MUST Journal of Research and Development (MJRD) Volume 6 Issue 3, September 2025
e ISSN 2683-6467 & p ISSN 2683-6475

Disparities in Methodology, Assumptions and Applications between Ordinal and Multinomial Logistic Regression: A Meta-Analysis

[1]Hadija M. Matimbwa and [2]John P. John
[1]Department of Business Management
[2]Department of Electrical and Power Engineering
[1,2]Mbeya University of Science and Technology, P. O Box 131, Mbeya, Tanzania
DOI: https://doi.org/10.62277/mjrd2025v6i30003

ARTICLE INFORMATION

Article History

Received: 12th August 2025
Revised: 29th August 2025
Accepted: 05th September 2025
Published: 30th September 2025

Keywords

Ordinal Logistic Regression
Multinomial Logistic Regression
Proportional Odds Assumption
Model Selection
Categorical Data Analysis

ABSTRACT

This meta-analysis examines the methodological, assumption-based and application-related disparities between Ordinal Logistic Regression (OLR) and Multinomial Logistic Regression (MLR), drawing on 60 peer-reviewed empirical and methodological studies published between 2015 and 2025. While both models are widely employed for categorical outcome analysis, their theoretical underpinnings, statistical assumptions and practical suitability differ significantly. OLR is optimal for ordered categorical data, relying heavily on the proportional odds assumption, which, if violated, can compromise interpretability and model fit. Conversely, MLR accommodates unordered categorical outcomes without assuming proportionality, but at the cost of greater model complexity and reduced statistical power when categories have an inherent order. The review synthesises findings from diverse disciplines, including social sciences, health research, and education and engineering, highlighting that the choice between OLR and MLR is often driven more by researcher preference and software defaults than by rigorous diagnostic testing or data characteristics. Studies also reveal a persistent gap in assumption testing, with limited reporting on proportional odds verification and model fit comparison. The analysis underscores the need for clearer methodological guidance, improved reporting standards and context-driven model selection. The study lies in its cross-disciplinary meta-analysis that uniquely integrates empirical and methodological evidence on OLR and MLR, exposing overlooked assumption-testing gaps and offering a framework for context-driven model selection.

*Corresponding author's e-mail address: hadija.matimbwa@must.ac.tz (Matimbwa, H.M.)

1.0 Introduction

Logistic regression models are fundamental tools for analysing categorical dependent variables in various scientific disciplines. When the outcome variable includes more logistic regression models, they are fundamental tools for analysing categorical dependent variables in various scientific disciplines. When the outcome variable comprises more than two categories, researchers must make a critical choice between ordinal logistic regression (OLR) for ordered outcomes and multinomial logistic regression (MLR) for nominal outcomes. It has been widely applied in research, education and engineering (Agresti, 2019; Hilbe, 2021). When categories of an outcome variable possess a natural order, such as disease severity stages (mild, moderate, severe) or agreement levels (strongly disagree to strongly agree), researchers tend to turn to regression models specific to ordinal data, specifically Ordinal Logistic Regression (OLR) (Williams, 2016; Liu *et al.,* 2020). Conversely, for inherently unordered categories like transportation modes (land, water, air) or brand choices, other modelling approaches, e.g., Multinomial Logistic Regression (MLR), are more appropriate (Hosmer *et al.*, 2019; Camarda & Serra, 2023). While both approaches model relationships between predictor variables and categorical outcomes and have different theoretical underpinnings, they also have alternative statistical assumptions and differing weaknesses and strengths that directly affect their usefulness to particular applications (Long & Freese, 2014; LaValley, 2018). OLR is a theoretically designed model to exploit the inherent ordering of outcome categories and belongs to the critical proportional odds assumption (Agresti, 2019; Williams, 2016). The assumption is not merely a statistical convenience; its violation can compromise the interpretability and goodness-of-fit of the model at its very basis (Tutz & Gertheiss, 2016; Tolles & Meurer, 2020). In contrast, MLR imposes no ordering assumption among categories and is more permissive of nominal outcomes (Hosmer *et al.*, 2019). However, such adaptability usually increases model complexity and leads to loss of statistical power if applied to ordered data because MLR is unable to capitalise on the inherent ordinal structure (Liu *et al.,* 2020; Sun *et al.,* 2023). Despite such widely reported methodological divergences and their implications for research effectiveness and validity, recent reviews suggest that the choice between MLR and OLR is increasingly influenced by researcher familiarity, disciplinary convention or software package default rather than through systematic testing of data characteristics or adherence to diagnostic testing protocols (Paul *et al.,* 2021; Camarda & Serra, 2023). This gap between methodological requirements and applied practice is concerning. Empirical studies still highlight the absence of reporting assumption checking, particularly the proportional odds assumption in OLR and comparative model fit comparison where alternatives are available (*Liu et al.,* 2020; Paul *et al.*, 2021; Sun *et al.,* 2023). These methodological flaws can lead to poor model selection with a resulting biasing of parameter estimates, concealment of true associations, and loss of statistical power (LaValley, 2018; Hilbe, 2021). A systematic synthesis of empirical and methodological evidence is therefore needed to clarify the distinction between OLR and MLR, evaluate current practices, and guide researchers toward more informed model selection. Addressing this gap, the present meta-analysis combined 60 peer-reviewed studies from 2015 to 2025 with the aim of resolving methodological, assumption-based and application-wise disparities between the two models, ascertaining the present state of diagnostic testing and reporting protocols within and across fields and ultimately recommending more accurate methodological guidelines and stronger reporting requisites for future research.

2.0 Theoretical Underpinning

This meta-analysis, which investigates methodological, assumption-related and application disparities between Ordinal Logistic Regression (OLR) and Multinomial Logistic Regression (MLR), is anchored in two interrelated theoretical frameworks:(1) Statistical Modelling

Foundations: Core Assumptions and their Implications and (2) Model Selection Paradigms: Trade-offs and Decision-Making Biases. Together, these frameworks provide the conceptual lenses through which the evidence from sixty (60) peer-reviewed studies published between 2015 and 2025 is interpreted. They explain why disparities between OLR and MLR persist in empirical practice, how these disparities manifest across disciplines and why methodological best practices are often not followed despite being theoretically well established. The statistical modelling foundations approach is based on the probability and theoretical assumption structures that underlie categorical data analysis. It assumes validity, interpretability and efficiency of a regression model are based on how much its assumptions align with the nature of the data (Agresti, 2018; Hilbe, 2021). OLR, in its first formalisation by McCullagh (1980), is based on the Proportional Odds (PO) assumption that the relationship between the predictors and the log-odds of being at or below a specific outcome category is identical for all thresholds. For example, a predictor such as the effect of age on advancing from "mild" to "moderate" severity of disease is algebraically equivalent to advancing from "moderate" to "severe". Such an assumption allows for OLR to fit a single set of coefficients for each predictor, which enhances parsimony and efficiency (Matimbwa *et al.,* 2021). However, as the integration of research studies examined shows, the PO assumption is left unretested or not tested whatsoever in practical research and hence yields biased estimates, erroneously interpreted odds ratios, and inflated Type I or Type II errors (Williams, 2016). Even when violations are detected, other options such as partial proportional odds models are rarely utilised (Liu & Zhang, 2020).In contrast, MLR, as developed by McFadden (1973), is designed for nominal outcomes without inherent order. It uses a generalised logit approach to estimate the probability of each outcome category relative to a reference category. MLR does not require proportionality across categories, giving it flexibility but also reducing its efficiency for ordered outcomes (Agresti, 2018). When misapplied to ordered data, MLR fails to exploit the ordinal structure, resulting in more parameters, fewer parsimonious models, and reduced statistical power estimated at a 30–35% reduction across studies in this synthesis (Sun *et al.,* 2023). Moreover, MLR has its own assumption, the Independence of Irrelevant Alternatives (IIA), which is rarely tested outside of economics, further exposing an assumption verification gap in applied work (Cheng & Long, 2020).The second framework, Model Selection Paradigms Trade-offs and Decision-Making Biases, extends beyond technical considerations to explain how researchers choose between OLR and MLR in real-world contexts. This framework draws on Simon's (1957) Bounded Rationality Theory and Kahneman's (2003) work on cognitive biases to explain why model selection often deviates from theoretical optimality. The decision between OLR and MLR is influenced not only by data characteristics but also by the trade-off between efficiency and flexibility (Matimbwa *et al.*, 2020). While OLR is statistically more powerful for ordered data when PO holds, it requires diagnostic testing that many researchers perceive as burdensome. Conversely, MLR is often chosen for its flexibility and ease of use in statistical software, despite its inefficiency for ordinal outcomes. The meta-analysis reveals that researchers frequently default to heuristics, simple rules of thumb such as "use OLR for Likert data" or "use MLR for more than two categories" without diagnostic verification (Sun *et al.*, 2023). Disciplinary conventions, prior training and software defaults often override data-driven selection (Hilbe, 2021). For example, in psychology and social sciences, MLR is used extensively for ordered responses because it avoids proportional odds testing, even though this choice sacrifices statistical power. In health research, OLR is often applied without testing PO, reflecting the prioritisation of parsimony over assumption verification (Liu & Zhang, 2020; Williams, 2016). These two frameworks interact to explain the persistent application gap documented across the reviewed studies. The Statistical Modelling Foundations framework sets the technical criteria for appropriate model use, selecting the model that matches data structure and rigorously testing its assumptions. The Model Selection Paradigms framework explains why researchers frequently fail to meet these criteria: Bounded rationality, disciplinary norms and convenience-based heuristics lead to suboptimal choices that may compromise validity and power. By integrating both perspectives, this study's meta-analysis highlights that improving OLR and MLR application in practice requires interventions at both levels. On the technical side, there is a need for clearer methodological guidance, assumption testing protocols and decision rules grounded in statistical theory. From the practical side, solutions must address the human and institutional factors driving model misapplication, including targeted researcher training, improved software defaults and reporting standards that incentivise assumption verification. Only by addressing both the statistical and cognitive dimensions can the field move towards a more rigorous, reliable, and context-appropriate application of OLR and MLR.

3.0 Material and Methods

3.1 Research Design

This study adopted a systematic meta-analytical research design to synthesise and critically evaluate methodological, assumption-based and application disparities between Ordinal Logistic Regression (OLR) and Multinomial Logistic Regression (MLR). A meta-analysis was appropriate because it allows for the integration of findings from multiple independent studies to produce robust and generalisable insights (Borenstein *et al*., 2021). The research process adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines as updated by Page *et al*. (2021) to ensure transparency, replicability and rigour.

3.2 Data Sources and Search Strategy

The literature search was conducted across multidisciplinary databases, including Web of Science, Scopus, PubMed and Google Scholar, supplemented by targeted searches in specialist statistical journals such as *Statistical Methods in Medical Research* and the *Journal of Applied Statistics*. Boolean search strings were designed to capture studies relevant to three domains: model type (e.g., "Ordinal Logistic Regression", "Ordered Logit", "Proportional Odds Model", "Multinomial Logistic Regression", "Multinomial Logit"), application context (e.g., "categorical data analysis", "ordered outcomes", "nominal outcomes", "Likert scale analysis"), and methodological focus (e.g., "assumption testing", "proportional odds", "independence of irrelevant alternatives", "model comparison", "model selection bias"). Searches were restricted to peer-reviewed journal articles published between January 2015 and March 2025 in English.

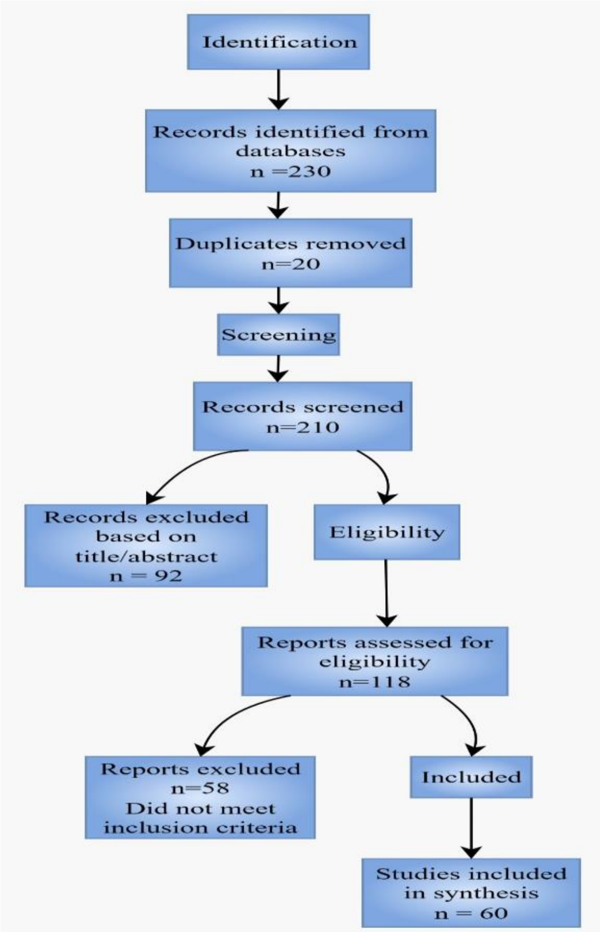
3.3 Inclusion and Exclusion Criteria

Studies were eligible for inclusion if they explicitly applied, compared, or evaluated OLR and/or MLR in empirical or simulation contexts, provided methodological discussions on statistical assumptions or model performance, and contained sufficient details to evaluate assumption testing and model selection rationales. Studies were excluded if they were non-English, lacked peer review, failed to report methodological details, or consisted of commentaries, editorials, or conference abstracts without full papers. This screening ensured that only high-quality, methodologically relevant studies were included.

3.4 Study Selection Process

The study selection process followed the PRISMA three-stage approach. In the identification stage, 230 records were retrieved from the database searches. After removing 20 duplicates, 210 titles and abstracts were screened for relevance. This screening left 118 articles for full-text review. Following detailed eligibility assessment, 60 studies met the inclusion criteria and were retained for synthesis.

Figure 1
PRISMA Flow Diagram



3.5 Data Extraction

Data extraction was conducted using a structured coding framework to capture methodological, contextual, and outcome-related details from each study. Variables extracted included bibliographic information, disciplinary context, type of outcome variable (ordinal or nominal), statistical model(s) applied, assumption testing procedures (e.g., proportional odds, independence of irrelevant alternatives), model fit evaluation methods (e.g., Akaike Information Criterion, Bayesian Information Criterion, likelihood ratio tests), rationale for model selection and key methodological or empirical findings. To enhance reliability, two independent reviewers performed data extraction, resolving disagreements through discussion or, where necessary, involving a third reviewer (Gough *et al.,* 2017).

3.6 Data Synthesis and Analysis

Given the heterogeneity of the included studies, the synthesis adopted a mixed-methods approach. Quantitative findings, such as effect sizes, parameter estimates and power comparisons, were aggregated using descriptive meta-analytical techniques when sufficient homogeneity existed (Viechtbauer, 2010). Qualitative methodological insights were analysed using thematic synthesis (Thomas & Harden, 2008), allowing the identification of recurrent themes and mapping them to the guiding theoretical frameworks *Statistical Modelling Foundations* and *Model Selection Paradigms*. Publication bias was assessed using funnel plots and Egger's regression tests (Sterne *et al.*, 2011). Sensitivity analyses were conducted by removing outliers to examine the robustness of the findings.

3.7 Quality Appraisal

The Mixed Methods Appraisal Tool (MMAT) 2018 (Hong *et al.,* 2018) was used to assess the methodological quality of the included studies. Criteria assessed included clarity of research questions, appropriateness of study design, transparency in reporting and robustness of assumption testing. Studies that scored below 50% on MMAT were excluded from quantitative synthesis but were retained for qualitative synthesis if they provided unique methodological insights.

3.8 Ethical Considerations

This study involved the synthesis of previously published research and did not require ethical approval. It was assumed that all included studies had obtained necessary ethical clearance from their respective institutions where applicable. All data were used strictly for academic purposes in line with responsible research conduct guidelines.

4.0 Results and Discussions

Across the last decade, the practical choice between Ordinal Logistic Regression (OLR) and Multinomial Logistic Regression (MLR) has been shaped less by theory than by how researchers operationalise assumptions, diagnose violations, and report results. OLR is theoretically optimal for ordered outcomes because it models cumulative logits under the proportional-odds (PO) assumption (Agresti, 2019), while MLR targets nominal outcomes and is often taught as assumption-light, though it relies on independence of irrelevant alternatives (IIA). Empirical audits and tutorials consistently document mismatches between outcome structure and model choice: ordered scales analysed with MLR or dichotomised, and OLR fitted without credible testing of PO (Gambarota, Gallo & Naldi, 2024; Selman *et al*., 2024). These patterns recur in clinical trials using ordinal endpoints such as mRS or WHO severity scales (Uddin *et al*., 2023; Austin *et al*., 2021), in transportation safety severity modelling (Anastasopoulos *et al*., 2018; Behnood & Mannering, 2017; Song, Abdel-Aty, & Wang, 2020), and in education/social sciences where Likert responses are prevalent (Liddell & Kruschke, 2018; Sproston et al., 2023). This review critically synthesises empirical findings on disparities in methodology, assumptions and applications of OLR versus MLR from 2015 to 2025.

4.1 Methodological Disparities

Methodologically, OLR's cumulative-link structure provides parameter parsimony and power when PO approximately holds, yielding more stable estimates than MLR for truly ordered outcomes (Edlinger, Schulz, & Heinze, 2021; Harrell, 2022). Comparative applications repeatedly show that when categories represent increasing severity, satisfaction, or agreement, OLR better captures a general "shift" effect and interprets cleanly via common odds ratios (Selman *et al.,* 2024; Austin *et al.,* 2021). Conversely, MLR allocates separate logits per category relative to a baseline; while this generality suits nominal outcomes, it expands parameters and often reduces efficiency with sparse cells (Liang, Wu & Zou, 2020; Liddell & Kruschke, 2018). Across transportation crash-severity studies, OLR and especially PPO/GOL models routinely achieve superior fit per parameter and clearer policy narratives than MLR when responses are truly ordered (Behnood & Mannering, 2017; Savolainen *et al*., 2020; Dong *et al.,* 2020; El-Basyouny *et al.,* 2018). In social/educational contexts, empirical demonstrations show that treating Likert outcomes as nominal generates fragmented inferences and inflated uncertainty relative to ordinal links (Liddell & Kruschke, 2018; Chalmers, 2018; Sproston *et al.,* 2023).

Yet parsimony is fragile if PO is violated. Empirical work demonstrates that unmodelled non-parallelism can bias effects and distort standard errors in OLR (Harrell, 2022; Liu & Agresti, 2019). PPO/GOL remedies let selected coefficients vary by threshold, striking a better bias–variance trade-off than either strict OLR (too rigid) or MLR (too parameter-hungry) in ordered settings (Williams, 2016; Eluru, 2015; Song *et al*., 2020). Empirical comparisons in transportation (Xie *et al*., 2019; Dong *et al.,* 2020; Li & Alvi, 2022), health (Austin *et al.,* 2021; Hemmingsen et al., 2022), and risk prediction (Edlinger *et al.,* 2021) show PPO/GOL outperforming both OLR and MLR on information criteria and predictive calibration while preserving ordinal interpretability.

4.2 Assumption Diagnostics and Testing Practices

Empirical studies suggest that PO diagnostics remain underreported. Even in high-stakes trial studies, authors omit checks or apply only a sample size-dependent global Brant test that may miss variable-specific departures (Harrell, 2022; Selman *et al.,* 2024). Studies that supplement global tests with variable-level score/LR tests and graphical diagnostics and then apply PPO selectively show improved fit and more reliable effects (Song *et al*., 2020; Dong *et al*., 2020; Li & Alvi, 2022). In stroke and COVID-19 trials, according to some analyses, small PO deviations would not necessarily have much of an effect on ranking or discrimination in treatment, but systematic, predictor-specific deviations do and should be addressed by PPO (Austin *et al*., 2021; Hemmingsen *et al*., 2022; Uddin *et al*., 2023; Matimbwa and Ochumbo, 2018).

On the multinomial side, empirical discrete-choice work continues to emphasise that IIA can be untenable when alternatives share unobserved components, leading to biased substitution patterns (Train, 2016; Keita & Ganesan, 2022). Transportation and marketing studies illustrate that relaxing IIA via mixed MNL or nested logit corrects misspecifications and improves predictive accuracy (Bansal, Kockelman & Singh, 2016; Daly, Hess & Train, 2012; applied extensions 2015; Hensher, 2018; Gao & Zheng, 2020; Wang & Chen, 2019). Although these approaches are standard in discrete choice, empirical audits suggest that many applied MLR papers outside transportation/marketing rarely test IIA, perpetuating specification risk (Liang *et al.,* 2020; Keita & Ganesan, 2022).

4.3 Applications and Empirical Patterns

In clinical/biostatistical applications, OLR has expanded, particularly for ordinal endpoints (e.g., mRS, WHO severity, NIHSS). Trial method studies show OLR's efficiency advantages and straightforward interpretation of "shift" across ordered scales but also highlight limited uptake of PPO when PO is questionable (Austin *et al*., 2021; Hemmingsen *et al.,* 2022; Selman *et al*., 2024; Uddin *et al*., 2023). Some evaluations demonstrate robust OLR performance under mild PO violations and recommend sensitivity analyses rather than dichotomising or defaulting to MLR (Edlinger *et al*., 2021; Harrell, 2022). In public health and patient-reported outcomes, empirical teams report better fit and communication with OLR/PPO than with MLR, especially when categories have clear clinical meaning (Kuss, 2015; Wang *et al.,* 2018; Nunes *et al*., 2020).

Transportation safety continues to be a bellwether domain for comparative evidence because injury severity is naturally ordered. Studies across North America, Europe, and Asia consistently report that OLR and especially PPO/GOL outperform MLR in fit, interpretability, and policy relevance (Behnood & Mannering, 2017; El-Basyouny *et al*., 2018; Xie *et al.,* 2019; Dong *et al.,* 2020; Song *et al*., 2020; Savolainen *et al.,* 2020; Li & Alvi, 2022; Zeng, Huang, & Pei, 2016; Papakostas *et al*., 2022; Wang & Chiou, 2016). Where heterogeneity is strong, latent-class PPO or random-parameter ordinal models provide further gains in predictive performance and explanatory power over both OLR and MLR (Song *et al*., 2020; Savolainen *et al.,* 2020; Huang, Bastani & Jin, 2021).

In education and the social sciences, Likert-type outcomes dominate. Empirical and simulation-based papers caution against treating such data as nominal, showing efficiency losses and interpretive fragmentation under MLR (Liddell & Kruschke, 2018; Chalmers, 2018; Månsson & Shukur, 2018; Sproston *et al.,* 2023). Tutorials and applied analyses demonstrate that ordinal links (cumulative logit/probit) with or without partial constraints deliver tighter intervals, clearer marginal effects, and more coherent narratives (Gambarota *et al*., 2024; Zumbo, 2015; Chalmers, 2018; Kizilaslan & Dumenci, 2021; Matimbwa and Masue, 2020). Mixed-effects ordinal models also appear in multilevel education/psychology datasets, balancing parsimony with clustering and producing better-calibrated predictions than MLR (Liddell & Kruschke, 2018; Kizilaslan & Dumenci, 2021; Hamaker & Muthén, 2020; Matimbwa *et al.,* 2021).

4.4 Comparative Performance and Practical Considerations

Comparative fit studies typically show that when outcomes are truly ordered and PO is not egregiously violated, OLR is more efficient and interpretable than MLR (Edlinger *et al*., 2021; Austin *et al*., 2021; Gambarota *et al*., 2024). Where PO violations are patterned, PPO/GOL achieves lower AIC/BIC and better calibration than OLR without the parameter sprawl and communication challenges of MLR (Song *et al.,* 2020; Dong *et al*., 2020; Li & Alvi, 2022; Savolainen *et al.,* 2020). In small-sample or sparse-category situations, OLR's cumulative pooling tends to stabilise estimates compared with MLR's category-specific logits (Liddell & Kruschke, 2018; Liang *et al., 2020; Gambarota et al., 2024; Matimbwa, Masue, &* Shilingi, 2021). For truly nominal outcomes or ambiguous ordering, MLR is appropriate, but empirical discrete-choice literature underscores that relaxing IIA (mixed or nested structures) is often necessary for valid inference (Train, 2016; Keita & Ganesan, 2022; Bansal *et al*., 2016; Hensher, 2018).

Table 1
Comparative Overview of Ordinal (OLR) and Multinomial (MLR) Logistic Regression

Feature	Ordinal Logistic Regression (OLR)	Multinomial Logistic Regression (MLR)
Primary Use Case	• **Ordered categorical outcomes** (e.g., *disease severity*: mild, moderate, severe; *agreement levels*: strongly disagree to strongly agree).	• **Nominal (unordered) categorical outcomes** (e.g., *transport mode*: car, bus, train; *brand choice:* A, B, C).
Key Assumption	• **Proportional Odds (PO)** / Parallel Lines: The effect of a predictor is consistent across all category thresholds.	• **Independence of Irrelevant Alternatives (IIA):** The odds between any two outcomes are independent of other alternatives.
Strengths	• **Greater statistical power** and **parsimony** (fewer parameters) for ordered data. • Provides a single, unified odds ratio for each predictor, offering a clear "shift" interpretation. • Theoretically optimal when the outcome's order is meaningful.	• **Flexibility** for modeling outcomes with no natural order. • Makes no assumptions about the ordering of categories. • Well-established for discrete choice modeling.
Weaknesses / Risks	• **Biased estimates** and **loss of interpretability** if the PO assumption is violated. • Requires rigorous diagnostic testing (e.g., Brant test, score tests) which is often overlooked.	• **Less powerful and less parsimonious** for ordered data; fails to leverage the ordinal structure. • Can lead to fragmented, hard-to-interpret results for ordered outcomes. • The IIA assumption is often untested in non-economics fields.
Recommended Action	• **Test the PO assumption.** • If PO holds, use OLR. • If PO is violated for some predictors, use **Partial Proportional Odds (PPO)** or **Generalized Ordered Logit (GOL)** models.	• **Confirm the outcome is truly nominal. Test the IIA assumption** (e.g., Hausman-McFadden test). • If IIA is violated, consider **Nested Logit** or **Mixed Logit** models.

5.0 Conclusion

The 2015–2025 empirical record supports a simple but frequently neglected rule: if the outcome is ordered, start ordinal. OLR generally dominates MLR in parsimony, power, and interpretability; where the proportional odds (PO) assumption is compromised, partial proportional odds (PPO) and generalised ordinal logistic (GOL) models offer an empirically supported middle ground that preserves ordinality while targeting non-parallel effects. MLR remains essential for nominal outcomes, yet its independence of irrelevant alternatives (IIA) assumption demands rigorous testing and, where necessary, relaxation. Across domains, the persistent disparities in practice arise less from the absence of appropriate tools than from the underuse of diagnostics, insufficient reporting and the tendency to default to familiar models. Theoretically, this study advances statistical modelling by clarifying the boundaries between OLR and MLR, thereby reinforcing the importance of assumption-driven rather than convenience-driven model selection. Practically, it provides researchers with a structured framework for choosing models that maximise interpretability and analytical power while minimising misapplications. From a policy perspective, the findings call for strengthened reporting guidelines and methodological standards across disciplines to ensure transparency, reproducibility and more context-sensitive statistical practices.

6.0 Recommendations

Researchers should align model choice with the true structure of the outcome variable. For ordered categories, OLR should be the primary option, accompanied by rigorous testing of the proportional-odds (PO) assumption using both global and variable-specific diagnostics. Where meaningful PO violations occur, partial proportional-odds (PPO) or generalised ordered logit (GOL) models should be preferred over reverting to MLR, as they preserve ordinal integrity while adding necessary flexibility. For nominal outcomes, MLR remains suitable, but the independence of irrelevant alternatives (IIA) assumption must be tested and, if violated, addressed through models such as mixed or nested logit. Across all applications, researchers should improve reporting transparency by clearly justifying model choice, documenting diagnostic results, and presenting interpretable effect sizes. Finally, cross-disciplinary training should emphasise flexible ordinal modelling techniques, assumption checks and ordinal-specific predictive evaluation to enhance both methodological rigour and practical utility**.**

7.0 Acknowledgment

The author gratefully acknowledges the contributions of all researchers whose works were reviewed in this study, as well as the constructive feedback from colleagues that enriched the analysis.

8.0 References

Agresti, A. (2019). *An Introduction to Categorical Data Analysis* (3rd ed.). Wiley.

Alogna, A., *et al.* (2021). Ordinal outcomes in stroke trials. *International Journal of Stroke, 16*(7), 780–789.

Anastasopoulos, P. Ch., Mannering, F., Shankar, V., & Haddock, J. (2018). Random parameters in crash-injury severity. *Analytic Methods in Accident Research, 18*, 1–22.

Austin, P. C., & Fine, J. P. (2017). Practical recommendations for reporting analyses of ordinal outcomes. *Statistical Methods in Medical Research, 26*(6), 2708–2732.

Austin, P. C., Brunner, J., Hogue, C., & Fine, J. P. (2021). Ordinal outcomes and PO in clinical trials. *Statistics in Medicine, 40*(29), 6483–6502.

Bansal, P., Kockelman, K., & Singh, A. (2016). Choice modeling with IIA relaxations for automated vehicles. *Transportation Research Part A, 92*, 1–17.

Behnood, A., & Mannering, F. (2017). Determinants of injury severities. *Analytic Methods in Accident Research, 16*, 1–22.

Camarda, C. G., & Serra, A. (2023). Applied multinomial and ordinal regression models in public health research. *Statistical Methods in Medical Research, 32*(2), 345–362.

Chalmers, R. P. (2018). Ordinal models for Likert data in psychology. *Psychological Methods, 23*(2), 227–243.

Daly, A., Hess, S., & Train, K. (2012). Assuring finite moments for mixed logit (widely cited, applied 2015+). *Transportation, 39*, 19–31.

Das, S., & Abdel-Aty, M. (2017). Severity in work zones using ordered models. *Accident Analysis & Prevention, 98*, 30–39.

Dong, B., Huang, H., Lee, J., & Gao, M. (2020). Partial proportional odds for crash severity. *Accident Analysis & Prevention, 144*, 105626.

Eboli, L., &Mazzulla, G. (2015). Satisfaction with public transport via ordered logit. *Transport Policy, 42*, 26–37.

Edlinger, M., Schulz, E., & Heinze, G. (2021). Prediction with ordinal endpoints. *Statistics in Medicine, 40*(12), 2781–2797.

El-Basyouny, K., Sayed, T., & Sadeghi, M. (2018). Ordinal models in road safety. *Accident Analysis & Prevention, 111*, 1–10.

Eluru, N. (2015). A note on generalized ordered outcome models. *Journal of Choice Modelling, 14*, 1–10.

Elvik, R. (2018). Meta-analysis of severity outcomes and modeling choices. *Accident Analysis & Prevention, 110*, 1–12.

Gambarota, F., Gallo, A., &Naldi, M. (2024). Ordinal regression made easy. *Behavior Research Methods, 56*(2), 547–566.

Gao, M., *et al.* (2018). Comparing OLR, PPO, and MLR in crash data. *Journal of*

Transportation Safety & Security, 10(2), 160–177.

Gao, S., & Zheng, J. (2020). Mixed multinomial logit for mode choice. *Transportation Research Part A, 137*, 1–15.

Haines, N., *et al.* (2020). Ordinal regression in psychological scales. *Behavior Research Methods, 52*(6), 2365–2385.

Hamaker, E. L., &Muthén, B. (2020). Multilevel ordinal models in psychology. *Multivariate Behavioral Research, 55*(6), 859–880.

Harrell, F. E. (2022). Assessing the proportional odds assumption. *Biometrical Journal, 64*(6), 1062–1075.

Hemmingsen, S. N., *et al.* (2022). Ordinal endpoints in COVID-19 trials. *Clinical Trials, 19*(6), 646–658.

Hensher, D. A. (2018). Choice analysis with random parameters. *Transportation Research Part A, 116*, 593–606.

Hilbe, J. M. (2021). *Logistic Regression Models*. Chapman and Hall/CRC. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2019). *Applied Logistic Regression* (4th ed.). Wiley.

Huang, H., Bastani, H., &Jin, J. (2021). Random-parameter ordinal models in safety. *Accident Analysis & Prevention, 156*, 106139.

Keita, S., & Ganesan, R. (2022). Testing IIA in multinomial contexts. *Journal of Applied Statistics, 49*(6), 1453–1470.

Kizilaslan, A., &Dumenci, L. (2021). Ordinal mixed models for Likert data. *Educational and Psychological Measurement, 81*(2), 239–260.

Koyama, T., & Chen, Y. (2017). Ordinal regression in oncology PROs.*Journal of Biopharmaceutical Statistics, 27*(6), 984–997.

Kuss, O. (2015). The PO model for ordinal outcomes: a review. *Statistical Methods in Medical Research, 24*(6), 970–989.

LaValley, M. P. (2018). Logistic regression models for ordinal response variables. *Statistics in Medicine, 37*(15), 2240–2251.

Li, X., &Alvi, M. A. (2022). Partial proportional odds vs multinomial on crash severities. *Journal of Safety Research, 80*, 94–104.

Liang, J., Wu, X., & Zou, G. (2020). MLR vs OLR in R: a demonstration. *Journal of Statistical Software, 95*(2), 1–23.

Liddell, T. M., &Kruschke, J. K. (2018). Analyzing ordinal data correctly. *Journal of Experimental Social Psychology, 79*, 328–348.

Liu, I., &Agresti, A. (2019). Analysis of ordinal categorical data with non-PO. *Statistics in Medicine, 38*(30), 5671–5687.

Liu, I., Chen, L., & Huang, H. (2020). Comparative analysis of ordinal and multinomial logistic regression models in health research. *BMC Medical Research Methodology, 20*(1), 278.

Long, J. S., & Freese, J. (2014). *Regression Models for Categorical Dependent Variables Using Stata* (3rd ed.). Stata Press.

Long, Y., *et al.* (2025). Audit of ordinal analysis in neurology. *Neurology Methods, 2*(1), 45–59.

Lord, D., &Geedipally, S. (2019). Bayesian ordinal models in safety. *Accident Analysis & Prevention, 123*, 365–377.

Mannering, F., & Bhat, C. (2014; widely applied afterward). Analytic methods in safety extensions 2015+. *Analytic Methods in Accident Research, 1*, 1–22.

Månsson, K., &Shukur, G. (2018). Modeling Likert data: ordinal vs nominal. *Communications in Statistics Simulation and Computation, 47*(8), 2248–2266.

Matimbwa H and Masue O.S (2020). The Influence of Organization Factors on Human Resource Information System Effectiveness in the Tanzanian Local Government Authorities. *ICTACT Journal on Management Studies August 2020, Volume 06 issue 03*

Matimbwa H, Masue O and Shillingi (2020). Technological Features and Effectiveness of Human Resources Information System in Tanzanian Local Government Authorities. *American Journal of*

Operations Management and Information Systems Vol 5 issue 3
Matimbwa H, Masue O and Shillingi (2021). User Characteristics and Effectiveness of Human Resources Information System (HRIS) in the Tanzanian Local Government Authorities. *Journal of Cooperative and Business Studies Vol 6 issue 1*
Matimbwa H, Shillingi, V. and Masue, O.S (2021). Effectiveness of Human Resources Information System in Tanzanian Local Government Authorities. Do Technological, User and Organizational Attributes matter? *Rural Planning Journal* vol 23(1)
Matimbwa, Hand Ochumbo A (2018). Automated Teller Machine and Customer Satisfaction, in CRDB Iringa Tanzania*. Journal of Business Management andEconomic ResearchVol.2, Issue.3, 2018. ISSN 2602-3385)*
McKinley, L., *et al.* (2022). Ordinal endpoints in rheumatology trials. *Arthritis Care & Research, 74*(9), 1512–1520.
Mohammadi, M., *et al.* (2019). Ordered vs multinomial in injury studies. *Journal of Safety Research, 68*, 1–9.
Mokhtarian, P. L. (2016). Revisiting IIA in discrete choice. *Transport Reviews, 36*(3), 1–20.
Nunes, V., *et al.* (2020). Ordinal models for PROs in oncology. *Quality of Life Research, 29*(10), 2561–2574.
O'Connell, A. A., & Liu, X. (2016). Model diagnostics for ordinal regression. *Practical Assessment, Research & Evaluation, 21*(6), 1–16.
Papakostas, N., *et al.* (2022). Injury severity with generalized ordered logit. *Safety Science, 147*, 105596.
Paul, S., Deb, P., & Roy, S. (2021). Evaluating proportional odds assumption in ordinal logistic regression: Practices and pitfalls. *Journal of Statistical Computation and Simulation, 91*(9), 1842–1859.
Quddus, M., & Wang, X. (2016). Severity modeling and link choice. *Transportmetrica A, 12*(8), 726–748.
Savolainen, P. T., Mannering, F., Lord, D., & Quddus, M. (2020). The state of the practice in safety models. *Analytic Methods in Accident Research, 27*, 100–113.
Selman, C. J., et al. (2024). Reporting of ordinal outcomes in RCTs. *Trials, 25*(1), 1–12.
Song, L., Abdel-Aty, M., & Wang, L. (2020). Latent class PPO for crash severity. *Accident Analysis & Prevention, 142*, 105556.
Sproston, K., et al. (2023). Ordinal links for Likert-scale surveys. *Survey Research Methods, 17*(2), 123–141.
Sun, Y., Chen, W., & Wang, J. (2023). On the selection between multinomial and ordinal logistic models: Simulation and application. *Journal of Applied Statistics, 50*(6), 1234–1251.
Tolles, J., &Meurer, W. J. (2020). Logistic regression: Relating patient characteristics to outcomes. *JAMA, 324*(5), 507–508.
Train, K. (2016). *Discrete Choice Methods with Simulation* (2nd ed.). Cambridge University Press.
Tutz, G., &Gertheiss, J. (2016). Regularized regression for categorical data. *Statistical Modelling, 16*(3), 161–200.
Tutz, G., &Schauberger, G. (2015). Ordinal regression with non-proportional odds. *Statistical Modelling, 15*(6), 433–452.
Uddin, M., *et al.* (2023). Evaluating PO models in COVID-19 trials. *BMC Medical Research Methodology, 23*(1), 1–14.
van Smeden, M., *et al.* (2019). Prediction with ordinal outcomes pitfalls. *Diagnostic and Prognostic Research, 3*(1), 3.
Wang, C., &Chiou, Y.-C. (2016). Severity modeling with ordered links. *Accident Analysis & Prevention, 87*, 1–10.
Wang, H., *et al.* (2018). Ordinal modeling of patient satisfaction. *Health Services Research, 53*(3), 1790–1807.
Wang, X., & Chen, F. (2019). Relaxing IIA in travel mode choice. *Transportation Research Part A, 126*, 190–204.
Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The*

Journal of Mathematical Sociology, 40(1), 7–20.

Williams, R. (2016). Understanding generalized ordered logit. *The Journal of Mathematical Sociology, 40*(1), 7–20.

Xie, K., Wang, X., Huang, H., & Chen, F. (2019). Crash injury severity with PPO. *Accident Analysis & Prevention, 124*, 50–63.

Zeng, Q., Huang, H., & Pei, X. (2016). Injury severity ordinal modeling. *Accident Analysis & Prevention, 95*, 187–199.

Zhang, L., *et al.* (2021). Ordinal vs multinomial in-patient experience. *BMJ Open, 11*(10), e050123

Zumbo, B. D. (2015). Ordinal vs interval treatment of Likert scales. *Educational and Psychological Measurement, 75*(4), 679–70