Aprendizado de Máquina versus Regressão Estatística para Predição de Risco após Síndromes Coronarianas Agudas

Allêh Nogueira

Sumário

1 Introdução

2 Revisão de literatura

2.1 Síndrome coronariana aguda – conceitos, significância, e direções futuras

2.2 De volta para o futuro: como desenvolver modelos de predição clínica?

2.3 Uma (breve) digressão sobre aprendizado de máquina

3 Objetivos

Técnicas de inteligência artificial difundem-se progressivamente em medicina. Contudo, carece-se de estudos prospectivos que comparem seu desempenho preditivo com o de métodos estatísticos tradicionais, especialmente em cardiologia. Nosso objetivo primário consiste em comparar a capacidade discriminatória de dois métodos de modelagem preditiva – a tradicional regressão estatística, e o incipiente aprendizado de máquina –, para predição de eventos clínicos após síndromes coronarianas agudas (SCA).

Há ainda dois objetivos secundários: (i) estimar o *trade-off* teórico entre o risco de eventos cardiovasculares e hemorrágicos de cada paciente; (ii) simular as consequências clínicas da aplicação de modelos estatísticos, e baseados em aprendizado de máquina. Justifica-se (i) porque a individualização das terapias antitrombóticas em SCA permanece subestudada, embora populacionalmente demostre-se robusto benefício em sua intensificação à medida que o risco cardiovascular aumente. Justifica-se (ii) porque a demostração de superioridade discriminatória de um modelo preditivo é condição necessária, mas insuficiente, para inferir incremento na utilidade clínica das predições.

4 Métodos

4.1 Delineamento e pacientes do estudo

Nosso Registro de Síndromes Coronarianas Agudas (SCA) é uma coorte prospectiva, alocada em um hospital soteropolitano terciário. A coorte foi delineada para avaliar os desfechos de pacientes consecutivamente internados em unidade coronariana intensiva, entre setembro de 2011 e julho de 2019. Durante o recrutamento, os investigadores não interferiram na admissão dos pacientes ou condução do estudo. No momento da admissão, todos os participantes proveram consentimento livre e esclarecido por escrito. O comitê de ética em pesquisa local aprovou o protocolo do estudo sob o certificado da apresentação de apreciação ética n.° 57161016.8.0000.5544 (apêndice X). Este estudo é reportado conforme as diretrizes CONSORT[@] e TRIPOD[@].

Incluímos todos os pacientes com quadro clínico sugestivo de SCA, que possuíam ao menos um dos seguintes critérios: (1) injúria miocárdica, definida por variação dos níveis séricos de troponinas cardíacas, dado que pelo menos um valor superou o percentil 99[@]; (2) evidências eletrocardiográficas de isquemia aguda em pelo menos duas derivações contíguas[@]; (3) obstrução coronariana $\geq 70\%$ em paciente portador de síndrome coronariana crônica, e/ou ondas Q patológicas em paciente com história sugestiva de infarto prévio.

4.2 Desfechos e preditores

Derivamos modelos para predizer a ocorrência de dois desfechos: um cardiovascular e um hemorrágico. Eventos cardiovasculares adversos graves foram definidos como um composto de morte por todas as causas e reinfarto não fatal. Sangramento grave foi definido como hemorragia tipos 3 ou 5, segundo a classificação do *Bleeding Academic Research Consortium* (BARC).[@]

À admissão, médicos cardiologistas colherem os dados clínicos, que foram sucessivamente armazenados em um conjunto de dados *on-line*. Com base em conhecimentos prévios,[@] selecionamos 23 variáveis como potenciais preditores. A quantidade máxima de preditores no modelo inicial de regressão logística foi determinada pela regra geral de pelo menos 10 eventos por parâmetro preditor. Na análise de associação univariada

com o desfecho, as variáveis com menor valor de P foram escolhidas como preditores, no modelo logístico inicial. O modelo logístico final foi derivado mediante um algoritmo *stepwise* bidirecional, e otimizado para minimização do critério de informação de Akaike, que mensura o erro preditivo e penaliza o modelo de acordo com sua quantidade de preditores. Como os classificadores de aprendizado de máquina não restringem a quantidade máxima de preditores e desempenham melhor à medida que há mais dados,[@] todas as 23 variáveis selecionadas foram utilizadas como preditores.

Dentre as 23 variáveis escolhidas, 15 referiam-se a características demográficas e relativas à história clínica: idade, sexo, índice de massa corporal (IMC), uso atual de tabaco, hipertensão arterial sistêmica, diabetes mellitus, dislipidemia, doença arterial obstrutiva periférica, estenose carotídea, doença arterial coronariana (DAC) prévia, infarto prévio, história familiar de DAC prematura – definida como DAC diagnosticada em familiar de 1.° grau com idade < 55 anos, em homens, ou < 65 anos, em mulheres –, acidente vascular encefálico prévio, doença renal crônica, e uso prévio de aspirina. As demais variáveis referem-se aos exames físico e complementar: frequência cardíaca, pressão arterial sistólica, classes II-IV de Killip, alterações isquêmicas no eletrocardiograma, injúria miocárdica, NT pró-BNP sérico, creatinina sérica, e hemoglobina sérica.

4.3 Derivação dos modelos preditivos

A coorte foi aleatoriamente dividida, segundo uma proporção de 4:1, em duas subamostras: (i) a coorte de derivação inclui 80% dos pacientes, onde se derivou os modelos estatísticos de regressão e treinou-se os modelos de aprendizado de máquina; (ii) a coorte de validação interna incluiu os 20% restantes, onde se avaliou o desempenho dos modelos.

Três modelos foram desenvolvidos para predizer a ocorrência de cada desfecho: (i) o modelo estatístico tradicional empregou regressão logística binária; (ii) o modelo básico de aprendizado de máquina empregou o classificador *support vector machine* (SVM) com *kernel* de função de base radial – um classificador simples, antigo, e amplamente utilizado em predição clínica; (iii) o modelo avançado de aprendizado de máquina empregou o classificador *eXtreme gradient boosting* (XGB) – um método sofisticado, relativamente moderno, mas pouco estabelecido para predição clínica.

Os hiperparâmetros de cada modelo de aprendizado de máquina foram otimizados por *tuning* para maximizar a acurácia preditiva. O método de seleção dos hiperparâmetros diferiu entre os classificadores: *repeated cross-validations* com 10 reamostragens e 100 iterações foram utilizadas para o SVM, dada sua menor complexidade computacional; *cross-validations* com 10 reamostragens foram utilizadas para o XGB, dada sua menor complexidade computacional.

4.4 Desempenho dos modelos preditivos

O desempenho de cada modelo foi estimado por métricas de discriminação, calibração e acurácia preditiva. Discriminação é a capacidade do modelo em classificar corretamente quem desenvolverá o desfecho.[] Avaliamo-la pela estatística c de Harrell, numericamente igual à área sob a curva *receiver operator characteristic* (ROC). c consiste na probabilidade de que a risco predito pelo modelo, para um paciente que desenvolveu o desfecho, supere o risco predito para um paciente que não o desenvolveu.[] A discriminação é perfeita quando $c = 1$, e irrelevante quando $c \leq 0{,}5$.[] As discriminações dos modelos foram comparadas pelos métodos de DeLong.[@]

Calibração denota a concordância entre os desfechos observados e as predições do modelo.[] Avaliamo-la pela inclinação e intercepto da curva de calibração,[] delineadas em um plano cuja abscissa é o risco predito e a ordenada é o risco observado de desenvolver o desfecho. A inclinação da curva mensura a dispersão dos riscos estimados.[] Idealmente, a curva tem inclinação unitária; inclinações < 1 indicam estimativas de risco extremadas – isto é, pacientes em maior risco tem predições superestimadas, enquanto aqueles em menor risco tem predições subestimadas –; inclinações > 1 indicam estimativas conservadoras – isto é, pacientes em maior risco tem predições subestimadas, enquanto aqueles em maior risco tem predições superestimadas –. O intercepto em y é estimado pelo *calibration-in-the-large*, cujo valor ideal é 0. Interceptos negativos indicam que, em geral, o risco predito é subestimado; interceptos positivos indicam que, em geral, o risco predito é superestimado.

A acurácia prognóstica foi avaliada por medidas de sensibilidade (Se), especificidade (Sp), valores preditivos e razões de probabilidade. Sensibilidade, ou taxa de verdadeiros positivos, consiste na proporção de desfechos classificados corretamente pelo modelo. Especificidade, ou taxa de verdadeiros negativos, consiste na proporção de não desfechos classificados corretamente pelo modelo. Predições altamente sensíveis descartam acuradamente a ocorrência do desfecho, pois implicam baixa taxa de falsos negativos. Predições altamente específicas confirmam acuradamente a ocorrência de desfecho, pois implicam baixa taxa de falsos positivos.

Razões de verossimilhança estimam a chance pré-teste da predição estar correta, combinando medidas de sensibilidade e especificidade. A razão de probabilidade positiva (RP_+) mensura em quantas vezes o resultado do modelo aumenta o risco do paciente desenvolver o desfecho, e é calculada pela razão entre a probabilidade de verdadeiros positivos (Se) e a probabilidade de falsos-positivos ($1-Sp$). A razão de probabilidade negativa ($RP\_$) mensura em quantas vezes o resultado do modelo diminui o risco do paciente desenvolver o desfecho, e é calculada pela razão entre a probabilidade de falsos negativos ($1-Se$) e a probabilidade de verdadeiros negativos (Sp).

Valor preditivo é a probabilidade pós-teste da predição estar correta. Valor preditivo positivo (VP_+) consiste na proporção de desfechos preditos que realmente ocorreram; e valor preditivo negativo ($VP\_$), na proporção de não desfechos preditos que realmente não ocorreram. Os valores preditivos dependem da incidência de desfechos ocorridos (I). Com base no teorema de Bayes, pode-se obter valores preditivos para incidências diferentes da deste estudo ao multiplicar-se a chance de ocorrer o desfecho ($C = I/[1-I]$) pela respectiva razão de probabilidade: $VP_+ = C * RP_+$; $VP\_ = C * RP\_$. Para todas as métricas de acurácia, o ponto de corte ótimo (c^*) para classificação dicotômica foi determinado pelo índice de Youden (J) – uma estimativa de máxima acurácia que integra sensibilidade e especificidade: $J = max_c\{Se + Sp - 1\}$.

4.5 Importância dos preditores

Para cada desfecho, determinamos os principais preditores de cada modelo de aprendizado de máquina. Nos modelos baseados em SVM, a importância dos preditores foi mensurada pelo R^2 das respectivas regressões polinomiais locais (suavizadores *loess*). O R^2 representa a proporção da variância na variável dependente explicada pela variável independente, portanto seu valor aumenta à medida que a importância da variável independente cresce. Utilizou-se *loess* porque a premissa de linearidade não pode ser generalizada para todas as variáveis.

Nos modelos baseados em XGB, a importância de cada preditor foi computada pelo aumento do erro preditivo após permutar seus valores. Um preditor é considerado importante se a permutação dos seus valores diminui a capacidade preditiva do modelo – que, portanto, depende fortemente desse preditor. Como ambas as métricas de importância são relativas a cada modelo, apresentamos a importância dos preditores em escala dimensionada. As importâncias dimensionadas são obtidas dividindo-se a métrica de importância de cada variável pela métrica de importância da principal variável, de modo a produzir estimativas comparáveis e facilmente interpretáveis, que variam de 0 a 100.

Os resultados dos modelos estatísticos de regressão foram descritos por razões de chances, seus respectivos intervalos de confiança (IC) ao nível de 95%, e valores de P. A razão de chances representa o incremento no risco de ocorrer o desfecho, a cada aumento de uma unidade na escala da variável independente (para variáveis dicotômicas, 0 e 1 representam ausência e presença da condição denotada, respectivamente). Uma razão de chances de 1 denota inalteração do risco basal, que é representada pela linha de nulidade ($x = 1$), no gráfico de floresta. Embora preditores independentes sejam definidos por um valor de $P < 0{,}05$ na análise multivariada, o modelo final pode incluir preditores dependentes, pois as regressões logísticas foram otimizadas para maximizar o desempenho preditivo – em vez de modeladas para inferir causalidade.

4.6 Utilidade clínica dos modelos

A utilidade clínica dos modelos foi examinada mediante curva de análise de decisão. Essa curva estima o benefício clínico líquido para cada modelo preditivo, comparando-o às estratégias de tratar todos ou nenhum dos pacientes. O benefício líquido difere de, e complementa, as tradicionais medidas de desepenho preditivo – discriminação e calibração –, porque incorpora as consquências das decisões tomadas com base nas

predições de um modelo. Na curva de análise de decisão, o benefício líquido é calculado para um intervalo de probabilidades limiares. Define-se probabilidade limiar (p_t) como o risco de desfecho minimamente necessário para se indicar uma intervenção adicional. Sejam P_d a prevalência de uma doença, e Se a sensibilidade do modelo, logo o benefício líquido é calculado por $Se * P_d - (1 - Se) * (1 - P_d) * w$, onde $w = p_t/(1 - p_t)$ é a chance de ocorrer um desfecho na probabilidade limiar.

Na curva de decisão, a ordenada representa o benefício e a abscissa, a preferência clínica. O benefício de um modelo será máximo se ele identificar corretamente quais pacientes terão e quais não terão o desfecho. Preferência refere-se a como médicos e pacientes valoram o desfecho em questão: se o desfecho for muito importante, a probabilidade limiar tenderá a 0, pois se preferirá indicar uma intervenção em vão, em vez de deixar de prevenir um desfecho; o oposto ocorrerá caso o desfecho seja irrelevante. Em muitos cenários, a estratégia mais comum – mas não mais adequada – é intervir em todos os pacientes. O benefício líquido também pode ser expresso em termos da quantidade de intervenções fúteis evitadas ao se incorporar um modelo preditivo, em comparação com uma estratégia de intervir em todos os pacientes.

4.7 Concordância e *trade-off* entre riscos cardiovascular e hemorrágico

Para cada desfecho, selecionamos o modelo com maior discriminação numérica. As predições desses modelos foram utilizadas para classificar os pacientes em quantis de risco cardiovascular e hemorrágico. Estratificamos os pacientes em tercis de risco e avaliamos a concordância entre as classes de risco cardiovascular e hemorrágico mediante estatística κ de Cohen; valores entre 0,6 e 0,8 foram considerados indicativos de concordância substancial.

Em SCA, ocorre paradoxo risco-tratamento quando pacientes em alto risco de eventos cardiovasculares recebem intervenções antitrombóticas mais intensivas que os pacientes em baixo risco. Para examinar este fenômeno, estimamos o *trade-off* entre os riscos cardiovascular e hemorrágico para cada estrato de risco – expresso pela diferença absoluta entre os riscos observados de eventos cardiovasculares adversos graves e sangramento grave.

4.8 Análise estatística

Variáveis categóricas foram descritas por frequências absolutas e relativas. Variáveis numéricas foram descritas por média e desvio padrão (DP), caso a distribuição fosse normal, ou mediana e intervalo interquartil (IIQ), caso a distribuição fosse não normal. A normalidade foi testada avaliando-se, conjuntamente: (i) o histograma das distribuições; (ii) estatísticas de assimetria e curtose; e (iii) resultados do teste de Shapiro-Wilk. Para estimar diferenças entre grupos, utilizamos teste t para variáveis contínuas paramétricas, teste U de Mann-Whitney para variáveis contínuas não paramétricas, teste χ^2 para variáveis categóricas, e teste exato de Fischer para tabelas 2x2.

A análise primária deste estudo consiste em comparar a discriminação do modelo logístico com a de dois modelos baseados em aprendizado de máquina. Aplicamos o método de Bonferroni para duas comparações múltiplas, a fim de corrigir a inflação do risco de erro tipo I. Para tratar dados faltantes, 40 conjuntos de dados foram multiplamente imputados com 100 iterações mediante equações encadeadas.[@mice]. A significância estatística foi definida por um valor de P bicaudal inferior a 0,05. Todas as análises foram conduzidas no R, versão 4.1.0 (*R Foundation for Statistical Computing*, Viena, Austria). Os modelos de aprendizado de máquina foram treinados e otimizados utilizando-se o pacote caret.[@]

4.9 Cálculos de tamanho amostral

A coorte completou-se previamente à concepção deste estudo, evidenciando incidências de 7%, para eventos cardiovasculares adversos graves, e de 5%, para sangramentos graves. Assumimos, como hipótese nula, que as estatísticas c dos modelos preditivos igualem-se a 0,73, para o desfecho cardiovascular, e a 0,71, para o desfecho hemorrágico, pois essas são as discriminações obtidas nas amostras de validação interna para a predição de morte por todas as causas ou reinfarto, pelo escore GRACE[@] (*Global Registry of Acute Coronary Events*), e para predição de sangramento grave, pelo escore CRUSADE[@] (*Can Rapid*

risk stratification of Unstable angina patients Suppress ADverse outcomes with Early implementation of the ACC/AHA Guidelines), respectivamente.

Para o desfecho cardiovascular, estimamos que 1.017 pacientes seriam necessários para detectar uma diferença de 0,12 entre as estatísticas c, com poder de 0,80 ao nível de significância de 0,05. Para o desfecho hemorrágico, estimamos que 1.047 pacientes seriam necessários para detectar uma diferença de 0,14 entre as estatísticas c, com poder de 0,80 ao nível de significância de 0,05. Esses tamanhos amostrais foram calculados pelos métodos de Obuchowski.[@]

5 Resultados

5.1 Características clínicas basais

Apresentamos as características clínicas da população estudada na tabela 1. As tabelas 2 e 3 exibem tais características estratificadas pela ocorrência de desfechos cardiovasculares e hemorrágicos, respectivamente. As características basais dos pacientes foram bem balanceadas dentre as coortes de derivação e validação dos modelos (tabela 1). Na coorte de derivação, eventos cardiovasculares adversos graves ocorreram em 73 (7,1%) pacientes, e sangramentos graves ocorreram em 50 (4,7%) pacientes, durante o período intra-hospitalar.

Em geral, a idade média foi de 65 ± 14 anos, e a maioria dos pacientes eram homens (60% [791/1.314]) sem coronariopatia previamente documentada (60% [518/1.311]), que, à admissão, apresentaram-se sem sinais clínicos de insuficiência cardíaca aguda (87% [1.131/1.295]) a despeito de evidências de injúria miocárdica (55% [714/1.309]), e isquemia ao eletrocardiograma (54% [706/1.297]). A frequência total de dados faltantes foi de 3,5% (1.156/32.850); o apêndice A apresenta detalhadamente a distribuição e o padrão de ocorrência de dados faltantes, bem como os diagnósticos de desempenho da imputação múltipla.

5.2 Importância dos preditores

Nos modelos baseados em aprendizado de máquina, a importância dos preditores variou conforme o algoritmo classificador e o desfecho (figuras 1 e 2). Os principais preditores de eventos cardiovasculares adversos graves foram, respectivamente, classe de Killip, NT pró-BNP, idade, e hemoglobina, para o classificador SVM, e IMC, NT pró-BNP, idade, e pressão arterial sistólica, para o classificador XGB. Os principais preditores de sangramento grave foram, respectivamente, NT pró-BNP, classe de Killip, idade, e injúria miocárdica, para o classificador SVM, e NT pró-BNP, IMC, frequencia cardíaca e idade, para o classificador XGB.

Na predição estatística de eventos cardiovasculares, incluimos no modelo inicial: idade, doença renal crônica, frequência cardíaca, classe de Killip, NT pró-BNP, creatinina, e hemoglobina (tabela 2). A única variável excluída do modelo final foi doença renal crônica. Excetuando-se creatinina, os demais preditores associaram-se independentemente a eventos cardiovasculares (figura X). Na predição estatística de sangramento grave, incluimos no modelo inicial: idade, classe de Killip, injúria miocárdica, NT pró-BNP, e hemoglobina (tabela 3). A única variável excluída do modelo final foi hemoglobina. Todos os preditores associaram-se independentemente a sangramento grave (figura X).

5.3 Desempenho discriminatório

A figura X representa a discriminação dos modelos para cada desfecho, expressa pelas curvas *receiver operator characteristic* das coortes de validação. Na coorte de derivação, as estatísticas c para predição de eventos cardiovasculares foram de 0,80 (IC 95%, 0,74–0,86), para o modelo estatístico, 0,99 (IC 95%, 0,98–1,00), para o modelo SVM, e 0,95 (IC 95%, 0,93–0,97), para o modelo XGB. Contudo, quando aplicados à coorte de validação interna, tais modelos produziram estatísticas c de 0,76 (IC 95%, 0,64–0,87), 0,69 (IC 95%, 0,53–0,84), e 0,65 (IC 95%, 0,47–0,82), respectivamente. Na validação, a discriminação do modelo XGB não superou a do modelo estatístico ($P = 0{,}08$), assim como a discriminação do modelo SVM ($P = 0{,}31$). Em análise exploratória, houve indiferença discriminativa entre os classificadores SVM e XGB para predição de eventos cardiovasculares ($P = 0{,}49$).

Para predição de sangramento grave, as estatísticas c na coorte de derivação foram de 0,83 (IC 95%, 0,77–0,88), para o modelo estatístico, 1,00 (IC 95%, 1,00–1,00), para o modelo SVM e também para o modelo XGB. Contudo, quando aplicados à coorte de validação interna, tais modelos produziram estatísticas c de 0,77 (IC 95%, 0,61–0,93), 0,57 (IC 95%, 0,38–0,75), e 0,67 (IC 95%, 0,49–0,86), respectivamente. Na validação, a discriminação do modelo XGB não superou a do modelo estatístico ($P = 0{,}15$), assim como a discriminação do modelo SVM ($P = 0{,}36$). Em análise exploratória, houve indiferença discriminativa entre os classificadores SVM e XGB para predição de sangramento grave ($P = 0{,}36$).

5.4 Calibração preditiva

As figuras X e Y representam as curvas de calibração dos modelos para predição de eventos cardiovasculares adversos graves e sangramento grave, respectivamente. Ao longo dos estratos de risco de eventos cardiovasculares, as predições estatísticas foram relativamente bem valoradas (intercepto, -0,05; IC 95%, -0,61–0,50) e balanceadas (inclinação, 0,89; IC 95%, 0,34–1,45); o modelo SVM mostrou resultados similares para valoração (intercepto, -0,24; IC 95%, -0,78–0,31) e dispersão (inclinação, 1,19; IC 95%, 0,30–2,09) dos riscos preditos. Apesar de relativamente bem valoradas (intercepto, -0,06; IC 95%, -0,64–0,51), as predições de risco cardiovascular do modelo XGB foram consistentemente mais extremadas (inclinação, 0,47; IC 95%, 0,03–0,91).

As predições estatísticas para risco de sangramento grave foram relativamente bem valoradas (intercepto, 0,43; IC 95%, -0,13–0,98) e balanceadas (inclinação, 0,94; IC 95%, 0,37–1,52). O modelo SVM apresentou comportamento similar quanto à valoração (intercepto, 0,25; IC 95%, -0,30–0,81) e dispersão (inclinação, 0,46; IC 95%, -0,15–1,07) das predições. Contudo, as predições do modelo XGB foram consistentemente mais extremadas (inclinação, 0,38; IC 95%, 0,06–0,07), embora adequadamente valoradas, em média (intercepto, 1,25; IC 95%, 0,65–1,85).

5.5 Consequências clínicas dos modelos

Comparada às demais estratégias, a regressão estatística demonstrou benefício líquido consistentemente superior para probabilidades limiares de eventos cardiovasculares adversos graves entre 0% e ~5%; acima deste limite, nenhum modelo foi hegemônico. Similarmente, houve superioridade do modelo estatístico para probabilidades limiares de sangramento grave entre 0% e ~7,5%; acima deste limite, nenhum modelo foi hegemônico. Para baixas probabilidades de ambos os desfechos, o modelo XGB foi o único cujo benefício líquido foi inferior ao da estratégia de intervir em todos os pacientes. A tabela X descreve as métricas de acurácia dos modelos.

5.6 Concordância e *trade-off* entre riscos cardiovascular e hemorrágico

6 Discussão

6.1 Principais resultados

- Indiferença entre discriminações; superioridade em calibração em relação e utilidade clínica
- concordancia e tradeoff entre riscos cv e hemorragico (?)

6.2 Significância do estudo

6.2.1 Aprendizado de máquina para predição de risco: potencial ou potência?

- modelos de ML são data hungry
- Otimismo dos modelos baseados em aprendizado de máquina
- Modelos estatísticos parecem ser mais robustos para dados *low-dimensional*
- deep learning e deep phenotiping são alternativas a serem testadas

6.2.2 Predição clínica baseada em valor

- interpretação da utilidade clínica, considerando como referencia os riscos preditos para cada quantil dos escores grace e crusade

6.2.3 Transpondo o paradoxo risco-tratamento

- identificação do paradoxo neste estudo
- o paradoxo é maléfico, pois o tradeoff de pacientes de alto risco cardiovascular ainda é sempre favorável ao uso de antitrombóticos

6.3 Limitações

- dados fantantes

7 Conclusão