

# **CLASSIFICAÇÃO E ANÁLISE DE DADOS**

*Métodos e Aplicações III - CLADMap III*



**CLAD**

## **Editores**

Helena Bacelar-Nicolau

Fernanda Sousa

Carlos Marcelo

Ana Sousa Ferreira

Paulo Infante

Adelaide Figueiredo

## A importância da adequabilidade do modelo no desempenho de cartas de controlo com risco ajustado

Maria João Inácio<sup>1</sup> · Paulo Infante<sup>2</sup> · Fernanda Otilia Figueiredo<sup>3</sup>

**Resumo** Várias cartas de controlo têm sido adaptadas para uma utilização mais eficiente na área da Saúde, incorporando na análise, um modelo que estima, para cada paciente, o risco de ocorrência do acontecimento que se pretende monitorizar. A adequabilidade do modelo é muito importante no delineamento destas cartas na medida em que afeta o desempenho das mesmas. Neste trabalho, estudámos o efeito do erro de estimação e predição no desempenho das cartas de controlo CUSUM adaptadas para risco ajustado.

**Palavras-chave:** ARL, Erro de Estimação dos Parâmetros do Modelo, Monitorização de Processos de Saúde, RACUSUM.

### 1 Introdução

As cartas de controlo foram desenvolvidas por Walter Shewhart em 1920 e posteriormente popularizadas por Deming, e ao longo dos anos têm sido feitas várias adaptações às cartas existentes e desenvolvidas novas metodologias, por forma a tornar esta ferramenta mais potente e/ou mais adequada em diversas situações (Montgomery, 2013).

Apesar de criadas inicialmente com o intuito de serem utilizadas em processos industriais, as técnicas de controlo estatístico do processo têm-se revelado muito úteis em todas as áreas.

No início dos anos 90, surge a ideia de aplicar cartas de controlo na área da saúde, para monitorizar o desempenho de instituições de saúde e dos seus profissionais. No entanto, ao contrário dos processos industriais, onde as matérias primas são relativamente homogéneas, em aplicações médicas os indivíduos (pacientes) frequentemente variam muito em termos do risco de ocorrência do acontecimento que se pretende monitorizar, e, portanto, o uso das cartas de controlo

---

<sup>1</sup> Instituto Politécnico de Tomar e CIMA/IIFA, [mjantunes@ipt.pt](mailto:mjantunes@ipt.pt)

<sup>2</sup> Universidade de Évora, DMAT/ECT e CIMA/IIFA, [pinfante@uevora.pt](mailto:pinfante@uevora.pt)

<sup>3</sup> Faculdade de Economia da Universidade do Porto e CEAUL, [otilia@fep.up.pt](mailto:otilia@fep.up.pt)

tradicionais, regra geral, não é adequado para monitorizar este tipo de processos. Neste contexto, e com intuito de controlar esta fonte de variação, evitando assim erros indesejados, foram desenvolvidas as cartas de controlo com risco ajustado.

Nas cartas de controlo com risco ajustado são incorporados modelos de regressão, usualmente designados por modelos de risco, que estimam, para cada paciente, o risco de ocorrência do evento que se está a monitorizar. Tipicamente, são ajustados modelos de regressão logística baseados em scores conhecidos, tais como o *Parsonnet Score*, o *APACHE* e o *Euroscore* (Cook *et al.*, 2008; Woodall *et al.*, 2015). Os resultados observados são comparados com o que se poderia esperar com base no modelo de risco. O objetivo é detetar períodos de tempo nos quais os resultados dos cuidados de saúde não correspondem às expetativas que se tinham no momento da admissão. Pontos fora dos limites de controlo significam que o processo não se ajusta às expetativas, dando assim indicação de alterações na qualidade dos serviços prestados e consequentemente alertam para a necessidade de se proceder a alguma intervenção. Se por um lado, a deteção antecipada de um deterioramento na qualidade dos serviços prestados permite a eliminação de problemas, por outro, a deteção de uma melhoria permite a implementação e o uso de melhores práticas clínicas. Assim, a monitorização destes processos permite uma melhoria na qualidade dos serviços prestados, quer por parte das instituições quer pelos próprios profissionais de saúde.

Ao longo dos últimos anos têm sido propostas várias cartas de controlo adaptadas para risco ajustado. Atendendo a que, em contexto médico, os eventos binários são os mais frequentes, são as cartas por atributos que mais se utilizam, e se têm desenvolvido. Entre as mais populares, salientamos as cartas *Risk-Adjusted P-chart* (Alemi *et al.*, 1996), *Variable Life-Adjust Display* ou *Cumulative Risk-Adjusted Mortality* (Lovegrove *et al.*, 1997 e Poloniecki *et al.*, 1998), *Risk-Adjusted Exponentially Weighted Moving Average* (Cook *et al.*, 2011), *Risk-adjusted sequential probability ratio tests* (Spiegelhalter *et al.*, 2003) e *Risk-Adjusted CUMulative SUM control chart* (Steiner *et al.*, 2000). Para uma visão geral sobre estas cartas podem ser consultados, por exemplo, os artigos de Grigg & Farewell (2004), Cook *et al.* (2008) e Woodall *et al.* (2015).

Com exceção da carta *Risk-Adjusted P-chart*, que designaremos por carta RAP, todas as restantes são atualizadas sempre que dispomos de informação sobre um paciente. Dada a natureza dos processos a monitorizar é extremamente importante que, caso ocorram alterações no processo, estas sejam detetadas rapidamente, e, a menos que disponhamos de muitos casos, o que não é usual nestes processos, as cartas indivíduo a indivíduo são preferíveis à carta RAP, na medida em que se aumenta a hipótese de detetar mais rapidamente alterações nos cuidados de saúde.

De entre as várias cartas adaptadas para risco ajustado, a mais popular (a que tem recebido mais atenção na literatura, sendo também a mais aplicada) é a carta *Risk-Adjusted CUMulative SUM*, RACUSUM. Novick *et al.* (2006) usaram cartas CUSUM tradicionais e RACUSUM para monitorizar os resultados de cirurgia de

revascularização miocárdia, e compararam os seus resultados, tendo concluído que a RACUSUM apresenta vantagens relativamente à carta CUSUM tradicional, na medida em que esta não emite sinal de deterioração do processo quando o risco pré-operatório é elevado. Chen *et al.* (2011) aplicaram a carta RACUSUM para monitorizar a taxa de mortalidade de pacientes com paragem cardíaca assistidos pelas equipas de emergência médica. Para outras aplicações podem ver-se, por exemplo, os artigos de Harris *et al.* (2005), Axelrod *et al.* (2006), Moore *et al.* (2007) e Bottle & Aylin (2008).

O ajuste do modelo de risco é uma questão importante no delineamento das cartas de controlo com risco ajustado, na medida em que afeta o desempenho das mesmas.

Cook *et al.* (2008) apresentam algumas considerações relacionadas com o modelo de risco, referindo, por exemplo, que este deve ter uma boa calibração e capacidade de discriminação. Calibração é a capacidade do modelo estimar o risco real do acontecimento que estamos a monitorizar. Um modelo com uma fraca calibração pode alterar os resultados. Por exemplo, admita-se que o acontecimento de interesse é a morte do paciente após uma intervenção cirúrgica. Se o modelo sobestimar o risco de morte de um paciente, podemos (erradamente) concluir que a qualidade dos serviços prestados melhorou; se, por outro lado, o modelo subestimar o risco de morte do paciente somos levados a concluir (erradamente) que existe uma deterioração do processo. Discriminação é a habilidade do modelo em atribuir uma probabilidade de morte mais elevada a um paciente que morre do que a um que sobrevive. De acordo com estes autores, se a capacidade de discriminação do modelo for baixa então o modelo é pouco melhor que o acaso, independentemente da sua calibração. Assim, se o modelo não for adequado, isto é, se não estimar de forma fidedigna o risco do evento que estamos a monitorizar, toda a análise pode estar comprometida.

De acordo com a nossa pesquisa, muito pouco tem sido feito para avaliar o impacto do uso de um modelo incorreto. Steiner *et al.* (2001) avaliaram o efeito no desempenho das cartas RACUSUM, considerando, na fase de monitorização, duas distribuições extremas de pacientes: uma constituída apenas pelos pacientes de menor risco e outra constituída apenas pelos pacientes de maior risco. Assumindo que o modelo de risco inicialmente obtido era o correto, avaliaram as alterações ocorridas no *Average Run Length* (ARL), para vários valores de *odds ratio* (OR), nas duas populações consideradas, tendo concluído que estes foram claramente afetados. Tian *et al.* (2014) seguindo uma abordagem semelhante, mas utilizando apenas os dados utilizados para definir o modelo de risco, estudam as alterações que ocorrem no ARL sob controlo para 5 subpopulações diferentes. No seu estudo mostram que o ARL sob controlo varia significativamente nas 5 subpopulações, e concluem ainda que o ARL sob controlo diminui à medida que a média dos *Parsonnet scores* aumenta. Estes dados foram também trabalhados por Jones & Steiner (2012), mas utilizando uma abordagem diferente. Estes avaliaram as

alterações decorrentes no ARL sob e fora de controlo, este último assumindo uma duplicação do OR, utilizando processos de reamostragem, considerando quer amostras de dimensão diferentes, quer diferentes valores do ARL sob controlo. Avaliaram ainda o desempenho da carta RACUSUM provocando alterações no termo independente do modelo por forma a obter diferentes taxas de mortalidade. Também estes autores concluíram que o erro de estimação tem um impacto significativo no desempenho destas cartas. Loke & Gan (2012), utilizando também uma base de dados reais, investigaram a sensibilidade das cartas RACUSUM para alterações na distribuição de risco, usando distribuições beta simuladas com diferentes parâmetros. Mostraram que o ARL sob controlo poderia ser até 13% menor ou 31% maior do que o especificado.

Neste trabalho, utilizando uma base de dados real, e uma abordagem diferente das apresentadas anteriormente, avaliamos o efeito do erro de estimação dos parâmetros e do erro de predição no desempenho das cartas de controlo RACUSUM.

## 2 Cartas de Controlo CUSUM com risco ajustado (RACUSUM)

A carta de controlo de somas acumuladas binária adaptada para risco ajustado, vulgarmente designada por carta RACUSUM, foi desenvolvida por Steiner *et al.* (2000), motivado pela necessidade de monitorizar apropriadamente o desempenho cirúrgico na presença de riscos diferentes de paciente para paciente.

Uma vez que cada paciente tem um nível de risco base diferente, na carta de controlo RACUSUM as hipóteses em teste são formuladas em termos do OR, assumindo a forma:

$$H_0: OR = R_0 \quad vs \quad H_1: OR = R_1.$$

Tipicamente define-se  $R_0 = 1$ , que reflete o atual desempenho médico, i.e., o desempenho definido pelo modelo de risco estimado. Na hipótese alternativa, estabelecemos  $R_1 > R_0$  para detetar um deterioramento na qualidade dos serviços prestados e  $R_1 < R_0$  para detetar uma melhoria dos mesmos. Para  $R_0 = 1$ , é comum definir-se  $R_1 = 2$  ou  $R_1 = 0,5$ , consoante se trate, respetivamente, de uma carta RACUSUM unilateral superior ou inferior.

Representando-se por  $p_j$  a probabilidade do evento ocorrer para o paciente  $j$ , estimada pelo modelo de risco, e por  $y_j$  o resultado observado para o paciente  $j$ , definido por:

$$y_j = \begin{cases} 1 & , \text{ se o evento ocorre para o paciente } j \\ 0 & , \text{ caso contrário} \end{cases}.$$

A estatística RACUSUM unilateral superior é definida por

$$S_j^+ = \max(0, S_{j-1}^+ + w_j) \quad \text{com } S_0^+ = 0,$$

e a estatística RACUSUM unilateral inferior por

$$S_j^- = \min(0, S_{j-1}^- - w_j) \quad \text{com } S_0^- = 0,$$

onde  $w_j$  são os pesos atribuídos aos pacientes, determinados com base no logaritmo da razão de verosimilhanças, definidos por

$$w_j = \ln\left(\frac{f(y_j|R_0)}{f(y_j|R_1)}\right) = \begin{cases} \ln\left(\frac{(1-p_j+R_0p_j)R_1}{(1-p_j+R_1p_j)R_0}\right) & \text{se } y_j = 1 \\ \ln\left(\frac{1-p_j+R_0p_j}{1-p_j+R_1p_j}\right) & \text{se } y_j = 0 \end{cases}$$

$$= \ln\left[\left(\frac{1-p_j+R_0p_j}{1-p_j+R_1p_j}\right)\left(\frac{R_1}{R_0}\right)^{y_j}\right]; \quad y_j = 0,1.$$

Estes pesos terão um valor absoluto maior no caso de termos um resultado não esperado, i.e., se o evento não ocorre para um paciente com elevada probabilidade de ocorrência *à priori* ou se ocorre para um paciente com baixa probabilidade de ocorrência *à priori*. No caso do resultado ser o esperado estes pesos terão, em módulo, um valor baixo. Assim, a estatística RACUSUM unilateral superior sofre um maior aumento quando o acontecimento ocorre para um indivíduo com baixa probabilidade de risco. Por outro lado, a estatística é limitada a valores não negativos, não permitindo acumular muita informação quando ocorre uma sucessão de casos bem sucedidos, sendo assim possível detetar mais rapidamente uma alteração de deterioração na qualidade dos serviços prestados, se esta vier a ocorrer. De forma semelhante, a estatística RACUSUM unilateral inferior sofre um maior decréscimo quando o acontecimento não ocorre para um indivíduo com alta probabilidade de risco. Neste caso, é limitada a valores não positivos, por forma a não acumular muita informação quando ocorre uma sucessão de casos mal sucedidos, detetando mais rapidamente uma alteração de melhoria na qualidade dos serviços prestados, se esta vier a ocorrer.

Moustakides (1986) mostrou que a escolha ótima dos pesos  $w_j$  para a carta CUSUM tradicional é baseada no logaritmo da razão de verosimilhanças. É frequente, na literatura, ver-se justificada a escolha destes pesos na carta RACUSUM com base no mesmo resultado, afirmando-se que desta forma a carta tem um desempenho ótimo, no sentido de que de entre todas as cartas com o mesmo ARL sob controlo, a escolha do logaritmo da razão de verosimilhanças para os pesos dá o menor ARL fora de controlo. Woodall *et al.* (2015) chamam a atenção de que este resultado não está provado, na medida em que os resultados de otimização são baseados na hipótese de uma sequência de observações independentes e identicamente distribuídas, e, no nosso caso, as observações são assumidas como independentes, mas não identicamente distribuídas, já que os pacientes têm diferentes fatores de risco.

A carta RACUSUM emite sinal quando  $S_j^+ > LSC$  ou  $S_j^- < LIC$ , consoante se trate respetivamente da carta unilateral superior ou inferior. Os limites de controlo

$LIC$  e  $LSC$  são definidos de modo a que a carta tenha um determinado ARL sob controlo pré-definido. É comum, neste tipo de aplicações, considerarem-se valores bastantes elevados para o ARL sob controlo, como no caso que apresentamos neste trabalho em que a carta é desenhada para ter um ARL sob controlo na ordem dos 7000 pacientes. Vários métodos podem ser usados para determinar os limites de controlo, incluindo o uso de cadeias de Markov ou métodos de simulação de Monte Carlo.

A técnica mais utilizada para aplicação da carta RACUSUM assenta na construção das duas cartas unilaterais, atrás definidas, permitindo desta forma detetar quer melhorias quer deterioramentos na qualidade dos serviços prestados. Atendendo a que  $S_j^+$  é sempre não-negativo e  $S_j^-$  é sempre não-positivo, podemos representá-las sobre o mesmo eixo horizontal.

Grigg & Farewell (2004) referem a utilidade da fórmula definida por Khan (1984) para obter uma aproximação do ARL da carta RACUSUM bilateral:

$$\frac{1}{ARL} \cong \frac{1}{ARL_S} + \frac{1}{ARL_I}$$

onde  $ARL_S$  e  $ARL_I$  representam, respetivamente o ARL da carta RACUSUM unilateral superior e inferior. Megahed *et al.* (2011) apresentam condições sob as quais esta relação é exata.

### 3 Metodologia e Resultados

Para avaliar o impacto do uso de um modelo incorreto no desempenho da carta RACUSUM vamos utilizar uma base de dados real. Estes dados já foram trabalhados, no âmbito de cartas de controlo com risco ajustados, por Steiner *et al.* (2000, 2001), Grigg *et al.* (2003), Grigg & Farewell (2004), Gombay (2011), Jones & Steiner (2012), Tian *et al.* (2014) e Woodall (2015). Esta base de dados contém informação relativa a todas as cirurgias, realizadas num centro de cirurgia cardíaca do Reino Unido no período de 1992 a 1998, com informação relativa ao paciente e ao tipo de intervenção realizada. De acordo com Steiner *et al.* (2000), durante este período não foi realizado nenhum processo de monitorização. O acontecimento de interesse é a morte do paciente nos 30 dias após a cirurgia. No período em estudo, foram realizadas um total de 6994 cirurgias, tendo-se observado que 461 pacientes faleceram nos primeiros 30 dias após a cirurgia, o que resulta numa taxa bruta de mortalidade de aproximadamente 6,6%.

Nos dois primeiros anos (1992 a 1993) foram realizadas 2218 cirurgias tendo-se verificado no *follow-up* 143 mortes, o que corresponde a uma taxa bruta de mortalidade de aproximadamente 6,4%. Com base nestes dados estimámos a probabilidade de morte nos primeiros 30 dias após a cirurgia, usando um modelo de

regressão logística. O método de seleção usado foi o *Backward* e o modelo obtido é definido por:

$$\text{logit}(p_j) = -3,68 + 0,077X_j$$

onde  $X_j$  representa o *Parsonnet Score* para o paciente  $j$ . O facto de apenas uma variável se ter revelado estatisticamente significativa não nos surpreendeu, na medida em que o *Parsonnet Score* é um índice definido com base em características individuais, tais como idade, género, diabetes e hipertensão, variáveis estas que também estavam disponíveis na nossa base de dados.

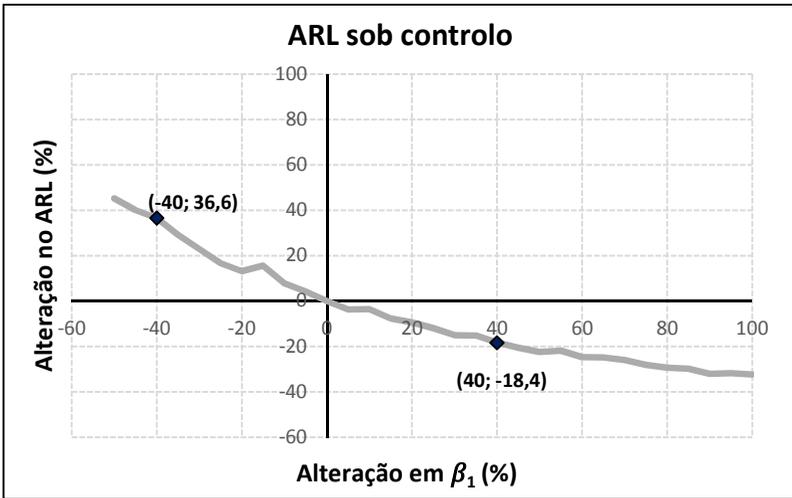
Admitimos que o processo de monitorização começou em 1994 e estudámos unicamente a carta RACUSUM unilateral superior. Provocando alterações quer no parâmetro do modelo quer na probabilidade estimada de morte, estudámos a sensibilidade destas cartas, aferindo o seu impacto no desempenho das mesmas. Neste trabalho, à semelhança do que é usual em cartas de controlo com risco ajustado, o desempenho das cartas é medido em termos do ARL, que representa o número médio de pacientes até que um sinal seja emitido. Idealmente, o ARL deve ser grande quando o processo está sob controlo e pequeno quando ocorre uma alteração significativa no processo.

A análise foi realizada no *software* R e o valor do ARL determinado via Cadeias de Markov, utilizando o *script* disponibilizado para o efeito pelo Professor Stefan Steiner na sua página pessoal (Steiner, 2001).

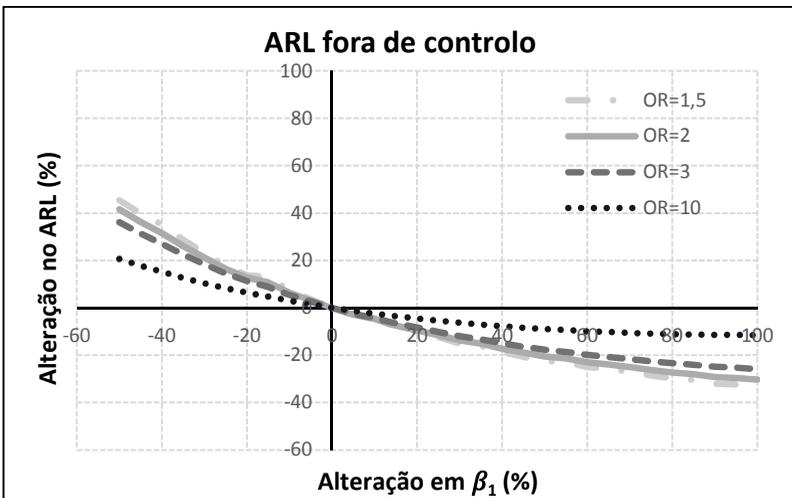
### 3.1 Erro de estimação dos parâmetros do modelo

Numa primeira fase avaliámos o efeito do erro de estimação dos parâmetros do modelo. Considerámos que a carta foi desenhada para detetar uma duplicação do OR de morte. O limite superior de controlo considerado para a carta foi de 4,5, o que se traduz num ARL sob controlo de aproximadamente 7000 pacientes. Atendendo a que por mês se efetuam aproximadamente 80 cirurgias, em média temos um falso alarme a cada 87,5 meses ( $\cong 7,2$  anos). Os ARL's fora de controlo, assumindo que o verdadeiro valor do OR é 1,5, 2, 3 e 10, são respetivamente, 533 ( $\cong 6,7$  meses), 202 ( $\cong 2,5$  meses), 91 ( $\cong 1,1$  meses) e 26 ( $\cong 10$  dias).

Provocámos alterações no parâmetro  $\beta_1$  (coeficiente da variável *Parsonnet Score*) que variaram entre -50% e 100%, considerando incrementos de 5%. Para cada uma das alterações provocadas, calculámos o ARL sob e fora de controlo, este último considerando que o verdadeiro OR era 1,5, 2, 3 e 10. Os resultados são apresentados na forma de um gráfico, onde no eixo das abcissas colocámos a alteração provocada no parâmetro e no eixo das ordenadas a correspondente alteração no ARL, ambas expressas em termos percentuais. Os resultados obtidos são apresentados nas Figuras 1 e 2.



**Figura 1** - Variação no ARL sob controle devido a alterações no parâmetro do modelo.



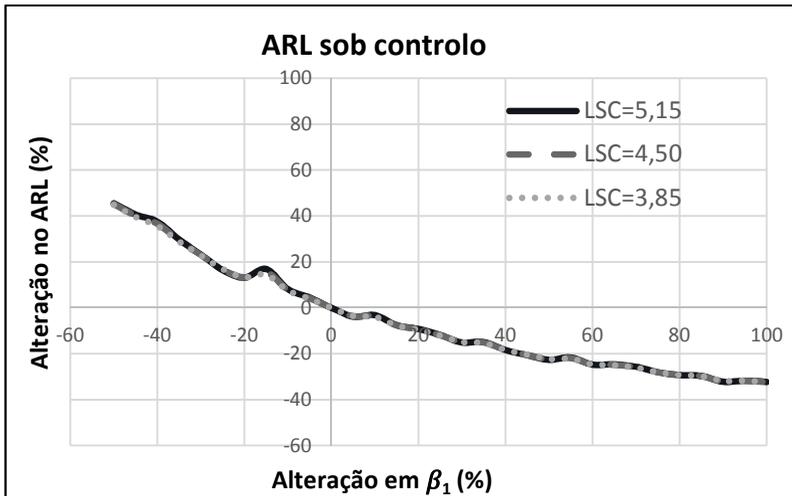
**Figura 2** - Variação no ARL fora de controle devido a alterações no parâmetro do modelo.

A análise destes gráficos permite-nos em primeiro lugar concluir que alterações no parâmetro provocam variações no ARL de sentido contrário. Podemos também observar que estimativas por defeito provocam maiores alterações no ARL do que estimativas por excesso. A título de exemplo foram assinalados dois pontos na Figura 1: podemos ver que uma diminuição de 40% em  $\beta_1$  provoca um aumento no

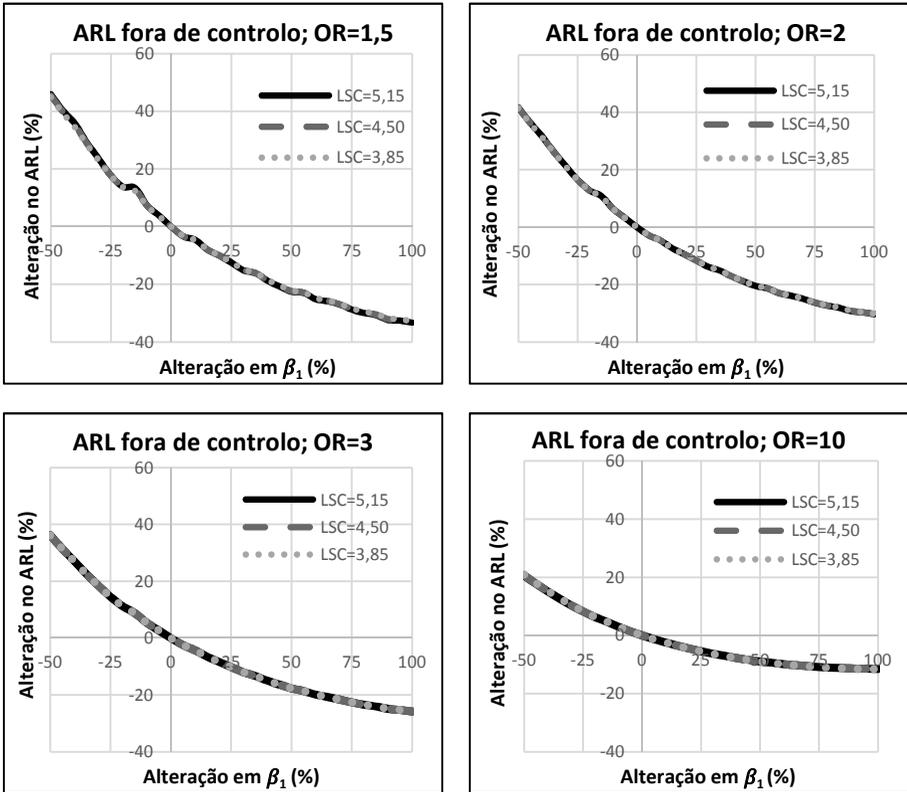
ARL na ordem dos 36,6%, enquanto que um aumento de 40% em  $\beta_1$  provoca uma diminuição no ARL na ordem dos 18,4%. A análise da Figura 2 permite-nos ainda concluir que quanto maior for o OR menores são as alterações ocorridas no ARL fora de controlo.

Estes primeiros resultados mostram-nos que a carta RACUSUM é sensível a alterações no parâmetro, mesmo quando o verdadeiro valor de OR é demasiado elevado, embora neste caso o impacto seja menor.

Com intuito de verificar se as alterações ocorridas no ARL dependiam do limite superior escolhido para a carta, o que é o mesmo que dizer, se dependiam do ARL sob controlo escolhido no delineamento da carta, e em caso afirmativo perceber como variavam, repetimos a análise anterior, mas considerando que a carta era concebida para ter um ARL sob controlo aproximadamente igual a 3500 ( $LSC=3,85$ ) e 14000 ( $LSC=5,15$ ), ou seja, respetivamente, metade e o dobro do valor considerado inicialmente. Os resultados obtidos apresentam-se nos gráficos seguintes (Figuras 3 e 4).



**Figura 3** - Comparação da variação no ARL sob controlo devido a alterações nos parâmetros do modelo, considerando cartas com diferentes limites superiores de controlo.



**Figura 4** - Comparação da variação no ARL fora de controlo devido a alterações no parâmetro do modelo, considerando cartas com diferentes limites superiores de controlo.

Como podemos observar as linhas praticamente sobrepõem-se, sendo que as reduzidas diferenças que existem se prendem com o facto de os valores serem obtidos por aproximação. Podemos assim concluir que o impacto do erro de estimação dos parâmetros no ARL da carta é independente do ARL sob controlo escolhido para a carta.

### 3.2 Erro de predição

Numa segunda fase, utilizando um procedimento análogo, avaliámos o impacto do erro de predição no desempenho da carta RACUSUM. Provocámos alterações na probabilidade de morte ( $p_j$ ) que variaram entre -50% e 100%, considerando incrementos de 5%.

Os resultados obtidos são similares aos obtidos anteriormente e naturalmente conduzem-nos às mesmas conclusões. A carta RACUSUM é sensível a alterações no modelo de risco sendo que o desempenho da carta sofre maiores alterações quando o modelo estima por defeito a probabilidade de morte do que quando estima por excesso, sendo o efeito tanto menor quanto maior é o verdadeiro OR. Além disso, também neste caso conseguimos mostrar que as alterações são independentes do ARL sob controle definido para a carta. Para melhor visualizar estes resultados, na Tabela 1 apresentamos apenas alguns dos valores obtidos nesta análise.

**Tabela 1** - Variação no ARL sob e fora de controle devido a alterações na probabilidade de morte estimada ( $p_j$ ).

Alteração na probabilidade de morte estimada ( $p_j$ )		ARL sob Controle (ARL <sub>0</sub> =7040)	ARL fora de controle		
			OR=1.5 (ARL <sub>1</sub> =533)	OR=2 (ARL <sub>1</sub> =201)	OR=3 (ARL <sub>1</sub> =91)
<b>-50%</b>	Variação (%)	11914 (69,2%)	919 (72,5%)	344 (70,9%)	152 (67,1%)
<b>-10%</b>	Variação (%)	7395 (5,0%)	568 (6,7%)	216 (7,2%)	97 (7,2%)
<b>10%</b>	Variação (%)	6566 (-6,8%)	497 (-6,8%)	188 (-6,6%)	85 (-6,2%)
<b>50%</b>	Variação (%)	5252 (-25,4%)	399 (-25,1%)	153 (-24,2%)	70 (-22,8%)
<b>100%</b>	Variação (%)	4411 (-37,3%)	336 (-37,0%)	129 (-23,9%)	60 (-33,9%)

## 4 Conclusões

O modelo de risco desempenha um papel importante no delineamento das cartas RACUSUM, na medida em que afeta o desempenho das mesmas. Neste trabalho, provocando alterações quer no parâmetro do modelo quer na probabilidade de morte estimada, avaliamos as alterações ocorridas no desempenho da carta RACUSUM, desempenho esse medido em termos dos ARL's sob e fora de controle. Observamos que a carta RACUSUM é sensível a alterações do modelo. Para ambas as situações analisadas, os resultados obtidos são concordantes e permitem-nos concluir que os ARL's diminuem à medida que aumenta a probabilidade de morte (e vice-versa); além disso, modelos que estimam por defeito a probabilidade de morte têm maior impacto no desempenho da carta do

que modelos que estimam por excesso. Foi ainda possível observar que as alterações são maiores quando ocorrem pequenas variações no processo. Por último, constatámos ainda que, o impacto foi independente do ARL sob controlo escolhido para a carta.

A monitorização da qualidade de processos de saúde é importante e sabemos que nenhum modelo é perfeito: assim, pensamos ser de grande importância o desenvolvimento de novas cartas, ou adaptações às já existentes, de modo a que estas se revelem mais robustas.

## Agradecimentos

Os autores agradecem ao professor Stefan Steiner da Universidade de Waterloo por ter facultado os dados que serviram de base a este estudo.

Este trabalho foi parcialmente financiado por fundos nacionais através da FCT - Fundação para a Ciência e a Tecnologia, projetos UID/MAT/04674/2019(CIMA) e UID/MAT/00006/2019(CEAUL).

## Referências

- ALEMI, F., ROM, W. & EISENSTEIN, E. (1996). Risk-Adjusted Control Charts for Health Care Assessment, *Annals of Operations Research*, 67, 45-60.
- AXELROD, D., GUIDINGER, M., METZGER, R., WIESNER, R., WEBB, R. & MERION, R. (2006). Transplant Center Quality Assessment Using a Continuously Updatable, Risk-Adjusted Technique (CUSUM), *American Journal of Transplantation*, 6, 313-323.
- BOTTLE, A. & AYLIN, P. (2008). Intelligent Information: A National System for Monitoring Clinical Performance, *Health Services Research*, 43, 1-31.
- CHEN, T., CHUNG, K., HU, F., FAN, C. & YANG, M. (2011). The use of statistical process control (risk-adjusted CUSUM, risk-adjusted RSPRT and CRAM with prediction limits) for monitoring the outcomes of out-of-hospital cardiac arrest patients rescued by the EMS system, *Journal of Evaluation in Clinical Practice*, 17, 71-77.
- COOK, D., COORY, M. & WEBSTER, A. (2011). Exponentially Weighted Moving Average Charts to Compare Observed and Expected Values for Monitoring Risk- Adjusted Hospital Indicators, *BMJ Quality and Safety*, 20, 469-474.
- COOK, D., DUKE, G., HART, G., PILCHER, D. & MULLANY, D. (2008). Review of the application of risk-adjusted charts to analyse mortality outcomes in critical care, *Critical Care and Resuscitation*, 10, 239-251.

- GOMBAY, E., HUSSEIN, A. & STEINER, S. (2011). Monitoring Binary Outcomes Using Risk-Adjusted Charts: A Comparative Study, *Statistics in Medicine*, 30, 2815–2826.
- GRIGG, O., FAREWELL, V. & SPIEGELHALTER, D. (2003). The Use of Risk-Adjusted CUSUM and RSPRT Charts for Monitoring in Medical Contexts, *Statistical Methods in Medical Research*, 12, 147–170.
- GRIGG, O. & FAREWELL, V. (2004). An overview of risk adjusted charts, *Journal of the Royal Statistical Society*, 167, 523-539.
- HARRIS, J., FORBES, T., STEINER, S., LAWLOR, K., DEROSE, G. & HARRIS, K. (2005). Risk-Adjusted Analysis of Early Mortality After Ruptured Abdominal Aortic Aneurysm Repair, *Journal of Vascular Surgery*, 42, 387-391.
- JONES, M. & STEINER, S. (2012). Assessing the effect of estimation error on risk-adjusted CUSUM chart performance, *International Journal for Quality in Health Care*, 24(2), 176-178.
- KHAN, R. (1984). On cumulative sum procedures and the SPRT with applications, *Journal of the Royal Statistical Society. Series B*, 46, 79–85.
- LOKE, C. & GAN, F. (2012). Joint Monitoring Scheme for Clinical Failures and Predisposed Risks, *Quality Technology and Quantitative Management*, 9(1), 3–21.
- LOVEGROVE, J., VALENCIA, O., TREASURE, T., SHERLAW-JOHNSON, C. & GALLIVAN, S. (1997). Monitoring the results of cardiac surgery by variable life-adjusted display, *Lancet*, 18, 1128-1130.
- MEGAHED, F., KENSLER, J., BEDAIR, K. & WOODALL, W. (2011). A Note on the ARL of Two-Sided Bernoulli-based CUSUM Control Charts, *Journal of Quality Technology*, 43(1), 43–49.
- MONTGOMERY, D.C. (2013). *Introduction to Statistical Quality Control*, 7th Edition. John Wiley & Sons, Inc, New York.
- MOORE, R., NUTLEY, M., CINA, C., MOTAMEDI, M., FARIS, P. & ABUZNADAH, W. (2007). Improved Survival After Introduction of an Emergency Endovascular Therapy Protocol for Ruptured Abdominal Aortic Aneurysms, *Journal of Vascular Surgery*, 4, 443-450.
- MOUSTAKIDES, G. (1986). Optimal Stopping Times for Detecting Changes in Distribution, *The Annals of Statistics*, 14, 1379–1387.
- NOVICK, R., FOX, S., STITT, L., FORBES, T. & STEINER, S. (2006). Direct comparison of risk-adjusted and non-risk-adjusted CUSUM analyses of coronary artery bypass surgery outcomes, *The Journal of Thoracic and Cardiovascular Surgery*, 132(2), 386-391.
- POLONIECKI, J., VALENCIA, O. & LITTLEJOHNS, P. (1998). Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery, *British Medical Journal*, 316, 1697-700.

- R CORE TEAM (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.r-project.org/>.
- SPIEGELHALTER, D., GRIGG, O., KINSMAN, R. & TREASURE, T. (2003). Sequential probability ratio tests: applications to Bristol, Shipmann and adult cardiac surgery, *International Journal for Quality Health Care*, 15, 1-7.
- STEINER, S., COOK, R., FAREWELL, V. & TREASURE, T. (2000). Monitoring Surgical Performance Using Risk Adjusted Cumulative Sum Charts, *Biostatistics*, 1, 441-452.
- STEINER, S., COOK, R. & FAREWELL, V. (2001). Risk-Adjusted Monitoring of Binary Surgical Outcomes, *Medical Decision Making*, 21(3), 163–169.
- STEINER, S. (2001). Software to support risk-adjusted CUSUM charts: Findarl.r. URL [https://uwaterloo.ca/statistics-and-actuarial-science/sites/ca.statistics\\_and-actuarial-science/files/uploads/files/findarl.r](https://uwaterloo.ca/statistics-and-actuarial-science/sites/ca.statistics_and-actuarial-science/files/uploads/files/findarl.r) (acedido em 12 de Janeiro de 2016).
- TIAN, W., SUN, H., ZANG, X. & WOODALL, W. (2014). The impact of varying patient populations on the in-control performance of the risk-adjusted CUSUM chart, *International Journal for Quality in Health Care*, 27(1), 31-36.
- WOODALL, W., FOGEL, S. & STEINER, S. (2015). The Monitoring and Improvement of Surgical-Outcome Quality, *Journal of Quality Technology*, 47(4), 383-399.