

Chao-Yuan Huang¹ , Fabian Güiza Grandas^{1,2},
Marine Flechet¹ , Geert Meyfroidt^{1,2}

Modelos de predição clínica para lesão renal aguda na unidade de terapia intensiva: uma revisão sistemática

Clinical prediction models for acute kidney injury in the intensive care unit: a systematic review

1. Laboratory of Intensive Care Medicine, Academic Department of Cellular and Molecular Medicine, Katholieke Universiteit Leuven - Leuven, Belgium.

2. Department of Intensive Care Medicine, University Hospitals Leuven - Leuven, Belgium.

RESUMO

Objetivo: Relatar os modelos atualmente disponíveis de predição para o desenvolvimento de lesão renal aguda em unidades de terapia para adultos heterogêneas.

Métodos: Foi realizada revisão sistemática dos modelos de predição de lesão renal aguda em unidades de terapia intensiva para pacientes adultos. Fizemos busca na base PubMed® quanto a publicações que relatassem desenvolvimento de um novo modelo de predição, validação de um modelo estabelecido, ou impacto de um modelo de predição existente para diagnóstico precoce de lesão renal aguda em unidades de terapia intensiva.

Resultados: Triamos 583 artigos potencialmente relevantes. Dentre os 32 artigos que restaram após a primeira seleção, apenas 5 cumpriram os critérios para inclusão. As adaptações não padronizadas feitas para definir o valor inicial de creatinina sérica, quando o valor desse exame antes da admissão estava faltando, levaram a definições heterogêneas do resultado. Preditores

comumente incluídos foram sepse, idade e nível de creatinina sérica. Os modelos finais incluíram entre 5 e 19 fatores de risco. As áreas sob a curva Característica de Operação do Receptor para predição do desenvolvimento de lesão renal aguda nas coortes de validação interna variaram entre 0,78 e 0,88. Apenas dois estudos tiveram validação externa.

Conclusão: Os modelos de predição clínica de lesão renal aguda podem ajudar na aplicação em tempo apropriado de intervenções preventivas para os pacientes adequados. Contudo, poucos modelos foram externamente validados para as populações da unidade de terapia intensiva. Além disto, definições heterogêneas para lesão renal aguda e critérios de avaliação, e a falta de análise do impacto comprometem uma comparação abrangente dos modelos existentes. São necessárias novas pesquisas para validar os modelos estabelecidos e analisar seu impacto clínico, antes que estes possam ser aplicados na prática clínica.

Descritores: Lesão renal aguda; Modelos teóricos; Unidades de terapia intensiva

Conflitos de interesse: Nenhum.

Submetido em 5 de março de 2019

Aceito em 11 de julho de 2019

Autor correspondente:

Geert Meyfroidt

Department of Intensive Care Medicine

University Hospitals Leuven

Herestraat 49 - B-3000 Leuven

Belgium

E-mail: geert.meyfroidt@uzleuven.be

Editor responsável: Jorge Ibrain Figueira Salluh

DOI: 10.5935/0103-507X.20200018

INTRODUÇÃO

Pacientes críticos necessitam de suporte específico para preservar as funções de seus órgãos vitais, razão pela qual são tratados em unidades de terapia intensiva (UTI).^(1,2) A predição das futuras condições de saúde desses pacientes é crucial, e os médicos tentam prever as alterações em suas condições clínicas o mais cedo possível, com a finalidade de ajustar os tratamentos, influenciar na eventual falência de órgãos em estágio inicial e evitar resultados desfavoráveis. A lesão renal aguda (LRA) é uma das mais comuns falências de órgão em pacientes críticos e afeta cerca de 40% das admissões à UTI.⁽³⁻⁶⁾ Define-se LRA como um rápido declínio da função excretora renal. Ela é classificada em três estágios, segundo o grau de aumento na creatinina sérica (CrS) e/ou o declínio do débito urinário (DU).⁽⁷⁾ Entretanto, ambos os critérios são tardios e constituem sinais não específicos do real dano renal subjacente. Para



detecção precoce da LRA, diversos marcadores plasmáticos ou séricos e urinários foram identificados, e estão disponíveis por meio de exames diagnósticos,⁽⁸⁻¹⁵⁾ muito embora ainda seja obscuro qual população de pacientes se beneficiaria dos exames de biomarcadores, qual o momento em que tais exames deveriam ser realizados, ou quais as consequências terapêuticas que um exame positivo poderia ter na ausência de uma terapêutica específica para a LRA. Além disso, estes exames de biomarcadores acarretam certos custos financeiros.

O ambiente da UTI é rico em dados. O típico paciente crítico, com monitoramento contínuo dos sinais vitais, dispositivos terapêuticos, exames de imagem radiológica e avaliações laboratoriais seriadas, produz grandes quantidades de dados. A implantação do registro eletrônico de dados em saúde vem crescendo no mundo,⁽¹⁶⁾ tornando estes dados disponíveis para análise, e, assim, a UTI passa a ser um ambiente ideal para transformar tais dados em valiosas informações para predição. Porém, essa grande e complexa base de dados de registros em saúde representa um desafio às tradicionais técnicas estatísticas. Algoritmos de aprendizado automático, também conhecidos como aprendizado da máquina, podem oferecer um melhor entendimento das complexas variabilidade e interações entre variáveis intrincadas.⁽¹⁷⁾

O objetivo da presente revisão foi investigar quais modelos de predição internamente validados para diagnóstico de LRA em UTI heterogêneas para pacientes adultos estão atualmente disponíveis.

MÉTODOS

Estratégia de busca

Fizemos busca na base de dados PubMed® quanto a artigos de revisão e relatos completos de estudos retrospectivos e prospectivos, com texto completo disponível, publicados em inglês no período entre 1° de janeiro de 2012 e 5 de junho de 2019. Em razão da falta de uma definição unificada de LRA antes dos critérios propostos pelo *Kidney Disease: Improving Global Outcomes* (KDIGO) em 2012,⁽⁷⁾ investigamos apenas artigos publicados nos últimos 7 anos.

Utilizamos uma combinação de três termos de pesquisa, em inglês: “*intensive care unit*”, “*acute kidney injury*” e “*prediction*”. Com base em uma revisão prévia,⁽¹⁸⁾ adaptamos estas palavras para termos e descritores do sistema *Medical Subject Heading* (MeSH) em uma busca quanto aos títulos e resumos (Tabela 1). Examinamos manualmente todos os estudos incluídos. Os artigos potencialmente identificados a partir de outras fontes e referências da literatura obtida também foram incluídos no exame.

Tabela 1 – Estratégia de busca (utilização dos termos em inglês)

Termo MeSH		Título/resumo
<i>Acute kidney injury, renal replacement therapy, renal dialysis, Renal insufficiency</i>	OR	<i>acute kidney injury, renal insufficiency*, acute renal failure, renal replacement therapy, dialysis*, peritoneal dialysis*, hemodialysis*, hemodiafiltration</i>
		AND
<i>Decision support techniques</i>	OR	<i>predict* model, predict* rule, predict* score, prognosis* model, nomogram*, decision rule, risk model*, risk algorithm*, validation, risk index, risk predict*, clinical model</i>
		AND
<i>Intensive care units, critical illness</i>	OR	<i>intensive care unit*, critical ill*</i>

Fonte: adaptado de Wilson T, Quan S, Cheema K, Zarnke K, Quinn R, de Koning L, et al. Risk prediction models for acute kidney injury following major noncardiac surgery: systematic review. *Nephrol Dial Transplant*. 2016;31(2):231-40. ⁽¹⁸⁾ MeSH - *Medical Subject Headings*.

Critérios de inclusão

Inicialmente, os revisores triaram todos os resumos e títulos de artigos recuperados pela estratégia descrita. Os artigos reconhecidos como relevantes foram lidos em sua totalidade se desenvolvessem um novo modelo de predição para diagnóstico precoce de LRA na UTI, validassem um modelo estabelecido para diagnóstico precoce de LRA na UTI, ou avaliassem o impacto de um modelo preditivo existente para diagnóstico precoce de LRA na UTI. Definimos um modelo de predição para diagnóstico precoce de LRA como um sistema que utilizasse mais de um fator de risco para estimar a probabilidade do desenvolvimento de LRA.

Critérios de exclusão

Excluimos os estudos que tivessem sido realizados em pacientes fora do ambiente da UTI; tivessem sido realizados em pacientes pediátricos; não tivessem passado por validação formal, interna ou externamente;⁽¹⁹⁾ tivessem foco exclusivo em biomarcadores e não contivessem um modelo de predição para LRA; não tivessem mensurações formais do desempenho (no mínimo tivessem o relato do cálculo da área sob a curva Característica de Operação do Receptor - ASC COR/estatística c).

Extração dos dados

Utilizamos um formulário de extração de dados em conformidade com revisões e diretrizes prévias.⁽¹⁸⁻²⁰⁾ Os itens extraídos incluíam o tipo de estudo (por exemplo, estudos prospectivo, retrospectivo caso-controle ou de coorte), população (por exemplo, pacientes cardíacos, sépticos ou cirúrgicos), método de modelamento (por exemplo, regressão logística multivariada), definição de

LRA (por exemplo, *Acute Kidney Injury Network* - AKIN - ou KDIGO), método de validação interna (por exemplo, divisão aleatória, validação cruzada ou *bootstrap*), número de pacientes, tipo e número de preditores, e desempenho do modelo (por exemplo, discriminação, calibração e análise de curva decisória). Quando disponíveis, relataram-se os resultados da validação externa.

Desempenho do modelo

A avaliação do desempenho do modelo se concentrou na discriminação, na calibração e na análise da curva decisória do modelo. Discriminação se refere à capacidade de um modelo para discriminar pacientes com sintomas específicos daqueles sem tais sintomas. Uma forma comum para relatar a discriminação do modelo é observar a curva COR e relatar a ASC COR.⁽²¹⁾ Calibração se refere à proporção de concordância entre probabilidades previstas e observadas na população.⁽²²⁾ Métodos comumente utilizados para avaliar a calibração do modelo incluem calibração em geral, curva de calibração e teste de Hosmer-Lemeshow.⁽²³⁾ Um modelo bem calibrado deveria ter valor do teste Hosmer-Lemeshow acima de 0,05, curva de calibração próxima a 1 e calibração em geral próxima a zero. É também aconselhável colocar em um gráfico as proporções observadas de desfecho em oposição às probabilidades previstas, juntamente de uma indicação da curva de calibração.⁽¹⁹⁾ Análise da curva decisória se refere à quantificação do benefício líquido do modelo de predição em uma variedade de diferentes possíveis limiares de classificação.⁽²⁴⁾ Para um limiar específico, o modelo de predição com o maior benefício líquido será o mais útil na prática clínica.

RESULTADOS

Identificamos 583 artigos potencialmente relevantes, dos quais 551 foram excluídos por meio da triagem dos títulos e resumos. Lemos integralmente os 32 estudos restantes, e retivemos cinco para comparação detalhada, já que satisfaziam os critérios de inclusão (Figura 1). Na tabela 2, mostramos um resumo comparativo dos modelos de predição clínica para LRA na UTI.

Malhotra et al.⁽²⁵⁾ conduziram ensaio prospectivo multicêntrico no qual desenvolveram e validaram tanto interna quanto externamente um escore de risco para diagnóstico precoce de LRA em pacientes de UTI. Primeiramente, candidatos a preditor dentro de 48 horas após entrada na UTI foram selecionados com base em estudos prévios. Em segundo lugar, utilizou-se uma seleção *stepwise* gradual para os fatores escolhidos. Em terceiro lugar, o modelo de predição foi construído em uma coorte de desenvolvimento com utilização de uma metodologia de regressão logística multivariada com validação cruzada cinco vezes para predizer

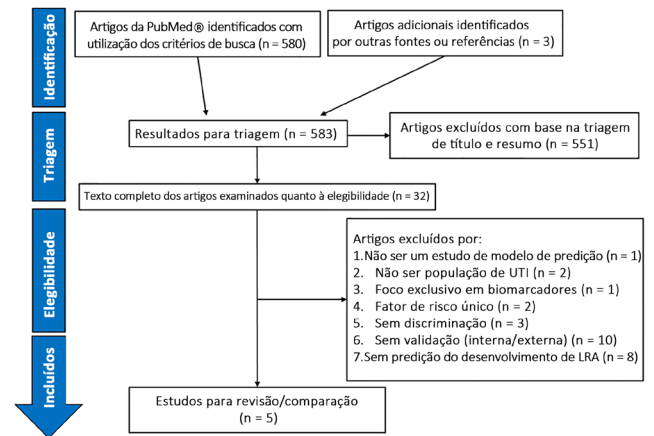


Figura 1 – Fluxograma *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) de quatro fases.⁽⁴⁰⁾ UTI - unidade de terapia intensiva; LRA - lesão renal aguda.

a ocorrência de LRA em todos seus estágios segundo os critérios KDIGO para CrS (estágios 1, 2 e 3) na primeira semana de permanência na UTI. Após isso, converteu-se o modelo de predição em regressão com base nos coeficientes para um escore baseado em risco variando de zero a 21. Os autores examinaram a possibilidade de generalizar o modelo de escore de risco por meio de validação interna do modelo em uma coorte de teste e procederam à validação externa em uma coorte independente. As ASC COR para a validação interna e externa foram, respectivamente, de 0,79 (intervalo de confiança de 95% - IC95% 0,70 - 0,89) e 0,81 (IC95% 0,78 - 0,83). A calibração foi relatada como valor para o teste Hosmer-Lemeshow de 0,293 para a validação interna, e com gráficos para as coortes de validação interna e externa. Como o valor do teste Hosmer-Lemeshow foi maior que 0,05 e houve concordância nas curvas de calibração entre as proporções observadas de resultado e probabilidades previstas, o modelo de predição foi considerado bem calibrado.

O AKI Predictor,⁽²⁶⁾ uma calculadora prognóstica *online* para diagnóstico precoce de LRA na UTI, foi criado considerando a base de dados do maior ensaio randomizado controlado multicêntrico.⁽²⁷⁾ Desenvolveram-se quatro modelos de predição com base nas informações clínicas disponíveis em diferentes momentos de avaliação: antes da admissão à UTI (modelo basal), quando da admissão à UTI (modelo admissão) e no primeiro dia após admissão à UTI (modelo dia 1), e 24 horas após a admissão à UTI (modelo dia 1+). Primeiramente, os candidatos a preditores foram considerados com base na literatura, na opinião de especialistas e sua disponibilidade na base de dados. Em segundo lugar, foi realizada uma seleção com utilização de uma análise de eliminação regressiva *bootstrap*. Em terceiro lugar, desenvolveram-se modelos em uma coorte para desenvolvimento com utilização de

Tabela 2 – Comparação de modelos de predição clínica para diagnóstico de lesão renal aguda em unidades de terapia intensiva para adultos

Característica	Subgrupo de característica	Malhotra et al. ⁽²⁵⁾	Flechet et al. ^{*(26)}	Deng et al. ^{†(28)}	Chiofalo et al. ^{‡(29)}	Zimmerman et al. ^{‡(30)}
Ano de publicação		2017	2017	2017	2019	2019
Tamanho da amostra	Coorte de desenvolvimento	573	2.123	1.084	4.572	19.160
	Coorte de validação interna	144	2.123	1.084	1.958	4.790
	Coorte de validação externa	1.300	2.367 (conjunto independente separado dos dados originais)	Não se aplica	Não se aplica	Não se aplica
Tipo de estudo	Prospectivo ou retrospectivo	Prospectivo	Retrospectivo	Prospectivo	Retrospectivo	Retrospectivo
	Centro único ou multicêntrico	Multicêntrico	Multicêntrico	Multicêntrico	Centro único	Centro único
População de pacientes		Pacientes adultos admitidos à UTI sem LRA conhecida na entrada	Pacientes adultos admitidos à UTI sem histórico de NT e CrS basal $\geq 4\text{mg/dL}$	Pacientes adultos admitidos à UTI	Pacientes adultos admitidos à UTI sem NT e histórico de LRA	Pacientes adultos admitidos à UTI sem LRC e LRA na admissão
Definição de LRA	Critérios	Critérios KDIGO CrS	Critérios KDIGO CrS	Critérios KDIGO CrS e/ ou DU	Critérios AKIN CrS e/ ou DU	Critérios CrS KDIGO
	CrS basal	Mensurações de CrS 7 a 365 dias antes da admissão à UTI	Mensurações da CrS 1 semana a 3 meses antes da admissão à UTI	Mensurações da CrS 3 a > 365 dias antes da admissão à UTI	Mensurações da CrS 180 dias antes da admissão à UTI	CrS mínima no dia 1
	CrS basal imputada	CrS na admissão à UTI	Cálculo retroativo MDRD	CrS na admissão à UTI ou seguimento até 365 dias	Cálculo retroativo MDRD	Não se aplica já que nenhum paciente teve falta da CrS basal
Incidência do desfecho (%)	Coorte de desenvolvimento	22	27,7	30,1	30	16,5
	Coorte de validação interna	24	27,7	30,1	30	16,5
	Coorte de validação externa	45	29,2 (conjunto independente separado dos dados originais)	Não se aplica	Não se aplica	Não se aplica
Valores faltantes de CrS antes da admissão (%)	Coorte de desenvolvimento	24	22,8	36,0	Não relatado	Não relatado
	Coorte de validação interna	24	22,8	36,0	Não relatado	Não relatado
	Coorte de validação externa	30	22,9 (conjunto independente separado dos dados originais)	Não se aplica	Não se aplica	Não se aplica
Tempo	Coleta de preditores de risco	Dentro de 48 horas após admissão à UTI	Dentro de 1 dia após admissão à UTI	Dentro de 1 dia após admissão à UTI	Até 1 dia antes do momento da predição	Dentro de 1 dia após admissão à UTI
	Predição	Dentro de 1 semana após entrada no estudo	Dentro de 1 semana após admissão à UTI	Dentro de 1 semana após entrada na UTI	A cada 15 minutos após admissão à UTI	Dentro de 72 horas após admissão à UTI
Number of risk factors for any AKI		10	12	5	19	16
Técnicas	Seleção de variáveis	Eliminação <i>stepwise forward</i>	Método <i>backward elimination</i>	Método <i>stepwise</i>	Algoritmo de regressão <i>stepwise</i>	Método <i>backward elimination</i>
	Modelação de algoritmo	MLR	<i>Random forest</i>	RLU e RLM	<i>Random forest</i>	RLM
	Validação interna	5-fold cross-validation	<i>Bootstrapping</i>	<i>Bootstrapping</i>	Separação ao acaso	Validação cruzada de 5 vias
	Validação externa	Independent prospective cohort	Divisão independente do conjunto de dados originais	Não se aplica	Não se aplica	Não se aplica
Discriminação (ASC COR)	Coorte de desenvolvimento	Não relatado	0,86 (0,86 - 0,86)	0,821 (0,792 - 0,850)	0,949 (0,943 - 0,954)	Não relatado
	Coorte de validação interna	0,792 (0,697 - 0,887)	0,86 (0,86 - 0,86)	0,821 (0,792 - 0,850)	0,882 (0,867 - 0,897)	0,78
	Coorte de validação externa	0,81 (0,78 - 0,83)	0,81 (0,81 - 0,81)	Não se aplica	Não se aplica	Não se aplica
Calibração	Coorte de desenvolvimento	Não relatado	Curva de calibração: 0,87 (0,87 - 0,88), Calibração geral: -0,00, e curva de calibração	Não se aplica	Valor de p no teste H-L: 0,3, e curva de calibração	Não se aplica
	Coorte de validação interna	Valor de p no teste de Hosmer-Lemeshow: 0,293, e curva de calibração	Curva de calibração: 0,87 (0,87 - 0,88) Calibração geral: -0,00 (-0,01 - 0,00) e curva de calibração	Não se aplica	Valor de p no teste de Hosmer-Lemeshow: 0,3, e curva de calibração	Não se aplica
	Coorte de validação externa	Curva de calibração	Curva de calibração: 0,78 (0,78 - 0,79) Calibração geral: -0,01 (-0,01 - -0,01), e curva de calibração	Não se aplica	Não se aplica	Não se aplica

UTI – unidade de terapia intensiva; LRA – lesão renal aguda; NT – nefropatia terminal; CrS – creatinina sérica; LRC – lesão renal crônica; KDIGO - *Kidney Disease: Improving Global Outcomes*; MDRD – *Modification of Diet in Renal Disease*; DU – débito urinário; AKIN - *Acute Kidney Injury Network*; RLM – regressão logística multivariada; RLU – regressão logística univariada; ASC COR – área sob a curva Característica de Operação do Receptor. * Dados de Flechet et al. só são relatados para o modelo de predição para qualquer lesão renal aguda no primeiro dia; † Dados de Deng et al. só são relatados para o modelo de regressão logística multivariada derivado com o método *backward elimination*; ‡ dados de Zimmerman et al. só são relatados para o modelo de regressão logística multivariada derivado com o método *backward elimination*.

um algoritmo floresta aleatória para prever a ocorrência de todos os estágios de LRA (critérios KDIGO baseados em CrS, estágios 1, 2 e 3; AKI-123) e estágios mais graves (critérios KDIGO baseados em CrS estágios 2 e 3; AKI-23) separadamente na primeira semana de permanência na UTI. Em quarto lugar, o desempenho do modelo clínico foi validado internamente por técnica de *bootstrap*. Subsequentemente, os modelos foram validados em uma coorte de validação independente não previamente vista, no qual o desempenho prognóstico foi comparado contra mensurações de lipocalina associada à gelatinase de neutrófilos (NGAL) obtidas quando da admissão à UTI. As ASC COR do modelo dia 1 para AKI-123 para as coortes de desenvolvimento e validação foram, respectivamente, de 0,86 (IC95% 0,86 – 0,86) e 0,81 (IC95% 0,81 – 0,81). A ASC COR de NGAL para AKI-123 para a coorte de validação foi de 0,67 (IC95% 0,67 – 0,67). As curvas de validação para o modelo dia 1 para AKI-123 para as coortes de desenvolvimento e validação foram, respectivamente, de 0,87 (IC95% 0,87 – 0,88) e 0,78 (IC95% 0,78 – 0,79). As curvas decisórias para o modelo dia 1 para AKI-123 mostrou utilidade clínica para limiares de classificação com ampla variação desde 4,2% até 72,4% na coorte de validação.

Dengetal.⁽²⁸⁾ conduziram estudo observacional prospectivo multicêntrico da eficácia de três biomarcadores individuais (cistatina sérica - sCysC), níveis urinários de N-acetil-β-D-glucosaminidase (uNAG) e a proporção albumina/creatinina (uACR) na urina, além de combinações do diagnóstico precoce de LRA em pacientes de UTI. Primeiramente, para prever a ocorrência de todos os estágios de LRA (critérios KDIGO de CrS estágios 1, 2, and 3; qualquer AKI) e dos estágios mais graves (critérios KDIGO de CrS estágios 2 e 3; AKI grave) separadamente na primeira semana de permanência do paciente na UTI, os autores contemplaram o desempenho dos três biomarcadores separadamente. Segundo, para melhorar o desempenho desses biomarcadores na detecção de LRA, foram criados três modelos, consistindo em diferentes combinações de biomarcadores, e examinados seus desempenhos. Em terceiro lugar, as variáveis clínicas que mostraram ser significativamente diferentes entre pacientes com e sem LRA (valor bilateral de p inferior a 0,1) na análise univariada dentro de 1 dia após admissão à UTI foram vistas como candidatas no modelo final. Em quarto lugar, utilizou-se uma seleção *stepwise* para selecionar as variáveis utilizadas nos modelos de predição. Em quinto, construíram modelos de predição para diagnóstico precoce de LRA com regressão logística multivariada, investigaram a contribuição dos biomarcadores para os modelos de predição clínica e validaram internamente o desempenho por meio de *bootstrap*. As ASC COR para os modelos multivariados de sCysC e uNAG

para qualquer AKI grave foram, respectivamente, de 0,756 (IC95% 0,723 – 0,789) e 0,863 (IC95% 0,827 – 0,900). As ASC COR para os modelos de predição para qualquer nível de AKI e AKI grave foram, respectivamente 0,821 (IC95% 0,792 – 0,850) e 0,908 (IC95% 0,881 – 0,934). As ASC COR para os modelos de predição combinados com sCysC e uNAG para qualquer nível AKI e AKI grave foram, respectivamente, 0,836 (IC95% 0,808 – 0,864) e 0,918 (IC95% 0,893 – 0,944).

Chiofolo et al.⁽²⁹⁾ conduziram estudo retrospectivo em centro único de uma coorte heterogênea de pacientes adultos admitidos à UTI, com o objetivo de desenvolver e validar um modelo de predição para o desenvolvimento de LRA em qualquer estágio (critérios AKIN com CrS e/ou DU sem inclusão de regras para diálise). Primeiramente, selecionaram preditores até 24 horas antes do momento da predição com ASC COR maior que 0,5 na análise univariada. Em segundo lugar, selecionaram preditores com base em suas ASC COR e em julgamento clínico. Em terceiro lugar, conduziu-se um algoritmo de regressão *stepwise* para confirmar a seleção de preditores. Em quarto, construíram-se modelos floresta aleatória com diferentes combinações. Em quinto lugar, o modelo final foi selecionado com base na ASC COR e percentagem de detecção precoce de LRA, tendo sido procedida à validação interna em uma coorte randomicamente separada dos dados originais. Finalmente, ajustando para diferentes limiares, o modelo de predição pôde ser utilizado para detectar qualquer estágio AKI ou apenas AKI moderada-grave. Para monitorar continuamente pacientes de UTI, o modelo de predição foi delineado para mostrar a probabilidade de desenvolver LRA a cada 15 minutos desde a admissão à UTI até a alta. As ASC COR nas coortes de desenvolvimento e validação foram, respectivamente 0,949 (IC95% 0,943 – 0,954) e 0,882 (IC95% 0,867 – 0,897). O desempenho bem calibrado foi confirmado por um teste Hosmer-Lemeshow com valor de p de 0,3. Além disto, com base no maior benefício líquido do modelo em uma ampla variedade de limiares de classificação (aproximadamente de 5% a 95%) na análise da curva decisória, verificou-se a utilidade clínica do modelo.

Zimmerman et al.⁽³⁰⁾ relataram estudo retrospectivo em centro único no qual desenvolveram e validaram modelos de predição para LRA em qualquer estágio (critérios KDIGO de CrS). Primeiramente, 1 dia após a admissão à UTI selecionaram-se fatores de risco para LRA identificados em estudos prévios, dados demográficos, clínicos e resultados de exames laboratoriais. Em segundo lugar, as variáveis com falta de mais que 20% de seus valores foram removidas, e os valores faltantes para as variáveis restantes foram imputados com a utilização de imputação multivariada por equações encadeadas (MICE). Em terceiro lugar, as variáveis que não tiveram

associação significativa com o resultado na análise univariada (valor de $p > 0,05$) foram excluídas. Em quarto, aplicou-se um método de eliminação para as variáveis significativamente associadas com o desfecho (valor de $p \leq 0,05$). Em quinto lugar, desenvolveram-se modelos para desenvolvimento de LRA dentro de 72 horas após admissão à UTI, com utilização de regressão logística multivariada, um método floresta aleatória, e uma *multilayer perceptron* com base nas variáveis selecionadas. Em sexto lugar, os modelos foram validados internamente por 10 processamentos de validação cruzada de cinco vias, e as mensurações tiveram suas médias calculadas nos 10 processamentos para modelos com e sem utilização do método de eliminação regressiva. A média das ASC COR na coorte de validação interna para regressão logística multivariada, método floresta aleatória e *multilayer perceptron* com eliminação regressiva foram 0,780, 0,772, e 0,792, enquanto as ASC COR médias sem o método de eliminação regressiva foram 0,783, 0,779, e 0,796, respectivamente.

DISCUSSÃO

Nesta revisão sistemática, identificamos cinco estudos nos quais foram construídos e validados modelos para o diagnóstico precoce de LRA em coortes heterogêneas de pacientes na UTI. Um estudo se focalizou na avaliação do desempenho com e sem a inclusão de biomarcadores em modelos de predição clínica.⁽²⁸⁾ Em quatro estudos, a definição de LRA foi feita com utilização de critérios do KDIGO,^(25,26,28,30) dentre os quais três não incluíram os critérios de DU,^(25,26,30) e um estudo utilizou critérios AKIN sem inclusão de regras de diálise.⁽²⁹⁾ As definições da CrS basal e os métodos para lidar com dados faltantes antes da UTI foram heterogêneos, impedindo comparação entre os modelos. Três modelos relataram medidas de calibração.^(25,26,29) Apenas dois modelos tiveram validação 5 cinco a 19, porém nenhum preditor foi utilizado de forma consistente em todos os estudos. Ao que sabemos, até aqui não ocorreu qualquer análise de impacto destes modelos de predição em ambiente clínico prático.

Definições AKI e severidade

Embora os modelos discutidos tenham se baseado nos mais recentes critérios KDIGO e AKI unificados,^(25,26,28,30) identificamos que as definições de LRA foram adaptadas quando não estavam disponíveis os dados necessários. Primeiramente, três estudos definiram LRA com utilização do critério de CrS do KDIGO, sem inclusão de DU.^(25,26,30) Esta definição é compreensível à vista da dificuldade de medir e registrar o DU a cada hora, o que demanda inserção manual dos dados. Contudo, a não utilização do DU, um importante componente da definição de LRA, fica aquém do ideal, já

que a modificação no DU poderia ser um marcador mais sensível de disfunção renal do que a CrS.⁽³¹⁾ Uma segunda, e talvez mais crucial questão, foi a falta de um método padronizado para definir a concentração basal de CrS, que é amplamente influenciada pela definição da CrS com base na alteração e na classificação da LRA. Para determinação da CrS basal, um estudo utilizou o nível mínimo de creatinina no dia 1,⁽³⁰⁾ enquanto os demais utilizaram valores de CrS antes da admissão. Com relação a dados de valores antes da admissão faltantes, dois estudos utilizaram a CrS da admissão,^(25,28) e os outros dois estudos estimaram a CrS basal por cálculo retroativo, utilizando a fórmula *Modification of Diet in Renal Disease* (MDRD) e assumindo uma taxa de filtração glomerular normal.^(26,29) Entretanto, nenhuma destas estimativas é isenta de problemas. Alguns estudos demonstraram que o cálculo retroativo com a MDRD pode superestimar a incidência de LRA, e o uso dos valores de CrS na admissão pode subestimá-la.⁽³²⁾

Com relação ao desfecho de predição da severidade da LRA, dois estudos investigaram todos os estágios AKI (estágios AKI 1, 2 ou 3),^(25,30) enquanto os outros dois se focalizaram em AKI moderada a grave (estágios AKI 2 ou 3) e todos os estágios AKI.^(26,28,29) A mais forte associação entre preditores e desfecho para LRA mais grave⁽³³⁾ poderia explicar o desempenho mais robusto dos modelos construídos para os escores AKI de severidade moderada a grave.

Validação, calibração e análise da curva decisória

Para evitar avaliações excessivamente otimistas do desempenho, um estudo utilizou divisão aleatória,⁽²⁹⁾ dois utilizaram validação cruzada,^(25,30) e os demais utilizaram *bootstrap*^(26,28) para a validação interna. É recomendada a utilização de *bootstrap* para estimar a validade interna, já que este método produz estimativas estáveis e reduz a ocorrência de vies.^(34,35) Dois estudos validaram externamente seus modelos de forma adequada,^(25,26) enquanto os demais não realizaram uma validação externa.⁽²⁸⁻³⁰⁾ A validação externa é uma fase necessária para que um modelo possa ser generalizável em condições previamente não vistas. Quanto mais os modelos forem validados externamente em diferentes condições, mais sua possibilidade de generalização será confiável.

Embora os cinco estudos tenham relatado boa discriminação, dois não examinaram o desempenho com relação à calibração.^(28,30) Um modelo de predição não será útil se não tiver utilidade prática, se obtiver alta discriminação mas uma calibração fraca, já que a calibração demonstra se o modelo se aplica ou não à população examinada. Mais ainda, em três estudos que incluíram uma mensuração da calibração, um estudo utilizou calibração em geral e curva de calibração,⁽²⁶⁾

e os demais utilizaram o teste Hosmer-Lemeshow com a curva de calibração.^(25,29) Notavelmente, mesmo que o teste Hosmer-Lemeshow tenha sido amplamente utilizado, ele é sensível ao tamanho da amostra e à incidência do resultado predito.⁽³⁶⁾ Assim, é necessária a cautela quando se avalia a calibração do modelo com o teste Hosmer-Lemeshow.

Além disto, dois dos cinco estudos conduziram análises da curva decisória.^(26,29) Ambos os estudos mostraram a utilidade clínica de seus modelos para uma ampla faixa de limiares de classificação. A análise da curva decisória é recomendada não apenas porque permite escolher um modelo com o maior benefício líquido, mas também porque proporciona uma faixa de limiares de classificação com utilidade clínica.

Delineamento do estudo, comparação com biomarcadores e preditores de risco comumente incluídos

Dois dos estudos foram prospectivos,^(25,28) três estudos foram multicêntricos com amostras superiores a 500 pacientes,^(25,26,28) e todos os estudos se basearam em bases de dados de uma mescla de pacientes críticos admitidos à UTI. As amostras grandes dos estudos proporcionaram compreensões cruciais do desenvolvimento de modelos de predição para o diagnóstico precoce de LRA em populações de UTI, o que complementou as revisões em populações de hospitais gerais,⁽³⁷⁾ pacientes após transplante de fígado,⁽³⁸⁾ pacientes após cirurgia cardíaca⁽³⁹⁾ e pacientes de cirurgias não cardíacas de grande porte.⁽¹⁸⁾ Os dados de estudos multicêntricos têm maior heterogeneidade, o que leva a modelos com baixas chances de ajuste excessivo. A incidência de LRA variou de 16,5% a 30,1%.

Embora tenha cumprido os critérios para inclusão por ter incluído pelo menos um modelo de predição, um dos estudos foi especificamente delineado para examinar o desempenho preditivo de biomarcadores, e não de modelos de predição.⁽²⁸⁾ Com o fim de investigar diagnóstico precoce de LRA em populações de UTI, os autores investigaram três biomarcadores (sCysC, uNAG e uACR) e suas combinações.⁽²⁸⁾ Em contraste, os outros quatro estudos se focalizaram em modelos de predição,^(25,26,29,30) dentre os quais um fez uma comparação entre modelos de predição com e sem inclusão de um biomarcador (NGAL).⁽²⁶⁾ A tabela 3 apresenta uma comparação entre biomarcadores, modelos de predição e modelos combinados no diagnóstico precoce de LRA em pacientes adultos na UTI. Nos dois estudos que compararam biomarcadores e modelos de predição,^(26,28) os modelos de predição tiveram desempenho acima dos biomarcadores, no que se refere à discriminação. É também digno de nota que, após combinar biomarcadores com os modelos de predição clínica, em ambos os estudos se relatou melhora da discriminação; contudo, ainda é discutível se o pequeno aumento na discriminação após inclusão do biomarcador compensa o aumento dos custos.

Embora todos os modelos tenham incluído uma variada metodologia de seleção, houve grande variabilidade entre os diferentes estudos com relação às variáveis finais utilizadas. Dentre as 47 variáveis identificadas nos cinco modelos de predição, nenhuma foi consistentemente utilizada em todos os estudos (Tabela 4). Entretanto, sepse, idade e nível de creatinina foram preditores comumente incluídos (presentes em mais de dois estudos).

Tabela 3 - Comparação entre biomarcadores, modelos de predição e modelos combinados para diagnóstico precoce de lesão renal aguda em unidades de terapia intensiva para pacientes adultos

Característica	Subgrupos de características	Malhotra et al. ⁽²⁵⁾	Flechet et al. ^{*(26)}	Deng et al. ^{†(28)}	Chiofalo et al. ⁽²⁹⁾	Zimmerman et al. ^{‡(30)}
Biomarcadores utilizados para comparação		Não se aplica	sNGAL	sCysC e uNAG	Não se aplica	Não se aplica
Discriminação dos biomarcadores (ASC COR)		Não se aplica	Para NGAL na coorte de validação de NGAL: 0,74 (0,74 - 0,74)	0,756 (0,723 - 0,789)	Não se aplica	Não se aplica
Discriminação dos modelos de predição (ASC COR)	Coorte de desenvolvimento	Não relatado	0,86 (0,86 - 0,86)	0,821 (0,792 - 0,850)	0,949 (0,943 - 0,954)	Não relatado
	Coorte de validação interna	0,792 (0,697 - 0,887)	0,86 (0,86 - 0,86)	0,821 (0,792 - 0,850)	0,882 (0,867 - 0,897)	0,78
	Coorte de validação externa	0,81 (0,78 - 0,83)	0,81 (0,81 - 0,81)	Não se aplica	Não se aplica	Não se aplica
Discriminação dos modelos de predição com biomarcadores (ASC COR)		Não se aplica	Para a combinação do modelo na coorte de validação de NGAL: 0,80 (0,80 - 0,80)	0,836 (0,808 - 0,864)	Não se aplica	Não se aplica

sNGAL – nível sérico de lipocalina associada a gelatinase de neutrófilo; sCysC – cistatina C no soro; uNAG – N-acetil-D-glucosaminidase urinária; ASC COR – área sob a curva Característica de Operação do Receptor; NGAL – lipocalina associada a gelatinase de neutrófilo. *Dados de Flechet et al. são relatados apenas para o modelo de predição de qualquer lesão renal aguda no primeiro dia; † dados de Deng et al. relatados apenas para o modelo de predição para qualquer lesão renal aguda; ‡ dados para Zimmerman et al. relatados apenas para o modelo de regressão logística multivariada derivado com seleção *backward*.

Tabela 4 - Fatores de risco utilizados nos modelos de predição clínica para lesão renal aguda na unidade de terapia intensiva entre os diferentes estudos

Características	Malhotra et al. ⁽²⁵⁾	Flechet et al. ^{*(26)}	Deng et al. ^{†(28)}	Chiofolo et al. ⁽²⁹⁾	Zimmerman et al. ^{‡(30)}	Total
Variáveis demográficas						
Idade		✓		✓	✓	3
Sexo			✓		✓	2
Etnia					✓	1
Variáveis crônicas						
CrS basal		✓				1
Hipertensão	✓					1
Diabetes		✓				1
Nefropatia crônica	✓					1
Hepatopatia crônica	✓		✓			2
Insuficiência cardíaca congestiva	✓					1
Doença coronária aterosclerótica	✓					1
Variáveis agudas						
Valor do pH	✓			✓		2
Ventilação mecânica	✓				✓	2
Nível de hemoglobina	✓				✓	2
Categoria cirúrgica		✓				1
Admissão planejada		✓	✓			2
Glicemia na admissão à UTI		✓				1
Suporte hemodinâmico na admissão à UTI		✓				1
Lactato máximo		✓				1
Bilirrubina		✓				1
Nível de creatinina		✓	✓		✓	3
APACHE II		✓				1
Fármacos nefrotóxicos	✓					1
Sepse	✓	✓	✓			3
Ureia plasmática				✓	✓	2
Pressão arterial diastólica não invasiva				✓		1
Temperatura				✓		1
Pressão arterial média não invasiva				✓		1
Hematócrito				✓		1
Nível de sódio				✓		1
Nível de potássio				✓	✓	2
Nível de cálcio					✓	1
TFG estimada				✓	✓	2
Débito urinário mediano após 12 horas				✓		1
Débito urinário mediano após 24 horas				✓		1
Índice de choque com base em pressão arterial diastólica não invasiva				✓		1
Índice de choque com base em pressão arterial média invasiva				✓		1
Pressão de pulso				✓		1
Volume corrente fornecido				✓		1
Proporção entre pressão parcial arterial de oxigênio e fração inspirada de oxigênio				✓		1
Balço hídrico líquido				✓		1
Dose acumulada de solução fisiológica normal				✓		1
Pressão arterial sistólica					✓	1
SpO ₂					✓	1
Nível de bicarbonato					✓	1
Contagem de plaquetas					✓	1
Tempo de tromboplastina parcial					✓	1
Proporção internacional normalizada					✓	1

CrS – creatinina sérica; UTI – unidade de terapia intensiva; APACHE – *Acute Physiology and Chronic Health Evaluation*; TFG – taxa de filtração glomerular; SpO₂: saturação do oxigênio. * Dados de Flechet et al. relatados apenas para o modelo de predição de qualquer lesão renal aguda no primeiro dia; † dados de Deng et al. relatados apenas para o modelo de predição de qualquer lesão renal aguda; ‡ dados de Zimmerman et al. relatados apenas para o modelo de regressão logística multivariada derivada com seleção *backward*.

Limitações

Primeiramente, a principal limitação desta revisão sistemática se refere ao pequeno número de estudos incluídos em razão dos rigorosos critérios de inclusão, o que pode nos ter impedido de identificar estudos menos formais, porém novos. Em segundo lugar, o protocolo da nossa revisão sistemática não foi antecipadamente publicado em uma base de dados internacional para revisões sistemáticas prospectivamente registradas, como a PROSPERO (<https://www.crd.york.ac.uk/prospero>). Contudo, procedemos à busca da literatura de forma sistemática, em conformidade com metodologia adequada baseada em diretrizes recentes.⁽⁴⁰⁾ Em terceiro lugar utilizamos apenas a base de dados PubMed® para busca de artigos potencialmente relevantes para esta revisão sistemática e, assim, podemos ter deixado de identificar manuscritos disponíveis em outras bases de dados. Em quarto lugar, só recuperamos estudos publicados nos últimos 7 anos; esta janela de tempo se justifica em razão das múltiplas definições de LRA antes da publicação dos critérios KDIGO para LRA em 2012. Ao avaliar apenas estudos concluídos após a publicação destes critérios unificados, queríamos evitar definições heterogêneas de LRA. Entretanto, ainda é necessário reconhecer que foram realizadas adaptações não padronizadas para definir o valor basal da CrS quando seu valor antes da admissão não estava disponível, o que levou a definições heterogêneas do resultado. Em quinto lugar, não recomendamos qualquer modelo em relação aos outros, já que apenas dois de cinco

estudos utilizaram uma coorte de validação externa, e nenhuma análise de impacto foi realizada. Finalmente, como não obtivemos dados dos participantes individuais destes estudos, não é possível realizar uma metanálise.

CONCLUSÃO

Ao longo dos últimos anos desenvolveram-se diversos modelos que podem prever a ocorrência de lesão renal aguda em pacientes admitidos à unidade de terapia intensiva. Como apenas dois desses modelos foram validados em uma coorte independente, seu uso clínico permanece limitado. Para tornar estes modelos suficientemente robustos para dar suporte a decisões na prática clínica, é necessário que se sigam três importantes recomendações. Primeiro, devem ser realizados estudos prospectivos cegos para uma validação independente. Em segundo lugar, como critérios de avaliação, estes estudos de validação devem utilizar análise de calibração e de curva decisória, além das análises de discriminação mais universalmente relatadas. Terceiro, para tornar os modelos mais generalizáveis, recomenda-se o uso dos critérios KDIGO para lesão renal aguda e as definições estabelecidas para definir o valor basal de creatinina sérica. É claro que a disponibilidade de grandes conjuntos de base de alta resolução combinada com novos instrumentos estatísticos e aprendizado da máquina abre oportunidades para desenvolver e validar modelos preditivos robustos para lesão renal aguda que tenham benefício potencial para cuidados ao paciente e estratificação do risco.

ABSTRACT

Objective: To report on the currently available prediction models for the development of acute kidney injury in heterogeneous adult intensive care units.

Methods: A systematic review of clinical prediction models for acute kidney injury in adult intensive care unit populations was carried out. PubMed® was searched for publications reporting on the development of a novel prediction model, validation of an established model, or impact of an existing prediction model for early acute kidney injury diagnosis in intensive care units.

Results: We screened 583 potentially relevant articles. Among the 32 remaining articles in the first selection, only 5 met the inclusion criteria. The nonstandardized adaptations that were made to define baseline serum creatinine when the preadmission value was missing led to heterogeneous definitions of the outcome. Commonly included predictors were sepsis,

age, and serum creatinine level. The final models included between 5 and 19 risk factors. The areas under the Receiver Operating Characteristic curves to predict acute kidney injury development in the internal validation cohorts ranged from 0.78 to 0.88. Only two studies were externally validated.

Conclusion: Clinical prediction models for acute kidney injury can help in applying more timely preventive interventions to the right patients. However, in intensive care unit populations, few models have been externally validated. In addition, heterogeneous definitions for acute kidney injury and evaluation criteria and the lack of impact analysis hamper a thorough comparison of existing models. Future research is needed to validate the established models and to analyze their clinical impact before they can be applied in clinical practice.

Keywords: Acute kidney injury; Models, theoretical; Intensive care units

REFERÊNCIAS

1. Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. *Lancet*. 2010;376(9749):1339-46.
2. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, et al. What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine. *J Crit Care*. 2017;37:270-6.
3. Gunst J, Vanhorebeek I, Casaer MP, Hermans G, Wouters PJ, Dubois J, et al. Impact of early parenteral nutrition on metabolism and kidney injury. *J Am Soc Nephrol*. 2013;24(6):995-1005.
4. Bellomo R, Kellum JA, Ronco C. Acute kidney injury. *Lancet*. 2012;380(9843):756-66.
5. Joannidis M, Metnitz B, Bauer P, Schusterschitz N, Moreno R, Druml W, et al. Acute kidney injury in critically ill patients classified by AKIN versus RIFLE using the SAPS 3 database. *Int Care Med*. 2009;35(10):1692-702.
6. Nisula S, Kaukonen KM, Vaara ST, Korhonen AM, Poukkanen M, Karlsson S, Haapio M, Inkinen O, Parviainen I, Suojaranta-Ylinen R, Laurila JJ, Tenhunen J, Reinikainen M, Ala-Kokko T, Ruokonen E, Kuitunen A, Pettilä V; FINNAKI Study Group. Incidence, risk factors and 90-day mortality of patients with acute kidney injury in Finnish intensive care units: The FINNAKI study. *Intensive Care Med*. 2013;39(3):420-8.
7. Khweja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract*. 2012;120(4):c179-84.
8. Ostermann M, Joannidis M. Biomarkers for AKI improve clinical practice: no. *Intensive Care Med*. 2015;41(4):618-22.
9. Bell M, Larsson A, Venge P, Bellomo R, Mårtensson J. Assessment of cell-cycle arrest biomarkers to predict early and delayed acute kidney injury. *Dis Markers*. 2015;2015:158658.
10. Prowle JR. Measurement of AKI biomarkers in the ICU: still striving for appropriate clinical indications. *Intensive Care Med*. 2015;41(3):541-3.
11. Haase-Fielitz A, Haase M, Devarajan P. Neutrophil gelatinase-associated lipocalin as a biomarker of acute kidney injury: a critical evaluation of current status. *Ann Clin Biochem*. 2014;51(Pt 3):335-51.
12. Vanmassenhove J, Vanholder R, Nagler E, Van Biesen W. Urinary and serum biomarkers for the diagnosis of acute kidney injury: an in-depth review of the literature. *Nephrol Dial Transplant*. 2013;28(2):254-73.
13. Kashani K, Al-Khafaji A, Ardiles T, Artigas A, Bagshaw SM, Bell M, et al. Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Crit Care*. 2013;17(1):R25.
14. Bihorac A, Chawla LS, Shaw AD, Al-Khafaji A, Davison DL, Demuth GE, et al. Validation of cell-cycle arrest biomarkers for acute kidney injury using clinical adjudication. *Am J Respir Crit Care Med*. 2014;189(8):932-9.
15. Hoste EA, McCullough PA, Kashani K, Chawla LS, Joannidis M, Shaw AD, Feldkamp T, Uettwiller-Geiger DL, McCarthy P, Shi J, Walker MG, Kellum JA; Sapphire Investigators. Derivation and validation of cutoffs for clinical use of cell cycle arrest biomarkers. *Nephrol Dial Transplant*. 2014;29(11):2054-61.
16. Boonstra A, Versluis A, Vos JF. Implementing electronic health records in hospitals: a systematic literature review. *BMC Health Serv Res*. 2014;14:370.
17. Flechet M, Grandas FG, Meyfroidt G. Informatics in neurocritical care: new ideas for Big Data. *Curr Opin Crit Care*. 2016;22(2):87-93.
18. Wilson T, Quan S, Cheema K, Zarnke K, Quinn R, de Koning L, et al. Risk prediction models for acute kidney injury following major noncardiac surgery: systematic review. *Nephrol Dial Transplant*. 2016;31(2):231-40.
19. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-73.
20. Hodgson LE, Sarnowski A, Roderick PJ, Dimitrov BD, Venn RM, Forni LG. Systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations. *BMJ Open*. 2017;7(9):e016591.
21. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36.
22. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115(7):928-35.
23. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: John Wiley & Sons; 2000.
24. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA*. 2015;313(4):409-10.
25. Malhotra R, Kashani KB, Macedo E, Kim J, Bouchard J, Wynn S, et al. A risk prediction score for acute kidney injury in the intensive care unit. *Nephrol Dial Transplant*. 2017;32(5):814-22.
26. Flechet M, Güiza F, Schetz M, Wouters P, Vanhorebeek I, Derese I, et al. AKIpredictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation and comparison to serum neutrophil gelatinase-associated lipocalin. *Intensive Care Med*. 2017;43(6):764-73.
27. Casaer MP, Mesotten D, Hermans G, Wouters PJ, Schetz M, Meyfroidt G, et al. Early versus late parenteral nutrition in critically ill adults. *N Engl J Med*. 2011;365(6):506-17.
28. Deng Y, Chi R, Chen S, Ye H, Yuan J, Wang L, et al. Evaluation of clinically available renal biomarkers in critically ill adults: a prospective multicenter observational study. *Crit Care*. 2017;21(1):46.
29. Chiofolo C, Chbat N, Ghosh E, Eshelman L, Kashani K. Automated continuous acute kidney injury prediction and surveillance: a random forest model. *Mayo Clin Proc*. 2019;94(5):783-92.
30. Zimmerman LP, Reyfman PA, Smith AD, Zeng Z, Zho A, Sanchez-Pinto LN, et al. Early prediction of acute kidney injury following ICU admission using a multivariate panel of physiological measurements. *BMC Med Inform Decis Mak*. 2019;19(Suppl 1):16.
31. Macedo E, Malhotra R, Claire-Del Granado R, Fedullo P, Mehta RL. Defining urine output criterion for acute kidney injury in critically ill patients. *Nephrol Dial Transplant*. 2011;26(2):509-15.
32. De Rosa S, Samoni S, Ronco C. Creatinine-based definitions: from baseline creatinine to serum creatinine adjustment in intensive care. *Crit Care*. 2016;20:69.
33. Sutherland SM, Chawla LS, Kane-Gill SL, Hsu RK, Kramer AA, Goldstein SL, Kellum JA, Ronco C, Bagshaw SM; 15 ADQI Consensus Group. Utilizing electronic health records to predict acute kidney injury risk and outcomes: workgroup statements from the 15(th) ADQI Consensus Conference. *Can J Kidney Health Dis*. 2016;3:11.
34. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-81.
35. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KG. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol*. 2003;56(5):441-7.
36. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Crit Care Med*. 2007;35(9):2052-6.
37. Hodgson LE, Sarnowski A, Roderick PJ, Dimitrov BD, Venn RM, Forni LG. Systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations. *BMJ Open*. 2017;7(9):e016591.
38. Caragata R, Wyssusek KH, Kruger P. Acute kidney injury following liver transplantation: a systematic review of published predictive models. *Anaesth Intensive Care*. 2016;44(2):251-61.
39. Huen SC, Parikh CR. Predicting acute kidney injury after cardiac surgery: a systematic review. *Ann Thorac Surg*. 2012;93(1):337-47.
40. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-9, W64.