

Tesi di Laurea di:
Enrico Villa
Matricola N° 839325

Anno Accademico 2022/2023

IN COLLABORAZIONE CON
IRCCS MARIO NEGRI,
RANICA (BG)

Ringraziamenti

In primo luogo vorrei ringraziare l'Istituto di Ricerche Farmacologiche Mario Negri di Ranica che mi concesso la possibilità di svolgere il progetto di tesi concludendo il mio percorso universitario magistrale. In particolare ci tengo a ringraziare il Dott. Stefano Finazzi che è stato un aiuto fondamentale per migliorarmi, fornendo un supporto attivo in ogni fase di questo progetto. Vorrei ringraziare il Dott. Maximiliano Mollura e il Dott. Demetrio Magatti per il supporto tecnico che mi hanno sempre garantito. Un ringraziamento sentito anche al Dott. Orlando Sagliocco, fautore dell'idea clinica che ha dato vita a questa tesi.

Infine ringrazio il Professor Vincenzo Bagnardi, che mi ha aiutato nella revisione e nella rifinitura della stessa e che mi ha accompagnato in questo percorso magistrale.

Indice

1. INTRODUZIONE	3
1.1 Ventilazione meccanica	4
1.2 ARDS	6
1.3 Target ossigenativi	8
1.4 Indicatori alternativi a PF	11
2. IPOTESI DI RICERCA	13
2.1 Quesito statistico	13
2.2 Database MargheritaTre	15
2.3 Selezione pazienti e centri ospedalieri	17
3. GESTIONE DATI	19
3.1 Variabili considerate	20
3.1.1 Selezione variabili indipendenti	20
3.1.2 Importanza del controllo di PvO_2 in ogni centro ospedaliero	26
3.1.3 Definizione della covariata intubati	27
3.2 Preparazione dataset	29
3.2.1 Estrazione e traposizione dei dati	29
3.2.2 Controllo dei valori plausibili	30
3.2.3 Gestione dati mancanti	36
3.2.4 Aggregazione righe temporalmente vicine con valori simili di PaO_2	37

4. METODI	41
4.1 Modelli proposti	41
4.1.1 Modello ad effetti misti	41
4.1.2 La variabile temporale nel modello ad effetti misti	49
4.1.3 Confronto costruzione modello LMER vs. LME	51
4.1.4 Modello logistico	53
4.1.5 Calibration belt	57
4.2 Validazione del modello	58
4.2.1 Definizione dell'errore: Median Absolute Error	59
4.2.2 Rappresentazioni grafiche dei residui	60
5. RISULTATI	61
5.1 Grafici andamenti variabile risposta	61
5.2 Risultati descrittivi dei modelli	68
5.3 Risultati sul test di validazione	76
5.4 Risultati modello logistico	83
6. CONCLUSIONI	87
7. POSSIBILI SVILUPPI FUTURI	91
Bibliografia	93

1. INTRODUZIONE

La terapia intensiva rappresenta una delle aree più complesse dell'assistenza sanitaria: il trattamento dei pazienti è particolarmente impegnativo sia in termini di scelte terapeutiche che in termini economici anche perché esige il monitoraggio costante di una serie di parametri i cui valori devono essere acquisiti con modalità invasive per il paziente e gravose per l'organizzazione ospedaliera e, di riflesso, per l'intero servizio sanitario.

Il presente studio si propone di implementare, con particolare riguardo ai soggetti affetti da distress respiratorio acuto e sottoposti a ventilazione meccanica, un modello predittivo avente come riferimento un parametro rilevante nell'ambito della terapia intensiva, la pressione parziale di ossigeno nel sangue, in quanto estremamente informativo della condizione clinica del paziente ed idoneo ad orientare alla massima efficacia l'intervento terapeutico.

1.1 Ventilazione meccanica

La ventilazione meccanica è la tecnica di supporto vitale maggiormente utilizzata e il suo corretto utilizzo è una delle problematiche più complesse in ambito medico. Patologie come l'insufficienza respiratoria acuta, la sindrome da distress respiratorio acuto (ARDS), la broncopneumopatia cronica ostruttiva (BPCO) e l'asma, che compromettono la funzione respiratoria naturale del corpo, possono richiedere la ventilazione meccanica [1].

Esistono diverse modalità di ventilazione, alcune utilizzate principalmente in situazioni di emergenza, altre adoperate come supporto a lungo termine alla funzione respiratoria del paziente. In base alla sua condizione e al tipo di patologia respiratoria, si sceglie la modalità di ventilazione più adatta; si parla, ad esempio, di ventilazione meccanica a pressione positiva, ventilazione meccanica controllata, ventilazione meccanica assistita e ventilazione meccanica spontanea.

La ventilazione è una parte essenziale dello scambio gassoso, basato su quattro processi - erogazione di ossigeno, escrezione di anidride carbonica, corrispondenza di ventilazione e perfusione e diffusione - che consentono al sistema respiratorio di mantenere normali pressioni parziali di ossigeno (PaO_2) ed anidride carbonica ($PaCO_2$) nel sangue arterioso [2]. Pertanto, la ventilazione è una tecnica fondamentale perché fornisce ossigeno, elimina anidride carbonica e determina i rapporti ventilazione-perfusione.

In generale, la ventilazione meccanica non è un trattamento a sè stante ma fa parte di una strategia globale; infatti, in un paziente ventilato è importante considerare molteplici aspetti come il trattamento eziologico, la gestione della sedazione, la fisioterapia e la prevenzione della perdita muscolare.

Sebbene la ventilazione meccanica sia un trattamento salvavita, è importante tenere in considerazione anche le complicazioni legate all'utilizzo di questa tecnica. È infatti definita nell'articolo "Mechanical Ventilation: State of Art" [3], *un male necessario*, in quanto può

portare a importanti potenziali complicazioni, tra cui quelle respiratorie, cardiovascolari, metaboliche e psicologiche. I rischi più significativi che possono emergere sono l'infezione polmonare, il danno polmonare indotto dalla stessa e la sindrome da distress respiratorio acuto (ARDS).

1.2 ARDS

La sindrome da distress respiratorio acuto (ARDS) è una causa comune di insufficienza respiratoria nei pazienti critici ed è definita dalla necessità di ventilazione meccanica. Questa sindrome è presente in circa il 10% di tutti i pazienti nelle unità di terapia intensiva in tutto il mondo [4], e si presenta più comunemente nel contesto di polmonite (batterica e virale), sepsi non polmonare, aspirazione di contenuto gastrico e/o orale ed esofageo che può portare ad una successiva infezione e traumi maggiori, quali ferite contudenti o ustioni. L'ARDS è stata definita nel 1994 dall'American-European Consensus Conference (AECC); da allora sono emerse questioni riguardanti l'attendibilità e la validità di questa definizione.

Nel 2011, un gruppo di esperti (un'iniziativa della Società Europea di Medicina Intensiva e approvata dalla Società Americana di Terapia Toracica e dalla Società di Medicina Intensiva) ha sviluppato la Definizione di Berlino [5], concentrandosi su fattibilità, affidabilità, validità e una valutazione obiettiva delle sue prestazioni.

Secondo la definizione, *la sindrome viene identificata sulla base dei seguenti aspetti:*

- *Tempistica: insufficienza respiratoria entro 1 settimana da un insulto noto o sintomi respiratori nuovi e/o in peggioramento*
- *Origine: insufficienza respiratoria non completamente spiegata dalla funzione cardiaca o dal sovraccarico di volume (è necessario un criterio oggettivo come l'ecocardiografia per escludere l'edema idrostatico se non è presente alcun fattore di rischio)*
- *Imaging: opacità bilaterali alla radiografia del torace o alla TC non completamente spiegate da versamento, collasso o noduli*
- *Ossigenazione: insorgenza acuta di ipossiemia definita come $PaO_2/FiO_2 < 300$ mm-Hg su almeno PEEP 5 cmH₂Oa (positive end-expiratory pressure). In particolare si propongono 3 categorie di ARDS che si escludono a vicenda in base al grado di ipos-*

siemia: lieve ($200 \text{ mm Hg} < PaO_2/FIO_2 ; \leq 300 \text{ mm Hg}$), moderata ($100 \text{ mm} < PaO_2/FIO_2 \leq 200 \text{ mm Hg}$) e grave ($PaO_2/FIO_2 \leq 100 \text{ mm Hg}$).

Si considerano anche 4 variabili ausiliarie per ARDS grave: gravità radiografica, compliance del sistema respiratorio ($\leq 40 \text{ mL/cmH}_2O$), pressione positiva di fine espirazione ($\geq 10 \text{ cmH}_2O$) e volume espiratorio corretto al minuto (≥ 10 litri/minuto).

Il trattamento per l'ARDS non implica l'utilizzo di terapie farmacologiche specifiche ma si concentra sulla ventilazione polmonare protettiva; nonostante alcuni miglioramenti, la mortalità rimane alta, circa al 30-40% [6].

La maggior parte dei pazienti con ARDS necessita di sedazione, intubazione e ventilazione durante il trattamento della lesione sottostante. È possibile utilizzare qualsiasi modalità del ventilatore, secondo la Surviving Sepsis Clinical Practice Guideline e la rete ARDS del National Heart, Lung, and Blood Institute (ARDSNet). Vengono impostati la frequenza respiratoria, il tempo espiratorio, la pressione positiva di fine espirazione (PEEP) e la FiO_2 , in accordo con i protocolli ARDSNet.

Le impostazioni vengono regolate per mantenere una saturazione di ossigeno dall'88% al 95% e una pressione di plateau di 30 cmH_2O o meno per evitare il barotrauma¹ [7], una complicanza della ventilazione meccanica correlata con l'aumento della morbilità e della mortalità [5]. Il meccanismo naturale della respirazione negli esseri umani dipende da pressioni intratoraciche negative. Al contrario, i pazienti in ventilazione meccanica ventilano con pressioni positive [8] e dato che tali pressioni non sono fisiologiche, ciò può portare a complicazioni come il barotrauma.

¹Il barotrauma è un danno al tessuto corporeo secondario alla differenza di pressione nelle cavità chiuse all'interno del corpo. Comunemente osservato nei subacquei, negli apneisti o anche nei passeggeri degli aerei durante la salita e la discesa. Gli organi più comunemente colpiti dal barotrauma sono l'orecchio medio (barotrauma otico), i seni paranasali (sinus barotrauma) e i polmoni (barotrauma polmonare)

1.3 Target ossigenativi

Nell'ambito della terapia intensiva le decisioni cliniche quali le variazioni del setting ventilatorio o la modulazione della frazione inspiratoria di ossigeno (FiO_2) sono prese in accordo con la linea guida riferita a valori di pressione parziale di O_2 (PaO_2) o derivati (rapporto PF: PaO_2/FiO_2)[9].

Il parametro PaO_2 indica la quantità di ossigeno presente nel sangue arterioso. In una soluzione in equilibrio con una miscela di gas contenente O_2 , la concentrazione di ossigeno in soluzione dipenderà, in generale, dal PaO_2 .

La concentrazione di ossigeno in una soluzione in equilibrio con una miscela di gas contenente ossigeno dipenderà in modo significativo dalla pressione parziale dell'ossigeno nella miscela di gas per il principio della legge di Henry. Essa stabilisce che la quantità di un gas che si dissolve in un liquido è direttamente proporzionale alla pressione parziale di quel gas sopra la superficie del liquido e descrive questa relazione con l'equazione: $C = k * PaO_2$ dove C rappresenta la concentrazione di ossigeno nella soluzione e k è una costante di proporzionalità che dipende dalla temperatura e dalle caratteristiche della soluzione [10].

PaO_2 è un indicatore importante della funzione polmonare e del grado di ossigenazione del sangue, misurato attraverso un prelievo di sangue arterioso ed espresso in millimetri di mercurio (mmHg); il valore ottimale per un paziente sano si attesta fra 80 e 100 mmHg. Valori superiori a 100 mmHg possono essere causati da una respirazione iperventilata o da una terapia con ossigeno ad alta concentrazione; tali valori possono portare a complicazioni a lungo termine poiché l'eccesso di ossigeno può causare danni tissutali [11].

I valori inferiori a 75 mmHg possono essere causati da una varietà di condizioni, tra cui l'insufficienza respiratoria, l'asma, la bronchite cronica, la polmonite o la sindrome da distress respiratorio acuto (ARDS). In questa situazione si ha una ridotta capacità dei polmoni nello scambio ossigeno-anidride carbonica.

Valori molto bassi possono essere pericolosi per la vita del paziente e possono richiedere un intervento medico immediato per aumentare l'apporto di ossigeno al corpo.

La letteratura [12] propone come target ossigenativo, in associazione al PaO_2 , il rapporto PF ovvero il rapporto tra PaO_2 e FiO_2 (frazione inspirata di ossigeno), espresso in millimetri di mercurio. Una riduzione del valore del PF può essere un indicatore di insufficienza respiratoria e di necessità di supporto respiratorio.

Gli autori dello studio denominato “LOCO2 trial (Liberal Oxygenation Versus Conservative Oxygenation in Patients With Acute Respiratory Distress Syndrome: Impact on Mortality)” [13] riportano la tesi per la quale la terapia con ossigeno conservativa non dovrebbe essere considerata come la strategia di prima scelta per la gestione degli ARDS. Si è confrontata, come in precedenza, l'efficacia della strategia ossigenativa conservativa contro quella convenzionale in pazienti con diagnosi di ARDS. In questo caso si evidenzia l'aumento di mortalità a 90 giorni nei pazienti con terapia ossigenativa conservativa associato ad un incremento del rischio di ischemia mesenterica.

Dopo 3 mesi, il 44.4% dei pazienti sottoposti a terapia conservativa è deceduto, mentre solo il 30.4% del gruppo sottoposto a terapia liberale muore.

Ad oggi sono diversi gli studi in corso di svolgimento che si occupano di target ossigenativi nell'ambito della terapia intensiva tra cui: “Blood Pressure and OXYgenation Targets After Out-of-Hospital Cardiac Arrest (BOX) (NCT03141099), conservative versus CONventional oxygenation target in Intensive Care Patients” (ICONIC) (NTR7376) [14], studio clinico randomizzato che confronta due approcci di ossigenazione in pazienti sottoposti ad arresto cardiaco extra-ospedaliero. Il trial prevede il confronto tra due gruppi di pazienti: uno che riceve una terapia di ossigenazione conservativa e uno che riceve una terapia di ossigenazione convenzionale. L'obiettivo dello studio è di determinare se ci sia una differenza significativa nel risultato clinico tra i due gruppi.

Sempre con lo stesso scopo, in letteratura si incontrano diversi altri studi randomizzati per

definire una scelta tra ossigenazione liberale e restrittiva tra cui il progetto “Strategy to Avoid Excessive Oxygen for Critically Ill Trauma Patients” [15] che cerca di valutare se una strategia di ossigenazione restrittiva, che mira ad evitare l’ossigenazione eccessiva, possa migliorare l’esito del paziente.

Maggiore rilevanza può essere attribuita al “Mega Randomised Registry Trial Comparing Conservative vs. Liberal Oxygenation” (Mega-ROX) trial [16], attualmente in fase di reclutamento che prevede una numerosità campionaria molto elevata, pari a 40000 pazienti.

Lo studio si concentra sulla valutazione dell’effetto dell’ossigenoterapia conservativa rispetto alla terapia liberale sulla mortalità a 90 giorni ed altri esiti clinici nei pazienti con insufficienza respiratoria acuta, introducendo il concetto di tailoring² del target terapeutico. Un significativo avanzamento di questo trial è rappresentato dalla possibilità di identificare un “target ideale di PaO_2 ”.

In conclusione, è emerso come la letteratura riporti opinioni contrastanti sui livelli ottimali di ossigenazione per i pazienti. Questo suggerisce che i livelli intermedi di PaO_2 o SpO_2 possono causare problemi di ipossia³ ed iperossia⁴ nei tessuti. Pertanto, è necessario adattare la terapia ossigenativa ai singoli pazienti (tailoring), con l’obiettivo di mantenere un livello di ossigenazione adeguato. Tuttavia questo richiede un monitoraggio attento dei livelli di PaO_2 per evitare l’esposizione involontaria e pronunciata all’iperossia [17].

Si conferma dunque la centralità delle misurazioni e delle registrazioni dei valori di pressione parziale di ossigeno nel sangue nel delineare il quadro clinico del paziente. Mediante le indicazioni che il valore di pressione parziale è in grado di fornire, è possibile intervenire modificando il setting ventilatorio per migliorare le condizioni cliniche del paziente.

²La personalizzazione della terapia a seconda delle esigenze del singolo paziente

³Insufficiente disponibilità o utilizzazione dell’ossigeno da parte di tessuti o di organi, per cause varie (asfissia, diminuzione del numero dei globuli rossi, ecc.) ma non per ridotto afflusso di sangue arterioso (nel qual caso si parla di ischemia)

⁴Aumento della pressione parziale di ossigeno nell’organismo, conseguente ad aumento di pressione dell’ossigeno nell’aria inspirata (si verifica, tra l’altro, in seguito a somministrazione di ossigeno puro o comunque a pressione superiore a quella esistente nell’aria atmosferica)

1.4 Indicatori alternativi a PF

Nell'orientarsi tra la letteratura riguardante la pressione parziale di ossigeno, sono emersi molteplici studi che fanno riferimento a nuovi indicatori alternativi al PF, surrogati dello stesso PF come la saturazione periferica di ossigeno (pulsossimetria SpO_2) [18]. Recenti studi, come quello condotto da Sheetal Babu, Kundavaram PP Abhilash e Mahasampath Gowri, hanno confrontato il rapporto SpO_2/FiO_2 , misurato tramite l'ossimetria del polso, con il rapporto PF, rilevato mediante emogasanalisi arteriosa. Lo studio ha coinvolto 122 pazienti che hanno ricevuto ossigeno tramite varie modalità, tra cui cannula nasale, maschera intera e ventilazione meccanica invasiva. I risultati hanno evidenziato una forte correlazione tra il rapporto SpO_2/FiO_2 e il rapporto PF, in particolare nelle situazioni in cui sono utilizzate diverse modalità di somministrazione di ossigeno. I dati di questo studio supportano l'utilizzo del rapporto SpO_2/FiO_2 come indicatore alternativo del PF in determinate situazioni cliniche [19].

Ulteriori studi sperimentali [20] si sono concentrati su altri indicatori, in particolare sulla previsione della mortalità basata sul rapporto SaO_2/FiO_2 (Saturazione arteriosa di ossigeno / frazione inspirata di ossigeno) utilizzando l'analisi del database eICU.

Gli autori hanno esaminato i dati di oltre 190000 pazienti adulti ricoverati in terapia intensiva in circa 200 unità di terapia intensiva in tutto il mondo, arrivando alla conclusione che il rapporto SaO_2/FiO_2 risulta essere un forte indicatore della mortalità. Tale rapporto può essere utilizzato come un parametro di monitoraggio per i pazienti ricoverati in terapia intensiva ed appare utile nella gestione del trattamento.

Studi recenti riguardanti la funzione dei target ossigenativi nell'ambito della terapia intensiva non sono stati in grado di fornire risultati concordi su quale parametro utilizzare come target. Nello studio randomizzato monocentrico Oxygen-ICU [21] si prende in considerazione come target ossigenativo la saturazione di ossigeno nel sangue (SpO_2). È stata condotta un'analisi

si su pazienti ricoverati in terapia intensiva con l'obiettivo di valutare l'effetto della terapia conservativa dell'ossigeno rispetto alla terapia convenzionale.

La terapia conservativa consiste nel mantenere i livelli di saturazione dell'ossigeno tra il 94% e il 98% con valori corrispondenti di PaO_2 pari a 70-100 mmHg, mentre la terapia convenzionale prevede di mantenere i livelli di saturazione sopra il 97% (valori corrispondenti di PaO_2 fino a 150 mmHg). La relazione tra PaO_2 e SpO_2 è approssimativamente descritta dalla curva di dissociazione ossiemoglobinica, che è sigmoide. Tale curva mostra che per PaO_2 più bassi, la SpO_2 diminuisce in modo più rapido, mentre per PaO_2 più elevati, il valore di SpO_2 aumenta in modo più graduale [22]. Emerge che per un valori di SpO_2 maggiore del 90% si correla un PaO_2 maggiore di 60 mmHg per oltre il 94% delle rilevazioni effettuate [23].

L'endpoint dello studio Oxygen-ICU [21] prevede di considerare la mortalità: i risultati emersi mostrano un tasso di mortalità significativamente più basso nel braccio ossigenoterapia conservativa (11.6% vs 20.2%) e i soggetti sottoposti a tale terapia (SpO_2 target 94-98% (PaO_2 70-100 mm Hg) presentano un numero minore di episodi di shock, insufficienza epatica e batteriemia⁵.

Lo studio condotto da Girardis et al. conclude che la terapia conservativa dell'ossigeno può essere considerata una strategia efficace e sicura per i pazienti ricoverati in terapia intensiva che richiedono ossigenoterapia.

È emerso, però, che parte della letteratura non sia concorde nel sostenere che metodologie meno invasive, quali SpO_2 , siano in grado di definire in modo univoco i target ossigenativi [24].

⁵Condizione morbosa, rilevata dall'emocoltura, caratterizzata dalla presenza transitoria nel sangue di germi patogeni

2. IPOTESI DI RICERCA

2.1 Quesito statistico

In un contesto come quello descritto fino a questo momento, nel quale è emersa una spiccata componente di incertezza, è necessario individuare target ossigenativi validi in grado fornire informazioni valide e continue al clinico.

La misurazione di PaO_2 è emerso essere un indicatore significativo per valutare i danni causati dall'iperossia (eccesso di ossigeno) nel corpo. La pressione parziale di ossigeno nel sangue risulta essere un fattore predittivo importante per valutare l'entità dei danni causati dall'eccesso di ossigeno, specialmente se si tiene conto del fatto che l'esposizione prolungata può aumentare il rischio di danni significativi. Appare fondamentale offrire un monitoraggio continuo di PaO_2 che consenta di rilevare tempestivamente le variazioni nel livello di ossigeno nel sangue e di intervenire prontamente per prevenire o limitare eventuali danni. Si prefigura, dunque, come una misura di monitoraggio essenziale per garantire la sicurezza dei pazienti e prevenire l'insorgenza di complicanze [25].

Nella pratica clinica, è comune monitorare quattro volte al giorno il valore di pressione parziale di ossigeno nelle Unità di Terapia Intensiva. Ciò è dovuto al fatto che il parametro di interesse è misurato tramite un prelievo arterioso, maggiormente invasivo ed oneroso, per cui si preferisce non sottoporvi il paziente troppo frequentemente ricorrendo ad intervalli regolari più ampi, a meno di necessità cliniche. Tuttavia, le misurazioni di altri parametri vengono effettuate con maggiore frequenza, sebbene non offrano lo stesso livello di informazioni fornite dalla PaO_2 .

L'idea di partenza consiste nell'utilizzare anche le misurazioni di pressione registrate con una minore frequenza al fine di ottenere dei punti di riferimento nel tracciato della PaO_2 per ciascun paziente durante tutto il periodo di permanenza nell'unità intensiva.

L'obiettivo principale è ottenere un sistema predittivo che consenta di stimare PaO_2 in base ad una serie di misurazioni più frequenti e continue, integrando tali informazioni con i punti di riferimento ottenuti dalle misurazioni a frequenza ridotta. Ciò permetterebbe di avere una visione più completa e dettagliata delle variazioni dei livelli di ossigeno nel sangue nel tempo per ciascun individuo, contribuendo a un monitoraggio più accurato delle condizioni cliniche.

2.2 Database MargheritaTre

La terapia intensiva rappresenta uno dei più complessi e costosi componenti dell'Assistenza Sanitaria Nazionale.

Il GiViTI (Gruppo italiano per la Valutazione degli interventi in Terapia Intensiva) è una rete di terapie intensive italiane nata nel 1991 che si pone come scopo principale promuovere e realizzare progetti di ricerca indipendenti, focalizzati sulla valutazione e il miglioramento della qualità dell'assistenza e sull'utilizzo razionale delle risorse. Il GiViTI coinvolge circa 250 unità di terapia intensiva in Italia, rappresentando un'importante iniziativa per l'avanzamento dell'assistenza sanitaria critica nel paese.

In particolare, il progetto MargheritaTre si propone di favorire la collaborazione tra diverse unità di terapia intensiva con l'obiettivo di migliorare la sicurezza dei pazienti ricoverati e valutare costantemente le prestazioni del centro ospedaliero, confrontando esiti e procedure, attraverso l'adozione di una cartella clinica elettronica comune. Inoltre, il progetto contribuisce alla diffusione delle migliori pratiche cliniche e alla riduzione di errori clinici.

Il software MargheritaTre è stato implementato in collaborazione con specialisti intensivisti ed infermieri dell'area critica e dell'Istituto di Ricerche Farmacologiche Mario Negri.

Attualmente utilizzato in circa 70 reparti di Terapia Intensiva, è stato progettato con l'obiettivo di integrare la pratica clinica con la ricerca. Il software consente la raccolta di ogni tipo di informazione clinica, rispettando pienamente la normativa privacy vigente garantendo un alto livello di qualità. Questi dati vengono utilizzati per condurre analisi statistiche volte a confrontare i processi clinici ed identificare punti di forza e debolezze nelle prestazioni, oltre a poter fornire un importante fonte per la ricerca sui diversi parametri clinici raccolti nelle unità di terapia intensiva. Data la vastità e la complessità dei dati raccolti, essi sono stati organizzati in tabelle e viste per renderli più facilmente accessibili e gestibili. Le tabelle rappresentano la struttura dati fondamentale, mentre le viste, definite come viste virtuali o proiezioni di tabelle,

offrono una prospettiva semplificata di specifiche porzioni dei dati, rendendo più agevole il recupero delle informazioni desiderate. Poiché ogni tabella o vista contiene solo una parte dei dati, ad esempio raccoglie le informazioni sui parametri raccolti oppure sui dati anagrafici, può essere necessario eseguire operazioni di *join* per combinare le informazioni provenienti da diverse fonti ed ottenere una visione più completa ed integrata dei dati.

Nel contesto di questo studio, il linguaggio SQL (Structured Query Language) si rivela particolarmente comodo per la visualizzazione dei dati, poiché consente di effettuare interrogazioni specifiche e ricavare informazioni precise senza dover scaricare l'intera mole di dati. L'utilizzo di SQL offre una maggiore efficienza e rapidità nella ricerca di dati specifici, facilitando la comprensione della struttura dati che verrà poi sfruttata in un altro ambiente per svolgere le analisi mediante il linguaggio di programmazione R.

2.3 Selezione pazienti e centri ospedalieri

Passaggio preliminare fondamentale è selezionare alcuni centri ospedalieri per addestrare dei modelli successivamente proposti. La scelta di tali centri è uno dei passaggi da considerarsi forse più semplice, ma di grandissima importanza poiché la qualità dei dati raccolti può essere in grado di influenzare sensibilmente i risultati finali.

Dei centri ospedalieri a disposizione nel progetto MargheritaTre è stato importante selezionare delle strutture che rispettassero criteri di qualità del dato e di numerosità: è stata considerata la validità e la precisione con cui tali centri sono soliti raccogliere i parametri dei pazienti critici in terapia intensiva. Queste informazioni sono state fornite dal supporto e dall'esperienza dei ricercatori del Laboratorio di Clinical Data Science dell'Istituto di Ricerche Farmacologiche Mario Negri di Bergamo, i quali hanno indicato i centri maggiormente adatti allo scopo dello studio.

In questa prima analisi, la lista dei centri ospedalieri è stata ridotta a cinque strutture considerando un numero più limitato al fine di evitare un eccessivo carico computazionale. Inoltre, è stato essenziale assicurarsi che le strutture selezionate fossero in grado di raccogliere un numero adeguato di misurazioni per la costruzione di modelli predittivi affidabili.

Infine, il numero appropriato di centri selezionati non è unicamente determinato dal totale di pazienti ricoverati, ma è influenzato dalla durata del periodo in rianimazione. Dal momento che vengono effettuate misurazioni ripetute per lo stesso individuo durante la permanenza nel reparto intensivo, si hanno misurazioni multiple per ogni paziente.

Tali misurazioni multiple consentono di avere un'ampia copertura di osservazioni per ogni soggetto, comprendendo diverse condizioni ed andamenti clinici che possono manifestarsi durante il periodo di ricovero intensivo.

A seguito di tali premesse, sono 5 i centri ospedalieri selezionati qui riportati:

Centro ospedaliero	Numero osservazioni	Numero pazienti
1	15266	199
2	30646	2341
3	2705	3087
4	17523	555
5	1386	1617

Tabella 1: Centri ospedalieri considerati

La selezione dei pazienti è stata attentamente pianificata al fine di garantire l'omogeneità dei dati: sono stati presi in considerazione soggetti ricoverati in unità di terapia intensiva presso i centri selezionati. Inoltre, come criterio di inclusione è stata considerata l'età dei pazienti, limitando l'analisi ad individui di età superiore ai 18 anni. Tale decisione è utile per garantire l'uniformità dei dati, evitando la variazione potenziale dei parametri di interesse associata all'età poiché tali valori dei parametri potrebbero differire significativamente da quelli degli adulti, complicando l'analisi e l'interpretazione dei risultati.

È importante sottolineare che, per quanto riguarda l'intubazione dei pazienti, questa non è stata considerata come criterio di selezione, bensì come descritto in seguito (cfr. 3.1.3) è stata trattata come covariata, prendendo in considerazione il suo potenziale impatto sulle variabili di interesse.

3. GESTIONE DATI

La pulizia dei dati è un processo centrale per svolgere un'analisi statistica. Il processo consiste nella revisione, nella correzione e nell'eliminazione di dati incompleti, duplicati, errati o incoerenti all'interno del dataset. È fondamentale poiché garantisce la qualità e l'affidabilità delle informazioni raccolte.

Da tale processo dipende l'adeguata costruzione di un dataset con caratteristiche definite: per ogni riga si ha un paziente con il codice identificativo personale e del centro ospedaliero di appartenenza, con tutte le misurazioni rilevate in uno preciso momento durante la permanenza nel reparto intensivo. È opportuno che vengano rimossi i dati mancanti (cfr. 3.2.3) e si abbia una colonna del dataset corrispondente ad ogni parametro misurato. Il mantenimento di un dataset pulito ed aggiornato consente di ottenere una base solida su cui costruire le analisi di interesse, anche a scopo previsivo.

3.1 Variabili considerate

3.1.1 Selezione variabili indipendenti

Il database MargheritaTre raccoglie una grande quantità di dati legati al periodo in terapia intensiva di ogni individuo, riportando le informazioni anagrafiche e della cartella clinica elettronica, in particolare, i farmaci somministrati, i presidi utilizzati e i parametri clinici vitali misurati.

I dati sono inseriti ogni qualvolta vengono registrati: operazione che si svolge in tempo reale, seguendo una scansione temporale dettagliata; ad esempio, ad ogni somministrazione di farmaci o utilizzo di un presidio (come la tracheostomia), tali eventi vengono annotati nella cartella clinica elettronica del paziente. Ogni misurazione di parametri clinici vitali, come la pressione arteriosa, la frequenza cardiaca, la temperatura corporea e altri dati rilevanti, viene registrata e validata da medici o infermieri nel momento stesso in cui è posta in essere.

Data la disponibilità di una grande mole di dati e di informazioni legate ad ogni paziente, è importante definire quali variabili indipendenti considerare. Tale scelta viene effettuata basandosi su aspetti prettamente clinici, considerando i parametri che possano essere in grado di descrivere l'andamento della pressione parziale di ossigeno nel sangue. In collaborazione con i clinici che hanno supportato il progetto è stato possibile redigere una lista di parametri da estrarre dal database MargheritaTre, riportati nella Tabella 2:

Nome parametro	ID parametro
Emoglobina	74
EtCO ₂	28
FCOHb (carbossiemoglobina)	75
FiO ₂	29
FMetHb (metaemoglobina)	76
Frequenza cardiaca	9
Lattati	51
PaO ₂	47
PEEP	35
pH arterioso	45
Pressione arteriosa diastolica	6
Pressione arteriosa media	7
Pressione arteriosa media calcolata	8
Pressione arteriosa sistolica	5
PvO ₂	55
Saturazione periferica SpO ₂	27
Temperatura corporea	103
Tipo respiro	30

Tabella 2: Tabella parametri considerati

In ordine ai parametri presenti nella tabella di cui sopra, si osserva quanto segue sottolineando il legame clinico delle covariate con la variabile risposta PaO_2 :

- **Emoglobina:** è una proteina contenuta nei globuli rossi responsabile del trasporto dell'ossigeno dai polmoni ai tessuti del corpo. La sua capacità di legarsi all'ossigeno influenza direttamente la quantità di ossigeno trasportata nel sangue. Una concentrazione bassa di emoglobina può ridurre la capacità del sangue di trasportare ossigeno influenzando sulla PaO_2 [26].
- **End-tidal CO₂:** è la concentrazione di anidride carbonica espirata alla fine di un ciclo respiratorio. Viene utilizzata come indicatore dell'efficacia della ventilazione polmonare e del metabolismo cellulare. Cambiamenti significativi nei livelli di $EtCO_2$ possono suggerire problemi respiratori, alterazioni dell'equilibrio acido-base, fino a descrivere uno stato di shock. L' $EtCO_2$ può essere correlata alla PaO_2 poiché entrambi

sono coinvolti nel processo di scambio di gas a livello polmonare e tissutale ma offrono una lettura diversa a livello clinico in quanto marcatori di due meccanismi differenti di scambio di ossigeno e CO_2 . [27].

- **FCO_{Hb}**: la carbossiemoglobina, indicata in forma abbreviata *COHb*, è un complesso stabile formato da monossido di carbonio (CO) ed emoglobina all'interno dei globuli rossi. Elevati livelli possono ridurre la disponibilità di ossigeno trasportato nel sangue, influenzando negativamente la pressione parziale di ossigeno nel sangue. La misurazione dei livelli di *FCO_{Hb}* è importante per valutare l'esposizione a gas tossici o situazioni in cui l'anidride carbonica può essere presente in eccesso nell'ambiente o nel corpo [28].
- **FiO₂**: la frazione inspirata di ossigeno indica la percentuale di ossigeno presente nell'aria inspirata da un paziente. La *FiO₂* è controllata nelle terapie in cui viene somministrato ossigeno mediante maschere o ventilatori. La misurazione accurata della *FiO₂* è cruciale per ottimizzare l'ossigenazione dei pazienti con problemi respiratori o con ipossiemia. Una frazione adeguata può contribuire a migliorare la *PaO₂* e l'ossigenazione dei tessuti. Essa è in funzione della condizione del paziente: in condizioni normali, la *FiO₂* nell'aria è di circa 21%, pari all'ossigeno presente nell'aria che normalmente si respira. Nei pazienti che richiedono una supplementazione di ossigeno, i valori variano. Ad esempio, nei pazienti con lieve ipossiemia, la frazione può essere aumentata al 24%-28%. Per un'ipossiemia grave, *FiO₂* può essere aumentata al 30%-50% ricorrendo a maschere ad alta concentrazione o ventilatori. Nei pazienti ventilati meccanicamente in terapia intensiva, i valori di *FiO₂* variano dal 40% al 100% a seconda della gravità dell'insufficienza respiratoria [29].
- **FMetHb**: la metaemoglobina è una forma di emoglobina in cui il ferro è ossidato e non può legarsi all'ossigeno. Livelli elevati possono ridurre la capacità del sangue

di trasportare ossigeno, influenzando la PaO_2 . L'accumulo di metaemoglobina può verificarsi a seguito di esposizione a sostanze ossidanti o a causa di condizioni genetiche e la sua misurazione risulta essenziale per la diagnosi e la gestione delle condizioni associate ad alterazioni emoglobiniche [30].

- **Frequenza cardiaca:** riflette l'efficienza del cuore nell'erogare ossigeno ai tessuti, valutando il numero di battiti in un minuto. Una frequenza cardiaca elevata può indicare uno sforzo del cuore per compensare una ridotta ossigenazione del sangue o una maggiore richiesta metabolica da parte dei tessuti. Il monitoraggio della frequenza cardiaca è importante per valutare l'efficacia del sistema circolatorio e può fornire indizi sulla PaO_2 , poiché il cuore risponde all'ipossiemia aumentando la frequenza dei battiti per cercare di garantire un adeguato apporto di ossigeno ai tessuti [31].
- **Lattati:** prodotti del metabolismo anaerobico, il loro aumento nel sangue può essere un segno di ipossiemia o insufficienza di ossigeno ai tessuti. La misurazione dei lattati è utile per valutare la presenza di acidosi lattica, un'alterazione acido-base che può verificarsi in caso di ipossiemia prolungata o in altre condizioni critiche. Si pensa che i lattati possano essere correlati alla PaO_2 poiché la loro presenza può indicare difficoltà di ossigenazione dei tessuti [32].
- **PEEP:** la pressione positiva nelle vie aeree finale (Positive End-Expiratory Pressure) è una terapia di supporto respiratorio utilizzata in pazienti con insufficienza respiratoria. Con l'ausilio di questo trattamento, viene mantenuta una pressione costante nelle vie aeree durante la fase di espirazione, al fine di migliorare l'ossigenazione e prevenire un collasso alveolare. La PEEP è cruciale per migliorare la PaO_2 in pazienti con compromissione polmonare e ridotta ossigenazione del sangue. L'articolo di "Mechanical Ventilation: State of the Art" di Pham et al. [33] discute l'uso della PEEP nella gestione dei pazienti critici.

- **pH arterioso:** rappresenta l'acidità del sangue arterioso e riflette l'equilibrio acido-base del corpo. Alterazioni del pH possono indicare problemi di ossigenazione e ventilazione polmonare. Un pH arterioso alterato potrebbe influenzare la PaO_2 e il trasporto di ossigeno nel sangue. L'articolo "Interpretation of arterial blood gas" di Puri [34] offre una visione dettagliata sull'interpretazione dei gas del sangue arterioso e la sua rilevanza clinica.
- **Pressione arteriosa diastolica:** rappresenta la pressione nel sistema arterioso durante la fase di rilassamento del cuore. È un parametro importante per valutare l'efficacia del sistema circolatorio nel fornire ossigeno ai tessuti. Una ridotta pressione arteriosa diastolica può indicare problemi nella perfusione dei tessuti ed influenzare la PaO_2 [35].
- **Pressione arteriosa media:** indica la pressione media nel sistema arterioso durante un ciclo cardiaco. Essa rappresenta un'importante misura di perfusione tissutale e ossigenazione. Un valore adeguato è essenziale per garantire un apporto sufficiente di ossigeno ai tessuti, per cui si può supporre che cambiamenti significativi possano influenzare la PaO_2 e possano essere indicativi di problemi circolatori.
- **Pressione arteriosa media calcolata:** si tratta della stima di pressione arteriosa media ottenuta attraverso una formula matematica basata sui valori di pressione sistolica e diastolica. Si tratta di un parametro clinico importante per valutare la perfusione tissutale e la funzione circolatoria. [36]. In presenza sia della pressione arteriosa media (PAM) effettivamente misurata che della PAM calcolata, si è osservato che la frequenza di rilevazione del valore calcolato è significativamente maggiore. Pertanto, si considera di utilizzare esclusivamente la PAM calcolata per le analisi e le valutazioni cliniche successive.

- **Pressione arteriosa sistolica:** rappresenta la pressione nel sistema arterioso durante la fase di contrazione del cuore. Una pressione arteriosa sistolica adeguata può essere un utile parametro per comprendere se l'ossigenazione dei tessuti è sufficiente [37].
- **PvO₂:** la pressione parziale di ossigeno nel sangue venoso rappresenta la quantità di ossigeno disciolto nel sangue venoso. L'importanza di questo parametro verrà trattato diffusamente nel paragrafo 3.1.2.
- **SpO₂:** la saturazione periferica rappresenta la percentuale di emoglobina nel sangue arterioso che è legata all'ossigeno. La misurazione di SpO₂ viene eseguita con un ossimetro e offre un'indicazione immediata dell'ossigenazione del sangue. Un valore inferiore al 95% è spesso considerato un valore associato ad una condizione di ipossiemia. Questo parametro è comunemente utilizzato per monitorare l'ossigenazione dei pazienti [38].
- **Temperatura corporea:** riflette il bilancio termico del corpo, influenzando il metabolismo cellulare e l'ossigenazione dei tessuti. Considerando una delle misurazioni più semplici nell'ambito della terapia intensiva, in questa prima fase, è stata inserita tra le variabili indipendenti selezionate [39].
- **Tipo respiro:** questo parametro non viene utilizzato direttamente nell'ambito dei modelli proposti, ma ha un ruolo importante nell'individuazione dei soggetti intubati in ventilazione meccanica (cfr. 3.2.6).

Si può supporre che la pressione parziale di ossigeno nel sangue sia influenzata da numerosi fattori, di tipo fisiologico e clinico. La misurazione accurata e la comprensione di questi parametri sono fondamentali per valutare l'ossigenazione del paziente e per la gestione di condizioni mediche complesse. La conoscenza della relazione tra queste variabili e la PaO₂ può aiutare ad individuare e trattare velocemente i pazienti critici.

3.1.2 Importanza del controllo di PvO_2 in ogni centro ospedaliero

Nel paragrafo precedente è stato citato il parametro di pressione venosa di ossigeno nel sangue (PvO_2). Non in tutti i centri ospedalieri sono importate come due parametri distinti le misurazioni di PaO_2 e PvO_2 , causando un possibile errore metodologico nella procedura di selezione della variabile risposta.

Al fine di evitare questa criticità, è stata effettuata una verifica preliminare su ogni centro selezionato. Si è controllato che la colonna corrispondente a PvO_2 non fosse sempre mancante, ma avesse alcune misurazioni. Questa semplice verifica ha permesso di accertare in quali centri la variabile di interesse PaO_2 fosse identificata correttamente, per escludere gli altri centri, garantendo una maggiore affidabilità ed accuratezza dei dati raccolti. Tale approccio metodologico contribuisce a garantire la validità dei valori di pressione arteriosa considerati.

3.1.3 Definizione della covariata intubati

A seguito della definizione di PvO_2 , è necessario focalizzarsi su un'ulteriore variabile rilevante. Essa si riferisce all'identificazione della condizione respiratoria del soggetto nell'istante in cui viene effettuata la misurazione. Si tratta di una variabile binaria che classifica i pazienti in intubato o non intubato. In una prima fase, si è pensato di utilizzare tale informazione come criterio di selezione dei soggetti da includere nell'analisi, ma dopo attente valutazioni, si è esclusa questa possibilità poiché avrebbe comportato una sensibile diminuzione della numerosità campionaria. Inoltre, è importante considerare che lo stato di intubazione di un paziente può variare nel corso della permanenza in terapia intensiva. In forza di ciò, si è deciso di inserire l'informazione sulla tipologia di respirazione come variabile indipendente in corrispondenza di ogni rilevazione dei parametri. In questo modo, si potrà tenere conto dell'effetto della ventilazione meccanica sull' PaO_2 senza dover imporre vincoli rigidi sulla selezione dei dati.

Per identificare i soggetti intubati, è stato sviluppato un algoritmo basato sulle informazioni ricavate dal database riferito ai presidi di ogni paziente. In particolare, si è estratto il codice identificativo dei soggetti che presentano i presidi di tubo orotracheale e tracheostomia. In aggiunta a questa informazione, nella fase iniziale di estrazione dei parametri dal database PARAMETRI è stata considerata anche la variabile *tipo respiro*; sono stati selezionati tutti i pazienti che hanno associato al tipo respiro una delle seguenti condizioni respiratorie:

- Controllata
- Assistita
- CPAP⁶ (Continuous Positive Airway Pressure)[40]

⁶In ambito ospedaliero, è una terapia respiratoria non invasiva utilizzata per trattare pazienti con problemi respiratori, fornendo una pressione costante di aria o ossigeno ai pazienti durante l'inspirazione e l'espirazione. Questa pressione continua mantiene le vie aeree aperte, riducendo il collasso delle vie aeree durante l'espirazione e migliorando lo scambio di gas nei polmoni

È stata considerata la data di inizio e fine intubazione che è stata confrontata con l'occasione temporale in cui è avvenuta la misurazione. Se la rilevazione dei parametri rientra nell'intervallo del periodo di intubazione, allora il soggetto è intubato (variabile `intubato = 1`). Dal punto di vista operativo la definizione dei soggetti intubati per ogni misurazione rilevata nel tempo è stata realizzata mediante il codice che segue:

```
dati <- dati %>%  
  mutate(intubato =  
    ifelse(!is.na(DATAINIZIO_intub) &  
           DATA >= DATAINIZIO_intub &  
           DATA <= DATAFINE_intub, 1, 0))
```

3.2 Preparazione dataset

3.2.1 Estrazione e trapiosizione dei dati

La fase di estrazione e preparazione dei dati per l'analisi è finalizzata ad ottenere un dataset adeguatamente strutturato e pronto per l'analisi statistica per modelli ad effetti misti. In essa è compresa l'estrazione delle informazioni anagrafiche e dei parametri clinici dai centri ospedalieri selezionati, la creazione di un unico data frame unendo tali informazioni, e la trasposizione dei parametri di interesse per ottenere una struttura dati adatta alle analisi longitudinali.

L'estrazione dei dati anagrafici e dei parametri clinici viene effettuata solo per i centri ospedalieri selezionati, attraverso funzioni R messe a disposizione dal laboratorio di Clinical Data Science. Una particolare attenzione è dedicata alla manipolazione dei dati temporali. Alcune misurazioni nel data frame presentano una data di "1-01-01 00:00:00", che è considerata una data non congrua. Per affrontare questa problematica, tali misurazioni vengono sostituite con l'ora di ingresso in terapia intensiva del paziente. Inoltre, viene calcolata la distanza in minuti tra la data di ogni misurazione e l'ora di ingresso in TI.

La variabile *distanza minuti* verrà utilizzata successivamente come variabile temporale nei modelli; dal punto di vista clinico ha un ruolo importante poiché si suppone possa essere un indice di severità.

Il dataset ottenuto dall'unione delle informazioni anagrafiche e dei parametri clinici presenta righe distinte per lo stesso soggetto, in base alla data in cui è stata effettuata la misurazione: tuttavia, le informazioni sui parametri di interesse sono contenute in un'unica colonna NOMEPARAMETRO con i relativi valori in VALORENUMERICO. Per rendere il dataset utilizzabile per le analisi statistiche, è necessario eseguire una trasposizione, in modo da ottenere una colonna per ogni parametro clinico e i corrispondenti valori per ogni soggetto e occasione temporale distinta.

3.2.2 Controllo dei valori plausibili

Allo scopo di garantire la qualità e l'affidabilità dei dati utilizzati nell'analisi, è opportuno identificare ed escludere eventuali valori estremi ed incoerenti che potrebbero compromettere l'affidabilità dei risultati. Durante questa fase, verranno elaborati alcuni grafici per individuare eventuali punti anomali ed outlier che potrebbero derivare da errori di rilevazione, strumentali o altre fonti di irregolarità; inoltre, verrà esplicitata la strategia adottata per la gestione dei dati mancanti, in modo da minimizzare l'impatto di tali mancanze sulla precisione e rappresentatività dell'analisi. A tale riguardo per ogni parametro selezionato, è prodotto un boxplot, per comprendere quale sia la variabilità dei parametri ed individuare eventuali valori anomali. Il primo boxplot (Figura 1) riporta la sintesi di età, frequenza cardiaca, End-Tidal CO_2 e la saturazione di ossigeno.

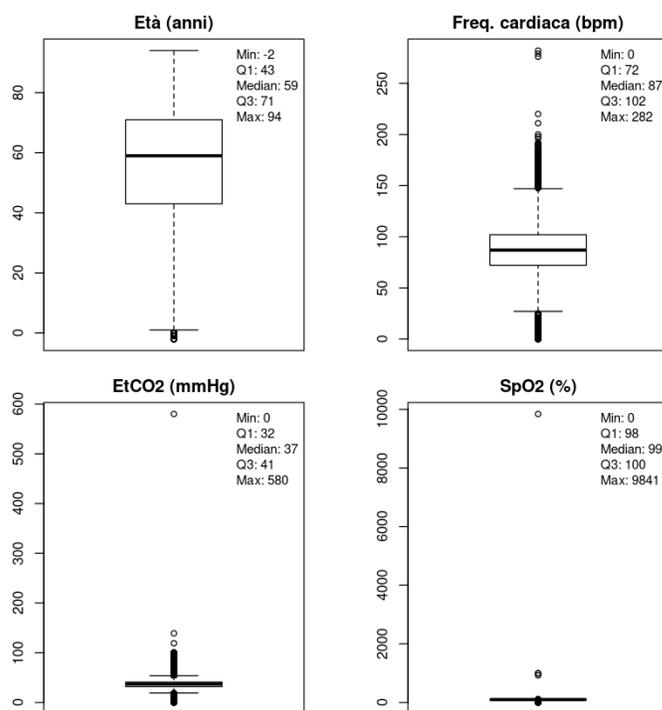


Figura 1: Boxplot età, frequenza cardiaca, $EtCO_2$, SpO_2

Si noti come per l'età del paziente vi sia un valore minimo non plausibile, dato che è negativo: ciò porta ad escludere tale valore, non accettando soggetti con età inferiore a 18 anni come già riportato nei criteri di inclusione (cfr. 2.3).

La frequenza cardiaca copre un ampio intervallo di valori, che possono variare notevolmente in diverse situazioni, da normali condizioni di riposo fino a situazioni cliniche critiche. In questo intervallo, rientrano tutte le misurazioni raccolte nei diversi centri ospedalieri selezionati.

L'End-Tidal CO_2 e la saturazione di ossigeno SpO_2 presentano valori racchiusi tra il valore minimo ed il terzo quartile; sono poche, in particolare per SpO_2 le misurazioni superiori al valore del 75-esimo percentile, le quali sono state eliminate, poiché considerati valori non plausibili.

Valori così estremi, totalmente non compatibili con valori fisiologici (anche critici), possono provenire da errori di inserimento del dato o da misurazioni effettuate da uno strumento non collegato correttamente o non tarato.

Il boxplot sottostante (Figura 2) riporta i valori di questi 4 parametri: metaemoglobina, emoglobina, carbossiemoglobina e la pressione positiva di fine espirazione (PEEP). Tutte le misurazioni dei parametri considerati risultano ragionevoli e non mostrano valori estremi o incongruenti dal punto di vista clinico con la condizione, seppur critica, del paziente.

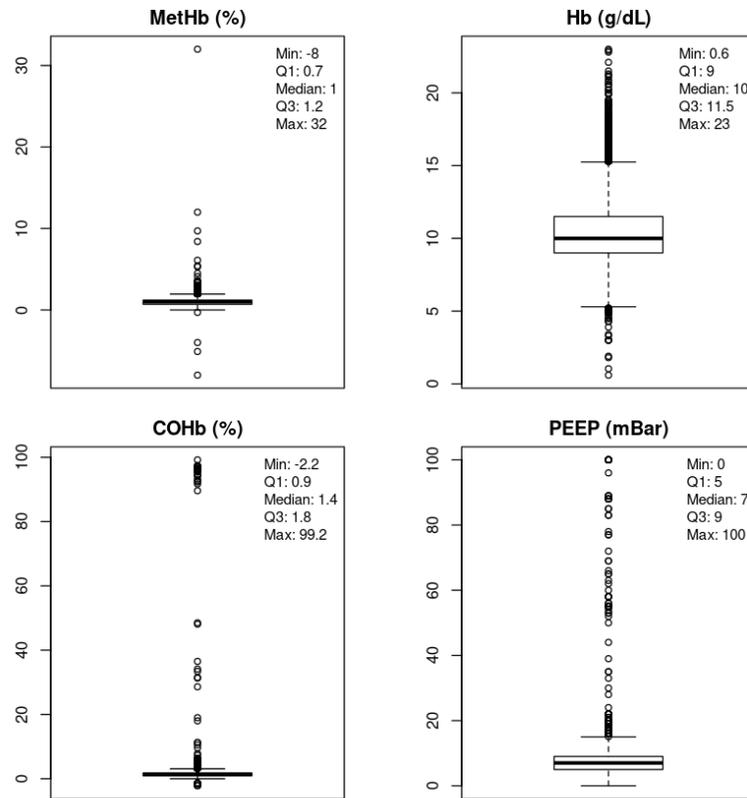


Figura 2: Boxplot MetHb, Hb, COHb, PEEP

I boxplot riportati di seguito (Figura 3) mostrano i differenti parametri relativi alla pressione registrati (sistolica, diastolica, media calcolata) e FiO_2 (frazione inspirata di O_2) Riguardo alle pressioni, si noti come per la sistolica vi sia un massimo di 1607 mmHg, valore estremo non compatibile con valori plausibili: non è possibile risalire al valore corretto, dunque si

preferisce escludere questa misurazione.

Valori di pressione diastolica così elevati (182 mmHg), ma comunque inferiori a 300 mmHg [41] sono difficilmente raggiungibili ma, in un contesto intensivo, ciò può essere associato a una crisi ipertensiva, ictus, infarto miocardico o altre gravi condizioni, per cui viene considerato in parte plausibile e per questo mantenuto.

Il range di FiO_2 si riferisce alla percentuale di ossigeno presente nell'aria che una persona sta inspirando, per cui è espresso in un valore compreso tra 0 e 100: i valori al di fuori di questo intervallo sono esclusi e considerati mancanti.

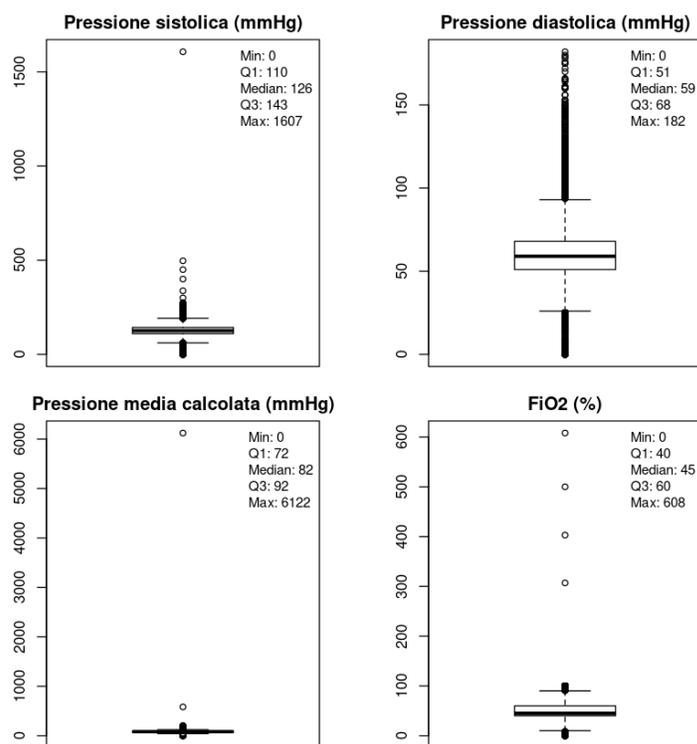


Figura 3: Boxplot pressione sistolica, diastolica, media calcolata, FiO_2

Nell'ultima finestra grafica (Figura 4) vengono mostrati i boxplot relativi a pH, lattati, temperatura e alla variabile risposta, PaO_2 .

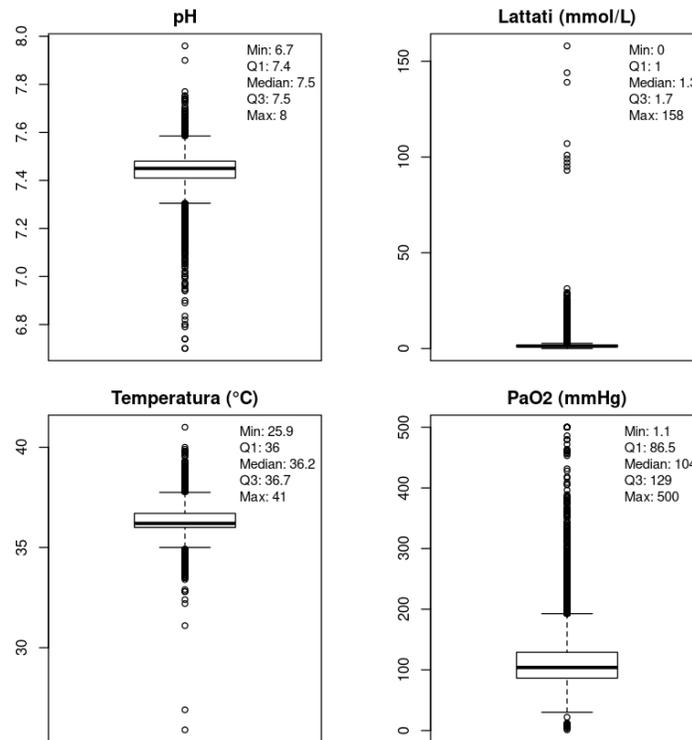


Figura 4: Boxplot pH, lattati, temperatura, PaO_2

Il pH è un parametro che può variare da 0 a 14. In campo clinico, alcuni studi hanno suggerito la possibilità di limitare questo intervallo [42], tuttavia, per evitare errori viene mantenuto l'intero range di valori possibili. I lattati [43] riportano alcuni valori attorno a 100 mmol/l, valori non plausibile anche se i pazienti fossero sottoposti ad un intenso sforzo [44]. Come ultimo grafico, viene proposto quello della variabile risposta, pressione arterosa di ossigeno nel sangue, per cui tutti i valori vengono considerati plausibili, poiché valori estremi appaiono come indicatori di una situazione critica del paziente.

Parametro	Soglia inferiore	Soglia superiore
EtCO ₂ (mmHg)	0	200
SpO ₂ (%)	0	100
FC (bpm)	0	300
MetHb (%)	0	25
Hb (g/dl)	0	25
COHb (%)	0	100
PEEP (mBar)	0	100
Pressione sistolica (mmHg)	0	500
Pressione diastolica (mmHg)	0	200
Pressione media calcolata (mmHg)	0	300
FiO ₂ (%)	0	100
pH	0	14
Lattati (mmol/L)	0	40
Temperatura (°C)	20	50
PaO ₂ (mmHg)	0	600

Tabella 3: Definizione soglie per i parametri selezionati

A seguito dell'analisi grafica dei valori anomali e, tenuto conto delle indicazioni cliniche, è stato definito un intervallo di valori plausibili per ciascun parametro, compresa la PaO_2 . La tabella riepilogativa (Tabella 3) riporta le soglie inferiori e superiori per ogni parametro selezionato.

3.2.3 Gestione dati mancanti

Dopo aver escluso i valori inverosimili e classificandoli come dati mancanti (*NA*), è opportuno considerare come gestire l'elevata quantità di celle con dati mancanti: tale situazione si verifica perché le misurazioni dei parametri di interesse vengono effettuate in momenti diversi e non sempre tutti i parametri sono registrati simultaneamente.

Sono state valutate diverse soluzioni per gestire i dati mancanti, inclusa l'imputazione dei dati mancanti cercando di stimare i valori mancanti in base ad altre informazioni disponibili, tuttavia, si è verificato che tale imputazione non restituisce risultati precisi ed accurati come nelle attese, pertanto si è scelto di non effettuare alcuna imputazione e di mantenere solo le righe che contengono tutti i parametri di interesse. Si è preferito sacrificare la numerosità delle osservazioni da utilizzare per prediligere dati più precisi e completi. Seguendo questa indicazione sono state escluse dall'analisi alcune covariate che presentano una proporzione di dati mancanti superiore al 70%. I parametri esclusi sono i seguenti:

- PEEP (Positive End-Expiratory Pressure)
- $EtCO_2$ (End-Tidal CO_2)
- Temperatura

3.2.4 Aggregazione righe temporalmente vicine con valori simili di PaO_2

Viene definito un dataset longitudinale in cui ogni riga rappresenta diverse misurazioni dei parametri di interesse, registrate in momenti temporali differenti. Ogni paziente ricoverato ha più righe associate, ognuna delle quali corrisponde a una diversa rilevazione delle variabili durante il periodo di permanenza nel reparto intensivo. Il dataset, dunque, segue il percorso temporale dei pazienti e registra le variazioni dei parametri nel corso del tempo durante il ricovero in terapia intensiva. Allo scopo di rendere maggiormente sintetico il dataset utilizzato, ma allo stesso tempo ponendosi l'obiettivo di non perdere informazioni, si è pensato di aggregare le righe temporalmente vicine nel caso in cui esse presentassero valori simili della variabile risposta.

In questo contesto, le misurazioni vengono considerate vicine nel tempo quando sono state effettuate entro un intervallo di 10 minuti l'una dall'altra. La vicinanza nel tempo delle misurazioni permette di analizzare eventuali variazioni o cambiamenti rapidi nei parametri clinici. Oltre alla vicinanza nel tempo, per considerare due record "aggregabili", viene valutata la vicinanza dei valori di PaO_2 . In questo caso, si considerano "simili" i valori che differiscono tra loro per un massimo di 5 mmHg. Unire le righe con valori di PaO_2 simili, calcolando la media, è utile per sintetizzare misurazioni vicine nel tempo che non offrono informazioni aggiuntive riguardo il valore di pressione parziale di ossigeno nel sangue, eliminando delle righe superflue. Sotto riportato un estratto del dataset in fase di pulizia dati in cui viene mostrata la situazione di aggregazione di 2 righe poiché rispettano i criteri sopra definiti e poi alcune righe che nonostante la vicinanza temporale, presentano valori distanti di PaO_2 e per questo motivo non sono state aggregate. Nell'estratto del dataset in fase di pulizia dati si illustrano situazioni in cui due righe sono state aggregate perché rispettano i criteri sopra definiti riguardo alle misurazioni del parametro di interesse registrato in momenti temporali vicini⁷ (Tabella 4).

⁷Distanza minuti: tempo trascorso dall'ingresso in Terapia Intensiva alla misurazione

IDCENTRORICOVERO	SpO₂	FC	MetHb	Hb	FiO₂	PaO₂	distanza minuti
4-100	98	85	0.1	9.0	40	72	80
4-100	98	91	0.5	9.8	42	76	87

Tabella 4: Tabella record da aggregare

La tabella precedente presenta valori della variabile risposta entro la soglia definita di 5 mmHg e una distanza in minuti ravvicinata, entro 10 minuti da una misurazione all'altra. Per questo motivo i due record vengono aggregati, utilizzando l'operatore media per le variabili continue (Tabella 5).

IDCENTRORICOVERO	SpO₂	FC	MetHb	Hb	FiO₂	PaO₂	distanza minuti
4-100	98	88	0.3	9.4	41	74	83.5

Tabella 5: Tabella record aggregati

Tuttavia, sono anche presenti alcune righe che, nonostante la vicinanza temporale, mostrano valori significativamente diversi per il parametro PaO_2 . A causa di questa differenza sostanziale nei valori del parametro, queste righe non sono state aggregate e considerate come misurazioni separate anche se sono state effettuate in momenti ravvicinati nel tempo. La decisione di non aggregare alcune righe, nonostante la loro vicinanza temporale, è stata presa per evitare di combinare misurazioni che presentano cambiamenti ravvicinati nella condizione clinica del paziente.

Queste righe hanno mostrato variazioni importanti nei valori di PaO_2 , indicando che la condizione clinica del paziente si è rapidamente modificata in modo significativo.

Un esempio di questa situazione è riportata nell'estratto di dataset che segue:

IDCENTRORICOVERO	SpO2	FC	MetHb	Hb	FiO2	PaO2	distanza minuti
4-100	100	85	0.7	10.7	60	80	820
4-100	99	75	0.3	8.9	62	55	825

Tabella 6: Tabella record da mantenere separati

Nel contesto presentato, la distanza tra le misurazioni della pressione parziale di ossigeno nel sangue per lo stesso paziente è di soli 5 minuti, inferiore alla soglia di 10 minuti stabilita. I valori di pressione parziale di ossigeno nel sangue hanno subito modifiche significative nel corso di pochi minuti. Questi risultati indicano che le condizioni cliniche del paziente stanno cambiando rapidamente e per questo motivo si preferisce mantenere separati i record.

Il diagramma di flusso, rappresentato nella Figura 5, riassume le numerosità coinvolte. Emerge come la vasta quantità di dati presenti nel database MargheritaTre non possa essere completamente utilizzata a causa del grande numero di informazioni mancanti. Questo si traduce nel fatto che solo un numero relativamente limitato di record ha tutti i dati disponibili.

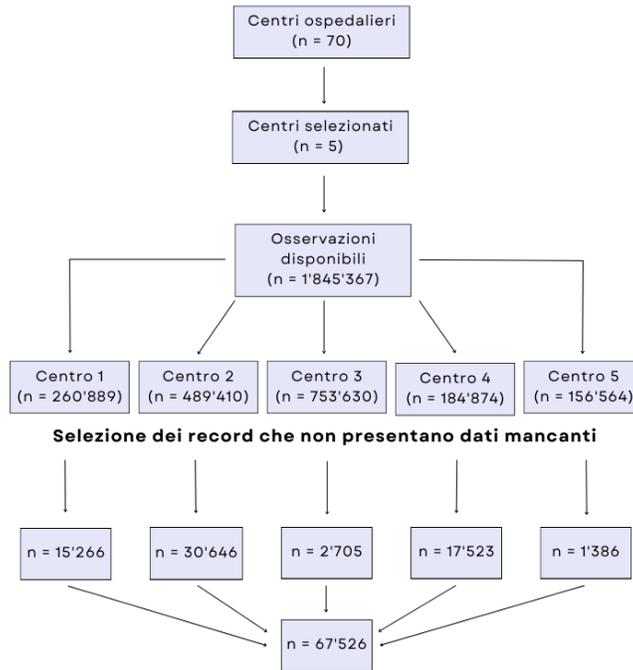


Figura 5: Flowchart selezione misurazioni

È importante sottolineare che la scelta riguardo la selezione delle righe piene e dell'aggregazione secondo le modalità presentate è stata del tutto discrezionale e, durante la discussione dei risultati, verranno valutati i vantaggi e gli svantaggi di questa strategia.

4. METODI

4.1 Modelli proposti

4.1.1 Modello ad effetti misti

Posto come obiettivo quello di ottenere una previsione per la pressione parziale di ossigeno a diversi istanti temporali, è necessario riflettere su quale possa essere la metodologia migliore da implementare. Si tratta di un outcome di tipo continuo che vuole essere spiegato mediante diverse covariate, per lo più continue con la presenza di alcune binarie. A tale scopo il modello statistico che maggiormente si adatta è il modello ad effetti misti [45]. Esso si adatta perfettamente all'utilizzo di dati longitudinali e tiene conto degli effetti fissi delle covariate ed attribuisce una variabilità propria per ogni livello della variabili che viene inserita come effetto casuale, in questo caso l'identificativo del centro. In base all'appartenenza ad un centro, vi sarà un effetto causale associato; è stato scelto di non considerare un effetto casuale sul paziente in modo da ridurre la variabilità. Ciò significa che si assume che la variabilità tra i pazienti sia già catturata efficacemente dalle covariate fisse e dall'effetto casuale dei centri.

Alla base dei modelli lineari ad effetti misti vi è l'assunto che sottoinsiemi di parametri di regressione variano casualmente da un individuo all'altro, tenendo conto delle fonti di eterogeneità naturale nella popolazione. Si assume che gli individui nella popolazione abbiano le proprie traiettorie di risposta medie nel tempo. La caratteristica distintiva dei modelli lineari ad effetti misti è che la risposta media è modellata come una combinazione di caratteristiche della popolazione che si presume siano condivise da tutti gli individui e degli effetti casuali che rappresentano la variabilità residua tra individui non modellata dagli effetti fissi. I primi sono, appunto, gli effetti fissi mentre i secondi sono effetti casuali. Il termine misti è utilizzato in questo contesto per indicare che il modello contiene sia effetti fissi che casuali.

Sebbene il modello lineare ad effetti misti assuma che le risposte longitudinali dipendano da una combinazione di parametri di regressione della popolazione ed effetti specifici del soggetto, esso porta comunque a un modello per la media marginale della risposta (media sulla distribuzione degli effetti casuali) che può essere espresso nella forma $E(Y_{ij}|X_i) = X_i\beta$. Tuttavia, l'introduzione degli effetti casuali induce una covarianza tra le risposte e $Cov(Y_{ij}|X_i) = \sigma_i$ ha una struttura distintiva di effetti casuali. Con l'inclusione degli effetti casuali, le covarianze tra le misure ripetute possono essere espresse come funzioni del tempo.

I modelli lineari ad effetti misti distinguono in modo esplicito le fonti di variabilità tra i centri e all'interno delle strutture ospedaliere. Inoltre, non solo è possibile stimare i parametri che descrivono come cambi la risposta media nella popolazione di interesse, ma è anche possibile prevedere il cambiamento delle traiettorie di risposta individuali nel tempo. Ad esempio, proprio come obiettivo del progetto, questi modelli possono essere utilizzati per ottenere previsioni delle traiettorie individuali.

Il modello ad effetti misti più semplice è quello per cui l'effetto del soggetto varia casualmente. Si ipotizza che ogni centro abbia un livello sottostante di risposta che persiste nel

tempo. Questo effetto della struttura ospedaliera viene incorporato nel modello lineare ad effetti misti considerandolo casuale, ottenendo:

$$Y_{ij} = X'_{ij}\beta + b_i + \epsilon_{ij} \quad (1)$$

dove b_i è l'effetto casuale del singolo e ϵ_{ij} sono considerati errori di misura.

Si può generalizzare l'equazione precedente consentendo ad ulteriori coefficienti di regressione di variare casualmente. La premessa sottostante al modello è che un sottoinsieme dei coefficienti di regressione vari casualmente da un centro all'altro. Nel caso più semplice è l'intercetta a variare casualmente, inducendo una covarianza tra le misure ripetute. In questo modo, sono disponibili diverse strutture di covarianza per gli effetti casuali.

Il seguente modello lineare ad effetti misti presenta intercette e pendenze che variano casualmente tra i centri. Per t_{ij} occasione temporale relativa all' j -esimo istante di misurazione dell' i -esimo centro, si assume che:

$$Y_{ij} = \beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \epsilon_{ij}, \quad j = 1, \dots, n_i \quad (2)$$

Utilizzando la notazione vettoriale e matriciale, il modello lineare a effetti misti può essere espresso come:

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i \quad (3)$$

dove β è un vettore $p \times 1$ di effetti fissi, b_i è un vettore $q \times 1$ di effetti casuali, X_i è una matrice $n_i \times p$ di covariate e Z_i è una matrice $n_i \times q$ di covariate, con $q \leq p$. Qui Z_i è una matrice nota che collega il vettore degli effetti casuali b_i a Y_i .

Gli effetti casuali b_i sono assunti indipendenti dalle covariate X_i con distribuzione multivariata Normale con media zero e matrice di covarianza G , $(E(b_i) = 0, Cov(b_i) = G$.

In linea teorica, si potrebbe essere assumere una distribuzione multivariata a media nulla per

b_i ; a livello pratico si predilige assumere sempre che b_i abbia una distribuzione multivariata Normale.

Una distinzione importante nel modello lineare ad effetti misti è quella tra la media condizionale e la media marginale di Y_{ij} . La media condizionale o specifica del soggetto di Y_i , dato b_i , è

$$E(Y_{ij} | b_i) = X_{ij}\beta + Z_{ij}b_i \quad (4)$$

mentre la media marginale o media di popolazione di Y_i , mediata sulla distribuzione degli effetti casuali b_i , è

$$E(Y_i) = \mu_i = E(E(Y_{ij} | b_i)) = E(X_{ij}\beta + Z_{ij}b_i) = X_i\beta + Z_iE(b_i) = X_i\beta \quad (5)$$

Dato che $E(b_i) = 0$, il vettore dei coefficienti di regressione β si ipotizza essere lo stesso per tutti i centri; ha interpretazioni in termini medi, ma tiene conto degli effetti fissi sulla popolazione. Al contrario, il vettore b_i , in combinazione con gli effetti fissi corrispondenti, è composto da coefficienti di regressione specifici del centro ospedaliero. Infine, il vettore degli errori ε_i di dimensione $n_i \times 1$ è supposto essere indipendente da b_i e avere anche una distribuzione multivariata Normale con media zero e matrice di covarianza R_i . Si assume, inoltre, che R_i sia la matrice diagonale $\sigma^2 I_{n_i}$, dove I_{n_i} è una matrice identità $n_i \times n_i$. In tal caso, ε_{ij} e ε_{ik} sono incorrelate, con varianza uguale e gli ε_{ij} possono essere considerati come errori di campionamento o di misurazione.

Per chiarire la notazione vettoriale e matriciale introdotta finora, si considera il modello lineare ad effetti misti con intercette e pendenze che variano casualmente tra i centri selezionati.

Per l' i -esima struttura ospedaliera al j -esimo istante di misurazione:

$$Y_{ij} = \beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \varepsilon_{ij}, \quad j = 1, \dots, n_i \quad (6)$$

Facendo ricorso alla notazione vettoriale e matriciale, il modello può essere espresso come:

$$Y_i = X_i\beta + Z_ib_i + \varepsilon_i \quad (7)$$

dove

$$X_i = \begin{bmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{i,n_i} \end{bmatrix}, \quad Z_i = \begin{bmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{i,n_i} \end{bmatrix},$$

e

$$b_i = \begin{bmatrix} b_{1i} \\ b_{2i} \end{bmatrix}, \quad \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{i,n_i} \end{bmatrix}.$$

In questo modello, gli effetti casuali sono le componenti del vettore b_i (cioè b_{1i} e b_{2i}), mentre gli errori di misurazione sono gli elementi del vettore ε_i ($\varepsilon_{i1}, \varepsilon_{i2}, \dots$)

In riferimento alla struttura di covarianza indotta dagli effetti casuali

$$Cov(Y_i) = Z_i G Z_i^T + \sigma^2 I_{n_i} \quad (8)$$

consente di analizzare esplicitamente le fonti di variabilità tra individui (matrice G) e all'interno degli individui (matrice R_i). Questo è un aspetto distintivo del modello lineare a effetti misti rispetto ai modelli di struttura di covarianza. I modelli di struttura di covarianza considerano tutte le fonti di variabilità che influenzano la covarianza, ma non distinguono tra le diverse fonti di variabilità. Il modello lineare a effetti misti, al contrario, è in grado di distinguere in modo esplicito le fonti di variabilità tra soggetti e all'interno dei soggetti.

Vi è un'attenzione particolare alla stima degli effetti casuali, poiché di fondamentale importanza per procedere con la previsione di un nuovo soggetto. Dato un nuovo paziente appartenente ad una delle cinque strutture selezionate, conoscere l'effetto casuale per quel soggetto permette di ottenere una previsione che tenga conto non solo dell'effetto fisso delle covariate, ma anche dell'effetto casuale del centro ospedaliero di appartenenza.

In generale, in molte applicazioni sui dati longitudinali, l'inferenza è focalizzata sugli effetti fissi. Questi parametri di regressione hanno un'interpretazione in termini di variazione nella risposta media nel tempo e la loro relazione con le covariate. In alcune situazioni, tuttavia, lo scopo è quello di predire i profili di risposta specifici di ciascun individuo; ad esempio vi sono studi in cui l'obiettivo è identificare gli individui che hanno mostrato il maggior aumento o diminuzione della risposta nel tempo. Dal momento che il modello lineare ad effetti misti distingue esplicitamente tra effetti fissi e casuali, è possibile anche predire traiettorie di risposta

specifiche per ciascun soggetto nel tempo.

Si può dimostrare che la predizione di una variabile casuale equivale a predire la sua media condizionale, sulla base dei dati disponibili. Quindi, il miglior predittore di b_i è la media condizionale di b_i , dato il vettore delle risposte Y_i e β è:

$$E(b_i | Y_i) = GZ_i^T \Sigma_i^{-1} (Y_i - X_i \hat{\beta}) \quad (9)$$

Questo è noto come “miglior predittore lineare non distorto” (BLUP). Questo predittore di b_i dipende dalla covarianza sconosciuta tra le Y_i . Quando i parametri di covarianza non conosciuti sono sostituiti con le loro stime REML (restricted maximum likelihood) o di massima verosimiglianza, il predittore:

$$\hat{b}_i = \hat{G}Z_i^T \hat{\Sigma}_i^{-1} (Y_i - X_i \hat{\beta}) \quad (10)$$

è indicato come “BLUP empirico”, \hat{b}_i , il quale permette di ottenere anche il profilo di risposta previsto per l’i-esimo soggetto:

$$\hat{Y}_i = X_i \beta + Z_i \hat{b}_i \quad (11)$$

Si noti che il profilo di risposta previsto per l’i-esimo soggetto può anche essere espresso come una media pesata tra il profilo di risposta media stimato della popolazione, $X_i \beta$, e il profilo di risposta osservato dell’iesimo soggetto, Y_i .

La quantità di “*contrazione*” verso la media della popolazione dipende dalla magnitudine relativa di R_i e Σ_i ($\Sigma_i = Cov(Y_i) = Z_i G Z_i' + R_i$) R_i caratterizza la variabilità all’interno del soggetto, mentre Σ_i incorpora sia la variabilità all’interno del soggetto che tra i soggetti. Quando R_i è relativamente grande e la variabilità nel soggetto è maggiore rispetto alla variabilità tra soggetti, viene assegnato più peso al profilo di risposta media stimato della popolazione piuttosto che alle risposte osservate dell’i-esimo individuo. D’altra parte, quando

la variabilità tra i soggetti è grande rispetto a quella all'interno del soggetto, viene assegnato più peso alle risposte osservate dell' i -esimo soggetto, Y_i .

Dunque, meno peso è attribuito alle risposte osservate dell' i -esimo soggetto quando la variabilità all'interno del soggetto è relativamente grande e la popolazione è relativamente omogenea.

La quantità di “*contrazione*” verso la media della popolazione dipende anche da n_i , il numero di osservazioni sull' i -esimo soggetto. Vi è una maggiore contrazione verso la media della popolazione quando n_i è piccolo.

4.1.2 La variabile temporale nel modello ad effetti misti

La variabile temporale, in questo caso identificata dalla distanza in minuti di ogni misurazione dall'ingresso in terapia intensiva del paziente, svolge un ruolo fondamentale nell'utilizzo di un modello ad effetti misti. La covariata `distanza.minuti` cattura le variazioni sistematiche e l'evoluzione dei dati nel corso del tempo, con l'obiettivo di comprendere come i soggetti modificano la propria situazione clinica nel corso delle osservazioni ripetute.

Inoltre, il tempo è considerato come indicatore della severità della condizione clinica del paziente. Ad esempio se un soggetto presenta elevati valori associati alla `distanza.minuti`, ciò suggerisce una permanenza nel reparto intensivo più lunga, il che può essere legato ad una condizione maggiormente critica. Quanto più il valore della variabile `distanza.minuti` è alto, tanto maggiore è la probabilità che il paziente abbia affrontato una situazione di salute più complessa o grave, richiedendo un periodo di monitoraggio e assistenza più prolungato nella struttura intensiva.

Il tempo in un'analisi longitudinale può essere codificato in modo differente alle necessità e agli obiettivi dell'analisi. In questo caso è stato proposto in forma lineare e mediante una regressione polinomiale a tratti (piecewise polynomials regression o regressione splines). In ultimo si è scelto di proporre anche un modello che non contempli la variabile temporale, per comprendere se vi potessero essere delle differenze significative in base alla presenza o meno di `distanza.minuti`.

La regressione splines è una tecnica statistica utilizzata per modellare in modo flessibile le relazioni non lineari tra una variabile indipendente e una variabile dipendente in un modello di regressione. La funzione `splines::bs` viene utilizzata per creare una rappresentazione di `distanza.minuti` sotto forma di splines cubica. Questo consente di catturare eventuali

andamenti non lineari tra la variabile temporale e la variabile dipendente PaO_2 .

La regressione è basata su funzioni spline formate da polinomi di grado n definiti in specifici intervalli delimitati da nodi e che si legano con continuità nei punti d'incontro. Questo garantisce la fluidità della curva di spline e permette di evitare discontinuità o salti nelle stime. L'uso di splines nella regressione permette di ottenere una maggiore flessibilità nel modellare la forma della curva di relazione, consentendo di adattarsi meglio ai dati senza dover assumere una forma funzionale specifica a priori.

Nella figura sotto riportata, emergono c_1, c_2, c_3 , i 3 nodi della regressione splines in analisi.

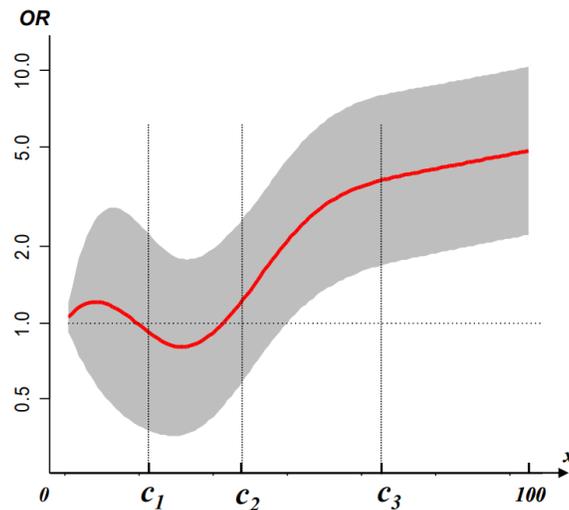


Figura 6: Regressione splines cubica

I risultati della regressione splines verranno confrontati con il modello che considera il tempo linearmente per valutare eventuali differenze significative nelle capacità predittive.

4.1.3 Confronto costruzione modello LMER vs. LME

Nella stima di modelli lineari misti esistono due funzioni ampiamente utilizzate nel linguaggio R: `lme` del pacchetto `nlme` e `lmer` (pacchetto `lme4`). Queste due funzioni implementano metodologie leggermente diverse per la stima dei parametri.

La principale differenza tra le due funzioni riguarda il metodo di stima utilizzato. La funzione `lme` utilizza il metodo della massima verosimiglianza (maximum likelihood) per stimare i parametri del modello lineare misto. Questo metodo si pone l'obiettivo di individuare i valori dei parametri che massimizzano la probabilità di ottenere i dati osservati sotto l'assunzione del modello. D'altra parte, la funzione `lmer` utilizza il metodo dei minimi quadrati ridotti (restricted maximum likelihood, REML) per stimare i parametri. Il REML è una versione ridotta della massima verosimiglianza, in cui i parametri fissi vengono stimati condizionatamente alla varianza casuale. Questo approccio si concentra sulla stima della varianza casuale e rimuove il bias introdotto dalla stima dei parametri fissi, producendo stime che possono risultare meno distorte.

Riguardo le funzioni di ottimizzazione utilizzate per massimizzare la verosimiglianza, la funzione `lme` utilizza algoritmi quali il metodo di Newton-Raphson [46] o il metodo di Fisher-Scoring. Questi algoritmi cercano di trovare il punto di massimo della funzione di verosimiglianza, corrispondente alle stime dei parametri del modello. La funzione `lmer`, invece, utilizza un algoritmo basato sulla scomposizione della varianza totale in varianza fissa e casuale. Tale approccio tiene conto della struttura gerarchica dei dati e dei diversi livelli di variabilità.

La scelta tra `lme` e `lmer` dipende dalle specifiche del problema di ricerca e dalla natura dei dati analizzati. Spesso, se il focus dell'analisi è sulla stima dei parametri fissi del modello e si vuole evitare il bias introdotto dalla varianza casuale, allora si può ricorrere a `lmer` con il metodo REML che potrebbe portare ad un risultato maggiormente appropriato. Se l'obiettivo

è ottenere stime dei parametri che massimizzino la verosimiglianza completa, lme potrebbe essere preferibile.

Nel contesto in questione, non è chiaramente specificato quale delle due funzioni sia la migliore da utilizzare. È ragionevole presumere che le differenze nelle stime calcolate siano molto piccole, e per lo scopo previsivo del progetto non vi sia alcuna differenza significativa nel scegliere una funzione rispetto all'altra.

Nel capitolo dei risultati, verrà presentato un estratto in cui le stime ottenute tramite le due diverse funzioni verranno confrontate per un'analisi comparativa.

4.1.4 Modello logistico

In aggiunta al modello lineare, che fornisce una stima del valore di pressione parziale di ossigeno, è proposto anche un modello logistico.

Esso si pone un obiettivo meno ambizioso, ma particolarmente interessante dal punto di vista medico, perché il suo scopo è quello di identificare un paziente con una situazione critica al fine di offrire celermente un trattamento adeguato.

Per svolgere questa analisi l'outcome deve essere categorizzato, definendo una soglia di criticità. La letteratura [11] [47] riporta come critico un paziente con un valore di PaO_2 inferiore a 60 mmHg.

Il modello logistico risponde a problemi di classificazione, assegnando ad una delle due classi (nel caso in esame paziente critico o meno) ogni individuo. Per fare ciò, è necessario stabilire una regola decisionale per la quale, secondo le informazioni dell'individuo, questi venga assegnato ad una delle classi. Il modello di regressione logistica tra q fattori X_1, \dots, X_q ed un output Y risulta

$$Y \sim B(1, p), \quad g(p) = a_1 X_1 + \dots + a_q X_q + b. \quad (12)$$

Nella regressione logistica l'output Y è dicotomico: 0 o 1, mentre i predittori assumono valori reali generici, come nella regressione lineare multipla tradizionale. Si suppone che, dati i valori dei predittori, l'output Y sia casuale ma con legge univocamente determinata dai predittori.

La variabile di risposta è una variabile aleatoria di Bernoulli e dipende da una combinazione lineare dei predittori. Questa combinazione lineare è chiamata "predittore lineare" (η) ed è

espressa come:

$$\eta = a_1x_1 + \dots + a_px_p + b \quad (13)$$

Il parametro di probabilità (p) che rappresenta la probabilità di successo è influenzato da η attraverso una funzione che garantisce che p rimanga nell'intervallo $[0, 1]$. La relazione tra p e η non è semplicemente $p = \eta$, ma viene trasformata usando una funzione logistica per garantire che p sia una probabilità valida. Si adotta il modello di cui sopra dove g è una funzione definita in $[0, 1]$ a valori reali, invertibile, in modo che sia

$$p = g^{-1}(a_1x_1 + \dots + a_px_p + b). \quad (14)$$

Una scelta molto comune è la funzione detta logit:

$$g(p) = \log\left(\frac{p}{1-p}\right). \quad (15)$$

Per $p \rightarrow 0$ essa tende a $-\infty$, mentre per $p \rightarrow 1$ tende a $+\infty$, ed è strettamente crescente, oltre che regolare. La sua funzione inversa è (Figura 7):

$$p = g^{-1}(\eta) = \frac{\exp(\eta)}{1 + \exp(\eta)}. \quad (16)$$

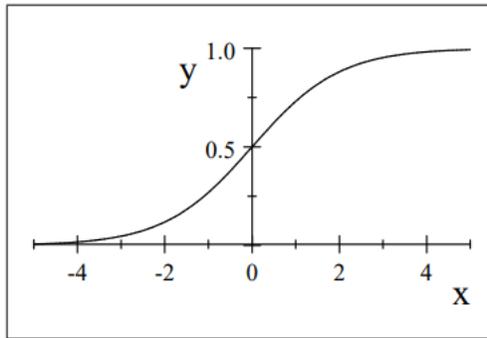


Figura 7: Funzione logistica

In definitiva, il modello è $Y \sim B(1, p)$ con $p = \frac{\exp(\eta)}{1 + \exp(\eta)}$, dove $\eta = a_1 x_1 + \dots + a_p x_p + b$.

Se la probabilità p associata ad un individuo è molto elevata, si ha un certo livello di sicurezza nell'assegnare il soggetto alla classe 1 ($Y = 1$), mentre se è molto bassa, si suppone sia $Y = 0$. In base all'obiettivo specifico, è possibile modificare la soglia di classificazione: si può porre la soglia a 0.5, ma non è detto che tale valore minimizzi falsi positivi e falsi negativi. È lecito stabilire $p > \frac{1}{2}$ o $p < \frac{1}{2}$ in modo arbitrario per cercare di minimizzare i falsi positivi (favorendo la specificità) oppure rendendo minimo il numero di falsi negativi (maggior sensibilità).

Per stimare i coefficienti si ricorre al metodo della massima della massima verosimiglianza. Noti i valori x_1, \dots, x_p dei predittori di un individuo la funzione di verosimiglianza si può scrivere:

$$P(Y = y) = p^y (1 - p)^{1-y} \quad (17)$$

Generalizzando per più soggetti:

$$P(Y^{(1)} = y^{(1)}, \dots, Y^{(n)} = y^{(n)}) = \prod_{i=1}^n p^{(i)y^{(i)}} (1 - p^{(i)})^{1-y^{(i)}} \quad (18)$$

dove $p^{(i)} = g^{-1}(\eta^{(i)})$ e $\eta^{(i)} = a_1 x_1^{(i)} + \dots + a_p x_p^{(i)} + b$.

Questa è la verosimiglianza del campione sperimentale, funzione di a_1, \dots, a_p, b . Il metodo di massima verosimiglianza consiste nel cercare i valori di a_1, \dots, a_p, b che la rendono massima. Pertanto, occorre massimizzare la funzione:

$$\mathcal{L}(a_1, \dots, a_p, b) = \sum_{i=1}^n y^{(i)} \log p^{(i)} + (1 - y^{(i)}) \log(1 - p^{(i)}) \quad (19)$$

Il modello logistico appare come un ottimo strumento per affrontare la sfida clinica di prevedere la condizione critica di un soggetto. È fondamentale porre grande attenzione nella scelta della soglia di classificazione per definire un paziente critico poiché questa soglia determina il punto di separazione tra le due classi, riflettendosi su sensibilità e specificità del modello. La sensibilità rappresenta la capacità del modello di identificare correttamente i soggetti veramente critici (veri positivi), evitando falsi negativi. D'altro canto, la specificità riflette la capacità del modello di riconoscere correttamente i soggetti veramente non critici (veri negativi), evitando falsi positivi.

L'aspetto più complesso sta nel trovare un compromesso nella scelta della soglia, minimizzando sia i falsi positivi che i falsi negativi, sulla base delle esigenze cliniche. Se l'obiettivo è evitare di classificare erroneamente soggetti critici come non critici (falsi negativi), si potrebbe valutare una soglia più bassa, in modo da aumentare la sensibilità del modello. Ciò garantirebbe una più semplice identificazione dei soggetti veramente critici (veri positivi), ma potrebbe aumentare anche il numero di falsi positivi. D'altra parte, se lo scopo principale è minimizzare i falsi positivi, si potrebbe scegliere una soglia più alta, aumentando la specificità del modello. Si avrebbe una riduzione di falsi positivi, ma potrebbe diminuire la capacità del modello di individuare tutti i soggetti veramente critici, ovvero la metrica di sensibilità.

4.1.5 Calibration belt

L'efficacia di un modello logistico è soggetta alla definizione arbitraria della soglia di probabilità utilizzata per separare le classi. È possibile migliorare questa valutazione considerando il valore effettivo di probabilità che il modello attribuisce a ciascuna osservazione, in quanto tale valore è in grado di rivelare la capacità predittiva dell'evento binario.

In questo contesto, la *calibration belt* costituisce uno strumento di estrema utilità per verificare quanto le previsioni riflettano le osservazioni reali, contribuendo così a determinare la qualità delle previsioni del modello. In particolare, la *calibration belt* viene utilizzata per valutare la calibrazione del modello, ossia la sua capacità di stimare in modo accurato le probabilità di eventi binari. Essa viene costruita tracciando una curva che rappresenta la relazione tra le probabilità previste dal modello e le probabilità effettive di osservare l'evento di interesse. Per esempio, se il modello prevede una probabilità del 70% per un evento e l'evento si verifica effettivamente nel 60% dei casi, ciò indica la distanza tra previsione e realtà. La cintura di calibrazione può includere anche fasce di confidenza per rappresentare l'incertezza nelle previsioni del modello. Queste fasce di confidenza forniscono una stima della variabilità tra le previsioni del modello e le osservazioni effettive. Il gruppo di ricerca dell'istituto di Ricerche Farmacologiche Mario Negri ha realizzato un pacchetto R (*givitir*) che raccoglie diverse funzioni per valutare la calibrazione di modelli binari di previsione, come i modelli logistici, utilizzando la "*GiViTI calibration belt*".

La funzione prende il nome dal "Gruppo Italiano per la valutazione degli interventi in Terapia Intensiva" (GiViTI) che ha lavorato al progetto. Questo approccio mira a valutare la calibrazione dei modelli, ossia la loro capacità di stimare in modo affidabile i tassi di evento. Si consideri che un modello è ben calibrato se le probabilità previste corrispondono accuratamente alle proporzioni osservate della risposta. Il pacchetto R viene utilizzato per migliorare la precisione nella rappresentazione dei risultati di un modello logistico.

4.2 Validazione del modello

La validazione di un modello è un processo estremamente importante finalizzato a valutare l'efficacia e la capacità predittiva del modello stesso. La tecnica di validazione utilizzata in questo progetto è la *hold-out validation* che prevede la suddivisione del dataset completo in due parti distinte: un set di addestramento e un set di validazione. Il set di addestramento è utilizzato per addestrare il modello (in questo caso 80% delle osservazioni disponibili), mentre il set di validazione, il restante 20%, viene utilizzato per valutare le prestazioni del modello. A seguito dell'addestramento del modello sui dati di training, viene valutato su nuovi dati, ovvero sul set di validazione.

Mantenere il set di validazione separato dal set di addestramento è fondamentale per evitare l'overfitting, ovvero l'eccessivo adattamento del modello ai dati di addestramento con la conseguente perdita della capacità di generalizzare su nuovi dati.

Una pratica comune in questo tipo di validazione sta nello scegliere casualmente quali osservazioni inserire nel validation set. Ciò assicura che entrambi i dataset (training e validation) siano rappresentativi del dataset completo e che le prestazioni del modello siano valutate su dati diversi da quelli utilizzati per l'addestramento.

Per valutare le performance di un modello, si ricorre a diverse metriche che forniscono una valutazione delle sue prestazioni.

4.2.1 Definizione dell'errore: Median Absolute Error

Il Median Absolute Error è una metrica che può essere utilizzata per quantificare le prestazioni di un modello. Questa metrica ha un'interpretazione leggermente più complessa rispetto al Mean Absolute Error. Tuttavia, il principale vantaggio nell'utilizzare la mediana d'errore è la sua capacità di non essere sensibile ad eventuali outlier a differenza di Mean Absolute Error o Mean Squared Error. Il MedAE calcola la mediana delle differenze tra le predizioni del modello e i valori reali, ignorando il segno. Questo evita che le differenze positive e negative si annullino tra loro: considerando il segno, l'errore calcolato potrebbe risultare notevolmente inferiore alla vera differenza tra modello e dati.

È espresso come:

$$\text{Median Absolute Error} = \text{Mediana}(|y_i - \hat{y}_i|) \quad (20)$$

dove y_i rappresenta le osservazioni reali e \hat{y}_i rappresenta le predizioni del modello. Questo valore fornisce una stima robusta della dispersione delle predizioni rispetto ai valori reali ed è particolarmente utile quando i dati contengono valori anomali che potrebbero influenzare eccessivamente le metriche di errore.

L'errore deve essere accompagnato da una stima della variabilità, al fine di fornire un'indicazione più precisa sulla capacità del modello di adattarsi ai dati e di effettuare delle previsioni sul dataset di validazione. A tale scopo verrà sempre fornita la deviazione standard, ma è possibile ottenere anche una stima intervallare dell'errore, in particolare riportando i quartili della distribuzione i quali offrono un maggiore dettaglio sulla distribuzione della mediana d'errore. Nel caso in cui primo e terzo quartili presentino valori vicini alla mediana d'errore, si può sostenere che non vi sia dispersione attorno al valore dell'indice di posizione. Mediante un boxplot è possibile rappresentare graficamente la situazione descritta.

4.2.2 Rappresentazioni grafiche dei residui

La rappresentazione grafica è un elemento molto utile per valutare la bontà dei modelli statistici e la loro capacità di previsione. Tra i grafici più comuni utilizzati per questa valutazione si hanno:

- *Valori Predetti vs. Valori Osservati*: questo tipo di grafico fornisce una visualizzazione diretta della capacità del modello di predire i valori reali. Gli osservati sono posizionati sull'asse delle ascisse, rappresentando i dati osservati, mentre i valori stimati dal modello sono collocati sull'asse delle ordinate. Un modello altamente accurato produrrà una dispersione dei punti intorno a una linea diagonale, che rappresenta una previsione perfetta. La deviazione da questa linea indica la differenza tra le previsioni del modello e i dati reali.
- *Valori Osservati vs. (Osservati - Stimati)*: tale grafico è uno strumento estremamente valido per valutare la capacità del modello di catturare la variazione nei dati osservati. Sul grafico, sull'asse delle ascisse vengono riportati i dati osservati, mentre sull'asse delle ordinate si ha la differenza tra i dati osservati e quelli stimati dal modello. Questo grafico è utile per identificare se il modello tende a sottostimare o sovrastimare i dati e se esiste un bias sistematico nelle previsioni. Se il modello è accurato, i punti dovrebbero distribuirsi intorno alla retta $y=0$.
- *Valori Predetti vs. Errore (Osservato - Stimato)*: permette di evidenziare gli errori residui del modello, cioè la differenza tra i dati osservati e quelli stimati. Sull'asse delle ascisse sono posizionati i valori predetti dal modello, mentre sull'asse delle ordinate sono rappresentati gli errori residui. È possibile identificare eventuali pattern o tendenze negli errori residui. Se gli errori residui sono casuali e distribuiti intorno allo zero, il modello sta catturando efficacemente la variabilità nei dati se, invece, emergono schemi nei residui potrebbe essere necessario apportare miglioramenti al modello.

5. RISULTATI

5.1 Grafici andamenti variabile risposta

A seguire si presentano i risultati dell'analisi delle traiettorie e della distribuzione della variabile risposta in diversi pazienti. La rappresentazione grafica permette di esplorare le dinamiche e gli andamenti della pressione di ossigeno nel sangue nel tempo. Questa analisi grafica preliminare rappresenta un passo importante per comprendere le relazioni tra le variabili coinvolte e identificare eventuali pattern, trend o cambiamenti significativi.

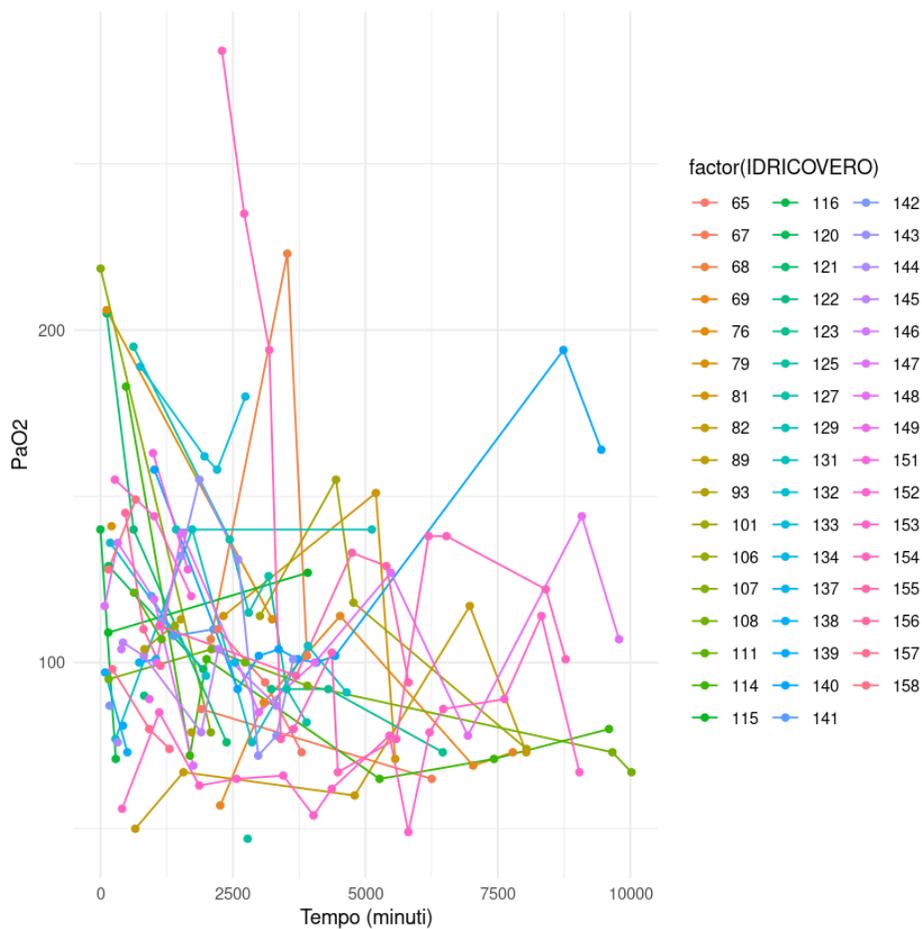


Figura 8: Andamento PaO_2 per un sottogruppo casuale di pazienti

La figura 8 riporta lo spaghetti plot⁸ di un sottogruppo di 50 pazienti selezionato in modo casuale nel corso della prima settimana in terapia intensiva. Si noti come non sia possibile identificare in modo univoco un andamento della variabile risposta del tempo a causa dell'elevata variabilità tra soggetti. Nonostante ciò è opportuno considerare nel modello il tempo, in quanto esso fornisce una scansione temporale e, dal punto di vista clinico, è una covariata utile per indicare la severità della condizione del paziente.

L'analisi della distribuzione di frequenza della variabile dipendente, rappresentata nel grafico sottostante (Figura 9), mette in evidenza la complessità nell'identificare una forma funzionale chiara per la variabile PaO_2 . La distribuzione non mostra una simmetria tipica della distribuzione normale e, sebbene vi sia un andamento simile ad essa, sovrapponendo una curva gaussiana con media e varianza calcolate dai dati empirici, non si ottiene una perfetta coincidenza. Questo suggerisce che la variabile di risposta potrebbe essere influenzata da diverse dinamiche e fattori non lineari, rendendo difficile la sua approssimazione con un modello statistico semplice.

⁸Il termine spaghetti plot si riferisce al grafico per la visualizzazione delle traiettorie per singoli soggetti. Il nome è dovuto alla forma che il grafico assume

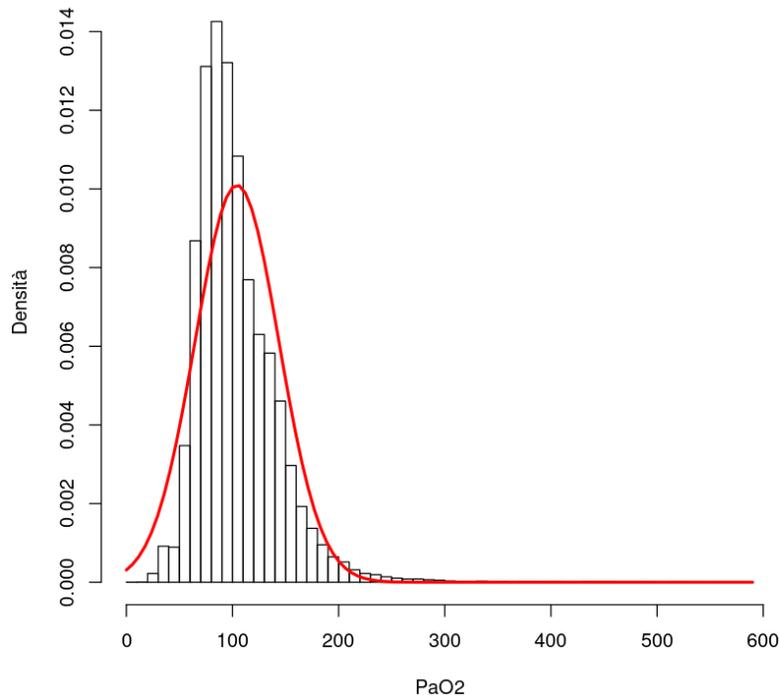


Figura 9: Distribuzione PaO_2

Di seguito si riporta il grafico di un paziente selezionato casualmente (Figura 10), ricoverato presso il centro ospedaliero 4. Non è possibile conoscere dai dati disponibili l'esito del periodo in terapia intensiva, se si sia concluso con un miglioramento della condizione clinica oppure abbia avuto esito negativo. Nei 16 giorni di ricovero intensivo il paziente ha 26 misurazioni della pressione parziale di ossigeno nel sangue.

Si rappresenta sia la variabile risposta che tutte le covariate per provare a comprendere se vi possa essere, per questo soggetto, una certa relazione tra l'andamento della pressione e le variabili indipendenti.

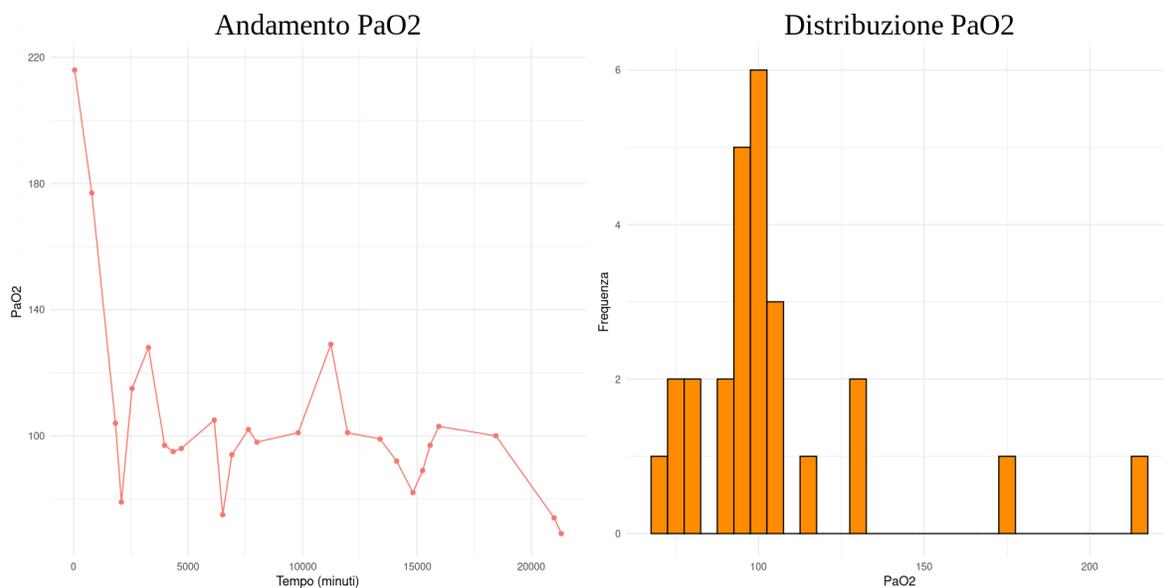


Figura 10: Andamento e distribuzione di PaO_2 per un paziente del Centro 4

Per quanto riguarda la variabile risposta, il paziente in analisi all'ingresso in terapia intensiva presenta un elevato valore di PaO_2 che nel corso del tempo diminuisce e si stabilizza all'incirca in valori attorno a 100 mmHg, con un decremento sensibile dopo circa 14 giorni. La distribuzione di frequenza conferma che il soggetto abbia per lo più misurazioni con valori simili che si attestano attorno ai 100 mmHg, ma presenta anche rilevazioni più distanti con valori prossimi a valori più bassi (~ 80 mmHg) e superiori alla soglia critica di 150 mmHg [48].

Seguono i grafici per descrivere gli andamenti delle covariate continue di questo paziente: sarebbe auspicabile individuare un pattern generalizzabile poi a tutti i pazienti. I grafici degli andamenti delle variabili indipendenti (Figura 11, 12, 13) mostrano il legame tra con la pressione parziale di ossigeno nel sangue. Una diminuzione di PaO_2 è associata ad un aumento nel tempo della frequenza cardiaca, in linea con quanto esposto al paragrafo 3.1.1. Si ha una diminuzione dei valori di emoglobina e valori lattati e ph crescenti con il trascorrere

del tempo dall'ingresso in terapia intensiva. Nonostante ciò, è possibile sostenere che data l'elevata variabilità tra pazienti, risulta difficile individuare un pattern generalizzabile. Questa rappresentazione grafica è comunque utile per mostrare l'evoluzione clinica del paziente.

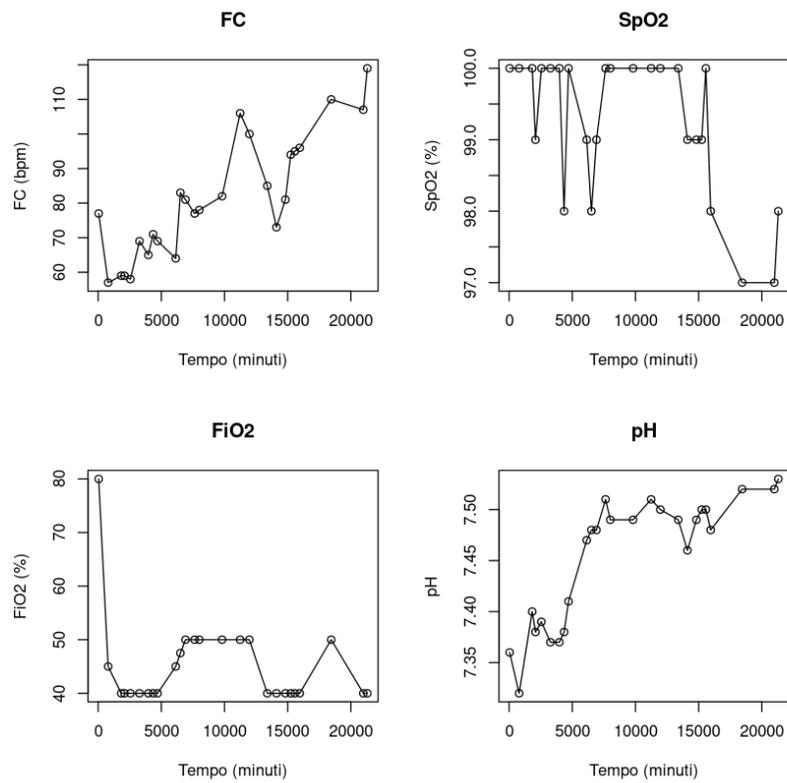


Figura 11: Andamento FC, SpO_2 , FiO_2 , pH per il paziente in esame

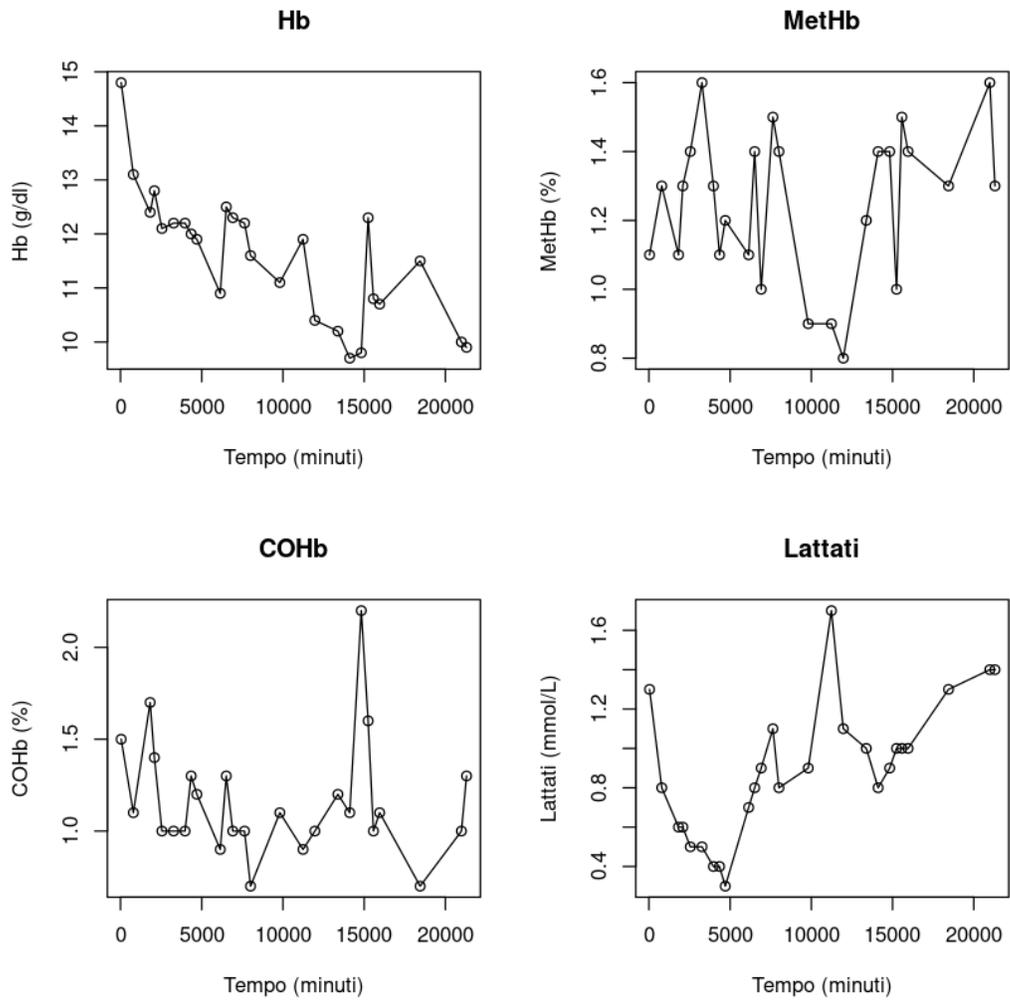


Figura 12: Andamento Hb, MetHb, COHb, lattati per il paziente in esame

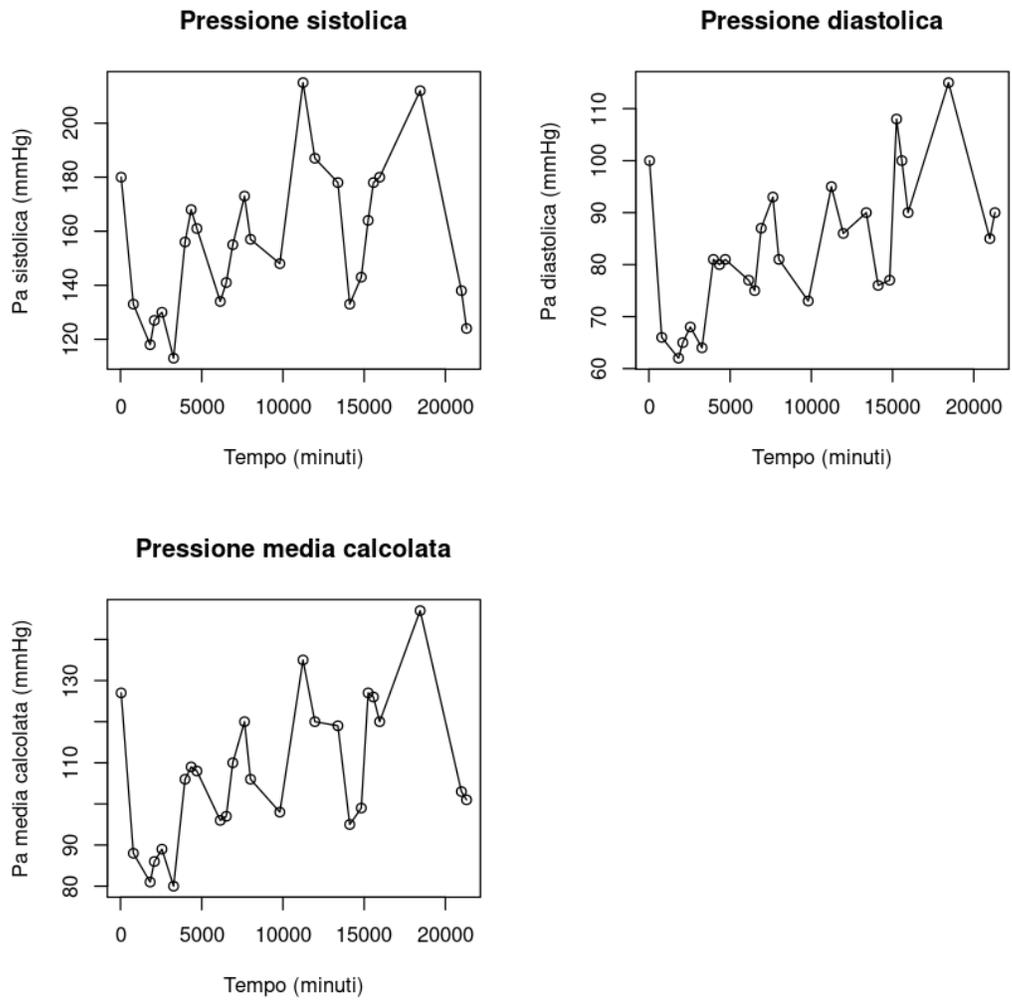


Figura 13: Andamento pressione sistolica, diastolica e calcolata per il paziente in esame

5.2 Risultati descrittivi dei modelli

Il modello statistico di maggiore rilevanza in questo progetto è il modello lineare ad effetti misti. Prima di utilizzarlo come modello predittivo, è opportuno valutare i risultati descrittivi per comprendere meglio l'effetto ed il ruolo delle variabili indipendenti in relazione alla variabile risposta PaO_2 . Come descritto precedentemente (cfr. 4.1.2) sono proposti 3 differenti modelli misti, ognuno dei quali considera diversamente la covariata legata al tempo: viene valutata linearmente, mediante una regressione splines cubica e, in ultima analisi, viene esclusa dalle variabili considerate.

	Value	Sd Error	DF	t-value	p-value	
(Intercetta)	-284.0755	18.1445	47762	-15.6563	<0.0001	*
Età (anni)	-0.1265	0.0104	47762	-12.1583	<0.0001	*
Genere (Uomini)	-2.8856	0.3343	47762	-8.6320	<0.0001	*
SpO2 (%)	4.5071	0.0562	47762	80.2208	<0.0001	*
Frequenza cardiaca (bpm)	-0.1013	0.0083	47762	-12.1976	<0.0001	*
MetHb (%)	4.2705	0.5243	47762	8.1454	<0.0001	*
Hb (g/dl)	-0.3790	0.0932	47762	-4.0671	<0.0001	*
COHb (%)	-1.7499	0.1677	47762	-10.4372	<0.0001	*
Pressione sistolica (mmHg)	-0.0404	0.0262	47762	-1.5448	0.1224	
Pressione diastolica (mmHg)	0.1232	0.0522	47762	2.3602	0.0183	*
Pressione media calcolata (mmHg)	-0.0444	0.0752	47762	-0.5898	0.5553	
FiO2 (%)	0.3353	0.0110	47762	30.4860	<0.0001	*
pH	-10.9635	2.3454	47762	-4.6745	<0.0001	*
Lattati (mmol/L)	1.6455	0.1413	47762	11.6419	<0.0001	*
Distanza dall'ingresso in TI (gg.)	-0.2133	0.0151	47762	-14.1232	<0.0001	*
Lag 1° ordine PaO2 (mmHg)	0.1820	0.0042	47762	43.0034	<0.0001	*
Lag 2° ordine PaO2 (mmHg)	0.1355	0.0042	47762	32.3390	<0.0001	*
Intubato (Si)	3.6536	0.3689	47762	9.9036	<0.0001	*

Tabella 7: Risultati modello con distanza.giorni lineare

In primo luogo è opportuno concentrare l'attenzione sull'intercetta; la significatività (p-value < 0.001) indica che il valore stimato dell'intercetta è statisticamente diverso da zero, ma non appare utile e commentabile nel modello in questione. Infatti, il coefficiente risulta essere pari a -284.0755. Questo rappresenta il valore medio stimato di PaO_2 quando tutte le altre variabili indipendenti sono pari a zero, situazione clinica che non è fisiologicamente possibile. Il coefficiente associato all'età è -0.1265. Indica che con un aumento di un'unità nell'età, è attesa una diminuzione di 0.1265 mmHg di PaO_2 . Gli uomini hanno un valore medio di Y inferiore di 2.8856 mmHg; la differenza tra uomo e donna è significativa dal punto di vista statistico (p-value < 0.001).

Il coefficiente della metemoglobina è 4.2705; di conseguenza un aumento di un'unità percentuale di metemoglobina (MetHb) corrisponde un aumento di 4.2705. La sua significatività suggerisce che la MetHb possa essere un fattore rilevante per spiegare PaO_2 .

La pressione sistolica presenta un coefficiente negativo per cui all'aumentare del valore si avrebbe una diminuzione della pressione parziale di ossigeno nel sangue. Questa diminuzione non ha però significatività statistica (p-value = 0.1224). La frazione di ossigeno inspirata, considerato uno dei parametri clinici più rilevanti, riporta un coefficiente di 0.3353. Un aumento di un'unità percentuale nella FiO_2 , porta un aumento di 0.3353 di pressione di ossigeno nel sangue. Dato il suo valore di significatività (p-value < 0.001) si hanno evidenze per sostenere che la frazione di ossigeno inspirata possa essere un fattore rilevante per spiegare la variabile risposta.

Una delle variabili con maggiore importanza clinica è la distanza (in giorni) dall'ingresso in terapia intensiva. Il coefficiente associato risulta contenuto (-0.2133): una variazione di un'unità temporale (un giorno) porta ad una diminuzione limitata di PaO_2 . Inoltre, poiché la forma funzionale del tempo è stata considerata lineare nel modello, può risultare difficile

descrivere un'associazione più evidente tra la distanza temporale e PaO_2 . Una scelta diversa nella rappresentazione del tempo potrebbe rivelare un'associazione più significativa tra la distanza dall'ingresso in terapia intensiva e la variabile di risposta.

Nel modello sono stati inseriti anche i lag di primo e secondo ordine della variabile risposta. In questo modo si vuole dare rilevanza ai valori che PaO_2 ha assunto nelle due misurazioni precedenti. Il coefficiente associato al lag di primo ordine è 0.182. Un aumento unitario nel valore in mmHg della PaO_2 precedente porta ad un aumento significativo (p-value < 0.001) di 0.182 della pressione parziale di ossigeno. In ultimo viene riportata la variabile binaria "intubato" che definisce la presenza di presidi per la ventilazione meccanica (cfr. 3.2.4). Il coefficiente è pari a 3.6536. Questo valore indica che, rispetto a chi non è intubato, il paziente intubato ha un valore medio di PaO_2 maggiore di 3.6536. La significatività del coefficiente porta a sostenere l'ipotesi per cui questa variabile sia significativa per spiegare le differenze nella pressione parziale.

L'output che segue (Tabella 8) riporta i risultati del modello ad effetti misti in cui è stato inserito la regressione spline. L'attenzione maggiore verrà posta proprio sui coefficienti relativi al tempo.

	Value	Sd Error	DF	t-value	p-value	
(Intercetta)	-334.9483	18.2833	47760	-18.3199	<0.0001	*
Età (anni)	-0.1326	0.0104	47760	-12.7926	<0.0001	*
Genere (Uomini)	-2.5818	0.3334	47760	-7.7436	<0.0001	*
SpO2 (%)	4.4508	0.0561	47760	79.3966	<0.0001	*
Frequenza cardiaca (bpm)	-0.0888	0.0083	47760	-10.6833	<0.0001	*
MetHb (%)	3.8559	0.5229	47760	7.3736	<0.0001	*
Hb (g/dl)	-0.6345	0.0938	47760	-6.7642	<0.0001	*
COHb (%)	-1.7522	0.1671	47760	-10.4891	<0.0001	*
Pressione sistolica (mmHg)	-0.0300	0.0261	47760	-1.1512	0.2497	
Pressione diastolica (mmHg)	0.1255	0.0520	47760	2.4129	0.0158	*
Pressione media calcolata (mmHg)	-0.0522	0.0749	47760	-0.6961	0.4864	
FiO2 (%)	0.3569	0.0110	47760	32.3995	<0.0001	*
pH	-2.7030	2.3766	47760	-1.1374	0.2554	
Lattati (mmol/L)	1.3510	0.1417	47760	9.5374	<0.0001	*
Splines 1 distanza dall'ingresso in TI (gg.)	-42.8518	2.0184	47760	-21.2305	<0.0001	*
Splines 2 distanza dall'ingresso in TI (gg.)	45.9850	4.6944	47760	9.7957	<0.0001	*
Splines 3 distanza dall'ingresso in TI (gg.)	-41.8683	6.6747	47760	-6.2727	<0.0001	*
Lag 1° ordine PaO2 (mmHg)	0.1790	0.0042	47760	42.4224	<0.0001	*
Lag 2° ordine PaO2 (mmHg)	0.1333	0.0042	47760	31.9244	<0.0001	*
Intubato (Si)	4.0916	0.3683	47760	11.1104	<0.0001	*

Tabella 8: Risultati modello con splines(distanza.giorni)

La regressione splines permette di modellare in modo flessibile la relazione tra la distanza dall'ingresso in terapia intensiva e la variabile di risposta. In questo caso si è utilizzato una regressione cubica. Saranno dunque tre i coefficienti proposti dal modello ad effetti misti associati alle splines. È importante sottolineare che questi coefficienti tengono conto anche dell'influenza simultanea delle altre variabili covariate presenti nel modello. I coefficienti positivi o negativi evidenziano come la relazione tra la distanza in minuti dall'accesso in terapia intensiva e PaO_2 possa variare in base alle specifiche trasformazioni quadratiche e

cubiche considerate. Il primo coefficiente splines per il tempo ha un valore di -42.8518. Rappresenta la pendenza stimata per la spline al punto di riferimento (riferimento della distanza dall'ingresso in TI). La sua significatività (p-value < 0.0001) indica che il valore stimato della pendenza per la spline 1 è statisticamente diverso da zero, suggerendo un ruolo significativo nella spiegazione delle variazioni di PaO_2 in relazione alla distanza dall'ingresso in terapia intensiva.

Il coefficiente associato alla spline 2 è 45.985, il quale indica la pendenza stimata per la spline al punto di riferimento. Il segno in questo caso è opposto, quindi in questo intervallo di tempo intermedio si ha un aumento del valore della pressione parziale di ossigeno.

Infine il coefficiente associato all'ultimo intervallo costruito dalla regressione spline risulta essere -41.8683, per cui in questo periodo temporale si ha una diminuzione del valore medio della variabile risposta, simile a quello riportato nel primo intervallo temporale. Per ogni coefficiente della regressione splines è associato un p-value significativo (p-value < 0.0001). L'utilizzo della regressione splines consente di catturare in modo più flessibile le dinamiche di PaO_2 nel tempo, tenendo conto delle variazioni nella distanza dall'ingresso in terapia intensiva e fornendo un'analisi più accurata delle influenze temporali sulla variabile di risposta.

	Value	Sd Error	DF	t-value	p-value	
(Intercetta)	-282.2722	18.1642	47763	-15.5400	<0.0001	*
Età (anni)	-0.1284	0.0104	47763	-12.3247	<0.0001	*
Genere (Uomini)	-3.1111	0.3346	47763	-9.2980	<0.0001	*
SpO2 (%)	4.5293	0.0563	47763	80.4852	<0.0001	*
Frequenza cardiaca (bpm)	-0.1125	0.0083	47763	-13.5859	<0.0001	*
MetHb (%)	4.4772	0.5246	47763	8.5341	<0.0001	*
Hb (g/dl)	-0.0791	0.0909	47763	-0.8698	0.3844	
COHb (%)	-1.8023	0.1679	47763	-10.7323	<0.0001	*
Pressione sistolica (mmHg)	-0.0408	0.0262	47763	-1.5543	0.1201	
Pressione diastolica (mmHg)	0.1112	0.0523	47763	2.1254	0.0336	*
Pressione media calcolata (mmHg)	-0.0360	0.0754	47763	-0.4775	0.633	
FiO2 (%)	0.3346	0.0110	47763	30.3698	<0.0001	*
pH	-12.1651	2.3487	47763	-5.1795	<0.0001	*
Lattati (mmol/L)	1.8181	0.1411	47763	12.8855	<0.0001	*
Lag 1° ordine PaO2 (mmHg)	0.1845	0.0042	47763	43.5288	<0.0001	*
Lag 2° ordine PaO2 (mmHg)	0.1379	0.0042	47763	32.8763	<0.0001	*
Intubato (Si)	3.3362	0.3690	47763	9.0417	<0.0001	*

Tabella 9: Risultati modello senza distanza.giorni

Il modello ad effetti misti che non presenta la variabile distanza.giorni (Tabella 9) evidenzia coefficienti molto simili ai modelli precedenti. Nel dettaglio si noti, ad esempio, come per la covariata “intubato” il coefficiente sia pari a 3.3362 e risulti significativo. Nel primo modello proposto (Tabella 7) il coefficiente associato alla presenza di una ventilazione non spontanea è pari a 3.6536. I due valori sono molto vicini e ciò si riscontra anche per gli altri coefficienti.

Nella parte teorica (cfr. 4.1.3), si sono analizzati due pacchetti del software R a cui si ricorre per stimare i coefficienti di un modello. Dal punto di vista teorico vi sono delle differenze nelle stime poiché sono differenti le modalità della stima stessa. Per comprendere meglio questa situazione, è necessario valutare se queste differenze siano significative. Ciò sarà di supporto per decidere se i due modelli siano effettivamente equivalenti o se uno dei due dovrebbe essere preferito rispetto all'altro.

	Value	Sd Error	t-value	p-value	
(Intercetta)	-284.0755	18.1445	-15.6563	<0.0001	*
Età	-0.1265	0.0104	-12.1583	<0.0001	*
Genere (Uomini)	-2.8856	0.3343	-8.6320	<0.0001	*
SpO2	4.5071	0.0562	80.2208	<0.0001	*
Frequenza cardiaca	-0.1013	0.0083	-12.1976	<0.0001	*
MetHb	4.2705	0.5243	8.1454	<0.0001	*
Hb	-0.3790	0.0932	-4.0671	<0.0001	*
COHb	-1.7499	0.1677	-10.4372	<0.0001	*
Pressione sistolica	-0.0404	0.0262	-1.5448	0.1224	
Pressione diastolica	0.1232	0.0522	2.3602	0.0183	*
Pressione media calcolata	-0.0444	0.0752	-0.5898	0.5553	
FiO2	0.3353	0.0110	30.4860	<0.0001	*
pH	-10.9635	2.3454	-4.6745	<0.0001	*
Lattati	1.6455	0.1413	11.6419	<0.0001	*
Distanza dall'ingresso in TI (min.)	-0.0001	0.0000	-14.1232	<0.0001	*
Lag 1° ordine PaO2	0.1820	0.0042	43.0034	<0.0001	*
Lag 2° ordine PaO2	0.1355	0.0042	32.3390	<0.0001	*
Intubato (Si)	3.6536	0.3689	9.9036	<0.0001	*

Tabella 10: Risultati modello stimato con la funzione LMER

La tabella precedente (Tabella 10) mostra i risultati per il modello che considera la variabile temporale lineare stimato mediante la funzione LMER. Confrontando tali risultati con il corrispondente modello stimato con LME (Tabella 9) è possibile notare come non vi sia alcuna differenza alla quarta cifra decimale per quanto riguarda i coefficienti e le significatività corrispondenti.

Si può dunque sostenere che non vi sia una scelta consigliata tra le due tipologie di funzione ma nella situazione in esame, riportano risultati identici. Per questo motivo nel prosieguo si farà riferimento sempre ai modelli stimati con LME, totalmente analoghi a quelli prodotti con il pacchetto LMER.

5.3 Risultati sul test di validazione

Dopo l'analisi descrittiva, va posta maggiore attenzione sull'efficacia e l'accuratezza dei modelli statistici adottati. Mediante il dataset di validazione (cfr. 4.2) è possibile verificare come il modello si comporta nel prevedere la variabile di risposta su dati indipendenti e non utilizzati durante la fase di addestramento.

Le performance dei modelli sono state valutate utilizzando la metrica Median Absolute Error che misura la mediana degli errori assoluti tra le previsioni e i valori reali della variabile risposta. Il valore di MedAE pari a 16.69 mmHg (Tabella 11) indica che il 50% degli errori assoluti è inferiore a 16.69 mmHg.

La stima puntuale dell'errore mostra un risultato relativamente buono, discostandosi di poco più di 16 mmHg dal valore reale, tuttavia, l'obiettivo iniziale di minimizzare questo errore per ottenere risultati altamente accurati appare ancora distante. Sebbene la stima puntuale possa essere considerata in parte accettabile, essa perde di valenza se associata alla corrispondente deviazione standard, infatti, si hanno valori dell'indice di variabilità molto elevati. Ne deriva che il modello non è in grado di adattarsi bene ai dati e non permette di attuare una previsione accurata della pressione parziale di ossigeno.

Si noti, peraltro, come i risultati dei tre modelli proposti, i quali differiscono per la gestione della covariata associata al tempo, abbiano risultati pressoché identici. La variabile legata al tempo non sembra avere un impatto significativo sulla previsione di PaO_2 . In seguito verranno riportati solo i risultati ottenuti con il modello con la regressione splines del tempo.

Modello	MedAE	Dev. std
Modello distanza giorni lineare	16.6917	24.5169
Modello splines(distanza giorni)	16.6609	24.4124
Modello senza distanza giorni	16.5581	24.6234

Tabella 11: Median Absolute Error e relativa deviazione standard

I risultati sono poi considerati per ogni centro ospedaliero per valutare eventuali differenze significative tra le singole strutture (Tabella 12).

Centro ospedaliero	MedAE	Dev. std
1	14.6240	24.9676
2	16.9052	22.7288
3	13.2461	29.2682
4	18.4556	26.4777
5	13.7266	13.4988

Tabella 12: Median Absolute Error e relativa deviazione standard per centro ospedaliero

L'analisi stratificata per centro ospedaliero evidenzia valori puntuali inferiori per due centri specifici. Questo significa che la previsione del modello è più accurata in questi due centri, con un errore mediano assoluto di circa 13 mmHg. È importante notare che la variabilità tra i due centri è molto diversa. Il centro ospedaliero 3 ha una deviazione standard molto elevata, pari a 29.26, indicando una maggiore dispersione dei dati intorno alla stima puntuale. Al contrario, l'ospedale identificato con il numero 5 ha una variabilità più contenuta, con una deviazione standard inferiore. La differenza nella variabilità tra i centri potrebbe essere dovuta a diversi fattori. In particolare, la minore variabilità tra i soggetti presenti in alcune

strutture ospedaliere potrebbe contribuire alla maggiore precisione del modello nei centri con una deviazione standard più bassa. È possibile valutare anche graficamente come i risultati abbiano una elevata variabilità. Viene riportato come primo grafico (Figura 14) il boxplot della distanza tra il valore osservato ed il valore stimato.

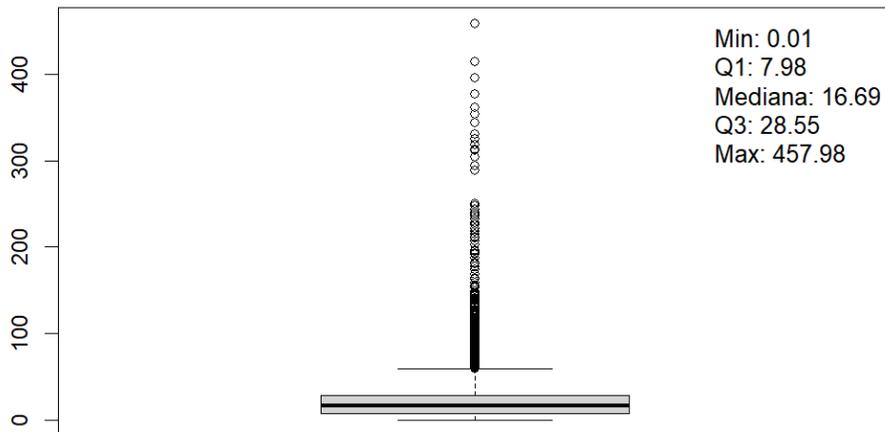


Figura 14: Boxplot Median Absolute Error

Nel boxplot sopra rappresentato emerge chiaramente la presenza di numerosi valori anomali che corrispondono a previsioni estremamente errate. Questi valori anomali contribuiscono a comprimere la rappresentazione grafica dei dati intorno alla mediana.

Nonostante questa compressione grafica, la dispersione dei dati attorno alla mediana è in realtà notevole, come evidenziato dalla notevole distanza interquartile tra il primo quartile e il terzo quartile. Questo aspetto conferma quanto precedentemente osservato riguardo alla deviazione standard, indicando che vi sono considerevoli variazioni nelle previsioni.

Va sottolineato che il valore massimo estremamente elevato evidenzia situazioni in cui il

modello di previsione non è in grado di fornire stime affidabili per il valore osservato di PaO_2 , per cui il modello non è in grado di gestire con precisione tali situazioni con rilevanti risvolti clinici.

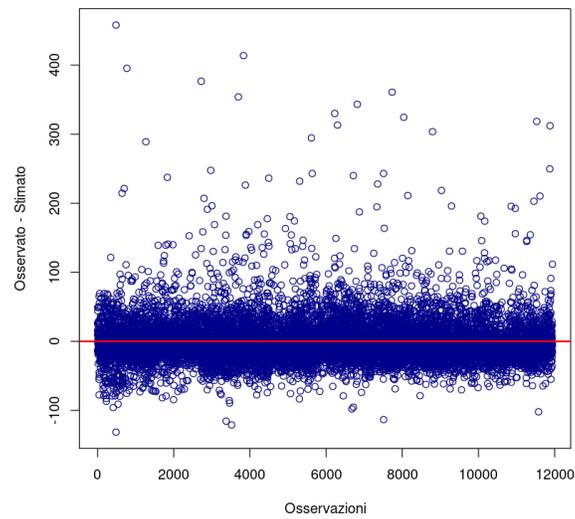


Figura 15: Median Absolute Error

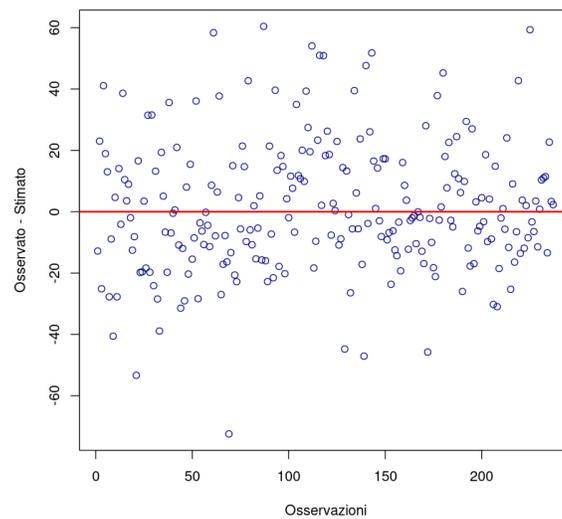


Figura 16: Median Absolute Error per il centro 5

Dai grafici sopra riportati dei residui (Figura 15, 16) non emergono pattern particolari, ma si nota l'elevata variabilità delle previsioni ottenute. Vi sono residui maggiori di 100, i quali indicano una differenza tra il valore stimato ed il valore reale estremamente elevata.

I risultati assumono una maggiore stabilità se si considera il solo centro 5 (Figura 16): in questo caso i residui sono si raccolgono maggiormente attorno alla retta orizzontale $y = 0$.

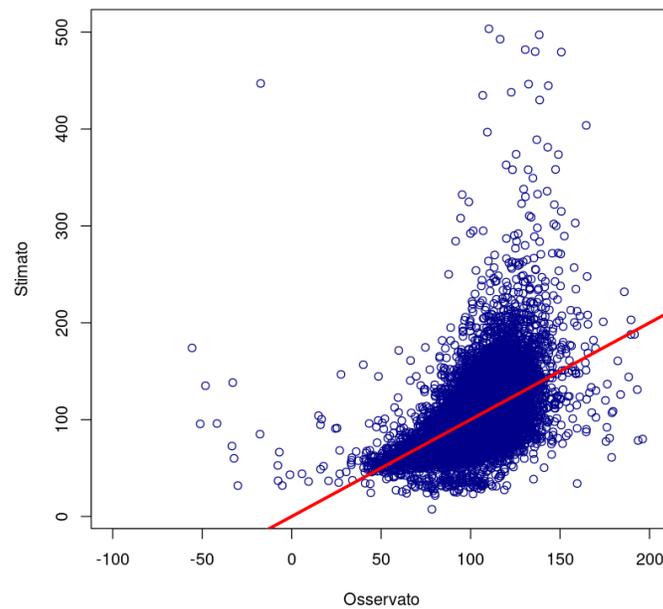


Figura 17: Grafico valori predetti vs. valori osservati

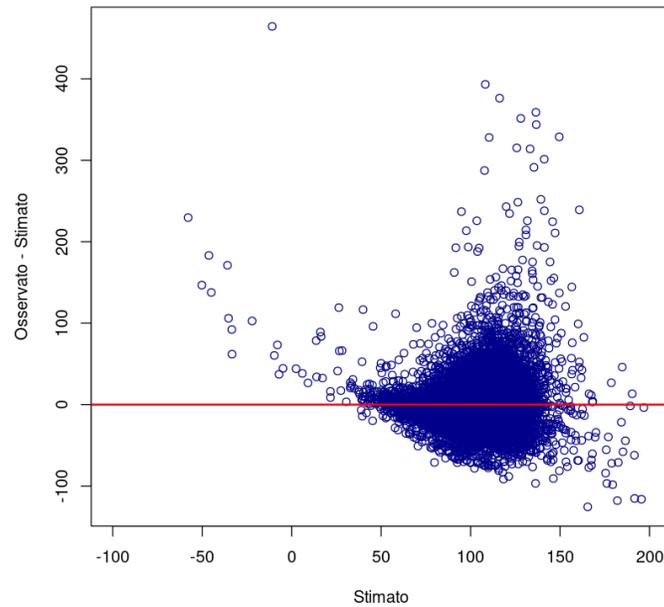


Figura 18: Grafico valori predetti vs. errore

Nella figura 17 viene riportato il grafico dei valori predetti vs. valori osservati. I punti dovrebbero tendere ad allinearsi lungo la bisettrice per poter sostenere che le previsioni del modello siano simili ai valori osservati. In questa circostanza è difficile sostenere che il modello abbia una buona capacità di predire la variabile di risposta. Il grafico dei valori predetti vs. errori (Figura 18) in una situazione ideale dovrebbe presentare i punti distribuiti in modo casuale intorno a zero. Questa distribuzione indica che il modello non ha un bias sistematico nella predizione, e gli errori sono distribuiti in modo equo sopra e sotto zero. È auspicabile anche l'assenza di pattern, la quale può suggerire che il modello non stia catturando completamente le relazioni tra le variabili o che ci siano altre variabili non considerate che influenzano la variabile di risposta. In questo caso si nota un leggero pattern nella struttura, oltre alla mancanza di una distribuzione omogenea lungo l'asse delle ascisse con $y = 0$.

In aggiunta ai risultati predittivi dei modelli, aventi come obiettivo quello di identificare il valore di pressione parziale di ossigeno nel sangue, va considerata anche un'analisi supplementare per comprendere se la direzione di tale previsione è corretta.

Va calcolata, infatti, una differenza (delta) che rappresenta il confronto tra le direzioni dei cambiamenti delle variabili predette rispetto ai cambiamenti della variabile osservata, tenendo conto delle osservazioni precedenti. Non si tiene dunque conto della vicinanza tra il valore osservato ed il valore stimato, ma si valuta solamente la direzione della stima.

Errati	Corretti
34.55 %	65.45 %

Tabella 13: Andamenti predetti

La tabella riassuntiva (Tabella 13) mostra la percentuale di andamenti correttamente predetti. Questa percentuale è relativamente bassa e conferma la limitata capacità del modello ad effetti misti di adattarsi ai dati e prevedere con accuratezza la pressione parziale di ossigeno.

5.4 Risultati modello logistico

Il modello logistico è proposto per identificare i soggetti critici come tali. L'obiettivo non è più quello di stimare il valore puntuale di pressione parziale di ossigeno bensì, dopo aver stabilito una soglia clinica, di individuare i soggetti con un valore al di sotto della soglia, da identificarsi come soggetti critici e dunque oggetto di attenzione. Come già discusso (cfr 4.1.4), la letteratura riporta una soglia di criticità per PaO_2 per valori inferiori a 60 mmHg. Utilizzando tale riferimento, si osservano 663 soggetti critici sui 11943 che costituiscono il dataset di validation. In base all'interesse clinico si pone una maggiore attenzione sui falsi negativi o i falsi positivi. In particolare, la curva ROC permette di mettere in relazione la sensibilità e la specificità sulla base delle diverse soglie e si configura come uno dei migliori strumenti per valutare le performance di un modello predittivo.

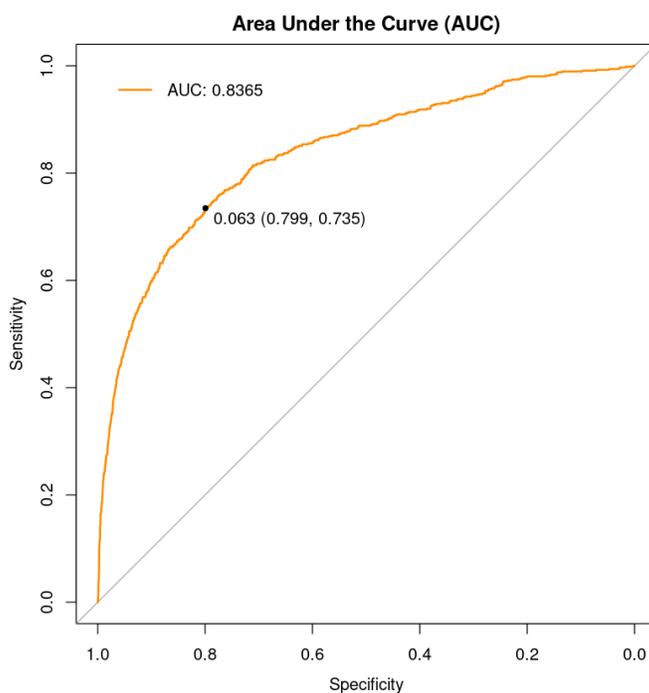


Figura 19: Curva ROC modello logistico

Nel caso descritto, la curva ROC presenta un AUC (Area Under Curve) superiore a 0.83 (0.8365). L'AUC è un indicatore chiave: se fosse 1, indicherebbe una perfetta capacità di discriminazione tra classi, mentre un valore di 0.5 significherebbe una discriminazione casuale, quindi un valore superiore a 0.835 suggerisce una buona capacità predittiva.

Mediante l'opzione `print.thres = TRUE` è possibile ricavare il valore di soglia migliore in grado di offrire i valori combinati più elevati di sensibilità e specificità. A seguito della definizione di questa soglia pari a 0.063, si calcola la matrice di confusione (Tabella 14).

	Non critico	Critico
Non critico	9024	2256
Critico	177	486

Tabella 14: Matrice di confusione

Nella matrice di confusione vi sono diversi individui misclassificati come falsi negativi: sono 177 i soggetti che il modello predice come non critici quando, invece nella realtà hanno una condizione clinica critica. Ciò comporta una sensibilità pari a 0.73.

Il valore di specificità, che identifica i veri negativi, minimizzando i falsi positivi ha un valore leggermente superiore pari a 0.8. Questo valore indica una buona capacità del modello logistico nell'identificare come non malati i soggetti che non lo sono realmente.

La scelta della miglior soglia proposta dal grafico della curva ROC porta a buoni risultati, con una discreta capacità del modello sia di identificare i pazienti critici (sensibilità 0.73) che di identificare i pazienti non critici con una specificità elevata (0.8).

La funzione `ci.sp` calcola la specificità ed il relativo intervallo di confidenza per un valore dato di sensibilità. Di default vengono calcolati intervalli di confidenza al 95% con 2000

replicazioni di bootstrap stratificato.

	2.5%	50%	97.5%
0.8	0.6659508	0.7213652	0.7576285
0.9	0.3850133	0.4627660	0.5425532

Tabella 15: Specificità e relativi intervalli di confidenza dati i valori di sensibilità

Dati i valori di sensibilità pari all'80% e al 90% (Tabella 15), emergono delle differenze significative: si ha rispettivamente un valore di specificità pari a 0.72 e 0.46. Nel secondo caso (sensibilità 0.9) si ha una notevole diminuzione della specificità a favore di un aumento contenuto di sensibilità. La decisione su quale valore di sensibilità sia migliore dipende principalmente dal contesto clinico che aiuta a determinare se porre maggiore attenzione alla riduzione dei falsi positivi o alla capacità di ridurre al minimo il rischio di diagnosticare erroneamente un paziente sano come critico.

La calibration belt (cfr. 4.1.5) si conferma rilevante nell'approfondimento dello studio di un modello logistico. Dal grafico qui riportato (Figura 20), emerge una ridotta capacità del modello di identificare la corretta probabilità di ogni soggetto.

Questa analisi sottolinea una discrepanza significativa tra le probabilità previste dal modello e le probabilità effettivamente osservate nei soggetti analizzati. In particolare, per i pazienti per i quali il modello stima una probabilità di essere pazienti critici compresa tra il 16% e il 46%, si osserva che la probabilità reale di essere in condizioni critiche risulta essere superiore. Tale differenza fornisce un'evidenza del fatto che il modello tende a sottostimare la probabilità che un paziente si trovi in condizione critica. Questo risultato

assume una notevole rilevanza clinica poiché indica che il modello logistico presenta una tendenza sistematica a sottostimare il rischio di gravi condizioni nei pazienti. Per probabilità estremamente basse e per valori di probabilità associati a pazienti critici compresi tra il 92% ed il 100% si osserva che la probabilità reale di essere in condizioni critiche risulta essere inferiore, per cui il modello in tali circostanze sovrastima il rischio di condizione critica.

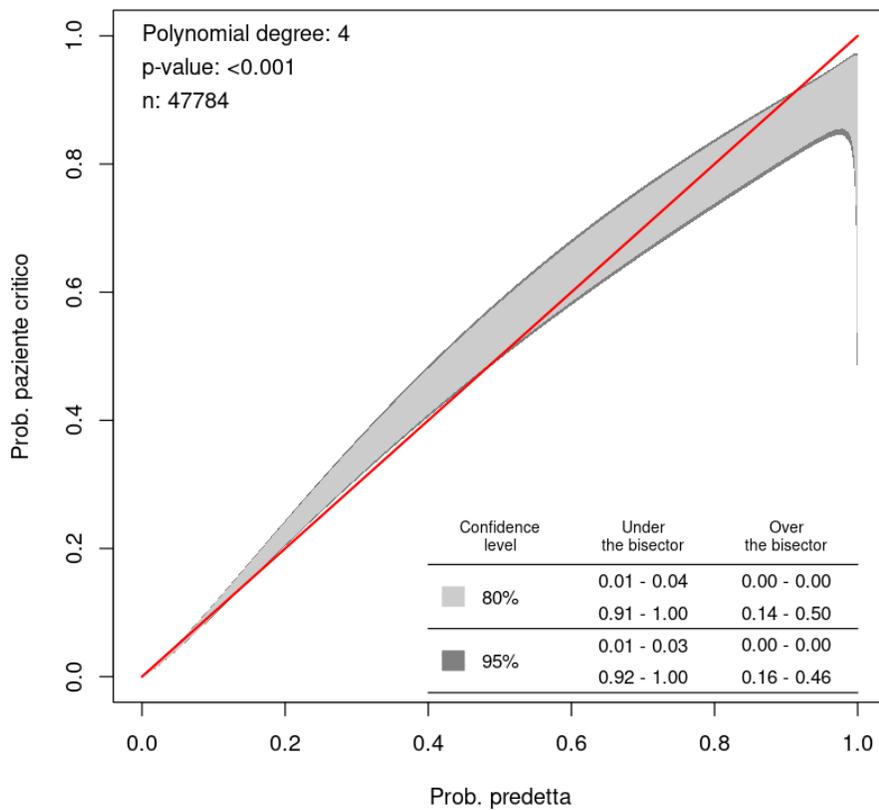


Figura 20: Calibration Belt

6. CONCLUSIONI

Il database MargheritaTre si è mostrato estremamente flessibile e completo di molte informazioni sulle condizioni cliniche dei pazienti ricoverati in terapia intensiva. La sua capacità di raccogliere una grande mole di dati provenienti da centri ospedalieri di tutta Italia lo qualifica un database ottimo per svolgere analisi complesse in ambito intensivo.

Considerato che il parametro PaO_2 viene rilevato meno frequentemente rispetto ad altri, poiché richiede l'esame dell'emogasanalisi, esame più invasivo e costoso rispetto al monitoraggio di altri parametri quali la saturazione del sangue (SpO_2), l'interesse clinico era quello di catturare la relazione tra le variabili dipendenti e la variabile risposta per fornire una previsione di PaO_2 anche negli istanti in cui essa non viene misurata. Ciò aiuterebbe a supportare il clinico nelle decisioni fornendo una stima del valore di pressione parziale sulla base di altre informazioni note. In altri termini, l'ambizioso scopo del presente progetto era quello di prevedere i valori della pressione parziale di ossigeno nel sangue sulla base di altre covariate rilevate nel corso del tempo, utilizzando centri ospedalieri selezionati. I risultati dello studio, tuttavia, non sono stati all'altezza delle aspettative, infatti sul test di validazione emerge un errore mediano (MedAE) per il modello ad effetti misti con regressione splines pari a 16.66 mmHg, valore abbastanza elevato che porta a sostenere che il modello non sia adatto ad effettuare una previsione accurata. Questa conclusione risulta confermata dall'elevata variabilità associata all'errore (deviazione standard 24.41), rendendo in gran parte inutilizzabile il modello a fini previsivi. Il migliore risultato stratificato per centro ospedaliero non permette in ogni caso di sostenere la bontà dei risultati e quindi esclude la generalizzazione per ogni struttura. Indubbiamente i risultati del centro ospedaliero 5 sono i più incoraggianti con un errore mediano di 13.72 mmHg e deviazione standard 13.49, ma rimangono ancora lontani dall'accuratezza richiesta dando ulteriore prova di come l'elevata variabilità dei soggetti comporti stime poco precise.

Nonostante i numerosi parametri vitali misurati e registrati, non è emersa una stretta relazione tra le variabili indipendenti e PaO_2 : questa circostanza spinge ad interrogarsi sulla selezione delle covariate stesse e sull'identificazione di un sottogruppo maggiormente rappresentativo sul quale riproporre le analisi svolte. Inoltre, è apparso marginale il ruolo della covariata legata al tempo, che non fornisce alcuna informazione aggiuntiva sulla condizione del paziente. Si tratta di un risultato che non sorprende eccessivamente: indica che i parametri misurati sul paziente sono sufficienti a descrivere le sue condizioni e non risulta fondamentale sapere le tempistiche del ricovero. Nel quadro delineato può essere utile riflettere sulla gestione dei dati mancanti che va certamente migliorata ed approfondita per fornire una maggiore qualità delle informazioni che verranno poi utilizzate nei modelli. L'attuale gestione, infatti prevedendo l'esclusione di un record se non presenta tutte le misurazioni, costituisce un limite in quanto non si utilizzano molte informazioni cliniche che pur vengono raccolte.

Il progetto ha evidenziato quanto sia cruciale la fase iniziale di raccolta dati, mettendo in luce l'importante ruolo svolto dalla preparazione accurata dei dati e dalla loro qualità: ne deriva la necessità per i centri ospedalieri di raccogliere i parametri rilevanti con modalità diverse da quelle attuali perfezionando ed ottimizzando il processo. Questa linea di azione può contribuire a creare un database più completo, fattore che può condurre analisi con evidenze più solide.

In ultimo è opportuno un breve riferimento alla selezione degli intubati. Tale passaggio, molto importante nella fase preliminare, non ha condotto ai risultati auspicati in quanto non vi è stato un miglioramento in termini predittivi del modello. È emerso come la variabile `intubati` riporti un coefficiente positivo (4.0916) che determina un aumento significativo ($p\text{-value} < 0.0001$) del valore medio di pressione parziale di ossigeno nel sangue a parità delle altre covariate. I soggetti intubati presentano in media un valore leggermente più elevato di PaO_2 dovuto alla respirazione meccanica a cui sono sottoposti per difficoltà respiratorie.

Considerare la ventilazione meccanica come covariata e non come criterio di selezione ha

permesso di mantenere una numerosità campionaria più elevata e, allo stesso tempo, ha consentito di comprendere le differenze tra il trattamento respiratorio e la respirazione autonoma. Appare dunque un valido punto di partenza per eventuali studi futuri una maggiore attenzione sui soggetti intubati.

Per quanto concerne invece l'analisi supplementare riguardante il modello logistico, essa ha dimostrato di avere una buona capacità di individuare correttamente sia i casi critici che quelli non critici. In base all'obiettivo clinico, è possibile porre una maggiore attenzione nell'identificazione dei soggetti critici rispetto ai non critici. Se lo scopo è quello di ridurre al minimo i falsi negativi al fine di identificare ed intervenire sui soggetti realmente positivi si predilige una maggiore sensibilità. Al contrario se non si vuole sottoporre ad esami invasivi soggetti che non sono realmente critici si valuterà una maggiore specificità. Il modello logistico proposto è stato in grado di rispondere con un buon grado di precisione a tale obiettivo. La calibration belt non conferma del tutto questi risultati in quanto, confrontando le probabilità stimate con quelle reali, si ha una sottostima da parte del modello logistico per probabilità predette fra 16 e 46%. Si ha, invece, una sovrastima per probabilità predette molto basse (< 3%) o elevate (92%). Emerge una capacità limitata del modello logistico di predire accuratamente le probabilità effettive di condizioni critiche nei pazienti, mettendo in luce quindi l'importanza di un'ulteriore revisione e perfezionamento del modello al fine di migliorare la sua precisione predittiva nell'ambito clinico. In determinate situazioni, tale sottostima può essere attribuita alla mancata inclusione di interazioni tra le variabili covariate considerate. Pertanto, sarebbe opportuno riesaminare il modello logistico valutando la possibilità di includere tali interazioni al fine di migliorare la sua capacità predittiva.

La scelta della soglia di classificazione è una decisione complessa che richiede un bilanciamento tra la capacità di identificare tutti i casi critici (sensibilità) e la capacità di eliminare i casi non critici (specificità), considerando l'obiettivo clinico specifico.

Relativamente all'analisi della curva ROC è emerso un risultato soddisfacente di AUC di

0.8365. Il modello logistico, dunque, è in grado di classificare correttamente i soggetti critici, mostrandosi strumento efficace nel supportare il personale medico nell'identificazione dei pazienti a maggiore rischio. Ciò nonostante il modello ha mostrato delle criticità nella stima continua della probabilità di essere un paziente critico.

Il progetto non ha condotto esattamente all'obiettivo verso il quale era indirizzato: malgrado ciò lo studio si rivela utile spunto per l'approfondimento della tematica legata alla pressione di ossigeno nel sangue con il fine ultimo portare un aiuto concreto al personale medico nel supporto al trattamento dei pazienti ricoverati in terapia intensiva.

7. POSSIBILI SVILUPPI FUTURI

Alla luce di quanto esposto fino ad ora, emerge chiaramente la necessità di ampliare il progetto di ricerca al fine di ottenere delle evidenze più solide. Nel gettare le basi per una migliore comprensione delle dinamiche legate alla pressione arteriosa parziale di ossigeno nel sangue, vi sono molteplici aspetti da approfondire e migliorare.

In particolare, facendo riferimento alle analisi che considerano il modello ad effetti misti come elemento principale, si ritiene possa essere di grande interesse lavorare sui dati mancanti. Nello specifico, le assunzioni fatte per i dati mancanti (selezione dei soli record che non presentassero alcun valore mancante) hanno diminuito la numerosità campionaria e la frequenza di rilevazioni nel tempo per ogni soggetto. Una diversa gestione dei dati mancanti potrebbe portare ad un dataset di analisi più accurato e completo, portando a migliori risultati predittivi. Per tali ragioni è allo studio, con il supporto clinico, l'implementazione di un algoritmo di imputazione dei valori mancanti, in modo tale da non perdere un numero elevato di record a causa della mancanza di pochi parametri rilevati. Inoltre, è in fase di valutazione la selezione di un sottoinsieme di variabili indipendenti rilevati con maggiore frequenza in terapia intensiva e più legate alla variabile risposta (ad esempio PEEP, End-Tidal CO_2 , FiO_2) per i soli soggetti intubati. Sarebbe auspicabile utilizzare come variabili indipendenti i parametri registrati con maggiore frequenza, in modo tale da fornire informazioni aggiuntive al modello e offrire un migliore tracciato dell'andamento di PaO_2 in funzione delle covariate. Questi diversi criteri di selezione potrebbero ridurre la numerosità del campione, aumentando però il numero di misurazioni nel tempo e riducendo al minimo la presenza di dati mancanti. Il miglioramento della qualità del dato, dunque, si presenta come la direzione da intraprendere per ottenere una previsione più accurata.

In un progetto di più ampio respiro sarebbe possibile mettere in discussione la scelta del modello ad effetti misti come modello predittivo e proporre degli strumenti più sofisticati nel-

l'ambito del machine learning per ottenere una predizione della variabile di interesse. Al riguardo, si riscontrano numerosi articoli [49] [50] che utilizzano le reti neurali per rispondere a tale necessità; si ricorre in particolare ad una rete neurale LSTM (Long Short-Term Memory), rete di tipo ricorrente specializzata nell'elaborare sequenze di dati, tipicamente utilizzata per analizzare e predire dati temporali in pazienti ricoverati in reparti intensivi. Le reti neurali LSTM sono progettate per affrontare le limitazioni delle reti neurali ricorrenti tradizionali, le quali riscontrano diverse criticità nel catturare dipendenze a lungo termine nei dati longitudinali.

In conclusione, il presente studio non è che la prima tappa nel percorso verso l'approfondimento di una tematica strategica nell'ambito della terapia intensiva, percorso che deve essere sempre guidato dall'intento di supportare le decisioni cliniche con strumenti di maggior efficacia nell'interesse ultimo della salvaguardia del paziente.

Bibliografia

1. Kreit JW. Gas exchange. Kreit JW, editor. Oxford University Press; 2017. doi:10.1093/med/9780190670085.003.0002
2. Melody Bishop BSc C RRT. Basic principles of mechanical ventilation. Robinson AB, editor. 2022.
3. Pham T, Brochard LJ, Slutsky AS. Mechanical ventilation: State of the art. Mayo Clinic Proceedings. 2017;92: 1382–1400. doi:10.1016/j.mayocp.2017.05.004
4. Grasselli G, al. CSC et. ESICM guidelines on acute respiratory distress syndrome: Definition, phenotyping and respiratory support strategies. Intensive Care Medicine. 2023;49: 727–759. doi:10.1007/s00134-023-07050-7
5. Acute respiratory distress syndrome: The berlin definition. JAMA. 2012;307. doi:10.1001/jama.2012.5669
6. Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, et al. Acute respiratory distress syndrome. Nature Reviews Disease Primers. 2019;5. doi:10.1038/s41572-019-0069-0
7. Saguil A. FM. Acute respiratory distress syndrome: Diagnosis and management. American Family Physician. 2012;85.
8. Diaz R HD. Barotrauma and mechanical ventilation. 2022. Available: <https://www.ncbi.nlm.nih.gov/books/NBK545226/>
9. Pelosi P, Brazzi L, Gattinoni L. Prone position in acute respiratory distress syndrome. European Respiratory Journal. 2002;20: 1017–1028. doi:10.1183/09031936.02.00401702
10. Avishay DM, Tenny KM. Henry’s law. StatPearls. 2023.

11. Demiselle J, al. EC et. Target arterial PO₂ according to the underlying pathology: A mini-review of the available data in mechanically ventilated patients. *Annals of Intensive Care*. 2021;11. doi:10.1186/s13613-021-00872-y
12. Feiner JR, Weiskopf RB. Evaluating pulmonary function. *Critical Care Medicine*. 2017;45: e40–e48. doi:10.1097/ccm.0000000000002017
13. Barrot L, Asfar P, Mauny F, Winiszewski H, Montini F, Badie J, et al. Liberal or conservative oxygen therapy for acute respiratory distress syndrome. *New England Journal of Medicine*. 2020;382: 999–1008. doi:10.1056/nejmoa1916431
14. Diepen S van, Tavazzi G, Morrow DA. Blood pressure and oxygenation targets after out-of-hospital cardiac arrest-trial (BOX). *European Heart Journal Acute Cardiovascular Care*. 2022;11: 714–715. doi:10.1093/ehjacc/zuac109
15. Available: <https://clinicaltrials.gov/ct2/show/NCT04534959>
16. Available: <https://www.anzics.com.au/current-active-endorsed-research/mega-rox/>
17. Young PJ, Hodgson CL, Rasmussen BS. Oxygen targets. *Intensive Care Medicine*. 2022;48: 732–735. doi:10.1007/s00134-022-06714-0
18. Semler MW, Casey JD, Lloyd BD, Hastings PG, Hays MA, Stollings JL, et al. Oxygen-saturation targets for critically ill adults receiving mechanical ventilation. *New England Journal of Medicine*. 2022;387: 1759–1769. doi:10.1056/nejmoa2208415
19. Gowri M, Babu S. Association between SpO₂/FiO₂ ratio and PaO₂/FiO₂ ratio in different modes of oxygen supplementation. *Indian Journal of Critical Care Medicine*. 2021;25: 1001–1005. doi:10.5005/jp-journals-10071-23977

20. Patel S, Singh G, Zarbiv S, Ghiassi K, Rachoin J-S. Mortality prediction using SaO₂/FiO₂ ratio based on eICU database analysis. Tran QK, editor. *Critical Care Research and Practice*. 2021;2021: 1–9. doi:10.1155/2021/6672603
21. Girardis M, Busani S, Damiani E, Donati A, Rinaldi L, Marudi A, et al. Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit. *JAMA*. 2016;316: 1583. doi:10.1001/jama.2016.11993
22. Madan A. Correlation between the levels of SpO₂ and PaO₂. *Lung India*. 2017;34: 307. doi:10.4103/lungindia.lungindia_106_17
23. Goodgame B, Zeserson E, al. JDH et. 414. *Critical Care Medicine*. 2012;40: 1–328. doi:10.1097/01.ccm.0000424632.30897.65
24. Thijssen M, Janssen L, Noble J le, Foudraine N. Facing SpO₂ and SaO₂ discrepancies in ICU patients: Is the perfusion index helpful? *Journal of Clinical Monitoring and Computing*. 2019;34: 693–698. doi:10.1007/s10877-019-00371-3
25. Helmerhorst HJF, Arts DL, Schultz MJ, Voort PHJ van der, Abu-Hanna A, Jonge E de, et al. Metrics of arterial hyperoxia and associated outcomes in critical care. *Critical Care Medicine*. 2017;45: 187–195. doi:10.1097/ccm.0000000000002084
26. Silverman TA, Weiskopf RB. Hemoglobin-based oxygen carriers. *Anesthesiology*. 2009;111: 946–963. doi:10.1097/aln.0b013e3181ba3c2c
27. Seganti N MA et al Imbriaco G. Analisi del corretto utilizzo del monitoraggio EtCO₂ nel percorso emergenza-urgenza ed in terapia intensiva. *SCENARIO - Rivista del Gruppo Italiano Tracheostomizzati*. 2023. Available: <https://scenario.aniarti.it/index.php/scenario/article/view/2>
28. Kao LW, Nañagas KA. Carbon monoxide poisoning. *Emergency Medicine Clinics of North America*. 2004;22: 985–1018. doi:10.1016/j.emc.2004.05.003

29. Fuentes S, Chowdhury YS. Fraction of inspired oxygen. StatPearls [Internet]. 2023. Available: <https://www.ncbi.nlm.nih.gov/books/NBK560867/>
30. Wright RO, Lewander WJ, Woolf AD. Methemoglobinemia: Etiology, pharmacology, and clinical management. *Annals of Emergency Medicine*. 1999;34: 646–656. doi:10.1016/s0196-0644(99)70167-8
31. Lara-Cantón I, Badurdeen S, Dekker J, Davis P, Roberts C, Pas A te, et al. Oxygen saturation and heart rate in healthy term and late preterm infants with delayed cord clamping. *Pediatric Research*. 2022. doi:10.1038/s41390-021-01805-y
32. Pino RM, Singh J. Appropriate clinical use of lactate measurements. *Anesthesiology*. 2020;134: 637–644. doi:10.1097/aln.0000000000003655
33. Pham T, Brochard LJ, Slutsky AS. Mechanical ventilation: State of the art. *Mayo Clinic Proceedings*. 2017;92: 1382–1400. doi:10.1016/j.mayocp.2017.05.004
34. Puri S, Paul G, Sood P. Interpretation of arterial blood gas. *Indian Journal of Critical Care Medicine*. 2010;14: 57–64. doi:10.4103/0972-5229.68215
35. Chu Y, Tang K, Hsu Y-C, Huang T, Wang D, Li W, et al. Non-invasive arterial blood pressure measurement and SpO₂ estimation using PPG signal: A deep learning framework. *BMC Medical Informatics and Decision Making*. 2023;23. doi:10.1186/s12911-023-02215-2
36. James GD, Gerber LM. Measuring arterial blood pressure in humans: Auscultatory and automatic measurement techniques for human biological field studies. *American Journal of Human Biology*. 2017;30: e23063. doi:10.1002/ajhb.23063
37. Michard F. Changes in Arterial Pressure during Mechanical Ventilation. *Anesthesiology*. 2005;103: 419–428. doi:10.1097/00000542-200508000-00026

38. Tusman G, Bohm SH, Suarez-Sipmann F. Advanced uses of pulse oximetry for monitoring mechanically ventilated patients. *Anesthesia & Analgesia*. 2017;124: 62–71. doi:10.1213/ane.0000000000001283
39. Giuliano K AJ et al. Temperature measurement in critically ill adults: A comparison of tympanic and oral methods. *American journal of critical care : an official publication, American Association of Critical-Care Nurses*. 2000;9: 254–261.
40. Pinto V, Sharma S. Continuous positive airway pressure. *StatPearls [Internet]*. 2023. Available: <https://www.ncbi.nlm.nih.gov/books/NBK482178/>
41. Kilgannon JH. Association between arterial hyperoxia following resuscitation from cardiac arrest and in-hospital mortality. *JAMA*. 2010;303: 2165. doi:10.1001/jama.2010.707
42. O B-F, C. KH. The normal pH value of human skin. *Der Hautarzt; Zeitschrift fur Dermatologie, Venerologie, und verwandte Gebiete*. 1986;37: 126–129.
43. Marikar D, Babu P, Fine-Goulden M. How to interpret lactate. *Archives of disease in childhood - Education edition*. 2020;106: 167–171. doi:10.1136/archdischild-2020-319601
44. Hall MM, Rajasekaran S, Thomsen TW, Peterson AR. Lactate: Friend or foe. 2016;8: S8–S15. doi:10.1016/j.pmrj.2015.10.018
45. Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, editors. *Longitudinal data analysis*. Chapman; Hall/CRC; 2008. pp. 189–234. doi:10.1201/9781420011579
46. Yuan R, Lazaric A, Gower RM. Sketched newton-raphson. *SIAM Journal on Optimization*. 2022;32: 1555–1583. doi:10.1137/21m139788x
47. Mirabile VS, Shebl E, Sankari A, al. et. Respiratory failure. *StatPearls*. 2023. Available: <https://www.ncbi.nlm.nih.gov/books/NBK526127/>

48. Baekgaard JS, al. P-SA et. Early hyperoxemia is associated with lower adjusted mortality after severe trauma: Results from a french registry. *Critical Care*. 2020;24. doi:10.1186/s13054-020-03274-x
49. Ren S, Zupetic JA, Tabary M, DeSensi R, Nouraie M, Lu X, et al. Machine learning based algorithms to impute PaO₂ from SpO₂ values and development of an online calculator. *Scientific Reports*. 2022;12. doi:10.1038/s41598-022-12419-7
50. Lin Y-W, Zhou Y, Faghri F, Shaw MJ, Campbell RH. Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. Moskovich R, editor. *PLOS ONE*. 2019;14: e0218942. doi:10.1371/journal.pone.0218942