

**Table 1.** Characteristics of the 36 included studies

<b>Study characteristics</b>	<b>Number (%) of studies</b> <i>N=36</i>
Study location	
United States	10 (27·8)
Italy	6 (16·7)
Australia	3 (8·3)
China	3 (8·3)
Other	14 (38·9)
Funding source(s)	
Governments, universities, and/or other public bodies	19 (52·8)
Private research foundations or medical centres	7 (19·4)
Not reported	26 (27·8)
Study design	
Prospective cohort	18 (50·0)
Retrospective cohort	18 (50·0)
Study design	
Prospective cohort	18 (50·0)
Retrospective cohort	18 (50·0)
Measurement of mortality	
Based on data retrieved from government registries, insurance databases, or electronic health records	30 (83·3)
Based on data obtained from follow-up with participants or family members	2 (5·6)
Not reported	4 (11·1)
Use of collected data	
Model development and validation	19 (52·8)
Model validation only	16 (44·4)
Model recalibration	1 (2·8)
Number of participants	
<1,000	11 (30·6)
1,000 to 9,999	16 (44·4)
≥10,000	9 (25·0)

**Table 2.** Characteristics of the 74 model validations done in included studies

Model validation characteristics	Number (%) of model validations <i>N</i> =74
Time-horizon for prediction*	
≤3 months	33 (44·6)
6 months	13 (17·6)
1 year	46 (62·2)
>1 year	13 (17·6)
Source of data <sup>†</sup>	
Electronic health records	44 (59·5)
Questionnaires administered by staff	29 (39·2)
Clinical assessments	28 (37·8)
Previous survey or trial data	11 (14·9)
Self-report assessments	3 (4·1)
Time of prediction	
Any time	28 (37·8)
During ED attendance	19 (25·7)
During inpatient admission	23 (31·1)
During CGA	4 (5·4)
Evaluation of discrimination*	
AUC	46 (62·2)
Harrell's <i>c</i> -statistic	21 (28·4)
Other <sup>‡</sup>	10 (13·5)
Not reported	7 (9·5)
Level of discrimination (where reported as AUC or Harrell's <i>c</i> -statistic)*	
Excellent (≥0·80)	12 (16·2)
Acceptable (0·70 to 0·79)	60 (81·1)
Poor (0·50 to 0·69)	33 (44·6)
Not reported	7 (9·5)
Evaluation of calibration*	
Calibration plot analysis with explicit interpretation	12 (16·2)
Calibration plot analysis without explicit interpretation	11 (14·9)
Hosmer-Lemeshow test results	4 (5·4)
Calibration-in-the-large	1 (1·4)
Other <sup>§</sup>	4 (5·4)
Not reported	50 (67·7)
Level of calibration (based on authors' interpretations or Hosmer-Lemeshow test) <sup>¶</sup>	
Adequate	19 (25·7)
Poor	1 (1·4)
Not reported	50 (67·6)

AUC: Area under the receiver operating characteristic curve; CGA: Comprehensive geriatric assessment; ED: Emergency department.

\*As validation could be done at multiple time-horizons or report multiple measures of discrimination or calibration, the sum of the counts exceeds the number of model validations (*N*=74)

<sup>†</sup>As validation could be done in more than one source of data, the sum of the counts may exceed the number of included model (*N*=74)

<sup>‡</sup>Other measures of discrimination included sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios

<sup>§</sup>Other measures of calibration included calibration intercept, calibration slope, calibration error, observed-to-expected ratio, and correct classification

<sup>¶</sup>A *p*-value of ≥0·05 in the Hosmer-Lemeshow test was taken to indicate adequate calibration

**Table 3.** Predictors used by the 64 unique prediction models examined

<b>Predictor</b>	<b>Frequency (%)</b> <i>N=64 unique models</i>
Any morbidity	45 (70.3)
Cardiovascular diseases	39 (60.9)
Cancer	37 (57.8)
Neurological/psychiatric diseases	35 (54.7)
Respiratory diseases	34 (53.1)
Renal diseases	33 (51.6)
Metabolic diseases	30 (46.9)
Gastrointestinal diseases	28 (43.8)
Liver diseases	27 (42.2)
Urogenital/sexually transmitted diseases	26 (40.6)
Rheumatological diseases	25 (39.1)
Haematological diseases	14 (21.9)
Age	33 (51.6)
Sex/gender	30 (46.9)
Blood test & urinalysis	17 (26.6)
Physical status	14 (21.9)
Previous admissions/length of stay	13 (20.3)
Mental & cognitive status	13 (20.3)
ADL & IADL	12 (18.8)
Medications	11 (17.2)
Race & ethnicity	9 (14.1)
Alcohol intake/alcoholism	9 (14.1)
Weight	8 (12.5)
Residency status	8 (12.5)
Vital signs	8 (12.5)
Number of comorbidities	7 (10.9)
Marital status	6 (9.4)
Smoking	6 (9.4)
Number of predictors	
1 to 10	28 (43.8)
11 to 20	17 (26.6)
>20	18 (28.1)
Not reported	1 (1.6)

ADL: Activities of Daily Living; IADL: Instrumental Activities of Daily Living.

**Table 4.** Model characteristics associated with prediction models having acceptable discrimination

Model characteristic	Acceptable discrimination* Number (%) of models	RR (95%CI)	P-value
<b>1 month time-horizon (N=27)</b>			
Model uses morbidities as a predictor			
No	6 (54.5)	Reference	
Yes	9 (56.3)	1.03 (0.52, 2.06)	0.931
Model uses age as a predictor			
No	7 (70.0)	Reference	
Yes	8 (47.1)	0.67 (0.35, 1.28)	0.229
Model uses sex as a predictor			
No	7 (46.7)	Reference	
Yes	8 (66.7)	1.43 (0.73, 2.80)	0.299
Model adopts clinical assessment data			
No	8 (66.7)	Reference	
Yes	7 (46.7)	0.70 (0.36, 1.37)	0.299
Model adopts questionnaire or self-report data			
No	7 (58.3)	Reference	
Yes	8 (53.3)	0.91 (0.47, 1.79)	0.794
Number of predictors			
1 to 10	5 (41.7)	Reference	
11 to 20	3 (50.0)	1.20 (0.42, 3.41)	0.738
>20	7 (77.8)	1.87 (0.88, 3.97)	0.105
Time-point for prediction			
During inpatient admission	6 (46.2)	Reference	
During ED attendance	4 (57.1)	1.24 (0.52, 2.95)	0.630
Any time	5 (71.4)	1.55 (0.73, 3.28)	0.255
<b>1 year time-horizon (N=46)</b>			
Model uses morbidities as a predictor			
No	8 (61.5)	Reference	
Yes	24 (72.7)	1.18 (0.73, 1.91)	0.493
Model uses age as a predictor			
No	5 (41.7)	Reference	
Yes	27 (79.4)	1.91 (0.96, 3.80)	0.067
Model uses sex as a predictor			
No	11 (50.0)	Reference	
Yes	21 (87.5)	1.75 (1.12, 2.73)	<b>0.014</b>
Model adopts clinical assessment data			
No	27 (77.1)	Reference	
Yes	5 (45.5)	0.59 (0.30, 1.15)	0.123
Model adopts questionnaire or self-report data			
No	28 (71.8)	Reference	
Yes	4 (57.1)	0.80 (0.41, 1.56)	0.505
Number of predictors			
1 to 10	10 (52.6)	Reference	
11 to 20	7 (70.0)	1.33 (0.74, 2.40)	0.342
>20	15 (88.2)	1.68 (1.06, 2.66)	<b>0.028</b>
Time-point for prediction			
During inpatient admission	7 (53.8)	Reference	
During ED attendance	2 (33.3)	0.62 (0.18, 2.14)	0.448
Any time	21 (84.0)	1.56 (0.92, 2.65)	0.101
During CGA	2 (100)	1.86 (1.12, 3.07)	<b>0.016</b>

CGA: Comprehensive geriatric assessment; CI: Confidence interval; ED: Emergency department; RR: Relative risk.

\*Acceptable discrimination indicates an area under the receiver operating characteristic curve or Harrell's *c*-statistic  $\geq 0.70$