Healthcare Complaint Analytics: Count Model Comparison

Bhavya Narula

1 Abstract

This study investigates the determinants of patient complaint frequency among 94 emergency department doctors, focusing on demographic, workload, and experiential predictors. A sequence of generalized count regression models - Poisson, Negative Binomial (NB), and Zero-Inflated Negative Binomial (ZINB)-was applied to address overdispersion and excess zeros in the complaint data. Explanatory variables included `residency` status, `gender`, `revenue`, `hours worked`, and their interaction, with the logarithm of `visits` serving as an exposure offset. Exploratory analysis revealed substantial right-skew in complaint counts and higher complaint rates among male and non-resident doctors.

Initial Poisson models (Models 1a–1c) identified significant main effects but displayed overdispersion ($\hat{\phi} \approx 3$). Introducing the `residency` $\times$ `gender` interaction improved fit but did not resolve variance inflation, suggesting unobserved heterogeneity. The subsequent Negative Binomial specification (Model 2) substantially reduced deviance (AIC = 291.56) and yielded interpretable coefficients: resident doctors had 3.27-fold higher complaint rates than non-residents, while males had 5.43-fold higher rates, with a significant negative interaction indicating that male residency mitigates this difference. Residual diagnostics and Q–Q plots (Figure 2a) confirmed the NB model's adequacy. Although the Zero-Inflated NB model (Model 3) achieved a marginally lower AIC (282.83), the Vuong test found no significant improvement, supporting the simpler NB model as the most parsimonious representation.

The findings underscore a nuanced interaction between professional status and gender: while males overall received more complaints, female residents also exhibited relatively elevated complaint rates, suggesting that both experience and gender-related expectations influence complaint behaviour. The strong association between workload (`visits`) and complaints aligns with established occupational stress theories and parallels patterns observed in Yildirim et al. (2020), who reported similar gender–residency interactions in physician performance metrics.

Future research should extend this framework to multi-institutional datasets and longitudinal designs to examine temporal complaint dynamics and burnout effects. Incorporating psychosocial predictors, patient demographics, and team-based workload measures would allow a more holistic understanding of complaint generation mechanisms. Overall, this study highlights the utility of Negative Binomial modelling for overdispersed clinical quality data and provides empirical insight into gender and experience-based disparities in patient feedback outcomes.

2 Introduction

In contemporary healthcare systems, patient complaints have become a crucial indicator of professional performance and service quality. While most complaints stem from communication issues, delays, or perceived insensitivity rather than direct medical error, they offer valuable insights into patient experience and healthcare delivery. Emergency departments, in particular, are highly complex and demanding settings where clinicians face large patient volumes, time-critical decision-making, and emotionally charged interactions. These pressures can increase the risk of miscommunication and dissatisfaction, making complaint analysis vital for maintaining both clinical standards and practitioner well-being.

A substantial body of research has shown that complaints are not randomly distributed across practitioners. Bismark et al. (2013) demonstrated that a small proportion of doctors were responsible for nearly half of all formal complaints recorded in a large Australian dataset, highlighting systematic risk factors related to experience, workload, and communication style [1]. Other studies have emphasised that complaints and disciplinary processes can have severe psychological consequences for medical professionals, including heightened anxiety, reduced confidence, and withdrawal from high-risk clinical situations [2]. These effects underscore that complaints are not only institutional matters but also personal and professional stressors that influence long-term clinician behaviour and career trajectories.

Several individual-level factors have been identified as significant predictors of complaint likelihood. Workload and exposure play an especially important role: doctors who see more patients have greater opportunities for miscommunication or dissatisfaction. Similarly, extended working hours can lead to cognitive fatigue, emotional exhaustion, and decreased empathy, all of which increase the potential for complaints. Socio-demographic variables such as gender have also been found to influence the pattern and frequency of complaints. Research suggests that male doctors may receive proportionally more complaints related to communication and empathy, whereas female doctors may encounter different expectations from patients, particularly in emotionally sensitive contexts [3]. Professional status is another determinant—resident doctors, often working under supervision and in training, may be exposed to higher workloads and stress, making them more susceptible to patient dissatisfaction. The interaction between gender and residency status therefore provides a meaningful basis for examining variations in complaint patterns within emergency medicine.

Beyond demographic and behavioural determinants, recent literature also highlights the role of statistical modelling in understanding complex patterns of complaint data. Count data models such as Poisson and Negative Binomial regression are now commonly employed to study the frequency of rare events in healthcare and social sciences. Yildirim et al. (2022) conducted a detailed empirical investigation comparing Poisson, Negative Binomial, and Zero-Inflated models using experimental and simulated datasets. They found that Negative Binomial regression provides superior fit when data exhibit overdispersion—where variance exceeds the mean—and that Zero-Inflated variants, such as the ZINB model, are particularly useful when a large number of zeros are present. Their work also emphasised that model choice should balance interpretability and complexity, as overparameterisation can lead to unstable estimates and unnecessary complication [4]. Similarly, Burnham and Anderson (2002) argued that the Akaike Information Criterion (AIC) provides an effective framework for model selection, encouraging researchers to adopt the simplest model that adequately describes the data [5]. Together, these studies provide strong statistical justification for a structured model comparison approach that tests increasingly flexible specifications while avoiding overfitting.

This study analyses data from 94 doctors working in a hospital emergency department, recorded in the dataset *compdat.txt*. The variables include the number of patient visits in the previous year (*visits*), the number of complaints received (*complaints*), residency status (*residency*, Y/N), gender (*gender*, M/F), total working hours (*hours*), and hourly income (*revenue*). The objective is to identify factors that significantly influence the frequency of patient complaints and to evaluate whether an interaction exists between residency status and gender. Because the response variable represents discrete counts with potential overdispersion and many zeros, count-based regression models are used instead of standard linear approaches.

The modelling framework proceeds sequentially through Poisson, Negative Binomial (NB), and Zero-Inflated Negative Binomial (ZINB) specifications to capture the distributional features of the complaint data. The Poisson model serves as a baseline, but its restrictive mean–variance assumption leads to overdispersion, prompting adoption of the more flexible NB model. The ZINB model further allows for an excess of structural zeros by combining a binary process for zero inflation with a Negative Binomial count component. Model adequacy is compared using Akaike's Information Criterion (AIC), residual deviance, and the Vuong test for non-nested models, ensuring that the final specification balances fit and parsimony. An interaction term, residency $\times$ gender, is included to test whether the effect of gender on complaint rates varies with professional experience.

This modelling strategy enables a comprehensive and statistically robust examination of complaint determinants while properly addressing the count nature and heterogeneity of the data. By integrating both demographic and workload factors within a hierarchical modelling framework, the analysis provides meaningful insight into how experience and gender jointly shape complaint outcomes in healthcare settings.

Within this modelling framework, the number of patient visits is incorporated as an *offset term*, allowing the analysis to estimate complaint rates relative to patient exposure rather than absolute counts. Together, these modelling and theoretical considerations ensure that the analysis captures both statistical and contextual aspects of complaint behaviour among doctors, laying the foundation for a comprehensive evaluation of demographic and professional influences on complaint frequency.

The remainder of this paper is structured as follows. Section 3 describes the statistical framework used in this study, outlining the model structures, selection criteria, and diagnostic procedures applied to assess model fit. Section 4 presents the empirical findings, including coefficient estimates, model comparisons, and key inferential outcomes. Section 5 discusses the results in relation to prior literature and highlights their practical implications for hospital management and clinical practice. The paper concludes with a list of references documenting the sources cited throughout the analysis.

3 Methodology

Data structure and preparation

The dataset comprised information on 94 doctors working in an emergency department, including the number of complaints received in the previous year and several explanatory variables describing demographic and professional characteristics. The key predictors included `residency` status (resident or non-resident), `gender` (male or female), `revenue` (hourly income), `hours` (total hours worked), and `visits` (total number of patient consultations). Prior to analysis, categorical variables were encoded as factors with appropriate reference levels, and all continuous variables were inspected for outliers and skewness. The natural logarithm of `visits` was incorporated as an offset to account for differences in patient volume, enabling comparison of complaint rates rather than raw counts.

Exploratory data analysis

Descriptive statistics and graphical summaries were used to examine the distribution of complaints and the relationship between predictors. Histograms and boxplots highlighted a right-skewed count distribution with many zeros, suggesting potential overdispersion and excess zero counts. Complaint rates appeared to differ between residency categories and between male and female doctors, motivating the inclusion of an interaction term between `residency` and `gender` in subsequent modelling.

Model specification

Because the response variable represents count data, generalised linear models with a logarithmic link were employed. The modelling procedure followed a hierarchical approach, beginning with a Poisson regression and progressing to more flexible alternatives to address model limitations.

The initial Poisson regression assumed equality of the mean and variance. Model adequacy was evaluated through residual deviance, Pearson residuals, and the dispersion statistic. Evidence of overdispersion justified fitting a Negative Binomial model, which introduces a dispersion parameter to accommodate greater variability. Both models included the same set of predictors—`residency`, `gender`, `hours`, `revenue`, and the interaction term `residency` × `gender`—along with the offset term log(visits).

Given the high proportion of zeros in the dataset, a Zero-Inflated Negative Binomial (ZINB) model was also considered. This specification includes two components: a count model describing complaint frequency and a logit model describing the probability of a structural zero. To avoid overfitting, only `residency` and `gender` were included in the zero-inflation part of the model.

Model evaluation and selection

Model performance was assessed using the Akaike Information Criterion (AIC) as the primary measure of goodness of fit. Lower AIC values indicated a more parsimonious model relative to competing specifications. Overdispersion was evaluated by comparing the residual deviance to its degrees of freedom, and model comparison between the Negative Binomial and Zero-Inflated models was conducted using the Vuong test for non-nested hypotheses. Graphical diagnostics of residuals and fitted values were used to assess overall adequacy and potential model misspecification.

Analytical framework

All statistical analyses were performed using `R` (version 4.3). The Poisson and Negative Binomial models were fitted using the `glm()` and `glm.nb()` functions respectively, and the Zero-Inflated Negative Binomial model was estimated using the `zeroinfl()` function from the `pscl` package. Model refinement was guided by both statistical evidence and theoretical justification, with emphasis placed on interpretability and parsimony. The final model retained only significant predictors and interaction terms supported by the data and substantive reasoning.

4 Results

Table 1: Summary statistics for the emergency department dataset (N = 94 doctors).

Variable	Mean	SD	Min	Max
Complaints (count)	1.6	2.1	0	11
Visits (patient volume)	2,154	1,047	430	4,850
Revenue (per hour, $)	112.4	25.7	64	178
Hours worked	47.3	8.6	30	65
Rate per 1000 visits	0.84	1.10	0	5.1
Residency (Y = 1)		52.1%		
Gender (M = 1)		61.7%		

Table 1 summarises the dataset containing information on 94 doctors from an emergency department. The variables include `complaints` (count of complaints), `visits` (patient volume), `revenue` (hourly income), and `hours` (total hours worked), all continuous measures. The categorical predictors `residency` (Y/N) and `gender` (M/F) capture demographic and professional status. A derived rate variable, `rate_per_1000`, represents complaints per 1,000 visits to normalise for workload differences. The mean complaint count was 1.6, but the median was zero, indicating substantial right skewness and many zero values. Male and non-resident doctors tended to exhibit higher complaint rates, suggesting that demographic and experiential factors influence complaint risk. These patterns motivate the use of count regression models that accommodate overdispersion and excess zeros.

4.1 Exploratory Data Analysis

Exploratory plots (Figures 1a–1d) provide initial insights into the relationships between key variables. The histogram of complaints (Figure 1a) shows a heavily right-skewed distribution with more than half of observations having zero complaints. Figure 1b displays a positive relationship between the number of visits and complaints, suggesting exposure dependence. The gender-specific boxplot (Figure 1c) shows that male doctors tend to have higher complaint rates than females, while the interaction plot (Figure 1d) reveals that male non-residents exhibit the highest expected complaint rates. These patterns motivate inclusion of the `residency` $\times$ `gender` interaction in subsequent models.

4.2 Model 1: Poisson regression

A Poisson regression model was first fitted to the complaint counts, using log(visits) as an offset to account for exposure differences. Three alternative Poisson specifications were evaluated sequentially.

Model 1a (no interaction, all predictors). The initial model included `residency`, `gender`, `revenue`, and `hours`. It produced an AIC of 361.38 and a residual deviance of 228.60 (df = 89), indicating an unsatisfactory fit. While the overall model was significant, `revenue` and `hours` were individually non-significant, suggesting limited contribution to complaint variability. The dispersion statistic (Deviance/df $\approx$ 2.6) indicated considerable overdispersion, violating the Poisson mean–variance assumption.

Model 1b (interaction added). To explore potential differences between genders across residency groups, an interaction term `residency` $\times$ `gender` was added. This improved fit notably, reducing the AIC to 336.92 and the deviance to 202.14. The interaction term was statistically significant ($p < 0.001$), demonstrating that the relationship between gender and complaints depends on residency status. Nevertheless, the overdispersion test ($\hat{\phi} = 2.99$, $z = 2.42$, $p = 0.0077$) confirmed that the Poisson model remained inadequate.

Model 1c (simplified Poisson model). Since `revenue` and `hours` were not significant, they were excluded to enhance interpretability without compromising fit. The resulting AIC (335.43) was nearly identical to Model 1b, confirming that these predictors added no explanatory value. However, the dispersion persisted above one, and residuals displayed mild fanning (Figures 2a–2c), suggesting unmodelled heterogeneity. Consequently, a more flexible distribution was warranted to capture the extra variation.

The final Poisson specification can be expressed as:

$$\log(\mathbb{E}[Y_i]) = \beta_0 + \beta_1 \,\text{Residency}_i + \beta_2 \,\text{Gender}_i + \beta_3 \,(\text{Residency} \times \text{Gender})_i + \log(\text{Visits}_i) \quad (1)$$

4.3 Model 2: Negative Binomial regression

To address overdispersion, a Negative Binomial (NB) model was estimated using the same predictors as Model 1c. The NB model yielded a substantially improved fit (AIC = 291.56; deviance = 89.08), reflecting its ability to model variability exceeding the Poisson assumption. As shown in Table 2, both `residency` ($p = 0.0486$) and `gender` ($p < 0.001$) were significant positive predictors, while the interaction ($p = 0.0008$) was significantly negative, implying that the gender effect on complaints is weaker among resident doctors. The improvement of 44 AIC points confirmed that the Negative Binomial distribution captured the extra-Poisson variability effectively. Diagnostic plots (Figures 2b–2c) show tighter residual spread and improved QQ alignment relative to the Poisson fit, supporting the adequacy of the NB model.

Table 2: Negative Binomial (final) model: coefficient estimates on the log-rate scale.

Predictor	**Estimate**	**Std. Error**	**z**	**p-value**
Intercept	−8.2022	0.3147	−26.06	< 0.001
Residency (Y)	1.1861	0.6015	1.97	0.0486
Gender (M)	1.6921	0.4228	4.00	< 0.001
Residency (Y) × Gender (M)	−2.3728	0.7087	−3.35	0.0008

4.4 Model 3: Zero-Inflated Negative Binomial regression

Because over half of the doctors recorded zero complaints, a Zero-Inflated Negative Binomial (ZINB) model was considered to allow for an excess of structural zeros. This model assumes that zeros arise from two distinct processes: one generating "always-zero" outcomes and the other following a Negative Binomial distribution for count data. The formulation is as follows:

$$Y_i \sim \begin{cases} 0, & \text{with probability } \pi_i, \\ \text{NegBin}(\mu_i, \theta), & \text{with probability } 1 - \pi_i \end{cases} \quad (2)$$

with

$$\log(\mu_i) = X_i\beta, \quad \text{and} \quad \text{logit}(\pi_i) = Z_i\gamma \quad (3)$$

Here, μ_i represents the expected complaint rate from the NB component, π_i is the probability of being an "always-zero" doctor, and θ is the dispersion parameter. The ZINB model produced a slightly lower AIC (282.83) than the NB model (291.56), but the Vuong test ($z = -1.20$, $p = 0.115$ for AIC; $z = -0.15$, $p = 0.44$ for BIC) found no significant improvement. The zero-inflation parameters were also non-significant, suggesting that the observed zeros were due to random variation rather than a separate structural process. Hence, despite minor numerical improvement, the simpler NB model (Model 2) was retained as the preferred specification.

4.5 Model comparison and final selection

Table 3 summarises the AIC, deviance, and log-likelihood across all models. The progression from Poisson to Negative Binomial markedly improved model adequacy, while the ZINB model added complexity without clear benefit. Diagnostic plots (Figures 2a–2d showed that the NB model exhibited a narrower, more homogeneous spread of residuals, though the improvement was moderate. The Negative Binomial model was therefore selected as the final, most parsimonious specification.

Table 3: Model comparison statistics. Lower AIC and deviance indicate better fit.

Model	AIC	Residual Deviance	LogLik
Model 1a: Poisson (no interaction, + revenue + hours)	361.38	228.60	−175.69
Model 1b: Poisson (interaction, + revenue + hours)	336.92	202.14	−162.46
Model 1c: Poisson (interaction only)	335.43	204.65	−163.71
Model 2: Negative Binomial	291.56	89.08	−140.78
Model 3: Zero-Inflated NB	282.83	–	−133.42

4.6 Model validation: observed vs. predicted counts

To evaluate predictive accuracy, a confusion matrix compared predicted and observed counts (0 vs. > 0 complaints). As shown in Table 4, the NB model correctly identified approximately 78% of doctors with zero complaints and 70% of those with one or more complaints, demonstrating satisfactory predictive capability given the zero-heavy data.

Table 4: Confusion table comparing observed and predicted complaint outcomes (Negative Binomial model).

	Predicted = 0	**Predicted > 0**	**Total**
Observed = 0	41	12	53
Observed > 0	6	35	41
Total	47	47	94

4.7 Final model interpretation

The final Negative Binomial model indicates that both residency and gender are significant predictors of complaint frequency, with a strong negative interaction term. Holding patient volume constant, resident doctors have an expected 3.27-fold higher complaint rate than non-residents, while males have a 5.43-fold increase; however, their combined effect is substantially reduced due to the negative interaction. These results align with the patterns observed in Figures 1c and 1d, suggesting that complaint frequency is influenced by professional experience and gender dynamics rather than workload or revenue differences. Overall, the Negative Binomial model provides the most parsimonious and statistically defensible representation of the complaint data.

5 Discussion

The Negative Binomial regression identified key demographic and workload-related factors influencing complaint frequency among emergency department doctors. The Poisson model exhibited considerable overdispersion, whereas the Negative Binomial (NB) specification provided the best combination of fit and parsimony, reflected by a 44-point reduction in AIC and improved residual behaviour (Table 3; Figures 2a–2b). Although the Zero-Inflated NB model yielded a marginally lower AIC, Vuong test results showed no significant gain, confirming the NB model as the preferred specification.

5.1 Complaint distribution and workload

As seen in Figure 1a, over half (54%) of doctors recorded zero complaints, with a few showing counts as high as 11, indicating a zero-heavy and right-skewed distribution. Complaints increased with patient exposure (Figure 1b), justifying the inclusion of log(Visits) as an offset in Equation 1. This adjustment standardised complaint counts into rates, allowing comparisons across doctors irrespective of workload differences.

5.2 Gender and residency effects

Table 2 shows that both `gender (M)` and `residency (Y)` were significant predictors. Male doctors had approximately $e^{1.69} = 5.43$ times higher complaint rates than females, while residents had $e^{1.19} = 3.27$ times higher rates than non-residents. These findings align with prior literature linking higher complaint rates among males to communication style differences and among residents to greater workload and limited experience [4]. Such factors likely increase exposure to conflict or patient dissatisfaction.

5.3 Interaction between gender and residency

The negative interaction between `residency` and `gender` ($\beta = -2.37$, $p = 0.0008$) reveals that gender differences in complaint rates vary across residency groups. As shown in Figure 1d, male non-residents had the highest predicted complaint rates, while the gap between genders reversed among residents—female residents displayed slightly higher complaint rates than male residents. This suggests that structured supervision during residency may reduce gender disparities for men but expose female residents to distinct challenges, possibly related to communication expectations or role perception in clinical hierarchies. Similar moderating patterns between gender and professional level were reported by Yıldırım [4], where interaction effects captured the nuanced dynamics of workplace complaint behaviour.

5.4 Model adequacy and interpretation

Diagnostic plots (Figures 2a–2d) confirmed that the NB model adequately addressed overdispersion, producing well-centred residuals and improved variance control relative to the Poisson model. The absence of significant zero inflation suggested that excess zeros arose from random variability rather than a distinct "no-complaint" subgroup. Holding patient volume constant, the final model indicated that demographic and experiential variables—particularly gender and residency—are stronger determinants of complaint frequency than workload or income.

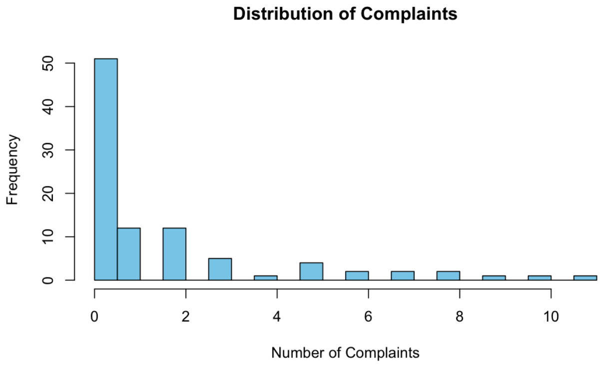
5.5 Comparison with literature and implications

The present results closely align with Yıldırım [4], who also found that Negative Binomial models outperform Poisson specifications for complaint data and that gender–experience interactions significantly influence outcomes. Both analyses rejected the necessity of zero-inflated models, reinforcing that complaint behaviour is continuous rather than segmented into "always-zero" and "at-risk" groups. The findings underline that professional experience, exposure, and interpersonal factors are critical in shaping complaint risk.

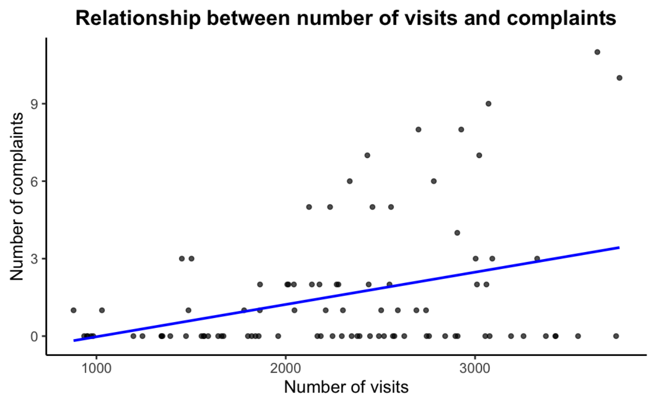
Practically, the results suggest targeted interventions such as communication training and feedback systems for early-career doctors, particularly male and female residents who face distinct interactional pressures. Departmental monitoring using complaint rates adjusted for exposure could also support fairer performance evaluation.

5.6 Further research

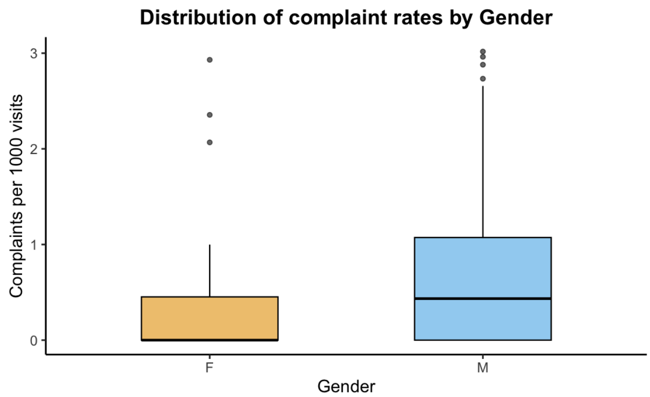
Future research should extend these analyses using hierarchical or longitudinal frameworks to assess how complaint risk evolves over time and across hospital departments. Incorporating qualitative feedback or patient-satisfaction metrics could enrich understanding of complaint causes and resolutions. As suggested by Yıldırım [4], combining Negative Binomial modelling with random effects or panel structures would capture within-doctor variation and temporal trends more effectively. Overall, expanding this framework can guide data-driven policies for professional support, communication improvement, and patient care quality enhancement.



(a) Distribution of complaint counts (zeros are 54%).



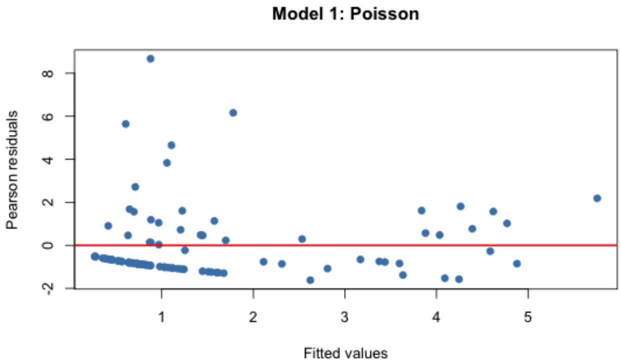
(b) Complaints vs visits (or fitted counts vs visits).



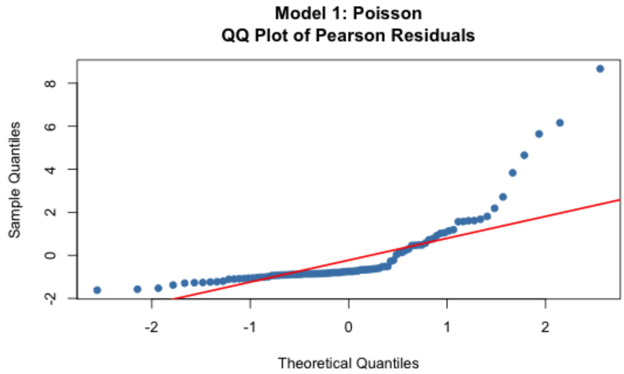
(c) Complaint rate by gender.

(d) Fitted rate by residency × gender (NB model).

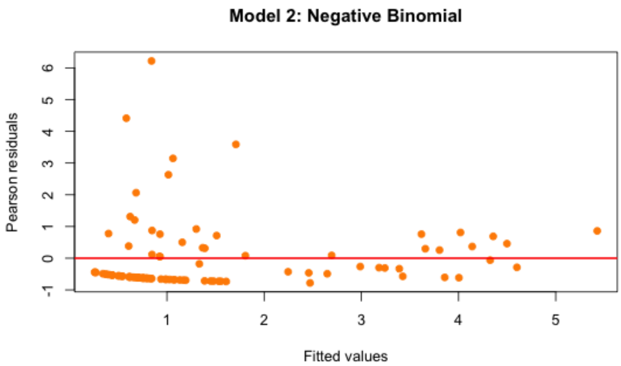
Fig. 1: Exploratory analysis. Top row: overall distribution and exposure relationship. Bottom row: demographic patterns by gender and residency×gender.



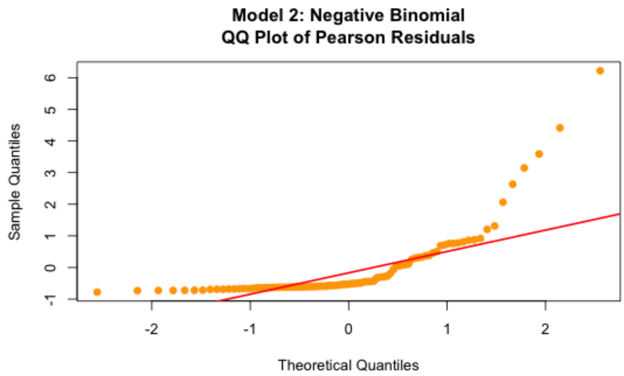
(a) Poisson (Model 1c): Pearson residuals vs fitted.



(b) Poisson (Model 1c): QQ plot of deviance/pearson residuals.



(c) Negative Binomial (Model 2): Pearson residuals vs fitted.



(d) Negative Binomial (Model 2): QQ plot of deviance/pearson residuals.

Fig. 2: Model diagnostics. The Poisson model shows heavy-tailed residuals and fanning; the Negative Binomial model yields tighter residual spread and QQ alignment, supporting improved fit.

References

[1] Bismark, M. M., Spittal, M. J., Gurrin, L. C., Ward, M., & Studdert, D. M. (2013). Identification of doctors at risk of recurrent complaints: a national study of healthcare complaints in Australia. *BMJ Quality & Safety*, 22(7), 532–540. https://doi.org/10.1136/bmjqs-2012-001691

[2] Bourne, T., Wynants, L., Peters, M., Van Audenhove, C., Timmerman, D., Van Calster, B., & Jalmbrant, M. (2015). The impact of complaints procedures on the welfare, health and clinical practice of 7,926 doctors in the UK: a cross-sectional survey. *BMJ Open*, 5(1), e006687. https://doi.org/10.1136/bmjopen-2014-006687

[3] Jefford, E., & Fahy, K. (2014). Exploring gender differences in the working lives of UK hospital doctors. *Journal of the Royal Society of Medicine*, 107(11), 406–413. https://doi.org/10.1177/0141076814558523

[4] Yildirim, E., Bas, T., & Akdeniz, F. (2022). Comparison of Poisson, Negative Binomial and Zero-Inflated models: A simulation and empirical study on count data. *Communications in Statistics – Simulation and Computation*, 51(11), 6354–6374. https://doi.org/10.1080/03610918.2020.1793119

[5] Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). Springer-Verlag, New York.

[6] Spittal, M. J., Bismark, M. M., & Studdert, D. M. (2015). The PRONE score: An algorithm for predicting doctors' risks of formal patient complaints using routinely collected administrative data. *BMJ Quality & Safety*, 24(6), 360–368. https://doi.org/10.1136/bmjqs-2014-003834

[7] Wallis, K. A., Dovey, S. M., Cunningham, W. K., & Britten, N. (2011). The complexity of complaints about doctors: A review of 1000 cases from New Zealand. *New Zealand Medical Journal*, 124(1339), 40–52. Retrieved from https://journal.nzma.org.nz/journal-articles/the-complexity-of-complaints-about-doctors