Prediction of Walking and Arm Recovery after Stroke: A Critical Review

Li Khim Kwah [1],* and Robert D. Herbert [2]

[1] Discipline of Physiotherapy, Graduate School of Health, University of Technology Sydney, Ultimo, NSW 2007, Australia

[2] Neuroscience Research Australia (NeuRA), Randwick, NSW 2031, Australia; r.herbert@neura.edu.au

* Correspondence: likhim.kwah@uts.edu.au; Tel.: +61-2-9514-4281

Abstract: Clinicians often base their predictions of walking and arm recovery on multiple predictors. Multivariate prediction models may assist clinicians to make accurate predictions. Several reviews have been published on the prediction of motor recovery after stroke, but none have critically appraised development and validation studies of models for predicting walking and arm recovery. In this review, we highlight some common methodological limitations of models that have been developed and validated. Notable models include the proportional recovery model and the PREP algorithm. We also identify five other models based on clinical predictors that might be ready for further validation. It has been suggested that neurophysiological and neuroimaging data may be used to predict arm recovery. Current evidence suggests, but does not show conclusively, that the addition of neurophysiological and neuroimaging data to models containing clinical predictors yields clinically important increases in predictive accuracy.

Keywords: prediction; prognosis; recovery; walking; arm function; stroke

1. Introduction

It would be useful to be able to predict recovery of walking and arm after stroke. Accurate predictions are needed so that clinicians can provide patients with prognoses, set goals, select therapies and plan discharge [1–4]. For example, if it was possible to predict with some certainty that a particular patient would be unable to walk independently at six months, the clinicians providing that patient with acute and subacute care might work toward a discharge goal of safe transfers. Therapy might involve carer training and equipment prescription rather than intensive gait training. The ability to make accurate predictions could reduce the length of stay in hospitals and enable efficient utilization of stroke care resources [4,5].

Several systematic reviews have identified strong predictors of walking and arm recovery after stroke [2,3,6]. In one systematic review of prognostic studies on walking, clinical variables such as age, severity of paresis and leg power were found to be strong predictors of walking after stroke (based on five studies, each of between 197 and 804 patients) [2]. In another systematic review of prognostic studies on arm recovery, clinical, neurophysiological and neuroimaging data were found to be strong predictors of arm recovery after stroke (based on 58 studies of 9–1197 patients) [3]. These clinical, neurophysiological and neuroimaging data included measures of upper limb impairment, upper limb function, lower limb impairment, motor and somatosensory evoked potentials, and measures obtained with diffusion tensor imaging [3].

In practice, clinicians base their predictions about clinical outcomes on multiple variables [7–9]. If multiple predictors are to be used to make prognoses, there needs to be a proper accounting of

the independent (incremental) predictive value of each predictor variable. Therefore the most useful information about prognosis is likely to come from multivariate prediction models [7–9].

The research which underpins establishment of clinically useful multivariate prediction models involves several steps. First 'development studies' are conducted to build the multivariate prediction models [7]. Subsequently the predictive accuracy of the models is tested on new cohorts [7,10]. These studies are known as 'validation studies' [7]. It is recommended that prediction models should not be used in clinical practice until both development and validation studies have been conducted [7,10]. Once development and validation studies have been conducted, impact studies may be conducted, although the reality is that few reports of impact studies are published. Impact studies resemble clinical trials; they test the efficacy of use of prediction models on patient outcomes, clinician behaviour and cost-effectiveness of care [7,11]. Recent narrative reviews have provided updates on the prediction of motor recovery after stroke [5,12] but these reviews have not focused on development and validation studies of models for predicting walking and arm recovery.

This review provides a critical review of prediction models of walking and arm recovery after stroke. Studies were identified using the search strategy and inclusion criteria in the Appendix A. The review begins in the second section with the definitions and measurements of walking and arm recovery. The third section provides a detailed description of the recommended process for developing and validating a prediction model because this process provides a benchmark against which prediction modelling studies of walking and arm recovery can be evaluated. The fourth section critically appraises development and validation studies of walking and arm recovery with the aim of identifying multivariate models that could potentially be implemented in clinical practice. Much has been written about the role of neurophysiological and neuroimaging data in predicting arm recovery. The fifth section considers whether neurophysiological and neuroimaging data provide additional predictive value over clinical data alone in predicting arm recovery. We conclude with a summary and recommendations for future prediction modelling studies.

2. Definitions and Measurements of Walking and Arm Recovery

Most studies investigating recovery of walking after stroke have focused on the ability to walk independently [13–22]. The operationalization of this construct varies slightly with regard to the distance covered, the type of aids needed and the type of surfaces. For example, independent walking has been defined as "independence of gait and stair climbing with or without the use of some form of assistive device" [16], "walks 5 m without any assistance or walks with aid" [22] and "imperfect gait, but could be independently ambulant, usually with a walking aid" [15]. Various measurement tools have been used to measure walking in stroke. Some authors have used walking-specific measurement tools like the Functional Ambulation Category [14,16,17,20,23], the Rivermead Mobility Index [19,24] or the walking items of scales that measure stroke severity (e.g., Scandinavian Stroke Scale) [22], motor function (e.g., Motor Assessment Scale) [25] or activities of daily living (e.g., the Barthel Index or the Functional Independence Measure) [21,26–33]. Despite the heterogeneity of these measurement tools, it is easy to extract data on walking from them. Independent walking is equivalent to a score of ≥4/5 on the Functional Ambulation Category [14,16,17,20,23], a score of ≥3/6 on item 5 of the Motor Assessment Scale [25], a score of ≥9/12 on the gait item of Scandinavian Stroke Scale [22] and a score of 15/15 on the mobility item of the Barthel Index [28,29]. Studies that measured walking outcomes were included in this review.

For arm recovery, there is less consensus on an operational definition. It is generally agreed that measures of arm recovery should reflect a patient's ability to use his or her arm in everyday life [6,34–36]. This is certainly considered an important goal to patients after stroke [37]. It is difficult, however, to define arm recovery due to the wide range of actions or tasks associated with functional use of the arm. Therefore the term arm recovery in this review will be used broadly to include measures of body functions or structure (e.g., muscle strength) or activities (e.g., ability to carry a tray of food across a crowded room) [6,38]. Various measurement tools used in the literature to measure arm

recovery include arm-specific measurement tools like the Action Research Arm Test [39–43], the Nine Hole Peg Test [43–46] and the upper limb items of scales that measure body functions or structures (e.g., Motricity Index [45], Fugl-Meyer [45,47]) or activities (e.g., Motor Assessment Scale [25,30,48]). In prognostic studies in stroke, arm recovery has been defined by cut-off scores or change scores on these scales. Examples include cut-off scores of 10/56 [41,42], 35/56 [39], 50/56 [49] on the Action Research Arm Test, 5/6 on upper limb items 7 and 8 of Motor Assessment Scale [25], 18/18 on upper limb items 6 to 8 of Motor Assessment Scale [48] and change scores of upper limb items of Motricity Index [45] and Fugl-Meyer [45,47]. The problem with attributing a score to the definition of arm recovery is that readers have to be familiar with the scale to know what the score represents. Some have argued that a low score (e.g., 10/56 on the Action Research Arm Test) may not be reflective of meaningful arm function [5,39]. Some authors have also used activities of daily living scales (e.g., the feeding, dressing and toileting items of the Barthel Index) to measure arm recovery [50,51]. These scales are, however, insensitive to recovery of function in the hemiparetic arm since they allow use of the intact arm [6,34,52]. For this reason, tests that allow use of both arms (e.g., the Wolf Motor Function Test, the Frenchay Arm Test, the Arm Motor Ability Test and the Bell Test) may be less preferred as measures of arm recovery [52]. Studies that measured outcomes of hemiparetic arm recovery were included in this review. Studies that used activities of daily living scales or tests that allow use of both arms were not included in this review.

3. How Should Prediction Models Be Developed and Validated?

In recent years, several excellent guidelines have been published to guide the conduct and reporting of prognostic studies [53–55] and prediction modelling studies [8,9,56]. These guidelines make recommendations about the processes of model development and validation. Adherence to the processes recommended in these guidelines should reduce bias in predictions and maximise the applicability of the models to clinical practice. For model development, we focus on three elements recommended in all guidelines. They are: (a) recruitment of representative cohorts; (b) use of clearly defined, standardised and easily accessible predictors; and (c) careful selection of predictors for inclusion in the model. For model validation, we focus on two key elements: (a) the need for external validation; and (b) the reporting of predictive accuracy in a way that allows for comparison of the performance of models across cohorts in both development and validation studies.

To obtain representative cohorts it is necessary to sample an inception cohort consisting of consecutive cases. This means that all (or nearly all) eligible participants must be recruited at a uniform time point in the course of the disease [57,58]. Non-consecutive sampling in cohort studies can result in selective sampling (so that, for example, only patients with atypically good or atypically bad prognoses are included in the study) which potentially leads to the development of models that cannot be applied to the wider population of interest [55,56]. The second element to consider when developing a model is the use of clearly defined and standardised predictors. Heterogeneity in the definition and measurement of predictors potentially reduces the accuracy of predictions and may cause otherwise potentially useful predictors to be excluded from the final prediction mode [56]. It is also important to ensure that predictors are easily accessible and available at the time the model is to be used, as this enhances the generalisability and applicability of models to clinical practice [7]. The third element to consider is the careful selection of predictors when building a multivariate model. While there are many variable selection procedures, and none is considered universally superior to others [59] it is clear that the use of simplistic variable selection procedures (e.g., forward selection techniques) can lead to predictor selection bias [56], particularly if the sample size is small or the number of potential predictors is large [60,61], if continuous predictors are dichotomised [62,63], or if there is no accounting in the selection process for the number of alternative models considered by the variable selection algorithm [56,59]. To reduce the risk of overfitting, contemporary variable selection techniques such as shrinkage techniques and bootstrapping are recommended to adjust the

weights assigned to the selected predictors [56,64,65]. These procedures should increase the accuracy of predictions made when the model is applied to new cohorts [56,64,65].

Validation of a development model involves testing the model on a new group of patients, preferably patients with similar characteristics located in a different hospital or centre [10]. This is called external validation. External validation is more rigorous than internal validation (using data splitting techniques or bootstrapping techniques to determine the sensitivity of model parameters to randomness in the sample) or temporal validation (applying the model to new patients in same hospital or centre) [10]. If comparisons are to be made between the development cohort and the validation cohort, the calibration and discrimination of the models must be quantified [56]. Calibration refers to how well observed probabilities agree with predicted probabilities. These probabilities are often plotted against each other so that, with perfect predictions (i.e., observed probabilities matching predicted probabilities), the observations fall along a line [65]. Discrimination refers to how well the model can distinguish between patients with good and bad outcomes and is often quantified with the concordance (c) statistic or the area under the receiver operating characteristic (ROC) curve [65]. Other measures of model performance such as the Hosmer-Lemeshow test (for calibration) and sensitivity and specificity (for discrimination) are widely used but are less useful for quantifying model performance [56,65].

4. Critical Appraisal of Development and Validation Studies of Walking and Arm Recovery

4.1. Walking and Arm Recovery: Critical Appraisal of Development Studies

It is important that models are developed on representative acute stroke cohorts. The cohorts may be obtained from registries [13,22,29] or may consist of consecutive patients admitted to hospitals with stroke [25,44,45,66–68]. However a large proportion of prognostic studies/models on walking and arm recovery have been developed on patients included in clinical trials [16,41,46,69–73] or referred for rehabilitation [14,17–19,23,24,26–28,33,40,51,74–83]. This is a problem because clinical trials often select participants using strict inclusion criteria that do not render the sample representative of the wider stroke population, and rehabilitation cohorts may have prognoses that are not typical of all patients. Use of these samples potentially biases predictive models, or at least limits the predictions to populations like those sampled in these studies. Models developed on trial cohorts or rehabilitation cohorts might still be useful if clinicians are trying to predict outcomes of patients in rehabilitation, but they are less useful for predicting outcomes in the early stages after stroke.

If a model is to be applicable in clinical practice, the predictors included in the model must be reproducible and easily obtained or available at the time the model is to be used [7,56]. Therapy intensity and length of stay have been shown to predict recovery of walking [24,28,66,84]. However, it is not possible to obtain data on therapy intensity or length of stay early after stroke. Therapy intensity may be interpreted differently in various stroke units, especially in different countries [24,28,66] which may cause the models to be poorly calibrated when applied internationally. In some prognostic models of upper limb function, predictors such as active range of motion and muscle strength of the upper limbs have been identified [72,81,82,85]. However these predictors were measured with 3D tracking systems or dynamometers which might not be available in routine clinical practice. Scales such as the Orpington Prognostic Score and the Orgogozo Score predict recovery in walking and upper limb function [44,77] but these scales are less commonly used than scales such as the National Institutes of Health Stroke Scale (NIHSS) [86] and so they may not be immediately available for making prognoses in many clinical settings.

Few studies that have developed prediction models of walking and arm recovery have used contemporary, robust variable selection processes. Most of these studies have either not conducted multivariate analyses [26,29,33,40,46,51,68,77–79] or have not provided a clear description of the variable selection process used to build their multivariate models [17,22,70,80]. In many studies where the variable selection process is clearly described, there is a large ratio of predictors to outcomes

(e.g., there may be <10 events/predictor) [16,41], continuous predictors are dichotomised [14,24,41,76], or simplistic selection procedures (e.g., forward selection techniques) are used without accounting for the number of models considered in the selection process [16,24,41,45,76]. As a result, many models of walking and arm recovery prediction models may be susceptible to overfitting and predictor selection bias. It is likely that some of the predictors in these models will have spurious associations with the outcomes and that, when these models are used on new cohorts of patients they will produce inaccurate predictions [56,64,65]. Ultimately, the strongest protection against these problems comes from validation of the models prior to their use in clinical practice.

4.2. Arm Recovery: Critical Appraisal of Validation Studies

To date, none of the prediction models for walking has been validated. However one model for arm recovery—the proportional recovery model—has been validated [47]. This model was developed on 41 patients with a stroke and arm deficits (i.e., patients with Fugl Meyer-Upper Extremity (FM-UE) scores <66). Patients were recruited from a single inpatient stroke service in the USA. The model has subsequently been tested in two other cohorts in the USA and the Netherlands [87,88]. The outcome predicted in the proportional recovery model is the change in FM-UE scores (i.e., final FM-UE measured approximately 3 or 6 months after stroke minus initial FM-UE measured 24–72 h after stroke). The predictors included initial FM-UE, subcortical lesion volume, age and time to reassessment, though the final model only required input of initial FM-UE scores as the mean values of subcortical lesion volume, age and time to reassessment of the development sample was used in the final model in the development [47] and validation studies [87,88]. We have two concerns about the validation of the proportional recovery model. One is that the predictive accuracy of the proportional recovery model was inflated by excluding outliers, both in the development and validation stages. The development study reported high predictive accuracy ($r^2 = 0.90$) after excluding 7 of 41 with severe initial impairment on the grounds that these were 'outliers' [47]. With the outliers included, the original adjusted r^2 was only 0.47 [47]. The same issue arose in the validation studies, which reported high predictive accuracy after excluding 65 of 211 patients [87] and 7 of 30 patients [88]. Another issue is that, in one of the validation studies, the validation cohort included some participants who had been in the development cohort (14 of the 30 participants) [88]. The authors have consistently claimed that the proportional recovery model can accurately predict arm recovery in people who have had mild or moderate strokes, but not severe strokes. That claim would be stronger if the decision to apply the model only to people with mild or moderate strokes was explicitly made prior to the validation, and if the validation study was conducted on a completely independent sample to the development sample.

It appears that the proportional recovery model predicts outcomes well in some patients and not in others. Further analyses in the validation studies have revealed that patients with more severe stroke deficits (e.g., severe arm impairment [88], absent finger extension, facial palsy, severe leg impairment and total or partial anterior cerebral infarction [87]) had recovery profiles that did not fit the proportional recovery model. In a third study (not a validation study per se), Byblow et al. investigated the prognostic value of motor evoked potentials when used with the proportional recovery model [73]. The authors found the proportional recovery model had excellent predictive accuracy ($r^2 = 0.88–0.95$) in 82 patients with present motor evoked potentials [73]. However, there were no head-to-head comparisons of the predictive accuracy of the model with and without neurophysiological measures at 6, 12 and 26 weeks and only a small number of patients with absent motor evoked potentials ($n = 11$) were included in the study. Further validation of the proportional recovery model with motor evoked potentials is required in a larger and more heterogeneous sample. Nonetheless, these results suggest that the inclusion of additional clinical or neurophysiological predictors might improve performance of the proportional recovery model in all patients with stroke.

4.3. Walking and Arm Recovery: Models That Might Be Ready for External Validation

In this section, we identify five well-developed models that might be ready for external validation. These models (two on walking [20,25] and three on arm recovery [25,42]) were developed in the Netherlands [20,42] and Australia [25]. For the purpose of this review, we consider models which have recruited representative cohorts early after stroke, used clearly defined and standardised predictors in their models, and/or used robust methods to select predictors. The models we consider used standardised outcome measures and have high rates of follow-up, study features that suggest the findings are at low risk of bias [8,9,55,56]. All of these models reported the predictive accuracy of their models, which facilitates external validation.

4.3.1. Walking

In the study by Veerbeek and colleagues [20], a sample of 154 non-ambulatory patients were recruited from nine hospital stroke units in the Netherlands. The predicted outcome was walking measured with the Functional Ambulation Category. The Functional Ambulation Category has been shown to be a reliable and valid tool for measuring walking ability in people who have had a stroke [89–91]. The study had a high follow-up rate (92% of survivors at 6 months). The authors found that early presence of lower limb muscle strength and sitting ability predicted recovery of walking six months after stroke [20]. These predictors were measured with the Motricity Index for leg (MI leg) and the Trunk Control Test item 3: sitting balance (TCT-s). These assessment tools are widely used in Europe and elsewhere [14,41,69,71] so it is possible to externally validate the model and it would ultimately be possible to use the model in these regions. The prediction models are as follows:

Model used <72 h:

$$\text{Probability of walking at 6 months} = 1/(1 + e^{-0.982 + 2.691 \, \text{TCT-s} + 2.083 \, \text{MI leg}}) \tag{1}$$

Model used on day 5:

$$\text{Probability of walking at 6 months} = 1/(1 + e^{-1.236 + 2.815 \, \text{TCT-s} + 1.609 \, \text{MI leg}}) \tag{2}$$

Model used on day 9:

$$\text{Probability of walking at 6 months} = 1/(1 + e^{-2.226 + 3.629 \, \text{TCT-s} + 1.854 \, \text{MI leg}}) \tag{3}$$

Note that TCT-s and MI leg were dichotomised at 25. Therefore, TCT-s and MI leg have values of 0 (if the score is ≤25) or 1 (score >25). The sensitivity of the models ranged from 0.93 (95% CI = 0.86–0.96) to 0.94 (95% CI = 0.87–0.97) and specificity ranged from 0.63 (95% CI = 0.43–0.78) to 0.83 (95% CI = 0.64–0.93) [20].

In the second study on walking, by Kwah et al. [25] 114 non-ambulatory patients were recruited. Participants were sampled only from a single hospital in Australia. Sampling of consecutive patients reduced the risk of bias. It was found that age and severity of stroke (measured with NIHSS) predicted independent walking six months after stroke [25]. Age and NIHSS data are relatively easy data to obtain. If NIHSS scores are not available they can also be substituted with Canadian Neurological Scale (CNS) scores using the following regression equation [92]: *NIHSS = 23 − 2 × CNS*. The risk of bias in this study was low because of the recruitment of a representative inception cohort, use of a standardised outcome measure (Motor Assessment Scale item 5 for walking) [93,94], high follow-up rates (95% of survivors at 6 months), consideration of only a small number of potential predictors, and the use of bootstrapping to shrink parameter estimates. The prediction model is:

Model used within 4 weeks of stroke:

$$\text{Probability of walking at 6 months} = 1/(1 + e^{-11.02852 − 0.1053 \, \text{age} − 0.2436 \, \text{NIHSS}}) \tag{4}$$

The model had good discrimination (AUC of 0.84, 95% CI 0.77 to 0.92). Calibration was assessed using the Hosmer-Lemeshow test and showed no statistical significance for the model (0.70), indicating that there was no evidence of a failure of fit.

4.3.2. Arm Recovery

The methods used in the study by Nijland and colleagues [42] were similar to methods used in the study by Veerbeek and colleagues [20] in that a representative acute stroke cohort (188 patients with hemiparesis) was recruited and predictors were standardised and easily collected in stroke units. The study also used a standardised outcome measure (the Action Research Arm Test) [95–97] to measure arm recovery and had high follow-up rates (93% of survivors). It was found that early presence of finger extension and shoulder abduction predicted arm recovery six months after stroke [42]. Predictors included the Fugl-Meyer finger extension (FM-FE) and the Motricity Index shoulder abduction (MI-SA). The multivariate prediction models are as follows:

Model used on day 2:

$$\text{Probability of arm recovery at 6 months} = 1/(1 + e^{-1.119 + 2.807\,\text{FM-FE} + 2.149\,\text{MI-SA}}) \tag{5}$$

Model used on day 5:

$$\text{Probability of arm recovery at 6 months} = 1/(1 + e^{-1.874 + 3.070\,\text{FM-FE} + 3.075\,\text{MI-SA}}) \tag{6}$$

Model used on day 9:

$$\text{Probability of arm recovery at 6 months} = 1/(1 + e^{-1.815 + 3.224\,\text{FM-FE} + 2.449\,\text{MI-SA}}) \tag{7}$$

Note that FM-FE was dichotomised at 1 and MI-SA was dichotomised at 9. Therefore, if a patient has scores of FM-FE of less than 1 and MI-SA of less than 9, one inputs "0" for these variables. The sensitivity of the models ranged from 0.89 (95% CI = 0.85–0.92) to 0.95 (95% CI = 0.91–0.97) and specificity ranged from 0.83 (95% CI = 0.72–0.90) to 0.83 (95% CI = 0.74–0.89) [42].

The last two models of arm recovery were developed by Kwah et al. [25]. The outcomes were measured using items 7 and items 8 of the Motor Assessment Scale [36,93,94]. One model predicted the probability of being able to move a cup across a table six months after stroke (*N* = 65 patients) and the other model predicted the probability of being able to feed with a spoonful of liquid using the affected arm (*N* = 69 patients). The prediction models are:

Model used within 4 weeks of stroke:

$$\text{Probability of moving a cup across the table at 6 months} = 1/(1 + e^{-4.8167 - 0.0533\,\text{age} - 0.1240\,\text{NIHSS}}) \tag{8}$$

$$\text{Probability of feeding oneself with spoonful of liquid at 6 months} = 1/(1 + e^{-2.0855 - 0.2262\,\text{NIHSS}}) \tag{9}$$

The AUCs for the prediction models were 0.73 (95% CI 0.59 to 0.87) for moving a cup and 0.82 (95% CI 0.70 to 0.94) for feeding. The Hosmer-Lemeshow test was not statistically significant for any model (0.74 for moving a cup, 0.38 for feeding oneself), so there was no evidence of a failure of fit.

It is important to note that these well-developed models are not without methodological limitations.

For example, in the models developed in the Netherlands, cohorts did not appear to be recruited consecutively. There is also a risk of overfitting or predictor selection bias in these models because continuous predictors were dichotomised and the number of predictors entered into the model exceeded the number of outcome events (thereby not fulfilling the widely used criterion of 10 outcome events per candidate predictor [60,61]). In the models developed in Australia, the cohorts were recruited from a single site which reduces the generalisability of the results. The sample sizes used in the arm recovery cohorts were also small (*n* < 70) which precluded the consideration of many predictors. The Australian studies used bootstrapping techniques to 'shrink' their models, but the bootstrapped

models had poor calibration so subsequently only simplistic backwards selection models were used. Ultimately, these models must be externally validated before they can be used in clinical practice.

5. Arm Recovery: Do Neurophysiological and Neuroimaging Data Provide Additional Predictive Value over Clinical Data?

Neurophysiological and neuroimaging variables have been used to predict arm recovery. The most widely used neurophysiological predictor is the presence of motor evoked potentials (MEPs). MEPs are obtained using transcranial magnetic stimulation (TMS) which is a non-invasive form of brain stimulation. TMS of the primary motor cortex on the same side as the lesion activates muscles on the contralateral limb via the corticospinal tract, and the resulting MEPs are measured with surface electromyography [12]. Neuroimaging predictors include functional MRI and diffusion weighted MRI [5,12]. A common approach is to measure fractional anisotropy (FA) in specific sites in the brain, often the posterior limb of the internal capsules, using diffusion weighted MRI [49,73,98–101]. FA measures the directional preference of diffusion in white matter. Disruption of the corticospinal tracts decreases FA [98]. FA of the lesioned and non-lesioned hemispheres are often measured and presented as a ratio to predict arm recovery. This ratio is termed the FA asymmetry index. FA asymmetry index values of <0.15 and <0.25 have been used to predict arm recovery [49,73,98]. Recent studies that have used neuroimaging predictors to predict arm recovery have also measured "early fiber number ratio" [102] and "weighted corticospinal lesion load" [103,104] as markers of corticospinal tract integrity.

Two systematic reviews have been conducted on the prediction of arm recovery [3,6] and several more systematic reviews have investigated the prognostic value of neurophysiological and neuroimaging predictors on arm recovery [105–107]. The reviews conclude there is sufficient evidence to suggest neurophysiological and neuroimaging data have the potential to predict arm recovery [3,6,105–107]. However it remains debatable whether neurophysiological and neuroimaging data provide sufficient additional predictive information, compared to clinical data alone, to justify their use for prediction of arm recovery in clinical practice. This consideration is particularly important because there are already well-developed models that can predict arm recovery accurately using low-cost, readily available and easily collected clinical data [25,42].

A premise of studies that use neurophysiology and neuroimaging data to predict arm recovery is that clinical data alone do not provide adequate predictions of arm recovery, particularly in patients with initial severe arm impairment [73,103,108,109]. When this claim is made explicitly it is often supported by reference to the proportional recovery model or poorly conducted prognostic studies, rather than models which have shown good predictive accuracy [25,42]. In addition, comparisons of predictive accuracy of models based on clinical data and models based on neurophysiological or neuroimaging data consistently demonstrate that prediction models based on clinical data have similar or higher predictive accuracy in predicting arm recovery [75,99,103,104,110]. When neurophysiological or neuroimaging data are added to models containing clinical data the increase in predictive accuracy is small [75,99,110] (e.g., increments in r^2 of 5%–9% [75,99])—probably too small to justify the expense and inconvenience of these investigations. (Note that in the study by Byblow et al., no comparisons of predictive accuracy were made between the proportional recovery model and the proportional recovery model with motor evoked potentials. So the additional predictive value of motor evoked potentials in predicting arm recovery in *all* patients has not yet been unambiguously established [73].)

A notable prediction model is the PREP algorithm [49]. This model uses clinical predictors that are similar to those identified by Niljand et al. [42] (muscle strength of shoulder abduction and finger extension) as well as neurophysiological (presence of MEPs) and neuroimaging (FA asymmetry index <0.15) predictors. The algorithm involves a step-by-step approach. First the strength of the shoulder abductor and finger extensor muscles is measured using the Medical Research Council muscle grades and, if the combined score is less than 8, TMS is used. If MEPs are absent, MRI is used. The PREP model had a positive predictive value of 88% and a negative predictive value of 83%

(specificity 88% and sensitivity 73%; positive likelihood ratio 6.1. negative likelihood ratio 0.31) on the development set for complete recovery at 3 months. This is a potentially useful level of predictive accuracy, particularly also because the model predicted a high level of arm recovery (ARAT ≥ 50/56) compared to that in Niljand et al. (ARAT ≥ 10/56) [42]. However, the predictive accuracy of the models at each step of the algorithm was not reported so it is not clear that the increase in predictive accuracy provided by TMS and DTI is sufficient to make the acquisition of these time- and labour-intensive measurements worthwhile. It is also worth noting that the sensitivity and specificity of the PREP model in predicting complete recovery were actually lower than the sensitivity and specificity of the model developed by Niljand et al. [42], which brings into question whether neurophysiological and neuroimaging data are necessary for accurate predictions of arm recovery.

A potential advantage of basing predictions on neurophysiology and neuroimaging data rather than clinical data is that it is possible to collect neurophysiology and neuroimaging data from patients with cognitive or neuropsychological deficits [12,75,110]. In these patients it may be hard to conduct clinical assessments that involve self-reporting or performance of complex tasks. However, studies that have predicted arm recovery with neurophysiological and neuroimaging data have often excluded patients with cognitive deficits [49,73,98–101,103,104,109]. Moreover, patients with previous strokes or contraindications to TMS or MRI are also often excluded from TMS and MRI studies [49,73,75,98–103,109,110]. Consequently it is not yet clear that neurophysiological and neuroimaging data can aid prognosis in cognitively or neuropsychologically impaired patients.

6. Conclusions and Recommendations for Future Studies

If a prediction model is to be used in clinical practice, it must be well developed and externally validated prior to use in clinical practice. While many models predicting death and severe disability after stroke have been well developed and externally validated [111–114], few models for predicting walking and arm recovery have been well developed. One model for arm recovery (the proportional recovery model) has been externally validated though results from development and validation studies suggest that the model does not predict well in *all* patients with stroke. This model does appear to predict outcomes well in people with less severe strokes. In this review we have identified five models [20,25,42] that are potentially ready for external validation.

It will benefit future researchers in prognostic research to have access to large datasets from stroke registries such as the Australian Stroke Clinical Registry [115]. Stroke registries can provide representative cohorts and large data sets that permit rigorous external validation of models. Regarding the choice of predictors, future researchers should consider the National Institute of Neurological Disorders and Stroke (NINDS) Stroke Common Data Element (CDE) project that recommends a common set of data to collect in studies involving people with stroke and transient ischemic attacks [116]. That would ensure data on predictors are standardised and more easily obtained in routine clinical practice, and it would help to facilitate external validation of models in other stroke cohorts.

Although there is no single recommended approach to building a multivariate prediction model, several techniques are known to lead to overfitting and predictor selection bias. Researchers conducting prognostic research should use contemporary techniques and refer to guidelines on the reporting and conduct of prediction modelling studies [8,9,56] in order to develop robust prediction models. A good example of a rigorously developed and validated prediction model is the iScore model used to predict mortality after stroke112]. The iScore model has been externally validated and has an acceptable predictive accuracy (c-statistics ≥0.78) [111,117]. It can be implemented with a web application [118] or a smartphone application [119]. Distribution of software for calculating prognoses using web sites or smartphones will facilitate use of prediction models in practice.

Neurophysiology and neuroimaging data have an important role to play in diagnosis and in understanding mechanisms of arm recovery [120–122]. While many studies suggest neurophysiology and neuroimaging data can be helpful in predicting arm recovery [3,6,105–107], there is no strong

evidence that models containing neurophysiology or neuroimaging data predict arm recovery *better* than models containing clinical data, or that the addition of neurophysiology or neuroimaging data to models containing clinical predictors provide sufficient increase in predictive accuracy to warrant their use. Researchers who develop prediction models based on neurophysiological or neuroimaging data should routinely investigate whether the neurophysiological or neuroimaging data provide better predictive accuracy than models based on clinical data alone. Exemplary studies of prediction models for stroke include the neuroimaging studies conducted by Johnston et al. [123–125] and Schiemanck et al. [126]. In their studies predicting outcomes [123–125] and independence in activities of daily living [126] after stroke, neuroimaging data (i.e., CT or MRI scans of infarct volume) either did not improve the predictive accuracy of the model with clinical predictors [123,124,126] or produced only small and unimportant increases in predictive accuracy [125].

Readers should be aware that we have only provided a brief critical appraisal of development studies in walking and arm recovery. Other factors such as outcome measurement, sample size, missing data, blinding of predictor measurement, timing of predictor measurement and handling of predictors in the model can also influence the risk of bias and/or applicability of a prediction model [56].

Author Contributions: L.K.K. and R.D.H. conceived the review. L.K.K. wrote the first draft of the review; R.D.H. and L.K.K. edited subsequent drafts. Both L.K.K. and R.D.H. made an equal intellectual contribution to the review. L.K.K. and R.D.H. approved the final version of the review before submission.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Search strategy: references for this review were identified by searching PubMed from January 1970 to July 2016. The search syntax included (predict* OR prognos* OR probability) AND ((walk OR gait OR ambulat* OR mobility) OR (arm OR upper limb)) AND (stroke). Additional references were identified by inspecting reference lists of relevant articles and authors' personal libraries. Studies were included if they were cohort studies, measured outcomes of walking and arm recovery and were published in English.

References

1. Houwink, A.; Nijland, R.H.; Geurts, A.C.; Kwakkel, G. Functional recovery of the paretic upper limb after stroke: Who regains hand capacity? *Arch. Phys. Med. Rehabil.* **2013**, *94*, 839–844. [CrossRef] [PubMed]
2. Craig, L.E.; Wu, O.; Bernhardt, J.; Langhorne, P. Predictors of poststroke mobility: Systematic review. *Int. J. Stroke* **2011**, *6*, 321–327. [CrossRef] [PubMed]
3. Coupar, F.; Pollock, A.; Rowe, P.; Weir, C.; Langhorne, P. Predictors of upper limb recovery after stroke: A systematic review and meta-analysis. *Clin. Rehabil.* **2012**, *26*, 291–313. [CrossRef] [PubMed]
4. Veerbeek, J.M.; Kwakkel, G.; van Wegen, E.E.; Ket, J.C.; Heymans, M.W. Early prediction of outcome of activities of daily living after stroke: A systematic review. *Stroke* **2011**, *42*, 1482–1488. [CrossRef] [PubMed]
5. Stinear, C. Prediction of recovery of motor function after stroke. *Lancet Neurol.* **2010**, *9*, 1228–1232. [CrossRef]
6. Chen, S.-Y.; Winstein, C.J. A systematic review of voluntary arm recovery in hemiparetic stroke: Critical predictors for meaningful outcomes using the international classification of functioning, disability, and health. *J. Neurol. Phys. Ther.* **2009**, *33*, 2–13. [CrossRef] [PubMed]
7. Moons, K.G.; Royston, P.; Vergouwe, Y.; Grobbee, D.E.; Altman, D.G. Prognosis and prognostic research: What, why, and how? *BMJ* **2009**, *338*, b375. [CrossRef] [PubMed]
8. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ* **2015**, *350*, g7594. [CrossRef] [PubMed]

9. Moons, K.G.; Altman, D.G.; Reitsma, J.B.; Ioannidis, J.P.; Macaskill, P.; Steyerberg, E.W.; Vickers, A.J.; Ransohoff, D.F.; Collins, G.S. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann. Intern. Med.* **2015**, *162*, W1–W73. [CrossRef] [PubMed]

10. Altman, D.G.; Vergouwe, Y.; Royston, P.; Moons, K.G. Prognosis and prognostic research: Validating a prognostic model. *BMJ* **2009**, *338*, b605. [CrossRef] [PubMed]

11. Moons, K.G.; Altman, D.G.; Vergouwe, Y.; Royston, P. Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. *BMJ* **2009**, *338*, b606. [CrossRef] [PubMed]

12. Stinear, C.M.; Byblow, W.D.; Ward, S.H. An update on predicting motor recovery after stroke. *Ann. Phys. Rehabil. Med.* **2014**, *57*, 489–498. [CrossRef] [PubMed]

13. Dallas, M.I.; Rone-Adams, S.; Echternach, J.L.; Brass, L.M.; Bravata, D.M. Dependence in prestroke mobility predicts adverse outcomes among patients with acute ischemic stroke. *Stroke* **2008**, *39*, 2298–2303. [CrossRef] [PubMed]

14. Duarte, E.; Marco, E.; Muniesa, J.M.; Belmonte, R.; Aguilar, J.J.; Escalada, F. Early detection of non-ambulatory survivors six months after stroke. *NeuroRehabilitation* **2010**, *26*, 317–323. [PubMed]

15. Feigin, L.; Sharon, B.; Czaczkes, B.; Rosin, A.J. Sitting equilibrium 2 weeks after a stroke can predict the walking ability after 6 months. *Gerontology* **1996**, *42*, 348–353. [CrossRef] [PubMed]

16. Kollen, B.; Kwakkel, G.; Lindeman, E. Longitudinal robustness of variables predicting independent gait following severe middle cerebral artery stroke: A prospective cohort study. *Clin. Rehabil.* **2006**, *20*, 262–268. [CrossRef] [PubMed]

17. Masiero, S.; Avesani, R.; Armani, M.; Verena, P.; Ermani, M. Predictive factors for ambulation in stroke patients in the rehabilitation setting: A multivariate analysis. *Clin. Neurol. Neurosurg.* **2007**, *109*, 763–769. [CrossRef] [PubMed]

18. Mayo, N.E.; Korner-Bitensky, N.A.; Becker, R. Recovery time of independent function post-stroke. *Am. J. Phys. Med. Rehabil.* **1991**, *70*, 5–12. [CrossRef] [PubMed]

19. Van de Port, I.G.; Kwakkel, G.; Schepers, V.P.; Lindeman, E. Predicting mobility outcome one year after stroke: A prospective cohort study. *J. Rehabil. Med.* **2006**, *38*, 218–223. [CrossRef] [PubMed]

20. Veerbeek, J.M.; Van Wegen, E.E.; Harmeling-Van der Wel, B.C.; Kwakkel, G. Is accurate prediction of gait in nonambulatory stroke patients possible within 72 hours poststroke?: The EPOS study. *Neurorehabil. Neural Repair* **2011**, *25*, 268–274. [CrossRef] [PubMed]

21. Wade, D.T.; Hewer, R.L. Functional abilities after stroke: Measurement, natural history and prognosis. *J. Neurol. Neurosurg. Psychiatry* **1987**, *50*, 177–182. [CrossRef] [PubMed]

22. Wandel, A.; Jorgensen, H.S.; Nakayama, H.; Raaschou, H.O.; Olsen, T.S. Prediction of walking function in stroke patients with initial lower extremity paralysis: The Copenhagen Stroke Study. *Arch. Phys. Med. Rehabil.* **2000**, *81*, 736–738. [CrossRef]

23. Sanchez-Blanco, I.; Ochoa-Sangrador, C.; Lopez-Munain, L.; Izquierdo-Sanchez, M.; Fermoso-Garcia, J. Predictive model of functional independence in stroke patients admitted to a rehabilitation programme. *Clin. Rehabil.* **1999**, *13*, 464–475. [CrossRef] [PubMed]

24. Paolucci, S.; Grasso, M.G.; Antonucci, G.; Bragoni, M.; Troisi, E.; Morelli, D.; Coiro, P.; De Angelis, D.; Rizzi, F. Mobility status after inpatient stroke rehabilitation: 1-year follow-up and prognostic factors. *Arch. Phys. Med. Rehabil.* **2001**, *82*, 2–8. [CrossRef] [PubMed]

25. Kwah, L.K.; Harvey, L.A.; Diong, J.; Herbert, R.D. Models containing age and NIHSS predict recovery of ambulation and upper limb function six months after stroke: An observational study. *J. Physiother.* **2013**, *59*, 189–197. [CrossRef]

26. Chae, J.; Johnston, M.; Kim, H.; Zorowitz, R. Admission motor impairment as a predictor of physical disability after stroke rehabilitation. *Am. J. Phys. Med. Rehabil.* **1995**, *74*, 218–223. [CrossRef] [PubMed]

27. Hellstrom, K.; Lindmark, B.; Wahlberg, B.; Fugl-Meyer, A.R. Self-efficacy in relation to impairments and activities of daily living disability in elderly patients with stroke: A prospective investigation. *J. Rehabil. Med.* **2003**, *35*, 202–207. [PubMed]

28. Hu, M.H.; Hsu, S.S.; Yip, P.K.; Jeng, J.S.; Wang, Y.H. Early and intensive rehabilitation predicts good functional outcomes in patients admitted to the stroke intensive care unit. *Disabil. Rehabil.* **2010**, *32*, 1251–1259. [CrossRef] [PubMed]

29. Jorgensen, H.S.; Nakayama, H.; Raaschou, H.O.; Olsen, T.S. Recovery of walking function in stroke patients—The Copenhagen Stroke Study. *Arch. Phys. Med. Rehabil.* **1995**, *76*, 27–32. [CrossRef]

30. Loewen, S.C.; Anderson, B.A. Predictors of stroke outcome using objective measurement scales. *Stroke* **1990**, *21*, 78–81. [CrossRef] [PubMed]

31. Patel, A.T.; Duncan, P.W.; Lai, S.M.; Studenski, S. The relation between impairments and functional outcomes poststroke. *Arch. Phys. Med. Rehabil.* **2000**, *81*, 1357–1363. [CrossRef] [PubMed]

32. Shelton, F.D.; Volpe, B.T.; Reding, M. Motor impairment as a predictor of functional recovery and guide to rehabilitation treatment after stroke. *Neurorehabil. Neural Repair* **2001**, *15*, 229–237. [CrossRef] [PubMed]

33. Singh, R.; Hunter, J.; Philip, A.; Todd, I. Predicting those who will walk after rehabilitation in a specialist stroke unit. *Clin. Rehabil.* **2006**, *20*, 149–152. [CrossRef] [PubMed]

34. Ashford, S.; Slade, M.; Malaprade, F.; Turner-Stokes, L. Evaluation of functional outcome measures for the hemiparetic upper limb: A systematic review. *J. Rehabil. Med.* **2008**, *40*, 787–795. [CrossRef] [PubMed]

35. Coupar, F.; Pollock, A.; van Wijck, F.; Morris, J.; Langhorne, P. Simultaneous bilateral training for improving arm function after stroke. *Cochrane Database Syst. Rev.* **2010**, CD006432. [CrossRef]

36. Lannin, N. Reliability, validity and factor structure of the upper limb subscale of the Motor Assessment Scale (UL-MAS) in adults following stroke. *Disabil. Rehabil.* **2004**, *26*, 109–116. [CrossRef] [PubMed]

37. Barker, R.N.; Gill, T.J.; Brauer, S.G. Factors contributing to upper limb recovery after stroke: A survey of stroke survivors in Queensland Australia. *Disabil. Rehabil.* **2007**, *29*, 981–989. [CrossRef] [PubMed]

38. Langhorne, P.; Coupar, F.; Pollock, A. Motor recovery after stroke: A systematic review. *Lancet Neurol.* **2009**, *8*, 741–754. [CrossRef]

39. Au-Yeung, S.S.; Hui-Chan, C.W. Predicting recovery of dextrous hand function in acute stroke. *Disabil. Rehabil.* **2009**, *31*, 394–401. [CrossRef] [PubMed]

40. Broeks, J.G.; Lankhorst, G.J.; Rumping, K.; Prevo, A.J.H. The long-term outcome of arm function after stroke: Results of a follow-up study. *Disabil. Rehabil.* **1999**, *21*, 357–364. [CrossRef] [PubMed]

41. Kwakkel, G.; Kollen, B.J.; van der Grond, J.; Prevo, A.J.H. Probability of regaining dexterity in the flaccid upper limb—Impact of severity of paresis and time since onset in acute stroke. *Stroke* **2003**, *34*, 2181–2186. [CrossRef] [PubMed]

42. Nijland, R.H.M.; van Wegen, E.E.H.; Harmeling-van der Wel, B.C.; Kwakkel, G.; Investigators, E. Presence of finger extension and shoulder abduction within 72 hours after stroke predicts functional recovery. Early Prediction of Functional Outcome After Stroke: The EPOS Cohort Study. *Stroke* **2010**, *41*, 745–750. [CrossRef] [PubMed]

43. Pandyan, A.D.; Cameron, M.; Powell, J.; Stott, D.J.; Granat, M.H. Contractures in the post-stroke wrist: A pilot study of its time course of development and its association with upper limb recovery. *Clin. Rehabil.* **2003**, *17*, 88–95. [CrossRef] [PubMed]

44. Meldrum, D.; Pittock, S.J.; Hardiman, O.; Ni Dhuill, C.; O'Regan, M. Recovery of the upper limb post ischaemic stroke and the predictive value of the Orpington Prognostic Score. *Clin. Rehabil.* **2004**, *18*, 694–702. [CrossRef] [PubMed]

45. Smania, N.; Paolucci, S.; Tinazzi, M.; Borghero, A.; Manganotti, P.; Fiaschi, A.; Moretto, G.; Bovi, P.; Gambarin, M. Active finger extension: A simple movement predicting recovery of arm function in patients with acute stroke. *Stroke* **2007**, *38*, 1088–1090. [CrossRef] [PubMed]

46. Sunderland, A.; Tinson, D.; Bradley, L.; Hewer, R.L. Arm function after stroke—An evaluation of grip strength as a measure of recvery and a prognostic indicator. *J. Neurol. Neurosurg. Psychiatry* **1989**, *52*, 1267–1272. [CrossRef] [PubMed]

47. Prabhakaran, S.; Zarahn, E.; Riley, C.; Speizer, A.; Chong, J.Y.; Lazar, R.M.; Marshall, R.S.; Krakauer, J.W. Inter-individual variability in the capacity for motor recovery after ischemic stroke. *Neurorehabil. Neural Repair* **2008**, *22*, 64–71. [CrossRef] [PubMed]

48. Williams, B.K.; Galea, M.P.; Winter, A.T. What is the functional outcome for the upper limb after stroke? *Aust. J. Physiother.* **2001**, *47*, 19–27. [CrossRef]

49. Stinear, C.M.; Barber, P.A.; Petoe, M.; Anwar, S.; Byblow, W.D. The PREP algorithm predicts potential for upper limb recovery after stroke. *Brain* **2012**, *135*, 2527–2535. [CrossRef] [PubMed]

50. Nakayama, H.; Jorgensen, H.S.; Raaschou, H.O.; Olsen, T.S. Recovery of upper extremity function in stroke patients: The Copenhagen Stroke Study. *Arch. Phys. Med. Rehabil.* **1994**, *75*, 394–398. [CrossRef]

51. Olsen, T.S. Arm and leg paresis as outcome predictors in stroke rehabilitation. *Stroke* **1990**, *21*, 247–251. [CrossRef] [PubMed]

52. Croarkin, E.; Danoff, J.; Barnes, C. Evidence-based rating of upper-extremity motor function tests used for people following a stroke. *Phys. Ther.* **2004**, *84*, 62–74. [PubMed]

53. Vandenbroucke, J.P.; Von Elm, E.; Altman, D.G.; Gøtzsche, P.C.; Mulrow, C.D.; Pocock, S.J.; Poole, C.; Schlesselman, J.J.; Egger, M. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Med.* **2007**, *4*, e297. [CrossRef] [PubMed]

54. Von Elm, E.; Altman, D.G.; Egger, M.; Pocock, S.J.; Gotzsche, P.C.; Vandenbroucke, J.P. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *J. Clin. Epidemiol.* **2008**, *61*, 344–349. [CrossRef] [PubMed]

55. Hayden, J.A.; van der Windt, D.A.; Cartwright, J.L.; Cote, P.; Bombardier, C. Assessing bias in studies of prognostic factors. *Ann. Intern. Med.* **2013**, *158*, 280–286. [CrossRef] [PubMed]

56. Moons, K.G.; de Groot, J.A.; Bouwmeester, W.; Vergouwe, Y.; Mallett, S.; Altman, D.G.; Reitsma, J.B.; Collins, G.S. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *PLoS Med.* **2014**, *11*, e1001744. [CrossRef] [PubMed]

57. Fletcher, R.H.; Fletcher, S.W. *Clinical Epidemiology: The Essentials*, 4th ed.; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2005.

58. Herbert, R.; Jamtvedt, G.; Mead, J.; Hagen, K. *Practical Evidence-Based Physiotherapy*, 2nd ed.; Elsevier Butterworth Heinemann: Oxford, UK, 2011.

59. Royston, P.; Moons, K.G.; Altman, D.G.; Vergouwe, Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ* **2009**, *338*, b604. [CrossRef] [PubMed]

60. Wynants, L.; Bouwmeester, W.; Moons, K.G.; Moerbeek, M.; Timmerman, D.; Van Huffel, S.; Van Calster, B.; Vergouwe, Y. A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. *J. Clin. Epidemiol.* **2015**, *68*, 1406–1414. [CrossRef] [PubMed]

61. Peduzzi, P.; Concato, J.; Kemper, E.; Holford, T.R.; Feinstein, A.R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **1996**, *49*, 1373–1379. [CrossRef]

62. Altman, D.G.; Royston, P. The cost of dichotomising continuous variables. *BMJ* **2006**, *332*, 1080. [CrossRef] [PubMed]

63. Royston, P.; Altman, D.G.; Sauerbrei, W. Dichotomizing continuous predictors in multiple regression: A bad idea. *Stat. Med.* **2006**, *25*, 127–141. [CrossRef] [PubMed]

64. Harrell, F.E., Jr.; Lee, K.L.; Mark, D.B. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **1996**, *15*, 361–387. [CrossRef]

65. Steyerberg, E. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*; Springer: Rotterdam, The Netherlands, 2010.

66. Scrivener, K.; Sherrington, C.; Schurr, K. Amount of exercise in the first week after stroke predicts walking speed and unassisted walking. *Neurorehabil. Neural Repair* **2012**, *26*, 932–938. [CrossRef] [PubMed]

67. De Weerdt, W.; Lincoln, N.B.; Harrison, M.A. Prediction of arm and hand function recovery in stroke patients. *Int. J. Rehabil. Res.* **1987**, *10*, 110–112. [CrossRef] [PubMed]

68. Heller, A.; Wade, D.T.; Wood, V.A.; Sunderland, A.; Hewer, R.L.; Ward, E. Arm function after stroke—Measurement and recovery over the 1st 3 months. *J. Neurol. Neurosurg. Psychiatry* **1987**, *50*, 714–719. [CrossRef] [PubMed]

69. Kollen, B.; van de Port, I.; Lindeman, E.; Twisk, J.; Kwakkel, G. Predicting improvement in gait after stroke: A longitudinal prospective study. *Stroke* **2005**, *36*, 2676–2680. [CrossRef] [PubMed]

70. Prescott, R.J.; Garraway, W.M.; Akhtar, A.J. Predicting functional outcome following acute stroke using a standard clinical examination. *Stroke* **1982**, *13*, 641–647. [CrossRef] [PubMed]

71. Kwakkel, G.; Kollen, B. Predicting improvement in the upper paretic limb after stroke: A longitudinal prospective study. *Restor. neurol. neurosci.* **2007**, *25*, 453–460. [PubMed]

72. Wagner, J.M.; Lang, C.E.; Sahrmann, S.A.; Edwards, D.F.; Dromerick, A.W. Sensorimotor impairments and reaching performance in subjects with poststroke hemiparesis during the first few months of recovery. *Phys. Ther.* **2007**, *87*, 751–765. [CrossRef] [PubMed]

73. Byblow, W.D.; Stinear, C.M.; Barber, P.A.; Petoe, M.A.; Ackerley, S.J. Proportional recovery after stroke depends on corticomotor integrity. *Ann. Neurol.* **2015**, *78*, 848–859. [CrossRef] [PubMed]

74. Feys, H.; De Weerdt, W.; Nuyens, G.; van de Winckel, A.; Selz, B.; Kiekens, C. Predicting motor recovery of the upper limb after stroke rehabilitation: Value of a clinical examination. *Physiother. Res. Int.* **2000**, *5*, 1–18. [CrossRef] [PubMed]

75. Feys, H.; Van Hees, J.; Bruyninckx, F.; Mercelis, R.; De Weerdt, W. Value of somatosensory and motor evoked potentials in predicting arm recovery after a stroke. *J. Neurol. Neurosurg. Psychiatry* **2000**, *68*, 323–331. [CrossRef] [PubMed]

76. Paolucci, S.; Bragoni, M.; Coiro, P.; De Angelis, D.; Fusco, F.R.; Morelli, D.; Venturiero, V.; Pratesi, L. Quantification of the probability of reaching mobility independence at discharge from a rehabilitation hospital in nonwalking early ischemic stroke patients: A multivariate study. *Cerebrovasc. Dis.* **2008**, *26*, 16–22. [CrossRef] [PubMed]

77. Petrilli, S.; Durufle, A.; Nicolas, B.; Pinel, J.F.; Kerdoncuff, V.; Gallien, P. Prognostic factors in the recovery of the ability to walk after stroke. *J. Stroke Cerebrovasc. Dis.* **2002**, *11*, 330–335. [CrossRef] [PubMed]

78. Hatakenaka, M.; Miyai, I.; Sakoda, S.; Yanagihara, T. Proximal paresis of the upper extremity in patients with stroke. *Neurology* **2007**, *69*, 348–355. [CrossRef] [PubMed]

79. Katrak, P.; Bowring, G.; Conroy, P.; Chilvers, M.; Poulos, R.; McNeil, D. Predicting upper limb recovery after stroke: The place of early shoulder and hand movement. *Arch. Phys. Med. Rehabil.* **1998**, *79*, 758–761. [CrossRef]

80. Wade, D.T.; Langton-Hewer, R.; Wood, V.A.; Skilbeck, C.E.; Ismail, H.M. The hemiplegic arm after stroke: Measurement and recovery. *J. Neurol. Neurosurg. Psychiatry* **1983**, *46*, 521–524. [CrossRef] [PubMed]

81. Mirbagheri, M.M.; Rymer, W.Z. Time-course of changes in arm impairment after stroke: Variables predicting motor recovery over 12 months. *Arch. Phys. Med. Rehabil.* **2008**, *89*, 1507–1513. [CrossRef] [PubMed]

82. Canning, C.G.; Ada, L.; Adams, R.; O'Dwyer, N.J. Loss of strength contributes more to physical disability after stroke than loss of dexterity. *Clin. Rehabil.* **2004**, *18*, 300–308. [CrossRef] [PubMed]

83. Koh, C.L.; Pan, S.L.; Jeng, J.S.; Chen, B.B.; Wang, Y.H.; Hsueh, I.P.; Hsieh, C.L. Predicting recovery of voluntary upper extremity movement in subacute stroke patients with severe upper extremity paresis. *PLoS ONE* **2015**, *10*, e0126857. [CrossRef] [PubMed]

84. Bohannon, R.W.; Ahlquist, M.; Lee, N.; Maljanian, R. Functional gains during acute hospitalization for stroke. *Neurorehabil. Neural Repair* **2003**, *17*, 192–195. [CrossRef] [PubMed]

85. Beebe, J.A.; Lang, C.E. Active range of motion predicts upper extremity function 3 months after stroke. *Stroke* **2009**, *40*, 1772–1779. [CrossRef] [PubMed]

86. Kasner, S.E. Clinical interpretation and use of stroke scales. *Lancet Neurol.* **2006**, *5*, 603–612. [CrossRef]

87. Winters, C.; van Wegen, E.E.; Daffertshofer, A.; Kwakkel, G. Generalizability of the proportional recovery model for the upper extremity after an ischemic stroke. *Neurorehabil. Neural Repair* **2015**, *29*, 614–622. [CrossRef] [PubMed]

88. Zarahn, E.; Alon, L.; Ryan, S.L.; Lazar, R.M.; Vry, M.S.; Weiller, C.; Marshall, R.S.; Krakauer, J.W. Prediction of motor recovery using initial impairment and fMRI 48 h poststroke. *Cereb. Cortex* **2011**, *21*, 2712–2721. [CrossRef] [PubMed]

89. Collen, F.M.; Wade, D.T.; Bradshaw, C.M. Mobility after stroke: Reliability of measures of impairment and disability. *Int. Disabil. Stud.* **1990**, *12*, 6–9. [CrossRef] [PubMed]

90. Mehrholz, J.; Wagner, K.; Rutte, K.; Meissner, D.; Pohl, M. Predictive validity and responsiveness of the functional ambulation category in hemiparetic patients after stroke. *Arch. Phys. Med. Rehabil.* **2007**, *88*, 1314–1319. [CrossRef] [PubMed]

91. Holden, M.K.; Gill, K.M.; Magliozzi, M.R.; Nathan, J.; Piehl-Baker, L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. *Phys. Ther.* **1984**, *64*, 35–40. [PubMed]

92. Nilanont, Y.; Komoltri, C.; Saposnik, G.; Côté, R.; Di Legge, S.; Jin, Y.; Prayoonwiwat, N.; Poungvarin, N.; Hachinski, V. The Canadian Neurological Scale and the NIHSS: Development and validation of a simple conversion model. *Cerebrovasc. Dis.* **2010**, *30*, 120–126. [CrossRef] [PubMed]

93. Carr, J.H.; Shepherd, R.B.; Nordholm, L.; Lynne, D. Investigation of a new motor assessment scale for stroke patients. *Phys. Ther.* **1985**, *65*, 175–180. [PubMed]

94. Poole, J.L.; Whitney, S.L. Motor assessment scale for stroke patients: Concurrent validity and interrater reliability. *Arch. Phys. Med. Rehabil.* **1988**, *69*, 195–197. [PubMed]

95. Hsieh, C.L.; Hsueh, I.P.; Chiang, F.M.; Lin, P.H. Inter-rater reliability and validity of the action research arm test in stroke patients. *Age ageing* **1998**, *27*, 107–113. [CrossRef] [PubMed]

96. Van der Lee, J.H.; De Groot, V.; Beckerman, H.; Wagenaar, R.C.; Lankhorst, G.J.; Bouter, L.M. The intra- and interrater reliability of the action research arm test: A practical test of upper extremity function in patients with stroke. *Arch. Phys. Med. Rehabil.* **2001**, *82*, 14–19. [CrossRef] [PubMed]

97. Nijland, R.; van Wegen, E.; Verbunt, J.; van Wijk, R.; van Kordelaar, J.; Kwakkel, G. A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test. *J. Rehabil. Med.* **2010**, *42*, 694–696.

98. Stinear, C.M.; Barber, P.A.; Smale, P.R.; Coxon, J.P.; Fleming, M.K.; Byblow, W.D. Functional potential in chronic stroke patients depends on corticospinal tract integrity. *Brain.* **2007**, *130*, 170–180. [CrossRef] [PubMed]

99. Puig, J.; Pedraza, S.; Blasco, G.; Daunis-i-Estadella, J.; Prados, F.; Remollo, S.; Prats-Galino, A.; Soria, G.; Boada, I.; Castellanos, M.; et al. Acute damage to the posterior limb of the internal capsule on diffusion tensor tractography as an early imaging predictor of motor outcome after stroke. *AJNR Am. J. Neuroradiol.* **2011**, *32*, 857–863. [CrossRef] [PubMed]

100. Jang, S.H.; Ahn, S.H.; Sakong, J.; Byun, W.M.; Choi, B.Y.; Chang, C.H.; Bai, D.; Son, S.M. Comparison of TMS and DTT for predicting motor outcome in intracerebral hemorrhage. *J. Neurol. Sci.* **2010**, *290*, 107–111. [CrossRef] [PubMed]

101. Kwon, Y.H.; Son, S.M.; Lee, J.; Bai, D.S.; Jang, S.H. Combined study of transcranial magnetic stimulation and diffusion tensor tractography for prediction of motor outcome in patients with corona radiata infarct. *J. Rehabil. Med.* **2011**, *43*, 430–434. [PubMed]

102. Bigourdan, A.; Munsch, F.; Coupé, P.; Guttmann, C.R.; Sagnier, S.; Renou, P.; Debruxelles, S.; Poli, M.; Dousset, V.; Sibon, I.; et al. Early fiber number ratio is a surrogate of corticospinal tract integrity and predicts motor recovery after stroke. *Stroke* **2016**, *47*, 1053–1059. [CrossRef] [PubMed]

103. Feng, W.; Wang, J.; Chhatbar, P.Y.; Doughty, C.; Landsittel, D.; Lioutas, V.A.; Kautz, S.A.; Schlaug, G. Corticospinal tract lesion load: An imaging biomarker for stroke motor outcomes. *Ann. Neurol.* **2015**, *78*, 860–870. [CrossRef] [PubMed]

104. Doughty, C.; Wang, J.; Feng, W.; Hackney, D.; Pani, E.; Schlaug, G. Detection and predictive value of fractional anisotropy changes of the corticospinal tract in the acute phase of a stroke. *Stroke* **2016**, *47*, 1520–1526. [CrossRef] [PubMed]

105. Kumar, P.; Kathuria, P.; Nair, P.; Prasad, K. Prediction of upper limb motor recovery after subacute ischemic stroke using diffusion tensor imaging: A systematic review and meta-analysis. *J. Stroke* **2016**, *18*, 50–59. [CrossRef] [PubMed]

106. Bembenek, J.P.; Kurczych, K.; Karli Nski, M.; Czlonkowska, A. The prognostic value of motor-evoked potentials in motor recovery and functional outcome after stroke—A systematic review of the literature. *Funct. Neurol.* **2012**, *27*, 79–84. [CrossRef] [PubMed]

107. Buma, F.E.; Lindeman, E.; Ramsey, N.F.; Kwakkel, G. Functional neuroimaging studies of early upper limb recovery after stroke: A systematic review of the literature. *Neurorehabil. Neural Repair* **2010**, *24*, 589–608. [CrossRef] [PubMed]

108. Buch, E.R.; Rizk, S.; Nicolo, P.; Cohen, L.G.; Schnider, A.; Guggisberg, A.G. Predicting motor improvement after stroke with clinical assessment and diffusion tensor imaging. *Neurology* **2016**, *86*, 1924–1925. [CrossRef] [PubMed]

109. Marshall, R.S.; Zarahn, E.; Alon, L.; Minzer, B.; Lazar, R.M.; Krakauer, J.W. Early imaging correlates of subsequent motor recovery after stroke. *Ann. Neurol.* **2009**, *65*, 596–602. [CrossRef] [PubMed]

110. Van Kuijk, A.A.; Pasman, J.W.; Hendricks, H.T.; Zwarts, M.J.; Geurts, A.C. Predicting hand motor recovery in severe stroke: The role of motor evoked potentials in relation to early clinical assessment. *Neurorehabil. Neural Repair* **2009**, *23*, 45–51. [CrossRef] [PubMed]

111. Saposnik, G.; Kapral, M.K.; Liu, Y.; Hall, R.; O'Donnell, M.; Raptis, S.; Tu, J.V.; Mamdani, M.; Austin, P.C. IScore: A risk score to predict death early after hospitalization for an acute ischemic stroke. *Circulation* **2011**, *123*, 739–749. [CrossRef] [PubMed]

112. Saposnik, G.; Raptis, S.; Kapral, M.K.; Liu, Y.; Tu, J.V.; Mamdani, M.; Austin, P.C. The iScore predicts poor functional outcomes early after hospitalization for an acute ischemic stroke. *Stroke* **2011**, *42*, 3421–3428. [CrossRef] [PubMed]

113. O'Donnell, M.J.; Fang, J.; D'Uva, C.; Saposnik, G.; Gould, L.; McGrath, E.; Kapral, M.K. The PLAN score: A bedside prediction rule for death and severe disability following acute ischemic stroke. *Arch. Intern. Med.* **2012**, *172*, 1548–1556. [CrossRef] [PubMed]

114. Reid, J.M.; Dai, D.; Christian, C.; Reidy, Y.; Counsell, C.; Gubitz, G.J.; Phillips, S.J. Developing predictive models of excellent and devastating outcome after stroke. *Age ageing* **2012**, *41*, 560–564. [CrossRef] [PubMed]

115. Cadilhac, D.A.; Lannin, N.A.; Anderson, C.S.; Levi, C.R.; Faux, S.; Price, C.; Middleton, S.; Lim, J.; Thrift, A.G.; Donnan, G.A. Protocol and pilot data for establishing the Australian Stroke Clinical Registry. *Int. J. Stroke* **2010**, *5*, 217–226. [CrossRef] [PubMed]

116. Saver, J.L.; Warach, S.; Janis, S.; Odenkirchen, J.; Becker, K.; Benavente, O.; Broderick, J.; Dromerick, A.W.; Duncan, P.; Elkind, M.S.; et al. Standardizing the structure of stroke clinical and epidemiologic research data: The National Institute of Neurological Disorders and Stroke (NINDS) Stroke Common Data Element (CDE) project. *Stroke* **2012**, *43*, 967–973. [CrossRef] [PubMed]

117. Zhang, N.; Liu, G.; Zhang, G.; Fang, J.; Wang, Y.; Zhao, X.; Pan, Y.; Guo, L.; Wang, Y. External validation of the iScore for predicting ischemic stroke mortality in patients in China. *Stroke* **2013**, *44*, 1924–1929. [CrossRef] [PubMed]

118. Ischemic Stroke Predictive Risk Score. Available online: http://www.sorcan.ca/iscore/ (accessed on 29 July 2016).

119. IScore—Ischemic Stroke Predictive Risk Score. Available online: https://itunes.apple.com/ca/app/iscore-ischemic-stroke-predictive/id423473762?mt=8 (accessed on 29 July 2016).

120. Carey, L.M.; Seitz, R.J. Functional neuroimaging in stroke recovery and neurorehabilitation: Conceptual issues and perspectives. *Int. J. Stroke* **2007**, *2*, 245–264. [CrossRef] [PubMed]

121. Ward, N.S. Does neuroimaging help to deliver better recovery of movement after stroke? *Curr. Opin. Neurol.* **2015**, *28*, 323–329. [CrossRef] [PubMed]

122. Dimyan, M.A.; Cohen, L.G. Contribution of transcranial magnetic stimulation to the understanding of functional recovery mechanisms after stroke. *Neurorehabil. Neural Repair.* **2010**, *24*, 125–135. [CrossRef] [PubMed]

123. Johnston, K.C.; Wagner, D.P.; Haley, E.C., Jr.; Connors, A.F., Jr. Combined clinical and imaging information as an early stroke outcome measure. *Stroke* **2002**, *33*, 466–472. [CrossRef] [PubMed]

124. Johnston, K.C.; Barrett, K.M.; Ding, Y.H.; Wagner, D.P. Clinical and imaging data at 5 days as a surrogate for 90-day outcome in ischemic stroke. *Stroke* **2009**, *40*, 1332–1333. [CrossRef] [PubMed]

125. Johnston, K.C.; Wagner, D.P.; Wang, X.Q.; Newman, G.C.; Thijs, V.; Sen, S.; Warach, S. Validation of an acute ischemic stroke model: Does diffusion-weighted imaging lesion volume offer a clinically significant improvement in prediction of outcome? *Stroke* **2007**, *38*, 1820–1825. [CrossRef] [PubMed]

126. Schiemanck, S.K.; Kwakkel, G.; Post, M.W.; Kappelle, L.J.; Prevo, A.J. Predicting long-term independency in activities of daily living after middle cerebral artery stroke: Does information from MRI have added predictive value compared with clinical information? *Stroke* **2006**, *37*, 1050–1054. [CrossRef] [PubMed]