Nephrol Dial Transplant (2015) 30: 1244–1249
doi: 10.1093/ndt/gfv071
Advance Access publication 25 March 2015

Analyzing hospitalization data: potential limitations of Poisson regression

Colin G. Weaver[1], Pietro Ravani[1,2], Matthew J. Oliver[3,4], Peter C. Austin[5,6] and Robert R. Quinn[1,2]

[1]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada, [2]Department of Medicine, University of Calgary, Calgary, Alberta, Canada, [3]Department of Medicine, University of Toronto, Toronto, Ontario, Canada, [4]Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada, [5]Institute for Clinical Evaluative Sciences, University of Toronto, Toronto, Ontario, Canada and [6]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Correspondence and offprint requests to: Robert R. Quinn; E-mail: rob.quinn@albertahealthservices.ca

ABSTRACT

Background. Poisson regression is commonly used to analyze hospitalization data when outcomes are expressed as counts (e.g. number of days in hospital). However, data often violate the assumptions on which Poisson regression is based. More appropriate extensions of this model, while available, are rarely used.

Methods. We compared hospitalization data between 206 patients treated with hemodialysis (HD) and 107 treated with peritoneal dialysis (PD) using Poisson regression and compared results from standard Poisson regression with those obtained using three other approaches for modeling count data: negative binomial (NB) regression, zero-inflated Poisson (ZIP) regression and zero-inflated negative binomial (ZINB) regression. We examined the appropriateness of each model and compared the results obtained with each approach.

Results. During a mean 1.9 years of follow-up, 183 of 313 patients (58%) were never hospitalized (indicating an excess of 'zeros'). The data also displayed overdispersion (variance greater than mean), violating another assumption of the Poisson model. Using four criteria, we determined that the NB and ZINB models performed best. According to these two models, patients treated with HD experienced similar hospitalization rates as those receiving PD {NB rate ratio (RR): 1.04 [bootstrapped 95% confidence interval (CI): 0.49–2.20]; ZINB summary RR: 1.21 (bootstrapped 95% CI 0.60–2.46)}. Poisson and ZIP models fit the data poorly and had much larger point estimates than the NB and ZINB models [Poisson RR: 1.93 (bootstrapped 95% CI 0.88–4.23); ZIP summary RR: 1.84 (bootstrapped 95% CI 0.88–3.84)].

Conclusions. We found substantially different results when modeling hospitalization data, depending on the approach used. Our results argue strongly for a sound model selection process and improved reporting around statistical methods used for modeling count data.

Keywords: hospitalization, negative binomial, Poisson, zero-inflated negative binomial, zero-inflated Poisson

INTRODUCTION

Hospitalization rates are of relevance to patients, health care providers and policymakers because they are a measure of burden of disease and a key driver of health care costs [1]. Hospitalization rates (i.e. some event count over time) can be expressed as number of days spent in hospital per unit time, as we do in this study, or as number of admissions per unit time. Poisson regression is commonly used to estimate the effect of different interventions or risk factors on the rate of hospitalization [2–12].

However, Poisson regression requires that certain assumptions be met [13]. First, Poisson regression assumes that, conditional on a given set of covariate values, the variance of the distribution of the number of events is equal to the mean. Poisson regression may not perform well in situations where the conditional variance is greater than the conditional mean, a phenomenon known as overdispersion [14]. Second, in situations where a large proportion of patients are never hospitalized (i.e. have a zero count), data do not fit a Poisson distribution well. Both of these problems—too much variability and excess zeros—often occur in practice, but studies that have used Poisson regression to model hospitalization data have rarely reported whether these assumptions were tested or met [2–12]. While alternative modeling techniques have been developed to address these issues [13], they have not been widely used in nephrology research. The use of inappropriate models may have led to misleading results, as study conclusions often vary depending on the extent to which the requirements of the chosen model are satisfied [15, 16].

This paper used data from a previously published study [17] that compared hospitalization rates between patients

treated with hemodialysis (HD) and peritoneal dialysis (PD). In this re-examination, we compared the results of Poisson regression with three other models for count data that address overdispersion and/or excess zeros to determine if there was a meaningful impact on results and conclusions.

MATERIALS AND METHODS

Study design and patient population

We used data from a study that enrolled patients who were eligible for both HD and PD and who began outpatient dialysis in one of four Canadian dialysis programs (the Manitoba Renal Program, Sunnybrook Health Sciences Centre, London Health Sciences Centre, and Halton Healthcare Services) between 21 July 2007 and 30 April 2010 [17]. Research ethics board approval was obtained at all four sites. Patients were classified according to the initial outpatient dialysis modality. All patients were followed for at least 6 months or until the occurrence of kidney transplantation, death or the end of the follow-up period (4 November 2010). We sought to determine the impact of the initial outpatient modality on subsequent hospitalization rates.

Statistical analysis

Patient demographic characteristics, comorbidities and laboratory results were compared between those on HD and PD using Fisher's exact test for categorical characteristics, and either two-sided, two-sample *t*-tests or Wilcoxon rank-sum tests for continuous characteristics.

We used four regression models to estimate the effect of dialysis modality on the rate of hospitalization (count of days spent in hospital per year of patient follow-up). To allow for comparability, all models were adjusted for the same patient demographic characteristics (age, sex), comorbidities (diabetes mellitus, coronary artery disease, congestive heart failure, peripheral vascular disease, a history of gastrointestinal bleeding), baseline laboratory results (estimated glomerular filtration rate, serum albumin), whether the patient received at least 4 months of pre-dialysis care, and whether dialysis started in hospital [2–12]. We screened all variables for multi-collinearity.

The Poisson model. We first fit a Poisson model to estimate the effect of dialysis modality on the rate of hospitalization. We then estimated the rate ratio (RR) comparing the hospitalization rate for HD compared with that for PD. We used graphical methods to compare observed rates to those predicted by the model.

Extensions of the Poisson model. Three common extensions of Poisson regression can address problems of overdispersion and excess zeros (see Supplementary data for visual representations). Negative binomial (NB) regression accounts for overdispersion by adding an additional dispersion (variance) parameter to the Poisson model. This model can accommodate increased variability [13]. The zero-inflated Poisson (ZIP) model deals with excess zeros by combining a logistic regression model with a traditional Poisson model [13, 18]. The Poisson portion models the count of days that individuals will spend in hospital, including the possibility of zero days. Then the logistic model estimates the probability—above and beyond that estimated by the Poisson model—an individual will spend zero days in hospital. The expected hospitalization rate for an individual is calculated by multiplying the probability of being hospitalized from the logistic model by the expected hospitalization rate from the Poisson model. Finally, zero-inflated negative binomial (ZINB) models can accommodate overdispersion and excess zeros, as they incorporate a logistic model and an NB model [13]. In our analyses, we adjusted for the same patient characteristics in the Poisson or NB and logistic models of the zero-inflated regressions.

The 'summary RR'. Zero-inflated regression involves the estimation of two models, and as a consequence reports both odds ratios (ORs) and RRs. In our case, there is both an OR and an RR associated with the effect of HD versus PD on the rate of hospitalization. This makes zero-inflated models difficult to interpret and challenging to compare with the single summary measures (the RRs) reported by Poisson or NB models. However, a summary estimate of the overall effect of a covariate on the outcome of interest can be derived by combining information from both models [19]. This is done by estimating two expected hospitalization rates for every individual: one assuming he or she is on HD, and one assuming he or she is on PD. (In both of these calculations, the expected hospitalization rate is the product of the probability of not having a zero from the logistic model and the expected rate from the Poisson or NB model.) The mean of the HD rates divided by the mean of the PD rates gives a value that approximates the effect of HD versus PD on the outcome, and is comparable with the RRs generated by the Poisson and NB models [19]. A confidence interval (CI) for this summary RR can be calculated via bootstrap resampling [20] or the delta method. In our analysis, we used 1000 bootstrap replicates to estimate non-parametric bootstrapped CIs using a normal approximation, implemented via a Stata ado file (colinweaver.net/summaryRR). For comparison, we also calculated bootstrapped CIs for the NB and Poisson RRs.

Model assessment. We used four methods to assess the relative performance of the models (Supplementary data). First, we used a modified likelihood ratio test [21] to examine whether the NB model provided a significantly better fit of the data than the Poisson model. The same test was used to compare ZINB regression to ZIP regression [13]. A significant result indicated that the data were overdispersed and that the NB model (or ZINB model) was preferred. Second, we used a Vuong test to compare Poisson and NB regression with their zero-inflated counterparts [using both the Akaike's information criterion (AIC) and Bayesian information criterion (BIC) adjustments] [22, 23]. The Vuong test compares the overall performance of non-nested models, taking into account model fit and parsimony (simplicity); in this case, a significant result indicates that the zero-inflated model fits the data better. Third, with a similar approach to the Vuong test, we used AIC and BIC values [13] to compare all models; a lower value indicates a better fit of the data after accounting for model

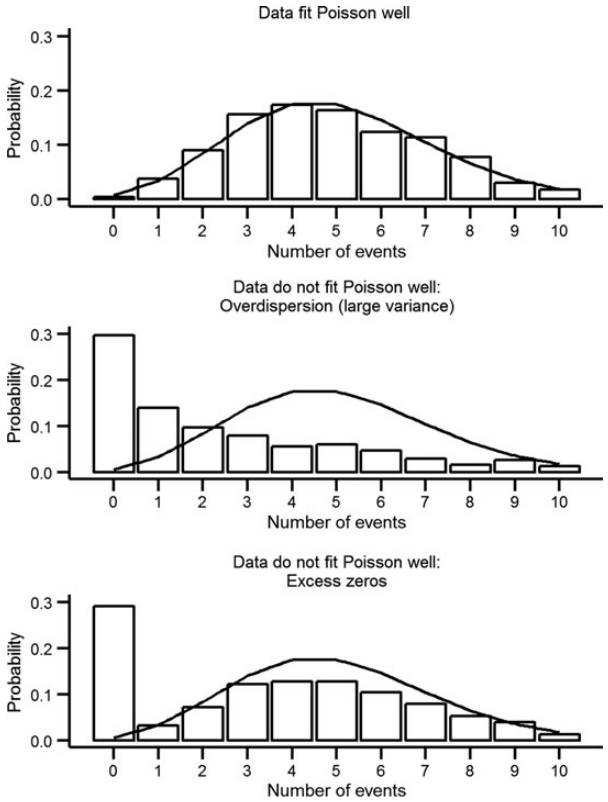

FIGURE 1: Assessing the models visually. Three randomly generated datasets with theoretical means of 5 are compared against a Poisson distribution with a mean of 5 (black line). Such a comparison can show if and how data deviate from a Poisson distribution—either via overdispersion or excess zeros. Data were generated from the following distributions: Poisson (top), negative binomial (middle) and zero-inflated Poisson (bottom).

complexity (i.e. the number of model parameters). Lastly, we visually assessed the models by plotting observed hospitalization rate frequencies versus those expected by each model [13]. Figure 1 demonstrates the rationale behind this visual comparison technique. We used Stata (Version 11.2, StataCorp., College Station, TX) for all analyses.

RESULTS

Complete data were available for 313 of 314 study participants. The first outpatient treatment was PD for 107 patients (34%) and HD for 206 patients (66%) (Table 1). The total number of days spent in hospital per patient-year of follow-up was 7.2 for the HD group and 4.2 for the PD group [unadjusted $RR_{HD:PD}$: 1.71; standard 95% (CI) 1.58–1.85].

The Poisson model

According to the standard Poisson model, HD patients had significantly higher hospitalization rates than PD patients (adjusted $RR_{HD:PD}$: 1.93; standard 95% CI 1.78–2.10). However, 183 patients (58%) were never hospitalized—a high frequency of zero counts when compared with the 13.6 (4%) expected from the Poisson model.

Table 1. Patient characteristics by dialysis treatment

	Patient group		P value
	PD	HD	
Number of patients	107	206	
Follow-up in days, median (IQR)	765 (625–882)	742 (637–861)	0.60
Started dialysis as inpatient, *n* (%)	15 (14)	115 (56)	<0.001
Age, mean (range)	61.2 (19–90)	62.5 (22–95)	0.48
Female, *n* (%)	49 (46)	87 (42)	0.55
Albumin (g/L), mean	34.1	30.3	<0.001
eGFR (mL/min/1.73 m^2), mean	8.1	8.2	0.32
Diabetes mellitus, *n* (%)	53 (50)	115 (56)	0.34
Congestive heart failure, *n* (%)	14 (13)	62 (30)	<0.001
Coronary artery disease, *n* (%)	21 (20)	57 (28)	0.13
Peripheral vascular disease, *n* (%)	8 (7)	34 (17)	0.035
Gastrointestinal bleeding, *n* (%)	7 (7)	21 (10)	0.40
Pre-dialysis care (>4 months), *n* (%)	97 (91)	137 (67)	<0.001

PD, peritoneal dialysis; HD, hemodialysis; IQR, interquartile range; eGFR, estimated glomerular filtration rate.

Model assessment

The likelihood ratio test indicated that the NB model fit the data significantly better than the Poisson model ($P < 0.001$). The three other methods used to assess appropriateness were inconclusive as to whether ZINB or NB provided a better fit for these data. The AIC and BIC Vuong tests were contradictory as to which of these two models was best ($P = 0.070$ and $P = 0.998$, respectively). The information criteria were also not in agreement, with AIC/BIC values of (lower is better): 1553/1654 (ZINB), 1567/1620 (NB), 5489/5586 (ZIP) and 10 924/10 973 (Poisson). The visual assessment shows ZINB to fit the data slightly better than NB, with both being visibly better than ZIP and Poisson (Figure 2).

Extensions of the Poisson model

When the data were modeled using ZIP regression, the adjusted summary $RR_{HD:PD}$ was 1.84 (bootstrapped 95% CI 0.88–3.84). (The bootstrapped 95% CI for the Poisson model was 0.88–4.23, which, unlike the standard CI, no longer shows a significant difference between the treatments.) Unlike the Poisson and ZIP models, when the data were modeled using NB and ZINB models, the RR and summary RR were much closer to 1 (Figure 3); the NB model had an adjusted $RR_{HD:PD}$ of 1.04 (bootstrapped 95% CI 0.49–2.20), and the adjusted summary $RR_{HD:PD}$ from the ZINB model was 1.21 (bootstrapped 95% CI 0.60–2.46).

DISCUSSION

In this study, we compared the results of Poisson regression with other approaches for modeling hospitalization data that address overdispersion and/or excess zeros. We found substantially

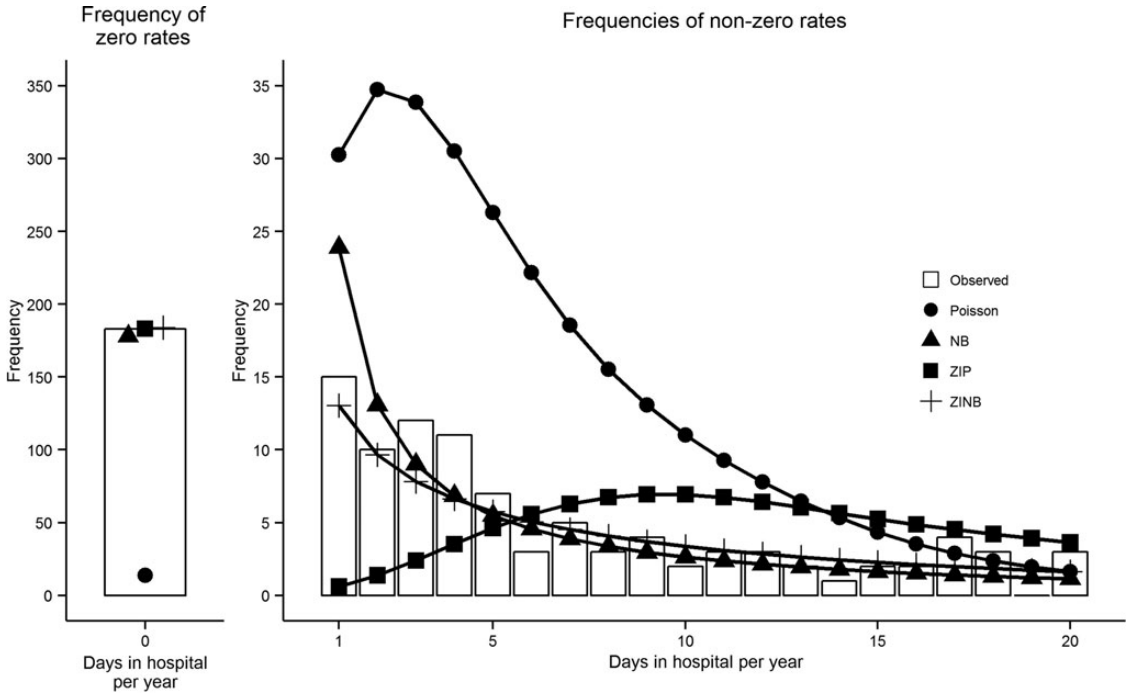

FIGURE 2: Comparison of the observed (actual) hospitalization rate frequencies with those expected from the models. This plot shows the number of patients (frequency; *y*-axis) that experienced a given number of days in hospital per year (observed; *x*-axis; bars). The lines show how closely the counts predicted by the four different models mirror the actual data. Eleven percent ($n = 35$) of patients had hospitalization rates of >20 days per year and are not shown. NB, negative binomial; ZIP, zero-inflated Poisson; ZINB, zero-inflated negative binomial.

Model	Probability of not hospitalized OR (standard 95% CI)	Length of stay RR (standard 95% CI)	Overall effect summary RR (standard / bootstrapped 95% CI)
Poisson	---	---	1.93 (1.78 - 2.10 / 0.88 - 4.23)
NB	---	---	1.04 (0.48 - 2.27 / 0.49 - 2.20)
ZIP	1.21 (0.69 – 2.11)	2.02 (1.85 – 2.19)	1.84 (NA / 0.88 - 3.84)
ZINB	1.31 (0.69 – 2.49)	1.37 (0.76 – 2.47)	1.21 (NA / 0.60 - 2.46)

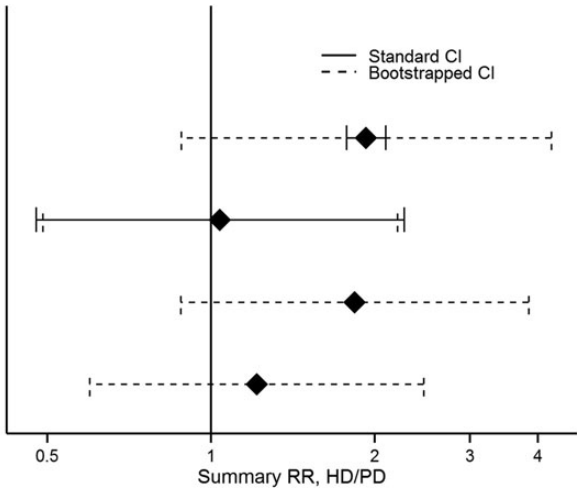

FIGURE 3: Estimated, adjusted effects of dialysis modality on hospitalization rate, by model. RRs and summary RRs were different according to the approach used to model hospitalization rates, with estimates much closer to 1 using NB and ZINB. NB, negative binomial; ZIP, zero-inflated Poisson; ZINB, zero-inflated negative binomial; OR, odds ratio; RR, rate ratio; HD, hemodialysis; PD, peritoneal dialysis; CI, confidence interval; NA, not applicable.

different estimates of the effect of dialysis modality on hospitalization rates, depending on the approach used. Traditional Poisson regression did not fit our data well and not only led to a higher relative rate of hospitalization among HD patients compared with patients treated with PD, but a statistically significant difference in hospitalization rates between groups. Models that accounted for the overdispersion contradicted this finding.

Our data are consistent with other studies. Point estimates have been shown to be less precise in situations where an inappropriate model is used [15, 16]. In addition, standard CIs are narrower in Poisson compared with NB models [14, 15], and tend to be narrower in ZIP compared with ZINB models [16]. This can lead to erroneously significant findings in the presence of overdispersion or excess zeros. In our example, the standard CIs generated by Poisson regression were much narrower than those generated by bootstrap methods or by NB regression. This resulted in the observation of a significant difference in the relative hospitalization rates among patients treated with HD compared with PD that was not present when more appropriate models were used.

Our results highlight the need to follow a sound model selection process when analyzing count data. A brief review of the literature comparing hospitalization rates between patients treated with HD or PD highlights the need to improve reporting around the statistical methods used [2–12]. None of the papers reviewed that used Poisson regression reported checking the assumptions underlying the approach or exploring other modeling strategies. Given the potential consequences of incorrect model choice, particularly in the presence of overdispersed data or an excess of zero counts, improved reporting and transparency around the statistical methods used is needed so that studies can be critically reviewed.

Despite their inherent appeal, there are important challenges when employing zero-inflated count models (ZIP and ZINB). In particular, communicating the results of an analysis that used a zero-inflated model can be difficult. These models produce two sets of results—one from the logistic portion of the model (ORs), and the other from the Poisson or NB portion of the model (RRs). While it is possible to calculate a summary RR of the overall effect of a covariate on the count outcome, in practice this takes some time and knowledge. Our analyses were done in Stata using an ado file we wrote to calculate the summary RR and its CI. The ado file is available online at colinweaver.net/summaryRR. In other statistical packages, predicted rates need to be used to calculate the summary RR, and a CI must be computed [19]. A CI can be calculated via bootstrap resampling or the delta method (the latter can be done with the NLMIXED command in SAS, see colinweaver.net/summaryRR). Simulation analyses comparing bootstrapping to the delta method for the summary RR CI have not been done; however, analyses using more complex models might look to bootstrapping [19], which is unfortunately the more computationally intensive method. Fitting our ZINB model on 1000 bootstrap resamples (a suggested number [13]) took 12 min on an Intel Xeon X5560 processor, with 313 individuals and 12 independent variables in each of the two models. Computing bootstrapped CIs with larger datasets and more variables could take prohibitively long.

Our study has several limitations. First, our dataset was relatively small and it is possible that larger or different datasets would give different results. However, previous studies would seem to support our conclusions regarding the modeling of count data [14–16]. At a minimum, our results reinforce the need for a sound model selection process and careful reporting of statistical methods in such analyses. Second, there is debate as to the best way to use zero-inflated models in practice. For example, there is a lack of consensus as to whether superior model fit is a sufficient reason to use a zero-inflated model. Some argue that there must also be a theoretical justification for two processes generating zero counts. For example, in a study where the outcome is number of pregnancies, zero counts could occur because the subject was male and unable to get pregnant, or because a woman did not become pregnant. If the sex of the subject was not observed, it would be logical to use a zero-inflated model. However, others argue that the presence of two processes generating zero counts is not a requirement for using zero-inflated models [13, 24]. We did not examine two other approaches often used when modeling count data: the hurdle model and the use of other calculation methods for CIs (besides bootstrap CIs). Hurdle models are very similar to zero-inflated models, containing a logistic model to address excess zeros and a truncated Poisson or NB model for non-zero counts (a summary RR can also be calculated for hurdle models) [13, 25]. Poisson and ZIP CIs can be modified to give more appropriate (usually wider) ranges when minor overdispersion is present [15, 16]: approaches include quasi-Poisson (Poisson regression only), bootstrap, robust and jackknife methods. It is important to note that these four techniques only affect the estimation of the variance, and therefore the width of the CIs, but not the point estimate (RRs) of the models.

We compared the results of Poisson regression with other approaches for modeling hospitalization data that address overdispersion and/or excess zeros. We found substantially different results about the effect of dialysis modality on hospitalization rates, depending on the approach used. Our results argue strongly for a sound model selection process and improved reporting around statistical methods used for modeling count data.

SUPPLEMENTARY DATA

Supplementary data are available online at http://ndt.oxfordjournals.org.

ACKNOWLEDGEMENTS

The authors thank Adriane Lewin for manuscript editing. The Physician Services Foundation, Incorporated, funded the Ontario portion of the original study that produced the data we used for this re-analysis. C.G.W. was supported by studentships from Alberta Innovates Health Solutions and the Markin Undergraduate Student Research Program. P.C.A. was supported in part by a Career Investigator Award from the Heart and Stroke Foundation, Ontario Office.

CONFLICT OF INTEREST STATEMENT

M.J.O. and R.R.Q. are co-leads for the International Society of Peritoneal Dialysis, North American Research Consortium which has received funding from Baxter Corporation. R.R.Q. received an honorarium from Baxter to speak at Canadian PD University which was donated to charity. M.J.O. and R.R.Q. are co-inventors of the Dialysis Measurement Analysis and Reporting (DMAR) system.

The results presented in this paper have not been published previously in whole or part, except in abstract format.

REFERENCES

1. Lee H, Manns B, Taub K *et al.* Cost analysis of ongoing care of patients with end-stage renal disease: the impact of dialysis modality and dialysis access. Am J Kidney Dis 2002; 40: 611–622

2. Bovbjerg RR, Diamond LH, Held PJ *et al.* Continuous ambulatory peritoneal dialysis: preliminary evidence in the debate over efficacy and cost. Health Aff (Millwood) 1983; 2: 96–102
3. Carlson DM, Duncan DA, Naessens JM *et al.* Hospitalization in dialysis patients. Mayo Clin Proc 1984; 59: 769–775
4. Gokal R, Jakubowski C, King J *et al.* Outcome in patients on continuous ambulatory peritoneal dialysis and haemodialysis: 4-year analysis of a prospective multicentre study. Lancet 1987; 2: 1105–1109
5. Burton PR, Walls J. A selection adjusted comparison of hospitalization on continuous ambulatory peritoneal dialysis and haemodialysis. J Clin Epidemiol 1989; 42: 531–539
6. Williams AJ, Nicholl JP, el Nahas AM *et al.* Continuous ambulatory peritoneal dialysis and haemodialysis in the elderly. Q J Med 1990; 74: 215–223
7. Singh S, Yium J, Macon E *et al.* Multicenter study of change in dialysis therapy-maintenance hemodialysis to continuous ambulatory peritoneal dialysis. Am J Kidney Dis 1992; 19: 246–251
8. Habach G, Bloembergen WE, Mauger EA *et al.* Hospitalization among United-States dialysis patients—hemodialysis versus peritoneal-dialysis. J Am Soc Nephrol 1995; 5: 1940–1948
9. Charytan C, Spinowitz BS, Galler M. A comparative study of continuous ambulatory peritoneal dialysis and center hemodialysis. Efficacy, complications, and outcome in the treatment of end-stage renal disease. Arch Intern Med 1986; 146: 1138–1143
10. Murphy SW, Foley RN, Barrett BJ *et al.* Comparative hospitalization of hemodialysis and peritoneal dialysis patients in Canada. Kidney Int 2000; 57: 2557–2563
11. Harris SA, Lamping DL, Brown EA *et al.* Clinical outcomes and quality of life in elderly patients on peritoneal dialysis versus hemodialysis. Perit Dial Int 2002; 22: 463–470
12. United States DoHaHS, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, US Renal Data System (USRDS). 2008 Annual Data Report. USRDS, 2008
13. Long JS, Freese J. Regression Models for Categorical Dependent Variables Using Stata, 2nd edn. College Station, TX: StataCorp LP, 2006
14. Hilbe JM. Negative Binomial Regression, 2nd edn. New York, NY: Cambridge University Press, 2011
15. Cox DR. Some remarks on overdispersion. Biometrika 1983; 70: 269–274
16. Perumean-Chaney SE, Morgan C, McDowall D *et al.* Zero-inflated and overdispersed: what's one to do? J Stat Comput Simul 2013; 83: 1671–1683
17. Quinn RR, Ravani P, Zhang X *et al.* Impact of modality choice on rates of hospitalization in patients eligible for both peritoneal dialysis and hemodialysis. Perit Dial Int 2014; 34: 41–48
18. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 1992; 34: 1–14
19. Albert JM, Wang W, Nelson S. Estimating overall exposure effects for zero-inflated regression models with application to dental caries. Stat Methods Med Res 2011; 23: 257–278
20. Efron B, Tibshirani R. An Introduction to the Bootstrap. New York, NY: Chapman & Hall, 1993
21. Gutierrez RG, Carter S, Drukker DM. sg160: on boundary-value likelihood-ratio tests. Stat Tech Bull 2001; 60: 15–18
22. Desmarais BA, Harden JJ. Testing for zero inflation in count models: bias correction for the Vuong test. Stat J 2013; 13: 810–835
23. Vuong QH. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 1989; 57: 307–333
24. Vittinghoff E, Shiboski S, McCulloch CE, Glidden DV. Regression Methods in Biostatistics. New York, NY: Springer, 2005
25. Mullahy J. Specification and testing of some modified count data models. J Econom 1986; 33: 341–365

Received for publication: 26.12.2014; Accepted in revised form: 18.2.2015