

Igor Tona Peres¹, Silvio Hamacher¹, Fernando Luiz Cyrino Oliveira¹, Fernando Augusto Bozza², Jorge Ibrain Figueira Salluh³

1. Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro - Rio de Janeiro (RJ), Brasil.

2. Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz - Rio de Janeiro (RJ), Brasil.

3. Instituto D'Or de Pesquisa e Ensino - Rio de Janeiro (RJ), Brasil.

Conflitos de interesse: Nenhum.

Jorge Ibrain Figueira Salluh é cofundador da Epimed Solutions, empresa de análises com base na nuvem.

Submetido em 29 de outubro de 2020

Aceito em 10 de fevereiro de 2021

Autor correspondente:

Igor Tona Peres

Departamento de Engenharia Industrial
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - Gávea
CEP: 22451-900 - Rio de Janeiro (RJ), Brasil
E-mail: igor.peres.puc@gmail.com

Editor responsável: Gilberto Friedman

DOI: 10.5935/0103-507X.20210025

Predição do tempo de permanência em unidades de terapia intensiva: uma revisão concisa

Prediction of intensive care units length of stay: a concise review

INTRODUÇÃO

O tempo de permanência (TDP) na unidade de terapia intensiva (UTI) é uma das métricas mais comumente utilizadas para avaliar a qualidade dos serviços de saúde. Apesar de suas potenciais limitações, o TDP na UTI é fácil de medir e reprodutível, assim como pode ser utilizado como marcador substituto para avaliar o uso de recursos, custos e eficiência.⁽¹⁾ Mais ainda, trata-se de um desfecho centrado no paciente, de forma que interessa a diferentes partes, como familiares dos pacientes, gestores, pagadores e equipe da UTI (Figura 1). Contudo, na maior parte das circunstâncias, o TDP na UTI é utilizado de forma retrospectiva, para avaliar a eficiência da UTI ou compreender os pacientes com maior TDP e, conseqüentemente, custos mais elevados. Assim, a predição do TDP no momento da admissão à UTI poderia ajudar a coordenar os cuidados, implantar medidas preventivas e manter uma melhor comunicação com gestores, pagadores e familiares, estabelecendo expectativas realistas.

A capacidade de prever o TDP para um paciente individual também poderia levar à melhora nas práticas de *benchmarking*, ao comparar o TDP de um único paciente com o de pacientes com diagnósticos similares (englobando, dessa forma, pontos fora da curva e identificando alvos para melhoria da qualidade). Permitiria também analisar os resultados gerais do TDP predito em comparação com o TDP observado (real) para todos os pacientes na UTI (ou subgrupos, como sepse, síndrome do desconforto respiratório agudo etc.). Esse tipo de análise (de forma similar ao que se faz com mortalidade por meio de taxas padronizadas) poderia ser uma mensuração robusta da eficiência da UTI.⁽²⁾ A quantidade maciça de dados diariamente gerados pela UTI, em combinação com recentes avanços tecnológicos e de métodos estatísticos, torna viável desenvolver modelos preditivos que possam ajudar os médicos na gestão diária da UTI e melhorar a qualidade dos cuidados e a eficiência.⁽³⁾

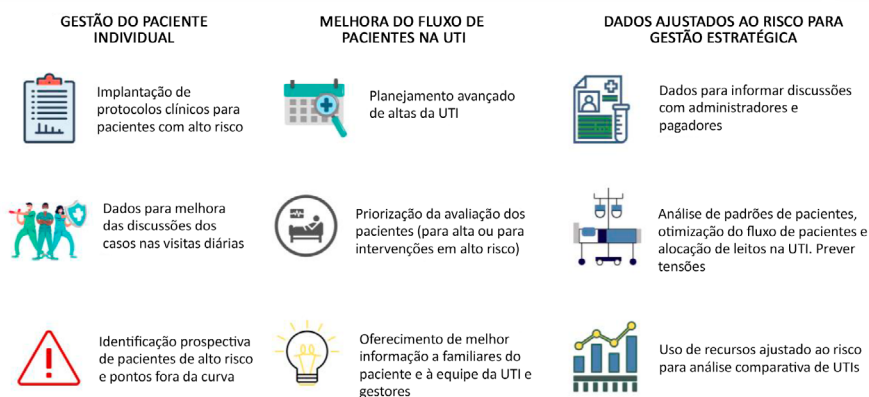


Figura 1 - Principais benefícios da predição do tempo de permanência na unidade de terapia intensiva. UTI - unidade de terapia intensiva.

O que se sabe sobre métodos para prever o tempo de permanência na unidade de terapia intensiva?

Em sua maioria, os estudos que avaliaram modelos para predição do TDP na UTI⁽⁴⁻⁸⁾ utilizaram regressão linear multivariada e não testaram outras abordagens para comparar sua precisão. A regressão linear tem a vantagem de uma interpretação mais clara, contudo, sua principal limitação é o fato de assumir um relacionamento linear com as covariáveis. Verburg et al.⁽⁹⁾ realizaram, em 2017, uma revisão sistemática, concluindo que o modelo desenvolvido por esses estudos não satisfaz as exigências gerais para a predição do TDP na UTI, tanto com relação ao planejamento de alocação de recursos, quanto para identificação do TDP para um paciente individual.

Procedemos a buscas na literatura para encontrar estudos que comparassem diferentes modelos de predição do TDP. Realizamos buscas nas bases de dados MEDLINE®, Embase e Scopus desde o início até 6 de outubro de 2020. A busca se limitou à língua inglesa e aos tipos de publicação “*article*”, “*article in press*” e “*review*”. A busca compreendeu os campos título e descritores, não se fazendo restrições quanto ao período de publicação. Usamos as seguintes questões: (“*ICU*” or “*Intensive Care*”) and (“*length of stay*”) and (“*predict**”). A seleção de estudos envolveu quatro aspectos: formulação de critérios de elegibilidade; leitura do resumo e seleção do texto completo para leitura; leitura do texto completo e seleção; e inclusão de novos estudos por busca retrógrada e anterógrada. Consideramos os seguintes critérios de elegibilidade para inclusão de estudos: estudos que incluíssem e comparassem modelos para predição do TDP na UTI, relatos estatísticos em termos de erro médio quadrático da previsão (*root mean*

square error - RMSE), erro médio absoluto (EMA) ou R². Não se incluíram estudos inferenciais. Encontramos cinco estudos de predição, cujas características são apresentadas de forma resumida na tabela 1.

Verburg et al.⁽¹⁰⁾ compararam seis modelos de regressão para predição do TDP na UTI para um conjunto de dados de 32.667 admissões à UTI. Os melhores modelos foram o modelo linear generalizado (MLG) com distribuição Gaussiana e o MLG com distribuição de Poisson; o pior modelo foi a regressão de Cox. O estudo testou a log-transformação do TDP na UTI para reduzir a distorção e melhorar a simetria de distribuição das variáveis, o que apresentou melhores resultados. Os autores também testaram truncamento do TPD aos 30 dias, o que melhorou o desempenho do modelo.

Moran et al.⁽⁴⁾ compararam sete modelos de regressão para prever o TDP na UTI para um conjunto de dados de 11.663 admissões à UTI. O melhor modelo foi o modelo linear misto (MLM). Os autores também testaram log-transformação do TDP na UTI, que obteve melhores resultados.

Houthoof et al.⁽¹¹⁾ compararam diferentes modelos baseados em dados para prever o TDP na UTI em pacientes no dia 5 na UTI. O modelo com melhor desempenho foi regressão de vetor suporte (SVR), e o pior foi uma rede neural artificial (RNA). Os autores incluíram na metodologia uma log-transformação do TDP na UTI, normalização, e seleção de variáveis (utilizando *random forest importance list* e procedimento *backward elimination* com SVR).

Li et al.⁽¹²⁾ criaram um modelo de predição com uso de técnicas de pré-processamento, uma análise de dados exploratórios e o algoritmo *least absolute shrinkage and*

Tabela 1 - Características de cada estudo de predição

Autor	Coorte	TDP na UTI ou no hospital	Delineamento	Predição	RMSE	EMA	R ² (%)	TDP em UTI em escala logarítmica	Truncamento aos 30 dias	Excluídos	Termos quadráticos/cúbicos	Termos de interação	Normalização	Divisão	Validação cruzada
Moran et al. ⁽⁴⁾	111.663	TDP em UTI	131 UTIs mistas	Admissão	4,50	2,30	22,0	Sim	Não	TDP em UTI > 60 dias	Sim	Sim	Não	N/A	N/A
Verburg et al. ⁽¹⁰⁾	32.667	TDP em UTI	83 UTIs mistas	Admissão	7,28	3,43	15,4	Sim	Sim	TDP em hospital > 365 dias	Não	Não	Não	Bootstrap	N/A
Houthoof et al. ⁽¹¹⁾	14.480	TDP em UTI	14 UTIs clínicas	Dia 5	N/A	1,79	21,9	Sim	Não	TDP em UTI > 40 dias	Não	Não	Sim (método não informado)	60/40	N/A
Li et al. ⁽¹²⁾	1214	TDP em UTI	Uma UTI	Admissão	0,88	0,87	35,0	Sim	Não	Não	Não	Não	Score-z	70/30	10 vezes
Muhlestein et al. ⁽¹³⁾	41.222	TDP em hospital	1.000 hospitalais	Admissão	0,56*	N/A	N/A	Não	Não	Não	Não	Não	Score-z	80/20	5 vezes
Caetano et al. ⁽¹⁴⁾	26.431	TDP em hospital	Um hospital	Admissão	0,47*	0,22*	81,3	Sim	Não	Não	Não	Não	Score-z	N/A	5 vezes

TDP - tempo de permanência; UTI - unidade de terapia intensiva; RMSE - erro médio quadrático da previsão; EMA - erro médio absoluto; N/A - não informado; * Erro médio quadrático da previsão/erro médio absoluto calculado utilizando log-transformação.

selection operator (LASSO). Na metodologia de pré-processamento, os autores incluíram tratamento de dados faltantes, transformação de Box-Cox (para TDP na UTI e variáveis com coeficientes de assimetria acima de 0,5) e normalização pelo *score-z*. Na fase de análise de dados exploratórios, o estudo explorou novas variáveis e analisou a colinearidade entre as variáveis existentes.

Encontramos também dois artigos sobre predição que compararam modelos baseados em dados para o TDP no hospital e apresentaram resultados significantes. Muhlestein et al.⁽¹³⁾ desenvolveram um novo método para, de forma sistemática, ranquear, selecionar e combinar diferentes algoritmos baseados em dados, construindo um modelo que prediz o TDP após craniotomia por tumores cerebrais. Os algoritmos com melhor desempenho foram *árvore gradient boosted tree* (GBT) e SVR. Esses modelos foram combinados com *elastic net*, para criar um modelo de conjunto. A metodologia de pré-processamento incluiu o tratamento para dados faltantes e normalização pelo *score-z*. Caetano et al.⁽¹⁴⁾ utilizaram um método baseado em dados para prever o TDP no hospital para um conjunto de dados de 26.431 admissões. O melhor modelo foi o *random forest* (RF), e os piores modelos foram a regressão por mínimos quadrados ordinários (MQO) e a *árvore de decisão* (DT). A metodologia considerou uma estratégia de pré-processamento que incluiu imputação de *k-nearest neighbors* (k-NN) para tratar de valores faltantes e normalização pelo *score-z* para colocar valores numéricos em uma mesma escala. Além disto, aplicou-se uma log-transformação à covariável “TDP prévio” e à variável de desfecho “TDP”.

A métrica de desempenho mais comumente utilizada para comparar os modelos de predição foi o RMSE, seguida por EMA e pelo coeficiente de determinação (R^2). Ao avaliar a tabela 1, podemos observar que os estudos

com melhor desempenho foram os de Caetano et al.,⁽¹⁴⁾ Muhlestein et al.⁽¹³⁾ e Li et al.⁽¹²⁾ Os resultados obtidos podem ser explicados pelo desenvolvimento de uma metodologia estruturada baseada em dados. Esses estudos incluíram, na fase de pré-processamento, o tratamento dos dados faltantes, a log-transformação (ou transformação por Box-Cox) para TDP na UTI e a normalização pelo *score-z*. Além disso, sua metodologia incluiu a separação do conjunto de dados em coortes de treinamento e teste, além da utilização de uma fase de validação cruzada para analisar sobreajuste do modelo. Li et al.⁽¹²⁾ também exploraram novas variáveis e analisaram a colinearidade entre as variáveis existentes.

Com relação aos tipos de modelos testados em cada estudo, podemos separá-los em modelos estatísticos e modelos baseados em dados, como apresentamos na tabela 2. Podemos observar que o SVR, um modelo atualizado com base em resultados, superou outros modelos em dois estudos. Outros modelos que apresentaram bons resultados foram GBT, RF, MLG, MLM e LASSO. Assim, sugerimos que futuros estudos levem em consideração as seguintes fases para obter predição razoável do TDP na UTI: extração dos dados e engenharia dos resultados; tratamento dos dados faltantes e pontos fora da curva; divisão dos dados em treinamento e teste; pré-processamento dos dados, incluindo análise de colinearidade, seleção de variáveis, transformações para resolver assimetrias, e normalização; validação cruzada para avaliar sobreajuste; e separação dos dados em grupos de treino e teste, comparando os diferentes tipos de modelos, inclusive modelos baseados em dados. Mais ainda, futuros estudos deveriam relatar seus resultados em termos de predição de erro (RMSE e EMA), que podem ajudar os pesquisadores a obter conclusões a respeito dos melhores modelos e fazer novas recomendações.

Tabela 2 - Modelos incluídos em cada estudo de predição

Autor	Modelo APACHE	Modelos estatísticos									Modelos com base em dados							
		MQO	MLG	MLM	MLG estendido	SN	FMM	TE	CPH	LASSO	RNA	k-NN	RF	SVR	RVR	DT	GBT	NB
Moran et al. ⁽⁴⁾		X	X	X	X	X	X	X										
Verburg et al. ⁽¹⁰⁾	X	X	X								X							
Houthoof et al. ⁽¹¹⁾												X	X	X	X	X	X	
Li et al. ⁽¹²⁾												X						
Muhlestein et al. ⁽¹³⁾		X										X	X	X	X		X	X
Caetano et al. ⁽¹⁴⁾		X										X	X	X	X			

APACHE - *Acute Physiology and Chronic Health Evaluation*; MQO - mínimos quadrados ordinários; MLG - modelo linear generalizado; MLM - modelo linear misto; SN - *skew normal/skew-t*; FMM - modelo mistura finito; TE - efeito do tratamento; CPH - risco proporcional de Cox; LASSO - *least absolute shrinkage and selection operator*; RNA - rede neural artificial; k-NN - *k-nearest neighbors*; RF - *random forest*; SVR - regressão de vetor suporte; RVR - regressão por vetores de relevância; DT - *árvore decisória*; GBT - *árvore gradient boosted tree*; NB - *naive bayes*.

Quanto à distribuição do TDP na UTI, a maioria dos autores testaram a log-transformação para reduzir distorções da distribuição. O truncamento dos dados de TDP na UTI é uma medida comum para evitar valores extremos. Assim, o truncamento em percentuais elevados (95% ou 99%) é uma alternativa para identificação de pontos fora da curva. Entretanto, o truncamento pode ser inadequado, porque podem ocorrer diferenças substanciais nos valores truncados, assim como melhorias mais acentuadas da eficiência podem ser alcançadas em pacientes com TDP na UTI mais longos. Assim, recomendamos cuidado ao comparar modelos que utilizaram dados truncados com aqueles que utilizaram dados originais.

Claramente não existe um modelo único que deve ser usado em todas as situações. Os melhores resultados dependerão de cada conjunto de dados, assim como os modelos devem ser treinados especificamente para cada caso. Mais ainda, é crucial extrair da base de dados da UTI as covariáveis relevantes. Alguns estudos demonstraram que existe uma relação não linear entre TDP na UTI e grau de severidade do paciente. Em outras palavras, pacientes mais graves tendem a ter um TDP mais longo. Contudo, os pacientes mais graves correm riscos mais elevados de óbito, o que pode diminuir o TDP na UTI previsto. Assim, é importante incluir aspectos relacionados à gravidade do paciente. Peres et al.⁽¹⁵⁾ sugeriram uma lista de fatores de risco para TDP na UTI que deveriam ser incluídos nos modelos de predição, como comorbidades, intervenções invasivas, marcadores laboratoriais e razão para admissão à UTI. Os modelos baseados em dados serão capazes de compreender esse relacionamento não linear, se aspectos relevantes forem incluídos na análise. A inclusão de variáveis irrelevantes pode aumentar a dimensão do problema, o que pode prejudicar os resultados do modelo. Por outro lado, a exclusão antecipada de aspectos relevantes pode gerar

resultados aquém do ideal. Assim, a extração de variáveis do conjunto de dados deve ser realizada com cautela.

Nosso estudo tem algumas limitações. Primeiramente, os trabalhos de Houthoof et al.⁽¹¹⁾ e Li et al.⁽¹²⁾ analisaram populações restritas. Houthoof et al.⁽¹¹⁾ examinaram uma coorte que incluía apenas pacientes clínicos, enquanto Li et al.⁽¹²⁾ investigaram uma coorte de uma única UTI. Em segundo lugar, incluímos dois artigos que focalizaram sua análise em predição do TDP no hospital e não na UTI. A forma da distribuição do TDP no hospital tende a ser similar à distribuição do TDP em UTI, porém alguns dos pressupostos devem ser distintos entre eles. Em terceiro lugar, um artigo⁽¹¹⁾ realizou a predição do TDP na UTI no dia 5, em vez de tê-lo feito na admissão. A distribuição do TDP na UTI após o dia 5 não foi a mesma comparativamente ao TDP original, o que pode ter afetado a análise comparativa.

CONCLUSÃO

Embora prever o tempo de permanência na unidade de terapia intensiva possa ser valioso para diversas partes (como médicos, pacientes, familiares e administradores), atualmente a maioria dos modelos publicados apresenta limitações para avaliar o tempo de permanência individual ou o desempenho geral da unidade de terapia intensiva. Futuros estudos que objetivem desenvolver e validar modelos para prever o tempo de permanência na unidade de terapia intensiva deveriam incluir em seus testes modelos com base em dados, especialmente para os desenvolvidos com grandes conjuntos de dados. Adicionalmente, esses modelos devem ser apresentados em tempo real e com interface amigável para o usuário, para, assim, permitir que as informações sejam utilizadas no local do atendimento e possam ter um impacto positivo nos desfechos clínicos.

REFERÊNCIAS

1. Dongelmans DA, Pilcher D, Beane A, Soares M, Del Pilar Arias Lopez M, Fernandez A, et al. Linking of global intensive care (LOGIC): an international benchmarking in critical care initiative. *J Crit Care.* 2020;60:305-10.
2. Ramos FJ, Salluh JI. Data-driven management for intensive care units. *ICU Manag Pract.* 2019;19(1):20-3.
3. Carra G, Salluh JI, da Silva Ramos FJ, Meyfroidt G. Data-driven ICU management: using Big Data and algorithms to improve outcomes. *J Crit Care.* 2020;60:300-4.
4. Moran JL, Solomon PJ; ANZICS Centre for Outcome and Resource Evaluation (CORE) of the Australian and New Zealand Intensive Care Society (ANZICS). A review of statistical estimators for risk-adjusted length of stay: analysis of the Australian and New Zealand Intensive Care Adult Patient Data-Base, 2008-2009. *BMC Med Res Methodol.* 2012;12:68.
5. Kramer AA, Zimmerman JE. A predictive model for the early identification of patients at risk for a prolonged intensive care unit length of stay. *BMC Med Inform Decis Mak.* 2010;10:27.
6. Niskanen M, Reinikainen M, Pettilä V. Case-mix-adjusted length of stay and mortality in 23 Finnish ICUs. *Intensive Care Med.* 2009;35(6):1060-7.
7. Vasilevskis EE, Kuzniewicz MW, Cason BA, Lane RK, Dean ML, Clay T, et al. Mortality probability model III and Simplified acute physiology score II: assessing their value in predicting length of stay and comparison to APACHE IV. *Chest.* 2009;136(1):89-101.
8. Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL. Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV. *Crit Care Med.* 2006;34(10):2517-29.
9. Verburg IW, Atashi A, Eslami S, Holman R, Abu-Hanna A, de Jonge E, et al. Which models can I use to predict adult ICU length of stay? A systematic review. *Crit Care Med.* 2017;45(2):e222-31.

10. Verburg IW, de Keizer NF, de Jonge E, Peek N. Comparison of regression methods for modeling intensive care length of stay. *PLoS One*. 2014;9(10):e109684.
11. Houthoofd R, Ruysinck J, van der Hertten J, Stijven S, Couckuyt I, Gadeyne B, et al. Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores. *Artif Intell Med*. 2015;63(3):191-207.
12. Li C, Chen L, Feng J, Wu D, Wang Z, Liu J, et al. Prediction of length of stay on the intensive care unit based on least absolute shrinkage and selection operator. *IEEE Access*. 2019;7:110710-21.
13. Muhlestein WE, Akagi DS, Davies JM, Chambless LB. Predicting inpatient length of stay after brain tumor surgery: developing machine learning ensembles to improve predictive performance. *Neurosurgery*. 2019;85(3):384-93.
14. Caetano N, Laureano RM, Cortez P. A data-driven approach to predict hospital length of stay - a portuguese case study. In: *ICEIS. 2014: 16th International Conference on Enterprise Information Systems*; 2014. p. 407-14. Disponível em: <https://www.scitepress.org/Papers/2014/48922/48922.pdf>
15. Peres IT, Hamacher S, Oliveira FL, Thomé AM, Bozza FA. What factors predict length of stay in the intensive care unit? Systematic review and meta-analysis. *J Crit Care*. 2020;60:183-94.