Visit us at stand #1028

numerus

Practical Challenges in the Implementation of ML-NMR for Rate Outcomes:

Lessons from a Recent European HTA Submission

Poster code **MSR50**

Authors

Jason Wilson (pictured), Senior Statistician

Lawrence Pont, Principal Statistician

Louis Alderton, Statistician

Numerus Ltd



Introduction

Multilevel network meta-regression (ML-NMR) is increasingly preferred by Health Technology Assessment (HTA) agencies for anchored population-adjusted indirect comparisons. During a recent European HTA submission in a rare indication, an initial fixed-effect network meta-analysis (NMA) faced agency requests for an ML-NMR to address baseline treatment effect modifier imbalances across trial populations. However, some endpoints of interest were count outcomes. Currently, ML-NMR is performed using the multinma R package available on CRAN. As of October 2026, for count and rate outcomes, multinma only natively supports modelling using the Poisson likelihood, which presents a significant limitation when handling overdispersed count data.

Methods

The primary rate outcome was initially modelled using Poisson likelihoods for both individual patient data (IPD) and aggregate-level data (AgD) trials. This approach led to severe convergence issues and uninterpretable credible intervals (CrIs), likely driven by overdispersion in the IPD trial.

Solution

To address this, the underlying Stan code of multinma (v0.7.2) was manually modified to incorporate the negative binomial (NB) distribution. Two custom approaches were evaluated: a dual-NB model (applying an NB likelihood to both IPD and AgD models, sharing a size parameter informed by the IPD) and a hybrid NB-Poisson model (applying an NB likelihood to the IPD model alongside a Poisson likelihood for the AgD model). The prior chosen for the size parameter was a half-Cauchy distribution with a mean of 0 and a scale of 5.

Table 1
Data simulation summaries by treatment arm

Study	Treatment	N	Age - years, mean (SD)	Sex– male, %	Baseline severity, mean (SD)	# Events	Exposure time
Study 1 - IPD	Active 1	249	55.8 (14.0)	44%	55.2 (9.8)	524	1
	Placebo	253	55.5 (14.0)	43%	55.3 (9.6)	920	1
Study 2 - ALD	Active 2	252	63.8 (11.1)	59%	47.5 (10.6)	775	1.5
	Placebo	250	64.2 (10.5)	58%	45.3 (11.3)	955	1.5
Study 3 - ALD	Active 3	114	57.2 (12.1)	45%	59.2 (9.1)	264	2
	Placebo	112	59.5 (13.5)	51%	64.8 (8.4)	368	2
Study 4 - ALD	Active 4	218	59.7 (10.3)	59%	52.4 (10.5)	289	1.5
	Placebo	220	60.3 (9.5)	67%	50.1 (11.8)	450	1.5

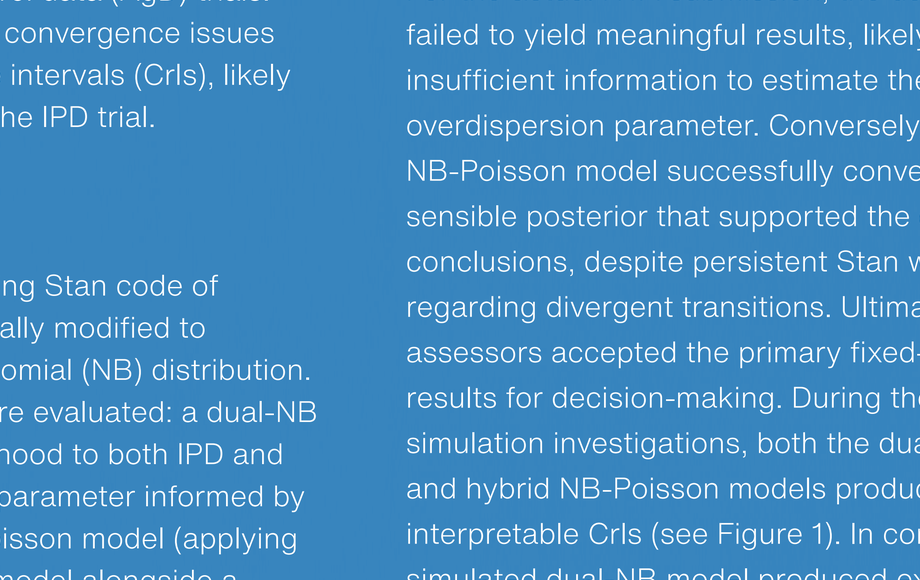
Investigation

Post-submission, these methods were further investigated using a simulated star-shaped network anchored by a placebo. Event counts for 502 subjects were simulated using a Poisson distribution, incorporating a subject-specific random effect to induce overdispersion alongside prognostic and treatment-modifying covariates (age, sex, and baseline disease severity score). Simulated summary data is presented in Table 1.

Results

For the actual HTA submission, the dual-NB model failed to yield meaningful results, likely due to insufficient information to estimate the AgD overdispersion parameter. Conversely, the hybrid NB-Poisson model successfully converged to a sensible posterior that supported the original NMA conclusions, despite persistent Stan warnings regarding divergent transitions. Ultimately, the HTA assessors accepted the primary fixed-effect NMA results for decision-making. During the subsequent simulation investigations, both the dual-Poisson and hybrid NB-Poisson models produced stable, interpretable CrIs (see Figure 1). In contrast, the simulated dual-NB model produced extremely wide CrIs for the AgD studies, confirming that the model struggled to explore the parameter space and achieve convergence.

Figure 1
Comparison of ML-NMR methods for count data



Discussion

This case study underscores the critical need for native negative binomial extensions within the ML-NMR framework to robustly handle overdispersed count data. It also highlights major operational challenges for HTDs under compressed timelines (e.g., 100 days for JCA dossier), where long run times and non-convergence issues must be anticipated.

Findings

Simulation results showed that a dual-Poisson model can converge in overdispersed settings given a simpler network and larger IPD sample size, indicating that the non-convergence originally observed was likely driven by a combination of overdispersion, small sample size, and network complexity. Despite likelihood inconsistencies, the hybrid model successfully increased variance of treatment contrasts within the network whilst having minimal impact on the point estimates.

Limitations

Several caveats remain for the hybrid approach; because the sum of independent NB distributions is not Poisson, individual- and aggregate-level likelihoods are not entirely analogous (Phillippo et al., 2020). Furthermore, if network exchangeability holds, overdispersion should logically be present across both IPD and AgD trials. The hybrid approach does not account for overdispersion in the AgD trials making a network-wide hyperparameter theoretically superior.

Conclusion

While a hybrid NB-Poisson model can operationally bypass Poisson-driven convergence issues in overdispersed ML-NMR networks, its theoretical likelihood inconsistencies highlight a critical need for native negative binomial framework extensions. Ultimately, when complex Bayesian computational bottlenecks arise under compressed HTA timelines, alternative population-adjustment methods like MAIC and STC remain vital sensitivity analyses to safeguard submissions.

References: Phillippo DM, Dias S, Ades AE, Belger M, Brnabic A, Schacht A, Saure D, Kadziola Z, Welton NJ. Multilevel network meta-regression for population-adjusted treatment comparisons. J R Stat Soc Ser A Stat Soc. 2020 Jun;183(3):1189-1210. doi: 10.1111/rssa.12579. Epub 2020 Jun 7. PMID: 32684669; PMCID: PMC7362893.

numerus

United Kingdom Numerus Ltd
+44 1189 770 077
Germany Numerus AG
+49 7121 1372 936

Get in touch
bd@numerus.com
www.numerus.com