

Estudo dinâmico da mortalidade por tuberculose no Estado de São Paulo, Brasil: uma abordagem bayesiana

Dynamic study of tuberculosis mortality in São Paulo State, Brazil: a Bayesian approach

James Dean Oliveira dos Santos Junior ¹
Basílio de Bragança Pereira ²

Abstract

This study used dynamic linear models to analyze the dynamics of tuberculosis (TB) mortality in São Paulo State, Brazil. The study confirms previous findings of a decline in TB mortality in State in the last decade, while pointing to an increase in deaths beginning in 2005. Analysis of the two most frequent underlying causes pointed to HIV/TB co-infection. Forecasts for the next two years were performed, providing little evidence to support a continuing downward trend in the number of TB-related deaths.

Tuberculosis; Mortality; Validity of Tests; Reproducibility of Tests

Introdução

Em razão de diversos sintomas, a tuberculose (TB) só foi identificada como doença na segunda metade do século XIX. Embora exista cura, ela ainda aflige a humanidade, permanecendo um problema de saúde pública, particularmente em países em desenvolvimento. Mesmo que houvesse a perspectiva de que a doença poderia ser erradicada, a partir da década de 90, ocorreu um aumento significativo de casos por conta da coinfecção com o HIV. Em 1993, a Organização Mundial da Saúde (OMS) declarou estado de emergência global.

No Brasil, em 2006, a incidência de TB foi estimada em 50 casos por 100 mil habitantes, enquanto a prevalência foi estimada em 55 casos por 100 mil habitantes. O Relatório Global de Controle da Tuberculose de 2008 ¹ coloca o Brasil na 16ª posição entre os países com maior incidência da doença. Além do advento da AIDS, outros fatores contribuem para o agravamento do quadro da TB no Brasil, como o empobrecimento da população e o sucateamento da Rede Pública de Assistência à Saúde ². Outro agravante tem sido o elevado percentual de abandono do tratamento da TB que, nos últimos vinte anos, tem-se mantido em torno de 14% ³. O abandono do tratamento pode levar ao desenvolvimento da TB multirresistente, associada à elevada taxa de mortalidade ⁴. Os altos níveis de TB multirresistente detectados no estudo realizado por Cunha

¹ Departamento de Estatística, Universidade Federal do Amazonas, Manaus, Brasil.

² Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

Correspondência

J. D. O. Santos Junior
Departamento de Estatística,
Universidade Federal do Amazonas.
Rua Dom Emanuel Gomes 37,
apto. 202, Rio de Janeiro, RJ
21940-350, Brasil.
james@ufam.edu.br

et al.⁵ indicam a falta de medidas públicas para o controle de TB no Brasil.

O estudo da dinâmica da mortalidade de determinada doença constitui ferramenta epidemiológica fundamental, servindo para estudar quantitativamente o comportamento da doença na população e avaliar as intervenções. Para um estudo dinâmico, modelos de previsão são indispensáveis. No Brasil, uma série de estudos foi realizada com propósito de entender a dinâmica da mortalidade por TB. Em Bierrenbach et al.⁶, os autores realizaram um estudo descritivo das taxas de mortalidade anuais para cada estado, cobrindo os anos entre 1984 e 2004. Outro estudo descritivo pode ser visto em Hino et al.⁷, no qual foram examinadas as taxas de mortalidade entre os anos de 1980 a 2001. Um estudo interessante utilizando análise de séries temporais foi realizado por Antunes & Wadlmam⁸, no qual os autores analisaram a série de taxas de mortalidade para a cidade de São Paulo. No último, modelos autorregressivos de média móvel (ARMA) foram utilizados.

O presente estudo objetiva examinar as mudanças ocorridas na mortalidade bruta ao longo do tempo para o Estado de São Paulo, utilizando modelos lineares dinâmicos. Tal abordagem possui a vantagem de os parâmetros – como a média – serem variáveis aleatórias, podendo avaliar as mudanças ao longo do tempo. Além disso, esses modelos possuem a propriedade de se adaptar rapidamente às mudanças na série temporal analisada. Este tipo de modelagem já foi utilizado, por exemplo, para prever o número de casos novos de AIDS no Brasil⁹ e para previsão de resultados de partidas de futebol¹⁰. Na seção *A Série Temporal do Número de Óbitos por Tuberculose no Estado de São Paulo*, discutimos a escolha dos dados no estudo. Na seção *Modelos Lineares Dinâmicos*, apresentamos a metodologia utilizada e a escolha do modelo apropriado. Na seção *Resultados*, discutimos os resultados obtidos.

A série temporal do número de óbitos por tuberculose no Estado de São Paulo

Os dados sobre mortalidade por estado podem ser obtidos *online* pelo Sistema de Informações sobre Mortalidade (SIM. <http://www.datasus.gov.br>). O SIM foi criado em 1975 pelo Ministério da Saúde para prover estatísticas sobre mortalidade e indicadores de saúde. Até este presente trabalho, os dados disponíveis datavam entre os anos de 1979 e 2006.

Neste trabalho, optamos por utilizar os dados referentes ao Estado de São Paulo. A motivação da escolha advém dos estudos sobre esse estado

realizados anteriormente. Por exemplo, em Antunes & Waldam⁸, um grande esforço foi empreendido para reunir informações sobre a mortalidade por TB entre nos anos de 1900 a 1997, a fim de realizar uma análise temporal estatística utilizando a classe de modelos ARMA. Além disso, os resultados em Antunes & Waldam⁸ concordam com um estudo descritivo mais recente, feito em Bierrenbach et al.⁶, no qual as taxas de mortalidade para o Estado de São Paulo apresentam declínio até o fim do ano de 2004.

Preferimos também utilizar dados mensais, pois nesse nível de detalhamento podemos caracterizar o efeito sazonal da série – tal efeito não pode ser observado com a série anual. Resolvemos, também, estudar a dinâmica do número total de óbitos registrados por mês, em vez das taxas de mortalidade. Com isso, podemos prever quantitativamente o número de óbitos em um dado horizonte no tempo.

Por causa da mudança da 9ª Classificação Internacional de Doenças (CID-9) para CID-10, utilizaram-se os dados a partir de 1996. Os códigos de causa do óbito foram utilizados seguindo Bierrenbach et al.⁶: (1) TB pulmonar (CID-9: 100, 109-119; CID-10: A150-A153, A160-A162 e A169); (2) TB respiratória extrapulmonar (CID-9: 101-108, 120-129; CID-10: A154-A159, A163-A168); (3) TB extrapulmonar (CID-9: 130-189; CID-10: A170-A179); (4) TB congênita (CID-10: P370); (5) sequelas de TB (CID-10: B90); (6) doença pelo HIV resultando em TB (CID-10: B20).

Entre os anos de 1996 e 2006, a TB pulmonar e a doença por HIV resultando por TB foram responsáveis por 97,23% dos óbitos (por TB) registrados.

Modelos lineares dinâmicos

Os modelos de previsão bayesianos que utilizam modelos lineares dinâmicos foram introduzidos em Harrison & Stevens¹¹. A abordagem é fundamentalmente composta por: (1) uma definição de um modelo sequencial; (2) estruturação de modelos paramétricos com parametrização significativa para o modelador; (3) representação probabilística da informação sobre os parâmetros; (4) previsões calculadas com base numa distribuição de probabilidade.

Um caso particular dentre os modelos de previsão bayesianos é modelo linear dinâmico, baseado na distribuição normal. Essa escolha é bastante razoável, já que grande parte dos modelos estatísticos de previsão é baseada na suposição de normalidade. Além disso, a distribuição normal possui diversas vantagens, como fácil conjugação e uma distribuição multivariada fá-

cil de lidar. O modelo linear dinâmico permite incorporar componentes de tendências e sazonalidade, permitindo previsões adequadas ao comportamento dos dados. Neste trabalho, utilizamos a teoria de modelos lineares dinâmicos para analisar a dinâmica do número de óbitos por TB. Para a análise, usamos os seguintes componentes:

a) Tendência linear: componente que dá a direção das previsões futuras, baseada na informação corrente. Ela é dividida em duas partes – nível e tendência. O nível representa o intercepto da previsão, ou seja, a média atual das observações, desconsiderando outras componentes. A tendência informa sobre o crescimento/declínio dos valores futuros da média dos dados. Assim, uma tendência positiva (negativa) implica crescimento (declínio) da média. Uma tendência próxima de zero implica que as previsões futuras estarão próximas do nível atual dos dados, a menos de outras componentes.

b) Efeito sazonal: componente que controla o efeito sazonal dos dados. Um efeito sazonal é um comportamento que se repete após um período de tempo (como as estações do ano, por exemplo). Existem duas formas usuais para caracterizar as oscilações do efeito sazonal. Utilizaremos a sazonalidade trigonométrica, que decompõe o efeito sazonal em funções de cossenos. Cada uma dessas funções é denominada harmônico.

Com a modelagem dinâmica, pode-se obter:

a) Estimativas suavizadas das componentes: essa técnica utiliza toda a série temporal para estimar novamente as componentes do modelo. A ideia básica sobre esse tipo de técnica é que não só o passado e o presente ajudam a prever o futuro, mas também a nossa compreensão do presente nos auxilia a entender o passado.

b) Previsões: as previsões podem ser feitas tanto para as componentes quanto para as observações futuras. Como esperado, ao tentar prever um horizonte longe no tempo, a precisão das previsões fica comprometida. Portanto, previsões em curto prazo sempre serão mais precisas e podem ser utilizadas para verificar se o modelo é adequado aos dados, como será mostrado na próxima seção.

O modelo linear dinâmico é apresentado formalmente na Figura 1. Para uma leitura complementar, sugerimos Pole & Harrison¹² e Petris et al.¹³.

Resultados

Analisamos o número total de óbitos registrados por mês que tiveram por causa básica a TB. A análise também foi realizada, em separado, para

os dois tipos de causas básicas mais frequentes: TB pulmonar e doença por HIV seguido de infecção por TB.

O modelo utilizado foi o linear dinâmico com tendência linear e sazonalidade trigonométrica – a escolha desse modelo está justificada em nota, na Figura 1. A Figura 2 mostra o número total de óbitos registrados e o previsto para cada mês. Pode-se ver que os valores previstos para cada mês se aproximam do valor observado com o passar dos meses. Note ainda que a maioria dos pontos observados está no intervalo de predição com 90% de credibilidade, o que sugere a adequação do modelo para previsões em curto prazo.

Na Figura 3, mostram-se as estimativas suavizadas para o nível e a tendência. Podemos constatar que a tendência permaneceu abaixo de zero entre os anos de 1996 e 2004. Isso caracteriza um período com tendência ao declínio do número total de óbitos registrados por mês. Nesse mesmo período, o nível médio da série também decresceu. Tais resultados reforçam a análise descritiva dada em Bierrenbach et al.⁶. Entretanto, a partir de 2005, a tendência voltou a crescer, assim como o seu nível. Em 2006, a tendência se estabilizou em torno de um, enquanto o nível da série continuou a crescer (Figura 3, tendência).

As estimativas suavizadas para a sazonalidade também são mostradas na Figura 3. Os dados sugerem que o efeito sazonal é nulo nos meses de abril e outubro. Um efeito positivo, i.e, um aumento no número total de óbitos é encontrado nos meses maio, junho, julho, agosto e setembro. Um efeito negativo (diminuição do número total de óbitos) ocorre no restante dos meses.

As previsões para os anos de 2007 e 2008 são mostradas na Figura 4. Pode-se notar que, ainda que exista uma expectativa de aumento do número total de óbitos por mês para a TB, os limites de credibilidade permitem um cenário otimista, caracterizado pelo limite inferior do intervalo de 90% de credibilidade, tendo resultados semelhantes aos de 2004 e 2005. Entretanto, o cenário esperado é o crescimento do nível dessa série, mantendo o padrão de 2006 (ou aumentando). É importante ressaltar que um modelo dinâmico tende a levar em consideração os dados mais recentes, logo as previsões estão mais relacionadas com os dados de 2005 e 2006 que os demais anos.

Como discutido anteriormente, mais de 97% dos óbitos por TB registrados entre 1996 e 2006 foram ou por TB pulmonar ou por coinfeção HIV/TB. Por esse motivo, torna-se interessante analisar separadamente ambas. O mesmo modelo foi aplicado para as séries temporais das duas causas básicas de óbito citadas. As estimativas suavizadas para o nível e tendência de ambas

Figura 1

Modelo linear dinâmico utilizado.

Definição 1: o modelo linear dinâmico é caracterizado pela quádrupla: $\{H, G, V, W\}_t = \{H_t, G_t, V_t, W_t\}$ e, para cada t , tem-se:

- H_t é um vetor conhecido ($n \times 1$);
- G_t é uma matriz conhecida ($n \times n$);
- V_t é uma variância (escalar) conhecida;
- W_t é uma matriz de variâncias conhecida ($n \times n$).

Esta quádrupla define o modelo relacionando Z_t ao vetor de parâmetros θ_t ($n \times 1$) no tempo t , e a evolução de θ_t através do tempo, via a especificação das distribuições:

$$(Z_t | \theta_t) \sim N(H_t^T \theta_t, V_t)$$

$$(\theta_t | \theta_{t-1}) \sim N(G_t^T \theta_{t-1}, W_t)$$

Com a escolha certa para H_t e G_t , pode-se construir a maioria das funções de previsão da literatura. Os seguintes casos serão usados neste trabalho:

- Modelo de tendência polinomial de primeira ordem:
- Modelo de tendência polinomial de segunda ordem (ou de tendência linear):

$$H_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

- Modelo sazonal com componentes harmônicas de período p (par):

$$H_t = \begin{pmatrix} E_2 \\ E_2 \\ \vdots \\ E_2 \\ 1^T \end{pmatrix}, G_t = \begin{pmatrix} J_2(1, \omega) & 0 & \cdots & 0 & 0 \\ 0 & J_2(1, \omega) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & J_2(1, h\omega - \omega) & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

onde $E_2 = (1, 0, \dots, 0)$, $1^T = (1, 1, \dots, 1)$, $\omega = 2\pi/p$ e

$$J_2(1, \omega) = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix}$$

Esses modelos podem ser superpostos (West & Harisson 14, Teorema 6.4., p. 187). As previsões um passo à frente serão realizadas através da função de previsão.

$$f_t = E(Z_{t+1} | D_t) = H_{t+1}^T G_{t+1} m$$

Nota: Por simplicidade, denotaremos os modelos por $Mod(o, h)$, onde $o (> 0)$ é a ordem da componente polinomial e h é a quantidade de harmônicos utilizada ($h = 0$ implica que o modelo não possui a parte sazonal). *A priori*, foi assumido que a média do nível era de 140 e a média das outras componentes foi tomada como nula. A variância dos termos foi tomada como 900 (ou seja, *a priori* os dados estão entre 140 ± 60 com probabilidade aproximada de 95%). A variância W_t foi estimada com base na função de verossimilhança (Petris et al. 13). A Figura 6 mostra a raiz do erro quadrático médio (EQM) e o erro médio absoluto (EAM) para previsões realizadas pelos melhores modelos (o cálculo dessas medidas foi realizado retirando as 12 primeiras observações, pois elas são altamente influenciadas pela *priori*). Utilizando o critério de minimização dos erros, optamos pelo modelo $Mod(2, 1)$, ou seja, pelo modelo de tendência linear (segunda ordem) com um harmônico.

estão na Figura 5. É importante notar que, para o número total de óbitos por TB respiratória, tanto o nível quanto a tendência decresceram ao longo dos anos. Para os anos de 2005 e 2006, podemos verificar que o nível aparenta certa estabilidade (70 óbitos por mês). Além disso, a ten-

dência para esses mesmos anos está próxima de zero, o que implica previsões para os próximos meses em torno do nível médio 70 óbitos por mês. Todavia, o mesmo não ocorre com a causa básica coinfeção HIV/TB. Para essa causa, notamos aumento no nível, começando em 2004.

Figura 2

Intervalo de credibilidade e previsões um mês à frente para o número total de óbitos por tuberculose.

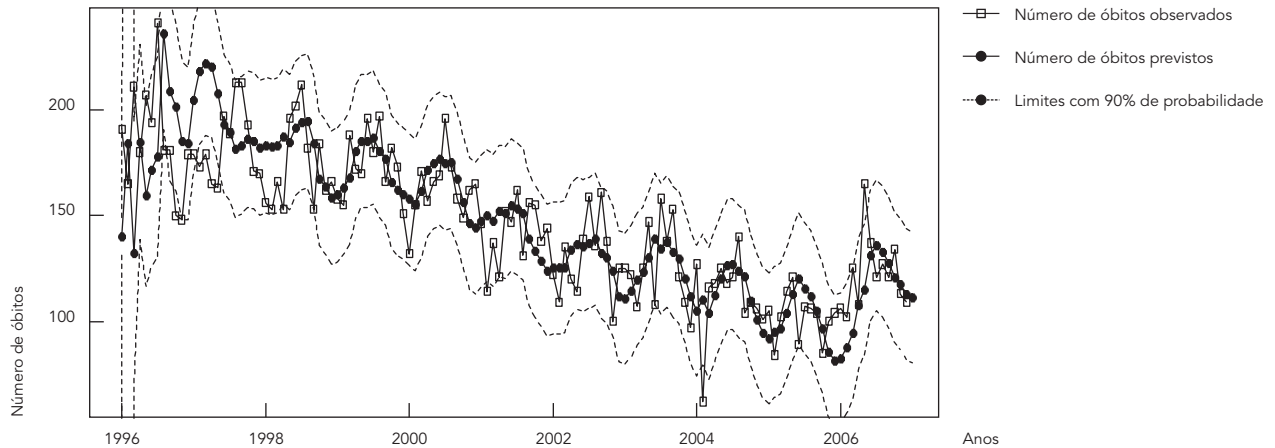
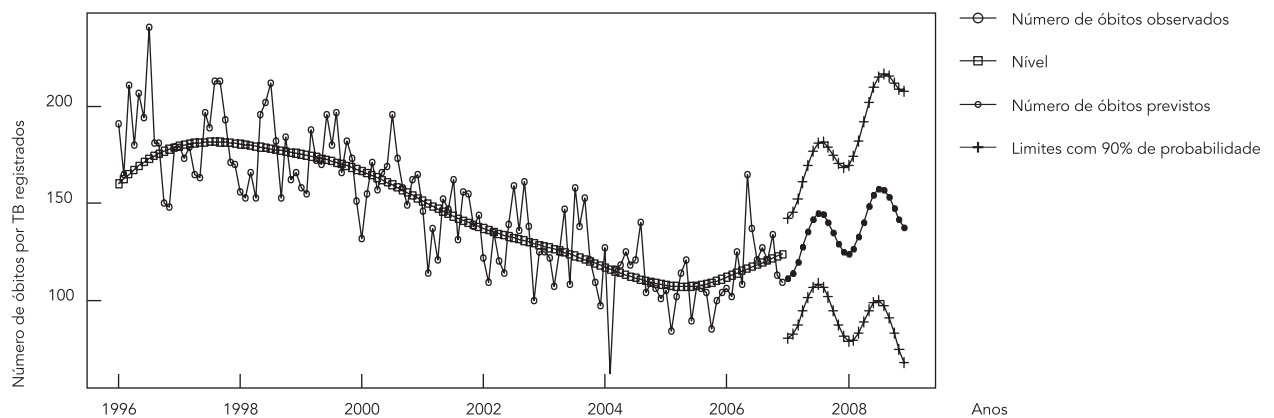


Figura 3

Decomposição da função de previsão em nível, tendência e sazonalidade.



TB: tuberculose.

Por volta do fim de 2004, a tendência começou a ficar positiva. Tanto o aumento do nível quanto da tendência indica que o aumento observado na série geral de óbitos por TB registrados é devido a essa causa básica. O cenário atual sugere que o aumento estimado para os anos de 2007 e 2008 para o número de óbitos por TB será por causa da coinfeção HIV/TB.

Considerações finais

Os modelos lineares dinâmicos se mostraram aptos a recuperar informações já reveladas em estudos descritivos dos dados de mortalidade por TB. A decomposição em nível/tendência/sazonalidade ajuda a entender melhor o comportamento da série, podendo, por exemplo, facilitar a análise

Figura 4

Série histórica e previsões para o número médio de registros de óbitos por tuberculose para os anos 2007 e 2008.

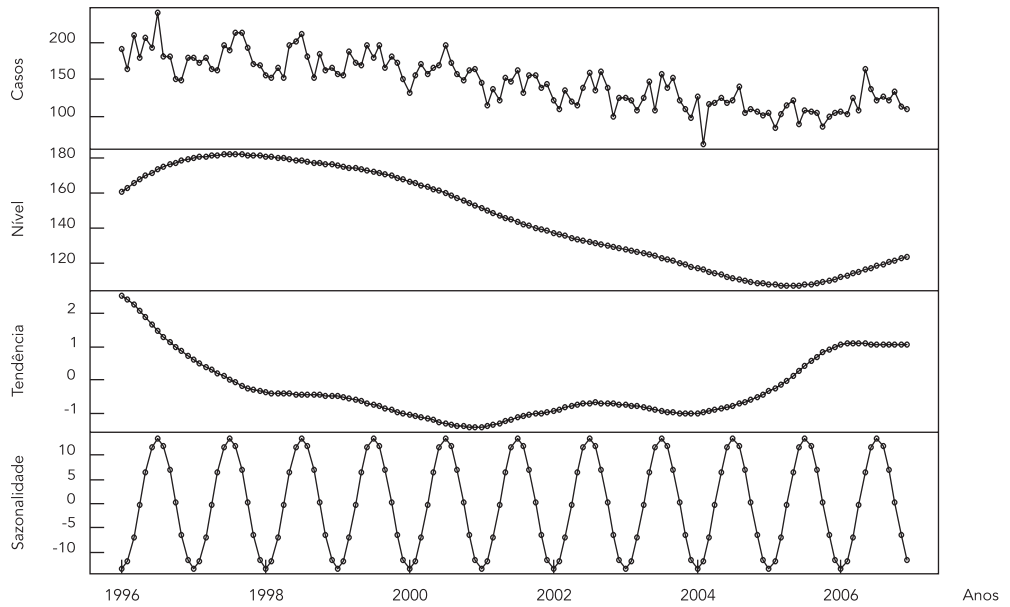
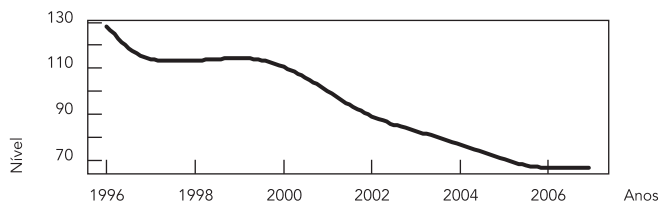


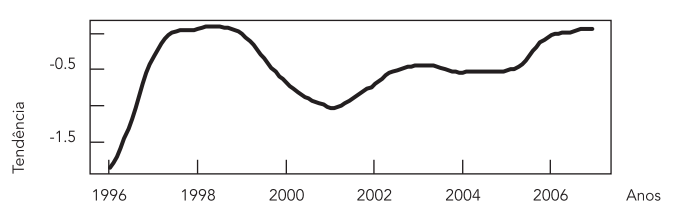
Figura 5

Estimativas suavizadas para o nível e tendência das séries históricas de óbitos por tuberculose (TB) registrados, em que as causas básicas foram TB pulmonar e coinfeção HIV/TB.

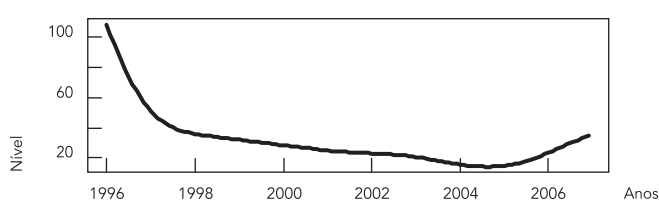
5a) TB pulmonar



5b) TB pulmonar



5c) Coinfeção HIV/TB



5d) Coinfeção HIV/TB

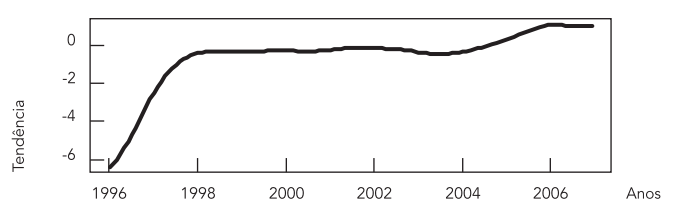
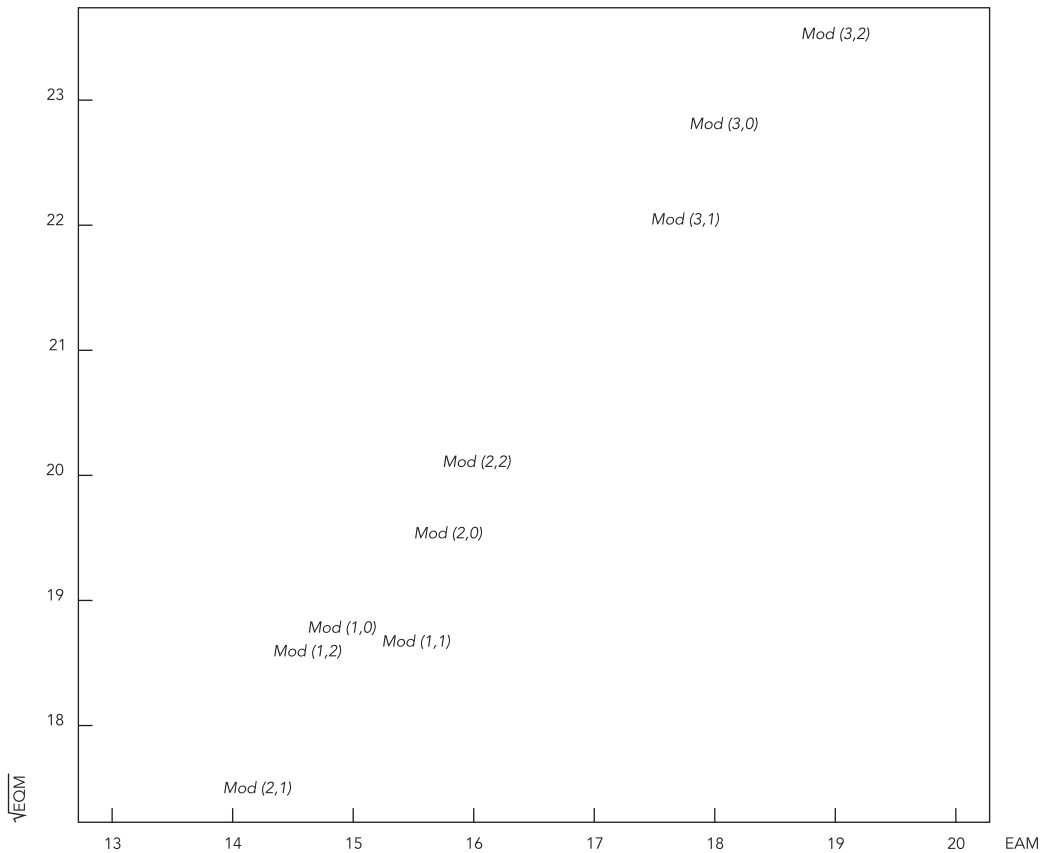


Figura 6

Erro quadrático médio e erro absoluto médio de previsão para os modelos propostos.



EAM: erro médio absoluto; EQM: erro quadrático médio.

do gestor público sobre a eficácia de uma nova medida de intervenção. Além disso, os modelos dinâmicos lineares permitem que o gestor possa estimar com razoável precisão o comportamento futuro da série, servindo como ferramenta de monitoração/intervenção.

Pela análise obtida neste trabalho, levantaram-se evidências de que o Estado de São Paulo vivenciou um período de baixa no número total de óbitos por TB, concordando com os estudos empíricos citados na seção *A Série Temporal do*

Número de Óbitos por Tuberculose no Estado de São Paulo. Entretanto, esta mesma análise, para os anos de 2005 e 2006 incluídos, revela um aumento no número total de óbitos. Uma análise mais detalhada que levou em conta apenas as duas causas básicas com maior número de óbitos revelou que tal aumento se deveu à coinfeção HIV/TB. Previsões para 2007 e 2008 indicam que o número total de óbitos por essa causa básica aumentou.

Resumo

Neste trabalho, utilizamos a técnica bayesiana de modelos lineares dinâmicos para estudar a dinâmica da mortalidade por tuberculose (TB) no Estado de São Paulo, Brasil. O estudo confirma os resultados empíricos da literatura, que indicam o decaimento da mortalidade para esse estado na última década. Este trabalho também indica um crescimento no número de óbitos a partir de 2005. A análise das duas causas básicas com maior ocorrência revelou que tal aumento é em razão da coinfeção HIV/TB. Previsões para os próximos dois anos foram realizadas, dando poucas evidências para sustentar a continuidade de tendência decrescente do número de óbitos por TB.

Tuberculose; Mortalidade; Validade dos Testes; Reprodutibilidade dos Testes

Colaboradores

J. D. O. Santos Junior participou da concepção do trabalho, organização dos dados, revisão bibliográfica, análise estatística e elaboração do artigo. B. B. Pereira colaborou na concepção inicial do trabalho, revisão da análise estatística e revisão do trabalho final.

Agradecimentos

Os autores agradecem o auxílio financeiro recebido pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pelo Projeto de Treinamento ICHORTA-AIDS/TB.

Referências

1. World Health Organization. Global tuberculosis control report 2008. http://www.who.int/tb/publications/global_report/2008/update/en/index.html (acessado em 01/Jan/2010).
2. Ruffino Netto A. Doenças emergentes ou reemergentes? Medicina (Ribeirão Preto) 1997; 30:405.
3. Almeida SAD, Honer MR. Abandono do tratamento da tuberculose em duas unidades de referência de Campo Grande, MS – 2002 e 2003. Bol Pneumol Sanit 2006; 14:167-71.
4. Kritski AL. Emergência de tuberculose resistente: renovado desafio. J Bras Pneumol 2010; 36:157-8.
5. Marques M, Cunha EAT, Ruffino-Netto A, Andrade SMO. Drug resistance profile of *Mycobacterium tuberculosis* in the state of Mato Grosso do Sul, Brazil, 2000-2006. J Bras Pneumol 2010; 36:224-31.
6. Bierrenbach AL, Gomes ABF, Noronha EF, Souza MFM. Tuberculosis incidence and cure rates, Brazil, 2000-2004. Rev Saúde Pública 2007; 41 Suppl 1:24-33.
7. Hino P, Costa-Júnior ML, Sasaki CM, Oliveira MF, Villa TCS, Santos CB. Time series of tuberculosis mortality in Brazil (1980-2001). Rev Latinoam Enferm 2007; 15:936-41.
8. Antunes JLE, Waldman EA. Tuberculosis in the twentieth century: time-series mortality in São Paulo, Brazil. Cad Saúde Pública 1999; 15:463-76.
9. Gamerman D, Migon HS. Forecasting the number of AIDS cases in Brazil. Statistician 1991; 40:427-42.
10. Owen A. Dynamic bayesian forecasting models of football match outcomes with estimation of the evolution variance parameter. IMA Journal of Management Mathematics 2011; 22:99-113.
11. Harrison PJ, Stevens CF. Bayesian forecasting. J R Stat Soc Series B Stat Methodol 1976; 38:205-47.
12. Pole AW, Harrison PJ. Applied bayesian forecasting and time series analysis. New York: Chapman-Hall; 2001.
13. Petris G, Petrone S, Campagnoli P. Dynamic linear models with R. New York: Springer-Verlag; 2010.
14. West M, Harrison, J. Bayesian forecasting and dynamic models. 2nd Ed. New York: Springer-Verlag; 1999.

Recebido em 18/Set/2009

Versão final reapresentada em 01/Mar/2011

Aprovado em 26/Abr/2011