

# Some findings on zero-inflated and hurdle Poisson models for disease mapping\*

Corpas-Burgos, Francisca  
*Health Inequalities Research Area  
FISABIO Foundation, Valencia, Spain.*

Garcia-Donato, Gonzalo  
*University of Castilla La Mancha, Albacete, Spain.*

Martinez-Beneito, Miguel A.  
*Department of Statistics and Operations Research,  
University of Valencia, Valencia, Spain.*

## Abstract

Zero excess in the study of geographically referenced mortality data sets has been the focus of considerable attention in the literature, with zero-inflation being the most common procedure to handle this lack of fit. Although hurdle models have also been used in disease mapping studies, their use is more rare. We show in this paper that models using particular treatments of zero excesses are often required for achieving appropriate fits in regular mortality studies since, otherwise, geographical units with low expected counts are oversmoothed. However, as also shown, an indiscriminate treatment of zero excess may be unnecessary and has a problematic implementation. In this regard, we find that naive zero-inflation and hurdle models, without an explicit modeling of the probabilities of zeroes do not fix zero excesses problems well enough and are clearly unsatisfactory. Results sharply suggest the need for an explicit modeling of the probabilities that should vary across areal units. Unfortunately, these more flexible modeling strategies can easily lead to improper posterior distributions as we prove in several theoretical results. Those procedures have been repeatedly used in the disease mapping literature and one should bear these issues in mind in order to propose valid models. We finally propose several valid modeling alternatives according to the results mentioned that are suitable for fitting zero excesses. We show that those proposals fix zero excesses problems and correct the mentioned oversmoothing of risks in low populated units depicting geographic patterns more suited to the data.

---

\*This is a preprint version of the paper: Corpas-Burgos, F; García-Donato, G; Martínez-Beneito, MA. Some findings on zero-inflated and hurdle poisson models for disease mapping. *Statistics in Medicine*, 2018, 37.

# 1 Introduction

Zero excesses have been frequently addressed within the disease mapping literature, see for example Ugarte et al. (2004); Song et al. (2011); Nieto-Barajas and Bandyopadhyay (2013); Musenge et al. (2013); Arab (2015). We consider this problem from a Bayesian perspective, a paradigm frequently adopted in this context (the last four references above are Bayesian). This topic has received considerable attention in recent years. For example, popular Bayesian software such as INLA Rue et al. (2009) has included up to 5 different functions that implement specific models to handle situations of zero excesses. This issue is not exclusive to disease mapping problems but, on the contrary, is related to any type of data taking in general positive integer values (including 0), such as for example Poisson, binomial or negative-binomial distributed data. Zero excesses are a source of overdispersion caused by a disagreement between the data and the distribution assumed: we have more zeroes in our data set than the proposed distribution could reasonably explain. As a consequence, zero excesses are features inherent to particular combinations of distributions (or models in general) and data sets, but not intrinsic to particular data sets. The presence of a large number of zeroes is symptomatic of a zero excess situation, but not necessarily indicative of one since observing many zeroes could be perfectly compatible with a Poisson distribution with a low expected value. Therefore an indiscriminate use of models dealing with zero excesses is, in principle, not necessary. In this sense several procedures have been developed for assessing zero excesses in specific problems like Van Der Broek Van Der Broek (1995) or Bayarri et al. Bayarri et al. (2008) which deal with this issue on Poisson data with constant or covariate-dependent expected cases.

Many disease mapping studies have incorporated zero excesses modeling strategies in the analysis of mortality spatial data. Nevertheless, to our knowledge, it has not been extensively tested whether zero-specific treatments should be routinely used in this context or if, on the contrary, the standard Poisson assumption (with spatially varying random effects) fits regular mortality data well enough. Moreover, it is rarely the case that the pursued positive effect of such treatments is checked with the unexpected possible consequence that the original data misfit, in terms of zero counts, still remains. A motivating aspect of this research is to shed some light on these two relevant questions using a real extensive setting with 540 areal units and 46 geographical patterns corresponding to roughly 27 different causes of mortality. In particular we consider the zero-inflated and hurdle Poisson models, the most popular models in the related literature. With respect to the first question, in roughly 15 patterns out of the 46 considered (barely 32% of the cases) we have observed a serious departure from the number of zeroes predicted with traditional disease mapping models, while the need for specific zero excess treatment for the rest is questionable. Our findings for the second question are more worrisome from a practical point of view. As we report, a preventive extra zero modeling may be totally innocuous for the zero-inflated approach without a particular modeling of the zero-specific component. For hurdle models the situation is even worse, since the estimations of the underlying

ing risks can be dramatically influenced by spurious circumstances like the spatial distribution of the population along the region of study. The consequence is relevant since, for many cases, we could be reporting nonsense estimations based on an unneeded zero excess treatment.

The results observed in the real application indicate that for regular zero-inflated and hurdle Poisson proposals a specific modeling of the probability of zero-excess is needed in order to construct satisfactory methods. This is admittedly the path followed by many applied works in the literature (references will be given). Nevertheless, as we prove, such modeling has an unforeseen important difficulty, namely that conditions for impropriety of the posterior distribution (an invalidating fact for many not so formal related approaches) are very soft. These theoretical results make the assignment of the prior distributions a very delicate issue, preventing the use of highly popular “casual” non-informative priors frequently implemented by-default in specialized Bayesian software. Our result is quite general and affects several components of the model (like fixed effects or variances of the random effects) and many of the link functions (e.g. logit or probit). Additionally, we propose alternative modeling strategies that, as we argue, are safer in terms of validity of the results.

This paper is divided into 6 sections. Section 2 introduces the BYM model Besag et al. (1991), the most popular proposal for disease mapping and two specific refinements, zero-inflation and hurdle Poisson modeling, in order to cope with zero excesses. Section 3 shows the performance of these proposals in the analysis of the Valencian Mortality Dataset. Section 4 contains the main theoretical results about conditions for impropriety of posterior and presents some valid proposals to overcome the problems encountered. Section 5 illustrates the dangers of using vague prior distributions on some particular variables of models treating zero excesses and reassesses the behavior of the proposals made in Section 4 on the previous Valencian Mortality Dataset. Finally, Section 6 draws some conclusions from the results derived in this paper.

## 2 Some proposals for treating zero excesses in disease mapping

The goal of disease mapping is dealing with the sparse information in the observed counts of some health outcome over a set of areal units. In general these units are small in statistical terms, with frequent low observed counts, that makes them noisy and weakly informative of the underlying risk of the disease for many of them. Thus, statistical modeling is needed for drawing acceptable risk estimates in those units. The models used for this task mainly rely on spatial conditional autoregressive random effects to induce geographical dependence on the risk estimates and therefore to increase the amount of information used to estimate them. Among the models using these random effects we highlight one that is particularly popular, the Besag, York and Mollié’s model Besag et al. (1991), BYM henceforth. For this model, data  $\{O_i : i = 1, \dots, I\}$  representing observed

counts on the areal units are modeled as

$$O_i | R_i \sim \text{Poisson}(E_i R_i), \quad i = 1, \dots, I$$

where  $E_i$  are the expected counts for each unit, typically calculated by means of some age standardization, and  $R_i$  are the corresponding risks that we would like to estimate. Regarding the modeling of this last term, BYM defines the log-risks as:

$$\log(R_i) = \mu + \phi_i + \theta_i, \quad (1)$$

where  $\mu$  stands for an intercept modeling the mean of the log-risks and the two subsequent terms are Gaussian random effects. The term  $\phi$  follows an intrinsic conditional autoregressive (ICAR) distribution, i.e. their components are assumed to have the following prior conditional distributions:

$$\phi_i | \phi_{-i}, \sigma_\phi \sim N \left( n_i^{-1} \sum_{j \sim i} \phi_j, n_i^{-1} \sigma_\phi^2 \right), \quad i = 1, \dots, I$$

where  $n_i$  stands for the number of neighboring areas of unit  $i$ , the subindex in  $\phi_{-i}$  indicates all terms in  $\phi$  excepting its  $i$ -th component and the subindex  $j \sim i$  denotes all those units  $j$  which are neighbors of  $i$ . This definition can be further elaborated introducing some parameters in order to weight the contribution of some units with respect to others, although we will not use that option. This term induces spatial dependence on  $\mathbf{R}$  and accounts for those factors of regional scope which take effect on several contiguous units, making them similar. In contrast, the term  $\theta$  in expression (1) accounts for risk factors of very limited geographical scope that take an effect just on isolated areal units and make their risks different to those of their surrounding units. The terms introducing independent variability on the risks are modeled as independent Gaussian random effects, i.e

$$\theta_i | \sigma_\theta \sim N(0, \sigma_\theta^2), \quad i = 1, \dots, I.$$

The amount of spatial dependence in  $\mathbf{R}$  depends on the balance between  $\sigma_\phi$  and  $\sigma_\theta$ . If the first has higher (respectively lower) values, in comparison to the second, the final pattern will show substantial spatial dependence (independence).

Besides the spatial modeling that could be done with the BYM model the data available may require a specific treatment of the observed zero counts if the model fitted could not explain the amount of observed zeroes in the data set. The most used tool for dealing with zero excesses is zero-inflation. Specifically, in case of modeling observed counts with a Poisson likelihood, the resulting model is known as Zero-inflated Poisson (ZIP) Lambert (1992). In its simplest form ZIP models assume the observed counts to follow a mixture of a degenerate distribution with all its mass at zero and a  $\text{Poisson}(\lambda)$  distribution, with weights  $1 - \pi^Z$  and  $\pi^Z$ , respectively. This inflates the amount of zeroes expected by the Poisson distribution as a function of  $\pi^Z$ .

ZIP models for disease mapping fuse the simplest ZIP approach just introduced with spatial models (such as BYM). This yields flexible ZIP

models with different (and dependent)  $\lambda_i$ s, acknowledging that the studied data set may have more zeroes than those reproduced by BYM. Being more precise, a ZIP version of the BYM model could be formulated as follows: The observed data are assumed to follow a Poisson distribution of mean  $E_i R_i Z_i$ , where  $E_i$  stands for the expected cases,  $R_i$  for the spatially-varying risks in the Poisson distribution of the BYM model and  $Z_i$  for a binary variable modeling if the observed counts correspond to an extra-Poisson zero ( $Z_i = 0$ ) or correspond to a value coming from the Poisson distribution ( $Z_i = 1$ ). The risks  $R_i$  would be modeled as in equation (1) and the  $Z_i$ s would follow a *Bernoulli*( $\pi^Z$ ) distribution, with unknown  $\pi^Z$ . For this model the smoothed Standardized Mortality Ratios (SMR) would be computed as  $R_i Z_i$ , i.e. a mixture of the BYM-based risks and 0. Examples of applications that adopt this modeling approach include Gschlößl and Czado (2008); Song et al. (2011); Musenge et al. (2013)

As an alternative to ZIP, data sets showing zero excesses are sometimes modeled as hurdle Poisson models Mullahy (1986), simply hurdle models henceforth. This proposal assumes the data to follow a mixture of a degenerate distribution with all its mass at zero and a Zero-truncated Poisson distribution. That is, in contrast to ZIP models, all observed zeroes in hurdle models are assumed to come from the zero-degenerate distribution. Thus, the parameter  $1 - \pi^H$  in hurdle models represents the probability that a given areal unit has zero observed cases instead of the percentage of extra-Poisson zeroes, the interpretation of  $1 - \pi^Z$  in ZIP. As for ZIP, hurdle models are combined with specific disease mapping proposals, such as BYM, in order to yield flexible spatial models accounting for zero excesses.

More specifically, for a hurdle version of the BYM model the observed counts  $O_i$  are assumed

$$P(O_i | \pi^H, \boldsymbol{\mu}) = (1 - \pi^H)^{I_{\{0\}}(O_i)} \left( \pi^H \left( \exp(-\mu_i) \frac{\mu_i^{O_i}}{O_i!} (1 - \exp(-\mu_i))^{-1} \right) \right)^{I_{(0, \infty)}(O_i)}$$

where  $\mu_i = E_i R_i$  and  $I_{\Omega}(x)$  is the indicator function for the condition  $x \in \Omega$ . The risks  $R_i$ s in this model would follow expression (1). For this proposal the smoothed SMR for the  $i$ -th unit should be computed as  $\pi^H (\mu_i / (1 - \exp(-\mu_i))) / E_i$  Neelon et al. (2013) where  $\pi^H$  is the probability of belonging to the truncated Poisson component and  $\mu_i / (1 - \exp(-\mu_i))$  is the expected value given that the observation belongs to that component. This term is divided by the expected cases  $E_i$  since  $\pi^H (\mu_i / (1 - \exp(-\mu_i)))$  would be the mean of  $O_i$  but we want to draw an estimate of  $O_i / E_i$  instead.

Both ZIP and hurdle versions of the BYM model, as introduced above, are posed under a Bayesian approach since BYM is also originally formulated from a Bayesian point of view. As a consequence all the parameters in BYM, ZIP and hurdle in this paper will have their own prior distribution. We will discuss prior distributions for these models more in depth in Section 4. Nevertheless, for now, we will not pay them further attention as they will be mostly irrelevant for the issues discussed in the next Section. Anyway, the prior distributions used in our analyses could be considered as regular prior choices for these models in the literature. Full

details on the priors used can be found at Annex C of this paper in the supplementary material, which contains all the code used for its analyses.

In the next section, we implement these three different approaches in a real extensive setting in order to assess their practical utility. As we will see, the results are far from being as satisfactory as expected.

### 3 An initial analysis of the Valencian Mortality Data: A motivating application

Now that we have introduced the BYM model and two potential tools to cope with zero excesses, we are going to test their performance in an extensive real setting. We will pay particular attention to their fit in terms of the number of predicted zeroes in comparison to those actually observed. Our particular data set for this task is the mortality data used in the Spatio-temporal Mortality Atlas of the Valencian Region (1987-2006) Zurriaga et al. (2010) in which we have ignored the temporal component. This atlas studies 46 geographical patterns corresponding to the distribution of 27 causes of mortality for each sex, excepting some particular combinations without enough deaths or without biological sense (e.g. prostate cancer in women). Mortality is disaggregated at the municipal level in a total of 540 municipalities of very different sizes, ranging from 22 to about 750,000 inhabitants (year 2000). Thus, observed deaths are expected to show substantial variability between municipalities, with some locations showing systematically 0 deaths for most of the causes.

The number of observed zeroes for the 46 geographical patterns analyzed ranges from 4 to 243. As we mentioned, such numbers, although sometimes high do not necessarily mean zero excesses. They can simply represent low mortality for any of those causes or low population for some municipalities. Thus, for assessing zero excesses with regard to the models introduced in Section 2 we have run each of them on the available data. For each model and cause of death we have sampled values from the posterior predictive distribution of the observed deaths for each municipality and we have compared those samples against the observed values. Specifically, we have compared the number of zeroes observed for each cause of mortality and those predicted by the models from the MCMC.

Table 1 shows the results obtained for some causes of deaths, specifically the first 10 causes. The full table with all 46 analyses made is annexed as supplementary material to this paper (Annex B [PROVISIONALLY AVAILABLE AT <http://www.uv.es/mamtnez/Zeroes.htm> DUE TO THE SIZE OF THIS DOCUMENT]). The second column of Table 1 contains the number of zeroes observed for each data set meanwhile the next 3 columns correspond to that same number as predicted by each model run. Namely, we have run the BYM model without any particular treatment of zeroes as well as ZIP and hurdle versions of that same model. Bold fonts in Table 1 denote those combinations of models and data sets evidencing zero excesses according to their predictive intervals.

All models in this paper were run in WinBUGS and the code for each of them can be found as annex material at Annex C. An R-markdown

document with all the analysis carried out can be found in that Annex. Three chains were run for each model and data set with 50,000 iterations, whose first 5,000 iterations were used as burn-in period. Of these, one of every 135 iterations was saved yielding a final sample size of 1,002 iterations. Convergence was assessed by means of the Brooks-Gelman-Rubin statistic (we required this to be lower than 1.1 for each variable in the model) and the effective sample size (required to be at least 100 for every variable in the model) implemented with the `R2WinBUGS` package of R.

[INSERT TABLE 1 HERE]

Table 1 (and in more detail the full table in Annex B) shows how BYM may fit quite poorly the number of zeroes for certain data sets. Namely, for 15 out of the 46 data sets considered the 95% posterior predictive intervals for the number of zeroes in BYM did not contain the real observed zero counts and for 5 additional data sets the upper limit of that interval coincided with the observed zeroes –this seems excessive since we would expect a priori just 2 or 3 of the observed zeroes to lay outside of the predictive intervals–. The main conclusion is a substantial lack of fit for BYM in terms of the number of zeroes predicted and therefore a general advice for specific treatment of those cases. On the contrary, BYM seems to accommodate well the number of zeroes in the rest of datasets (26), making it questionable the need for particular treatments of excess of zeroes in those settings.

With respect to the approaches with a particular treatment of zeroes, the results are not satisfactory for different reasons. Surprisingly, ZIP does not help much in fitting more zeroes and 11 out of the 46 original data sets showed 95% posterior predictive interval which do not contain the real observed number of zeroes and in 1 occasion the upper limit of the interval coincided with those zeroes. This performance, although better than that of BYM is also unacceptable since the number of predictive intervals that do not contain the corresponding observed value is far above of that corresponding to the nominal probability of the interval. On the contrary, for hurdle, all intervals contained the observed number of zeroes. Nevertheless, this better fit of the proportion of zeroes has a pernicious effect on the estimations of the SMRs that make them barely reliable. To understand this effect, we have represented in Figure 1 choropleth maps for the SMRs fitted for all three models in Table 1 for rectum cancer in males, one of the cases where the presence of a zero excess for BYM and ZIP is evident.

[INSERT FIGURE 1 HERE]

We first highlight that the maps for BYM and ZIP are quite similar for this data set and in general for all diseases fitted (maps not shown). This is not surprising according to the fit of the  $\pi^Z$  parameter in ZIP for all the causes. The posterior mean for this parameter, which measures the weight of the Poisson side of ZIP models, for all 46 data sets ranges from 0.973 to 0.998. Thus even though ZIP models should be able to fit zero excesses, they refuse to do it by minimising the weight of the zero-specific component. This may be a consequence of the implementation

where the  $\pi^Z$  parameter is common to all municipalities. So, decreasing  $\pi^Z$  for making room to more zeroes in smaller municipalities also entails an increase in the probability of observing zeroes in large cities where that probability is virtually zero. Since the amount of information available in large municipalities is much higher than that in the smaller ones, ZIP decides to reject the zero-specific term as its contribution is more harmful, in likelihood terms, for the large municipalities than beneficial for the smaller ones (those with potential lack of zeroes).

The SMR map for hurdle shows a weird pattern completely different to BYM. This map shows a polarized pattern with high SMRs in the smaller municipalities and low SMRs for the rest. This pattern is systematically repeated for most of the data sets analyzed (maps not shown), being more evident for those data sets with more observed zeroes. In our opinion, this is also an effect of having a common  $\pi^H$  parameter for all the municipalities. In contrast to ZIP now  $\pi^H$  for the different data sets is not so close to 1, being its posterior mean always very close to the proportion of non-zero observed counts for each data set. Nevertheless, as mentioned in the previous Section,  $\pi^H$  takes also an effect on the calculation of the SMRs for this model, decreasing the mean of the Poisson component in that same proportion. For the small cities this makes the number of predicted zeroes to be increased but, alternatively, for the non-small cities this makes the SMRs to be underestimated as evidenced in Figure 1.

These results suggest that in our case both zero-specific treatments using these naive proposals which put the same zero-specific probabilities to all units do not seem a good choice. At least in our extensive analysis, ZIP does not seem to have a clear effect with regard to the baseline BYM model. For hurdle the particular (naive) treatment of zeroes makes misleading the corresponding SMRs map. Since considering a common probability for the zero-specific side seems to be the cause of these problems, we will explore from now on the opportunities and benefits that the modeling of those probabilities could bring.

## 4 Modeling of the probability of observing a zero

One of the most valuable advantages of Bayesian hierarchical models is the possibility of modeling particular features of the data that we could be interested in. Nevertheless, that ability is not always good as it can lead us to models which are not necessarily well formulated and therefore to misleading or plainly wrong results.

As introduced in the previous section, both ZIP and hurdle models require a particular treatment of the assignment of the observed counts to any of the two processes intervening in each of them. That assignment follows a binary process which, up to now, has depended on a single parameter  $\pi$  common to all areal units. We will denote  $\pi$  when we refer indistinctly to either  $\pi^Z$  or  $\pi^H$ . The obvious alternative to a common probability is modeling unit-specific  $\pi_i$ s by means of, for example, logistic regression. This has been repeatedly done in the disease mapping context



for both ZIP Dalrymple et al. (2003); Gschlößl and Czado (2008); Neelon et al. (2010); Musenge et al. (2013); Nieto-Barajas and Bandyopadhyay (2013) and hurdle models Dalrymple et al. (2003); Neelon et al. (2010, 2013); Uppfill-Brown et al. (2014); Neelon et al. (2014); Arab (2015). That is, following several of the proposals in the literature, for both ZIP and hurdle models we will consider from now on

$$\text{logit}(\pi_i) = \mathbf{x}_i\boldsymbol{\beta} + \varphi_i \quad (2)$$

where  $\boldsymbol{\beta}$  model the effect of some set of covariates  $\mathbf{X}$  and  $\boldsymbol{\varphi}$  is a vector of (possibly spatial) Gaussian random effects modeling the effect of those factors that cannot be explained by  $\mathbf{X}$ .

In the next subsection we introduce a series of results of great interest for the models that we want to explore now. Namely, we have found important posterior propriety problems in hurdle and ZIP models when the vector of probabilities  $\boldsymbol{\pi}$  is modelled with either fixed or random effects. This makes that modeling quite tricky and caution has to be taken in order to avoid flawed modeling proposals. These results will determine some ZIP and hurdle specific proposals that should be avoided in general. We will discourage the use of those models particularly in a non-informative or objective setting. Additionally, Subsection 4.1 will allow us to focus on some valid proposals with different  $\pi_i$ s, that will be later developed at Subsection 4.2.

#### 4.1 Some theoretical results warning against the use of certain popular casual non-informative priors

Once a suitable model is specified, when it comes the need to assign the prior distribution, the applied literature is flooded with casual possibilities that include, for example, a uniform prior on fixed effects parameters or its ‘proper’ counterpart of a normal density with an arbitrarily large variance. Obviously, these proposals are valid (in the sense that results are covered by laws of probability) as far as the associated posterior distribution is proper (see below the comment on the ‘vague’ counterparts), a property that it is rarely checked in practice. We do so here and conclude that the conditions for propriety of the posterior are quite severe and are not fulfilled by many popular non-informative choices.

We start by introducing some results for hurdle models which consider  $\boldsymbol{\pi}^H$  as proposed in (2). The proofs and full formulation of the results introduced in this subsection are provided in Annex A of the supplementary material to this paper.

First, we have shown that the hurdle model with  $\boldsymbol{\pi}^H$  modeled as in (2) is problematic since some issues arise on the use of both fixed and random effects in that expression. As stated in Corollary 1 (Annex A) the use of random effects with improper prior distributions for  $\sigma$ , the standard deviation of the random effects, yields an improper posterior distribution regardless of the other elements in the model. This means that the use of random effects in (2) with many of the default prior choices in the literature for their variability should be avoided. Besides, if for the

$j^*$  column of  $\mathbf{X}$ ,  $x_{ij^*}$  is positive for every  $i$  with  $O_i > 0$  and negative otherwise (or vice versa) and the prior distribution of  $\beta_{j^*}$  is improper for large positive (respectively negative) values then the posterior distribution is also improper. So, we could also have posterior impropriety problems using fixed effects for modeling  $\boldsymbol{\pi}^H$ . Fortunately, this condition (although just a sufficient condition, not necessary, for impropriety) will not be fulfilled easily since it depends in a binary manner on all the (random) values of the outcome of the model. That binary condition should be fulfilled for all the observed outcomes which is not that easy, especially for regions with a large number of units. Additionally, Corollary 1 is very general since as stated there, these results above would hold equally for other common link functions in (2) such as probit or tobit; they would also hold for non-Poisson based likelihoods such as binomial or negative binomial and for other different spatial structures (with positive-definite covariance matrices) besides BYM.

The situation for ZIP models is not better. As stated in Corollary 2 (Annex A) the use of random effects in (2) for ZIP models is as problematic as for hurdle models, since it yields an improper posterior distribution under the same premises. Moreover, the results for ZIP are equally general since they also apply for different link functions, likelihood families and spatial structures for the mean of the non-zero process. Nevertheless, the case of fixed effects is substantially different (worse) for ZIP since these yield posterior impropriety more easily than hurdle models. Thus, we have found that a sufficient condition for posterior impropriety in ZIP would be that for any column  $j^*$  of  $\mathbf{X}$ ,  $x_{ij^*} > 0$  (respectively  $x_{ij^*} < 0$ ) for all  $i$  and  $\beta_{j^*}$  to diverge for large positive (respectively negative) values. This condition is much more general since this could be fulfilled by design of the covariates, regardless of the observed counts  $\mathbf{O}$ . In principle we could easily get rid of this issue by, for example, subtracting the mean of any of the covariates in the model but the problem would remain for the intercept. The intercept is positive for all the units in the model so any improper prior distribution on its corresponding term in  $\boldsymbol{\beta}$  would yield an improper posterior distribution, independently of the additional problems that the rest of covariates in the model could also entail.

One could be tempted to use vague proper prior distributions, instead of improper priors, as a possible strategy to avoid impropriety issues. This is a procedure frequently found in the literature, supposedly to avoid MCMC convergence problems. Nevertheless, according to the results stated above, these 'convergence problems' are a numerical manifestation of the more worrisome fact of having an improper limiting posterior distribution. Berger Berger (2006) argues that the use of a vague prior mimicking an improper prior with an associated improper posterior can only hide but not solve the problem. In our context, this of course invalidates the use of standard approaches like a vague Normal priors on each component of  $\boldsymbol{\beta}$ , vague Gamma priors on the precision of the random effects or Uniform prior distributions with a large upper limit on their standard deviations. Interestingly though, a tentative use of vague proper priors could serve as a diagnostic test to detect possible underlying problems of posterior impropriety. For instance, results assuming a uniform prior on the standard deviation of  $\boldsymbol{\varphi}$  with an arbitrary large upper

limit show high sensitivity to such upper limit, warning clearly about the possible impropriety of the posterior distribution.

## 4.2 Some valid proposals for modeling $\pi$

The previous subsection has stated some procedures to be avoided when modeling the probabilities  $\pi$  in both ZIP and hurdle models. One option would be to use informative prior distributions for  $\beta$  and  $\sigma$ . In this sense Agarwal et al. (2002) have made one proposal of informative prior distributions for  $\beta$  for ZIP models. Nevertheless, we would rather avoid informative prior distributions. So, we will propose some (non-informative) procedures for modeling  $\pi$  that do not fulfill the conditions for posterior impropriety stated above. Regretfully, we do not have a proof for the posterior propriety of these proposals since the impropriety conditions formulated are just sufficient but not necessary. In any case, these new proposals do not fall into the premises of those results, in contrast to many of the proposals formulated in the literature. Moreover, in our experience, these new proposals do not seem to show at all any of the MCMC convergence problems appearing when one of the models yielding improper posterior (according to the conditions stated in the previous subsection) were used. We formulate now 3 separate modeling proposals.

### Fixed effects modeling:

Although, as described above, the use of random effects for modeling  $\pi$  is quite problematic, the use of fixed effects for modeling  $\pi^H$  in hurdle seems a much less troublesome option. Thus, a potentially valid modeling proposal (we will refer to this as *FE* (Fixed Effects) henceforth) would be to consider a hurdle model as defined in Section 2 with

$$\text{logit}(\pi^H) = \mathbf{X}\beta.$$

A suitable proposal that could be used in principle for any disease mapping model would be:  $\mathbf{X} = [\mathbf{1}_I, \log(\mathbf{E})]$ , where  $\mathbf{E}$  stands for the vector of expected values used in the Poisson likelihood of hurdle models. We have taken the logarithm of the expected values to avoid any potential effect of the usually skewed distribution of this variable caused by the presence of very few large cities. According to the results above this could yield an improper posterior distribution if  $O_i$  is positive for each region with  $E_i > 1$  and  $O_i = 0$  otherwise (or vice versa). But, for a reasonably high number of areal units this condition seems very unlikely to be fulfilled.

This proposal models the logit of the probabilities of non-zeroes as a function of the expected observations at each areal unit. This seems quite reasonable since units with lower expected counts would show more easily zero observed counts meanwhile those larger units will show positive counts in general. This could be achieved for  $\beta_2$  (the coefficient corresponding to the log-expected cases) taking positive values. For this proposal we will consider an improper uniform prior distribution for each component of  $\beta$ . This is because we specifically want to avoid the use of vague prior distributions that could hide posterior impropriety problems

into just MCMC convergence problems due to the almost impropriety of posterior distributions.

Interestingly, note the link between this proposal and the EZIP1 proposal in Song et al. (2011) Song et al. (2011). In that paper a ZIP model with  $\pi_i^Z = \frac{E_i}{\delta + E_i}$  is proposed. A logit transformation of this expression yields  $\text{logit}(\pi_i^Z) = \log(E_i) - \log(\delta)$  which would be a ZIP version of the FE model just proposed. However, note that this model is valid since  $\delta$  is assumed to have a *Unif*(0, 1) prior distribution in the EZIP1 model which yields an exponential distribution of mean 1 on  $-\log(\delta)$ . This proper prior obviously avoids any potential impropriety on the posterior distribution.

### **Nested fixed effects modeling:**

The use of expected values as a surrogate of the (population) size of the areal units in the FE modeling seems quite reasonable. Nevertheless, this does not depend at all on the probabilities of non-zeroes resulting from the Poisson side of hurdle models:  $\pi^P = 1 - \exp(-\mu)$ . Although these probabilities have been evidenced to produce some misfit in the data in terms of zero excesses, they could be also used as sensible covariates for modeling the probabilities of zeroes  $\pi^H$ , instead of just  $\mathbf{E}$ . This approach was already introduced in the zero-altered model of Heilbron (1994) Heilbron (1994). These probabilities  $\pi^P$  would not just take into account the size of the areal units, through the expected counts  $\mathbf{E}$ , but also the risk attributed to any of them by the Poisson side of the model. These risks could be an additional source of information making considerable improvements as compared to the use of simple expected counts. Thus, our second proposal for modeling  $\pi^H$  in Hurdle models would be

$$\text{logit}(\pi_i^H) = \text{logit}(\pi_i^P) + \gamma.$$

This would be an alternative fixed effects logistic modelling of  $\pi^H$  using  $\text{logit}(\pi^P)$  as an offset. The values of that offset would be leveraged by  $\gamma$  so that if it takes values close to 0 this model would reproduce the probabilities in the Poisson layer, even for zero-counts, meanwhile for  $\gamma < 0$  the zero-specific probabilities would be inflated in regards to the Poisson model. Note that in case of adapting this modeling to the hurdle-BYM model in Section 2, the original (uninflated) BYM model could be reproduced within this proposal by making  $\gamma = 0$ , thus we will henceforth refer to this model as *NFE*, Nested Fixed Effects model. Once again we will consider an improper uniform prior distribution for  $\gamma$  so that any potential posterior impropriety problem in this model appears.

### **Geometric modeling:**

Since resorting to logit (or probit, tobit) regression has proved to bring lots of problems into ZIP and hurdle models, we could try to avoid those transformations in order to make sensible proposals. Thus, making

$$\pi_i = 1 - (1 - \pi^G)^{E_i}$$

seems a reasonable proposal for both ZIP and Hurdle models. For this proposal we would have that the probability of observing a zero count for a unit with  $n$  expected cases is  $(1 - \pi^G)^n$ , where  $1 - \pi^G$  is that same probability for a unit with 1 expected case. This geometric progression also holds for the Poisson process where the probability of observing zeroes with  $n$  expected cases  $exp(-n \lambda) = exp(-\lambda)^n$  follows that same relationship. Thus, the probabilities of zero counts for this proposal are in agreement with the Poisson side of the model. For  $\pi^G$ , which can be interpreted as the probability of observing a positive count for units with one expected case, we set a Uniform prior distribution between 0 and 1. Since this prior is proper we avoid any posterior impropriety coming from this term. One of the main advantages of this model is that since the modeling of  $\pi$  does not rely on any improper prior distribution this model could be also set up for ZIP modeling. This is contrast to the previous proposals whose ZIP counterparts would be discouraged since they rely on fixed effects logit modeling. We will refer to the ZIP and hurdle versions of this model henceforth as *ZGeo* and *HGeo* respectively.

## 5 Empirical illustration of the modeling proposals introduced

We start this Section by illustrating the problems induced on ZIP and hurdle models by arbitrary prior vagueness. With this section we seek to make clear how prior problems are not just present for improper prior distributions but also for vague proper priors, which are commonly used in ZIP and hurdle disease mapping models. Finally, we will show how the modelling proposals introduced in the previous section perform with the same datasets used in Section 3 where naive ZIP and Hurdle models showed a deficient performance.

### 5.1 An illustration of the prior vagueness problems in ZIP and hurdle models

We are going to illustrate the dangers of using vague proper priors, instead of improper priors, for modeling  $\pi^Z$  and  $\pi^H$  in ZIP and hurdle models. We have already proved that using improper priors for some variables in these models would yield improper posteriors but we want to evidence that using vague proper prior does not seem to be a safe option in any case. Thus, we have run two separate models in this study: a ZIP model with  $logit(\pi_i^Z) = \alpha$  for  $i = 1, \dots, I$  and a hurdle Model with  $logit(\pi_i^H) = \alpha + \gamma_i$  and  $\gamma_i \sim N(0, \sigma_\gamma)$  for  $i = 1, \dots, I$ . These models are somewhat naive, indeed, as mentioned in the paper the ZIP model proposed will not fit in general any risk excess, and additional regressors could be used for modelling both  $\pi^Z$  and  $\pi^H$  in order to improve them. Nevertheless, we have preferred to keep these models as simple as possible in order to illustrate the prior specification problems that they show. We have run these two models on the rectum cancer data set that has also illustrated the results in Section 3. All models were run in WinBUGS.

Regarding the ZIP model mentioned, we have run it for several different prior choices for  $\alpha$ :  $\alpha \sim N(0, \sigma_\alpha^2)$  for  $\sigma_\alpha^2$  equal to 10, 100, 1000 and 10000. For the first of these choices  $\alpha$  has a posterior mean of 5.26 and a 95% posterior credible interval of [3.22,8.68]. For  $\sigma_\alpha^2 = 100$  we obtain a posterior mean of 10.85 and a credible interval of [3.98,24.33]. For  $\sigma_\alpha^2 = 1000$  we obtain a posterior mean of 18.87 and a credible interval of [5.18,35.15]. Finally, for  $\sigma_\alpha^2 = 10000$  WinBUGS finds a numerical error (TRAP 66) and is not able to run this model. As we see, posterior inference on  $\alpha$  completely depends on the prior distribution set for this parameter. None of the models run, excepting that with  $\sigma_\alpha^2 = 10000$ , show any evident convergence problem. Thus, someone fitting these models without an additional sensitivity analysis, such as ours, will accept as good the results for any of the models run, when these models are just hiding the impropriety problems of an hypothetical improper prior choice for  $\alpha$ . Note that as we increase  $\sigma_\alpha^2$ ,  $\alpha$  increases steadily, giving zero probability to the zero-specific component. This reinforces the idea that naive ZIP proposals with logit modeling of  $\pi^Z$  do not fit appropriately zero excesses.

Regarding the random effects hurdle model, we have run it also with different prior distributions for  $\sigma_\gamma$ :  $\sigma_\gamma \sim Unif(0, U_\gamma)$  for  $U_\gamma$  equal to 2, 10 and 100. For  $U_\gamma$  equal to 2 the posterior mean of  $\sigma_\gamma$  is equal to 1.1 with 95% posterior credible interval [0.1,2.0]. For  $U_\gamma$  equal to 10 the posterior mean of  $\sigma_\gamma$  is equal to 6.4 with 95% posterior credible interval [0.7,9.8]. Finally, for  $U_\gamma$  equal to 100 the posterior mean of  $\sigma_\gamma$  is equal to 69.2 with 95% posterior credible interval [17.0,99.0]. Note how the upper limits of the posterior credible intervals for  $\sigma_\gamma$  are always very close to  $U_\gamma$ , pointing out the informativeness of these supposedly uninformative choices. Thus, in summary, we see how the posterior distribution of  $\sigma_\gamma$  heavily depends on the (arbitrary) vagueness of its prior distribution, which makes unadvisable the use of arbitrary vague proper priors for  $\sigma_\gamma$  as a safe substitute of an improper prior distribution.

## 5.2 A re-analysis of the Valencian Mortality Dataset

We turn back once again to the analysis of the Valencian Mortality Dataset in Section 3. We have run all 4 models proposed in the previous subsection: FE, NFE, HGeo and ZGeo, on the diseases considered there. First, we have assessed their fit in terms of the number of zero counts reproduced, i.e. the equivalent of Table 1 but for these new models. Table 2 in Annex B of the supplementary material to this paper shows, for all of them, the posterior medians and 95% credible intervals for the number of predicted zeroes. As a summary, in contrast to the results shown in Table 1, the posterior predictive distribution for the number of zeroes in all 4 models agree with those numbers observed for the real data sets. Namely, all 3 hurdle models yield similar results to the hurdle model in Table 1 with the posterior predictive median for the number of zeroes in the data sets always very close to the real observed zeroes. The modeling of the probabilities of zeroes in ZGeo has made a great improvement over naive ZIP models since for ZGeo the predictive posterior median for the number of zeroes is always very close to the real observed zeroes. All 95% credible intervals for the number of predicted zeroes for all diseases and models

contain the real observed zeroes as would be expected in models which are performing an explicit modeling of that particular feature in the data.

Second, we have also compared the fit of these models in general terms by using the Deviance Information Criterion (DIC) proposed by Spiegelhalter et al. Spiegelhalter et al. (2002). The DICs for all models and 46 data sets, with their corresponding deviances and number of effective parameters, can be found at Table 3 of Annex B. Regarding the FE model its DIC is higher than that of the BYM model for 43 out of 46 data sets so its performance in general does not seem very satisfactory. Although the FE model is more complex than BYM (has two additional parameters) the deviances obtained are in general substantially higher than those of BYM models. This suggests that the modeling proposed in FE is worse than that of the BYM model, thus maybe a linear function of  $\log(\mathbf{E})$  is not as good as it could seem in principle. As a consequence we will not pay further attention to this model from now on. The NFE model attains better DICs for 11 out of the 15 data sets identified as having zero excesses. Meanwhile, for just 5 out of the remaining 31 data sets with no evidence of zero excess NFE was better in terms of DIC, as could be expected since BYM is less complex than NFE and for these data sets NFE should not yield any improvement. Thus, NFE attains in general lower DICs in those settings where it would be expected. Regarding HGeo, it attained 6 out of 15 DICs lower than BYM for those data sets needing a particular treatment for zeroes and 2 out of 31 times was lower for those data sets that did not need that treatment in principle. Finally, ZGeo also obtained similar results to HGeo, improving BYM in 5 out 15 times where zero excesses were evidenced and 8 out of 31 times when these were not so evident. Thus the results of Geometric models are overall satisfactory although not as good, in terms of DIC, as those of NFE.

Regarding the estimates of the parameters in the models proposed, those of NFE showed a particularly coherent performance. Thus, for all data sets needing zero treatment the parameter  $\gamma$  in the model attained a 95% posterior credible interval completely below 0 (we mentioned that  $\gamma < 0$  should be a sign of zero correction with respect to BYM). On the contrary, for only 1 out of the 31 data sets not showing zero excesses the 95% credible interval for  $\gamma$  was completely below zero. Posterior means and 95% credible intervals for  $\gamma$  for all 46 data sets can be found at Table 4 in Annex B to the paper. We do not find anything particularly interesting in the  $\pi^G$  estimates obtained in the Geometric models. These parameters have a cumbersome interpretation since they are referred as the probability of the zero-specific term for units having  $E = 1$ , but each data set and spatial unit have different expected values. Thus, no particularly intuitive result is drawn from their estimates.

[INSERT FIGURE 2 HERE]

Figure 2 shows choropleth maps with the smoothed SMRs for rectal cancer in men for BYM, NFE, HGeo and ZGeo. Recall that this pattern was one of those needing some zero treatment. Both hurdle maps (NFE and HGeo) are similar as their modeling of the probabilities of zeroes is also similar, as mentioned in Section 4. They mainly modify the risks in those regions less populated and more prone to zeroes (upper-left side of

the maps) decreasing their risks in order to get those extra zeroes needed. As seen in Section 3 this differential performance of the low populated areas could not be achieved with the naive models introduced in Section 2. In contrast, regions having high SMRs hardly show any change. Thus hurdle models mainly modify the left tail of the distribution of the SMRs in order to fit the zero excess, but leaves the right tail of the distribution mostly unchanged. ZGeo introduces more differences with regard to BYM in both tails of the distribution. New regions with both high and low risks have emerged in this map. Several regions of very low risk have emerged in the upper-left side of the map. This result of ZGeo is very common and can also be seen in many of the diseases studied (see Annex B for seeing all 4 maps for the whole set of diseases).

Although the choice of a particular model for treating zero excesses is not a goal of this work, we would recommend to use NFE as benchmark proposal between all those introduced in this paper. We have found particularly satisfactory that NFE shows a better performance in terms of DIC than the rest of models and the estimates of its  $\gamma$  parameter seems very coherent. Moreover, this model seems to yield conservative results in that the change in their geographic patterns compared to BYM is milder than that for the rest of models, yet enough to correct the original zero excesses in the BYM models. Finally, the logit formulation of NFE makes it particularly well suited for further modelling  $\pi^H$  if needed in contrast to the Geometric proposals. Thus we overall find NFE a convenient proposal for modeling data sets showing zero excesses.

## 6 Conclusions

Disease mapping models with zero-specific treatment can be considered as enhanced disease mapping models controlling overdispersion in the observed counts inducing also dependence on the underlying risks. Nevertheless, overdispersion fitting procedures in general may not be enough for solving zero excesses problems, which are a unique kind of overdispersion. Thus, specific models are needed to deal with this problem. As shown in this paper zero excesses are present in certain data sets concerning mortality data, at least for the Valencian Mortality Dataset. A relevant proportion of the diseases studied have been found to show zero excesses, even after accounting for overdispersion with disease mapping models. Thus, as evidenced, zero excesses require attention for mortality geographic studies in general.

The Valencian Mortality Dataset is somewhat particular in some senses due to the high demographic variability of the units of study. This could have made naive ZIP and hurdle models (without modeling of the probabilities of the zero-specific component) seem particularly bad as zero-specific components also put substantial probability to zero counts in large cities. As a consequence the zero-specific components are discarded. Nevertheless, we expect substantial differences in the expected cases for regular disease mapping studies since otherwise those expected values would not be omnipresent in so many studies. In any case the Valencian Mortality Dataset is comprehensive enough and representative of real mortality



data so that the need of zero-treatment evidenced in this dataset could be a signal of a general fact in mortality data sets of other regions.

Maybe one reason why mortality data may show frequent zero excesses when smoothing the SMRs is inherent to the smoothing process. Smoothing procedures usually combine information on the observed data and the prior structure defined by the model. When that observed information is low (small units) the shrinkage towards the prior structure is stronger. As a consequence the risks in the smaller units may be easily oversmoothed towards the mean, or a local mean, yielding conservative risk estimates. Models treating zero excesses with a different probability of the zero-specific component solve this problem by decreasing the risk in the smaller units (those which are more likely to show zeroes) and therefore increasing the number of zeroes predicted. Nevertheless, a similar oversmoothing could exist in small units showing high risks. In that case their SMRs should be higher but they are oversmoothed towards the mean because of the small information in each of them. Proposals modeling zero excesses in no way would fix this issue which only alleviates the oversmoothing of small units showing low risks.

In our opinion the theoretical results in Subsection 4.1 are also of high importance from an applied point of view. They show that proposals leading to wrong (improper) results have been frequently proposed in the literature. These problems can have different consequences such as plain improper posterior distributions or, if arbitrary vague prior distributions are used, arbitrary posterior distributions which are extremely sensitive (possibly unnoticedly) to prior parameters. These problems are often interpreted in the literature as simple MCMC convergence problems. Although this may seem obvious, we would advice modelers to pay further attention to those convergence problems. In our experience those problems have been an excellent guidance for formulating the theoretical results in Subsection 4.1 since they clearly warn that something suspicious could be happening. In our opinion this is an additional advantage of MCMC inference since convergence problems can be treated, at least in this context, as a trace of problems in model formulations instead of simple drawbacks inherent to MCMC as an inferential tool.

The main purpose of this paper has not been to propose a particularly suitable model for dealing with zero excesses. Besides showing the high prevalence of zero excesses problems in regular mortality data, which would deserve further epidemiological research, the purpose of this paper is double. On one hand, we pretend to show some theoretical pointing out wrong procedures in this area. In our opinion this is quite important in order to avoid works proposing flawed models. The main value of this side of the paper is warning modelers on what procedures not to do instead of setting what to do with zero excesses. Our results are just sufficient, not necessary, conditions for posterior impropriety in these models. On the other hand, this paper illustrates several ‘valid’ proposals for modeling zero excesses, i.e. we wanted to illustrate suitable proposals for handling this particular issue that were admissible in light of the results shown in Section 4. It would be desirable to have a proof of the posterior propriety of these proposals, or even better necessary conditions for the posterior impropriety of zero-specific models in general. Regretfully we do not have

that proof but anyway the value of the results proved still remain since they guide us on what procedures not to follow which is a valuable guidance according to many models already proposed in the literature.

A more thorough comparison of the models in Subsection 4 and possibly some further models would be greatly advisable although that comparison is beyond the scope of this paper. We have found more interesting to illustrate several modeling proposals instead of exposing just one of them. Nevertheless, the results of this paper suggests the need for further research but also a high dose of caution when making modeling proposals in this area.

Finally, we would like to point out that according to Natarajan and McCulloch Natarajan and McCulloch (1995) the conditions stated there for posterior impropriety in the modeling of binary data are similar to those formulated in Albert and Anderson (1984) Albert and Anderson (1984) for non-existence of MLE in logit frequentist modeling. Indeed, the conditions for posterior impropriety in the Bayesian approach are more restrictive than those for non-existence of the MLEs in the frequentist context. The conditions set at Natarajan and McCulloch Natarajan and McCulloch (1995) have been those also set as conditions for posterior impropriety for the modeling of the probability of the zero-specific component with random effects in our work. Thus, the frequentist formulation of ZIP and hurdle models from a frequentist setting could be in principle as problematic as that same formulation from a Bayesian point of view. The conditions under which frequentist ZIP and hurdle models yield valid (or unvalid) MLEs should be further explored but the results of this paper and the work of Albert and Anderson Albert and Anderson (1984) shed some doubts on those formulations from a frequentist point of view.

## Acknowledgments

The authors acknowledge the support of the research grants MTM2016-77501-P from the Spanish Ministry of Economy and Competitiveness, BA15/00003 of Instituto de Salud Carlos III and predoctoral contract UGP-15-156 of FISABIO.

## References

- Agarwal, D. K., Gelfand, A. E., and Citron-Pousty, S. (2002). Zero-inflated models with application to spatial count data. *Environmental and ecological statistics*, 9:341–355.
- Albert, A. and Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*, 71(1):1–10.
- Arab, A. (2015). Spatial and spatio-temporal models for modeling epidemiological data with excess zeros. *International Journal of Environmental Research and Public Health*, 12(9):10536–10548.
- Bayarri, M. J., Berger, J., and Datta, G. S. (2008). Objective Bayes testing of Poisson versus inflated Poisson models. In *Pushing the Limits*

- of *Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, volume 3, pages 105–121. Institute of Mathematical Statistics.
- Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, 1:385–402.
- Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43:1–21.
- Dalrymple, M. L., Hudson, I. L., and Ford, R. P. K. (2003). Finite mixture, zero-inflated Poisson and hurdle models with application to SIDS. *Computational Statistics & Data Analysis*, 41(3-4):491–504.
- Gschlößl, S. and Czado, C. (2008). Modelling count data with overdispersion and spatial effects. *Statistical Papers*, 49(3):531–552.
- Heilbron, D. C. (1994). Zero-altered and other regression models for count data with added zeros. *Biometrical Journal*, 36:531–547.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to detects on manufacturing. *Technometrics*, 34:1–14.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33:341–365.
- Musenge, E., Freeman Chirwa, T., Kahn, K., and Vounatsou, P. (2013). Bayesian analysis of zero inflated spatiotemporal HIV/TB child mortality data through the INLA and SPDE approaches: Applied to data observed between 1992 and 2010 in rural North East South Africa. *International Journal of Applied Earth Observation and Geoinformation*, 22:86–98.
- Natarajan, R. and McCulloch, C. E. (1995). A note on the existence of the posterior distribution for a class of mixed models for binomial responses. *Biometrika*, 82(3):639–643.
- Neelon, B., Chang, H. H., Ling, Q., and Hastings, N. S. (2014). Spatiotemporal hurdle models for zero-inflated count data: Exploring trends in emergency department visits. *Statistical Methods in Medical Research*.
- Neelon, B., Ghosh, P., and Loebis, P. F. (2013). A spatial Poisson hurdle model for exploring geographic variation in emergency department visits. *Journal of the Royal Statistical Society, Series A*, 176(2):389–413.
- Neelon, B. H., O’Malley, A. J., and Normand, S.-L. T. (2010). A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use. *Stat Modelling*, 10(4):421–439.
- Nieto-Barajas, L. and Bandyopadhyay, D. (2013). A zero-inflated spatial gamma process model with application to disease mapping. *Journal of Agricultural, Biological, and Environmental Statistics*, 18(2):137–158.

- Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.
- Song, H.-R., Lawson, A., D’Agostino Jr, R. B., and Liese, A. D. (2011). Modeling type 1 and type 2 diabetes mellitus incidence in youth: an application of Bayesian hierarchical regression for sparse small area data. *Spatial and Spatiotemporal Epidemiology*, 2(1):23–33.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64:583–641.
- Ugarte, M. D., Ibañez, B., and Militino, A. F. (2004). Testing for Poisson zero inflation in disease mapping. *Biometrical Journal*, 46:526–539.
- Upfill-Brown, A. M., Lyons, H. M., Pate, M. A., Shuaib, F., Baig, S., Hu, H., Eckhoff, P. A., and Chabot-Couture, G. (2014). Predictive spatial risk model of poliovirus to aid prioritization and hasten eradication in Nigeria. *BMC Medicine*, 12(92).
- Van Der Broek, J. (1995). A score test for zero inflation in a Poisson distribution. *Biometrics*, 51(2):738–743.
- Zurriaga, O., Martínez-Beneito, M. A., Botella-Rocamora, P., López-Quílez, A., Melchor, I., Amador, A., Vanaclocha, H., and Nolasco, A. (2010). Spatio-temporal mortality atlas of Comunitat Valenciana. (<http://www.geeitema.org/AtlasET/index.jsp?idioma=I>. Accessed: May, 2nd 2016).

Table 1: Observed zeroes for each data set and posterior predicted zeroes for each model and for the first 10 mortality causes. Values in the *Obs. zeroes* column correspond to the real observed zeroes for each data set. For the next 3 columns, numbers correspond to the posterior predictive medians for this same quantity for each model run and the corresponding unilateral 95% posterior predictive intervals. Bold fonts denote those combinations of models and data sets evidencing zero excesses according to their predictive intervals.

Sex & Cause	Obs. zeroes	BYM	ZIP	Hurdle
(Men, All tumours)	4	2 [0,5]	3 [0,5]	5 [0,11]
(Women, All tumours)	7	6 [0,10]	6 [0,10]	8 [0,15]
(Men, Mouth)	216	<b>196 [0,211]</b>	<b>199 [0,215]</b>	216 [0,242]
(Men, Stomach)	105	<b>91 [0,103]</b>	<b>92 [0,104]</b>	105 [0,127]
(Women, Stomach)	150	137 [0,151]	138 [0,152]	150 [0,173]
(Men, Colorectal)	73	<b>58 [0,68]</b>	<b>59 [0,69]</b>	74 [0,93]
(Women, Colorectal)	74	72 [0,82]	73 [0,83]	74 [0,93]
(Men, Colon)	96	<b>79 [0,91]</b>	84 [0,96]	96 [0,119]
(Women, Colon)	98	91 [0,102]	92 [0,104]	99 [0,119]
(Men, Rectum)	201	<b>180 [0,196]</b>	<b>183 [0,199]</b>	202 [0,228]
...	...	...	...	...