

Bárbara Vieira Carneiro¹ , Lucas Lonardoní Crozatti¹ , Pedro Vitale Mendes¹ , Antonio Paulo Nassar Júnior² , Leandro Utino Taniguchi¹ 

Comparação da acurácia de residentes, médicos seniores e decisores substitutos na previsão da mortalidade hospitalar de pacientes críticos

1. Disciplina de Emergência, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo - São Paulo (SP), Brasil.

2. Unidade de Terapia Intensiva, A.C. Camargo Cancer Center - São Paulo (SP), Brasil.

RESUMO

Objetivo: Comparar o desempenho preditivo de residentes, médicos seniores de unidades de terapia intensiva e decisores substitutos dos pacientes logo no início da internação na unidade de terapia intensiva e avaliar se diferentes apresentações de prognóstico (probabilidade de sobrevida versus probabilidade de óbito) influenciaram seus desempenhos.

Métodos: Os decisores substitutos e os médicos responsáveis pelos pacientes críticos foram questionados durante as primeiras 48 horas de internação na unidade de terapia intensiva sobre a probabilidade do desfecho hospitalar do paciente. O enquadramento da pergunta (isto é, a probabilidade de sobrevida versus a probabilidade de óbito durante a internação) foi randomizado. Para avaliar o desempenho preditivo, comparou-se a área sob a curva ROC para desfecho hospitalar entre as categorias decisores substitutos e médicos. Também estratificaram-se os resultados de acordo com o enquadramento da pergunta randomizado.

Resultados: Entrevistaram-se decisores substitutos e médicos sobre os desfechos hospitalares de 118 pacientes.

O desempenho preditivo dos decisores substitutos foi significativamente inferior ao dos médicos (área sob a curva de 0,63 para decisores substitutos, 0,82 para residentes, 0,80 para residentes de medicina intensiva e 0,81 para médicos seniores de unidade de terapia intensiva). Não houve aumento no desempenho preditivo quanto à experiência dos médicos (ou seja, médicos seniores não previram desfechos melhor que médicos juniores). Os decisores substitutos pioraram seu desempenho de previsão quando perguntados sobre a probabilidade de óbito ao invés da probabilidade de sobrevida, mas não houve diferença entre os médicos.

Conclusão: Observou-se desempenho preditivo diferente ao comparar decisores substitutos e médicos, sem qualquer efeito da experiência no prognóstico dos profissionais de saúde. O enquadramento da pergunta afetou o desempenho preditivo dos substitutos, mas não o dos médicos.

Descritores: Prognóstico; Internato e residência; Estado crítico; Tomada de decisões; Atenção à saúde; Hospitalização; Unidades de terapia intensiva; Mortalidade hospitalar

Conflitos de interesse: Nenhum.

Submetido em 26 de Agosto de 2021

Aceito em 20 de dezembro de 2021

Autor correspondente:

Leandro Utino Taniguchi
Hospital das Clínicas, Faculdade de Medicina,
Universidade de São Paulo
Av. Eneas de Carvalho Aguiar 255
CEP: 05403 - São Paulo, Brasil
E-mail: leandro.taniguchi@gmail.com

Editor responsável: Jorge Ibrain Figueira Salluh

DOI: 10.5935/0103-507X.20220019-pt

INTRODUÇÃO

A avaliação prognóstica de pacientes em estado grave é realizada na admissão à unidade de terapia intensiva (UTI) e diariamente, após a admissão, para tomar decisões sobre tratamentos invasivos e discuti-las com os decisores substitutos (geralmente um membro da família). Embora incertezas e imprecisões sejam frequentes nos prognósticos, a maioria dos decisores substitutos dos pacientes entende isso como inevitável e ainda deseja discutir os prognósticos incertos.⁽¹⁾

Entretanto, essa mesma incerteza pode impedir que alguns médicos se envolvam em discussões mais aprofundadas, especialmente aqueles em início de carreira ou em estágio. Desse modo, uma melhor compreensão da precisão dos médicos ao fazer prognósticos poderia ser esclarecedora.

Alguns estudos sugeriram que os médicos de UTI eram mais precisos que os escores prognósticos de gravidade da doença na previsão da mortalidade.⁽²⁾ Os substitutos talvez também sejam precisos, mas menos precisos que os médicos. Essa discrepância em relação ao prognóstico do paciente pode gerar e fomentar conflitos.⁽³⁾ Entretanto, a heterogeneidade desses estudos sobre o período da avaliação (de menos de 24 horas até 128 horas da admissão à UTI) e o desfecho (de mortalidade na UTI a mortalidade em 180 dias) torna essa comparação problemática.^(2,3) A influência de experiência prévia e treinamento no desempenho discriminante de médicos não foi avaliada apropriadamente. Finalmente, o formato de apresentação do prognóstico pode influenciar o decisor substituto quanto à percepção do prognóstico,⁽⁴⁾ mas é desconhecido se isso influencia a percepção do médico.

Assim, realizamos o presente estudo para comparar a capacidade prognóstica de residentes, médicos seniores de UTI, decisores substitutos e do escore de gravidade da doença durante a internação em UTI e avaliar se diferentes apresentações das informações prognósticas (probabilidade de sobrevida *versus* probabilidade de óbito) influenciam a capacidade prognóstica das categorias de médicos e decisores substitutos.

MÉTODOS

Desenho do estudo e população de pacientes

Este foi um estudo de coorte prospectivo de pacientes admitidos em UTI médico-cirúrgica (uma unidade de 14 leitos no Hospital das Clínicas, um hospital público terciário afiliado à Universidade de São Paulo, localizado em São Paulo) realizado de agosto de 2017 a dezembro de 2019. O modelo da equipe durante o dia incluía dois médicos seniores de UTI, dois residentes de medicina intensiva e seis residentes em medicina interna (que permaneceram por 1 mês na UTI durante seus estágios em medicina interna). Os critérios de inclusão foram pacientes admitidos cujos decisores substitutos pudessem ser entrevistados \leq 48 horas de sua admissão à UTI, fosse pessoa \geq 18 anos de idade e tivessem se identificado como decisores substitutos (não necessariamente um membro da família de primeiro grau). Durante o período do estudo, cada paciente poderia ser visitado por duas pessoas diariamente, das 16h às 17h, quando os médicos também as abordariam para discutir o estado clínico dos pacientes.

Os critérios de exclusão foram pacientes menores de 18 anos, grávidas e readmitidos na mesma hospitalização.

Coleta de dados e procedimentos do estudo

Incluímos pacientes cujos decisores substitutos os tivessem visitado pelo menos uma vez nas primeiras 48 horas após a admissão à UTI e tivessem discutido o estado clínico do paciente com seu médico (geralmente o residente de medicina interna e/ou o residente de medicina intensiva). Entrevistamos os participantes apenas em dias úteis (os médicos podiam ser entrevistados das 16h às 17h e os substitutos também, das 16h às 17h, durante o horário de visita). Os pesquisadores consultaram os decisores substitutos dos pacientes sobre a participação no estudo durante o período da visita. Assim como em estudo anterior, se fosse identificado mais de um possível decisor substituto durante a visita, era entrevistado aquele que se declarasse como tendo responsabilidade significativa pela tomada de decisões.⁽³⁾ Todos os substitutos e médicos participantes deram consentimento livre e esclarecido por escrito. Quanto aos dados dos pacientes, o Comitê de Ética renunciou ao consentimento livre e esclarecido, uma vez que todas as informações do estudo foram coletadas do banco de dados administrativo já existente. O Comitê de Ética local aprovou todos os procedimentos do estudo (aprovação do Comitê de Ética 2.222.797, CAAE 69174717.4.0000.0068), que foi realizado de acordo com a legislação federal (resolução 466, de 12 de dezembro de 2012, do Ministério da Saúde do Brasil).

Antes da entrevista, randomizou-se o enquadramento da pergunta (ou seja, a probabilidade de sobrevida do paciente à hospitalização ou a probabilidade de óbito do paciente durante a hospitalização) em uma sequência aleatória de duas possibilidades (ou seja, randomização simples). O procedimento da entrevista consistiu em pedir ao residente de medicina interna, ao residente de medicina intensiva e ao médico sênior da UTI responsável pelo paciente, bem como ao seu decisor substituto, que estimassem a probabilidade de sobrevida/óbito do paciente durante a hospitalização. Foi-lhes feita a pergunta validada: “Quais são as chances de sobrevida/óbito do paciente durante essa hospitalização?”⁽³⁾ Eles responderam usando uma escala de zero a 100% em estratos de 10% cada (ou seja, zero, 10%, 20% e assim por diante, até 100%).⁽³⁾ Todas as respostas foram cegadas para as respostas dos outros. O tempo para o pessoal da pesquisa obter todas as respostas não excedia 1 hora.

Foram extraídas as características basais e os desfechos dos pacientes de um banco de dados administrativo, que foi prospectivamente coletado por pessoal treinado em um banco de dados informatizado baseado em nuvem (Sistema Epimed™).⁽⁵⁾

Os dados registrados incluíam idade, sexo, data de admissão à UTI, *Simplified Acute Physiology Score III* (SAPS III),^(6,7) procedência do paciente antes da UTI, diagnóstico à admissão, procedimentos cirúrgicos antes da admissão, Índice de Comorbidade de Charlson,⁽⁸⁾ *Eastern Cooperative Oncology Group* (ECOG), utilização de recursos durante a internação na UTI (ventilação mecânica, drogas vasoativas ou terapia de reposição renal), estado oncológico (locorregional ou metastático) e mortalidade hospitalar. No momento da entrevista, também coletaram-se os dados demográficos (idade e sexo) dos médicos e dos decisores substitutos. Dos residentes e residentes de medicina intensiva, também coletaram-se informações sobre sua especialidade médica anterior, residência atual e anos desde a graduação. Todos os médicos seniores da UTI dos participantes eram intensivistas com treinamento completo por mais de 5 anos certificados pela *Associação de Medicina Intensiva Brasileira*.

Tamanho da amostra e análise estatística

Em revisão sistemática anterior, foi sugerido desempenho discriminante estimado por uma curva *receiver operating characteristic* (ROC) de 0,85 para médicos e 0,63 para escores de gravidade da doença.⁽²⁾ Estimou-se que seria necessária uma amostra de 108 entrevistas para detectar um desempenho discriminante significativo (isto é, significativamente diferente de 0,5) de 0,63, o valor mais baixo da curva ROC na revisão sistemática anterior, considerando poder estatístico de 80%, alfa de 0,05 e mortalidade hospitalar de 35% (mortalidade hospitalar da UTI estudada no ano anterior).⁽⁹⁾ Esse tamanho de amostra também é adequado para detectar uma diferença significativa entre os valores de curva ROC previstos (0,85 *versus* 0,63) com a mesma mortalidade hospitalar, considerando poder estatístico de 80% e alfa de 0,05 (nesse caso, seriam necessárias 94 entrevistas).⁽⁹⁾ Para compensar uma possível retirada de consentimento de pacientes ou substitutos, decidiu-se que deveriam ser inscritos mais 10%, para compor a estimativa de tamanho da amostra final de 118 pacientes.

As variáveis contínuas são apresentadas como mediana e intervalo interquartil (IQ) e foram avaliadas usando os testes de Wilcoxon ou de Mann-Whitney. As variáveis categóricas são apresentadas como números e percentuais e foram avaliadas usando o teste do qui-quadrado.

Para avaliar a acurácia do prognóstico de cada grupo, foram obtidas áreas sob as curvas (ASCs) ROC para desfecho hospitalar, que foram comparadas com o método DeLong sem correção para comparações múltiplas.⁽¹⁰⁾ A calibração foi avaliada pelo método cinto de calibração, conforme descrito pelo *Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva* (GiViTI).

Esse método aplica uma função logística polinomial generalizada entre o desfecho e a transformação logística da probabilidade prevista estimada, com os respectivos limites de intervalo de confiança (IC) de 95% e 80%. Um desvio estatisticamente significativo da bissetriz (a linha de calibração perfeita) ocorre quando os limites de IC95% do cinto de calibração não incluem a bissetriz.⁽¹¹⁾ Para calcular o cinto de calibração, a probabilidade de 0% foi convertida para 0,01% e a probabilidade de 100% foi convertida para 99,99%, uma vez que o método não aceita as probabilidades de 0% e 100% do questionário. Realizaram-se análise *post hoc* estratificada pelo uso de ventilação mecânica invasiva no momento da entrevista e comparação dos grupos com o prognóstico SAPS III usando a equação padrão para a probabilidade de óbito.⁽⁷⁾ As análises estatísticas foram realizadas usando o *software Statistical Package for the Social Sciences* - SPSS (IBM Corp., Armonk, NY, EUA), MedCalc para Windows, versão 19.6 (MedCalc Software, Ostend, Bélgica) e R (<http://www.r-project.org>).

RESULTADOS

Foram incluídos 118 pacientes entre agosto de 2017 e dezembro de 2019 (Tabela 1). Não houve retirada de consentimento. Na admissão, os pacientes tinham idade mediana de 54 anos (IQ 38,75 - 65,25), a mediana do SAPS III era de 53 (IQ 42,75 - 63), e 45 pacientes (38,1%) vieram a óbito no hospital. No momento da entrevista, a maioria necessitava de suporte invasivo de órgãos (46,6% em ventilação mecânica e 50,8% em uso de vasopressores), mas nenhuma discussão sobre cuidados de fim de vida havia ocorrido (Tabela 1). Dois terços dos decisores substitutos eram mulheres (74 de 118) com idade mediana de 44,5 anos (IQ 34 - 54). A maioria delas era cônjuge (31,4%) ou irmã (35,6%) (Tabela 2).

Foram entrevistados 18 residentes, oito residentes de medicina intensiva e cinco médicos seniores de UTI. Os residentes entrevistados tinham idade mediana de 26 anos (IQ 25 - 27), e a maioria (72,2%) era do sexo masculino com menos de 2 anos de formatura (84,7%). Os residentes de medicina intensiva tinham idade mediana de 28 anos (IQ 28 - 31), a maioria (75%) era do sexo masculino, 77,6% tinham menos de 4 anos de formatura, e todos tinham residência anterior em medicina interna. Os médicos seniores da UTI tinham idade mediana de 37 anos (IQ 33,5 - 41,5), a maioria (80,0%) era do sexo masculino e tinha tempo mediano desde a graduação de 12 anos (IQ 10 - 19).

Prognóstico de desfecho hospitalar

Não houve ocorrência de dados faltantes sobre os desfechos previstos. A probabilidade prevista de desfecho hospitalar diferiu entre aqueles que foram entrevistados (Figura 1 e Tabela 1S - Material suplementar).

Tabela 1 - Características dos pacientes

Idade (anos)	54 (38,75 - 65,25)
Sexo masculino	60 (50,8)
SAPS III	53 (42,75 - 63)
Tipo de admissão	
Médica	77 (65,3)
Cirurgia de emergência	39 (33,1)
Cirurgia eletiva	2 (1,7)
Tempo de internação hospitalar antes da admissão à UTI (dias)	1 (1 - 4)
Índice de Comorbidade de Charlson	2 (0 - 3,5)
<i>Eastern Cooperative Oncology Group</i>	
0 - 1	85 (72,1)
2	23 (19,5)
3 - 4	7 (5,9)
Tipo de doença em estágio avançado	
Insuficiência cardíaca crônica (NYHA IV)	19 (16,1)
Cirrose	4 (3,4)
Câncer metastático ou neoplasia hematológica	7 (5,9)
Sepse na admissão à UTI*	29 (24,6)
Ventilação mecânica na entrevista†	55 (46,6)
Vasopressores na entrevista‡	60 (50,8)
Terapia de reposição renal na entrevista‡	14 (11,9)
Decisões sobre a limitação do tratamento‡	
Nenhuma	98 (83,1)
Qualquer decisão	12 (10,2)
Mortalidade hospitalar	45 (38,1)

SAPS III - *Simplified Acute Physiology Score III*; UTI - unidade de terapia intensiva; NYHA - *New York Heart Association*.
 *A sepsis foi definida como infecção suspeita ou confirmada com falência de órgãos; † se o paciente recebeu esse suporte invasivo de órgãos quando se realizou a entrevista; ‡ faltavam os dados de oito pacientes. Resultados expressos como mediana (intervalo interquartil) ou n (%).

Puderam ser observadas duas modas para as respostas dos decisores substitutos: um percentual elevado de 100% de sobrevivência e um percentual elevado de 0% de morte. Observou-se, para os médicos, distribuição uniforme do tipo de resposta.

Em geral, todos os grupos demonstraram desempenho discriminante significativo (ou seja, diferente de 0,50), incluindo o SAPS III (Figura 2). A comparação em pares mostrou desempenho discriminante significativamente maior entre todas as categorias de médicos e decisores substitutos ($p < 0,001$), mas não entre as categorias dos médicos em si. Na presente análise *post hoc*, a presença de ventilação mecânica invasiva deteriorou o desempenho discriminante dos médicos (Figura 1SA e 1SB - Material suplementar).

Tabela 2 - Características dos decisores substitutos

Idade (anos)	44,5 (34 - 54)
Sexo masculino	44 (37,6)
Relação com o paciente	
Cônjuge	37 (31,4)
Filhos	42 (35,6)
Pais	14 (11,9)
Outro parente	25 (21,2)
Nível de educação	
Não frequentou escola	4 (3,4)
Frequentou, mas não completou a escola	38 (32,3)
Concluiu o ensino médio	45 (38,2)
Graduação	31 (26,3)
Preferência religiosa	
Católico	45 (38,1)
Evangélico	37 (31,4)
Nenhuma, agnóstico ou ateu	8 (6,8)
Espírita	9 (7,6)
Outro cristão	2 (1,7)
Outras	4 (3,4)
Não respondeu	4 (3,3)

Resultados expressos como mediana (intervalo interquartil) ou n (%).

Em todas as análises, o desempenho do SAPS III foi significativamente diferente de 0,50, mas foi diferente da estimativa de substitutos somente naquelas sem ventilação mecânica (Figura 1SB - Material Suplementar). Não houve diferença entre o prognóstico do SAPS III e quaisquer dos grupos de médicos.

A análise da calibração demonstrou que as estimativas dos médicos foram calibradas (Figura 2S - Material Suplementar). Também mostrou que as estimativas dos decisores substitutos foram mal calibradas em uma vasta faixa de probabilidades previstas. O SAPS III subestimou a mortalidade na faixa média-baixa das probabilidades previstas.

Influência do enquadramento da pergunta

O desempenho preditivo de todas as categorias de médicos foi diferente do acaso, independentemente do enquadramento da pergunta. Entretanto, quando se perguntou aos decisores substitutos sobre a probabilidade de óbito, eles não foram precisos. Os médicos seniores da UTI demonstraram ASC mais alta em ambos os enquadramentos da pergunta, quando comparados aos decisores substitutos (Tabela 3).

Tabela 3 - Área sob a curva receiver operating characteristic dos grupos entrevistados para prever sobrevivência ou óbito durante a internação

	Substituto		Residente de medicina interna		Residente de medicina intensiva		Médico sênior de UTI	
	Sobrevivência	Óbito	Sobrevivência	Óbito	Sobrevivência	Óbito	Sobrevivência	Óbito
ASC	0,69*	0,56	0,80*	0,82†*	0,80*	0,80†*	0,85‡*	0,79†*
IC95%	0,56 - 0,80	0,42 - 0,69	0,68 - 0,89	0,70 - 0,91	0,68 - 0,89	0,67 - 0,89	0,74 - 0,93	0,66 - 0,89

UTI - unidade de terapia intensiva; ASC - área sob a curva; IC95% - intervalo de confiança de 95%. * $p < 0,001$ para ASC significativamente maior que 0,50; † $p < 0,01$ em relação à previsão do substituto de óbito na internação; ‡ $p = 0,01$ em relação à previsão do substituto de sobrevivência na internação.

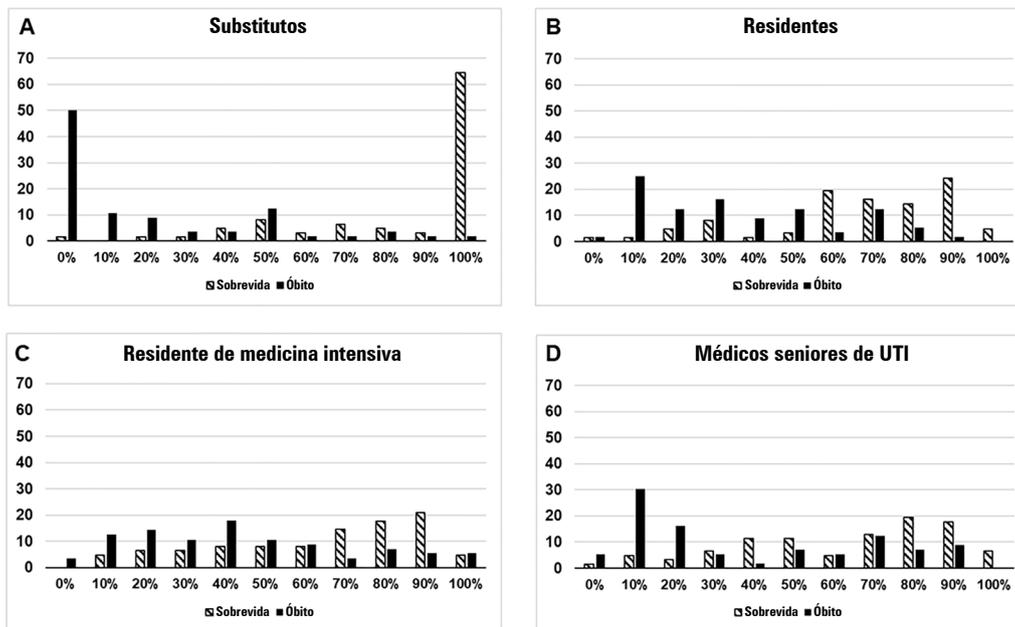
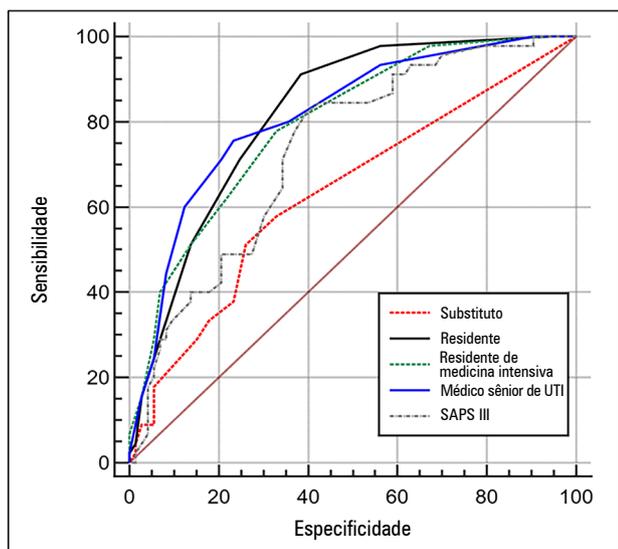


Figura 1 - Probabilidades de desfechos hospitalares estimadas para os pacientes por decisores substitutos (A), residentes de medicina interna (B), residentes de medicina intensiva (C) e médicos seniores da unidade de terapia intensiva (D).

UTI - unidade de terapia intensiva.



Grupo	ASC	IC 95%	Valor de p [†]
Substituto	0,63	0,54 - 0,72	0,016
Residente	0,82*	0,74 - 0,88	<0,001
Residente de medicina intensiva	0,80*	0,71 - 0,86	<0,001
Médico sênior de UTI	0,81*	0,73 - 0,88	<0,001
SAPS III	0,73	0,64 - 0,81	<0,001

*Significativamente diferente se comparado ao grupo substituto (p < 0,001).
 †Significância comparada a uma ASC não discriminatória (isto é, ASC de 0,50).

Figura 2 - Desempenho discriminante para desfechos hospitalares por substitutos, residentes de medicina interna, residentes de medicina intensiva, médicos seniores da unidade de terapia intensiva e *Simplified Acute Physiology Score III*.

UTI - unidade de terapia intensiva; SAPS III - *Simplified Acute Physiology Score III*; ASC - área sob a curva; IC95% - intervalo de confiança de 95%.

DISCUSSÃO

Os principais achados deste estudo são que o desempenho preditivo tanto dos decisores substitutos quanto o dos médicos em relação à mortalidade hospitalar em nossa coorte de pacientes críticos foi significativamente melhor que o acaso, mas todos os grupos de médicos (residentes de medicina interna, residentes de medicina intensiva e médicos seniores de UTI) tiveram melhor acurácia se comparados aos decisores substitutos; não houve relação entre a experiência dos médicos e melhor desempenho preditivo e diferentes enquadramentos da pergunta influenciaram o desempenho preditivo dos decisores substitutos, especialmente se perguntados sobre a possibilidade de óbito.

Sinuff et al. realizaram uma revisão sistemática de estudos que mostrou que o desempenho preditivo dos médicos era superior ao de escores de gravidade da doença em prever desfechos após admissão de pacientes individuais à UTI.⁽²⁾ Neste estudo, observou-se que o desempenho preditivo dos médicos situou-se na faixa excelente (0,80 - 0,90), como sugerido por Hosmer et al.⁽¹²⁾ Embora a maioria dos estudos seja dos anos 1990 - 2000, algumas avaliações recentes também demonstraram desempenho preditivo dos médicos semelhante ao deste estudo.⁽³⁾ Embora diferentes populações provavelmente foram estudadas (por exemplo, White et al. apenas estudaram pacientes no quinto dia de ventilação mecânica invasiva,⁽³⁾ e a revisão sistemática citada anteriormente mencionou estudos com mortalidade hospitalar variando entre 18% e 46%⁽²⁾), parece que as estimativas de prognóstico dos médicos são acuradas.

Embora não tenha sido possível demonstrar uma diferença estatisticamente significativa entre o desempenho do SAPS III e o desempenho dos médicos como descrito,⁽²⁾ isso pode ser devido à falta de poder nesta análise *post hoc*. Não obstante, em todas as condições, a estimativa pontual do desempenho dos médicos foi maior do que a observada com o SAPS III.

Não se observaram diferenças significativas entre médicos seniores e residentes juniores. Estudos anteriores não apontaram que médicos experientes em UTI demonstram desempenho preditivo superior ao de médicos inexperientes.⁽²⁾ Gusmão Vicente et al. não observaram diferenças estatisticamente significativas entre médicos seniores e juniores de UTI na previsão de internações mais longas em UTI e desfechos em UTI.⁽¹³⁾ Mesmo comparando apenas médicos em treinamento, Kruse et al. observaram a mesma curva ROC para residentes de especialidades, residentes e internos.⁽¹⁴⁾ Os presentes resultados são semelhantes, uma vez que médicos experientes (médicos seniores) e inexperientes (residentes de medicina intensiva e de medicina interna) tiveram desempenho preditivo semelhante, e todos eles foram bem calibrados.

Observou-se ainda neste estudo que o desempenho preditivo geral dos decisores substitutos foi inferior ao dos médicos. Em particular, a calibração foi insatisfatória em uma ampla faixa de estimativas previstas, como observado no cinto de calibração. Isso é digno de nota e bastante esperado, uma vez que não se espera que uma pessoa sem formação médica (que também tem sentimentos implicados pela pessoa que está representando) possa ser comparada a profissionais de saúde. No entanto, esse fato é clinicamente relevante, pois White et al. demonstraram que esse desempenho preditivo divergente está associado a divergências entre pontos de vista prognósticos.⁽³⁾ Em geral, membros de família são demasiadamente otimistas devido tanto à incompreensão das informações prognósticas transmitidas pelos médicos quanto aos vieses otimistas,^(3,15) e esses fatores podem levar a conflitos e ao uso excessivo de suporte invasivo inadequado. Assim, recomenda-se veementemente iniciar qualquer discussão séria, perguntando o ponto de vista do paciente e de seus familiares e suas percepções reais.⁽¹⁶⁾ Os presentes índices enfatizam essa eventual fonte de discordância, uma vez que também se observou que os substitutos geralmente transmitem respostas otimistas, com um alto índice de resposta “100% de chance de sobrevida” e “0% de chance de óbito”.

Um achado interessante deste estudo foi a influência do enquadramento da pergunta no desempenho preditivo dos decisores substitutos, mas não no dos médicos.

Anteriormente, foi demonstrado que o formato de apresentação de índices prognósticos poderia influenciar na percepção de risco de decisores substitutos, especialmente quando os índices são explicados qualitativamente,⁽⁴⁾ mas esse efeito não foi estudado em médicos. A maioria das publicações apresentou a pergunta considerando um quadro de sobrevida (ou seja, a probabilidade de sobrevida do paciente à hospitalização).^(14,17,18) Entretanto, o efeito do enquadramento da pergunta não foi previamente estudado. É bem conhecido que a percepção de risco é muito variável entre os sujeitos devido a múltiplos fatores.⁽¹⁹⁾ No paradigma psicométrico, o indivíduo julga quantitativamente o risco atual e o desejado.⁽²⁰⁾ Entretanto, em situações em que não existe experiência *a priori* (como admissão à UTI), é provável que as crenças e os sentimentos pessoais sofram influência do enquadramento das apresentações.⁽²⁰⁾ Acredita-se que isso deve ser considerado ao se apresentarem informações prognósticas aos substitutos e também que deve ser reconhecido ao se elaborarem futuros estudos qualitativos.

Outro achado inesperado foi a deterioração da precisão da categoria médicos quando confrontados com pacientes sob ventilação mecânica. Desconhecem-se as razões para a deterioração. Embora um erro tipo I não possa ser excluído nesta análise *post hoc*, foi demonstrada imprecisão no prognóstico para alguns aspectos de pacientes sob ventilação mecânica. Figueroa-Casas et al. demonstraram que a precisão dos médicos em prever a duração da ventilação mecânica foi limitada.⁽²¹⁾ Young et al. também sugeriram que a capacidade dos médicos de prever o prolongamento da ventilação e a necessidade de traqueostomia foi baixa.⁽²²⁾ Portanto, deve-se ter mais cuidado ao prever desfechos para pacientes sob ventilação mecânica, uma vez que pode haver algum viés de julgamento.

Este estudo apresenta algumas limitações. Primeiramente, é um estudo de um único centro, realizado em um hospital público afiliado a uma universidade. Desse modo, os resultados são influenciados pelo conjunto de casos locais de pacientes, médicos e decisores substitutos. Isso deve ser levado em consideração ao se compararem os presentes índices com outros. Em segundo lugar, não foram entrevistados outros profissionais de saúde além dos médicos para avaliar a sobrevida. Publicação anterior sugeriu que a percepção dos enfermeiros também é relevante.⁽²³⁾ Em terceiro lugar, a ausência de diferença de desempenho preditivo entre médicos seniores e residentes pode ser devido a uma redução do poder de detecção de diferenças menores do que a observada. Contudo, já foram documentados resultados semelhantes.^(13,14) Finalmente, apresenta-se a questão do prognóstico como estratos de risco numérico. O conceito numérico é crucial à compreensão correta de tal formato de apresentação^(24,25) e é particularmente relevante em países de baixa e média renda, onde os baixos níveis de instrução são infelizmente frequentes.

CONCLUSÃO

Observou-se um melhor desempenho preditivo de desfechos hospitalares entre médicos do que entre decisores substitutos, que geralmente são excessivamente otimistas. A experiência dos médicos participantes não contribuiu para o melhor desempenho preditivo. Por outro lado, o enquadramento da pergunta (ou seja, a probabilidade de óbito *versus* a probabilidade de sobrevida) influenciou o desempenho dos substitutos.

Contribuições dos autores

PV Mendes, AP Nassar Júnior e LU Taniguchi conceberam o estudo e participaram do desenho do estudo. LL Crozatti e BV Carneiro coletaram os dados. PV Mendes, AP Nassar Junior e LU Taniguchi coordenaram a análise e redigiram o manuscrito. Todos os autores leram e aprovaram o manuscrito final.

REFERÊNCIAS

1. Evans LR, Boyd EA, Malvar G, Apatira L, Luce JM, Lo B, et al. Surrogate decision-makers' perspectives on discussing prognosis in the face of uncertainty. *Am J Respir Crit Care Med.* 2009;179(1):48-53.
2. Sinuff T, Adhikari NK, Cook DJ, Schunemann HJ, Griffith LE, Rocker G, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. *Crit Care Med.* 2006;34(3):878-85.
3. White DB, Ernecoff N, Buddadhumaruk P, Hong S, Weissfeld L, Curtis JR, et al. Prevalence of and factors related to discordance about prognosis between physicians and surrogate decision makers of critically ill patients. *JAMA.* 2016;315(19):2086-94.
4. Chapman AR, Litton E, Chamberlain J, Ho KM. The effect of prognostic data presentation format on perceived risk among surrogate decision makers of critically ill patients: a randomized comparative trial. *J Crit Care.* 2015;30(2):231-5.
5. Zampieri FG, Soares M, Borges LP, Salluh JI, Ranzani OT. The Epimed Monitor ICU Database: a cloud-based national registry for adult intensive care unit patients in Brazil. *Rev Bras Ter Intensiva.* 2017;29(4):418-26.
6. Metnitz PG, Moreno RP, Almeida E, Jordan B, Bauer P, Campos RA, Iapichino G, Edbrooke D, Capuzzo M, Le Gall JR; SAPS 3 Investigators. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive Care Med.* 2005;31(10):1336-44.
7. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, Iapichino G, Edbrooke D, Capuzzo M, Le Gall JR; SAPS 3 Investigators. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med.* 2005;31(10):1345-55.
8. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40(5):373-83.
9. Goksuluk D, Korkmaz S, Zararsiz G, Karaagaoglu E. easyROC: an interactive web-tool for ROC curve analysis using R language environment. *R J.* 2016;8(2):213-30.
10. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44(3):837-45.
11. Finazzi S, Poole D, Luciani D, Cogo PE, Bertolini G. Calibration belt for quality-of-care assessment based on dichotomous outcomes. *PLoS One.* 2011;6(2):e16110.
12. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied Logistic Regression.* 3rd ed. New Jersey: John Wiley & Sons; 2013.
13. Gusmão Vicente F, Polito Lomar F, Melot C, Vincent JL. Can the experienced ICU physician predict ICU length of stay and outcome better than less experienced colleagues? *Intensive Care Med.* 2004;30(4):655-9.
14. Kruse JA, Thill-Baharozian MC, Carlson RW. Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. *JAMA.* 1988;260(12):1739-42.
15. Zier LS, Sottile PD, Hong SY, Weissfeld LA, White DB. Surrogate decision makers' interpretation of prognostic information: a mixed-methods study. *Ann Intern Med.* 2012;156(5):360-6.
16. Baile WF, Buckman R, Lenzi R, Gliber G, Beale EA, Kudelka AP. SPIKES-A six-step protocol for delivering bad news: application to the patient with cancer. *Oncologist.* 2000;5(4):302-11.
17. Brannen AL 2nd, Godfrey LJ, Goetter WE. Prediction of outcome from critical illness. A comparison of clinical judgment with a prediction rule. *Arch Intern Med.* 1989;149(5):1083-6.
18. Garrouste-Orgeas M, Montuclard L, Timsit JF, Misset B, Christias M, Carlet J. Triaging patients to the ICU: a pilot study of factors influencing admission decisions and patient outcomes. *Intensive Care Med.* 2003;29(5):774-81.
19. Slovic P. The psychology of risk. *Saude Soc.* 2010;19(4):731-47.
20. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making.* 2007;27(5):696-713.
21. Figueroa-Casas JB, Connery SM, Montoya R, Dwivedi AK, Lee S. Accuracy of early prediction of duration of mechanical ventilation by intensivists. *Ann Am Thorac Soc.* 2014;11(2): 182-5.
22. Young D, Harrison DA, Cuthbertson BH, Rowan K; TracMan Collaborators. Effect of early vs late tracheostomy placement on survival in patients receiving mechanical ventilation: the TracMan randomized trial. *JAMA.* 2013;309(20):2121-9.
23. Detsky ME, Harhay MO, Bayard DF, Delman AM, Buehler AE, Kent SA, et al. Discriminative accuracy of physician and nurse predictions for survival and functional outcomes 6 months after an ICU admission. *JAMA.* 2017;317(21):2187-95.
24. Peters E, Hart PS, Fraenkel L. Informing patients: the influence of numeracy, framing, and format of side effect information on risk perceptions. *Med Decis Making.* 2011;31(3):432-6.
25. Keller C, Siegrist M. Effect of risk communication formats on risk perception depending on numeracy. *Med Decis Making.* 2009;29(4):483-90.