NeuroRehabilitation
2026, Vol. 58(1) 3–16
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10538135251393516
journals.sagepub.com/home/nre
Mary Ann Liebert
A Part of Sage

Defining Severity Levels for Post-Stroke Upper Limb Motor Impairment and Activity Limitation: A Systematic Review

Elena Ierardi[1], Frederike van Wijck[2], Myzoon Ali[3,4], Catherine Best[5] and Fiona Coupar[1]

Abstract

Background: Stroke often results in upper limb motor impairment and activity limitation, however terminology to describe severity levels vary. This hinders data pooling from clinical trials to inform practice. There are no reviews that have synthesized severity levels of stroke-related upper limb motor impairment or activity limitation.

Objective: To systematically review published literature on descriptors of severity levels for post-stroke upper limb motor impairment and activity limitation.

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, we searched eight major databases. Inclusion criteria: primary research studies, adults post-stroke, severity of upper limb motor impairment and/or activity limitation described. We classified included papers by assessing descriptor precision: 'green' (studies including descriptors that used recommended outcome measures, cut-offs or central tendency and dispersion), 'red' (descriptors only), and 'amber' (remaining studies). Of the 'green' studies, we identified the most commonly reported descriptors and measures, and computed cut-off scores using non-parametric statistics.

Results: From 17,273 records, 750 studies were included. The most commonly used severity descriptors were 'mild, and/or moderate, and/or severe,' used in 580 (77%) of studies. For the Fugl-Meyer Assessment (Upper Extremity) (57 studies, 8% of the total number of studies included), 'severe' ranged from 0 to 25, 'moderate' from 26 to 50, and 'mild' from 51 to 66. Limited data from the remaining studies prevented further analysis.

Conclusions: Our review highlights a lack of standardization of the operationalization of 'severity' of post-stroke upper limb motor impairment and activity limitation. It provides a foundation for developing a standardized clinical language to describe severity levels to improve research and clinical practice.

Keywords

systematic review, upper limb, stroke, severity levels, motor impairment, activity limitation

Introduction

Approximately 33–66% of stroke survivors experience upper limb motor impairment and activity limitation in the acute phase (Langhorne et al., 2011). The severity of initial motor impairment is a predictor for recovery (Stinear et al., 2017) and informs clinical decision-making (Intercollegiate Stroke Working Party, 2023). Severity levels of upper limb motor impairment and activity limitation following stroke can be classified in a broad spectrum from minimal to full paresis (Lang et al., 2013). However, terminology to describe severity levels is heterogeneous. Classifications vary from the use of different levels (e.g., 'mild, moderate, severe' paresis, Brunner et al., 2017), to strata based on function (e.g., 'moderate or severe upper limb functional limitation', Rodgers et al., 2019), to impairment localization (e.g., proximal-distal impairment, whole arm, Daly et al., 2019). This contributes to a lack of uniformity in clinical language. A lack of standardization also complicates data pooling for systematic reviews and

[1]Department of Allied Health Professions, Glasgow Caledonian University, Glasgow, UK
[2]Department of Allied Health Professions, School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, UK
[3]Department of Nursing, Community Health Glasgow Caledonian University, Glasgow, UK
[4]School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, UK
[5]Health Sciences Stirling, University of Stirling, Stirling, UK

Corresponding Author:
Fiona Coupar, Govan Mbeki Building, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA, UK.
Email: Fiona.coupar@gcu.ac.uk

meta-analyses, the development of evidence-based clinical guidelines, clinical decision making, and the design of precision rehabilitation interventions.

Existing literature defining UL severity levels is sparse. A search for studies that classified levels of severity of motor impairment and activity limitation following stroke, as well as the International Prospective Register of Systematic Reviews (PROSPERO) database (PROSPERO, 2024) for review protocols revealed no systematic reviews (published or ongoing) on the topic. However a number of studies (Hoonhorst et al., 2015; Woodbury et al., 2013; Woytowicz et al., 2017) have used various statistical methods to define severity levels for upper limb motor impairment. Woodbury et al. (2013) included n = 512 stroke survivors within 0 to 145 days of stroke onset using the Item Response Theory (IRT) to calculate criterion-based and mathematically derived cut-off scores for severity levels of upper limb impairment based on the Fugl-Meyer Upper Extremity section FMA-UE. However, the authors acknowledged that the IRT model may have considerable error when used for individuals at the extremes of the distribution (Woodbury et al., 2013). Further, their sample was limited to individuals in early recovery stages (i.e., acute or subacute), reducing the generalizability of their findings (Woodbury et al., 2013).

Hoonhorst et al. (2015) involved a cohort of n = 460 patients at 6 months post-stroke, and employed Receiver Operating Characteristic (ROC) analysis, aligning FMA-UE and Action Research Arm Test (ARAT) scores to predict upper-limb capacity at six months post-stroke. The applicability of their results may therefore be limited to this time point, and further work is needed to determine optimal (FMA-UE) cut-off scores to predict upper-limb capacity categories for other post-stroke stages (Hoonhorst et al., 2015).

Woytowicz et al. (2017) included n = 247 participants in the chronic stage post-stroke (i.e., ≥6 months Kwakkel et al., 2017) using cluster analysis, also based on the FMA-UE (Woytowicz et al., 2017). However, because their study population was not equally distributed over the different categories, their findings might be biased towards chronic stroke survivors with moderate-to-severe impairment levels, restricting applicability of the findings.

Although characterized by rigorous statistical methods or mathematical models, each of the aforementioned studies (Hoonhorst et al., 2015; Woodbury et al., 2013; Woytowicz et al., 2017) highlighted limitations in defining severity cut-offs that were related to their sample size, clinical demographics and/ or methods used. Comparing or amalgamating these cut-offs would not be appropriate, given that the study populations were at different stages of recovery, exhibited varying levels of severity, with different methods used to determine cut-off points. Results from these studies (Hoonhorst et al., 2015; Woodbury et al., 2013; Woytowicz et al., 2017) highlight the need for further standardization of severity levels, by involving larger and more representative samples, refining and validating cut-off scores, and using data across all post-stroke stages.

From a biomedical perspective, upper limb function and activity limitation following stroke are influenced by the ability to control muscle force, as well as various other factors (e.g., spasticity, sensory deficits, pain). This review focuses on the voluntary aspect of upper limb motor function and activity, which involves intentional motor control, ranging from gross (e.g., abducting the shoulder) to fine motor control (e.g., writing and buttoning a shirt), (Lastash & Lestienne, 2006). Therefore, in this review, the term '*upper limb motor impairment and activity limitation*' refers to the reduced ability to intentionally control any of the upper limb joints, including hand and fingers.

The aim of this systematic review is to synthesize published descriptors of severity levels of upper limb motor impairment and activity limitation after stroke, to inform the development of a standardized clinical language.

Methods

A systematic review was undertaken, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Page et al., 2021). We established the review methods *a priori*, and registered our protocol on the International prospective register of systematic reviews (PROSPERO) (Ierardi et al., 2020).

Eligibility Criteria

Eligible studies included: participants of any sex, aged 18 years or older with a clinical diagnosis of stroke (or at least 75% of the participants diagnosed with stroke); primary research studies published in English of any quantitative design; descriptions of severity levels of upper limb motor impairment and/or activity limitation after stroke. Table 1 summarizes the inclusion and exclusion criteria.

Literature Searches

We searched CINAHL, AMED, MEDLINE, the Cochrane Library, ProQuest Health & Medical Collection, ProQuest Nursing & Allied Health Database, ProQuest Public Health Database and PEDro from inception to February 2024. The complete search strategies for all databases are available in an online depository (Ierardi, 2024a).

The search strategy was piloted prior to running the finalized searches (MacFarlane et al., 2022). Only studies written in English were searched. In consultation with a specialist academic librarian, built-in filters were not used (Ierardi & McCormick, 2020). Instead, we created a dedicated search string combining all key terms to exclude

Table 1. Overview of the Inclusion and Exclusion Criteria for This Systematic Review.

Inclusion Criteria	Exclusion Criteria
Participants and diagnosis	
• Participants of any gender; • Participants of aged 18 or more; • Individuals with stroke rather than patients with stroke (if at least 75% were stroke survivors).	• Non-human participants; • Participants younger than 18 years of age. • Conditions other than stroke or if less than 75% of participants were stroke survivors.
Study type	
• Primary research; • Quantitative study design; • Either intervention or non-intervention studies.	• Secondary research; • Qualitative study design or mixed methods
Impairment description	
• Any description of severity levels of upper limb motor impairment and activity limitation of any severity (e.g.,: 'mild to moderate' paresis, Borstad & Nichols-Larsen, 2016), or 'moderately recovered or well-recovered subgroups' (de Niet et al., 2007)	• Studies not providing a description of severity levels of upper limb motor impairment and activity limitation
Outcome measures	
Outcome measures related to upper limb, classified by International Classification of Functioning, Disability and Health (ICF) domain (Duncan Millar et al., 2021): • Impairment, e.g., Motor Assessment Scale; Medical Research Council Scale; Box and Blocks Test; • Activity limitation, e.g., Wolf Motor Function Test; Action Research Arm Test; Fugl-Meyer Assessment, etc.	• Any outcome measure not targeted at severity levels of upper limb motor impairment and activity limitation (e.g., Modified Ashworth Scale (MAS); or Hand Active Sensation Test (HASTe).
Language	
English only	Non-English studies

ineligible papers with the 'NOT' Boolean operator. Finally, we manually searched the eligible articles' reference lists for any further results.

Selection of Papers

We removed duplicates with Endnote (The EndNote Team, 2013). One author conducted the title screening phase and a second author cross-checked a random 10% of the titles.

We designed and applied a validated algorithm written in Python that enabled computer-aided identification of keywords within each paper, which reduced abstract screening time by 73% on average (Ierardi et al., 2023). Details of this method have been described in full elsewhere (Ierardi et al., 2023). Two independent review authors were involved at the abstract and full-text screening phases. Any discrepancies were resolved by a third reviewer. Reasons for rejection were recorded.

Precision of Data Reporting: The Traffic Light System

There were no appropriate critical appraisal tools for studies describing severity levels of upper limb motor impairment and activity limitation when this review was undertaken, although various tools (Cardiff University, 2023; Critical Appraisal Skills Programme (CASP), 2025; Nasser, 2020; UK EQUATOR Centre, 2023) are available to evaluate general methodological and reporting quality. Therefore, we developed a bespoke tool based on traffic light colors (adapted from Novak et al., 2020), to indicate the precision of data reporting of the severity descriptors. We classified each eligible paper based on the presence or absence of each of the following core elements, associated with descriptors of severity levels, as summarized in Table 2. Classification, using this traffic light tool, was undertaken by two independent co-authors. Findings were discussed until agreement was reached- with input from a third co-author if required.

The studies included were classified as condensed in Table 3.

Papers presenting information only as graphs or images were not eligible.

Data Extraction Tool

Two independent reviewers extracted data using a piloted, purpose-made spreadsheet to guarantee transparency and reproducibility (Godino, 2023). Discrepancies were discussed between reviewers until full agreement was found. Table 4 summarizes the contents of the data extraction tool.

If a paper was marked 'red', only the following information was extracted: author, year; study design; severity descriptor; number of participants, gender and time since onset (Kwakkel et al., 2017). We adopted Kwakkel et al.'s (2017) classification of post-stroke stages.

If a paper was marked 'amber', outcome measure and participant characteristics were also extracted. If a paper was

Table 2. Key Components for Describing Upper Limb Severity in Stroke Studies.

Component	Description
Severity Descriptor	• A descriptor of severity of upper limb motor impairment and/or activity limitation (e.g., low and high, or mild-moderate and severe).
Outcome Measure	• Published measures (either recommended-SMART (Duncan Millar et al., 2021) or not), e.g., Fugl-Meyer Assessment (FMA-UE), Action Research Arm Test (ARAT), Motricity Index (MI), used to assess stroke-related upper limb function.
Numerical Data	• Numerical data, i.e., quantitative expression of impairment/activity limitation with central tendency (e.g., median) and/or dispersion (e.g., interquartile range).
Participant Characteristics	• Clinical demographics or task-specific criteria related to upper limb status (e.g., "severe UE impairment and mild/moderate UE impairment were included if they were able to perform a drinking task with their affected arm", Persson et al., 2015)

marked 'green', all of the above information were extracted, in addition to numerical values relating to severity descriptors.

Data Analysis

We developed a bespoke algorithm that was robust and transparent to enable replication. Previous work established the recommended assessment instruments to measure upper limb impairment and activity limitation (Duncan Millar et al., 2021). We presented our data in light of these recommendations. The method for selecting papers for analysis and synthesis is described below and summarized in Figure 2, together with the results.

Categorizing Precision of Data Reporting. In order to describe the precision of data reporting within the body of included articles, we calculated the percentage of articles classified as 'red', 'amber', or 'green'. We then selected only the proportion of 'green' papers for further analysis, considering such articles to be of the highest precision of data reporting. As for 'amber' and 'red' papers, these remained in the narrative synthesis for describing severity levels.

Categorizing Types of Severity Level Descriptors. Because of the high variability of severity descriptors, we grouped the articles categorized as 'green' into the most commonly reported categories that emerged from this literature. These comprised: mild, and/or moderate, and/or severe; or: high/moderate/low functioning; or: prognosis/recovery descriptors. Within the mild and/or moderate and/or severe category, if an article mentioned only one or more (e.g., 'mild' (Sangole & Levin, 2009)) severity level descriptor, it was placed in the 'mild, moderate, severe' category. Descriptors for residual functioning, for example 'high functioning group' (FMA-UE = 45–66) and 'low functioning group' (FMA-UE = 10–27) (Raghavan et al., 2020), were grouped into the 'high/low' functioning category. Descriptors based on prognosis or recovery were grouped in the 'prognosis- or recovery-based' category (e.g., ARAT scores dichotomized "into '1' for those who regained some dexterity, $\geq$ 10 ARAT and '0' for those who did not regain any dexterity, < 10 ARAT" (Nijland et al., 2010)).

Identification of Cut-off Values for Each Severity Level. Of the "green" papers, we identified those that aligned with previous recommendations and presented those results (Duncan Millar et al., 2021). Then we selected those with discrete severity levels, defined by cut-off points (e.g., "…cut-offs indicating severe (FMA-UE 0-31), moderate (FMA-UE 32-57), and mild (FMA-UE 58-66) motor impairment" (Ghaziani et al., 2020)). Papers that did not establish the exact demarcation between 'moderate' and 'severe', could not be included in the analysis (e.g., "moderate to severe impairment of the upper limb defined by a score between 7 and 50 on the FMA-UE" (Menezes et al., 2018)).

Statistical Analysis. We selected only 'Green' papers that used one or more of the most commonly reported descriptors, accompanied by numerical data that expressed discrete severity levels in cut-off scores of single, full, recommended outcome measures (i.e., not composite outcome measures or sub-scales of measures) to enable data pooling (Figure 2). Data extracted indicated skewed distributions, which necessitated non-parametric descriptive statistics. We used Microsoft Excel to calculate the medians and interquartile ranges for the upper and lower boundary of each descriptor (i.e., the median and interquartile range for the lower limit of the 'mild' category, followed by the same for the upper limit of the 'mild' category etc.).

Results

Identification of Eligible Studies

The study selection process is summarized in the PRISMA 2020 flow diagram (Page et al., 2021) in Figure 1. The database searches resulted in a total of n = 17,273 hits, n = 750 papers were included. Manual screening of the eligible articles' reference lists did not retrieve any further results.

Table 3. Classification of Studies Based on Precision Data Reporting.

Category	Criteria	Example
Green (highest precision of data reporting)	Uses a (SMART =Standardizing Measurement in Arm Rehabilitation Trials (Duncan Millar et al., 2021)). outcome measure and provides numerical data.	"Individuals with severe upper limb paresis that presented a score below 30 on the Fugl-Meyer test were considered severely impaired."(Pereira et al., 2015)
Amber (intermediate precision of data reporting)	More than a descriptor only (as in the 'red' category), but lacking one or more criteria to qualify as a 'green' category.	Favorable prognosis: able to voluntarily extend the thumb and/or 2 or more fingers of the affected hand (10° or more), unfavorable prognosis: unable to voluntarily extend the thumb and/or 2 or more fingers of the affected hand (10° or more).(Kwakkel et al., 2016)
Red (lowest precision of data reporting)	Only provides a qualitative descriptor without numerical data or validated outcome measures.	"Severe, moderate, mild recovery groups"(Hui Yung Ang & Wai Kwong Man 2006)

Table 4. Data Extraction Tool (SMART = Standardizing Measurement in Arm Rehabilitation Trials (Duncan Millar et al., 2021)).

Information	Details
Author, Year	
Reference	
Traffic light color	Red Amber Green
Outcome measure	If SMART (Duncan Millar et al., 2021): select from drop down list any of the SMART (Duncan Millar et al., 2021) outcome measures. If non-SMART (Duncan Millar et al., 2021): select from drop down list
Study design	Intervention non-intervention
Severity descriptors	Minimized free text, i.e., "Very low, low, medium and high"(Edwards et al., 2019)
Score range	0 = not reported 1 = reported
Score range. Details, per level	Minimized free text, e.g., 10–29 = severe, 30–49 = moderate (Daly et al., 2005)
Number of included participants	
Total FEMALE	
Total MALE	
Time since onset	Early sub-acute 0–3 months (Kwakkel et al., 2017); Late sub-acute 3–6 months (Kwakkel et al., 2017); Chronic >6 months (Kwakkel et al., 2017); Mixed; Not reported
Rationale for defining the descriptor of the categories	0 = not reported; 1 = reported
Descriptor referenced	0 = not reported; 1 = reported

The complete lists of included and excluded full texts and detailed reasons for exclusion are available in an online repository (Ierardi, 2024b, 2024c). Moreover, data extracted from eligible papers are also accessible from the same resource (Ierardi et al., 2024a).

Characteristics of Included Studies

Of the eligible n = 750 articles, n = 419 had an intervention study design (55.8%) and n = 332 had a non-intervention study design (44.2%) and were published within the last 45 years. According to our traffic light system, n = 458 (61%) of included papers were of a high data reporting precision ('green'); n = 264 (35%) were of intermediate ('amber'), and n = 28 (4%) were of low data reporting precision ('red').

Study Population

The eligible articles involved a total of n = 36,429 stroke participants: n = 13,236 (36%) female; n = 13,236 (56%) male, while in n = 2,796 (8%) of participants their sex was not reported. The review included n = 13,884 (27%) of stroke survivors in the early sub-acute stage, n = 464 (2%) in the late sub-acute; n = 12,280 (52%) in the chronic stage, and n = 9,165 (18%) categorized as mixed stages, with an additional n = 636 (1%) left unreported.

Severity Descriptors

The most commonly used descriptors were mild, and/or moderate, and/or severe, adopted in n = 580 papers (77%), while the remaining 23% of articles used either prognosis-, recovery-based (n = 135 papers, 18%), or high-low functioning descriptors (n = 36 papers, 5%).

Outcome Measures

Out of the n = 750 included papers, n = 477 (63.5%) used a recommended outcome measure (Duncan Millar et al., 2021) to describe severity levels of upper limb motor impairment and activity limitation, while n = 246 (32.8%) used other outcome measures and n = 28 studies (4% of included articles) did not include any outcome measures.

From fifteen previously recommended outcome measures (Duncan Millar et al., 2021), we identified seven in our review. The: Fugl-Meyer Upper Extremity section (FMA-UE), was most commonly used (n = 383 papers; 51.0% of eligible papers) followed by the Action Research Arm Test (ARAT), (n = 42) 5.6%, the Box and Blocks Test (BBT) (n = 17 papers; 2.3%), Motricity index-Upper limb section (n = 15 papers; 2.0%), Motor Activity Log (MAL), (n = 12 papers; 1.6%), Wolf Motor Function Test (WMFT = 7), (n = 7 papers; 0.9%) and 9-Hole Peg Test, (n = 1 paper; 0.1%).

Papers Providing a Reference for Defining the Descriptor

Only n = 35 (4.7%) papers provided a reference to support their choice for severity level descriptors.

Statistical Analysis

Figure 2 summarizes the findings from the statistical analysis in this systematic review.

We screened the n = 750 eligible papers against the traffic light color, retaining n = 458 'green' studies. Among these, we examined the n = 360 studies that utilized the most commonly reported, discrete descriptors which were 'mild, and/ or moderate, and/or severe'. We then focused on the n = 342 papers which used the three most common recommended outcome measures (Duncan Millar et al., 2021) (FMA-UE, n = 308; ARAT, n = 23; MI, n = 11). We classified n = 281 papers ineligible for statistical analysis. We were able to include n = 61 papers in the statistical analysis (i.e., those reporting severity descriptors in terms of the FMA-UE, n = 57; MI, n = 2; ARAT, n = 2). However, there were not enough data to analyze data from the studies using the MI and the ARAT.

Severity Descriptors Defined by FMA-UE Cut-off Values

A total of n = 57 papers (n = 2,828 participants) employed the FMA-UE to define severity levels with cut-off values, representing 17% of the n = 342 'green' papers which used mild, and/or moderate and/or severe descriptors—i.e., n = 61 (8%) of all studies included in this review. The complete reference list of the 57 papers eligible for the statistical analysis is reported in an online repository (Ierardi et al., 2024b).

According to Fugl Meyer et al. (Fugl-Meyer et al., 1975) the minimum score on the FMA-UE is 0 and the maximum is 66 points. Figure 3 describes the range of FMA-UE values for each severity category. The severe category ranged from 0 (IQR: 0-0) to 25 (IQR: 10-35). The moderate category ranged from 26 (IQR: 15-34) to 50 (IQR: 38-59). The mild category ranges from 51 (IQR: 41-60) to 66 (IQR: 65-66).

However, there was considerable overlap in the demarcation of these categories (Figure 3), highlighting the challenge of establishing clear severity boundaries, even within a commonly used assessment tool like the FMA-UE.

Severity Descriptors Defined by ARAT Cut-off Values

Only two papers (Brunner et al., 2024; Waddell et al., 2016) (n = 105 participants) reported specific cut-off values of the ARAT, representing 0.6% of the n = 342 'green' papers which used discrete severity descriptors. Brunner et al. (2024) only indicated one level of impairment severity (severe) with cut-off values ranging from 0 to 13. Waddell et al. (2016) provided cut-off values for each severity level: their median 'severe' level ranged from 0 to 20, 'moderate' ranged from 21 to 39 and 'mild' ranged from 40 to 57. Insufficient data for this outcome measure precluded further analysis.

Severity Descriptors Defined by Motricity Index Cut off Values

Only n = 2 papers (Goffredo et al., 2019; Platz et al., 2009) (n = 212 participants) used the MI to define only two severity descriptors (mild and severe) representing 0.6% of the 'green' papers eligible for the statistical analysis. The 'severe' category ranged from 0 to 48, while the 'mild' category ranged from 49 to 99. Insufficient data for this outcome measure precluded further analysis.

Upper and Lower Boundaries for 'Mild, Moderate and Severe' Descriptors Defined by the Top-3 Recommended Outcome Measures

Table 5 summarizes the key findings of this review, including the lower and upper boundaries for each severity level based on one of the top-3 recommended (Duncan Millar et al., 2021) outcome measures, using data from 61

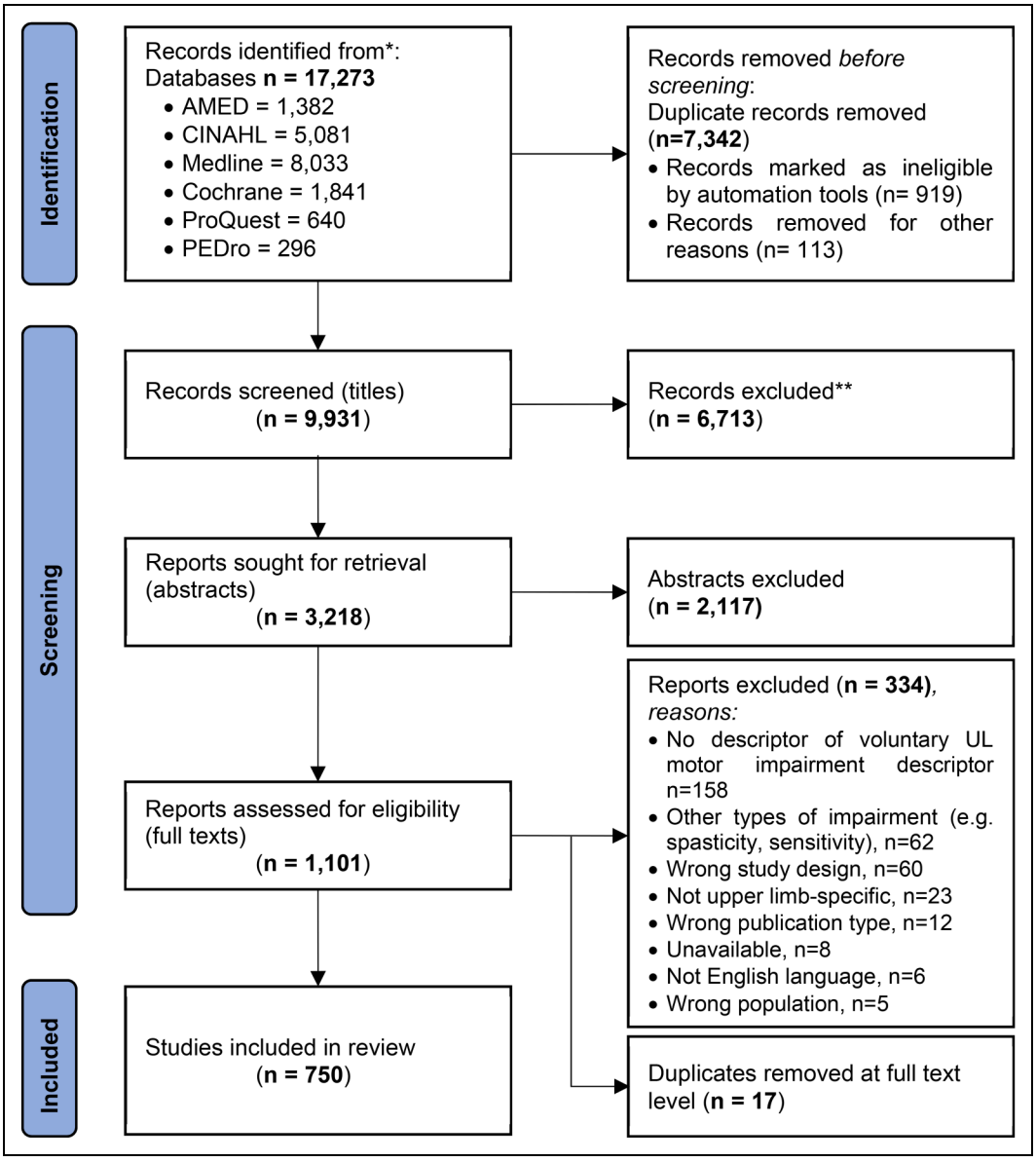

Figure 1. PRISMA 2020 Flow Diagram (Page et al., 2021) of the Study Selection Process for the Systematic Review.

papers (8% of articles included in this review) and 3,145 participants.

Of the 57 studies that used the Fugl-Meyer Assessment for the Upper Extremity to define levels of severity, 24 cited one or more references to justify their choice of cut-off scores. After removing duplicates, 13 unique papers remained (see Supplementary Table X). These can be categorized broadly into three categories: the first category of cited papers did not include any FMA-UE cut-offs (Anemaet, 2002; Gladstone et al., 2002; Kim, 2021; Nijland et al., 2010, 2013). The second category of cited papers provided cut-offs but no rationale (Michielsen et al., 2012); (Lum et al., 2002) or a reference to a paper reporting cut-offs (Velozo & Woodbury, 2011). The third category of cited papers provided cut-offs and various types of rationale, only four of which presented their underlying analysis; Luft et al. (2004) suggested cut-offs based on clinical observation, but did not provide any further information (Luft et al., 2004). Hoonhorst et al. (2015) based their cut-offs on Receiver Operating Characteristic curves from n = 460 stroke survivors (Hoonhorst et al., 2015). Kwakkel et al. (2003) developed a prognostic model using logistic regression analysis, based on factors such as upper limb dexterity and severity of paresis measured with the Action Research Arm Test, Motricity Index, and Fugl-Meyer motor scores (Kwakkel et al., 2003), involving n = 102 stroke patients. Woodbury et al. (2013)—the most commonly cited by 5 papers—employed Rasch analysis on the data of n = 512 to mathematically define cut-offs based on specific behavioural criteria

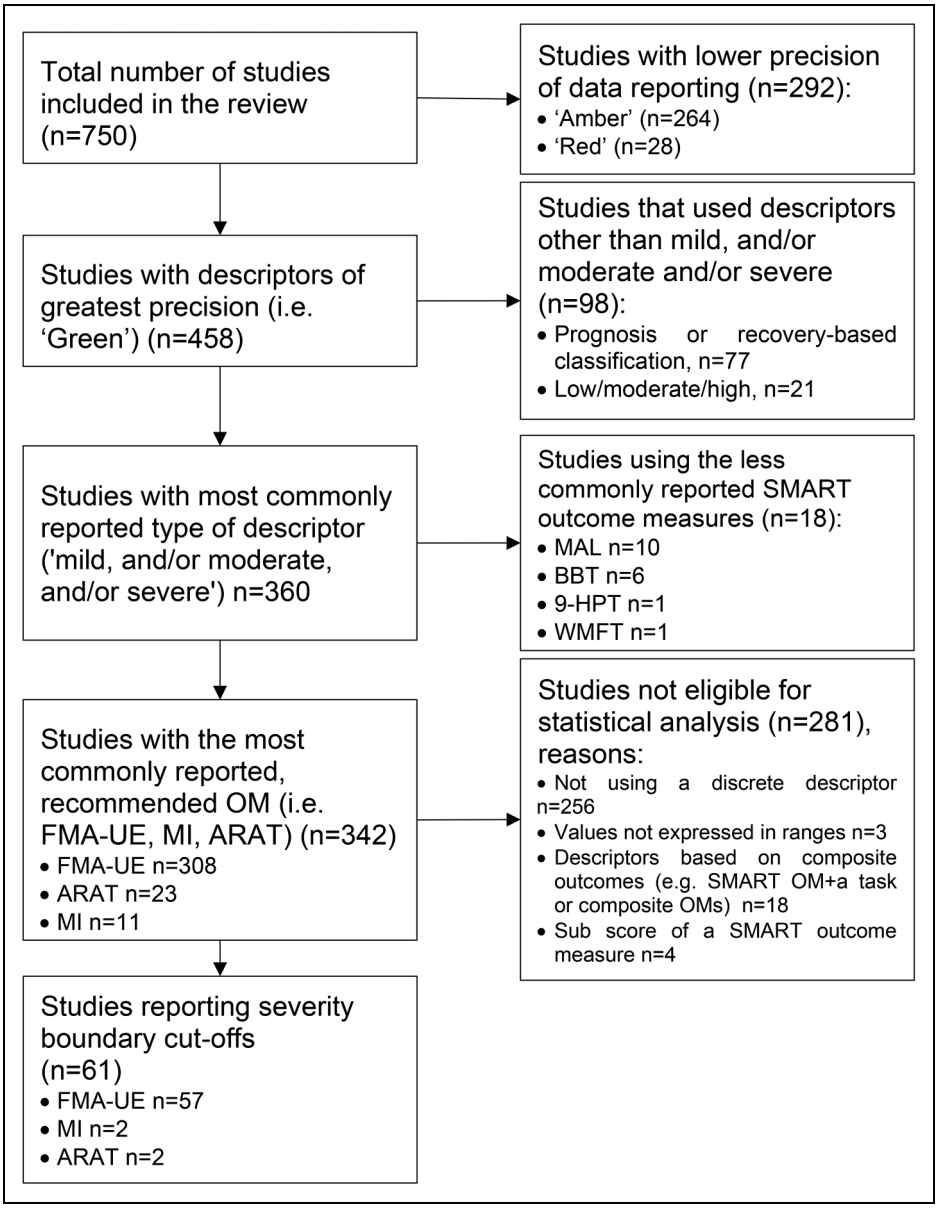

Figure 2. Decision Tree Diagram Outlining the Selection Process of Studies for Statistical Analysis in This Systematic Review and Findings. Papers Were Categorized Using a Bespoke Traffic Light System on the Precision of the Severity Level Descriptors: 'Green' Studies Reported a Recommended (Duncan Millar et al., 2021) Outcome Measure and Numerical Data: 'Red' Studies Included a Severity Level Descriptor Only; 'Amber' Studies Were Those not Classified as 'Green' or 'Red'. Abbreviations: OM = Outcome Measure; MAL = Motor Activity Log; BBT = Box and Block Test; 9-HPT = Nine-Hole peg Test; WMFT = Wolf Motor Function Test; FMA-UE = Fugl-Meyer Assessment-Upper Extremity; ARAT = Action Research Arm Test; MI = Motricity Index (Upper Limb Section); OM: Outcome Measure; SMART = Standardizing Measurement in Arm Rehabilitation Trials (Duncan Millar et al., 2021).

(Woodbury et al., 2013). Woytowicz et al. (2017) derived cut-off points based on clinical observation from a sample of n = 247 patients, followed by hierarchical cluster analysis (Woytowicz et al., 2017).

Discussion

We conducted the first systematic review of the literature that synthesized published severity level descriptors of upper limb motor impairment and activity limitation following a stroke. We developed a bespoke text-mining tool to accelerate abstract screening (Ierardi et al., 2023) and identified n = 750 eligible full texts.

Our review covered 45 years of evidence, highlighting the use of a broad spectrum of severity level descriptors of upper limb motor impairment and activity limitation, along with the use of a wide range of different outcome measures and cut-off values for each severity level.

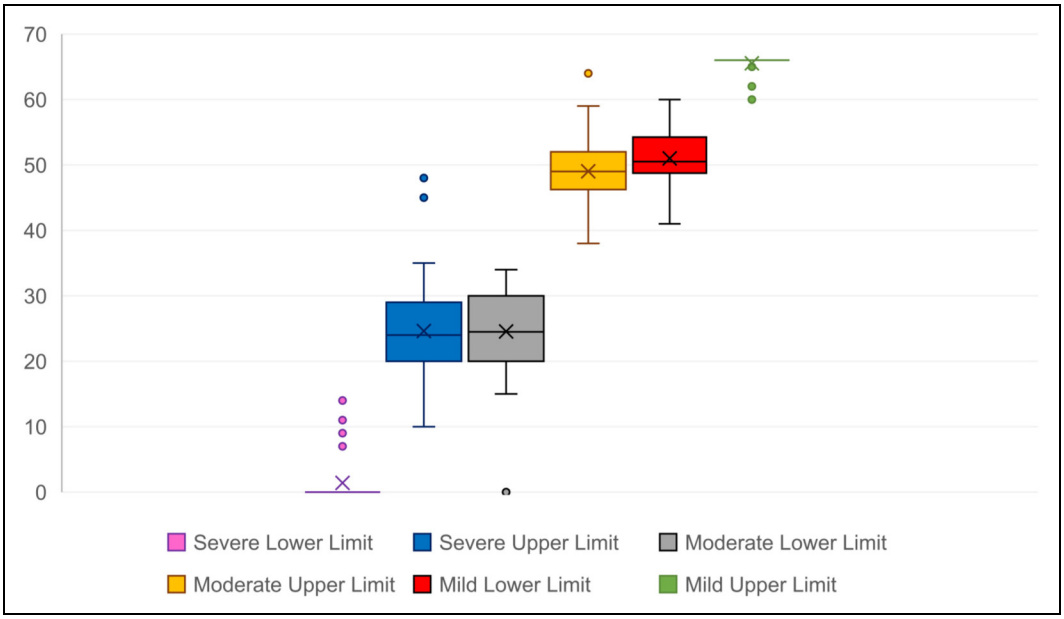

Figure 3. Box-and-Whisker Plot of Levels of Severity Defined by Fugl Meyer Assessment – Upper Extremity (FMA-UE) (Fugl-Meyer et al., 1975) Values (Ranges): Severe 0 (IQR: 0-0) to 25 (IQR: 10-35); Moderate: 26 (IQR: 15-34) to 50 (IQR: 38-59); Mild Ranges from 51 (IQR: 41-60) to 66 (IQR: 65-66).

Findings of our review showed a clear lack of standardization of severity level descriptors of upper limb motor impairment and activity limitation following stroke. The most frequently used descriptors are based on the classification of 'mild', 'moderate', and 'severe', as defined using the three most commonly recommended outcome measures: FMA-UE, ARAT, and MI.

The number of studies and participants contributing to each estimate are as follows: FMA-UE (57 papers, n = 2,828), ARAT (2 papers, n = 105), and MI (2 papers, n = 212). There were only sufficient data to derive cut-off values related to each severity level for the FMA-UE, based on a small percentage (8%) of all papers (n = 57 out of n = 750 total papers), which means that these severity level descriptors are currently dominated by this outcome measure only. Furthermore, there is an indication of blurring between the boundaries of severity levels within these articles. Such blurring may influence the consistency and comparability of severity levels across studies and suggests a need for further refinement of cut-off values. It will be important to consider the implications of statistical overlap between adjacent severity categories for clinical decision making. This would require further input from clinicians, which we recommend as a next stage in the development of the proposed severity categories.

In summary, the cited literature demonstrates a high level of heterogeneity in how FMA-UE severity cut offs are defined, while the justification of any cut-offs is based on a very small evidence base (n = 1,321 participants) from four studies (Hoonhorst et al., 2015; Kwakkel et al., 2003; Woodbury et al., 2013; Woytowicz et al., 2017). These inconsistencies suggest a need for more standardised, consensus-based criteria to support cross-study comparability and enhance the interpretation of outcomes in rehabilitation research.

The main findings of our systematic review highlight the urgent need for standardizing levels of severity of motor impairment and activity limitation, emphasizing the importance of agreed definitions and cut-off values based on recommended outcome measures.

Implications for Research and Clinical Practice

Future studies should aim to develop consensus, based on this evidence-based classification, as a standardized classification will provide researchers and clinicians with a common language and framework for categorizing severity levels of upper limb motor impairment and activity limitation. Such a framework has potential to improve stratification of severity levels and associated eligibility criteria for patients in future clinical trials of upper limb rehabilitation interventions, results of which will improve clinical guidelines. Improved stratification would also enable better informed clinical decision making about effective interventions tailored to specific severity levels, facilitating person-centred precision rehabilitation post-stroke (French et al., 2022).

Strengths and Limitations of This Review

Our systematic review (including n = 36,429 participants in total, of which n = 3,145 were included in the statistical

Table 5. Summary of key Findings: Lower and Upper Boundaries for Each Level of Impairment – or Activity Limitation Identified from the 62 Articles Using the top-3 SMART Outcome Measures (Duncan Millar et al., 2021); Number of Studies, Number of Participants. (FMA-UE = Fugl-Meyer, Upper Extremity Section; ARAT = Action Research Arm Test; MI = Motricity Index; IQRs = Interquartile Ranges; # Denotes Insufficient Data to Establish Boundary Median and IQR.

Measure	Severe Boundaries	Moderate Boundaries	Mild Boundaries	Number of Studies	Number of Participants
FMA-UE Lower median [IQR]	0 [0–0]	26 [15–34]	51 [41–60]	58	2,828
FMA-UE Upper median [IQR]	25 [10–35]	50 [38–59]	66 [65–66]		
ARAT#	0–16.5	21–39	40–57	2	105
MI#	0–48	No data available	49–99	2	212

analysis) builds on previous studies (Hoonhorst et al., 2015; Woodbury et al., 2013; Woytowicz et al., 2017), each of which employed different statistical methods to define severity levels for upper limb motor impairment and/ or activity limitation, based on individual patient data. As discussed earlier, while these studies highlighted the need for and contributed to the effort to standardize severity levels, they were limited by potential errors at the extremes of the data distribution (Woodbury et al., 2013), the reliance on data from single time points (Hoonhorst et al., 2015) and a sample predominantly representing chronic stroke survivors (Woytowicz et al., 2017) which limit the applicability of these results to a more representative stroke population. In contrast, our review includes data from a wider range of time points and severity levels, capturing a larger and more diverse patient population. However, even though this evidence synthesis comprised a sample of n = 36,429 stroke participants, there are limitations to its generalisability, as only 36% were female, 2% were in the late-subacute stage and 2% (n = 631) were described as having a mild level of severity or their upper limb impairment/ activity limitation.

We undertook this systematic review following rigorous methods to ensure full transparency and replicability: our protocol was registered before data extraction and analysis (Ierardi et al., 2020). We developed thorough searches consulting different specialist databases, employing advanced tools for managing references (Endnote™ (The EndNote Team, 2013), Rayyan (Ouzzani et al., 2016), a bespoke text-mining tool (Ierardi et al., 2023)). Two independent reviewers were involved at all screening and data extraction stages. We ensured a comprehensive synthesis of available evidence including all eligible papers, regardless of publication date.

Our work presents some limitations however: due to resource limitations, we restricted our searches to English-only publications, acknowledging the exclusion of any relevant non-English studies. We did not undertake any statistical or mathematical modeling of individual patient data, as this was not appropriate given the nature and distribution of the data included in our systematic review. This was the first systematic review that synthesized group-level data from published studies, providing an overview of the literature to capture the variations of definitions of severity levels, and highlighting areas where further refinement is needed. In future, Individual Patient Data—provided these follow recommendations for the type and timing post stroke (Duncan Millar et al., 2021; Kwakkel et al., 2017)—could be used to validate the severity classification described in this review. Moreover, Individual Patient Data analysis could enable the determination of more precise cut-offs, offer insights into severity classifications associated with specific clinical demographic variables and enable the detection of recovery trends over time.

Conclusions

In conclusion, the literature analyzed in this systematic review demonstrates a high level of heterogeneity in how FMA-UE severity cut offs are defined and justified. These inconsistencies suggest a need for more standardized, consensus-based criteria to support cross-study comparability and enhance the interpretation of outcomes in rehabilitation research.

Future research should first of all develop consensus on the classification presented in this review. Validating severity levels and cut-off values would not only improve the quality of stroke rehabilitation research but also support the development of precision rehabilitation for stroke survivors.

Acknowledgments

We are thankful to the late Chukwudi Martin Ogbueche (Glasgow Caledonian University) for his contribution to the article screening and data extraction phases, and to Emeritus Prof. Chris Eilbeck (Heriot-Watt University) for developing the data-mining tool to accelerate abstract screening for this systematic review.

ORCID iDs

Dr. Elena Ierardi https://orcid.org/0000-0002-7342-3958
Prof. Frederike van Wijck https://orcid.org/0000-0003-0855-799X
Dr. Myzoon Ali https://orcid.org/0000-0001-5899-2485

Ethical Approval and Informed Consent Statements

This systematic review did not involve the collection of new data from human or animal subjects, and therefore no ethical approval was required.

Funding

Dr. Elena Ierardi received a PhD studentship from Glasgow Caledonian University. The authors received no other financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interest

The authors declare no potential conflicts of interest.

Data Availability Statement

The authors are committed to facilitating openness, transparency and reproducibility of this research. All data extracted for this systematic review can be accessed from within the main article (directly or via links provided), or supplementary material.

Supplemental Material

Supplemental material for this article is available online.

References

Anemaet, W. K. (2002). Using standardized measures to meet the challenge of stroke assessment. *Topics in Geriatric Rehabilitation*, *18*(2), 47–62. https://doi.org/10.1097/00013614-200212000-00006

Borstad, A., & Nichols-Larsen, D. S. (2016). The brief kinesthesia test is feasible and sensitive: A study in stroke. *Brazilian Journal of Physical Therapy*, *20*(1), 81–86. https://doi.org/10.1590/bjpt-rbf.2014.0132

Brunner, I., Lundquist, C. B., Pedersen, A. R., Spaich, E. G., Dosen, S., & Savic, A. (2024). Brain computer interface training with motor imagery and functional electrical stimulation for patients with severe upper limb paresis after stroke: A randomized controlled pilot trial. *Journal of NeuroEngineering and Rehabilitation*, *21*(1), 1–12. https://doi.org/10.1186/s12984-024-01304-1

Brunner, I., Skouen, J. S., Hofstad, H., Aßmus, J., Becker, F., Sanders, A. M., Pallesen, H., Kristensen, L. Q., Michielsen, M., Thijs, L., & Verheyden, G. (2017). Virtual reality training for upper extremity in subacute stroke (VIRTUES): A multicenter RCT. *Neurology*, *89*(24), 2413–2421. https://doi.org/10.1212/WNL.0000000000004744

Cardiff University. (2023). *The Specialist Unit for Review Evidence (SURE)*. Online. https://www.cardiff.ac.uk/specialist-unit-for-review-evidence

Critical Appraisal Skills Programme (CASP). (2025). *Critical Appraisal Skills Programme (CASP)*. Online. https://casp-uk.net/

Daly, J. J., Hogan, N., Perepezko, E. M., Krebs, H. I., Rogers, J. M., Goyal, K. S., Dohring, M. E., Fredrickson, E., Nethery, J., & Ruff, R. L. (2005). Response to upper-limb robotics and functional neuromuscular stimulation following stroke. *Journal of Rehabilitation Research and Development*, *42*(6), 723–736. https://doi.org/10.1682/JRRD.2005.02.0048

Daly, J. J., McCabe, J. P., Holcomb, J., Monkiewicz, M., Gansen, J., & Pundik, S. (2019). Long-Dose intensive therapy is necessary for strong, clinically significant, upper limb functional gains and retained gains in severe/moderate chronic stroke. *Neurorehabilitation and Neural Repair*, *33*(7), 523–537. https://doi.org/10.1177/1545968319846120

de Niet, M., Bussmann, J. B., Ribbers, G. M., & Stam, H. J. (2007). The stroke upper-limb activity monitor: Its sensitivity to measure hemiplegic upper-limb activity during daily life. *Archives of Physical Medicine and Rehabilitation*, *88*(9), 1121–1126. https://doi.org/10.1016/j.apmr.2007.06.005

Duncan Millar, J., van Wijck, F., Pollock, A., & Ali, M. (2021). International consensus recommendations for outcome measurement in post-stroke arm rehabilitation trials. *European Journal of Physical and Rehabilitation Medicine*, *57*(1), 61–68. https://doi.org/10.23736/S1973-9087.20.06575-2

Edwards, D. J., Cortes, M., Rykman-Peltz, A., Chang, J., Elder, J., Thickbroom, G., Mariman, J. J., Gerber, L. M., Oromendia, C., Krebs, H. I., Fregni, F., Volpe, B. T., & Pascual-Leone, A. (2019). Clinical improvement with intensive robot-assisted arm training in chronic stroke is unchanged by supplementary tDCS. *Restorative Neurology and Neuroscience*, *37*(2), 167–180. https://doi.org/10.3233/RNN-180869

French, M. A., Roemmich, R. T., Daley, K., Beier, M., Penttinen, S., Raghavan, P., Searson, P., Wegener, S., & Celnik, P. (2022). Precision rehabilitation: Optimizing function, adding value to health care. *Archives of Physical Medicine and Rehabilitation*, *103*(6), 1233–1239. https://doi.org/10.1016/j.apmr.2022.01.154

Fugl-Meyer, A. R., Jääskö, L., Leyman, I., & Olsson, S. (1975). The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine*, *7*(1), 13–31. https://doi.org/10.2340/1650197771331

Ghaziani, E., Couppé, C., Siersma, V., Christensen, H., Magnusson, S. P., Sunnerhagen, K. S., Persson, H. C., & Alt Murphy, M. (2020). Easily conducted tests during the first week post-stroke can aid the prediction of arm functioning at 6 months. *Frontiers in Neurology*, *10*(1371), 1–13. https://doi.org/10.3389/fneur.2019.01371

Gladstone, D. J., Danells, C. J., & Black, S. E. (2002). The fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties. *Neurorehabil Neural Repair*, *16*(3), 232–240. http://search.ebscohost.com/login.aspx?direct=true&db=amed&AN=0042143&site=ehost-live&scope=site. https://doi.org/10.1177/154596802401105171

Godino, L. (2023). How to structure Microsoft Excel documents for systematic reviews. *Nurse Researcher*, *31*(1), 40–46. https://doi.org/10.7748/nr.2023.e1866

Goffredo, M., Mazzoleni, S., Gison, A., Infarinato, F., Pournajaf, S., Galafate, D., Agosti, M., Posteraro, F., & Franceschini, M. (2019). Kinematic parameters for tracking patient progress

during upper limb robot-assisted rehabilitation: An observational study on subacute stroke subjects. *Applied Bionics and Biomechanics*, *2019*(Article ID 4251089), 1–12. https://doi.org/10.1155/2019/4251089

Hoonhorst, M. H., Nijland, R. H., Van Den Berg, J. S., Emmelot, C. H., Kollen, B. J., & Kwakkel, G. (2015). How do fugl-Meyer arm motor scores relate to dexterity according to the action research arm test at 6 months poststroke? *Archives of Physical Medicine and Rehabilitation*, *96*(10), 1845–1849. https://doi.org/10.1016/j.apmr.2015.06.009

Hui Yung Ang, J., & Wai Kwong Man, D. (2006). The discriminative power of the Wolf motor function test in assessing upper extremity functions in persons with stroke. *International Journal of Rehabilitation Research*, *29*(4), 357–361. https://doi.org/10.1097/MRR.0b013e3280118e65

Ierardi, E. (2024a). *Complete search strategies run on all databases*. Online Repository. https://docs.google.com/document/d/1Jr9mWZcaYfsLz8hqm_BcXKyP0hAv7eRK/edit?usp=sharing&ouid=105408371782417730074&rtpof=true&sd=true

Ierardi, E. (2024b). *List of included papers*. Online Repository. https://drive.google.com/file/d/1SMZgMHiiQQoqoEJvvSq7DUZOqubmCfcr/view?usp=drive_link

Ierardi, E. (2024c, June 21). *List of excluded papers, with reasons*. Online Repository. https://drive.google.com/file/d/1aKDETDVDgnegcxi6sEJ2hk9BFhoBtarG/view?usp=drive_link

Ierardi, E., Coupar, F., Ali, M., & van Wijck, F. (2020). *A systematic review of descriptors of levels of severity of upper limb activity limitation and impairment after stroke*. PROSPERO International Prospective Register of Systematic Reviews. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=226244

Ierardi, E., Coupar, F., Ali, M., & van Wijck, F. (2024a, June 22). *Data extracted from eligible papers*. Online Repository. https://docs.google.com/spreadsheets/d/1oJv9iCZND2_4MvqgiV98St-tjOhnVc-W/edit?usp=sharing&ouid=105408371782417730074&rtpof=true&sd=true

Ierardi, E., Eilbeck, J. C., van Wijck, F., Ali, M., & Coupar, F. (2023). Data mining versus manual screening to select papers for inclusion in systematic reviews: A novel method to increase efficiency. *International Journal of Rehabilitation Research. Internationale Zeitschrift Fur Rehabilitationsforschung. Revue Internationale de Recherches de Readaptation*, *46*(3), 284–292. https://doi.org/10.1097/MRR.0000000000000595

Ierardi, E., & McCormick, K. (2020). Re: appointment request. In *Glasgow Caledonian University. Unpublished*.

Ierardi, E., van Wijck, F., Ali, M., Best, C., & Coupar, F. (2024b, June). *Complete reference list of the 58 papers eligible for the statistical analysis*. Online Repository. https://docs.google.com/document/d/1EPEMlricds_5A8dAXOJxNJEvj9EXHat4/edit?usp=drive_link&ouid=105408371782417730074&rtpof=true&sd=true

Intercollegiate Stroke Working Party. (2023). *National Clinical Guideline for Stroke for the United Kindgom and Ireland* (pp. 88–94). https://www.strokeguideline.org/app/uploads/2023/04/National-Clinical-Guideline-for-Stroke-2023.pdf

Kim, S. H. (2021). Effects of dual transcranial direct current stimulation and modified constraint-induced movement therapy to improve upper-limb function after stroke: A double-blinded, pilot randomized controlled trial. *Journal of Stroke and Cerebrovascular Diseases*, *30*(9), 105928. https://doi.org/10.1016/j.jstrokecerebrovasdis.2021.105928

Kwakkel, G., Kollen, B. J., Van der Grond, J. V., & Prevo, A. J. H. (2003). Probability of regaining dexterity in the flaccid upper limb: Impact of severity of paresis and time since onset in acute stroke. *Stroke*, *34*(9), 2181–2186. https://doi.org/10.1161/01.STR.0000087172.16305.CD

Kwakkel, G., Lannin, N. A., Borschmann, K., English, C., Ali, M., Churilov, L., Saposnik, G., Winstein, C., van Wegen, E. E. H., Wolf, S. L., Krakauer, J. W., & Bernhardt, J. (2017). Standardized measurement of sensorimotor recovery in stroke trials: Consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *International Journal of Stroke : Official Journal of the International Stroke Society*, *12*(5), 451–461. https://doi.org/10.1177/1747493017711813

Kwakkel, G., Winters, C., Van Wegen, E. E. H., Nijland, R. H. M., Van Kuijk, A. A. A., Visser-Meily, A., de Groot, J., de Vlugt, E., Arendzen, J. H., Geurts, A. C. H., & Meskers, C. G. M. (2016). Effects of unilateral upper limb training in two distinct prognostic groups early after stroke: The EXPLICIT-stroke randomized clinical trial. *Neurorehabilitation and Neural Repair*, *30*(9), 804–816. https://doi.org/10.1177/1545968315624784

Lang, C. E., Bland, M. D., Bailey, R. R., Schaefer, S. Y., & Birkenmeier, R. L. (2013). Assessment of upper extremity impairment, function, and activity after stroke: Foundations for clinical decision making. *Journal of Hand Therapy*, *26*(2), 104–115. https://doi.org/10.1016/j.jht.2012.06.005

Langhorne, P., Bernhardt, J., & Kwakkel, G. (2011). Stroke rehabilitation. *The Lancet (British Edition)*, *377*(9778), 1693–1702. https://doi.org/10.1016/S0140-6736(11)60325-5

Lastash, M. L., & Lestienne, F. (2006). *Motor Control and Learning*. Springer Science+Business Media, Inc. https://link.springer.com/book/10.1007/0-387-28287-4

Luft, A. R., McCombe-Waller, S., Whitall, J., Forrester, L. W., Macko, R., Sorkin, J. D., Schulz, J. B., Goldberg, A. P., & Hanley, D. F. (2004). Repetitive bilateral arm training and motor Cortex activation in chronic stroke. *JAMA*, *292*(15), 1853. https://doi.org/10.1001/jama.292.15.1853

Lum, P. S., Burgar, C. G., Shor, P. C., Majmundar, M., & Van der Loos, M. (2002). Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine and Rehabilitation*, *83*(7), 952–959. http://search.ebscohost.com/login.aspx?direct=true&db=amed&AN=0040532&site=ehost-live&scope=site. https://doi.org/10.1053/apmr.2001.33101

MacFarlane, A., Russell-Rose, T., & Shokraneh, F. (2022). Search strategy formulation for systematic reviews: Issues, challenges and opportunities. *Intelligent Systems with Applications*, *15*(200091), 1–10. https://doi.org/10.1016/j.iswa.2022.200091

Menezes, I. S., Cohen, L. G., Mello, E. A., Machado, A. G., Peckham, P. H., Anjos, S. M., Siqueira, I. L., Conti, J., Plow, E. B., & Conforto, A. B. (2018). Combined brain and peripheral nerve stimulation in chronic stroke patients with moderate to severe motor impairment. *Neuromodulation*, *21*(2), 176–183. https://doi.org/10.1111/ner.12717

Michielsen, M. E., Selles, R. W., Stam, H. J., Ribbers, G. M., & Bussmann, J. B. (2012). Quantifying nonuse in chronic stroke patients: A study into paretic, nonparetic, and bimanual upper-limb use in daily life. *Archives of Physical Medicine and Rehabilitation*, *93*(11), 1975–1981. https://doi.org/10.1016/j.apmr.2012.03.016

Nasser, M. (2020). New standards for systematic reviews incorporate population health sciences. *American Journal of Public Health*, *110*(6), 753–754. https://doi.org/10.2105/ajph.2020.305609

Nijland, R. H. M., van Wegen, E., Harmeling-van der Wel, B.,, & … Kwakkel, G., & Early Prediction of Functional Outcome After Stroke Investigators. (2013). Accuracy of physical therapists' early predictions of upper-limb function in hospital stroke units: The EPOS study. *Physical Therapy*, *93*(4), 460–469. http://search.ebscohost.com/login.aspx?direct=true&db=amed&AN=0166208&site=ehost-live&scope=site. https://doi.org/10.2522/ptj.20120112

Nijland, R. H. M., Van Wegen, E. E. H., Harmeling-Van Der Wel, B. C., & Kwakkel, G. (2010). Presence of finger extension and shoulder abduction within 72 h after stroke predicts functional recovery: Early prediction of functional outcome after stroke: The EPOS cohort study. *Stroke*, *41*(4), 745–750. https://doi.org/10.1161/STROKEAHA.109.572065

Novak, I., Morgan, C., Fahey, M., Finch-Edmondson, M., Galea, C., Hines, A., Langdon, K., Namara, M. M., Paton, M. C., Popat, H., Shore, B., Khamis, A., Stanton, E., Finemore, O. P., Tricks, A., te Velde, A., Dark, L., Morton, N., & Badawi, N. (2020). State of the evidence traffic lights 2019: Systematic review of interventions for preventing and treating children with cerebral palsy. In *Current neurology and neuroscience reports* (Vol. 20, Issue 2, pp. 1–21). Springer.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 1–10. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S.,, & … Mckenzie, J. E. (2021). PRISMA 2020 Explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *British Medical Journal*, *372*(160), 1–36. https://doi.org/10.1136/BMJ.N160

Pereira, N. D., Vieira, L., Pompeu, F. P., Menezes, I. D. S., Anjos, S. M. D., & Ovando, A. C. (2015). Translation, cultural adaptation and reliability of the brazilian version of the Graded Wolf Motor Function Test in adults with severe hemiparesis. *Fisioterapia Em Movimento*, *28*(4), 667–676. https://doi.org/10.1590/0103-5150.028.004.AO03

Persson, H. C., Alt Murphy, M., Danielsson, A., Lundgren-Nilsson, Å., & Sunnerhagen, K. S. (2015). A cohort study investigating a simple, early assessment to predict upper extremity function after stroke - a part of the SALGOT study. *BMC Neurology*, *15*(1), Article 92. https://doi.org/10.1186/s12883-015-0349-6

Platz, T., Van Kaick, S., Mehrholz, J., Leidner, O., Eickhof, C., & Pohl, M. (2009). Best conventional therapy versus modular impairment-oriented training for arm paresis after stroke: A single-blind, multicenter randomized controlled trial. *Neurorehabilitation and Neural Repair*, *23*(7), 706–716. https://doi.org/10.1177/1545968309335974

PROSPERO. (2024). *PROSPERO International prospective register of systematic reviews - Advanced Search*. https://www.crd.york.ac.uk/prospero/#searchadvanced

Raghavan, P., Bilaloglu, S., Ali, S. Z., Jin, X., Aluru, V., Buckley, M. C., Tang, A., Yousefi, A., Stone, J., Agrawal, S. K., & Lu, Y. (2020). The role of robotic path assistance and weight support in facilitating 3D movements in individuals with post-stroke hemiparesis. *Neurorehabilitation and Neural Repair*, *34*(2), 134–147. https://doi.org/10.1177/1545968319887685

Rodgers, H., Bosomworth, H., Krebs, H. I., van Wijck, F., Howel, D., Wilson, N., Aird, L., Alvarado, N., Andole, S., Cohen, D. L., Dawson, J., Fernandez-Garcia, C., Finch, T., Ford, G. A., Francis, R., Hogg, S., Hughes, N., Price, C. I., Ternent, L.,, & … Shaw, L. (2019). Robot assisted training for the upper limb after stroke (RATULS): A multicentre randomised controlled trial. *The Lancet*, *394*(10192), 51–62. https://doi.org/10.1016/S0140-6736(19)31055-4

Sangole, A. P., & Levin, M. F. (2009). Palmar arch modulation in patients with hemiparesis after a stroke. *Experimental Brain Research*, *199*(1), 59–70. https://doi.org/10.1007/s00221-009-1972-5

Stinear, C. M., Byblow, W. D., Ackerley, S. J., Barber, P. A., & Smith, M.-C. (2017). Predicting recovery potential for individual stroke patients increases rehabilitation efficiency. *Stroke*, *48*(4), 1011–1019. https://doi.org/10.1161/STROKEAHA.116.015790

The EndNote Team. (2013). *Endnote TM* (EndNote X9, Type: 64 bit). Clarivate. https://endnote.com/?language=en

UK EQUATOR Centre. (2023). *Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network*. The Centre for Statistics in Medicine (CSM), Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford. https://www.equator-network.org/

Velozo, C. A., & Woodbury, M. L. (2011). Translating measurement findings into rehabilitation practice: An example using fugl-Meyer assessment-upper extremity with patients following stroke. *Journal of Rehabilitation Research and Development*, *48*(10), 1211–1222. https://doi.org/10.1682/JRRD.2010.10.0203

Waddell, K. J., Birkenmeier, R. L., Bland, M. D., & Lang, C. E. (2016). An exploratory analysis of the self-reported goals of individuals with chronic upper-extremity paresis following stroke. *Disability and Rehabilitation*, *38*(9), 853–857. https://doi.org/10.3109/09638288.2015.1062926

Woodbury, M. L., Velozo, C. A., Richards, L. G., & Duncan, P. W. (2013). Rasch analysis staging methodology to classify upper extremity movement impairment after stroke. *Archives of Physical Medicine and Rehabilitation*, *94*(8), 1527–1533. https://doi.org/10.1016/j.apmr.2013.03.007

Woytowicz, E. J., Rietschel, J. C., Goodman, R. N., Conroy, S. S., Sorkin, J. D., Whitall, J., & McCombe Waller, S. (2017). Determining levels of upper extremity movement impairment by applying a cluster analysis to the fugl-Meyer assessment of the upper extremity in chronic stroke. *Archives of Physical Medicine and Rehabilitation*, *98*(3), 456–462. https://doi.org/10.1016/j.apmr.2016.06.023