Adaptive change-point mixed models applied to data on outpatient tetracycline use in Europe

Girma Minalu1, Marc Aerts1, Samuel Coenen2,3, Ann Versporten3, Arno Muller2, Niels Adriaenssens2,3, Philippe Beutels4, Geert Molenberghs1,5, Herman Goossens3 and Niel Hens1,4

1Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BIOSTAT), Hasselt University, Belgium
2Vaccine & Infectious Disease Institute (VAXINFECTIO), Centre for General Practice, University of Antwerp, Belgium
3Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Belgium
4Centre for Health Economics Research and Modeling Infectious Diseases Institute (CHERMID), Centre for the Evaluation of Vaccination, Vaccine and Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Belgium
5Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BIOSTAT), Catholic University of Leuven, Belgium

Abstract: In this paper, we propose a change-point mixed model to assess the change in the trend of outpatient antibiotic use in a Bayesian framework, where the change-points are unknown parameters of the model. Model selection using DIC indicates that the data supports the model with a country-specific change-point. The location of the change-points may be related to points in time where public health strategies aiming at increasing the awareness of the public to a more rational use of antibiotics or targeting to reduce overconsumption of antibiotics were initiated.

Key words: amplitude; antibiotic use; change-point model; non-linear model; phase shift; seasonal variation

Received January 2012; revised March 2013; accepted March 2013

1 Introduction

Antibiotics are drugs that inhibit or abolish the growth of bacteria. Antibiotic resistance is a major European and global public health problem and international efforts...
are needed to counteract the emergence of resistance. The increase in resistance rate of many important pathogens to currently most available antibiotics has now been recognized as a universal health hazard and potentially life-threatening problem. A large number of studies strongly suggest that this increase is directly related to the actual use of antibiotics. Antibiotic use is increasingly recognized as the main driver for resistance and differential selection pressure of antibiotic agents may be responsible for some of the observed differences (Goossens et al., 2005; Davey et al., 2008; Huttner et al., 2010).

Specific actions such as campaigns aimed at the public and general practitioners appeared essential, because antibiotic use in outpatients accounts for the main part of the overall antibiotic usage (>90%) (Goossens et al., 2005; Adriaenssens et al., 2011). Campaigns directed to the public in order to (i) inform about antibiotic resistance and to warn about the medical and general health issues related to the inappropriate use of antibiotics and (ii) foster the patient–physician and patient–pharmacist dialogue about the appropriate use of antibiotics, will increase the awareness of the public to a more rational use of antibiotics. In some European countries (e.g., in Belgium, France, Germany, Greece, Iceland, Italy, Luxembourg, Portugal, Spain and United Kingdom), campaigns were planned as part of a national strategy to reduce resistance to antimicrobial drugs. These strategies also included measures to promote appropriate use of antimicrobial drugs in hospitals, long-term care facilities and the agricultural sector (Huttner et al., 2010).

Longitudinal data on outpatient antibiotic use were available from 27 European countries for the period 1997–2009 within the European Surveillance of Antimicrobial Consumption (ESAC) project (Adriaenssens et al., 2011; Coenen et al., 2011; Minalu et al., 2011). Given that repeated measures were taken for each country, intra-country correlation has to be taken into account when analyzing the data. The main objective of the study is to develop an appropriate statistical model to assess the significance of country-specific trends in Europe and to identify possible change-points, while accounting for country-specific global use as well as seasonal effects.

In common regression, time series or longitudinal data analysis, the outcome variable is modelled as a linear function of explanatory variables and/or time. Sometimes it may happen that the relationship between the outcome and some explanatory variables and/or time is non-smooth (non-differentiable), showing one or more points where the effect on the response changes abruptly. These points are called break-points, change-points, transition-points or switch-points. To estimate the change-points, Bayesian (Smith, 1975; Carlin et al., 1992; Lange et al., 1992; Kiuchi et al., 1995; Slate and Turnbull, 2000; Ghosh and Vaida, 2007; Dominicus et al., 2008) or likelihood (Pastor and Guallar, 1998; Hall et al., 2000; Hall et al., 2003; Muggeo, 2003; Jacqmin-Gadda et al., 2006; Hens et al., 2010) methods may be used.

et al. (2006) proposed a random change-point model which combines a piecewise polynomial mixed model with a random change-point for the evolution of the cognitive test and a log-normal model depending on the random change-point for the time to dementia.

A fully Bayesian hierarchical structure for a mixed effects segmented regression model with one change-point was considered by Slate and Turnbull (2000), and applied to large data sets concerning prostate specific antigen as a serial marker for prostate cancer. Ghosh and Vaida (2007) proposed a change-point model with one random change-point for the analysis of longitudinal CD4 T-cell counts for HIV infected subjects following highly active antiretroviral treatment. And Dominicus et al. (2008) studied a Bayesian random change-point model with one random change-point to capture variability in measures of cognitive function. Hall et al. (2003) compared the Bayesian approach with the likelihood approach for modelling cognitive function over time, and pointed out that the Bayesian method has an advantage over the likelihood method in that it does not require all subjects to have the same change-point.

In this paper, an adaptive Bayesian linear spline model is proposed, where the number of knots (change-points) and their location are data driven and determined by the deviance information criterion (DIC). The presence and the location of the change-points is data driven and can vary across countries as random change-points. Latent country-specific indicators allow the model to switch off the change-points for particular countries.

The application of the model may yield new and important insights in the evolution of outpatient antibiotic use in Europe. We employ a fully Bayesian approach. The models are implemented in R using the R-package R2WinBUGS (Sturtz et al., 2005). The programs used for the analyses are available upon request from the first author. The program used to fit the change-point model with one unknown common change-point, one country-specific random change-point and a country-specific latent indicator for the change-point is included in Appendix III.

The paper is organized as follows. In Section 2, we describe the data on the total outpatient antibiotic use analyzed in the paper. In Section 3, we describe the models, the prior distributions for the parameters and we discuss how model comparison was applied. Results are presented in Section 4. Finally, discussions and concluding remarks are included in Section 5.

2 Outpatient antibiotic use data

Quarterly ESAC-NET data on total outpatient antibiotic use from 27 European countries were collected for the period 1997–2009 within ESAC-NET, an international network of surveillance systems. The methods of data collection and processing for the ESAC-NET project have been described in detail elsewhere (Adriaenssens et al., 2011; Coenen et al., 2011), and are also available on the ESAC-NET website (www.esac.ua.ac.be). Antibiotic use data is expressed as the number of defined
Figure 1 Observed country-specific evolutions for the quarterly use of tetracycline expressed in DID in 27 European countries

daily doses (DDD) per 1000 inhabitants per day (DID). This paper focuses on the outpatient use of tetracycline for the period 1997–2009, with the observed country-specific trends for the quarterly tetracycline use in DID shown in Figure 1.

As can be seen in Figure 1, there is variability across repeated measurements from the same country (i.e., within-country variability) as well as variability between countries (i.e., between-country variability), which suggests that country-specific intercepts and slopes should be incorporated into the model to account for heterogeneity across countries. The longitudinal profiles show clear seasonal variation of outpatient tetracycline use in all countries, with upward peaks in the winter season. Thus, a non-linear model needs to be adopted to take the seasonality into account. From the longitudinal profiles it can clearly be seen that countries with higher tetracycline use at the baseline (in 1997) have a higher amplitude (higher seasonal variation). Figure 1 also shows that not all longitudinal profiles are complete for all countries. Some profiles start later in time and others show intermediate missing parts. As the missingness mechanism is assumed to be missing completely at random (MCAR), all analyses were based on all available cases.

Statistical Modelling 2013; 13(3): 253–274
3 Adaptive change-point models

We start by introducing a non-linear mixed model with a sinusoidal component over time to account for the seasonal variation. We extend the non-linear mixed model by including fixed and random change-points to identify possible changes in the trend of tetracycline use in DID. We extend the existing approaches for random change-point models (e.g., Kiuchi et al., 1995; Ghosh and Vaida, 2007; Dominicus et al., 2008) by a general model building procedure where the number of knots and their location are data driven, and by taking into account a country-specific seasonal variation. The change-point models were also extended by including country-specific latent indicators, allowing the model to switch off the change-points for particular countries. All models are fitted in a fully Bayesian paradigm.

3.1 Non-linear mixed model

We applied the non-linear mixed model (3.1) to model the use of tetracycline in DID. An extension with known common change-points, unknown common change-points and country-specific random change-points is then considered. The non-linear mixed model is formulated as

\[
Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + (\beta_0^S + b_{0i}^S + \beta_1^S t_{ij}) \sin(\omega t_{ij} + \delta) + \varepsilon_{ij},
\]

where \(Y_{ij}\) is the total outpatient tetracycline use in DID for country \(i\) (\(i = 1, 2, \ldots, N\)) at time points \(t_{ij}\) (\(j = 1, 2, \ldots, n_i\)), \(n_i\) is the number of observations from the \(i\)th country, time = 1 corresponds to the start of the study (first quarter of 1997), \(\beta = (\beta_0, \beta_1, \beta_0^S, \beta_1^S, \delta)\) is a vector of fixed effects, \(\beta_0\) is the intercept, \(\beta_1\) is the regression coefficient describing the marginal linear time trend (\(t\)), \(\beta_0^S\) is the fixed amplitude, \(\beta_1^S\) is the amplitude varying over time, \(\omega\) (in radians) is the frequency which is a known constant (\(= 2\pi / T\) where \(T = 4\)) is the period for the sine curve, \(\delta\) (in radians) is the phase shift or phase angle which is an unknown parameter, \(b_i = (b_{0i}, b_{1i}, b_{0i}^S)\) is the country-specific vector of random effects where \(b_{0i}\) is the country-specific random intercept, \(b_{1i}\) is the country-specific random slope for time and \(b_{0i}^S\) is the country-specific random slope for amplitude and we assume \(b_i \sim N(0, D)\). The matrix \(D\) is a general covariance matrix with elements \(d_{ij} = d_{ji}\). \(\varepsilon_i\) is an \(n_i\)-dimensional vector of unexplained error terms \(\varepsilon_{ij}\). It is usually assumed that all \(\varepsilon_i\) are independent and normally distributed with mean vector zero and covariance matrix \(\Sigma_i\). Often, \(\Sigma_i\) is assumed equal to \(\sigma_i^2 I_{n_i}\), where \(I_{n_i}\) is the \(n_i\)-dimensional identity matrix.

Since no convergence was obtained when using an unstructured covariance matrix for the random effects, a diagonal covariance matrix was used.
3.2 Adaptive change-point model

Since there is no prior knowledge on the number of change-points, we gradually build up the model by first considering a change-point model with a known common change-point and then extending it by including unknown common and country-specific random change-points.

A general mixed model with country-specific mean can be written as

\[ Y_{ij} = \mu_i(t_{ij}) + \varepsilon_{ij}, \quad i = 1, 2, \ldots, N; \quad j = 1, 2, \ldots, n_i, \]

where \( Y_{ij} \) is the tetracycline use in DID for country \( i \) at time points \( t_{ij} \), \( \mu_i(t_{ij}) \) is the trend component, \( \varepsilon_{ij} \) is the measurement error which is assumed to be normally distributed with mean zero and constant variance \( \sigma^2 \). The country-specific mean components \( \mu_i(t_{ij}) \) are modelled as

\[ \mu_i(t_{ij}) = (\beta_0 + b_0^i) + (\beta_1 + b_1^i)t_{ij} + \mu_i^{CP}(t_{ij}), \]

where \( \mu_i^{CP}(t_{ij}) \) is a change-point component given by

\[ \mu_i^{CP}(t_{ij}) = \sum_{k=1}^{K} (\beta_{k+1} + b_{k+1}^i)(t_{ij} - K_{ki}), \]

where \( x_+ = \max(x, 0) \), \( K \) is the number of unknown change-points, \( K_{ki} = C_k \) or \( K_{ki} = C_k + c_k \) or \( K_{ki} = c_k \) where \( C_k \) denotes a global change-point and \( c_k \) a country-specific random change-point. If \( \mu_i^{CP}(t_{ij}) = 0 \) then there are no change-points and the model reduces to model (3.1).

Substituting equations (3.3) and (3.4) in equation (3.2) yields the model

\[ Y_{ij} = (\beta_0 + b_0^i) + (\beta_1 + b_1^i)t_{ij} + \sum_{k=1}^{K} (\beta_{(k+1)} + b_{(k+1)}^i)(t_{ij} - K_{ki}), \]

\[ + (\beta_0^S + b_0^S + \beta_1^S t_{ij}) \sin(\omega t_{ij} + \delta) + \varepsilon_{ij}, \]

where the fixed effects \( \beta_0, \beta_1, \beta_0^S, \beta_1^S, \omega \) and \( \delta \), and the random effects \( b_0^i, b_1^i \) and \( b_0^S \) are defined as before, \( K \) is the number of change-points, for \( k = 1, 2, \ldots, K \), \( \beta_{(k+1)} \) is the global difference in the linear trend before and after the change-point, \( b_{(k+1)}^i \) is the country-specific difference in the linear trend before and after the change-point and \( \varepsilon_{ij} \) is an unexplained error term. Random effects for the global level of use, the trend effects, the amplitude of the seasonal effect and the location of the change-point are used to account for heterogeneity across countries. The number of change-points \( K \) and the location of the change-point(s) are data driven.
Adaptive change-point mixed models applied to data on outpatient 259

In Equation (3.5) all countries are assumed to have a change in the trend of tetracycline use in DID, but this might not be true because some countries might not have a change in the trend of tetracycline use. To relax this assumption, we extend (3.5) by including country-specific latent indicators for the change-points,

\[ Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + \sum_{k=1}^{K} \left\{ (\beta_{(k+1)} + b_{(k+1)i})(t_{ij} - K_{ki}), I_{ki} \right\} + \epsilon_{ij}, \]

where \( I_{ki} \) is an unknown country-specific indicator for the change in the trend of tetracycline use in DID for country \( i \) for the \( k \)th change-point, \( k = 1, 2, \ldots, K \) where \( K \) is the number of change-points. Here, \( I_{ki} = 1 \) if there is a change at knot \( K_{ki} \) in the use of tetracycline over time in country \( i \), or \( I_{ki} = 0 \) if there is no change in the use of tetracycline over time in country \( i \).

As there are no prior information on the number of change-points in the study, the number of change-points \( K \) in Equations (3.5) and (3.6) has to be chosen prior to the data fitting, \( k = 1, \ldots, K \). We first start from the simplest model where there is only a known common change-point, i.e., \( K = 1 \). We gradually extend the model by including a known and an unknown common change-point. And later, we extended the model by including an additional unknown common change-point. Next to the common change-points, country-specific random change-points have also been included in the model.

3.3 Prior specification

The following uninformative prior distributions were used for the fixed effects:

\[ \beta_0, \beta_1, \beta_{(k+1)}, \beta_0^S, \beta_1^S, \delta \sim \text{Normal}(0, 1000), \text{ independently where } k = 1, \ldots, K, \]

\[ C_1 \sim \text{Uniform}(1, 52), \]

\[ C_2 \sim \text{Uniform}(C_1, 52). \]  

(3.7)

The normal priors on \( \beta_0, \beta_1, \beta_{(k+1)}, \beta_0^S, \beta_1^S \) and \( \delta \) have large variances, expressing our lack of knowledge about the regression coefficients. For the random effects, a normal prior distributions was used:

\[ b_{0i} \sim \text{Normal}(0, \sigma_{b0}^2), \]

\[ b_{1i} \sim \text{Normal}(0, \sigma_{b1}^2), \]

\[ b_{(k+1)i} \sim \text{Normal}(0, \sigma_{b(k+1)}^2), \]

\[ b_{0}^S \sim \text{Normal}(0, \sigma_{b0}^2), \]

\[ \epsilon_{ki} \sim \text{Normal}(C_k, \sigma_{\epsilon k}^2) \text{I}(1, 52). \]

(3.8)
A uniform prior distribution over the total range of time was also assumed for the country-specific random change-point:

\[ c_{ki} \sim \text{Uniform}(1, 52). \]  \hfill (3.9)

The country-specific indicator for the \( k \)th change-point (\( I_{ki} \)) is Bernoulli-distributed with probability \( P_k \), where the probability \( P_k \) is beta-distributed with shape parameters \( \alpha_p (=1) \) and \( \beta_p (=1) \):

\[ I_{ki} \sim \text{dbern}(P_k), \quad P_k \sim \text{dbeta}(1,1). \]  \hfill (3.10)

The hyperparameters in the prior distributions were chosen so that the priors are uninformative. An independent inverse gamma distribution with a shape parameter \( \alpha (=0.001) \) and a scale parameter \( \beta (=0.001) \) was used for the variance parameters.

\[ \sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_{b_{k+1}}^2, \sigma_{b_{k+1}}^2, \sigma_{c_k}^2, \sigma_{\epsilon}^2 \sim \text{IGamma}(0.001, 0.001), \text{independently}. \]  \hfill (3.11)

where \( x \sim \text{IGamma}(\alpha, \beta) \) means that \( 1/x \) has the Gamma distribution with mean \( \alpha/\beta \) and variance \( \alpha/\beta^2 \) (Ntzoufras, 2009).

### 3.4 Model selection

We use the DIC for model comparison (Spiegelhalter et al., 2002). The DIC can be represented as:

\[ \text{DIC} = p_D + \bar{D}. \]  \hfill (3.12)

DIC is a Bayesian equivalent to Akaike’s information criterion (AIC) and consists of two components, a term that measures goodness-of-fit (\( \bar{D} \), defined as the posterior expectation of the deviance) and a penalty term for model complexity (\( p_D \), defined as the difference between the posterior mean of deviance and the deviances evaluated at the posterior mean \( \bar{\theta} \) of the parameters). \( p_D = \bar{D} - D(\bar{\theta}) \). The smaller the DIC, the better the fit (Spiegelhalter et al., 2002; Gelman et al., 2004; Ghosh and Vaida, 2007; Dominicus et al., 2008).

There has been and there still is discussion on Bayesian model selection in general and on the specification of the prior for model selection and hypothesis testing, related to the Jeffreys-Lindley paradox (Lindley, 1957). See, for instance, Spiegelhalter et al. (2002) in which pros and cons of several approaches to Bayesian model selection are discussed by the authors and several discussants. Posterior model probabilities and Bayes factors might be considered to represent the gold standard in fully Bayesian model determination, but these quantities are sensitive to the choice of prior distribution in the case of specifying a default prior under weak prior information (Overstall and Forster, 2010). Criterion-based methods such as BIC or DIC do not give posterior model probabilities, and as such the issue of default prior specification is avoided. Here we opted for DIC as a criterion for model selection. An in-depth analysis of this issue with a comparison of the performance and characteristics of
adaptive change-point mixed models applied to data on outpatient 261
different model selection criteria is considered beyond the scope of this paper. For
more details and some recent developments on this issue related to Lindley’s paradox,
we refer to Casella et al. (2009), Mulder et al. (2009), Overstall and Forster (2010),
and references therein.
Plummer (2008) provided a justification for the DIC by demonstrating the link
between DIC and cross-validation. In his paper, DIC is shown to be an approximation
to a penalized loss function based on the deviance, with a penalty derived from
a cross-validation argument. This approximation is valid only when the effective
number of parameters in the model is much smaller than the number of independent
observations (i.e., \( p_D << n \)). A corrected DIC, \( \text{DIC}_c = \hat{D} + \sum_{i=1}^{n} p_{D_i} / (1 - p_{D_i}) \), was
suggested for generalized linear mixed models when the DIC cannot be justified as
approximation to the penalized plug-in deviance. To the best of our knowledge,
the use of the corrected DIC has not been studied for non-linear mixed models and
requires further research. Therefore, we do not pursue its use in this paper.
The quarterly tetracycline use data was analyzed in Minalu et al. (2011) using the
non-linear mixed model. The results of the non-linear mixed models were used as a
starting value for the MCMC algorithm. And for the additional change-point param-
eters, the locations of campaigns or policy changes in antibiotic use in most European
countries were used as starting values. To ensure adequate convergence all results
were obtained using two chains of 110 000 iterations, of which we discarded the first
10 000 (burn-in) and the chain was then thinned to every 5th sample as there was
autocorrelation for some parameters. Trace plots and the potential scale reduction
\( \hat{R} \) were used to check convergency of the MCMC algorithm (Gelman et al., 2004).

4 Results
We considered the following models, within the family (3.5):

Model 1: Non-linear mixed model without a change-point,
\[ \mu_{i_{CP}}(t_{ij}) = 0, \]
Model 2: Non-linear mixed model with a known common change-point \( (C_1 = 17) \),
\[ \mu_{i_{CP}}(t_{ij}) = (\beta_2 + b_{2i})(t_{ij} - 17), \]
Model 3: Non-linear mixed model with a known common change-point \( (C_1 = 29) \),
\[ \mu_{i_{CP}}(t_{ij}) = (\beta_2 + b_{2i})(t_{ij} - 29), \]
Model 4: Non-linear mixed model with one unknown common change-point
\( (C_1) \),
Model 5: Non-linear mixed model with two unknown common change-points \((C_1 \text{ and } C_2)\),
\[\mu_{CP}^i (t_{ij}) = (\beta_2 + b_{2i})(t_{ij} - C_1) +,\]
where ordering restriction was imposed for the common change-points (i.e., \(C_1 < C_2\)).

Model 6: Non-linear mixed model with one country-specific random change-point \((c_i)\),
\[\mu_{CP}^i (t_{ij}) = (\beta_2 + b_{2i})(t_{ij} - c_i) +,\]
where the country-specific random change-point is centred around the unknown point \(C_1\) and is restricted to lay within \([1,52]\), \(c_i \sim N(C_1, \sigma^2_c)(1, 52)\).

Model 7: Non-linear mixed model with one country-specific random change-point \((c_i)\),
\[\mu_{CP}^i (t_{ij}) = (\beta_2 + b_{2i})(t_{ij} - c_i) +,\]
where the country-specific random change-points are \(U(1, 52)\) distributed.

Model 1 without a change-point is first extended with known common change-points (Models 2 and 3). Because there were public campaigns in some of the European countries during the year 2000–01 (e.g., in Belgium, Germany and Greece) and during the year 2004–05 (e.g., in Portugal and United Kingdom), we used time \(= 17\) (first quarter of 2001) and time \(= 29\) (first quarter of 2004) as known common change-points in the trend of tetracycline use in DID, respectively, in Model 2 and Model 3. Next, we estimate the change-points by including unknown common and/or country-specific random change-points (Models 4–7). The non-linear mixed model (Model 1) was extended by including a non-linear trend and secondly an amplitude varying non-linearly over time (expressed as \(t_{ij}^\alpha\)). As these extended models did not outperform the change-point models, we only presented the results of the original non-linear and change-point models (Models 1–7). Various models with three change-points were applied too, but convergence could not be reached for any of these models.

For the unknown common change-points in Models 4–6, uniform prior distributions over the total range of time were used. A normal prior distribution with mean zero and variance \(\sigma^2_c\) was used for the country-specific random change-point in Model 6, while in Model 7 a uniform prior distribution over the total range of time was assumed for the country-specific random change-point. A summary of the posterior distributions of the model parameters in Models 1–7 is given in Table 1.

The results in Table 1 clearly indicate the need for one or more change-points. Indeed, Model 1 (no change-points) gets little support with the highest DIC = 391.6500. Including a known common change-point reduces the DIC considerably (Models 2
Table 1  Parameter estimates: posterior means (and standard errors), and model comparison (\(\bar{D}, pD\) and DIC values)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta_0)</td>
<td>2.6399 (0.2630)</td>
<td>2.7920 (0.3121)</td>
<td>2.5814 (0.2941)</td>
<td>2.6367 (0.2505)</td>
<td>2.7240 (0.2770)</td>
<td>2.5937 (0.2986)</td>
<td>2.6330 (0.2815)</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>-0.0087 (0.0041)</td>
<td>-0.0253 (0.0074)</td>
<td>-0.0118 (0.0067)</td>
<td>-0.0115 (0.0065)</td>
<td>-0.0212 (0.0078)</td>
<td>-0.0133 (0.0064)</td>
<td>-0.0146 (0.0071)</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>-0.0020 (0.0078)</td>
<td>0.0093 (0.0089)</td>
<td>0.0098 (0.0088)</td>
<td>0.0213 (0.0119)</td>
<td>0.0112 (0.0102)</td>
<td>0.0118 (0.0109)</td>
<td></td>
</tr>
<tr>
<td>(\beta_3)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(C_1)</td>
<td>-</td>
<td>17(^*)</td>
<td>29(^*)</td>
<td>29.375 (1.2912)</td>
<td>20.2353 (3.1875)</td>
<td>29.4144 (2.8802)</td>
<td></td>
</tr>
<tr>
<td>(C_2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>31.9461 (1.4710)</td>
<td>-</td>
</tr>
<tr>
<td>(\beta^0)</td>
<td>0.6176 (0.0629)</td>
<td>0.6098 (0.0622)</td>
<td>0.6112 (0.0614)</td>
<td>0.6109 (0.0618)</td>
<td>0.6083 (0.0616)</td>
<td>0.6100 (0.0612)</td>
<td>0.6113 (0.0618)</td>
</tr>
<tr>
<td>(\beta^1)</td>
<td>-0.0064 (0.0010)</td>
<td>-0.0062 (0.0009)</td>
<td>-0.0062 (0.0008)</td>
<td>-0.0062 (0.0008)</td>
<td>-0.0062 (0.0008)</td>
<td>-0.0062 (0.0008)</td>
<td>-0.0062 (0.0008)</td>
</tr>
<tr>
<td>(\delta)</td>
<td>0.4972 (0.0245)</td>
<td>0.5041 (0.0228)</td>
<td>0.4989 (0.0217)</td>
<td>0.4988 (0.0216)</td>
<td>0.5016 (0.0211)</td>
<td>0.5002 (0.0213)</td>
<td>0.5004 (0.0212)</td>
</tr>
</tbody>
</table>

| \(\sigma^2_{b0}\) | 1.8174 (0.5703) | 2.0457 (0.6508) | 1.9518 (0.6247) | 1.9091 (0.5979) | 1.8666 (0.5975) | 1.9835 (0.6359) | 2.0599 (0.6536) |
| \(\sigma^2_{b1}\) | 0.0004 (0.0001) | 0.0009 (0.0003) | 0.0009 (0.0003) | 0.0008 (0.0003) | 0.0009 (0.0003) | 0.0009 (0.0003) | 0.0009 (0.0004) |
| \(\sigma^2_{b2}\) | -            | 0.0010 (0.0004) | 0.0016 (0.0006) | 0.0016 (0.0006) | 0.0009 (0.0004) | 0.0020 (0.0008) | 0.0024 (0.0010) |
| \(\sigma^2_{b3}\) | -            | -            | -            | -            | 0.0018 (0.0007) | -            | -            |
| \(\sigma^2_{b4}\) | -            | -            | -            | -            | -            | 51.3065 (32.6960) | -            |
| \(\sigma^2_{bS}\) | 0.0777 (0.0249) | 0.0782 (0.0248) | 0.0788 (0.0246) | 0.0790 (0.0252) | 0.0792 (0.0250) | 0.0791 (0.0250) | 0.0791 (0.0249) |
| \(\sigma^2_{e}\) | 0.0806 (0.0038) | 0.0691 (0.0033) | 0.0623 (0.0030) | 0.0625 (0.0030) | 0.0597 (0.0029) | 0.0605 (0.0029) | 0.0603 (0.0030) |

| \(D\)      | 313.8719 | 162.3209 | 60.6070 | 62.9339 | 17.5635 | 30.5115 | 28.6827 |
| \(pD\)     | 77.7781 | 91.7648 | 96.3882 | 97.7551 | 102.4649 | 95.3957 | 56.2395 |
| \(DIC\)    | 391.6500 | 254.0857 | 156.9952 | 160.6891 | 120.0285 | 125.9073 | 84.9222 |

Note: *Because there were public campaigns in some of the European countries during the year 2000–01 (e.g., in Belgium, Germany and Greece) and during the year 2004–05 (e.g., in Portugal and UK), time = 17 and time = 29 are used as known common change-points.

Source: Authors' own.
and 3). There is no improvement when the known change-point 29 is replaced by an unknown common change-point (Model 4). There is, however, a further improvement when two unknown common change-points are included in the model (Model 5). In Models 2–5 all countries are assumed to have the same common change-point, while in Models 6–7 all countries have different change-points. Comparing Model 6 with Model 4 shows a reduction in DIC when including a country-specific random change-point next to the global change-point. A large improvement is achieved when a uniform prior distribution over the total range of time was used for the country-specific random change-point (Model 7). Scatter plots of country-specific estimates for the change-points in Models 6 and 7 are shown in Figures A1 and A2 of Appendix I.

The estimate for the unknown common change-point \( C_1 \) obtained from fitting Model 4 is 29.3975 (fourth quarter of 2003), which is quite close to the estimates for the common change-point obtained from fitting Model 6 \( (C_1 = 29.4144) \). The average for the estimated country-specific random change-points in Model 7 is 28.7451, which is very close to the estimate for the unknown common change-points in Models 4 and 6. From Model 5, the estimate for the first common change-point \( C_1 \) is 20.2353 (fourth quarter of 2001) and 31.9461 (fourth quarter of 2004) for the second common change-point \( (C_2) \).

The 95% quantile-based credible interval for \( \beta_1 \) \((-0.0270, -0.0003)\) indicates that there is a significant decrease in the global trend of tetracycline use in DID. The credible intervals for \( \beta^S_0 \) and \( \beta^S_1 \) do not include zero, indicating a significant overall seasonal variation and a significant overall seasonal variation trend over time, respectively.

The estimated linear trend (dashed line), the estimated change-point model (solid line) from Model 7 and the observed average DID for Europe are shown in Figure 2. The predicted mean is based on the predicted outcomes from the posterior distribution of the country-specific random effects. Figure 2 indicates that the model describes the data very well.

Models 2–7 assume that there are one or more trend changes of tetracycline use in all countries, but for some countries it might be better to have only one or even no change-point. To allow a data-adaptive selection of the number and location of the country-specific change-points, we extend Models 4–7 by including a latent country-specific indicator \( I_{ki} \) for the \( k \)th change-point, \( k = 1, 2, \ldots, K \) for country \( i \) \( (i = 1, 2, \ldots, N) \).

\textit{Model 4∗}: Non-linear mixed model with one unknown common change-point \( (C_1) \) and a country-specific indicator \( I_{1i} \),

\[
\mu_{i,CP}^{tij} = \{(\beta_2 + b_{2i})(t_{ij} - C_1)\} I_{1i},
\]

where \( I_{1i} \) is an unknown country-specific indicator for the change in the trend of DID for country \( i \). Here, \( I_{1i} = 1 \) if a change at \( C_1 \) in the use of tetracycline over time in country \( i \) is needed, or \( I_{1i} = 0 \) if no change in the use of tetracycline over time in country \( i \) is needed,
Adaptive change-point mixed models applied to data on outpatient

Figure 2  The observed mean DID (dots), the predicted mean profile (solid line) and the predicted linear trend (dashed line) obtained from fitting Model 7. Source: Authors’ own.

Model 5*: Non-linear mixed model with two unknown common change-points ($C_1$ and $C_2$) and two country-specific indicators ($I_{1i}$ and $I_{2i}$),

$$
\mu_{i}^{CP}(t_{ij}) = \{(\beta_2 + b_2 i)(t_{ij} - C_1)\} I_{1i} + \{(\beta_3 + b_3 i)(t_{ij} - C_2)\} I_{2i},
$$

where ordering restriction was imposed for the common change-points (i.e., $C_1 < C_2$).

Model 6*: Non-linear mixed model with one country-specific random change-point ($c_i$) and a country-specific indicator $I_{1i}$,

$$
\mu_{i}^{CP}(t_{ij}) = \{(\beta_2 + b_2 i)(t_{ij} - c_i)\} I_{1i},
$$

where the country-specific random change-point is centred around the unknown point $C_1$ and is restricted to lay within $[1,52]$, $c_i \sim N(C_1, \sigma^2_c)(1, 52)$.

Model 7*: Non-linear mixed model with a country-specific random change-point ($c_i$) and a country-specific indicator $I_{1i}$,

Statistical Modelling 2013; 13(3): 253–274
Table 2  Parameter estimates: posterior means and standard errors, and model comparison: $D$, $pD$ and DIC values obtained from fitting Models 4*, 6* and 7*

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 4*</th>
<th>Model 6*</th>
<th>Model 7*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>2.6488 (0.2734)</td>
<td>2.6618 (0.2890)</td>
<td>2.6322 (0.2527)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.0120 (0.0064)</td>
<td>-0.0140 (0.0066)</td>
<td>-0.0139 (0.0065)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0108 (0.0105)</td>
<td>0.0130 (0.0111)</td>
<td>0.0126 (0.0126)</td>
</tr>
<tr>
<td>$C_1$</td>
<td>29.4560 (1.2975)</td>
<td>29.1115 (2.8179)</td>
<td>-</td>
</tr>
<tr>
<td>$\rho^S_0$</td>
<td>0.6113 (0.0630)</td>
<td>0.6104 (0.0613)</td>
<td>0.6120 (0.0615)</td>
</tr>
<tr>
<td>$\rho^S_1$</td>
<td>-0.0062 (0.0008)</td>
<td>-0.0062 (0.0008)</td>
<td>-0.0062 (0.0008)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.4985 (0.0218)</td>
<td>0.4999 (0.0214)</td>
<td>0.5002 (0.0213)</td>
</tr>
<tr>
<td>$P_1$</td>
<td>0.8407 (0.1222)</td>
<td>0.8651 (0.1091)</td>
<td>0.8861 (0.1068)</td>
</tr>
</tbody>
</table>

| $\sigma^2_b0$ | 1.9527 (0.6229) | 2.0181 (0.6382) | 2.0580 (0.6472) |
| $\sigma^2_b1$ | 0.0009 (0.0003) | 0.0009 (0.0003) | 0.0009 (0.0003) |
| $\sigma^2_b2$ | 0.0018 (0.0006) | 0.0022 (0.0009) | 0.0026 (0.0011) |
| $\sigma^2_{c1}$ | - | 46.3954 (30.8729) | - |
| $\sigma^2_{b0S}$ | 0.0791 (0.0252) | 0.0788 (0.0248) | 0.0790 (0.0251) |
| $\sigma^2_{e1}$ | 0.0626 (0.0030) | 0.0606 (0.0030) | 0.0605 (0.0030) |

| $D$       | 65.1738 | 32.7773 | 31.0169 |
| $pD$      | 87.6091 | 92.3128 | 54.7426 |
| DIC       | 152.7831 | 125.0902 | 85.7595 |

Note: * Models 4, 6 and 7 are fitted with a country-specific latent indicator $I_{ki}$

Source: Authors’ own.

\[
\mu_{tij}^{CP} = \{(\beta_2 + b_{2i})(t_{ij} - c_i)\} I_{ki},
\]

where the country-specific random change points are $U(1, 52)$ distributed.

The parameter estimates for all parameters in Models 4*, 6* and 7* are given in Table 2. No convergence was obtained for Model 5*.

From the results given in Table 2, Model 7* has the lowest DIC value which is quite close to the DIC value of Model 7 (in Table 1). The parameter estimates given in Table 2 are also close to the corresponding parameter estimates given in Table 1. The parameter estimates for the country-specific latent indicators $I_{ki}$ are given in Table A1 in Appendix II. The posterior means for the change-point indicator $I_{ki}$ is greater than 0.5 for all countries, which indicates a change in the trend of tetracycline use for all countries.

The observed country-specific profiles and the predicted country-specific profiles from Model 7 for three selected countries (Iceland, Belgium and Austria) are shown in Figure 3. As can be seen from Figure 3, the predicted country-specific profiles follow closely the observed country-specific DID values. The bold dots indicate the estimated country-specific random change-points obtained from fitting Model 7.

A visual inspection of convergence diagnostics graphs for various model parameters showed that the posterior densities are smooth and unimodal shapes.
5 Discussion

This study was motivated by the need to assess the use of tetracycline in 27 European countries, to assess the change in the trend of tetracycline use over time, and to possibly relate any changes in antibiotics use due to campaigns and policy changes. The data have previously been analyzed based on a non-linear mixed model while taking into account the seasonal effects (Minalu et al., 2011). From the analysis, we have identified significant variation in total outpatient tetracycline use in Europe. Differences in tetracycline use between countries might be explained by variations in incidence of community acquired infections, culture and education, and differences in drug regulations and in the structure of the national pharmaceutical market (Goossens et al., 2005).
In this paper, we presented and discussed adaptive change-point Bayesian models to analyze the outpatient tetracycline use from 1997 to 2009. We considered the non-linear mixed model extended with known common change-points, unknown common change-points and country-specific random change-points. The change-point mixed model was also extended by including country-specific indicators for the change-points. A widely used statistic for comparing models in a Bayesian framework, the DIC, was used for model comparison. The model with country-specific change-points (Model 7) has the lowest value of DIC. There is some controversy on which criterion to use to compare Bayesian models. Gelman et al. (2004) suggested $p_V = \text{Var(Deviance)}/2$ as an estimate of the effective number of parameters in the model as an alternative to $p_D$. Note that using $p_V$ as an alternative measure of complexity, the change-point model with two unknown common change-points (Model 5) has the lowest DIC value.

The random change-point models have been applied in many applications (Kiuchi et al., 1995; Ghosh and Vaida, 2007; Dominicus et al., 2008). In this paper, we extended the existing approaches by a general model building procedure where the number of knots and their location are data driven. We also extended the previously proposed change-point models by taking into account a country-specific seasonal variation. The change-point models were also extended by including country-specific latent indicators, allowing the model to switch off the change-points for particular countries.

From the results obtained from fitting the change-point model with a country-specific change-point (Model 7), there is a significant decrease in the trend of tetracycline use in DID. There is a significant seasonal variation in the use of tetracycline and also a significant seasonal variation trend over time.

The adaptive change-point models can be extended with more change-points. But for the tetracycline use data, convergence was not reached by including more than two common change-points or more than one country-specific random change-point. We have conducted a small-scale simulation study under different scenarios to investigate the change-point model in more detail. The results of this small-scale simulation study show that the change-point model with two (resp. three) change-points fits the data best when the data are generated under the change-point model with two (resp. three) change-points. As for the analyses of the case study, the change-point model with three (resp. four) change-points did not converge. This simulation experiment confirms that the convergence issues we encountered in our data application for the model with three or more change-points are very likely attributable to the absence of three or more change-points.
Appendix I: Estimates for the country-specific change-point

Figure A1 Scatter plot of estimates for the country-specific change-points obtained from fitting Model 6. The vertical line indicates the estimated global change-point.
Source: Authors’ own.
Figure A2  Scatter plot of estimates for the country-specific change-points obtained from fitting Model 7. The vertical line indicates the average for the estimated country-specific random change-points.

Source: Authors’ own.
Appendix II: Estimates for the country-specific change-point indicators

Table A1  Parameter estimates: posterior means (and standard errors) for the country-specific indicators \( I_{ki} \) obtained from fitting Models 4\(^*\), 6\(^*\) and 7\(^*\)

<table>
<thead>
<tr>
<th>Country</th>
<th>Parameters</th>
<th>Model 4(^*)</th>
<th>Model 6(^*)</th>
<th>Model 7(^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>( I_1 )</td>
<td>0.9516(0.2147)</td>
<td>0.9693(0.1724)</td>
<td>0.9629(0.1890)</td>
</tr>
<tr>
<td>Belgium</td>
<td>( I_2 )</td>
<td>1.0000(0.0000)</td>
<td>1.0000(0.0000)</td>
<td>1.0000(0.0000)</td>
</tr>
<tr>
<td>Cyprus</td>
<td>( I_3 )</td>
<td>0.8434(0.3635)</td>
<td>0.8706(0.3356)</td>
<td>0.8794(0.3257)</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>( I_4 )</td>
<td>0.8026(0.3980)</td>
<td>0.8746(0.3111)</td>
<td>0.8963(0.3048)</td>
</tr>
<tr>
<td>Germany</td>
<td>( I_5 )</td>
<td>0.9696(0.1718)</td>
<td>0.9785(0.1450)</td>
<td>0.9740(0.1592)</td>
</tr>
<tr>
<td>Denmark</td>
<td>( I_6 )</td>
<td>0.9054(0.2926)</td>
<td>0.9194(0.2722)</td>
<td>0.9241(0.2648)</td>
</tr>
<tr>
<td>Estonia</td>
<td>( I_7 )</td>
<td>0.7104(0.4536)</td>
<td>0.7515(0.4321)</td>
<td>0.8006(0.3996)</td>
</tr>
<tr>
<td>Spain</td>
<td>( I_8 )</td>
<td>0.6616(0.4732)</td>
<td>0.7266(0.4457)</td>
<td>0.8132(0.3898)</td>
</tr>
<tr>
<td>Finland</td>
<td>( I_9 )</td>
<td>1.0000(0.0000)</td>
<td>1.0000(0.0000)</td>
<td>1.0000(0.0000)</td>
</tr>
<tr>
<td>Greece</td>
<td>( I_{10} )</td>
<td>0.8278(0.3776)</td>
<td>0.8634(0.3434)</td>
<td>0.8869(0.3167)</td>
</tr>
<tr>
<td>Croatia</td>
<td>( I_{11} )</td>
<td>0.9932(0.0819)</td>
<td>0.9930(0.0834)</td>
<td>0.9885(0.1066)</td>
</tr>
<tr>
<td>Hungary</td>
<td>( I_{12} )</td>
<td>0.9578(0.2016)</td>
<td>0.9569(0.1923)</td>
<td>0.9557(0.2057)</td>
</tr>
<tr>
<td>Ireland</td>
<td>( I_{13} )</td>
<td>0.8520(0.3551)</td>
<td>0.8604(0.3466)</td>
<td>0.8623(0.3222)</td>
</tr>
<tr>
<td>Israel</td>
<td>( I_{14} )</td>
<td>0.7719(0.4196)</td>
<td>0.8041(0.3969)</td>
<td>0.8389(0.3677)</td>
</tr>
<tr>
<td>Italy</td>
<td>( I_{15} )</td>
<td>0.8299(0.3757)</td>
<td>0.8457(0.3613)</td>
<td>0.8665(0.3401)</td>
</tr>
<tr>
<td>Lithuania</td>
<td>( I_{16} )</td>
<td>0.9705(0.1892)</td>
<td>0.9773(0.1489)</td>
<td>0.9756(0.1543)</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>( I_{17} )</td>
<td>0.7520(0.4319)</td>
<td>0.7994(0.4004)</td>
<td>0.8286(0.3769)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>( I_{18} )</td>
<td>0.9996(0.0194)</td>
<td>0.9996(0.0212)</td>
<td>0.9999(0.0308)</td>
</tr>
<tr>
<td>Portugal</td>
<td>( I_{19} )</td>
<td>0.8230(0.3446)</td>
<td>0.8908(0.3119)</td>
<td>0.8903(0.3125)</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>( I_{20} )</td>
<td>0.8067(0.3949)</td>
<td>0.8569(0.3502)</td>
<td>0.9025(0.2967)</td>
</tr>
<tr>
<td>Sweden</td>
<td>( I_{21} )</td>
<td>0.5899(0.4920)</td>
<td>0.6698(0.4703)</td>
<td>0.8136(0.3894)</td>
</tr>
<tr>
<td>Slovenia</td>
<td>( I_{22} )</td>
<td>0.7065(0.4554)</td>
<td>0.8022(0.3983)</td>
<td>0.9013(0.2982)</td>
</tr>
<tr>
<td>Slovakia</td>
<td>( I_{23} )</td>
<td>0.9998(0.0158)</td>
<td>0.9997(0.0180)</td>
<td>0.9996(0.0200)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>( I_{24} )</td>
<td>0.9768(0.1505)</td>
<td>0.9999(0.0087)</td>
<td>1.0000(0.0000)</td>
</tr>
<tr>
<td>Iceland</td>
<td>( I_{25} )</td>
<td>0.7986(0.4011)</td>
<td>0.8164(0.3872)</td>
<td>0.8408(0.3659)</td>
</tr>
<tr>
<td>Latvia</td>
<td>( I_{26} )</td>
<td>0.8389(0.3677)</td>
<td>0.8594(0.3476)</td>
<td>0.8678(0.3387)</td>
</tr>
</tbody>
</table>

Note: *Models 4, 6 and 7 are fitted with a country-specific latent indicator \( I_{ki} \).

Source: Authors’ own.

Appendix III: R code

The following WinBUGS code were used in R using the R-package R2WinBUGS to fit the change-point model with one unknown common change-point, one country-specific random change-point and a country-specific latent indicator for the change-point.

```R
# Model
def model{
  # Basic model
  for (i in 1:N){
    Y[i] ~ dnorm(mu[i],tau)
    mu[i] <- (B0 + b1[ID[i]]) + (B1 + b2[ID[i]])*T[i] + (B2 + b3[ID[i]])*
  }
}
```

Statistical Modelling 2013; 13(3): 253–274
\[(T[i]-(C1 + c1[ID[i]]))*\text{step}(T[i]-(C1 + c1[ID[i]]))*\text{change}[ID[i]] + (\alpha + b4[ID[i]] + \alpha\text{Time}*T[i])*\sin(\omega\text{Time}*T[i] + \delta)\]

# Priors for random effects
for (j in 1:M){
  b1[j] \sim \text{dnorm}(0,b0.\tau)
  b2[j] \sim \text{dnorm}(0,b1.\tau)
  b3[j] \sim \text{dnorm}(0,b2.\tau)
  b4[j] \sim \text{dnorm}(0,b3.\tau)
  c1[j] \sim \text{dnorm}(0,c1.\tau)
  \text{change}[j] \sim \text{dbern}(\text{changemean})
}

# Priors for fixed effects
B0 \sim \text{dnorm}(0,0.0001)
B1 \sim \text{dnorm}(0,0.0001)
B2 \sim \text{dnorm}(0,0.0001)
\alpha \sim \text{dnorm}(0,0.0001)
\alpha\text{Time} \sim \text{dnorm}(0,0.0001)
\delta \sim \text{dnorm}(0,0.0001)
C1 \sim \text{dunif}(1,52)
\text{changemean} \sim \text{dbeta}(1,1)

# Hyper priors
\tau \sim \text{dgamma}(0.001, 0.001)
b0.\tau \sim \text{dgamma}(0.001, 0.001)
b1.\tau \sim \text{dgamma}(0.001, 0.001)
b2.\tau \sim \text{dgamma}(0.001, 0.001)
b3.\tau \sim \text{dgamma}(0.001, 0.001)
c1.\tau \sim \text{dgamma}(0.001, 0.001)

sigma <- 1/\tau
sigma_b0 <- 1/b0.\tau
sigma_b1 <- 1/b1.\tau
sigma_b2 <- 1/b2.\tau
sigma_b3 <- 1/b3.\tau
sigma_c1 <- 1/c1.\tau
}

Acknowledgements

The authors would like to thank the Associate Editor and the reviewers for their valuable comments that have led to an improved version of the manuscript. We also gratefully acknowledge support from IAP research Network #P6/03 of the Belgian
Adaptive change-point mixed models applied to data on outpatient use

Government (Belgian Science Policy). The 1997–2005 data collection was funded by a grant from DG SANCO of the European Commission (Grant Agreement 2003211), whereas the 2006–09 data collection was funded by the ECDC (Grant Agreement 2007/001).

References


Girma Minalu et al.


*Statistical Modelling* 2013; 13(3): 253–274

NOT FOR COMMERCIAL USE