



Aalborg Universitet

AALBORG UNIVERSITY  
DENMARK

## Misuse of multinomial logistic regression in stroke related health research

*A systematic review of methodology*

Fitzhugh, Nicholas; Rasmussen, Line Ryberg; Simoni, Amalie Helme; Valentin, Jan Brink

*Published in:*  
European Journal of Neuroscience

*DOI (link to publication from Publisher):*  
[10.1111/ejn.16084](https://doi.org/10.1111/ejn.16084)

*Creative Commons License*  
CC BY 4.0

*Publication date:*  
2023

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*  
Fitzhugh, N., Rasmussen, L. R., Simoni, A. H., & Valentin, J. B. (2023). Misuse of multinomial logistic regression in stroke related health research: A systematic review of methodology. *European Journal of Neuroscience*, 58(4), 3116-3131. <https://doi.org/10.1111/ejn.16084>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# WILEY

## Frontiers in Flow Cytometry™

24 hour Virtual Event

September 13th, 2023

Frontiers in Flow Cytometry™ is for researchers across the globe looking for an opportunity to share and learn about current developments in flow cytometry. This 24 hour virtual event will feature keynote presentations by industry colleagues, webinars, demos, live networking opportunities and more.

### Key topics include:

- Spectral and conventional flow cytometry
- Immunophenotyping and Standardization
- Panel design and optimization
- Cancer Biology and Auto-immune Diseases
- Infectious diseases
- Advances in flow cytometry technology

[Register Now](#)

This event is sponsored by **ThermoFisher**  
SCIENTIFIC

# Misuse of multinomial logistic regression in stroke related health research: A systematic review of methodology

Nicholas Fitzhugh<sup>1,2</sup>  | Line Ryberg Rasmussen<sup>1</sup>  | Amalie Helme Simoni<sup>1</sup> | Jan Brink Valentin<sup>1</sup> 

<sup>1</sup>Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Gistrup, Denmark

<sup>2</sup>Danish Health Technology Council (Behandlingsrådet), Aalborg, Denmark

## Correspondence

Nicholas Fitzhugh, Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Selma Lagerlöfs Vej 249, Gistrup 9260, Denmark.

Email: [nfi@dcm.aau.dk](mailto:nfi@dcm.aau.dk)

Edited by: Guillaume Rousselet

## Abstract

Multinomial logistic regression (MLR) is often used to model the association between a nominal outcome variable and one or more covariates. The results of MLR are interpreted as relative risk ratios (RRR) and warrant a more coherent interpretation than ordinary logistic regression. Some authors compare the results of MLR to ordinal logistic regression (OLR), irrespective of the fact that these estimate different quantities. We aim to investigate the time trends in the use and misuse of MLR in studies including stroke patients, specifically the extent to which (1) the results are denoted as anything other than RRR, (2) comparisons are made of results with results of OLR and (3) results have been interpreted coherently. Secondly, we examine the use of model validation techniques in studies with predictive aims. We searched EMBASE and PubMed for articles using MLR on populations of stroke patients. Identified studies were screened, and information pertaining to our aims was extracted. A total of 285 articles were identified through a systematic literature search, and 68 of these were included in the review. Of these, 60 articles (88%) did not denote exponentiated coefficients of MLR as relative risk ratios but rather some other measure. Additionally, 63 articles (93%) interpreted the results of MLR in a non-coherent manner. Two articles attempted to compare MLR results with those of OLR. Nine studies attempted to use MLR for predictive means, and three used relevant validation techniques. From these findings, it is clear that the interpretation of MLR is often suboptimal.

## KEYWORDS

interpretation, modelling, odds ratio, relative risk ratio, risk ratio

## 1 | INTRODUCTION

Stroke research has increased rapidly over the past decades, leading to improved therapies and treatment guidelines (Goyal et al., 2016; Grefkes & Fink, 2020;

**Abbreviations:** IPTWs, inverse probability of treatment weights; MLR, multinomial logistic regression; OLR, ordinal logistic regression; OR, odds ratio; RR, risk ratio; RRR, relative risk ratios.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *European Journal of Neuroscience* published by Federation of European Neuroscience Societies and John Wiley & Sons Ltd.

Rasmussen et al., 2020; Schöenberger et al., 2019; Thomalla et al., 2018). However, if evidence-based guidelines are to improve patients' outcomes, it is crucial for such research that the statistical and epidemiological methodologies are applied properly and results are interpreted correctly. Previous reviews investigating the methodological conduct of prognosis studies conclude that the research quality within some medical fields is generally poor (Collins et al., 2011, 2014; Mallett et al., 2010). In the field of stroke research, outcomes are sometimes nominal or ordinal. For such outcome types, multinomial logistic regression (MLR) may be used, and ordinal logistic regression (OLR) may be used when the outcome is ordinal. MLR is a standard statistical model used for outcomes of nominal type and is useful for prediction, where the goal is prognosis or diagnosis (Covert et al., 2020; Uijl et al., 2020), estimating adjusted predicted probabilities risks (Wolfe et al., 1999), discrete choice models (Agresti, 2013) and can be used for the inverse probability of treatment weights (IPTWs) with multiple exposure groups (Boesgaard Graversen et al., 2021). However, when used to estimate the effect of exposure, the interpretation of the effect estimates can have implications, as the model parameters represent log relative risk ratios (RRR). Relative risk ratios have a more complex interpretation than odds ratios (ORs) because they are a doubly relative measure, though ORs themselves are not as easily interpretable as risk ratios (RR). While some sources reflect the correct use and interpretation (Duchon, 2014), critical assessment of study results and conclusions is difficult due to the complexity of the model. This leaves uncertainty as to how often MLR is used, with the interpretation of coefficient estimates being flawed. Moreover, healthcare professionals may have uncritically applied study results in clinical guidelines, which potentially leads to treatment choice and patient diagnosis, supported by flawed conclusions in research using MLR.

An alternative to MLR is the so-called OLR model, which can be used when the outcome variable is ordinal and can assume more than two values (Kleinbaum & Klein, 2010). OLR can be used to estimate ORs and relies on the so-called proportional odds assumption (Gelman & Hill, 2006; Kleinbaum & Klein, 2010), an assumption not used for MLR. Another misuse of MLR has been seen when it is used to validate the proportional odds assumption (Topriceanu et al., 2021). Furthermore, some authors compare the results of MLR and OLR, irrespective of the fact that these estimate different quantities.

In the current study, we seek to understand the reach of misconduct when applying MLR to comparative studies in stroke-related medical sciences. Hence, we aim to investigate the time trends in the use and misuse of MLR

in studies that include patients with stroke. Specifically, we aim to assess to what extent the included studies:

1. Denote the exponentiated coefficients as RRR or equivalent?
2. Compare the results of MLR to those of OLR?
3. Interpret the results of the MLR model coherently, such that the RRRs for the same exposure groups are interpreted together?

Additionally, we are interested in the use of MLR for prediction and examine whether studies with a predictive aim use any form of data split, external validation or similar to assess the models' predictive ability (Altman et al., 2009; Moons et al., 2009; Steyerberg & Vergouwe, 2014).

## 2 | BACKGROUND

Logistic regression models have been widely used in the medical sciences and still are. Using logistic regression, it is possible to model the effect of one or more binary and continuous explanatory variables on a binary outcome variable. Logistic regression offers ease of interpretation of the estimated effect, as both risk and ORs can be derived from the estimated coefficients. Logistic regression can be extended to the case where the outcome variable is nominal with more than two levels. This model is called the MLR model, and the extension results in a more complex model that has a wider application, but this comes at the expense of interpretability when comparing outcomes between exposure groups. The MLR model can, under some assumptions, be used to estimate the adjusted predicted probabilities of outcomes in observational studies where covariate adjustment is necessary; however, the model coefficients do not necessarily quantify the effect of an exposure on the outcome.

Suppose  $Y$  is a nominal outcome variable with  $k$  levels, then the MLR model can be expressed as,

$$\log\left(\frac{P(Y=i)}{P(Y=0)}\right) = \beta_{i,0} + \beta_{i,1}x_1 + \dots + \beta_{i,m}x_m$$

for  $i = 1, \dots, k-1$ , where  $x_1, \dots, x_m$  are explanatory variables (Gelman & Hill, 2006; Kleinbaum & Klein, 2010). Note that coefficients are specific to the  $i$ -th outcome level, and thus  $(k-1)$  different sets of  $m+1$  regression coefficients are estimated in the model. When using MLR, one must specify a reference level, which is  $Y=0$  above, but this choice is entirely arbitrary. In practice, any category may be chosen as the reference category, and the coefficients change accordingly.

## 2.1 | Interpretation of the coefficients

In the commonly used logistic regression model, where the outcome variable is binary, the exponentiated coefficients are interpreted easily as ORs. For MLR, this interpretation has been reused in some studies (Matsukawa et al., 2013; Rist et al., 2013; Zinkstok et al., 2014). However, the exponentiated coefficients cannot be interpreted as ordinary ORs but should rather be interpreted as a ratio of relative risks. This can be seen in the case where  $m + 1$ , because

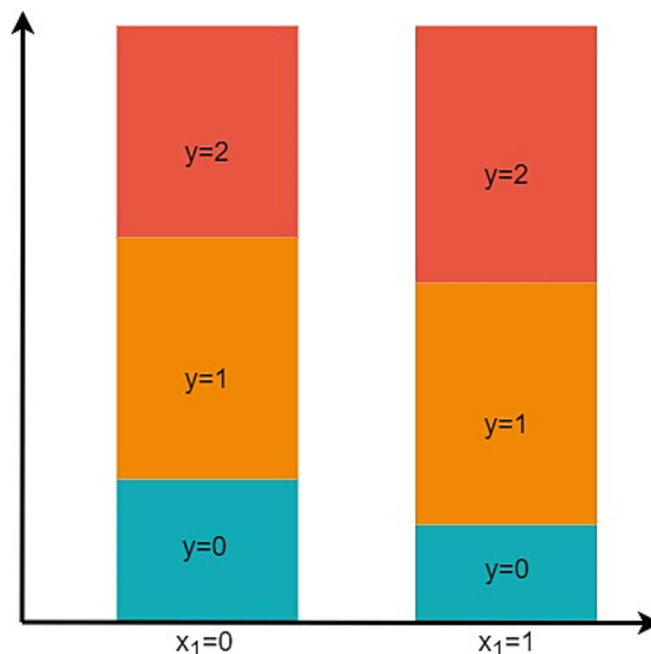
$$\frac{\left(\frac{P(Y=i|x_1=1)}{P(Y=0|x_1=1)}\right)}{\left(\frac{P(Y=i|x_1=0)}{P(Y=0|x_1=0)}\right)} = \frac{e^{\beta_{i,0} + \beta_{i,1} \cdot 1}}{e^{\beta_{i,0} + \beta_{i,1} \cdot 0}} = e^{\beta_{i,1}}$$

where the left-hand side is a ratio of the relative risks of being in the outcome category  $i$  versus in the reference category. As such, the exponentiated coefficients should be interpreted and referred to as RRRs. Denoting exponentiated coefficients as either OR, risk ratios (RR) or any other measure is therefore misleading and may lead to a faulty inference of effects. We instigate the extent of this denotation error in answering our first question.

Additionally, the complexity of MLR raises another issue connected to interpretation. A RRR above one may be interpreted as increased risk in the exposure group compared with the unexposed; however, a RRR above one might as well be caused by a reduced risk in the reference group as a result of an increased risk in one of the competing outcome groups. Most likely, the cause of the RRR above one is a combination of both scenarios, which warrants a coherent inherent interpretation of the RRRs of all outcome categories; that is, all coefficients of a variable must be interpreted together. The complexity of interpretation increases with the number of outcome groups, and even a RRR equal to one cannot uniquely be interpreted as equal risk between exposure groups. An example of one of these scenarios can be seen in Figure 1, where the risk of  $Y=1$  is the same in each exposure group but the RRR is above one.

This is a result of a reduction in risk for  $Y=0$  and an increase in risk for  $Y=2$ , respectively, between exposure groups, and thus any interpretation of the RRR for  $Y=1$  versus  $Y=0$  must be coherent with the interpretation of the RRR for  $Y=2$  versus  $Y=0$ . The histogram can be extended to arbitrarily many outcome levels, and interpretation would likewise be more complex. Moreover, any coherent interpretation of the results of MLR must recognize the doubly relative nature of RRR. In the case where  $Y$  is binary, MLR reduces to OLR.

We instigate the coherency of interpretation in answering our third question.



**FIGURE 1** This stacked histogram shows a scenario where the overall risk of  $Y = 1$  is the same in both the unexposed and exposed groups ( $x_1$  denotes the exposure group). Interpreting the exponentiated coefficients for  $x_1$  separately would lead researchers to infer an increased risk for  $Y = 1$  in the exposed group, even though the risk is the same in both groups. Interpreting coefficients for  $x_1$  for both  $Y = 1$  and  $Y = 2$  relative to the reference group together, one might be able to deduce that the decrease in RRR for  $Y = 1$  is caused by an increased risk for  $Y = 2$ .

## 2.2 | Predicted probabilities

MLR may be used to derive both adjusted predicted probabilities and conditional predicted probabilities. While conditional predicted probabilities are useful for estimating risk differences within certain strata, adjusted predicted probabilities are useful for estimating marginal risks for each exposure group and outcome class. Deriving predicted probabilities without adjusting for any covariates is equivalent to making a  $J \times K$  contingency table with  $J$  outcome classes and  $K$  exposure groups. We have deemed the use of MLR for estimating adjusted/conditional predicted probabilities legitimate but do not directly investigate the extent of this usage.

## 2.3 | OLR

When the outcome variable is ordinal, instead of MLR, one might consider the OLR model, from which ORs can be estimated. OLR can be used when an outcome variable is ordinal and has more than two levels. Suppose  $Y$  is an ordinal outcome variable with  $k$  levels. OLR can be expressed as:

$$\log\left(\frac{P(Y \geq i)}{P(Y < i)}\right) = \phi_i + \beta_1 x_1 + \dots + \beta_m x_m,$$

where  $i$  is some value of  $Y$  we will call a cutoff value and  $x_1, \dots, x_m$  are explanatory values (Kleinbaum & Klein, 2010). Contrary to MLR, OLR estimates the association between covariates and the log odds of  $Y$  being less than or equal to some cutoff  $i$ . Furthermore, note that the coefficients, except the intercept, are not dependent on the cutoff being considered. As such, the coefficients express a change in log odds for a unit change in a given covariate, irrespective of the cutoff value being considered, implying the assumption that this change is the same for each cutoff value. This assumption is called the proportional odds assumption, and it is central to OLR. Inspection of the proportional odds assumption can be done in various ways, including comparison of ORs for all possible cut-points and a Brant test (Brant, 1990). Inspection of the proportional odds assumption has also been attempted by comparing the ORs of the OLR with the RRR of the MLR (Topriceanu et al., 2021). However, such comparisons lead to faulty conclusions, as comparing coefficients that are not representations of the same measure of association is not meaningful.

An alternate model to OLR is a generalization called the generalized OLR model (GOLR), which may be used. This model does not rely on the proportional odds assumption and thus estimates multiple sets of coefficients, similar to MLR, though a comparison to MLR is still ill-advised. However, the increased complexity of GOLR complicates the interpretation of inferred effects.

## 2.4 | Illustrative cases

To further illustrate an important point of this review, namely, that the parameters of MLR and OLR cannot meaningfully be compared, we propose two examples. The first example uses registry data from the Danish National Patient Registries (usage approved by the Danish Data Protection Agency), and the second example involves a hypothetical situation. The contingency table for both examples is presented in Table 1. Analyses are performed in R (R Core Team, 2018).

Case 1: Association between mental disorders and stroke or death.

For the first example, we consider a cohort of 22,719 patients who have turned 60 years of age on or prior to 2009 and who have been diagnosed with atrial fibrillation (AF) (ICD-10: I48) at some point after turning 60. We then compare patients who have been diagnosed with severe mental disorders (ICD-10: F20, F22, F25, F30 or F31) with those who have no history of mental disorders on death and stroke (ICD-10: I60, I61, I63 or I64) within 2 years after they have been diagnosed with AF.

The proportions of death and stroke are summarized for this cohort using a contingency table (Table 1), from which we may compute the RRRs directly. Thus, we find  $RRR_{Stroke\ vs\ Healthy} = \frac{74.1\%/2.1\%}{70.3\%/1.8\%} = 0.90$  and  $RRR_{Dead\ vs\ Healthy} = \frac{24.8\%/2.1\%}{27.9\%/1.8\%} = 0.76$ . Equivalently, RRRs are found when utilizing a MLR model from the nnet package in R (Table 1). The  $RRR_{Stroke\ vs\ Healthy}$  of 0.9 is counterintuitive because patients with severe mental illness have an elevated risk of stroke but a reduced risk of

TABLE 1 Contingency tables and model parameter estimates for both cases.

Contingency table				Multinomial logistic regression	Generalized ordinal logistic regression	Ordinal logistic regression
Example from the Danish National Patient Registries		No mental disorder	Severe mental disorder	RRR	OR	OR
	Healthy (y = 0) (ref.)	1.8%	2.1%	-	-	-
	Stroke (y = 1)	70.3%	74.1%	.91	.86	.81
	Dead (y = 2)	27.9%	24.8%	.73	.81	.81
Hypothetical situation		Non-smoker	Smoker	RRR	OR	OR
	Healthy (y = 0) (ref.)	36%	11%	-	-	-
	Cancer (y = 1)	25%	67%	8.77	4.55	1.17
	Dead (y = 2)	39%	22%	1.85	.44	1.17

death. However, these results arise because of the larger group of patients that remain stroke-free and survive compared with the reference group.

In this example, the outcome variable is the patients' health state, which can assume progressively worse outcomes classes. As such, it can be considered as an ordinal variable, and we may model the associations between the outcome and mental health disorders with an OLR model from the ordinal package in R. Using the OLR model yields a single parameter estimate for the exposure, which is assumed to be equivalent for all cutoffs of the outcome variable (Table 1). This is often called the proportional odds assumption. If, instead, we use a GOLR model also from the ordinal package in R, which does not assume proportional odds, we can estimate the same number of parameters as with MLR (Table 1). Equivalent results are found by considering the contingency table, and doing so yields  $OR_{i='Stroke'} = \frac{(24.8\%+74.1\%)/2.1\%}{(27.9\%+70.3\%)/1.8\%} = 0.86$  and  $OR_{i='Dead'} = \frac{24.8\%/(2.1\%+74.1\%)}{27.9\%/(1.8\%+70.3\%)} = 0.84$  for the cutoffs of 'stroke' and 'dead', respectively. These estimates are identical to those of the GOLR model, with some rounding error. From the results of the GOLR, we see that the proportional odds assumption is likely to hold in this example, while the results of the MLR preserve the direction but not the magnitude. None of the analyses captures the elevated risk of stroke for patients with severe mental illness. This is only seen by considering the contingency table.

#### Case 2: A hypothetical example

In the second example, we see an elevated risk of cancer and a decreased risk of death in a group of smokers compared with non-smokers. Again, the RRR estimates of the MLR do not capture the decreased risk of death because both RRRs are above 1.0 (Table 1). Moreover, the example reveals that the RRR estimates of the MLR are not comparable to the OR estimates of the GOLR and OLR because the estimates differ in both magnitude and direction. Both examples highlight the difficulties in interpreting RRR. However, we stress that the MLR can be used to derive an adjusted contingency table using predicted probabilities.

## 3 | METHODS

We identified articles concerning the use of MLR in analysis, with stroke patients as the population, through a systematic literature search. The databases PubMed and EMBASE were initially searched on 29 August 2019, and a final search was conducted on 29 September 2021, using the search string described in Appendix A. Articles were included regardless of publication year.

### 3.1 | Inclusion/exclusion criteria

Initially, we included all articles that applied MLR to a population, which included stroke patients as either exposure or outcome. Moreover, studies that aimed at investigating or in any way reflected on the associations represented by the RRRs of MLR were subsequently included. Thus, prediction (prognosis or diagnosis) studies where the parameters of the MLR are not referred to as associations between an independent variable and the risk/odds of a response were excluded. Likewise, studies where the MLR was only used for IPTW were excluded. Studies deriving only adjusted/unadjusted predicted probabilities where no comparisons between exposure groups are made were also excluded. Studies in another language than Danish or English or studies with an unclear methodology were also excluded.

### 3.2 | Data extraction, analysis, and reporting

The titles and abstracts of all articles identified by the search string (Appendix A) were screened by two of the authors (LRR and AHS) to exclude articles not pertaining to multinomial logistic models or populations, including stroke patients. Studies not excluded after the screening of titles and abstracts were screened again on a full-text basis by the two other authors (JBV, NF). Lastly, some articles were identified by hand. These were screened in the same manner as the articles found in the systematic search (Figure 2).

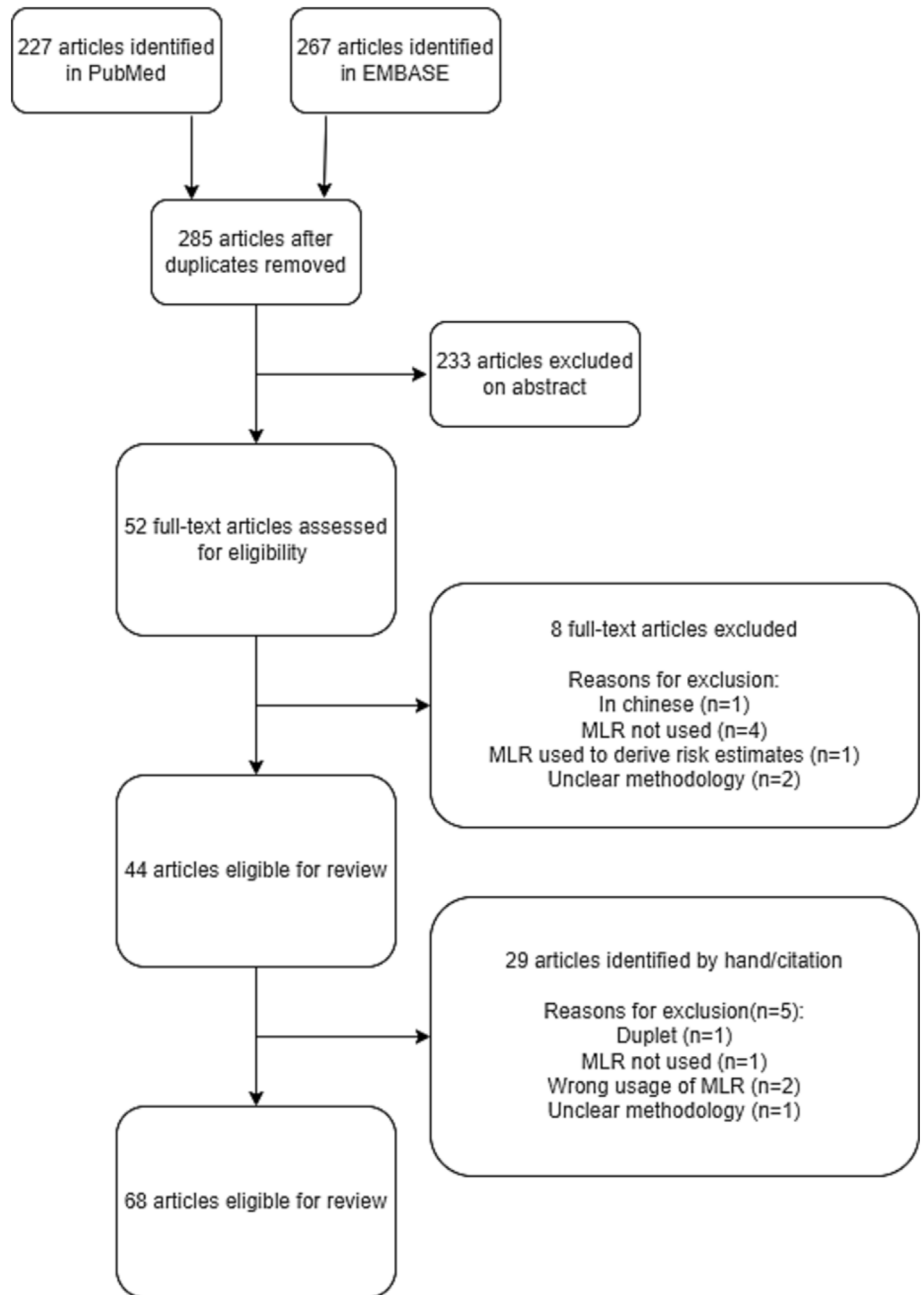
We have reported our systematic review in accordance with the PRISMA 2020 guidelines, except for the items relating to meta-analysis, which are not relevant in the context of this review.

The included articles were carefully examined for information about whether they:

1. Denoted exponentiated coefficients as anything other than RRR.
2. Compared MLR results to results of OLR.
3. Interpreted MLR results in a coherent manner.

Additionally, studies with predictive aims were reviewed to ascertain whether any form of data split or similar was used to assess the models' predictive ability. Using a random data split is in general ill-advised, as this method is inefficient and inferior to cross-validation and bootstrapping methods. We do not consider the use of apparent validation, that is, where a prediction model is validated on the same data as it has been trained, as a valid validation method.

**FIGURE 2** Flow diagram of the screening process for studies identified in the systematic literature search.



Furthermore, the publication year for each article was determined in order to assess the time trends in the use of MLR. The information extracted from each article is presented in Appendix B.

## 4 | RESULTS

The search string resulted in 227 articles in PubMed and 267 articles in EMBASE. After removing duplicates, our search yielded 285 articles (Figure 2).

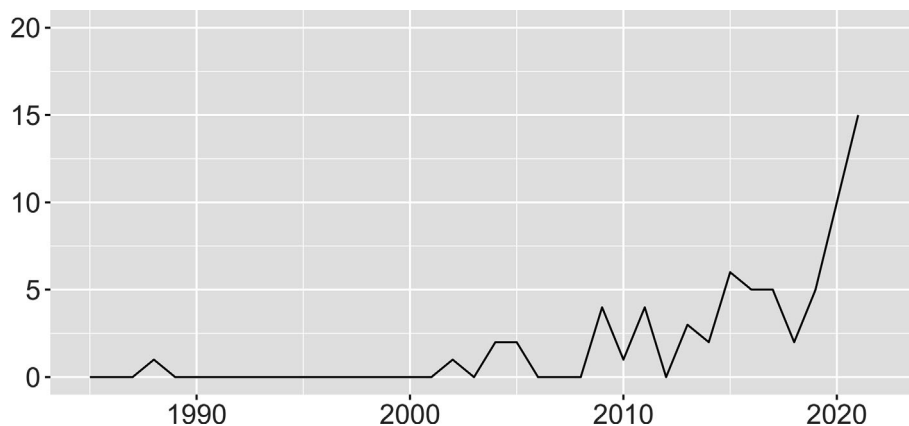
Following the screening of titles and abstracts, 52 articles were included for full-text screening. A further 28 articles were found by hand-searching reference lists.

In total, 68 studies met our inclusion criteria and were eligible for inclusion in the review. The included studies were published between 1988 and 2021, and 58 studies (85%) were published between January 2010 and September 2021 (Figure 3).

### 4.1 | Denoting the coefficients

As suspected, coefficients of MLR were seldom referred to as RRRs, with 60 articles (88%) referring to them as some other measure of risk (Table 2). OR interpretation was by far the most popular way to interpret coefficients, with 50 articles (74%) interpreting the exponentiated





**FIGURE 3** This graph shows the publication year of the included studies. As can be seen on the graph, the use of multinomial logistic regression (MLR) seems to be more frequent starting from 2012, with a large spike in usage in recent years.

**TABLE 2** Proportion of articles identified for each objective.

Domain	Number of studies (%) [ref]
<b>Primary objectives</b>	
Articles denoting exponentiated coefficients as other than RRR	60 (88%) (Anagnostou et al., 2020; Béjot et al., 2017; Buntin et al., 2005, 2009; Butcher et al., 2010; Choi et al., 2017, 2020; Chow et al., 2017; Clarke et al., 2011; Covert et al., 2020; Debette et al., 2011; Ellis et al., 2017; Fitzpatrick et al., 2020; Gocan et al., 2020; Gong et al., 2021; Hamaguchi et al., 2020; Hendrickx et al., 2021; Howrey et al., 2011; Huo et al., 2021; Jacob et al., 2021; Jaffre et al., 2014; Johnson et al., 2015; Jung et al., 2016; Kang et al., 2018; Kashani et al., 2021; Kim et al., 2004; Krajcoviechova et al., 2015; Kuohn et al., 2020; Ma et al., 2021; Macintosh et al., 2021; Matsukawa et al., 2013; Meurer et al., 2009; Minen et al., 2019; Mizrahi et al., 2005; Morrisey et al., 1988; Nagai et al., 2011; Nighoghossian et al., 2016; Park et al., 2013; Piayda et al., 2021; Pinto et al., 2021; Qiu et al., 2009; Quiñones et al., 2021; Rimmele et al., 2021; Rist et al., 2013, 2016; Rodríguez-Castro et al., 2018; Schlegel, 2004; Shigematsu et al., 2015; Singh et al., 2019; Song et al., 2015, 2017; Soroush et al., 2019; Stulberg et al., 2019; Tuttolomondo et al., 2015; Uijl et al., 2020; Ye et al., 2015; Zhao et al., 2019; Zinkstok et al., 2014; Zong et al., 2016; Zongo et al., 2019)
Articles comparing MLR results to the results of OLR	2 (3%) (Johnson et al., 2015; Lipsman et al., 2009)
Articles interpreting MLR results in a coherent manner	5 (7%) (Buntin et al., 2005, 2009; Hendrickx et al., 2021; Macintosh et al., 2021; Williams et al., 2021)
<b>Secondary objective</b>	
Articles with a predictive aim using data split or similar for predictive accuracy	3 of 9 (33.33%) (Covert et al., 2020; Hamaguchi et al., 2020; Uijl et al., 2020)

coefficients either fully (Béjot et al., 2017; Butcher et al., 2010; Choi et al., 2017; Clarke et al., 2011; Debette et al., 2011; Ellis et al., 2017; Jaffre et al., 2014; Kang et al., 2018; Matsukawa et al., 2013; Meurer et al., 2009; Mizrahi et al., 2005; Park et al., 2013; Rist et al., 2013; Tuttolomondo et al., 2015; Zinkstok et al., 2014; Zong et al., 2016) or in part as OR (Kim et al., 2004; Rist et al., 2016; Ye et al., 2015). In one article, the exponentiated coefficients were denoted as ORs but not explicitly interpreted (Nighoghossian et al., 2016).

Only eight articles included in this review denoted the exponentiated coefficients as relative risk ratios (Cupini et al., 2002; Daneshvari & Johansen, 2021; Loikas

et al., 2021; Rehman et al., 2020; Williams et al., 2021) while also interpreting them as such. Additionally, two articles denoted exponentiated coefficients as RRRs but frequently misinterpreted these as RRs (Chow et al., 2017; Loikas et al., 2021).

Two articles by the same author used MLR, but recognized the difficult interpretation of MLR and subsequently abandoned interpretation in favour of deriving adjusted predicted probabilities (Buntin et al., 2005, 2009). As such, these articles did not denote the exponentiated coefficients as either RRs, ORs or RRRs, although it should be noted that the results of the MLR were reported in supplements, which seem to be unavailable.

## 4.2 | Comparisons with OLR

Out of the 68 included articles, only two articles (3%) compare results from MLR to OLR with the purpose of validating the proportional odds assumption (Johnson et al., 2015; Lipsman et al., 2009).

## 4.3 | Coherency of interpretation

As seen in Figure 1 and earlier discussions, it is important to interpret the coefficients in a coherent manner and not interpret each coefficient independently. Of the 68 included articles, only five studies (7%) interpreted the coefficients coherently (Buntin et al., 2005, 2009; Hendrickx et al., 2021; Macintosh et al., 2021; Williams et al., 2021). Additionally, three studies (4%) did not explicitly interpret MLR results for unspecified reasons, although the studies aimed to do so (Lipsman et al., 2009; Nagai et al., 2011; Uijl et al., 2020).

Studies varied in the degree of coherency in their interpretation. As mentioned previously, two studies abandoned all interpretation of results in recognition of the difficult interpretation of MLR (Buntin et al., 2005, 2009). We counted these as having a coherent interpretation. Three studies are identified as having a coherent interpretation to some extent, as they recognized the reference outcome category (Hendrickx et al., 2021; Macintosh et al., 2021; Williams et al., 2021).

The remaining 60 studies (88%) interpreted the coefficients separately.

## 4.4 | Use of data split in studies with a predictive aim

As the secondary aim of this study, we examined the use of data splits or similar methods in studies with predictive aims. Among the included studies, nine aimed to use MLR for predictive purposes. Two studies applied data splits or other methods, where Covert et al. (2020) used cross-validation and Uijl et al. (2020) validated their model using an external dataset. The remaining studies applied Akaike's information criterion (Hamaguchi et al., 2020).

## 5 | DISCUSSION

In this review, we have formulated three problem areas regarding the use of MLR based on inadequacies seen previously in studies in this area. Investigation of these problems in studies involving stroke patients identified in

a systematic literature search shows that some of these problems are common when studies use MLR to compute association (Table 2).

The denotation of exponentiated coefficients from MLR in studies involving stroke patients as anything other than RRRs appears to be exceedingly common. We speculate that much of this misinterpretation is a legacy of the usage of binary logistic regression because the clear majority of studies denoted the results as OR. This is unfortunate, as the wrongful denotation of the exponentiated coefficients may lead to a wrongful interpretation of their meaning. Specifically, denoting the exponentiated coefficients as OR may lead to interpretations such as an increase or decrease in risk or odds, which we have shown is erroneous. The wrongful denotation of the results of MLR speaks to a misunderstanding of the model itself. Authors should be aware of the results of their model of choice, specifically what measures of effect these can be meaningfully interpreted as. As such, authors must be aware that MLR estimates RRRs and also be aware of their dually relative nature.

We specified a second problem area pertaining to comparisons with OLR because examples were presented where the results of MLR were compared with those of OLR. In the included studies involving stroke patients, this problem was not common, with two studies (Table 2) making this comparison. As previously discussed, making this comparison is not meaningful, and most authors did not attempt to do so.

As our third objective, we set out to investigate how the results of MLR were interpreted. We have described how the use of MLR necessitates a coherent interpretation of the estimated coefficients, as the method is significantly more complex than, for example, binary logistic regression. Coefficients are specific to the outcome category, resulting in multiple estimated coefficients for each covariate in the model. Each coefficient is therefore influenced by the coefficients of other outcome levels, and an interpretation taking this into account is necessary. Only five studies displayed coherency in interpretation, with the clear majority of studies interpreting the coefficients separately. This is problematic as single estimates provide limited information about risks, and thus these are given too much significance in many studies. In a worst-case scenario, this may lead to the wrong conclusions being drawn.

Two studies (Buntin et al., 2005, 2009) included in our review did not display any of the previously described problems. Both studies displayed an understanding of the complexities of MLR and, resultingly, abandoned the interpretation of the estimated coefficients. Instead, the authors opted to derive adjusted predicted probabilities, which can be done from the results of MLR

(Greene, 2012). The fact that only two studies did not display the problems we have described is not surprising, as authors familiar with the challenges of MLR expectedly tend to use other models. Contrastingly, one study (Soroush et al., 2019) interpreted the results of MLR only through *p*-values and made no attempt at interpreting the effect estimates or the relevant confidence intervals. Consequently, this study employs the least coherent interpretation seen in the studies included in our review.

It is not clear how these results relate to existing research; to our knowledge, no other study has examined the extent of the previously described challenges with MLR.

None of the studies identified in this review use MLR to derive conditional predicted probabilities or discrete choice models.

The secondary aim of our review was to examine the use of data splits or similar methods to validate MLR models used for prediction. In this review, only a few studies ( $n = 9$ ) used MLR for prediction, with only three studies (33.33%) validating the model using data split, external validation, cross-validation or similar methods. This proportion is lower than what has been found in other reviews (Bouwmeester et al., 2012; Collins et al., 2011, 2013, 2014), though methodology among these varies. One review (Collins et al., 2014) excludes studies that do not conduct external validation of prediction models, and it is evident from the flow diagram that 43 articles are excluded because they use random data splits, cross-validation or bootstrapping. Disregarding studies exclusively performing external validation of existing models and counting excluded studies using internal validation, the proportion of studies using data splits, external validation or similar would be more than half of the included articles. Similarly, a review regarding the methodology and reporting in the development of prediction models for type 2 diabetes (Collins et al., 2011) found that 38% of the included studies used internal validation of the model and 54% used external validation. Note that this review did not include studies performing external validation of an already existing model. A similar review of studies of chronic kidney disease (Collins et al., 2013) found that, of the 11 included studies, 73% used internal validation by data splits or similar methods, and additionally, 36% of studies used external validation of the model. This review did not include studies that performed external validation of an already existing model. Lastly, a review examining the reporting and methodology of studies in clinical prediction research (Bouwmeester et al., 2012), which included both model development studies and predictor-finding studies, found that the use of external validation was 27% and 0%, respectively. Likewise, the use of internal validation was

infrequent, but in total, the proportion of studies that used data splits, external validation or similar methods was larger than seen in our review.

The low proportion of studies using model validation methods seen in our review may be explained by multiple factors. Firstly, only a small number of predictive studies were identified by our review. Secondly, the use of data splitting techniques and external validations seems to vary between diseases (Collins et al., 2011, 2013). Lastly, the inclusion of either model development studies and/or studies identifying predictors seems to influence the findings (Collins et al., 2011).

## 5.1 | Limitations

Our systematic search and review were limited to articles in either English or Danish and only considered studies that were peer-reviewed. Additionally, we did not have the opportunity to do a full-text search for all articles, and it is likely that a large portion of studies that apply MLR do not state so in the title or abstract. As such, it cannot be ruled out that studies could have been missed. We expect that if these articles had been identified, the problems raised in this review might be more prevalent. Therefore, we may have underestimated the extent of these methodological problems.

This review detailed studies involving stroke patients as either a population or an exposure. It is unclear if the extent of the problems we have discussed in this review is similar when MLR is used on other patient populations, but we suspect this to be the case.

## 6 | CONCLUSIONS

This systematic review identifies 68 articles selected for inclusion. Based on the three questions that we sought to answer, we have found significant problems in the usage and interpretation of MLR. Study authors routinely mislabel and fail to coherently interpret the effects inferred by MLR.

Consequently, the risk of misinterpreting the results from applying MLR and drawing conclusions that the data does not support is high.

Contrastingly, the use of comparisons of MLR and OLR results is not widespread.

Our findings show that the use of MLR to estimate and interpret associations is difficult, and based on the results of this review, authors should be careful when interpreting the results of MLR. Considering the year of publication of the included studies, it is evident that most studies identified in this review have been published in

recent years and that the use of MLR seems to be increasing (Figure 3). Thus, the challenges we have identified are increasingly problematic and relevant.

It is important to state that while MLR may necessitate a too complex interpretation, the model still has merit when used for prognostic or diagnostic purposes, performing IPTW in data with a nominal exposure variable, using discrete choice models or computing adjusted predicted probabilities.

## AUTHOR CONTRIBUTIONS

Line Ryberg Rasmussen and Amalie Helme Simoni conducted the systematic literature search and screened the articles based on abstracts and titles. Nicholas Fitzhugh and Jan Brink Valentin conducted the full-text screening and data extraction. Nicholas Fitzhugh drafted the manuscript, and Jan Brink Valentin contributed with inputs. All authors subsequently reviewed the draft and provided feedback.

## ACKNOWLEDGEMENTS

Not applicable.

## CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

## PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ejn.16084>.

## DATA AVAILABILITY STATEMENT

The data supporting the conclusions is available in the article.

## ORCID

Nicholas Fitzhugh  <https://orcid.org/0000-0002-1491-3586>

Line Ryberg Rasmussen  <https://orcid.org/0000-0001-9317-7662>

Jan Brink Valentin  <https://orcid.org/0000-0002-8205-7179>

## REFERENCES

- Agresti, A. (2013). *Categorical data analysis (third)*. John Wiley & Sons Incorporated.
- Altman, D. G., Vergouwe, Y., Royston, P., & Moons, K. G. M. (2009). Prognosis and prognostic research: Validating a prognostic model. *BMJ*, 338(may28 1), b605. <https://doi.org/10.1136/bmj.b605>
- Anagnostou, E., Koutsoudaki, P., Tountopoulou, A., Spengos, K., & Vassilopoulou, S. (2020). Bedside assessment of vergence in stroke patients. *Journal of Neuro-Ophthalmology*, 41(4), 424–430. <https://doi.org/10.1097/wno.0000000000001035>
- Béjot, Y., Guilloteau, A., Joux, J., Lannuzel, A., Mimeau, E., Mislin-Tritsch, C., Fournel, I., Bonithon-Kopp, C., & the INDIA Study Group. (2017). Social deprivation and stroke severity on admission: A French cohort study in Burgundy and the West Indies–Guyana region. *European Journal of Neurology*, 24(5), 694–702. <https://doi.org/10.1111/ene.13271>
- Boesgaard Graversen, C., Brink Valentin, J., Lytken Larsen, M., Riahi, S., Holmberg, T., & Paaske Johnsen, S. (2021). Non-persistence with medication as a mediator for the social inequality in risk of major adverse cardiovascular events in patients with incident acute coronary syndrome: A nationwide cohort study. *Clinical Epidemiology*, 13, 1071–1083. <https://doi.org/10.2147/CLEP.S335133>
- Bouwmeester, W., Zuithoff, N. P. A., Mallett, S., Geerlings, M. I., Vergouwe, Y., Steyerberg, E. W., Altman, D. G., & Moons, K. G. M. (2012). Reporting and methods in clinical prediction research: A systematic review. *PLoS Medicine*, 9(5), e1001221. <https://doi.org/10.1371/journal.pmed.1001221>
- Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, 46(4), 1171–1178. <https://doi.org/10.2307/2532457>
- Buntin, M. B., Colla, C. H., & Escarce, J. J. (2009). Effects of payment changes on trends in post-acute care. *Health Services Research*, 44(4), 1188–1210. <https://doi.org/10.1111/j.1475-6773.2009.00968.x>
- Buntin, M. B., Garten, A. D., Paddock, S., Saliba, D., Totten, M., & Escarce, J. J. (2005). How much is postacute care use affected by its availability? *Health Services Research*, 40(2), 413–434. <https://doi.org/10.1111/j.1475-6773.2005.01366.x>
- Butcher, K., Christensen, S., Parsons, M., De Silva, D. A., Ebinger, M., Levi, C., Jeerakathil, T., Campbell, B. C. V. V., Barber, P. A., Bladin, C., Fink, J., Tress, B., Donnan, G. A., & Davis, S. M. (2010). Postthrombolysis blood pressure elevation is associated with hemorrhagic transformation. *Stroke*, 41(1), 72–77. <https://doi.org/10.1161/STROKEAHA.109.563767>
- Choi, J.-Y., Cha, J., Jung, J.-M., Seo, W.-K., Oh, K., Cho, K.-H., & Yu, S. (2020). Left ventricular wall motion abnormality is associated with cryptogenic stroke. *International Journal of Stroke*, 15(2), 188–196. <https://doi.org/10.1177/1747493019834181>
- Choi, N. G., DiNitto, D. M., Marti, C. N., & Choi, B. Y. (2016). Sociodemographic Characteristics and Health Status of Lifetime Abstainers, Ex-Drinkers, Bingers, and Nonbingers Among Baby Boomers and Older Adults. *Substance Use & Misuse*, 51(5), 637–648. <https://doi.org/10.3109/10826084.2015.1133645>
- Choi, N., Lee, J. Y., Sunwoo, J. S., Roh, H., Ahn, M. Y., Park, S. T., & Lee, K. B. (2017). Recently uncontrolled glycaemia in diabetic patients is associated with the severity of intracranial atherosclerosis. *Journal of Stroke and Cerebrovascular Diseases*, 26(11), 2615–2621. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.06.025>
- Chow, F. C., Price, R. W., Hsue, P. Y., & Kim, A. S. (2017). Greater risk of stroke of undetermined etiology in a contemporary HIV-infected cohort compared with uninfected individuals. *Journal of Stroke and Cerebrovascular Diseases*, 26(5), 1154–1160. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.02.010>
- Clarke, P. J., Blount, V., & Colantonio, A. (2011). Cognitive impairment predicts fatal incident stroke: Findings from a national sample of older adults. *Journal of the American Geriatrics Society*, 59(8), 1490–1496. <https://doi.org/10.1111/j.1532-5415.2011.03494.x>

- Collins, G. S., De Groot, J. A., Dutton, S., Omar, O., Shanyinde, M., Tajar, A., Voysey, M., Wharton, R., Yu, L. M., Moons, K. G., & Altman, D. G. (2014). External validation of multivariable prediction models: A systematic review of methodological conduct and reporting. *BMC Medical Research Methodology*, *14*(1), 40. <https://doi.org/10.1186/1471-2288-14-40>
- Collins, G. S., Mallett, S., Omar, O., & Yu, L. M. (2011). Developing risk prediction models for type 2 diabetes: A systematic review of methodology and reporting. *BMC Medicine*, *9*(1), 103. <https://doi.org/10.1186/1741-7015-9-103>
- Collins, G. S., Omar, O., Shanyinde, M., & Yu, L. M. (2013). A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *Journal of Clinical Epidemiology*, *66*(3), 268–277. <https://doi.org/10.1016/j.jclinepi.2012.06.020>
- Covert, S., Johnson, J. K., Stilphen, M., Passek, S., Thompson, N. R., & Katzan, I. (2020). Use of the activity measure for post-acute care “6 clicks” Basic mobility inpatient short form and National Institutes of Health stroke scale to predict hospital discharge disposition after stroke. *Physical Therapy*, *100*(9), 1423–1433. <https://doi.org/10.1093/ptj/pzaa102>
- Cupini, L. M., Pasqualetti, P., Diomed, M., Vernieri, F., Silvestrini, M., Rizzato, B., Ferrante, F., & Bernardi, G. (2002). Carotid artery intima-media thickness and lacunar versus non-lacunar infarcts. *Stroke*, *33*(3), 689–694. <https://doi.org/10.1161/hs0302.103661>
- Daneshvari, N. O., & Johansen, M. C. (2021). Associations between cerebral magnetic resonance imaging infarct volume and acute ischemic stroke etiology. *PLoS ONE*, *16*(8), e0256458. <https://doi.org/10.1371/journal.pone.0256458>
- Debette, S., Metso, T., Pezzini, A., Abboud, S., Metso, A., Leys, D., Bersano, A., Louillet, F., Caso, V., Lamy, C., Medeiros, E., Samson, Y., Grond-Ginsbach, C., Engelter, S. T., Thijs, V., Beretta, S., Béjot, Y., Sessa, M., Lorenza Muiesan, M., ... Cervical Artery Dissection and Ischemic Stroke Patients (CADISP) Group. (2011). Association of vascular risk factors with cervical artery dissection and ischemic stroke in young adults. *Circulation*, *123*(14), 1537–1544. <https://doi.org/10.1161/CIRCULATIONAHA.110.000125>
- Duchon, L. M. (2014). *Families and their health care after homelessness: Opportunities for improving access* (pp. 1–148). Routledge. <https://doi.org/10.4324/97811315805139>
- Ellis, C., Magwood, G., & White, B. M. (2017). Racial differences in patient-reported post-stroke disability in older adults. *Geriatrics*, *2*(2), 1–9. <https://doi.org/10.3390/geriatrics2020016>
- Fitzpatrick, T., Gocan, S., Wang, C. Q., Bourgoin, A., Taljaard, M., Cheng, W., Dowlatshahi, D., Stotts, G., & Shamy, M. (2020). Cohort study of features used by experts to diagnose transient ischemic attack. *The Neurohospitalist*, *10*(4), 245–249. <https://doi.org/10.1177/1941874420908485>
- Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>
- Gocan, S., Fitzpatrick, T., Wang, C. Q., Taljaard, M., Cheng, W., Bourgoin, A., Dowlatshahi, D., Stotts, G., & Shamy, M. (2020). Diagnosis of transient ischemic attack. *Stroke*, *51*(11), 3371–3374. <https://doi.org/10.1161/strokeaha.120.031510>
- Gong, P., Liu, Y., Gong, Y., Chen, G., Zhang, X., Wang, S., Zhou, F., Duan, R., Chen, W., Huang, T., Wang, M., Deng, Q., Shi, H., Zhou, J., Jiang, T., & Zhang, Y. (2021). The association of neutrophil to lymphocyte ratio, platelet to lymphocyte ratio, and lymphocyte to monocyte ratio with post-thrombolysis early neurological outcomes in patients with acute ischemic stroke. *Journal of Neuroinflammation*, *18*(1), 51. <https://doi.org/10.1186/s12974-021-02090-6>
- Goyal, M., Menon, B. K., van Zwam, W. H., Dippel, D. W. J., Mitchell, P. J., Demchuk, A. M., Dávalos, A., Majoie, C. B. L. M., van der Lugt, A., De Miquel, M. A., Donnan, G. A., Roos, Y. B. W. E. M., Bonafe, A., Jahan, R., Diener, H. C., Van den Berg, L. A., Levy, E. I., Berkhemer, O. A., Pereira, V. M., ... Jovin, T. G. (2016). Endovascular thrombectomy after large-vessel ischaemic stroke: A meta-analysis of individual patient data from five randomised trials. *The Lancet*, *387*(10029), 1723–1731. [https://doi.org/10.1016/S0140-6736\(16\)00163-X](https://doi.org/10.1016/S0140-6736(16)00163-X)
- Greene, W. H. (2012). Econometric analysis. In P. J. Boardman (Ed.), *Pearson (7th edition, Vol. 97)*. Pearson. [https://doi.org/10.1007/3-7908-1599-3\\_5](https://doi.org/10.1007/3-7908-1599-3_5)
- Grefkes, C., & Fink, G. R. (2020). Recovery from stroke: Current concepts and future perspectives. *Neurological Research and Practice*, *2*(1), 17. <https://doi.org/10.1186/s42466-020-00060-6>
- Hamaguchi, T., Yamada, N., Hada, T., & Abo, M. (2020). Prediction of motor recovery in the upper extremity for repetitive transcranial magnetic stimulation and occupational therapy goal setting in patients with chronic stroke: A retrospective analysis of prospectively collected data. *Frontiers in Neurology*, *11*-(October), 581186. <https://doi.org/10.3389/fneur.2020.581186>
- Hendrickx, W., Riveros, C., Askim, T., Bussmann, J. B. J. J., Callisaya, M. L., Chastin, S. F. M. M., Dean, C., Ezeugwu, V., Jones, T. M., Kuys, S. S., Mahendran, N., Manns, P. J., Mead, G., Moore, S. A., Paul, L., Pisters, M. F., Saunders, D. H., Simpson, D. B., Tiegies, Z., ... English, C. (2021). An exploration of sedentary behavior patterns in community-dwelling people with stroke: A cluster-based analysis. *Journal of Neurologic Physical Therapy*, *45*(3), 221–227. <https://doi.org/10.1097/NPT.0000000000000357>
- Howrey, B. T., Kuo, Y.-F., & Goodwin, J. S. (2011). Association of Care by Hospitalists on Discharge Destination and 30-day Outcomes After Acute Ischemic Stroke. *Medical Care*, *49*(8), 701–707. <https://doi.org/10.1097/mlr.0b013e3182166cb6>
- Huo, Y., Huang, S., Li, R., Gong, X., Zhang, W., Zhang, R., & Qin, X. (2021). Elevated hemoglobin is independently associated with enlarged perivascular spaces in the central semiovale. *Scientific Reports*, *11*(1), 2820. <https://doi.org/10.1038/s41598-021-82327-9>
- Jacob, B., Chan, V., Stock, D., Colantonio, A., & Cullen, N. (2021). Determinants of discharge disposition from acute care for survivors of hypoxic-ischemic brain injury: Results from a large population-based cohort data set. *Archives of Physical Medicine and Rehabilitation*, *102*(8), 1514–1523. <https://doi.org/10.1016/j.apmr.2021.01.083>
- Jaffre, A., Ruidavets, J. B., Calviere, L., Viguier, A., Ferrieres, J., & Larrue, V. (2014). Risk factor profile by etiological subtype of ischemic stroke in the young. *Clinical Neurology and Neurosurgery*, *120*(June 2006), 78–83. <https://doi.org/10.1016/j.clineuro.2014.02.017>
- Johnson, D. A., Brown, D. L., Morgenstern, L. B., Meurer, W. J., & Lisabeth, L. D. (2015). The association of neighborhood characteristics with sleep duration and daytime sleepiness. *Sleep Health*, *1*(3), 148–155. <https://doi.org/10.1016/j.sleh.2015.06.002>

- Jung, J.-M., Cha, J., Choi, J., Eun, M.-Y., Seo, W.-K., Oh, K., Cho, K.-H., & Yu, S. (2016). Prestroke antiplatelet agents in first-time ischemic stroke are related to subtypes. *European Neurology*, 75(1–2), 89–95. Portico. <https://doi.org/10.1159/000444295>
- Kang, K., Lee, W. W., Lee, J. J., Park, J. M., Kwon, O., & Kim, B. K. (2018). Association of higher waist circumference with milder stroke severity in acute ischaemic stroke. *Neurological Research*, 40(9), 785–794. <https://doi.org/10.1080/01616412.2018.1479346>
- Kashani, N., Cimflova, P., Ospel, J. M., Singh, N., Almekhlafi, M. A., Rempel, J., Fiehler, J., Chen, M., Sakai, N., Agid, R., Heran, M., Kappelhof, M., & Goyal, M. (2021). Endovascular device choice and tools for recanalization of medium vessel occlusions: Insights From the MeVO FRONTIERS International Survey. *Frontiers in Neurology*, 12. <https://doi.org/10.3389/fneur.2021.735899>
- Kim, H., Friedlander, Y., Longstreth, W. T., Edwards, K. L., Schwartz, S. M., & Siscovick, D. S. (2004). Family history as a risk factor for stroke in young women. *American Journal of Preventive Medicine*, 27(5), 391–396. <https://doi.org/10.1016/j.amepre.2004.08.008>
- Kleinbaum, D. G., & Klein, M. (2010). Logistic regression. In *Logistic regression*. Springer New York. <https://doi.org/10.1007/978-1-4419-1742-3>
- Krajcoviechova, A., Wohlfahrt, P., Mayer, O., Vanek, J., Hajkova, J., Hlinovsky, D., Kvasnicka, T., Tremblay, J., Hamet, P., Filipovsky, J., Kvasnicka, J., & Cifkova, R. (2015). Tobacco smoking strongly modifies the association of prothrombin G20210A with undetermined stroke: Consecutive survivors and population-based controls. *Atherosclerosis*, 240(2), 446–452. <https://doi.org/10.1016/j.atherosclerosis.2015.04.010>
- Kuohn, L. R., Leasure, A. C., Acosta, J. N., Vanent, K., Murthy, S. B., Kamel, H., Matouk, C. C., Sansing, L. H., Falcone, G. J., & Sheth, K. N. (2020). Cause of death in spontaneous intracerebral hemorrhage survivors. *Neurology*, 95(20), e2736–e2745. <https://doi.org/10.1212/wnl.0000000000010736>
- Lipsman, N., Tolentino, J., & Macdonald, R. L. (2009). Effect of country or continent of treatment on outcome after aneurysmal subarachnoid hemorrhage: Clinical article. *Journal of Neurosurgery*, 111(1), 67–74. <https://doi.org/10.3171/2008.10.JNS08184>
- Loikas, D., Linnér, L., Sundström, A., Wettermark, B., & von Euler, M. (2021). Post-stroke epilepsy and antiepileptic drug use in men and women. *Basic and Clinical Pharmacology and Toxicology*, 129(2), 148–157. <https://doi.org/10.1111/bcpt.13617>
- Ma, G., Sun, L., Qie, Z., He, J., & Cui, F. (2021). Factors associated with benzodiazepines prolonged-term use in post-stroke subjective sleep disturbance: A single-centre retrospective study from china. *Drug Design, Development and Therapy*, 15, 2469–2481. <https://doi.org/10.2147/dddt.s298552>
- Macintosh, B. J., Cohen, E., Colby-Milley, J., Fang, J., Zhou, L., Ouk, M., Wu, C. Y., Shah, B. R., Lanctôt, K., Herrmann, N., Linkewich, E., Law, M., Black, S. E., Swartz, R. H., Kapral, M. K., Edwards, J. D., & Swardfager, W. (2021). Diabetes mellitus is associated with poor in-hospital and long-term outcomes in young and midlife stroke survivors. *Journal of the American Heart Association*, 10(14), e019991. <https://doi.org/10.1161/JAHA.120.019991>
- Mallett, S., Royston, P., Dutton, S., Waters, R., & Altman, D. G. (2010). Reporting methods in studies developing prognostic models in cancer: A review. *BMC Medicine*, 8(1), 20. <https://doi.org/10.1186/1741-7015-8-20>
- Matsukawa, H., Shinoda, M., Fujii, M., Takahashi, O., Yamamoto, D., Murakata, A., & Ishikawa, R. (2013). Impact of body mass index on the location of spontaneous intracerebral hemorrhage. *World Neurosurgery*, 79(3), 478–483. <https://doi.org/10.1016/j.wneu.2011.10.038>
- Meurer, W. J., Sánchez, B. N., Smith, M. A., Lisabeth, L. D., Majersik, J. J., Brown, D. L., Uchino, K., Bonikowski, F. P., Mendizabal, J. E., Zahuranec, D. B., & Morgenstern, L. B. (2009). Predicting ischaemic stroke subtype from presenting systolic blood pressure: The BASIC Project. *Journal of Internal Medicine*, 265(3), 388–396. <https://doi.org/10.1111/j.1365-2796.2008.02022.x>
- Minen, M. T., Weissman, J., & Tietjen, G. E. (2019). The relationship between migraine or severe headache and chronic health conditions: A cross-sectional study from the national health interview survey 2013–2015. *Pain Medicine*, 20(11), 2263–2271. <https://doi.org/10.1093/pm/pnz113>
- Mizrahi, E. H., Fleissig, Y., Arad, M., & Adunsky, A. (2005). Plasma homocysteine level and functional outcome of patients with ischemic stroke. *Archives of Physical Medicine and Rehabilitation*, 86(1), 60–63. <https://doi.org/10.1016/j.apmr.2004.01.031>
- Moons, K. G. M., Royston, P., Vergouwe, Y., Grobbee, D. E., & Altman, D. G. (2009). Prognosis and prognostic research: What, why, and how? *BMJ (Online)*, 338(7706), 1317–1320. <https://doi.org/10.1136/bmj.b375>
- Morrissey, M. A., Sloan, F. A., & Valvona, J. (1988). Medicare Prospective Payment and Posthospital Transfers to Subacute Care. *Medical Care*, 26(7), 685–698. <https://www.jstor.org/stable/3765491>
- Nagai, K., Kozaki, K., Sonohara, K., Akishita, M., & Toba, K. (2011). Relationship between interleukin-6 and cerebral deep white matter and periventricular hyperintensity in elderly women. *Geriatrics and Gerontology International*, 11(3), 328–332. <https://doi.org/10.1111/j.1447-0594.2010.00686.x>
- Nighoghossian, N., Abbas, F., Cho, T. H., Geraldo, A. F., Cottaz, V., Janecek, E., Mechtouff, L., Bischoff, M., El Houry, C., Schott, A. M., Derex, L., Hermier, M., Tisserand, L. G., Amelie, R., Chamard, L., & Berthezene, Y. (2016). Impact of leukoaraiosis on parenchymal hemorrhage in elderly patients treated with thrombolysis. *Neuroradiology*, 58(10), 961–967. <https://doi.org/10.1007/s00234-016-1725-7>
- Park, J. M., Koo, J. S., Kim, B. K., Kwon, O., Lee, J. J. S., Kang, K., Lee, J. J. S., Lee, J. J. S., & Bae, H. J. (2013). Vertebrobasilar dolichoectasia as a risk factor for cerebral microbleeds. *European Journal of Neurology*, 20(5), 824–830. <https://doi.org/10.1111/ene.12075>
- Piayda, K., Afzal, S., Nielsen-Kudsk, J. E., Schmidt, B., Mazzone, P., Berti, S., Fischer, S., Lund, J., Montorfano, M., Hildick-Smith, D., Gage, R., Zhao, H., & Zeus, T. (2021). Length of stay following percutaneous left atrial appendage occlusion: Data from the prospective, multicenter Amplatzer Amulet Occluder Observational Study. *PLOS ONE*, 16(8), e0255721. <https://doi.org/10.1371/journal.pone.0255721>
- Pinto, C. A., Chua, G. N., Bridges, J. F. P., Brookes, E., Hyacinthe, J., & Tervonen, T. (2021). Comparing patient preferences for antithrombotic treatment during the acute and chronic phases of myocardial infarction: A discrete-choice experiment. *The*

- Patient - Patient-Centered Outcomes Research, 15(2), 255–266. <https://doi.org/10.1007/s40271-021-00548-6>
- Qiu, C., Cotch, M. F., Sigurdsson, S., Klein, R., Jonasson, F., Klein, B. E. K., Garcia, M., Jonsson, P. V., Harris, T. B., Eiriksdottir, G., Kjartansson, O., van Buchem, M. A., Gudnason, V., & Launer, L. J. (2009). Microvascular lesions in the brain and retina: The age, gene/environment susceptibility-Reykjavik study. *Annals of Neurology*, 65(5), 569–576. <https://doi.org/10.1002/ana.21614>
- Quiñones, A. R., Newsom, J. T., Elman, M. R., Markwardt, S., Nagel, C. L., Dorr, D. A., Allore, H. G., & Botoseneanu, A. (2021). Racial and Ethnic Differences in Multimorbidity Changes Over Time. *Medical Care*, 59(5), 402–409. <https://doi.org/10.1097/mlr.0000000000001527>
- R Core Team. (2018). R: A language and environment for statistical computing. <https://www.r-project.org/>
- Rasmussen, M., Schönenberger, S., Hendén, P. L., Valentin, J. B., Espelund, U. S., Sørensen, L. H., Juul, N., Uhlmann, L., Johnsen, S. P., Rentzos, A., Bösel, J., & Simonsen, C. Z. (2020). Blood pressure thresholds and neurologic outcomes after endovascular therapy for acute ischemic stroke. *JAMA Neurology*, 77(5), 622–631. <https://doi.org/10.1001/jamaneurol.2019.4838>
- Rehman, S., Chandra, R. V., Zhou, K., Tan, D., Lai, L., Asadi, H., Froelich, J., Thani, N., Nichols, L., Blizzard, L., Smith, K., Thrift, A. G., Stirling, C., Callisaya, M. L., Breslin, M., Reeves, M. J., & Gall, S. (2020). Sex differences in aneurysmal subarachnoid haemorrhage (aSAH): Aneurysm characteristics, neurological complications, and outcome. *Acta Neurochirurgica*, 162(9), 2271–2282. <https://doi.org/10.1007/s00701-020-04469-5>
- Rimmele, D. L., Schrage, T., Leberherz, L., Kriston, L., Gerloff, C., Härter, M., & Thomalla, G. (2021). Profiles of patients' self-reported health after acute stroke. *Neurological Research and Practice*, 3(1), 43. <https://doi.org/10.1186/s42466-021-00146-9>
- Rist, P. M., Buring, J. E., Kase, C. S., & Kurth, T. (2013). Effect of low-dose aspirin on functional outcome from cerebral vascular events in women. *Stroke*, 44(2), 432–436. <https://doi.org/10.1161/STROKEAHA.112.672451>
- Rist, P. M., Buring, J. E., Kase, C. S., & Kurth, T. (2016). Healthy lifestyle and functional outcomes from stroke in women. *American Journal of Medicine*, 129(7), 715–724.e2. <https://doi.org/10.1016/j.amjmed.2016.02.002>
- Rodríguez-Castro, E., López-Dequít, I., Santamaría-Cadavid, M., Arias-Rivas, S., Rodríguez-Yañez, M., Pumar, J. M., Hervella, P., López-Arias, E., da Silva-Candal, A., Estany, A., Piñeiro-Lamas, M., Sobrino, T., Campos, F., Portela, M., Vázquez-Lima, M., Castillo, J., & Iglesias-Rey, R. (2018). Trends in stroke outcomes in the last ten years in a European tertiary hospital. *BMC Neurology*, 18(1). <https://doi.org/10.1186/s12883-018-1164-7>
- Schlegel, D. J. (2004). Prediction of hospital disposition after thrombolysis for acute ischemic stroke using the national institutes of health stroke scale. *Archives of Neurology*, 61(7), 1061. <https://doi.org/10.1001/archneur.61.7.1061>
- Schönenberger, S., Hendén, P. L., Simonsen, C. Z., Uhlmann, L., Klose, C., Pfaff, J. A. R., Yoo, A. J., Sørensen, L. H., Ringleb, P. A., Wick, W., Kieser, M., Möhlenbruch, M. A., Rasmussen, M., Rentzos, A., & Bösel, J. (2019). Association of general anesthesia vs procedural sedation with functional outcome among patients with acute ischemic stroke undergoing thrombectomy: A systematic review and meta-analysis. *Jama*, 322(13), 1283–1293. <https://doi.org/10.1001/JAMA.2019.11455>
- Shigematsu, K., Watanabe, Y., & Nakano, H. (2015). Weekly variations of stroke occurrence: An observational cohort study based on the Kyoto Stroke Registry, Japan. *BMJ Open*, 5(3), e006294. <https://doi.org/10.1136/bmjopen-2014-006294>
- Singh, K., Patel, S. A., Biswas, S., Shivashankar, R., Kondal, D., Ajay, V. S., Anjana, R. M., Fatmi, Z., Ali, M. K., Kadir, M. M., Mohan, V., Tandon, N., Narayan, K. M. V., & Prabhakaran, D. (2019). Multimorbidity in South Asian adults: prevalence, risk factors and mortality. *Journal of Public Health*, 41(1), 80–89. <https://doi.org/10.1093/pubmed/fdy017>
- Song, T.-J., Cho, H.-J., Chang, Y., Youn, M., Shin, M.-J., Jo, I., Heo, J. H., & Kim, Y.-J. (2015). Low-Density-Lipoprotein particle size predicts a poor outcome in patients with atherothrombotic stroke. *Journal of Clinical Neurology*, 11(1), 80. <https://doi.org/10.3988/jcn.2015.11.1.80>
- Song, T., Pan, Y., Chen, R., Li, H., Zhao, X., Liu, L., Wang, C., Wang, Y., & Wang, Y. (2017). Is there a correlation between socioeconomic disparity and functional outcome after acute ischemic stroke? *PLOS ONE*, 12(7), e0181196. <https://doi.org/10.1371/journal.pone.0181196>
- Soroush, A., Sariaslani, P., Baharirad, N., Shams-Alizadeh, N., & Komasi, S. (2019). Partitioning stroke patients, determining related factors, and comparing derived clusters based on 12-month health outcomes. *Archives of Iranian Medicine*, 22(23), 708–715.
- Steyerberg, E. W., & Vergouwe, Y. (2014). Towards better clinical prediction models: Seven steps for development and an ABCD for validation. *European Heart Journal*, 35(29), 1925–1931. <https://doi.org/10.1093/eurheartj/ehu207>
- Stulberg, E. L., Dong, L., Zheutlin, A. R., Kim, S., Claflin, E. S., Skolarus, L. E., Morgenstern, L. B., & Lisabeth, L. D. (2019). Associations of Self-Reported History of Depression and Antidepressant Use Before Stroke Onset With Poststroke Post-Acute Rehabilitation Care—An Exploratory Study: The BASIC (Brain Attack Surveillance in Corpus Christi) Project. *Journal of the American Heart Association*, 8(16), e013382. <https://doi.org/10.1161/jaha.119.013382>
- Thomalla, G., Simonsen, C. Z., Boutitie, F., Andersen, G., Berthezene, Y., Cheng, B., Cheripelli, B., Cho, T.-H., Fazekas, F., Fiehler, J., Ford, I., Galinovic, I., Gellissen, S., Golsari, A., Gregori, J., Günther, M., Guibernau, J., Häusler, K. G., Hennerici, M., ... Gerloff, C. (2018). MRI-guided thrombolysis for stroke with unknown time of onset. *New England Journal of Medicine*, 379(7), 611–622. <https://doi.org/10.1056/NEJMoa1804355>
- Topriceanu, C. C., Wong, A., Moon, J. C., Hughes, A. D., Bann, D., Chaturvedi, N., Patalay, P., Conti, G., & Captur, G. (2021). Evaluating access to health and care services during lockdown by the COVID-19 survey in five UK national longitudinal studies. *BMJ Open*, 11(3), e045813. <https://doi.org/10.1136/bmjopen-2020-045813>
- Tuttolomondo, A., Casuccio, A., Buttà, C., Pecoraro, R., Di Raimondo, D., Della Corte, V., Arnao, V., Clemente, G., Maida, C., Simonetta, I., Miceli, G., Lucifora, B., Cirrincione, A., Di Bona, D., Corpora, F., Maugeri, R., Iacopino, G., & Pinto, A. (2015). Mediterranean Diet in patients with acute ischemic stroke: Relationships between Mediterranean Diet score, diagnostic subtype, and stroke

- severity index. *Atherosclerosis*, 243(1), 260–267. <https://doi.org/10.1016/j.atherosclerosis.2015.09.017>
- Uijl, A., Lund, L. H., Vaartjes, I., Brughts, J. J., Linssen, G. C., Asselbergs, F. W., Hoes, A. W., Dahlström, U., Koudstaal, S., & Savarese, G. (2020). A registry-based algorithm to predict ejection fraction in patients with heart failure. *ESC Heart Failure*, 7(5), 2388–2397. <https://doi.org/10.1002/ehf2.12779>
- Williams, N. L., Suarez, A., Negoita, S., Hillis, A. E., Gottesman, R. F., & Johansen, M. C. (2021). Cardiac structure and function is associated with hemispatial neglect severity. *Frontiers in Neurology*, 12(May), 666257. <https://doi.org/10.3389/fneur.2021.666257>
- Wolfe, C. D. A., Tilling, K., Beech, R., & Rudd, A. G. (1999). Variations in case fatality and dependency from stroke in western and central Europe. *Stroke*, 30(2), 350–356. <https://doi.org/10.1161/01.STR.30.2.350>
- Wysocki, N. A., Bambhroliya, A., Ankrom, C., Vahidy, F., Astudillo, C., Trevino, A., Malazarte, R., Cossey, T. C., Jagolino-Cole, A., Savitz, S., Wu, T.-C., & Sharrief, A. (2019). Outcomes Among Patients With Ischemic Stroke Treated With Intravenous tPA (Tissue-Type Plasminogen Activator) via Telemedicine. *Stroke*, 50(4), 895–900. <https://doi.org/10.1161/strokeaha.118.024703>
- Ye, F., Liu, J., Yang, S., & Guo, F. Q. (2015). Higher apolipoprotein B levels are associated with earlier onset of first-ever atherosclerotic stroke. *International Journal of Neuroscience*, 125(3), 186–190. <https://doi.org/10.3109/00207454.2014.951042>
- Zhao, D., Shridharmurthy, D., Alcusky, M. J., Yuan, Y., Nunes, A. P., Hume, A. L., Baek, J., & Lapane, K. L. (2020). The prevalence and factors associated with antiepileptic drug use in US nursing home residents. *Drugs & Aging*, 37(2), 137–145. <https://doi.org/10.1007/s40266-019-00732-2>
- Zinkstok, S. M., Beenen, L. F., Majoie, C. B., Marquering, H. A., De Haan, R. J., & Roos, Y. B. (2014). Early deterioration after thrombolysis plus aspirin in acute stroke: A post hoc analysis of the antiplatelet therapy in combination with recombinant t-PA thrombolysis in ischemic stroke trial. *Stroke*, 45(10), 3080–3082. <https://doi.org/10.1161/STROKEAHA.114.006268>
- Zong, L., Yao, M., Ni, J., Zhou, L., Yuan, J., Peng, B., Zhu, Y. C., & Cui, L. (2016). Kidney function is associated with severity of white matter hyperintensity in patients with acute ischemic stroke/TIA. *BMC Neurology*, 16(1), 193. <https://doi.org/10.1186/s12883-016-0714-0>
- Zongo, A., Simpson, S., Johnson, J. A., & Eurich, D. T. (2019). Change in trajectories of adherence to lipid-lowering drugs following non-fatal acute coronary syndrome or stroke. *Journal*

of the American Heart Association, 8(23), e013857. <https://doi.org/10.1161/jaha.119.013857>

**How to cite this article:** Fitzhugh, N., Rasmussen, L. R., Simoni, A. H., & Valentin, J. B. (2023). Misuse of multinomial logistic regression in stroke related health research: A systematic review of methodology. *European Journal of Neuroscience*, 58(4), 3116–3131. <https://doi.org/10.1111/ejn.16084>

## APPENDIX A: SEARCH STRINGS

### PubMed search string

((((((((((("Stroke"[Mesh]) OR "Brain Ischemia"[Mesh]) OR stroke\*[tw]) OR Brain Infarct\*[tw]) OR Cerebrovascular Accident\*[tw]) OR Apoplex\*[tw]) OR Cerebrovascular event\*[tw]) OR Cerebral Infarct\*[tw]) OR Brain Stem Infarct\*[tw]) OR Brain Ischemia\*[tw]) OR Cerebral Ischemia\*[tw]) OR Transient Ischemic Attack\*[tw])) AND ((((((multinomial regression\*[tw]) OR multinomial logit regression\*[tw]) OR multinomial logistic regression\*[tw]) OR multinomial regression\*[tw]) OR polytomous logistic regression\*[tw]) OR polytomous regression\*[tw]) OR multinomial logistic regression\*[tw]))

### EMBASE search string

('brain ischemia'/exp OR 'brain ischemia' OR 'cerebrovascular accident'/exp OR 'cerebrovascular accident' OR stroke\*:ti,ab,kw OR 'brain infarct\*':ti,ab,kw OR 'cerebrovascular accident\*':ti,ab,kw OR 'apoplex\*':ti,ab,kw OR 'cerebrovascular event\*':ti,ab,kw OR 'cerebral infarct\*':ti,ab,kw OR 'brain stem infarct\*':ti,ab,kw OR 'brain ischemia\*':ti,ab,kw OR 'cerebral ischemia\*':ti,ab,kw OR 'transient ischemic attack\*':ti,ab,kw) AND (((multinomial NEAR/3 regression\*):ti,ab,kw) OR ((polytomous NEAR/3 regression\*):ti,ab,kw)) NOT [conference abstract]/lim



## APPENDIX B: INCLUDED STUDIES

TABLE B1 Included studies.

Study	Exponentiated coefficients are denoted as ...	Does study compare MLR results with OLR?	Does the study conduct a coherent or separate interpretation of MLR results?	Does the study have a predictive aim?	Does the study conduct any form for cross validation?
Cupini et al. (2002)	RRR	No	Separate	Yes	No
Schlegel (2004)	RR	No	Separate	Yes	No
Kim et al. (2004)	OR, RR	No	Separate	No	No
Mizrahi et al. (2005)	OR	No	Separate	Yes	No
Meurer et al. (2009)	OR	No	Separate	Yes	No
Butcher et al. (2010)	OR	No	Separate	Yes	No
Clarke et al. (2011)	OR	No	Separate	No	No
Debette et al. (2011)	OR	No	Separate	No	No
Rist et al. (2013)	OR	No	Separate	No	No
Park et al. (2013)	OR	No	Separate	No	No
Matsukawa et al. (2013)	OR	No	Separate	No	No
Zinkstok et al. (2014)	OR	No	Separate	No	No
Jaffre et al. (2014)	OR	No	Separate	No	No
Ye et al. (2015)	OR, RR	No	Separate	No	No
Tuttolomondo et al. (2015)	OR	No	Separate	No	No
Jung et al. (2016)	OR	No	Separate	No	No
Rist et al. (2016)	OR, RR	No	Separate	No	No
Zong et al. (2016)	OR	No	Separate	No	No
Choi et al. (2017)	OR	No	Separate	No	No
Chow et al. (2017)	RRR, RR	No	Separate	No	No
Ellis et al. (2017)	OR	No	Separate	No	No
Béjot et al. (2017)	OR	No	Separate	No	No
Kang et al. (2018)	OR	No	Separate	No	No
Rodríguez-Castro et al. (2018)	OR	No	Separate	No	No
Soroush et al. (2019)	OR	No	Separate	No	No
Zongo et al. (2019)	OR	No	Separate	No	No
Choi et al. (2020)	OR	No	Separate	No	No
Stulberg et al. (2019)	OR	No	Separate	No	No
Anagnostou et al. (2020)	<sup>a</sup>	No	Separate	No	No
Hamaguchi et al. (2020)	OR	No	Separate	Yes	Yes
Covert et al. (2020)	OR	No	Separate	Yes	Yes
Rehman et al. (2020)	RRR	No	Separate	No	No
Fitzpatrick et al. (2020)	OR	No	Separate	No	No
Gocan et al. (2020)	OR	No	Separate	No	No
Kuohn et al. (2020)	OR	No	Separate	No	No
Williams et al. (2021)	RRR	No	Coherent	No	No

(Continued)

Study	Exponentiated coefficients are denoted as ...	Does study compare MLR results with OLR?	Does the study conduct a coherent or separate interpretation of MLR results?	Does the study have a predictive aim?	Does the study conduct any form for cross validation?
Daneshvari and Johansen (2021)	RRR	No	Separate	No	No
Macintosh et al. (2021)	OR	No	Coherent	No	No
Hendrickx et al. (2021)	- <sup>b</sup>	No	Coherent	No	No
Ma et al. (2021)	OR	No	Separate	No	No
Huo et al. (2021)	OR	No	Separate	No	No
Gong et al. (2021)	OR	No	Separate	Yes	No
Loikas et al. (2021)	RRR	No	Separate	No	No
Rimmele et al. (2021)	OR	No	Separate	No	No
Morrisey et al. (1988)	other <sup>c</sup>	No	Separate	No	No
Buntin et al. (2005)	- <sup>d</sup>	No	Coherent	No	No
Qiu et al. (2009)	OR	No	Separate	No	No
Buntin et al. (2009)	- <sup>d</sup>	No	Coherent	No	No
Nagai et al. (2011)	OR	No	- <sup>e</sup>	No	No
Howrey et al. (2011)	OR	No	Separate	No	No
Song et al. (2015)	OR	No	Separate	No	No
Shigematsu et al. (2015)	OR	No	Separate	No	No
Krajcoviechova et al. (2015)	OR	No	Separate	No	No
Johnson et al. (Zongo et al., 2019)	OR	Yes	Separate	No	No
Choi et al. (2016)	RRR	No	Separate	No	No
Nighoghossian et al. (2016)	OR	No	Separate	No	No
Singh et al. (2019)	other <sup>f</sup>	No	Separate	No	No
Wysocki et al. (2019)	RRR	No	Separate	No	No
Minen et al. (2019)	OR	No	Separate	No	No
Zhao et al. (2020)	OR	No	Separate	No	No
Uijl et al. (2020)	OR	No	- <sup>f</sup>	Yes	Yes
Piayda et al. (2021)	OR	No	Separate	No	No
Kashani et al. (2021)	RR	No	Separate	No	No
Quiñones et al. (2021)	OR	No	Separate	No	No
Pinto et al. (2021)	OR	No	Separate	No	No
Jacob et al. (2021)	OR	No	Separate	No	No
Song et al. (2017)	OR	No	Separate	No	No
Lipsman et al. (2009)	RRR	Yes	- <sup>g</sup>	No	No

<sup>a</sup> Exponentiated coefficients are not denoted in any way.

<sup>b</sup> Exponentiated coefficients are not denoted in any way.

<sup>c</sup> Coefficients are presented, but not denoted in any way.

<sup>d</sup> Concludes that interpretation is not easy and abandons this. Instead risk estimates are derived.

<sup>e</sup> Does not attempt to interpret results.

<sup>f</sup> Results not interpreted

<sup>g</sup> Result not interpreted, and not presented in article.