

## RESEARCH ARTICLE

# Measures of clustering and heterogeneity in multilevel Poisson regression analyses of rates/count data

Peter C. Austin<sup>1,2,3</sup>  | Henrik Stryhn<sup>4</sup> | George Leckie<sup>5</sup> | Juan Merlo<sup>6,7</sup>

<sup>1</sup>Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada

<sup>2</sup>Institute of Health Management, Policy and Evaluation, University of Toronto, Toronto, Ontario, Canada

<sup>3</sup>Schulich Heart Research Program, Sunnybrook Research Institute, Toronto, Ontario, Canada

<sup>4</sup>Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada

<sup>5</sup>Centre for Multilevel Modelling, University of Bristol, Bristol, UK

<sup>6</sup>Unit for Social Epidemiology, Faculty of Medicine, Lund University, Malmö, Sweden

<sup>7</sup>Center for Primary Health Care Research, Region Skåne, Malmö, Sweden

## Correspondence

Peter Austin, Institute for Clinical Evaluative Sciences, G106, 2075 Bayview Avenue, Toronto, Ontario, M4N 3M5 Canada.  
Email: peter.austin@ices.on.ca

## Funding information

Canadian Institutes of Health Research, Grant/Award Number: CRT43823, CTP 79847 and MOP 86508; Heart and Stroke Foundation of Canada, Grant/Award Number: Career Investigator Award; Swedish Research Council, Grant/Award Number: VR #2013-2484; Canadian Institutes of Health Research (CIHR), Grant/Award Number: MOP 86508

Multilevel data occur frequently in many research areas like health services research and epidemiology. A suitable way to analyze such data is through the use of multilevel regression models. These models incorporate cluster-specific random effects that allow one to partition the total variation in the outcome into between-cluster variation and between-individual variation. The magnitude of the effect of clustering provides a measure of the general contextual effect. When outcomes are binary or time-to-event in nature, the general contextual effect can be quantified by measures of heterogeneity like the median odds ratio or the median hazard ratio, respectively, which can be calculated from a multilevel regression model. Outcomes that are integer counts denoting the number of times that an event occurred are common in epidemiological and medical research. The median (incidence) rate ratio in multilevel Poisson regression for counts that corresponds to the median odds ratio or median hazard ratio for binary or time-to-event outcomes respectively is relatively unknown and is rarely used. The median rate ratio is the median relative change in the rate of the occurrence of the event when comparing identical subjects from 2 randomly selected different clusters that are ordered by rate. We also describe how the variance partition coefficient, which denotes the proportion of the variation in the outcome that is attributable to between-cluster differences, can be computed with count outcomes. We illustrate the application and interpretation of these measures in a case study analyzing the rate of hospital readmission in patients discharged from hospital with a diagnosis of heart failure.

## KEYWORDS

median incidence rate ratio, median rate ratio, multilevel analysis, Poisson regression, variance partition coefficient

## 1 | INTRODUCTION

Data with a multilevel nature occur frequently in public health, health services research, behavioral research, and in epidemiology. Examples include patients nested within hospitals, residents clustered within geographic regions, and

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2017 The Authors. *Statistics in Medicine* published by John Wiley & Sons Ltd.

employees nested within companies. The existence of multilevel information is relevant in the practice of epidemiology for both formal statistical and substantive epidemiological reasons. Analysts are increasingly aware that multilevel regression models are a suitable way to analyze clustered data.<sup>1,2</sup> A consequence of clustering of the subjects within clusters or higher-level units (eg, hospitals) is that subjects from the same cluster may have outcomes that are more similar than will subjects from different clusters. Multilevel regression models incorporate cluster-specific random effects that account for the dependency of the data by partitioning the total individual variance into variation due to the cluster and the individual-level variation that remains.<sup>3</sup>

From an epidemiological perspective, knowing the share of the total individual variation in the outcome that is attributable to the cluster level or the “variance partition coefficient” (VPC) is useful information. The higher the VPC, the more relevant the context appears to be for understanding individual health or disease outcomes.<sup>3,4</sup> The VPC informs on the existence of a *general contextual effect* (GCE), which is called “general” because it reflects the influence of the cluster context as a whole, without specifying any contextual characteristic other than the very boundaries that delimit the cluster.<sup>5</sup>

The VPC is easy to calculate and interpret in multilevel linear regression models with continuous outcomes.<sup>6</sup> In the case of discrete responses, the calculation and interpretation of the VPC is more complicated since, amongst other issues, the individual and cluster components of variance are modeled on different scales (the discrete response scale and the linear predictor scale, respectively). Furthermore, the components of variance depend on the covariates, so there is typically no unique VPC for models with discrete outcomes. Several alternative approaches for computing the VPC have been proposed when analyzing discrete outcomes. These include using a normal response approximation, the simulation method, Taylor series linearization, and the latent response method.<sup>3,7</sup>

Another useful characteristic for epidemiological interpretation is the intraclass correlation (coefficient) (ICC), which is the correlation between 2 subjects within the same cluster. For nonnormal models the ICC also depends on the covariate values of the 2 subjects considered, but when these are the same, the ICC in a 2-level model coincides with the VPC.

To avoid the interpretative technicalities of the VPC, Larsen et al introduced the concept of the median odds ratio (MOR) that can be used in the interpretation of the GCE when fitting a random effects logistic regression model.<sup>8,9</sup> The MOR indicates the median value of the odds ratios obtained when comparing the odds of the occurrence of the outcome in an individual from a randomly selected cluster with another individual with identical covariate values but randomly selected from a different cluster when the clusters are ordered by risk. In other words, to calculate the MOR, we should first measure the odds of the occurrence of the outcome for all randomly taken pairs of individuals from different clusters and, thereafter, compute the odds ratio for each pair of individuals having the individual from the cluster with the higher odds in the numerator and the individual from the cluster with the lower odds in the denominator. This would produce a distribution of odds ratios that are always equal to or higher than 1. The MOR is the median of this distribution of odds ratios.<sup>10</sup> An advantage to the MOR is that it permits the analyst to present the between-cluster variation as a measure of association (ie, an odds ratio) and thereby allows the comparison of the GCE with the fixed effects of the covariates in the model. More recently, the concept of the MOR has been extended to the analysis of survival or time-to-event outcomes, with the median hazard ratio (MHR) serving as a measure of the GCE.<sup>11-13</sup> The MHR has an interpretation similar to the MOR, with the exception that the hazard ratio replaces the odds ratio.

While measures of variance and the GCE have been well described for settings in which the outcome is continuous, binary, or time-to-event in nature, there is a paucity of information as to how to assess these quantities when outcomes are counts. This is despite the fact that count outcomes are common in epidemiological and medical research. Common examples include the number of hospitalizations or physician consultations during an observation period, the number of infections or injuries, and the number of days in hospital.

Having an explicitly expository focus, the objective to the current paper was two-fold. First, to describe the median rate ratio (MRR) to researchers in epidemiology and biostatistics. Second, to demonstrate how the VPC can be calculated for count outcomes. We will illustrate the utility of these 2 measures for assessing the magnitude of the general impact of the hospital context (ie, GCEs) on the number of hospital readmissions in patients discharged from hospital with a diagnosis of congestive heart failure (CHF). The paper is structured as follows. In Section 2, we describe methods for evaluating the VPC and for estimating the MRR when using a multilevel Poisson regression model for count outcomes. In Section 3, we provide a case study in which we illustrate the utility of these metrics for assessing the magnitude of the general contextual (ie, hospital) effects when analyzing hospital readmissions data. Finally, in Section 4, we summarize our findings.

## 2 | MEASURES OF VARIANCE AND OF THE GCE IN MULTILEVEL POISSON REGRESSION

In this section, we describe the multilevel Poisson regression model. We then describe 2 different methods by which the VPC can be estimated for count outcomes when fitting multilevel Poisson regression models. We also describe the concept of the median (incident) rate ratio (MRR) and describe a formula for its evaluation.

### 2.1 | The multilevel Poisson regression model

When the outcome is a count denoting the number of times that an event occurred (eg, number of hospital admissions or physicians consultations), a Poisson regression model can be used to relate the mean number of events to a set of explanatory variables using a logarithmic link function. The conventional Poisson regression model can be formulated as

$$\begin{aligned}\log(\lambda_i) &= \boldsymbol{\beta}\mathbf{X}_i \\ Y_i &\sim \text{Poisson}(\lambda_i),\end{aligned}\quad (1)$$

where  $\mathbf{X}_i$  denotes a  $p \times 1$  column matrix of covariates measured on the  $i$ th subject,  $Y_i$  denotes the count outcome measured on the  $i$ th subject, and  $\boldsymbol{\beta}$  denotes a  $1 \times p$  row matrix of regression coefficients. The parameter  $\lambda_i$  denotes the expected or mean number of events for the  $i$ th subject given their set of observed covariates. The model can be modified to include an *offset* term defined as the natural logarithm of the *exposure*, where the exposure could denote the time (ie, number of days) during which each subject is observed and therefore at risk for the occurrence of events (see later empirical example). The modification to include an offset term is

$$\begin{aligned}\log(\lambda_i) &= \boldsymbol{\beta}\mathbf{X}_i + \log(T_i) \\ Y_i &\sim \text{Poisson}(\lambda_i),\end{aligned}\quad (2)$$

where  $T_i$  denotes the exposure. Below, we shall often incorporate the offset into  $\mathbf{X}_i$ , since it may simply be viewed as an additional covariate whose regression coefficient is constrained to be equal to one.

The above formulation of the Poisson regression model assumes that the analyst is using individual subject-level data. In other words, there is one record for each subject. For each record, we have the vector of covariates and the count outcome for the subject associated with that record. In many settings, the Poisson regression model is used with aggregated data. For example, the responses may be neighborhood-level disease counts, in which case, the exposure would be the population at risk in each neighborhood. Alternatively, a given record may denote the aggregate data for all subjects with a given covariate pattern (we refer to all subjects with a given covariate pattern as a stratum). For instance, the record may contain the number of events observed amongst all females aged 55–64 years. In this case, the offset variable could be the number of subjects in the given stratum or the amount of person-time contributed by subjects in the given stratum.

The exponentiated regression coefficients arising from the fitted model are interpreted as rate ratios. They denote the relative change in the mean number of events that would be expected to occur for a one-unit increase in the relevant covariate. When an offset variable is included to denote the size of the population at risk, then the exponentiated regression coefficients denote the relative change in the rate of the events (ie, number of events per standardized population size). Frequently, when the outcome is an event that occurs at most once for each subject (eg, death or disease incidence) and an offset is included in the model, the exponentiated regression coefficient is referred to as the incidence rate ratio.

When subjects are nested or clustered in higher-level units, the conventional Poisson regression model can be modified through the inclusion of cluster-specific random effects to account for within-cluster correlation in the outcome. The simplest modification is that of the random intercept regression model in which the intercept is allowed to vary randomly across clusters:

$$\begin{aligned}\log(\lambda_{ij}) &= \boldsymbol{\beta}\mathbf{X}_{ij} + \alpha_j \\ Y_{ij} &\sim \text{Poisson}(\lambda_{ij}),\end{aligned}\quad (3)$$

where  $\alpha_j \sim N(0, \sigma^2)$  for the  $K$  clusters,  $j = 1, \dots, K$ . The model with an offset variable can be modified similarly. Throughout the manuscript, we consider the values of the predictors and offsets to be fixed. Note that the inclusion of random effects

means that the marginal distribution of  $Y_{ij}$  (averaged over  $\alpha_j$  but given  $\mathbf{X}_{ij}$ ) is not a Poisson distribution. Equation 3 should be interpreted to mean that  $Y_{ij}$  follows a Poisson distribution conditional on the covariates and the cluster-specific random effect (this is a similar interpretation as of a logistic model with random effects). The level 1 variance is the within-cluster variance, thus conditional on the cluster random effects in the model.

## 2.2 | VPC and ICC

The VPC denotes the proportion of the (unexplained) variation in the outcome that is due to between-cluster variation. In other words, the VPC denotes the proportion of variation that is beyond that explained by the fixed predictors that is due to between-cluster variation. As such, the VPC ranges from 0 to 1. A VPC of zero indicates that all of the variation in the outcome is due to within-cluster-between-subject variability. In this case, the correlation in outcomes between 2 randomly selected subjects from the same cluster will be zero and therefore, equal to the correlation in outcomes between 2 randomly selected subjects from different clusters. At the other extreme, a VPC of one indicates that all of the variation in the outcome is due to between-cluster variability and that all subjects within the same cluster display the same outcome. In this section, we describe 2 different methods for computing the VPC: exact calculation and a simulation-based approach. When outcomes are binary or ordinal, an approach based on assuming an underlying latent continuous response can be used.<sup>1,3</sup> However, there is no latent response formulation for Poisson models, and so this approach is not a possibility here. Goldstein et al explored a model linearization approach.<sup>3</sup> However, Stryhn et al found that this method performed poorly with count outcomes.<sup>14</sup> Accordingly, we do not consider the count version of this approach in the current study.

### 2.2.1 | Exact calculation

Given the random effects Poisson regression model described in Equation 2 or 3, Stryhn et al, in the proceedings of a symposium on veterinary epidemiology and economics, provided an exact calculation for the VPC and the ICC<sup>14</sup>:

$$\text{VPC} = \text{ICC} = \frac{\exp(2\beta\mathbf{X} + 2\sigma^2) - \exp(2\beta\mathbf{X} + \sigma^2)}{\exp(2\beta\mathbf{X} + 2\sigma^2) - \exp(2\beta\mathbf{X} + \sigma^2) + \exp(\beta\mathbf{X} + \sigma^2/2)}. \quad (4)$$

As the proof of this calculation has not been presented in the peer-reviewed literature, we include details of its derivation in Appendix A of this paper. Note that this quantity is simply the ratio between the cluster-level variance and the total variance, as in a linear multilevel model. Using this exact calculation, the VPC can be expressed directly as a function of the linear predictor and the variance of the random effects. Note that in this expression, the offset variable, if present, has been incorporated into the linear predictor. Note that in contrast to the VPC in 2-level random-intercept models for continuous responses, the VPC is a function of not only the variance of the random effects but also the linear predictor. The VPC will therefore take different values for each covariate pattern. In the following, we will, for simplicity, refer to the VPC only, although all statements related to the above calculations are equally valid for the ICC (under the assumption that the 2 individuals being randomly compared have the same covariate and offset values).

### 2.2.2 | Simulation-based estimation

Goldstein et al described a simulation-based approach to estimating the VPC when using a multilevel logistic regression model.<sup>3</sup> This algorithm can be easily modified for use with Poisson regression. To do so, one first fits the multilevel Poisson model (Equation 3 above) and estimates the distribution of the cluster-specific random effects. The algorithm then proceeds as follows:

1. Simulate a large number of cluster-level random effects from the random effects distribution that was obtained from fitting the multilevel Poisson regression model:  $\alpha_k \sim N(0, \hat{\sigma}^2)$ , for  $k = 1, \dots, M$ .
2. For a specific covariate pattern (and offset if present), use each of the simulated random effects drawn in step 1 to compute the predicted means using the regression coefficients from the fitted multilevel Poisson regression model:

$\hat{Y}_k = \exp(\hat{\beta}\mathbf{X} + \alpha_k)$ . For each of these computed count outcomes, compute the level 1 variance:  $v_k = \hat{Y}_k$  (by the equality of the mean and variance in the Poisson distribution).

3. The VPC is then evaluated as 
$$\frac{\text{Var}(\hat{Y}_k)}{\text{Var}(\hat{Y}_k) + \frac{1}{M} \sum_{k=1}^M v_k}$$
.

In the case study below, we used the above algorithm with  $M = 100\,000$  iterations. Note that the above algorithm provides an estimate of the VPC for a specific covariate pattern and size of the offset variable. As previously discussed, the above algorithm equally provides an estimate of the ICC.

### 2.3 | The median (incidence) rate ratio

Given the random effects Poisson regression model described in Equation 3, Rabe-Hesketh and Skrondal state that the median incidence rate ratio can be evaluated as<sup>15</sup>

$$\text{IRR}_{\text{median}} = \exp\left(\sqrt{2\sigma^2}\Phi^{-1}(0.75)\right), \quad (5)$$

where  $\Phi^{-1}$  denotes the inverse of the standard normal cumulative distribution function. This statistic measures the median relative change in the rate of the occurrence of the event when comparing identical subjects from 2 randomly selected different clusters that are ordered by rate. While Rabe-Hesketh and Skrondal use the term median incidence rate ratio, we will use the more general term median rate ratio (MRR). The primary reason for this is that an investigator may not be interested in a single incident event (eg, death or incidence of a specific disease). Instead, the investigator may be interested in the number of events that occurred when multiple events can occur per subject (eg, number of hospitalizations or episodes of unemployment). However, identical methods can be used regardless of whether the nature of the outcome is a single nonrecurring event or whether it is an event that can occur multiple times. As the definition of the median incidence rate ratio is a function of the variance of the random effects distribution, the same formula can be used regardless of whether one is examining incident events (eg, deaths or disease incidence) or the count of the number of events that an individual has experienced (eg, number of hospitalizations).

The corresponding measure of heterogeneity when outcomes are binary is the MOR,<sup>8-10</sup> while the MHR can be used with survival outcomes.<sup>11-13</sup> When using a logistic regression model with normally distributed random effects or a Cox model with normally distributed random effects, the MOR and the MHR can be evaluated using Equation 5. In other words, for the MRR, the MOR, and for the MHR when using a Cox model with normal random effects, the corresponding measure of heterogeneity is the same function of the variance of the random effects. This arises from the fact that when the random effects follow a normal distribution, then the distribution of  $|\alpha_j - \alpha_j|$  follows a half-normal distribution with variance equal to  $2\sigma^2$ . The median of this half-normal distribution is given by  $\sqrt{2\sigma^2}\Phi^{-1}(0.75)$ . As the measures of association for all 3 models are produced by exponentiating the regression coefficients and all 3 models incorporate normally distributed random effects, the formulas for estimating the MOR, the MHR, and the MRR are identical.

## 3 | CASE STUDY

We provide a case study to illustrate application of the VPC and the MRR for evaluating the hospital GCE on the number of hospital readmissions subsequent to discharge from hospital with a diagnosis of CHF.

### 3.1 | Data

The case study used patients from the Enhanced Feedback for Effective Cardiac Treatment (EFFECT) study, which was an initiative to improve the quality of care for patients with cardiovascular disease in Ontario.<sup>16</sup> The current study included 7162 patients discharged alive from hospital with a diagnosis of CHF between April 2004 and March 2005. These data have a multilevel structure, with patients nested within hospitals. The study sample consisted of 7162 patients treated at 96 hospitals. The number of patients per hospital ranged from 2 to 258, with a median of 74 (25th and 75th percentiles: 33 and 108, respectively). Due to the study inclusion and exclusion criteria, no patient had more than one

hospital discharge during the 1-year time frame of the study. We use the term index hospitalization to refer to the hospital admission that resulted in the patient being included in the EFFECT study.

The sample of patients discharged alive was linked to the Canadian Institute for Health Information Discharge Abstract Database (CIHI-DAD) using unique encoded patient identifiers. The CIHI-DAD contains records of all in-patient hospitalizations in the province of Ontario. The CIHI-DAD was used to determine the number of hospital admissions in the year subsequent to discharge from the index EFFECT hospitalization. The sample was also linked to the Registered Patients Database (RPDB), which contains basic demographic data on all residents of Ontario. In particular, it contains the date of death of all residents who have died since 1992. Using the RPDB, we were able to identify those subjects who died within 1 year of hospital discharge.

For each patient, the duration of time for which the patient was at risk of hospital readmission was the equal to the number of days between discharge from the index EFFECT hospitalization and the date of death. For those subjects who did not die during the first year subsequent to discharge, the duration of at-risk time was set equal to 365 (26.4% of patients had less than 365 days of at-risk time).

### 3.2 | Study outcome

The outcome for the case study was the number of hospital admissions in the 1 year after discharge from the index EFFECT hospitalization. The median number of hospital readmissions was 1, while the 75th percentile was 2, and the maximum number of readmissions was 14.

### 3.3 | Predictor variables

The EFFECT-HF mortality prediction models estimate the probability of death within 30 days and 1 year of hospitalization for CHF.<sup>17</sup> The model for predicting 1-year mortality uses 11 variables: age, systolic blood pressure on admission, respiratory rate on admission, low sodium serum concentration (<136 mEq/L), low serum hemoglobin (<10.0 g/dL), serum urea nitrogen, presence of cerebrovascular disease, presence of dementia, chronic obstructive pulmonary disease (COPD), hepatic cirrhosis, and cancer. While this model was developed for predicting mortality in patients hospitalized with heart failure, we use it in this case study for illustrative purposes, acknowledging that it was not developed for the purposes of modeling the number of hospital readmissions. In the current case study, each of the 4 continuous covariates (age, systolic blood pressure, respiratory rate, and serum urea nitrogen) were standardized to have mean zero and unit variance (this was achieved by subtracting the sample mean from each subject's value of that covariate and then dividing the resultant quantity by the standard deviation of that covariate).

### 3.4 | Statistical analyses

We fit 2 different multilevel Poisson regression models. First, we fit the null model, which included only hospital-specific random effects:

$$\log(\lambda_{ij}) = \beta_0 + \alpha_j + \log(T_{ij}). \quad (6)$$

$T_{ij}$  denotes the exposure for the  $i$ th patient in the  $j$ th hospital and was set equal to the at-risk time defined above. This model did not contain any patient characteristics and so the purpose of this analysis is to quantify the variation in the outcome and the proportion of this variation that is due to between-hospital differences before accounting for any patient characteristics.

Second, we fit a multilevel Poisson regression model that comprised the 11 variables in the EFFECT-HF mortality prediction model and the hospital-specific random effects:

$$\log(\lambda_{ij}) = \beta \mathbf{X}_{ij} + \alpha_j + \log(T_{ij}), \quad (7)$$

where the vector  $\mathbf{X}_{ij}$  denotes the vector of patient-level characteristics (including an intercept term) and  $\beta$  denotes the vector of associated regression coefficients. In each of these 2 models, we assumed that the distribution of the random effects was normal:  $\alpha_j \sim N(0, \sigma^2)$ .

For each of the 2 models, we computed the VPC using both the exact calculation and Goldstein's simulation-based approach that were described in Section 2. When using the simulation-based approach, we used 100 000 simulations to estimate the VPC. We computed the VPC for typical covariate patterns and values of the offset variable. In our sample, the 10th and 90th percentiles of the distribution of the at-risk time were 74 and 365 days, respectively. For each selected covariate pattern, we evaluated the VPC at all values of the offset variable ranging from 74 to 365 days, in increments of 1 day. We also computed the MRR using the method described in Section 2. One could also determine the mean or median offset over all subjects in the sample and compute a sample-specific "typical" VPC at the value of the mean or median offset (in our study sample, the mean and median offsets were 304 and 365, respectively). A caveat to this approach is that the mean offset may not correspond to a meaningful value (in our case, referring to subjects who had 304 days of at-risk time subsequent to hospital discharge). We would suggest that examining the VPC across a range of offsets allows for a greater examination of the variation in the VPC across different values of the offset and allows for estimation of the VPC at different meaningful values (eg, those who remained alive for 1 year postdischarge).

Unless otherwise noted, all statistical analyses were conducted using R, version<sup>18</sup> 3.1.2. The multilevel Poisson regression model was fit using the `glmer` function in the `lme4` package (version 1.1-7). R and Stata code for conducting these analyses are provided in Appendices B and C.

### 3.5 | Results

The median age of patients in the study sample was 78 years. The distribution of each of the 11 covariates in the regression model is summarized in Table 1.

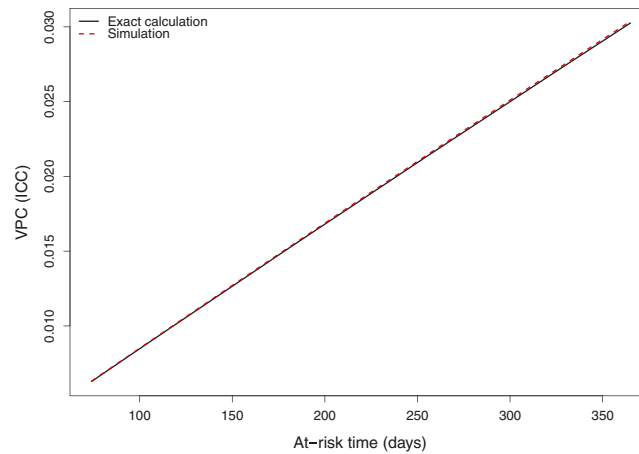
When fitting the null multilevel Poisson regression model, the estimated mean intercept was  $-5.60$ , while the variance of the hospital-specific random effects was estimated to be  $0.0226$ . The MRR was  $1.15$ . The linear predictor was equal to the mean intercept ( $-5.60$ ) plus the logarithm of the at-risk time. The values of the VPC obtained using exact calculations and the simulation-based approach are shown in Figure 1 for the different values of the at-risk time. The VPC ranged from  $0.006$  when the at-risk time was 74 day to  $0.030$  when the at-risk time was 365 days. There was very good agreement between the 2 approaches across the range of number of at-risk days. Amongst subjects who survived for 1-year postdischarge from the index EFFECT hospitalization (73.6% of patients), 3.0% of the variation in the number of readmissions was due to systematic differences between hospitals, while the remaining 97.0% was due to within-hospital between-patient differences.

Estimated rate ratios and associated 95% confidence intervals obtained from the multilevel Poisson regression model consisting of 11 patient characteristics are reported in Table 2. Seven of the 11 patient characteristics were significantly associated with the rate of hospitalization ( $P < .01$ ). Four variables (age, presence of cerebrovascular disease, presence of

**TABLE 1** Distribution of baseline covariates in the study sample

Variable	Median (IQR) or Percent With Condition
Continuous covariates	
Age (years)	70 (70-84)
Systolic blood pressure (mmHg)	145 (125-168)
Respiratory rate (breaths per minute)	24 (20-28)
Serum urea nitrogen (mmol/L)	8.2 (6.1-11.7)
Binary covariates	
Low sodium serum concentration (<136 mEq/L)	20.9%
Low serum hemoglobin (<10.0 g/dL)	12.8%
Cancer	11.3%
Chronic obstructive pulmonary disease (COPD)	27.1%
Cerebrovascular disease	17.9%
Hepatic cirrhosis	0.7%
Dementia	9.3%

Continuous covariates are summarized using the median (25th–75th percentiles), while binary covariates are summarized using the percentage of subjects in whom the condition is present. The sample consists of 7162 patients treated at 96 hospitals.



**FIGURE 1** Variance partition coefficient (VPC) (intra-class correlation coefficient [ICC]) for null multilevel Poisson model [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 2** Rate ratios and 95% confidence intervals for the multilevel Poisson regression model

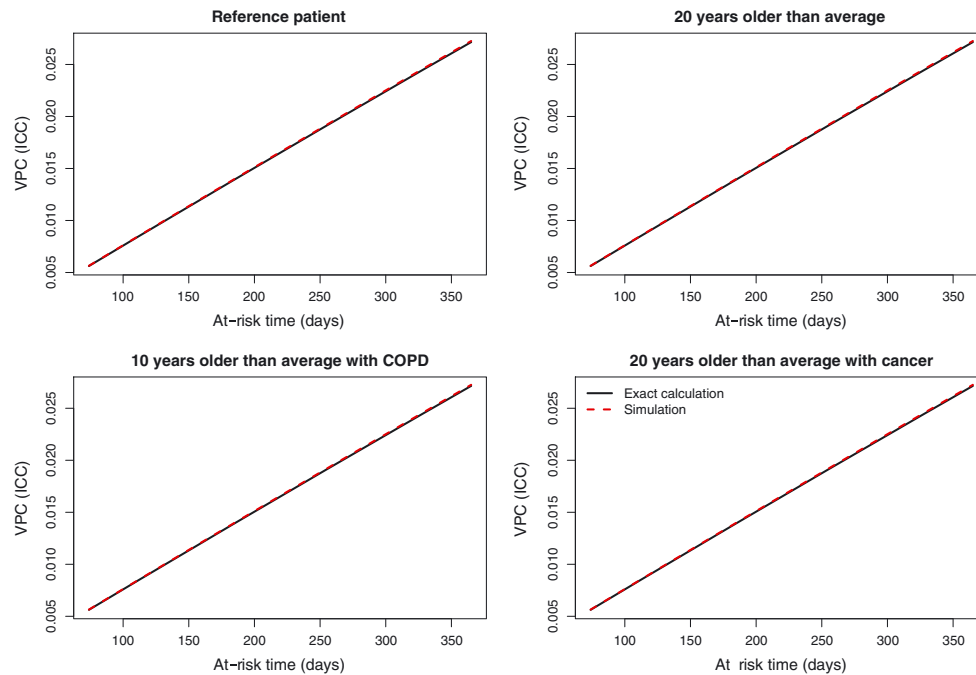
Variable	Rate Ratio	95% Confidence Interval	P Value
Fixed intercept			
	−5.69 (−5.74, −5.64)		
Fixed effects (rate ratios)			
Age	0.999	(0.976, 1.023)	$P = .9506$
Systolic blood pressure	0.912	(0.891, 0.933)	$P < .0001$
Respiratory rate	1.036	(1.013, 1.059)	$P = .0021$
Low sodium serum concentration (< 136 mEq/L)	1.098	(1.04, 1.159)	$P = .0008$
Low serum hemoglobin (< 10.0 g/dL)	1.151	(1.079, 1.227)	$P < .0001$
Serum urea nitrogen	1.167	(1.141, 1.194)	$P < .0001$
Cancer	1.239	(1.16, 1.324)	$P < .0001$
Chronic obstructive pulmonary disease (COPD)	1.109	(1.056, 1.165)	$P < .0001$
Cerebrovascular disease	1.021	(0.963, 1.082)	$P = .4923$
Hepatic cirrhosis	1.065	(0.827, 1.37)	$P = .6262$
Dementia	1.005	(0.924, 1.093)	$P = .9047$
Variance of random effect	0.0222		

The rate ratios for the 4 continuous variables denote the relative change in the rate of the outcome associated with a one-standard deviation change in the covariate. Age was standardized by centering at the mean age (76.4 years) and dividing by the standard deviation (11.6). For systolic blood pressure, the corresponding numbers were 147.0 and 29.2, respectively. For respiratory rate, the corresponding numbers were 24.9 and 6.1, respectively. For serum urea nitrogen, the corresponding numbers were 9.4 and 4.6, respectively. The sample consists of 7162 patients treated at 96 hospitals.

dementia, and hepatic cirrhosis) were not significantly associated with the rate of hospitalization ( $P > .49$ ). Note that the rate ratios for the 4 continuous covariates denote the relative change in the rate of hospital admission associated with a one standard deviation change in the covariate.

The relationship between VPC and the number of at-risk days are described in Figure 2 for 4 different patient profiles. We considered four different patient profiles: (i—top left panel) a patient whose 4 continuous covariates were set to the cohort average and whose 5 binary covariates were set to absent (or zero). We refer to this patient profile as a reference patient; (ii—top right panel) a patient who was 20 years older than average, but who was similar to the reference patient in all other aspects; (iii—bottom left panel) a patient who was 10 years older than average and who had COPD, but who was similar to the reference patient in all other aspects; (iv—bottom right panel) a patient who





**FIGURE 2** Variance partition coefficient (VPC) (intra-class correlation coefficient [ICC]) for multilevel Poisson model with patient characteristics [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

was 20 years older than average and who had cancer, but who was similar to the reference patient in all other aspects. For each of these four patient profiles, the simulation-based approach resulted in an estimate of the VPC that was very close to that provided by the exact calculation across the entire spectrum of at-risk days. When the number of at-risk days was equal to 365 (ie, amongst patients who survived for the entire year postdischarge), the VPC was approximately equal to 0.025 across the 4 patient profiles. Thus, amongst these patients, approximately 2.5% of the unexplained variation in the mean number of hospital readmissions was due to systematic differences between hospitals, while the remaining 97.5% was due to unmeasured patient characteristics.

To examine the dependency of the VPC on the linear predictor for a fixed exposure (at-risk time = 365 days), we computed the VPC using exact calculations for all 7157 unique covariate patterns observed in the study sample. The VPC ranged from 0.019 to 0.074. The median was 0.029, while the 25th and 75th percentiles were 0.025 and 0.034, respectively. Thus, across the observed range of patient profiles, we saw moderate variation in the VPC. However, half of the observed VPCs lay in a relatively narrow interval.

To examine further the dependency of the VPC on the linear predictor for a fixed exposure (at-risk time = 365 days), we computed the VPC using exact calculations for 80 000 hypothetical covariate patterns. For each of the four continuous covariates, we considered values that were 2 standard deviations above and below the mean, one standard deviation above and below the mean, and the mean value. For the 7 binary covariates, we considered both the presence and absence of the characteristic. This resulted in 80 000 unique covariate patterns ( $=5^4 \times 2^7$ ). Across the 80 000 covariate patterns, the VPC ranged from 0.015 to 0.085. The median was 0.037, while the 25th and 75th percentiles were 0.030 and 0.045, respectively. Thus, across a wide range of patient profiles, we saw moderate variation in the VPC. However, half of the estimated VPCs lay in a relatively narrow interval.

The MRR for the multilevel Poisson regression model with 11 patient characteristics was 1.153. Thus, if one were to repeatedly sample at random 2 subjects with the same covariates from different hospitals, then in half such comparisons, the rate ratio comparing the rate of hospital readmission between the subject at the hospital with the higher rate of hospitalizations and the subject at the hospital with the lower rate of hospitalizations (differences in rates are entirely quantified by the cluster-specific random effects) is at least 1.153.

One can better understand the magnitude of between-hospital variation in rates of readmission by comparing the MRR with the rate ratios for the patient-level characteristics. The MRR was 1.153 (the reciprocal of this MRR is  $1/1.153 = 0.867$ ). Nine of the 11 patient-level characteristics had a rate ratio that lay between 0.867 and 1.153. Two of

the patient characteristics (urea nitrogen and cancer) had rate ratios that exceeded 1.153 (1.167 and 1.239, respectively). Thus, the median effect of clustering on the rate of hospital readmission was greater than the effect of nine of the 11 patient characteristics. Note that this is dependent on how the rate ratio is defined for the four continuous covariates. In the current analyses, these are the effects for a one standard deviation change in the continuous predictor (but other changes could be considered).

For the multilevel Poisson model with patient characteristics, the estimated variance of the distribution of the random effects was 0.022. One can use a likelihood ratio test to compare the full approximate likelihood of the fitted multilevel Poisson regression model with that of a conventional Poisson model that does not include any random effects. Such a test is provided by the `mepoisson` function in Stata (version 13.1, College Station, TX). The chi-squared statistic associated with the likelihood ratio test was 72.77 (1 df), while the associated P value was less than .0001. Thus, even after adjustment for patient characteristics, we would reject the null hypothesis of no between-hospital variation in the rate of hospital readmission ( $P < .001$ ). This analysis was conducted using the `mepoisson` function in Stata as the likelihood ratio test is not formally provided in the `glmer` function of the `lme4` package of R (however, one could extract the deviance of the fitted model and manually compare it with the deviance of the conventional Poisson model without random effects).

## 4 | DISCUSSION

The objective of the current paper was to describe to researchers in epidemiology and medical research measures of variance and heterogeneity for use with multilevel Poisson regression models for count outcomes. The MRR allows one to quantify the magnitude of the GCEs (ie, the effect of clustering) on the rate of the occurrence of the outcome on the rate ratio scale. Furthermore, it also permits a comparison of the magnitude of this GCE with that of model covariates. The VPC permits the analyst to quantify the proportion of the variation in outcomes that is due to systematic differences between clusters.

In our case study, we used the MRR to quantify the magnitude of the effect of clustering within hospitals, that is, the “general contextual effect” on the rate of hospital admission subsequent to hospital discharge for CHF. We found that the MRR for a Poisson model that contained 11 patient characteristics was 1.15, indicating that, for 50% of possible pair-wise comparisons, the rate of readmission for a reference patient was no more than 15% greater when comparing a hospital with higher readmission rates to a hospital with lower readmission rates. Furthermore, the MRR, which measures the median effect of clustering on the rate ratio scale, was larger in magnitude than the rate ratios for nine of the 11 patient characteristics. This suggests that, compared to the effect of patient characteristics, clustering exerted only a moderate effect on the rate of hospital readmission. Reporting the MRR complements the reporting of the variance of the random effects distribution. The MRR provides a characterization of the magnitude of the effect of clustering that would not have been possible had we simply reported the variances of the random effects distributions.

The MOR was introduced by Larsen et al<sup>8</sup> in 2000. However, it began to be used more frequently only once it was introduced to the epidemiological literature<sup>9</sup> in 2005. The MHR was empirically applied in 2007 to examine geographic variation in ischemic heart disease mortality in Sweden.<sup>19</sup> Its formal derivation and interpretation was first published by Lanke in 2010 as a short appendix written in *Historical Methods*<sup>11,12</sup> to quantify the impact of the family-specific frailty in southern Sweden during the period<sup>11</sup> 1766 to 1895. However, it was only introduced to the biostatistical and biomedical researchers in 2017.<sup>13</sup> While the MRR was presented very briefly in a book published by Stata Press on multilevel analysis using Stata,<sup>15</sup> our hope is that the current publication will increase its use in the medical and epidemiological literature.

The VPC can be calculated simply when outcomes are continuous: it is just the ratio of the cluster-specific variance to the total individual variance (ie, the sum of the between and within cluster variance).<sup>1</sup> When outcomes are binary, a variety of methods are available.<sup>3</sup> Amongst these different methods, the latent variable approach appears to be the most commonly used. The latent variable approach assumes that underlying the observed binary outcome is an unobserved (latent) continuous version of that outcome.<sup>1</sup> However, no such latent variable approach exists for count outcomes. Instead, when analyzing count outcomes, one can use either the exact calculation provided by Stryhn et al or use a simulation-based approach based upon that described by Goldstein et al.<sup>3,14</sup> The VPC in each approach depends on the linear predictor. However, in Poisson regression, unlike with logistic regression, there is a potentially hidden variable in the linear predictor: the offset. Thus, while the use of the simulation-based approach for binary outcomes can result in a different value of the VPC for each covariate pattern, the use of the simulation (and exact) approach with count

outcomes can result in a different value of the VPC for each covariate pattern and each potential value of the offset variable (see Figures 1 and 2). In our case study, we found that for a given value of the at-risk time, the value of the VPC did not differ meaningfully across 4 patient profiles. However, within a given patient profile, the VPC varied meaningfully across the different durations of at-risk time. When the duration of at-risk time was low, the VPC was very low. However, when the duration of at-risk time was 365 (ie, in patients who survived the whole year subsequent to discharge from the index CHF hospitalization), the VPC was approximately 0.025, indicating that 2.5% of the variation in the number of hospital readmissions was due to systematic between-hospital differences. The formula for the exact calculation of the VPC shows that this observation was to be expected, as a larger offset always results in an increase in the VPC. The analyses in our case study included an offset variable due to the relatively large number of deaths during the first year posthospital discharge with a diagnosis of CHF (26.4% of patients). However, in many settings, the use of an offset variable would not be necessary. The strong effect of the duration of the at-risk period on the value of the VPC may limit its utility in some settings. Alternatively, the effect of the duration of the at-risk time on the VPC can provide meaningful information on hospital comparisons. For instance, the increasing VPC with increasing duration of at-risk time suggests that differences in the between-hospital number of hospital readmissions increases as the at-risk time increases. This suggests that differences in hospital outcomes are amplified as the duration of at-risk time increases.

Given the very close agreement between the simulation-based approach and the exact calculations, we would argue that, when using a multilevel Poisson regression model with random intercepts, there is no need to use the simulation-based approach and that the exact calculations be used.

We derived an exact calculation for the VPC (or ICC) in the case of the random intercept Poisson multilevel model. However, the methodology can be extended easily to the random slopes Poisson multilevel model, as outlined in Appendix A.

In conclusion, the MRR allows one to determine the median relative change in the rate of occurrence of the outcome between a subject in a cluster with a higher rate of the outcome and an identical subject in a cluster with a lower rate of the outcome. Such a measure permits for an intuitive description of the magnitude of the impact of clustering within hospitals or GCEs when analyzing clustered count data. The MRR can be used in conjunction with the VPC which quantifies the proportion of variation in the outcome that is due to systematic between-cluster differences. Together, the MRR and the VPC permit investigators to quantify variation and heterogeneity in count outcomes across clusters.

## ACKNOWLEDGMENTS

This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results, and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred. This research was supported in-part by an operating grant from the Canadian Institutes of Health Research (CIHR) (MOP 86508) and in-part by the Swedish Research Council (VR #2013-2484; PI: Juan Merlo). Dr Austin is supported in part by a Career Investigator award from the Heart and Stroke Foundation. This study was approved by the institutional review board at Sunnybrook Health Sciences Centre, Toronto, Canada. These datasets were linked using unique encoded identifiers and analyzed at the Institute for Clinical Evaluative Sciences (ICES). Parts of this material are based on data and/or information compiled and provided by CIHI. However, the analyses, conclusions, opinions, and statements expressed in the material are those of the author(s) and not necessarily those of CIHI. The Enhanced Feedback for Effective Cardiac Treatment (EFFECT) data used in the study was funded by a CIHR Team Grant in Cardiovascular Outcomes Research (Grant numbers CTP 79847 and CRT43823).

## ORCID

Peter C. Austin  <http://orcid.org/0000-0003-3337-233X>

## REFERENCES

1. Snijders T, Bosker R. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling (Second Edition)*. London: Sage Publications; 2012.
2. Goldstein H. *Multilevel Statistical Models*. West Sussex: John Wiley & Sons Ltd.; 2011.
3. Goldstein H, Browne W, Rasbash J. Partitioning variation in generalised linear multilevel models. *Underst Stat*. 2002;1:223-232.

4. Merlo J. Multilevel analytic approaches in social epidemiology: measures of health variation compared with traditional measures of association. *J Epidemiol Community Health*. 2003;57(8):550-552. <https://doi.org/10.1136/jech.57.8.550>
5. Merlo J, Ohlsson H, Lynch KF, Chaix B, Subramanian SV. Individual and collective bodies: using measures of variance and association in contextual epidemiology. *J Epidemiol Community Health*. 2009;63(12):1043-1048. <https://doi.org/10.1136/jech.2009.088310>
6. Merlo J, Chaix B, Yang M, Lynch J, Rastam L. A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. *J Epidemiol Community Health*. 2005;59(6):443-449. <https://doi.org/10.1136/jech.2004.023473>
7. Browne WJ, Subramanian SV, Jones K, Goldstein H. Variance partitioning in multilevel logistic models that exhibit overdispersion. *J Royal Stat Soc - Series A (Stat Soc)*. 2005;168(3):599-613.
8. Larsen K, Petersen JH, Budtz-Jorgensen E, Endahl L. Interpreting parameters in the logistic regression model with random effects. *Biometrics*. 2000;56(3):909-914. <https://doi.org/10.1111/j.0006-341X.2000.00909.x>
9. Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol*. 2005;161(1):81-88. <https://doi.org/10.1093/aje/kwi017>
10. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health*. 2006;60(4):290-297. <https://doi.org/10.1136/jech.2004.029454>
11. Bengtsson T, Dribe M. Quantifying the family frailty effect in infant and child mortality by using median hazard ratio (MHR). *Hist Methods*. 2010;43(1):15-27. <https://doi.org/10.1080/01615440903270299>
12. Lanke J. How to describe the impact of the family-specific frailty (appendix). *Hist Methods: J Quant Interdiscip Hist*. 2010;43(1):26-27.
13. Austin PC, Wagner P, Merlo J. The median hazard ratio: a useful measure of variance and general contextual effects in multilevel survival analysis. *Stat Med*. 2017;36(6):928-938. <https://doi.org/10.1002/sim.7188>
14. Stryhn H, Sanchez J, Morley P, Booker C, Dohoo IR. Interpretation of variance parameters in multilevel Poisson regression models. 11th International Symposium on Veterinary Epidemiology and Economics. Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics, 2006.
15. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata, Volume 2: Categorical Responses, Counts, and Survival*. Stata Press; 2012:1-939.
16. Tu JV, Donovan LR, Lee DS, et al. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. *JAMA*. 2009;302(21):2330-2337.
17. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting mortality among patients hospitalized for heart failure: derivation and validation of a clinical model. *JAMA*. 2003;290(19):2581-2587.
18. R Core Development Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2005.
19. Chaix B, Rosvall M, Merlo J. Assessment of the magnitude of geographical variations and socioeconomic contextual effects on ischaemic heart disease mortality: a multilevel survival analysis of a large Swedish cohort. *J Epidemiol Community Health*. 2007;61(4):349-355. <https://doi.org/10.1136/jech.2006.047597>
20. McCulloch CE, Searle SR, Neuhaus JM. *Generalized, Linear, and Mixed Models*. Second ed. New York, NY: John Wiley & Sons; 2008.
21. Spiegel MR. *Mathematical Handbook of Formulas and Tables*. New York, NY: McGraw-Hill; 1968.

**How to cite this article:** Austin PC, Stryhn H, Leckie G, Merlo J. Measures of clustering and heterogeneity in multilevel Poisson regression analyses of rates/count data. *Statistics in Medicine*. 2018;37:572-589. <https://doi.org/10.1002/sim.7532>

## APPENDIX A

### VARIANCE PARTITION AND INTRACLASS CORRELATION IN A MULTILEVEL POISSON REGRESSION MODEL

Formulae for the mean, variance, and covariance in a Poisson-Gaussian mixed model are derived in McCulloch et al.<sup>20(section 7.3)</sup> Here, we give a stand-alone version of the derivation and state the connection to the VPC and ICC. Our calculations will use the well-known integration formula (eg, Spiegel<sup>21(p98)</sup>),

$$\int_{-\infty}^{\infty} e^{-(ax^2+2bx+c)} dx = \sqrt{\frac{\pi}{a}} e^{(b^2-ac)/a}, \quad (\text{A.1})$$

for  $a > 0$  and arbitrary real-valued  $b$  and  $c$ . For  $\sigma > 0$  and  $u \sim N(0,1)$ , it therefore holds that

$$E[e^{\sigma u}] = \int_{-\infty}^{\infty} e^{\sigma u} \left( \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \right) du = e^{\sigma^2/2}, \quad (\text{A.2})$$

using Equation A.1 with  $a = 1/2$  and  $b = -\sigma/2$ ; this is the moment-generating function of a Gaussian variable. In the 2-level Poisson multilevel model with normally distributed random effects, Formula 3 in the paper, we omit the indices where not needed, and denote by  $b = \beta\mathbf{X}$  (or  $b = \beta\mathbf{X} + \log(T)$  with an offset) the linear predictor for a scalar observation  $Y$ , and write  $\alpha \equiv \sigma u$ . Then we have

$$\begin{aligned} E[Y] &= E[E[Y|u]] = E[e^{b+\sigma u}] = e^{b+\sigma^2/2}, \\ \text{Var}[Y] &= \text{Var}[E[Y|u]] + E[\text{Var}[Y|u]] \\ &= E[(E[Y|u])^2] - (E[E[Y|u]])^2 + E[E[Y|u]] \\ &= E[e^{2b+2\sigma u}] - (E[Y])^2 + E[Y] \\ &= e^{2b+2\sigma^2} - e^{2b+\sigma^2} + e^{b+\sigma^2/2}, \end{aligned} \quad (\text{A.3})$$

using the formulae for mean and variance of the Poisson distribution as well as Equation A.2 for the last step in both the mean and variance calculations.

The calculation shows that in the resulting Equation A.3, the last term should be understood as the lowest level (ie, within-cluster) variance, whereas the 2 first terms together comprise the upper (cluster-) level variance. For an alternative interpretation, we show below by similar calculations that the first 2 terms together also equal the covariance between 2 observations from the same cluster that have *the same linear predictor*  $b$ . Therefore, the VPC and the ICC for 2 such observations coincide and take the form, upon inserting  $b = \beta\mathbf{X}$ ,

$$\text{VPC} = \text{ICC} = \frac{e^{2\beta\mathbf{X}+2\sigma^2} - e^{2\beta\mathbf{X}+\sigma^2}}{e^{2\beta\mathbf{X}+2\sigma^2} - e^{2\beta\mathbf{X}+\sigma^2} + e^{\beta\mathbf{X}+\sigma^2/2}}, \quad (\text{A.4})$$

as stated in Equation 4 in the paper.

Next, we calculate for 2 observations, say  $Y_1$  and  $Y_2$ , from the same cluster with linear predictors  $b_1$  and  $b_2$ ,

$$\begin{aligned} \text{Cov}[Y_1, Y_2] &= E[Y_1 Y_2] - E[Y_1] E[Y_2] = E[E[Y_1 Y_2|u]] - E[Y_1] E[Y_2] \\ &= E[E[Y_1|u] E[Y_2|u]] - E[Y_1] E[Y_2] \\ &= E[e^{b_1+\sigma u} e^{b_2+\sigma u}] - E[Y_1] E[Y_2] \\ &= e^{b_1+b_2+2\sigma^2} - e^{b_1+b_2+\sigma^2}, \end{aligned}$$

which for  $b_1 = b_2$  corresponds to the first 2 terms in Equation A.3.

We finally outline how these calculations can be extended to a Poisson random slopes model. The method of calculating the mean and variance in A.3 still applies when the conditional linear predictor involves not only the term  $\sigma u$  for the random intercept, but also similar random term(s) for random slope(s). With an assumed multivariate Gaussian distribution for all cluster-level random terms, the linear combination of these in the linear predictor will be univariate Gaussian, and hence, after some algebraic manipulation, Equation A.2 will once more yield expressions for the resulting means (expectations) over the random effects.

## APPENDIX B

### R CODE FOR THE CASE STUDY ANALYSES WITH THE FULLY ADJUSTED REGRESSION MODEL

```
library(lme4)
```

```
zlist <- list(hospital=0,age=0,sbp=0,resp=0,sod136=0,hgb100=0,urea=0,
  cancer=0,copd=0,cvd=0,cirr=0,dem=0,readmissions=0,at.risk.time=0,
  log.at.risk.time=0)
```

```

chf.df <- data.frame(scan("cohort.chf.dat",zlist))

attach(chf.df)

# Standardize the continuous variables to have mean zero and unit variance.
age.s <- (age - mean(age))/sd(age)
sbp.s <- (sbp - mean(sbp))/sd(sbp)
resp.s <- (resp - mean(resp))/sd(resp)
urea.s <- (urea - mean(urea))/sd(urea)

# Fit multilevel Poisson regression model with hospital-specific
# random effects.
mlm <- glmer(readmissions ~ age.s + sbp.s + resp.s + sod136 + hgb100 + urea.s +
  cancer + copd + cvd + cirr + dem + (1|hospital),
  family=poisson(link="log"),offset=log.at.risk.time,nAGQ=7)

# Extract the variance of the random effects.
tau.info <- summary(mlm)$varcor$hospital
tau <- attr(tau.info,"stddev")
tau2 <- tau^2

# Extract the regression coefficients.
beta.coefs <- summary(mlm)$coefficients

#####
# Compute the Median Rate Ratio
#####

MRR <- exp(sqrt(2*tau2) * qnorm(0.75))

#####
# ICC using exact calculations
# We evaluate the ICC at four different covariate patterns.
# For each covariate pattern, we allow the offset term to range from
# 74 days (10th percentile in the data) to 365 (90th percentile in the data).
#####

# Vectors for storing the ICCs.
ICC.1 <- NULL
ICC.2 <- NULL
ICC.3 <- NULL
ICC.4 <- NULL

counter <- 0

for (at.risk.time in 74:365){
  counter <- counter + 1

  # Reference patient: Average of continuous covariates and
  # binary variables set to zero.
  XB <- beta.coefs[1] + log(at.risk.time)
  ICC.1[counter] <- (exp(2*XB + 2*tau2) - exp(2*XB + tau2)) /
    (exp(2*XB + 2*tau2) - exp(2*XB + tau2) + exp(XB + tau2/2))
}

```

```

# Patient 20 years older than average.
XB <- beta.coefs[1] + beta.coefs[2]*(20/sd(age)) + log(at.risk.time)
ICC.2[counter] <- (exp(2*XB + 2*tau2) - exp(2*XB + tau2)) /
  (exp(2*XB + 2*tau2) - exp(2*XB + tau2) + exp(XB + tau2/2))

# Patient 10 years older than average and with COPD.
XB <- beta.coefs[1] + beta.coefs[2]*(10/sd(age)) + beta.coefs[9] + log(at.risk.time)
ICC.3[counter] <- (exp(2*XB + 2*tau2) - exp(2*XB + tau2)) /
  (exp(2*XB + 2*tau2) - exp(2*XB + tau2) + exp(XB + tau2/2))

# Patient 20 years older than average with cancer.
XB <- beta.coefs[1] + beta.coefs[2]*(20/sd(age)) + beta.coefs[8] + log(at.risk.time)
ICC.4[counter] <- (exp(2*XB + 2*tau2) - exp(2*XB + tau2)) /
  (exp(2*XB + 2*tau2) - exp(2*XB + tau2) + exp(XB + tau2/2))
}

#####
# Compute the VPC using a simulation-based approach
#####

N.sims <- 100000
# Number of simulation iterations.

# Set random number seed for reproducibility.
set.seed(24042017)

uj <- rnorm(N.sims, mean=0, sd=tau)
# Generate random effects from the estimated random effects distribution.

VPC.1 <- NULL
VPC.2 <- NULL
VPC.3 <- NULL
VPC.4 <- NULL
counter <- 0

for (at.risk.time in 74:365){
  counter <- counter + 1

  # ICC for a reference patient (average of continuous variables) and binary
  # variables set to zero.
  LP <- beta.coefs[1] + uj + log(at.risk.time)
  Y <- exp(LP)
  # Predicted count of readmissions.

  var.Y <- Y
  # Poisson variance is equal to the mean.

  VPC.1[counter] <- (var(Y))/(var(Y) + mean(var.Y))

  # ICC for a patient 20 years older than average
  LP <- beta.coefs[1] + beta.coefs[2]*(20/sd(age)) + uj + log(at.risk.time)
  Y <- exp(LP)
  # Predicted count of readmissions.

```

```

var.Y <- Y
# Poisson variance is equal to the mean.

VPC.2[counter] <- (var(Y))/(var(Y) + mean(var.Y))

# ICC for a patient 10 years older than average and with COPD.
LP <- beta.coefs[1] + beta.coefs[2]*(10/sd(age)) + beta.coefs[9] + uj + log(at.risk.time)
Y <- exp(LP)
# Predicted count of readmissions.

var.Y <- Y
# Poisson variance is equal to the mean.

VPC.3[counter] <- (var(Y))/(var(Y) + mean(var.Y))

# ICC for a patient 20 years older than average and with cancer.
LP <- beta.coefs[1] + beta.coefs[2]*(20/sd(age)) + beta.coefs[8] + uj + log(at.risk.time)
Y <- exp(LP)
# Predicted count of readmissions.

var.Y <- Y
# Poisson variance is equal to the mean.

VPC.4[counter] <- (var(Y))/(var(Y) + mean(var.Y))

}

```

## APPENDIX C

### STATA CODE FOR FITTING THE MULTILEVEL POISSON REGRESSION MODEL

```

* Load the data.
infile hospital age sbp resp sod136 hgb100 urea cancer copd cvd cirr dem ///
  readmissions at_risk_time log_at_risk_time using "cohort.chf.dat"

* Standardize the continuous variables to have mean zero and unit variance.
egen age_s = std(age)
egen sbp_s = std(sbp)
egen resp_s = std(resp)
egen urea_s = std(urea)

* Fit multilevel Poisson regression model with hospital-specific random effects.
mepoisson readmissions age_s sbp_s resp_s sod136 hgb100 urea_s cancer copd ///
  cvd cirr dem, intmethod(mcaghermite) intpoints(7) ///
  offset(log_at_risk_time) || hospital:

* Extract the variance of the random effects.
scalar tau2 = _b[var(_cons[hospital]):_cons]
scalar list tau2

* Extract the regression coefficients.
matrix beta_coefs = e(b)
matrix list beta_coefs

```



\*\*\*\*\*

\* Compute the Median Rate Ratio

\*\*\*\*\*

scalar MRR = exp(sqrt(2\*tau2)\*invnorm(0.75))

scalar list MRR

\*\*\*\*\*

\* ICC using exact calculations

\* We evaluate the ICC at four different covariate patterns.

\* For each covariate pattern, we allow the offset term to range from

\* 74 days (10th percentile in the data) to 365 (90th percentile in the data).

\*\*\*\*\*

\* Store the SD of age.

summarize age

scalar sd\_age = r(sd)

\* Generate the exposure.

clear

set obs 292

generate at\_risk\_time = 73 + \_n

\* Reference patient: Average of continuous covariates and binary variables set

\* to zero.

generate XB1 = beta\_coefs[1,12] + log(at\_risk\_time)

generate ICC1 = (exp(2\*XB1 + 2\*tau2) - exp(2\*XB1 + tau2)) / ///  
(exp(2\*XB1 + 2\*tau2) - exp(2\*XB1 + tau2) + exp(XB1 + tau2/2))

\* Patient 20 years older than average.

generate XB2 = beta\_coefs[1,12] + beta\_coefs[1,1]\*(20/sd\_age) ///  
+ log(at\_risk\_time)

generate ICC2 = (exp(2\*XB2 + 2\*tau2) - exp(2\*XB2 + tau2)) / ///  
(exp(2\*XB2 + 2\*tau2) - exp(2\*XB2 + tau2) + exp(XB2 + tau2/2))

\* Patient 10 years older than average and with COPD.

generate XB3 = beta\_coefs[1,12] + beta\_coefs[1,1]\*(10/sd\_age) ///  
+ beta\_coefs[1,8] + log(at\_risk\_time)

generate ICC3 = (exp(2\*XB3 + 2\*tau2) - exp(2\*XB3 + tau2)) / ///  
(exp(2\*XB3 + 2\*tau2) - exp(2\*XB3 + tau2) + exp(XB3 + tau2/2))

\* Patient 20 years older than average with cancer.

generate XB4 = beta\_coefs[1,12] + beta\_coefs[1,1]\*(20/sd\_age) ///  
+ beta\_coefs[1,7] + log(at\_risk\_time)

generate ICC4 = (exp(2\*XB4 + 2\*tau2) - exp(2\*XB4 + tau2)) / ///  
(exp(2\*XB4 + 2\*tau2) - exp(2\*XB4 + tau2) + exp(XB4 + tau2/2))

\*\*\*\*\*

\* Compute the VPC using a simulation-based approach

\*\*\*\*\*

\* Number of simulation iterations.

expand 10000

```
bysort at_risk_time: generate iteration = _n
```

```
* Set random number seed for reproducibility.
```

```
set seed 24042017
```

```
* Generate random effects from the estimated random effects distribution.
```

```
generate uj = rnormal(0,sqrt(tau2))
```

```
* Reference patient: Average of continuous covariates and binary variables set
```

```
* to zero.
```

```
generate LP1 = XB1 + uj
```

```
generate Y1 = exp(LP1) // Predicted count of readmissions.
```

```
generate VAR1 = Y1 // Poisson variance is equal to the mean.
```

```
* Patient 20 years older than average.
```

```
generate LP2 = XB2 + uj
```

```
generate Y2 = exp(LP2)
```

```
generate VAR2 = Y2
```

```
* Patient 10 years older than average and with COPD.
```

```
generate LP3 = XB3 + uj
```

```
generate Y3 = exp(LP3)
```

```
generate VAR3 = Y3
```

```
* Patient 20 years older than average with cancer.
```

```
generate LP4 = XB4 + uj
```

```
generate Y4 = exp(LP4)
```

```
generate VAR4 = Y4
```

```
* Collapse data over 100000 iterations and generate relevant SDs and means.
```

```
collapse (sd) SD_Y1 = Y1 SD_Y2 = Y2 SD_Y3 = Y3 SD_Y4 = Y4 ///
```

```
(mean) VAR1 VAR2 VAR3 VAR4, by(at_risk_time)
```

```
* Generate variance of predicted counts.
```

```
generate VAR_Y1 = SD_Y1^2
```

```
generate VAR_Y2 = SD_Y2^2
```

```
generate VAR_Y3 = SD_Y3^2
```

```
generate VAR_Y4 = SD_Y4^2
```

```
* Generate ICCs.
```

```
generate VPC1 = VAR_Y1/(VAR_Y1 + VAR1)
```

```
generate VPC2 = VAR_Y2/(VAR_Y2 + VAR2)
```

```
generate VPC3 = VAR_Y3/(VAR_Y3 + VAR3)
```

```
generate VPC4 = VAR_Y4/(VAR_Y4 + VAR4)
```