Predicting Disability in Stroke—A Critical Review of the Literature

GERT KWAKKEL, ROBERT C. WAGENAAR, BOUDEWIJN J. KOLLEN, GUSTAAF J. LANKHORST

Summary

Research articles on the prognosis of stroke patients were analyzed to identify studies that met sound methodological principles of prognostic research as well as to identify variables capable of predicting functional outcome (ADL) after stroke. Data sources comprised a computer-aided search of published prognostic studies and references to literature used in prognostic studies. Seventy-eight studies were tested for adherence to the following key methodological criteria: reliability and validity of measurement instruments used to assess dependent and independent variables; inclusion of an inception cohort; adequate and uniform end-point of observation; control for drop-outs during period of observation; statistical testing of presumed relationship between dependent and independent variables; sufficient sample size in relation to number of determinants; control for multicollinearity; specification of patient characteristics (i.e. age, type, recurrent stroke and localization of stroke); description of interfering treatment effects during the period of observation; and cross-validation of the prediction model in a second independent group of patients.

Only three studies satisfied nine out of 11 criteria and ten studies eight criteria for the determination of valid prognostic research. The results of these studies indicate that the following variables are valid predictors for functional recovery after stroke: age; previous stroke; urinary continence; consciousness at onset; disorientation in time and place; severity of paralysis; sitting balance; admission ADL score; level of social support and metabolic rate of glucose outside the infarct area in hypertensive patients. This study supports the general opinion that not only are differences in objectives and heterogeneity in stroke patients responsible for the lack of accuracy in predicting functional outcome, but also the methodological flaws in published prognostic research.

Keywords: Cerebrovascular disorders, Literature review, Prognosis, Outcome, Activities of daily living, Physical disability.

Introduction

A Dutch prospective epidemiological study (TESS survey) demonstrated that 58% of the patients who survive a first stroke eventually regain independence in activities of daily life (ADL), while 82% of the stroke patients will walk independently again (with or without the use of a walking device or orthosis [1].) These results are confirmed in other epidemiological studies from Great Britain, Sweden, Norway and the United States [2-6].

Most functional recovery occurs within the first 2 months [7-10; see for reviews 11, 12]. Rehabilitation in this early post-stroke period is considered to be most beneficial and effective [9, 13-15; see for review 16]. There is less functional recovery at 4 to 5 months post-stroke and after 6 months little further functional recovery can be expected [7, 11, 12, 17, 18]. However, in some studies, improvements of hand function [19], ambulation [20] and ADL [8, 20, 21] have been observed after 6 months.

With the rising costs of stroke management [22, 23] and marked heterogeneity in stroke manifestation and recovery [24, 25], we need accurate and reliable predictors of functional recovery [24-27, 29]. We also need to know more about natural recovery patterns so that we can better assess the effectiveness of therapeutic interventions and their contribution to recovery [17, 29]. Early accurate prediction of outcome in stroke victims is important to: (1) set realistic and attainable therapeutic goals; (2) facilitate proper discharge planning; and (3) anticipate the need for home adjustments and community support.

Over 50 demographic, radiological, neurophysiological and neurological determinants claim individually or in combination to predict stroke outcome. The accuracy of the predictions varies from less than 40% [12, 30-32] to over 70% [33-36]. These differences may be caused by the choice of combination of predictors, differences in patient characteristics, timing of post-stroke observations, assessment instruments and applied methods of statistical analysis [24, 37-39]. Prediction models have not gained acceptance because
of doubts about their accuracy, their lack of applicability to other groups of patients, and practical difficulties in their use.

Gladman et al. determined the validity of five known prediction models [40]. Accuracy of prediction was found to be less than previously assumed. Differences in patient characteristics (e.g., age, stroke type and localization of stroke) may be responsible for this lack of accuracy; in particular with the application of complex multivariate models for the prediction of functional outcome [40]. This conclusion is supported by a prognostic study which validated the Edinburgh and the Guy’s prognostic score in 96 hospitalized stroke patients over 75 years of age [41].

Jongbloed [38] and Hier and Edelstein [37] recommend adequate stratification based on demographic and diagnostic data to increase external validity. Jongbloed chose six determinants which in her opinion successfully predicted functional outcome: these included level of disability on admission, age of patient, status following previous strokes, incontinence, as well as visual and spatial deficits [24]. No statistically significant relationship has been found between sex or side of stroke and functional outcome. It is unclear whether delay between onset of stroke and admission in hospital is a negative predictor for recovery [24].

Urinary incontinence, visual and spatial deficits, functional status on admission and a history of previous strokes indicate poor functional outcome [9]. Another seven factors may predict functional recovery: level of consciousness on admission, degree of hemiplegia, extent and localization of stroke on CT scan, cardiac abnormalities, multiple neurological disorders, attention and failure to develop any functional recovery within the first month post-stroke [9]. Hemianopia, old age, somatosensory deficiencies and side of stroke are possible negative predictors [9].

To ensure validity in prognostic studies that examine recovery in stroke patients two questions need to be addressed:
1. Which studies meet accepted methodological criteria for evaluating the quality of prognosis research in stroke?
2. Which variables successfully predict the recovery of disability after stroke?

**Method**

**Inclusion criteria:** The definition of stroke precludes 'transient ischaemic attacks' (TIAs) [42], and studies that include patients with TIAs will not be considered for further analysis. The present review evaluates disability corresponding to the codes 30–46 of the ICIDH (WHO [43] pages 157–162) referring to personal care and ambulation. Recovery from disability is defined as the final level of ADL independence, and improvement of ADL as the difference between functional status on admission and at a specific time after stroke. Prognostic studies that focus on extended ADL (EADL), which includes 'transport disability' and 'other locomotor disabilities' (codes 47, 48) [44, 45], and particular skills or focal disabilities such as upper extremity skills or ambulation, are not considered here.

**Literature search strategy:** The following sources were consulted for the collection of literature:

1. MEDLINE (CD-ROM), Excerpta Medica and Current Contents (1966–94) [46]. The following keywords were used: Cerebrovascular disorders, Activities of daily living, Self care, Physical disability, Prognosis and Outcome;
2. References to literature used in other (review) studies;
3. Indexes of journals which are not included in the MEDLINE system.

All studies were collected without consideration of research design. Subsequently, adherence of the studies to definitions of stroke and disability was established and a methodological scheme was applied that enabled the evaluation of threats to internal, statistical and external validity [47–51]. This scheme contains criteria which are recommended by the 'Task Force on Stroke Outcome Research of Impairments, Disabilities and Handicap' in order to improve the scientific quality and comparability of stroke outcome research [52]. The methodological guidelines are in agreement with the general recommendations for studying prognosis [49, 51, 53] and experimental literature in this field [16, 47].

**Internal validity:**

1. **Reliable or valid measurements**
   - Measurements of ADL function as well as determinants need to be reliable and valid [49]. In this study, reliability includes both reproducibility among observers and consistency among scale items. An assessment instrument is accepted as reliable if (a) the agreement between measurements exceeds a kappa statistic of 0.75; and (b) the inter-class correlation coefficient (ICC) has a p value below 0.05 [54, 55]. We consider (construct) validity demonstrated when relations with other tests have been established and have been tested for statistical significance (p < 0.05) [56].

2. **Inception cohort**
   - An inception cohort was identified if patients entered the study as early as possible within 2 weeks following the onset of stroke [52, page 11-68].

3. **Appropriate end-points of observation**
   - A minimal observation period of 6 months after stroke onset is considered to be necessary as functional recovery can be expected to be most significant within this period [16, 57].

4. **Control of patient drop-out**
   - Specification of patient withdrawal caused by migration, death or other reasons should be provided in order to prevent a biased outcome of results [49, 51, 58].

**Statistical validity:**

1. **Control for statistical significance**
   - Statistical validity is regarded as satisfactory when (i) the method for testing the relationship between dependent and independent variables is specified, and (ii) the explained variance is tested for statistical significance.

2. **Adequate estimation of sample size**
   - A sample size is considered to be adequate if at least ten patient cases are examined for each predictor [37, 51].

7. **Control for multicollinearity**
   - Control for multicollinearity is regarded as sufficient if in the multivariate analysis the relation between an outcome variable and two or more independent variables is partialled out for interaction [59, 60].
External validity:

8. Specification of inclusion and exclusion criteria
   The patient characteristics age, type, localization and number of sustained strokes should be specified [52].

9. Description of additional treatment effects during period of observation
   Paramedical and medical interventions should have been reported during the period of observation [51]. It should be noted that additional treatments are also a potential confounder for discovering discriminants for functional recovery after stroke.

10. Cross-validation of the prediction model
    The prediction model should be verified for validity and accuracy by cross-validating the derived model in a second study on an independent group of stroke patients [37, 51, 61].

Procedure:
The above mentioned factors concerning internal, statistical and external validity were used for the evaluation of the prognostic studies. All prognostic studies were scored according to the criteria listed in Table I. In total, ten methodological key principles were registered and used for 11 criteria. To each criterion a binary weight is attached (0/1) and a maximum score is set at 11 points for each study. The methodologic quality of the prognostic studies was assessed by one reviewer (G.K.). The results were converted into binary numerals and classified with the computer program SPSS.PC+ (Version 5.0).

Results
In total 142 prognostic studies were reviewed. Sixty-four studies did not meet the inclusion criteria for further analysis. Seventy-eight prognostic studies, involving 16149 patients with cerebrovascular disorders, were eligible for further analysis [3, 8, 19, 20, 30, 31, 33-36, 40, 41, 49, 62-126]. The results of the analysis of internal, statistical and external validity are summarized in Table II. Four of these involve validation studies of pre-existing prediction models [40, 41, 85, 99].

Internal validity: Reliable and valid instruments for ADL measurement were found in 41 studies. The Barthel Index [127] or its modification [128] proved to be the most popular scale to measure ADL independence (n = 29). Less than half of the studies (n = 31) employed ADL scales that lacked prior testing of

---

Table I. Binary outcome strategies rated 'adequate'

<table>
<thead>
<tr>
<th>Outcome strategies</th>
<th>Scale</th>
<th>Criteria</th>
<th>Criterion in Table II:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A &amp; B</td>
<td></td>
</tr>
<tr>
<td>To evaluate internal validity</td>
<td></td>
<td>Positive, if the prognostic study tested the reliability and validity of used measurements or referred to other studies which established reliability and validity</td>
<td>A &amp; B</td>
</tr>
<tr>
<td>1. Measurements reliable and valid?</td>
<td>Binary</td>
<td>Positive, if observation started within 2 weeks post-stroke</td>
<td>C</td>
</tr>
<tr>
<td>dependent variable (A)</td>
<td></td>
<td>Positive, if observation ended a minimal 6 months post-stroke</td>
<td>D</td>
</tr>
<tr>
<td>independent variable (B)</td>
<td></td>
<td>Positive, if drop-outs during period of observation are specified</td>
<td>E</td>
</tr>
<tr>
<td>2. Inception cohort during period of observation?</td>
<td></td>
<td>Positive, if relationship between dependent and independent variable is tested for statistical significance</td>
<td>F</td>
</tr>
<tr>
<td>3. Appropriate end-points of observation?</td>
<td></td>
<td>Positive, if ratio n : K exceeds 10 : 1</td>
<td>G</td>
</tr>
<tr>
<td>4. Control for drop-outs?</td>
<td></td>
<td>Positive, if interaction between two or more independent variables is tested in the prediction model</td>
<td>H</td>
</tr>
<tr>
<td>To evaluate statistical validity</td>
<td></td>
<td>Positive, if age, type, localization as well as number of strokes are specified in the cohort</td>
<td>I</td>
</tr>
<tr>
<td>5. Statistical validation of relationship between dependent and independent variables?</td>
<td></td>
<td>Positive, if information on (para)medical treatment was reported</td>
<td>J</td>
</tr>
<tr>
<td>6. Sample size (n) adequate in relation to the number of determinants (K)?</td>
<td></td>
<td>Positive, if the prediction model is validated in a second independent group of stroke patients</td>
<td>K</td>
</tr>
<tr>
<td>7. Control for multicollinearity?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To evaluate external validity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Specification of relevant patient characteristics? (i.e. age, type, number and localization of stroke)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Description of additional medical and paramedical interventions during observation?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Cross-validation of the prediction model in a second independent group?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table II. Scores for methodological characteristics of prognostic studies predicting functional outcome after stroke

<table>
<thead>
<tr>
<th>References</th>
<th>Criterion(^\star) internal</th>
<th>statistical</th>
<th>external</th>
<th>Score (max.(^\star) = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First author</td>
<td>Year</td>
<td>Ref. no.</td>
<td>n</td>
<td>A</td>
</tr>
<tr>
<td>Adams</td>
<td>1960</td>
<td>62</td>
<td>736</td>
<td>−</td>
</tr>
<tr>
<td>Allen</td>
<td>1984</td>
<td>63</td>
<td>148</td>
<td>+</td>
</tr>
<tr>
<td>Andrews</td>
<td>1984</td>
<td>64</td>
<td>135</td>
<td>−</td>
</tr>
<tr>
<td>Aron</td>
<td>1994</td>
<td>65</td>
<td>227</td>
<td>+</td>
</tr>
<tr>
<td>Basker</td>
<td>1995</td>
<td>66</td>
<td>330</td>
<td>+</td>
</tr>
<tr>
<td>Baxman</td>
<td>1987</td>
<td>70</td>
<td>109</td>
<td>+</td>
</tr>
<tr>
<td>Bender</td>
<td>1987</td>
<td>71</td>
<td>120</td>
<td>−</td>
</tr>
<tr>
<td>Besh</td>
<td>1987</td>
<td>72</td>
<td>106</td>
<td>−</td>
</tr>
<tr>
<td>Best</td>
<td>1980</td>
<td>73</td>
<td>80</td>
<td>−</td>
</tr>
<tr>
<td>Chest</td>
<td>1989</td>
<td>74</td>
<td>26</td>
<td>+</td>
</tr>
<tr>
<td>Carroll</td>
<td>1986</td>
<td>8</td>
<td>724</td>
<td>+</td>
</tr>
<tr>
<td>Davalos</td>
<td>1990</td>
<td>75</td>
<td>68</td>
<td>+</td>
</tr>
<tr>
<td>Davers</td>
<td>1991</td>
<td>76</td>
<td>166</td>
<td>+</td>
</tr>
<tr>
<td>Dignan</td>
<td>1986</td>
<td>8</td>
<td>724</td>
<td>+</td>
</tr>
<tr>
<td>Dow</td>
<td>1991</td>
<td>85</td>
<td>344</td>
<td>+</td>
</tr>
<tr>
<td>Drake</td>
<td>1993</td>
<td>86</td>
<td>555</td>
<td>−</td>
</tr>
<tr>
<td>Falcollo</td>
<td>1994</td>
<td>79</td>
<td>225</td>
<td>+</td>
</tr>
<tr>
<td>Feigenson</td>
<td>1992</td>
<td>80</td>
<td>224</td>
<td>−</td>
</tr>
<tr>
<td>Gil</td>
<td>1993</td>
<td>81</td>
<td>46</td>
<td>+</td>
</tr>
<tr>
<td>Gomperz</td>
<td>1988</td>
<td>84</td>
<td>361</td>
<td>+</td>
</tr>
<tr>
<td>Gordon</td>
<td>1988</td>
<td>86</td>
<td>226</td>
<td>+</td>
</tr>
<tr>
<td>Gowland</td>
<td>1984</td>
<td>88</td>
<td>36</td>
<td>+</td>
</tr>
<tr>
<td>Greger</td>
<td>1989</td>
<td>88</td>
<td>539</td>
<td>+</td>
</tr>
<tr>
<td>Heid</td>
<td>1993</td>
<td>89</td>
<td>119</td>
<td>+</td>
</tr>
<tr>
<td>Heinemann</td>
<td>1987</td>
<td>90</td>
<td>422</td>
<td>+</td>
</tr>
<tr>
<td>Heis</td>
<td>1991</td>
<td>91</td>
<td>76</td>
<td>−</td>
</tr>
<tr>
<td>Herstatt</td>
<td>1988</td>
<td>92</td>
<td>41</td>
<td>+</td>
</tr>
<tr>
<td>Johansson</td>
<td>1987</td>
<td>93</td>
<td>66</td>
<td>+</td>
</tr>
<tr>
<td>Johnson</td>
<td>1988</td>
<td>94</td>
<td>90</td>
<td>+</td>
</tr>
<tr>
<td>Kabi</td>
<td>1993</td>
<td>95</td>
<td>116</td>
<td>+</td>
</tr>
<tr>
<td>Kaplan</td>
<td>1989</td>
<td>97</td>
<td>87</td>
<td>−</td>
</tr>
<tr>
<td>Kaste</td>
<td>1991</td>
<td>98</td>
<td>32</td>
<td>+</td>
</tr>
<tr>
<td>Katz</td>
<td>1995</td>
<td>99</td>
<td>65</td>
<td>−</td>
</tr>
<tr>
<td>Kojima</td>
<td>1996</td>
<td>100</td>
<td>199</td>
<td>+</td>
</tr>
<tr>
<td>Kotli</td>
<td>1996</td>
<td>101</td>
<td>154</td>
<td>−</td>
</tr>
<tr>
<td>Lajoie</td>
<td>1988</td>
<td>102</td>
<td>68</td>
<td>−</td>
</tr>
<tr>
<td>Lincoln</td>
<td>1989</td>
<td>103</td>
<td>68</td>
<td>+</td>
</tr>
<tr>
<td>Linde</td>
<td>1990</td>
<td>104</td>
<td>60</td>
<td>+</td>
</tr>
<tr>
<td>Linder</td>
<td>1990</td>
<td>105</td>
<td>50</td>
<td>+</td>
</tr>
<tr>
<td>Mills</td>
<td>1983</td>
<td>106</td>
<td>110</td>
<td>−</td>
</tr>
<tr>
<td>Nakam</td>
<td>1994</td>
<td>107</td>
<td>515</td>
<td>+</td>
</tr>
<tr>
<td>Novack</td>
<td>1984</td>
<td>108</td>
<td>687</td>
<td>−</td>
</tr>
<tr>
<td>Olsen</td>
<td>1987</td>
<td>109</td>
<td>64</td>
<td>+</td>
</tr>
<tr>
<td>Pankh</td>
<td>1987</td>
<td>110</td>
<td>63</td>
<td>+</td>
</tr>
<tr>
<td>Parnes</td>
<td>1987</td>
<td>111</td>
<td>368</td>
<td>+</td>
</tr>
<tr>
<td>Pau</td>
<td>1986</td>
<td>112</td>
<td>130</td>
<td>−</td>
</tr>
<tr>
<td>Pett</td>
<td>1982</td>
<td>113</td>
<td>149</td>
<td>−</td>
</tr>
<tr>
<td>Pinto</td>
<td>1987</td>
<td>114</td>
<td>112</td>
<td>−</td>
</tr>
<tr>
<td>Rankin</td>
<td>1987</td>
<td>115</td>
<td>192</td>
<td>−</td>
</tr>
<tr>
<td>Reding</td>
<td>1988</td>
<td>116</td>
<td>95</td>
<td>−</td>
</tr>
<tr>
<td>Robinson</td>
<td>1988</td>
<td>117</td>
<td>103</td>
<td>−</td>
</tr>
<tr>
<td>Rose</td>
<td>1994</td>
<td>118</td>
<td>26</td>
<td>+</td>
</tr>
<tr>
<td>Sand</td>
<td>1989</td>
<td>119</td>
<td>22</td>
<td>+</td>
</tr>
<tr>
<td>Sali</td>
<td>1989</td>
<td>120</td>
<td>258</td>
<td>+</td>
</tr>
<tr>
<td>Sheikh</td>
<td>1983</td>
<td>121</td>
<td>900</td>
<td>+</td>
</tr>
<tr>
<td>Sinyor</td>
<td>1986</td>
<td>122</td>
<td>64</td>
<td>−</td>
</tr>
<tr>
<td>Smith</td>
<td>1985</td>
<td>123</td>
<td>128</td>
<td>−</td>
</tr>
<tr>
<td>Stone</td>
<td>1989</td>
<td>124</td>
<td>155</td>
<td>−</td>
</tr>
<tr>
<td>Sundet</td>
<td>1989</td>
<td>125</td>
<td>183</td>
<td>−</td>
</tr>
<tr>
<td>Taub</td>
<td>1992</td>
<td>126</td>
<td>225</td>
<td>+</td>
</tr>
<tr>
<td>Ueda</td>
<td>1996</td>
<td>127</td>
<td>155</td>
<td>−</td>
</tr>
<tr>
<td>Zeman</td>
<td>1989</td>
<td>128</td>
<td>35</td>
<td>+</td>
</tr>
</tbody>
</table>

**Footnotes:** see opposite.
validity and reliability. In 29 studies all the assessments of determinants were reliable and valid and included items such as age [31, 63, 64, 124], sex [121, 124], side of lesion [100, 121–124], level of ADL independence on admission [31, 34, 76, 90, 92], and consciousness [122, 124]. In another 36 studies (46%) only some of the determinants could be considered reliable and valid [31, 67, 68, 70, 76, 87, 91, 95, 96, 101, 103, 108, 110, 117]. However, when the criteria of reliability and validity were applied to dependent as well as independent variables in the prediction model, only 21 studies satisfied these criteria.

In 38 studies, the first observation took place within 2 weeks post-stroke. The timing of initial observation varied from several days [3, 34, 36, 63, 64, 67, 78, 84, 92, 96, 117, 122, 124] to several months [79, 80, 87, 102, 103, 121]. Only 21 studies, in which the initial observation occurred within the first 2 weeks post-stroke, observed a total period equal to or exceeding 6 months.

Drop-out details (such as death, migration) were provided in 62 studies.

When criteria to verify internal validity were applied to the 78 reviewed studies, only three satisfied all four criteria [84, 119, 124] and another 12 studies fulfilled three out of four requirements [41, 67, 68, 75, 76, 85, 89, 91, 92, 101, 117, 126]. However, in these latter studies either the assessment of determinants was not reliable or valid or the timing of the final observation was lacking.

Statistical validity: In 69 studies a statistical test was conducted to establish the relation between independent and dependent variables. Linear or logistic regression analysis was most popular and was used in 43 studies. One study indicated probability values, but failed to document information regarding the applied statistical procedure [72].

In six out of 69 studies the number of determinants was insufficient in relation to the number of patients observed [69, 74, 82, 92, 95, 112].

Forty-one studies, in which multivariate analysis was used, were corrected for interaction between predictors by differentiating partial and raw-correlation coefficients. In doing so they permitted insight into the explained variance of the determinants.

Thirty-six studies complied completely with all three criteria designed to test statistical validity.

External validity: First and recurrent strokes were differentiated in 43 studies, and 42 studies provided additional information regarding stroke type and localization, while the remaining 35 studies failed to document these details. However, none of the studies, in which the type and localization of stroke were specified, differentiated between diagnostic strata in their prediction model. Some studies confined their target group to, for instance, ischaemic strokes in the region supplied by the anterior cerebral arteries [33, 75, 96, 110, 115, 125], left or right hemispheres [69, 74, 97] or supratentorial haemorrhagic strokes [76, 108]. Altogether, only 31 studies provided information about stroke characteristics as well as recurrent strokes in the cohort.

Only 14 studies reported brief information about the nature of medical and paramedical interventions during the period of observation.

Finally, 13 studies validated a prediction model with that of a second stroke group or were specifically focused on validation [40, 41, 99] or improvement [85] of an already published prediction model in an independent cohort.

Overview of adherence to methodological principles: Table II presents the results of this critical review, indicating the binary scores for methodological characteristics for each prognostic study. None of the reviewed 78 studies complied completely with all 11 criteria set to test for internal, statistical and external validity. The three best studies obtained nine out of a maximum of 11 points [68, 84, 91] and ten studies satisfied eight criteria [35, 41, 67, 75, 85, 99, 105, 117, 119, 124].

Which variables successfully predict the recovery of disability after stroke? In spite of the above-mentioned shortcomings, eight cohort studies have been selected. These studies meet all criteria for internal and statistical validity (8 points or more), with the exception of the criteria for reliable and valid determinants. From these studies, only the findings based on valid and reliable determinants were selected. The latter studies demonstrate a trend towards the ability to predict functional recovery in stroke patients, i.e. disability on admission (within 2 weeks after stroke onset) [117, 119, 124]; urinary continence [68, 117, 119, 124]; degree of motor paresis or paralysis [68, 85, 117, 124]; old age [85, 91, 117, 124]; loss of consciousness within the first 48 hours post-stroke [85, 119, 124]; disorientation in time and place [67, 85, 124], poor sitting balance [124] as well as status following recurrent stroke [124]. There are strong indications that high level of perceived social support may be associated with faster and better functional outcome after stroke, but sample size limitations make it impossible to rule out other confounding factors [84]. New clinically introduced assessment techniques such as positron emission tomography (PET).
Demonstrate that functional outcome is significantly affected by cerebral metabolic rate outside the area of infarction in hypertensive subjects [91]. Different results are published for the predicting ability of visuospatial disturbances [compare 117 and 67, 124]. Homonymous hemianopia [67, 68, 85] and conjugate deviation of the eyes [68] are possible negative predictors for disability. Finally, sex [119, 124], ethnic origin [119, 124], and side of stroke [124] are unlikely to be relevant predictors.

Discussion

Only 78 of 142 collected studies predicted functional outcome with respect to ADL after stroke. Many studies (45%) measured destination after discharge, timing of destination, functional improvement, specific skills or perceived health status after stroke. In spite of considerable differences in their primary goals of prognostic research 24 out of 33 reviewed articles predicted functional status after stroke [24]. Only a few reports adhered to a minimum number of key methodological principles. None of the 78 studies complied completely with the main methodological criteria set for a prognostic study. As a consequence, an objective interpretation of putative relations between determinants and functional outcome is difficult. None the less, the 14 best studies used reliable and valid instruments to assess ADL and incorporated an inception cohort (with the exception of reference [99]). Because these articles were published in the past decade, an improvement in methodologic quality of prognostic research has been established.

Internal validity: Unreliable or invalid measurement instruments of ADL were responsible for a reduction in internal validity in half of the studies. This finding corresponds with Jongbloed's literature review [24] in which 63% of the studies measured ADL with instruments of no reported reliability or validity. There are no generally accepted criteria available for assessing ADL after stroke [129]. The Barthel Index appears to be the most frequently used assessment instrument (38%). Although only minor differences exist between popular ADL measurement scales such as the Barthel Index, Donaldson scale, Kenny Self-care Evaluation, and Katz Index, these tests do not correspond very well with the disability codes of the ICIDH concept [130, pages 69-71]. In order to obtain standardized assessment methods, ADL measurements containing the disability items evolved from ICIDH have been recommended recently [131, 130].

Clinically relevant determinants such as proprioception, sitting balance, spasticity, loss of consciousness, paresis, as well as some neuropsychological disorders (e.g. apraxia) are generally regarded as important predictors for functional outcome. However, the reliability and validity of the assessment of these determinants is not clear in all reviewed studies. However, the lack of standardization of most clinical assessment methods may have led to some factors not being reliably and/or validly assessed. Therefore the factors listed in this review as being significant may simply represent more easily measured clinical impairments or disabilities.

More than two-thirds of the studies failed to allow an adequate length of time for functional recovery. For instance, many studies are initiated in a rehabilitation facility. Referring a patient for stroke rehabilitation is not uniformly initiated and many be biased by non-medical factors such as regional admission policies. In many studies the final observation coincided with the timing of discharge, ranging from a month up to a year post-stroke [7, 20]. Studies confirm that destination after discharge and timing of discharge are often determined by non-medical factors [29] such as the presence of a healthy partner at home [133-135]. The lack of uniformity in timing of observations could explain this variance.

The interval between onset of stroke and initial observation was 37 days in the study by Novack et al. [102], but only 6 days in the study by Hertanu et al. [92]. Prognostic studies require an inception cohort, in which the first post-stroke observation occurs soon after the onset, followed by standardized intervals between subsequent observations [24, 38, 51-53]. Generally, the presence of neurological symptoms, such as urinary incontinence, impaired sitting balance or homonymous hemianopia, is temporary. The presence or severity of each symptom in its relation to disability is determined by the interval between onset of stroke and observation.

Finally, patient drop-out due to premature post-stroke death or recurrent stroke is likely to affect expected disability in early initiated prognostic research. In particular, haemorrhagic strokes can affect outcome of disability adversely, because of a high incidence of mortality in the acute post-stroke phase [136-139]. A number of early initiated post-stroke studies failed to specify the differences in patient numbers between initial and final observations.

Statistical validity: Lack of statistical validity is partly caused by failure to test found correlations for probability. As observation intervals are often not fixed [59], systematic differences and a possible biased relationship may result. To overcome the problem of a large individual variation in timing of initial observation, Reding recommended the use of survival analysis because this technique is not sensitive to the timing of observation [27]. Another way to minimize variance can be established by using non-linear regression [30] or log-linear regression [34, 76] methods. The graphic representation of post-stroke functional recovery is exponential [11] or sigmoidal [26] and most recovery occurs within the first few months post-stroke [7, 11, 26, 63]. Wagenaar et al. [26] found a predominantly logistic recovery pattern of skilled movements with multiple observations in time. However, the multiple observations that are required in time series experiments are the most likely reason for their lack of popularity.
It has been suggested by Leigh that the results of a stepwise regression procedure are arbitrary to some extent [140]. This may be reflected in the differences found in published explained variance of predictors in multivariate models, particularly when those clinical predictors are highly interrelated. In his opinion, stepwise routines lead to a problem when two variables (e.g. initial disability index and hemiplegia) are highly correlated. In a stepwise regression model, choosing one of these variables at an early stage will decrease the predictive power of the second variable, which ultimately may be excluded from the model. Hence, when using the two variables mentioned above, initial disability index may be included, while hemiplegia is not. However, the reverse selection may also be implemented.

In six out of 78 reviewed studies the number of observed patients was too small when taking the number of determinants into account. In this study, the three criteria for evaluation of statistical validity are not very specific. For example, another perceived problem pertains to the use of parametric regression methods in studies with ordinal dependent variables, which violates basic assumptions of proper regression analysis [141].

Neither the selected statistical methods for optimal estimation of variance nor (parametric) regression methods for ADL measurement are included for further consideration in this review of selected literature. However, these items need to be considered in future research.

External validity: Approximately 60% of prognostic studies did not specify patient characteristics such as type, location of stroke and number of recurrent strokes. Consequently the percentage of patients afflicted by subarachnoid haemorrhage, brainstem stroke or recurrent strokes could not be traced. Moreover, all the studies reviewed failed to distinguish between diagnostic strata (i.e. localization and extent of ischaemic or haemorrhagic stroke) in the prediction model. Reding [27] blamed this lack of diagnostic strata mainly on the lack of reliable and valid assessment instruments for specifying either the location or the extent of stroke. A more meaningful, distinct classification for different groups of stroke patients is required for future research. Only a few studies included pre-morbid data regarding pertinent patient characteristics [e.g. 20, 124]. Thus possible interactions may occur between stroke and cardiac disease [116], recurrent strokes [8, 20, 124] or other pre-existing degenerative diseases affecting mobility in elderly people. The more frequently occurring pre-existing disabilities in elderly patients may indeed be responsible for age being a negative predictor in recovery of disability after stroke [see 39, 41, 80]. However, there is no consensus on this subject [24, 80, 101, 142]. Some studies reported a negative influence of age on improvement [101, 143, 144] and outcome of disability [19, 20, 31, 63, 73, 85, 92, 101], but other studies did not [see, respectively, 145-149 and 80, 90]. A recent study suggested poor compensatory ability in elderly stroke patients as well as poor initial ADL-related aspects, but found no age-related differences in neurological improvement. In addition, a 10-year increase in age resulted in a 3-point decrease in Barthel Index gain [101].

Conclusion
This review of selected best studies has established the ability of a number of determinants to predict functional recovery after stroke. These include disability on admission, urinary continence, degree of motor paresis, age, level of consciousness within the first 48 hours post-stroke, orientation in time and place, status following recurrent stroke, sitting balance and level of perceived social support. Jongbloed identified five predictors [24], while Domboy et al. proposed ten poor prognostic indicators for disability [9]. These differences may be caused by a systematic approach to the application of relevant methodological criteria for qualifying prognostic research, the use of a computer-aided search for collecting articles regarding the prediction of disability after stroke, and the inclusion of newly published prognostic research of improved methodological quality in the past decade.

In future prognostic research, flaws in internal and statistical validity need to be addressed. The establishment of generally accepted standardized methodological criteria in prognostic stroke studies is required, because some criteria are considered arbitrary. Reaching consensus for all criteria is paramount. In addition, for the sake of closing the gap between research and practice there is a need to improve the applicability as well as accuracy of prediction models. In order to improve the applicability of prediction models for clinicians, measurements of relevant prognostic variables as well as algorithms of multivariate models should be simple and multivariate models should be suitable for clinical use. In order to improve the accuracy of prognostic research for relatively homogeneous patient groups, a better understanding is required of the individual time course and functional recovery patterns of ADL. To study natural recovery patterns in stroke patients, repeated measurements on the same subjects are recommended, particularly during the initial weeks [26]. Applying new statistical techniques such as growth curve analysis would enable us to model the time course of ADL. These regression techniques may result in the development of new models that combine determinants (e.g. initial value of a robust determinant such as disability at onset and the improvement of disability during initial weeks post onset).

Acknowledgements
The authors wish to thank Mrs E. van Deventer (Department of Neurology, Vrije Universiteit, Amsterdam) for collecting and copying relevant references. This project is part of a research project supported by a grant from the Netherlands Heart Foundation, reference no. 93.134.
References

42. Aho K, Harsma S, Hatano S, Markquardsen J, Smirnov VE, Strasser T. Cerebrovascular disease in the


Authors' addresses

G. Kwakkel, R. C. Wagenaar
Department of Physiotherapy,
G. J. Lankhorst
Department of Rehabilitation,
University Hospital Vrije Universiteit, PO Box 7057,
1007 MB Amsterdam,
The Netherlands

B. J. Kollen
Department of Physiotherapy,
Sophia Hospital,
Zwolle, The Netherlands

Received in revised form 26 April 1996