SUPPLEMENT

Supplemental Methods S1. Imputation of missing data

Because the NHATS data had a relatively small amounts of missing data, and because we judged that it was reasonably modeled on all of the measured variables, i.e., equivalent to an assumption of missing at random (MAR), we used multiple imputation as implemented in the SAS procedure MI.[1] We used fully conditional specification (FCS) for imputation of all variables with missing values.[2] For imputation of continuous variables, we used FCS regression and for dichotomous and ordinal variables we used FCS logistic and FCS discrim. There were no nominal variables with missing values. The multiple imputation procedure was performed on the complete set of person-years in the 2011 NHATS longitudinal data (N=39757). As discussed in the main text and as presented in Figure 1, pre- and post-ICU interviews that were missing all seven of the separate indicators of function, while included in the imputation process to help inform other variables, were not considered eligible for the final analytical sample. Predictor variables that were multiply imputed due to missingness are listed in Supplemental Table S2, and functional items in the outcome measure that were multiply imputed due to partial missingness in the functional outcome measure are listed in Supplemental Table S1.

Supplemental Table S1. Items that were multiply imputed due to partial missingness in the 7-item functional outcome measure*

Number of ADL items missing (range, 0-7)	Development cohort N=456	Validation cohort N=227
1	3	0
2	3	1
3	13	2
4	1	0
5	0	0
6	0	0

*Observations missing all 7 items in the functional measure were not imputed; see methods

Supplemental Table S2. Predictors that were multiply imputed due to missingness

Predictor	Development cohort N=456	Validation cohort N=227
Race	7	0
Education	7	0
Depression	8	5
Hearing	4	1
Vision	8	3
BMI	14	6
Social isolation	1	0
Pre-ICU disability	1 item missing in 1 observation	0

Supplemental Methods S2. Procedure for model selection

Because we started with 17 candidate variables, we used Bayesian Model Averaging (BMA) to directly select the variables for the final multivariable model. We used the BMA package by Raftery et al.[3] The 8 variables that exhibited an average posterior effect probability > 50% were retained. The BMA selection process was separately verified using 100 bootstrapped datasets that showed that only those 8 variables were chosen more than 50% of the time based on minimization of the Bayesian information criterion (BIC). Immediately below we provide R code illustrating how BMA was applied to the NHATS data:

```
#########################       Start of R code   #########################

## install package needed and then open for use
install.packages("BMA")
library(BMA)

## import data
PFDdata<-read.table(
  "PFDpenMLfull.csv",
  header=TRUE, sep=",", na.strings="NA", dec=".", strip.white=TRUE)

## get variable names
names(PFDdata)

attach(PFDdata) # this puts the contents of "PFDdata" on the search path

## save outcome as separate variable for use in BMA
PFD = PFDdata$PFD6m == 1
PFD
describe(PFD) ### single column with 264 values

## run the BMA on the data for a logistic model with typical and default parameter values
PFD.bic.glm <- bic.glm(PFDdata, PFD, glm.family="binomial", strict=FALSE, OR=20,
maxCol=70, Prior.param = c(rep(0.5, ncol(x))), nbest = 5)

## print output of BMA
PFD.bic.glm

## print summary of BMA
summary(PFD.bic.glm)
```

###################### End of R code ########################

This selection process was verified by selecting the model from 100 bootstrapped samples of the development data using backward-selection based on minimization of the Bayesian information criterion (BIC)[4] and calculating how many times each of the 17 candidate variables were selected. This process was facilitated by use of the SAS proc HPgenselect, which allows automated selection of generalized linear models with the Akaike Information Criterion (AIC), the BIC, and using p-values for entry and retention.

Immediately below we provide SAS code illustrating model selection a macro that automatically selects the model from bootstrapping of the development data and then saving indicators of each variable for easy summarization:

```
**The following statement create 100 samples of 683 caseID with replacement ;

proc surveyselect data = caseIDList method=urs outhits sampsize=683 rep = 100
seed = 020721 out = caseIDsample100; samplingunit caseID ; run;


data        caseIDsample100;
set         caseIDsample100 ;
BSobs = _n_;
run;    *** 68300 obs and 4 vars ;


proc datasets noprint;
delete allCaseIDbs;
run;
quit;

%macro BuildCaseBS;
%do i = 1 %to 68300;
data group; set caseIDsample100 ;
if BSobs =&i;
run;

data CaseBS_&i;
   merge group(in=row) Nhats2011imp1(in=all);
   by caseID ;
```

```
   if all and row ;
run;


proc append base=allCaseIDbs data=CaseBS_&i force;
run;

%end;
%mend BuildCaseBS;


%BuildCaseBS;


*** next run selection process on all BS samples;

proc datasets noprint;
delete allBootMVvars;
run;
quit;


%macro BootMVselNHATS;
%do i = 1 %to 100;
data group; set tem.allCaseIDbsFeb0721;
if replicate=&i;
run;

proc hpgenselect data= group ;

model decline =
age
bmi
dementia
depressed
female
Friedscore
hearingimp
hospLOS
icuLOS
LTHS
mdcaid_pastYr
mv967x
ncc
npriorhosp6m
preADLscore
socialisolation
visionimp
  / dist=binary link=logit ;
selection method=backward (choose=sbc);
ods output selectedeffects=Chosen_&i ;
run;


/* add code here selecting out specific variable names to set up variable-
specific indicators to facilitate counting of how many times each selected */
```

```
data BootVars_&i;
set         Chosen_&i ;
run;

data  BootVars_&i;
set         BootVars_&i;

I_age = (find(Effects,'age')>0);
I_bmi = (find(Effects,'bmi')>0);
I_dem = (find(Effects,'dementia')>0);
I_dep = (find(Effects,'depressed')>0);
I_fem = (find(Effects,'female')>0);
I_Fried = (find(Effects,'FriedScore')>0);
I_hearImp = (find(Effects,'hearingimp')>0);
I_hLOS = (find(Effects,'hospLOS')>0);
I_icuLOS = (find(Effects,'icuLOS')>0);
I_LTHS = (find(Effects,'LTHS')>0);
I_medicaid = (find(Effects,'mdcaid_pastYr')>0);
I_mechvent = (find(Effects,'mv967x')>0);
I_numCC = (find(Effects,'ncc')>0);
I_numPrHosp = (find(Effects,'npriorhosp6m')>0);
I_preADL = (find(Effects,'PreADLScore')>0);
I_socIso = (find(Effects,'socialisolation')>0);
I_visImp = (find(Effects,'visionimp')>0);

run;

proc append base=allBootMVvars data=BootVars_&i force;
run;

;

%end;
%mend BootMVselNHATS;


%BootMVselNHATS;


data tem.allBootMVvarsFeb0821;
set    allBootMVvars;
run;


proc means data=tem.allBootMVvarsFeb0821 mean ;
var
I_age
I_bmi
I_dem
I_dep
I_fem
I_Fried
I_hearImp
I_hLOS
I_icuLOS
I_LTHS
I_medicaid
```

```
I_mechvent
I_numCC
I_numPrHosp
I_preADL
I_socIso
I_visImp
;
run;
```

Five of the eight predictors (age, FriedScore, length of stay, number of prior hospitalizations, and depressive symptoms) were either count or continuous. We describe how linearity for each of these five variables was checked using FriedScore as an example. We calculated the natural log of the average rate of the outcome (logPFD) within each level of the Fried count, plotted the logPFD against the ordinal levels, and visually checked for linearity and symmetric distribution around that linear plot. Length of stay and number of prior hospitalizations required truncation of their maxima to ensure linearity and stability of estimation.

Supplemental Table S3. Operational definitions of the potential predictors

Potential predictor	NHATS study operational details
Age	in years
Female sex	1 = female; 0 = male
Race/ethnicity	1 = nonwhite race or Hispanic ethnicity; 0 = non-Hispanic white
Medicaid	Dual status indicator from Medicare administrative data, at any time during the previous 12 months of the ICU admission.
Less than a high school education	1 = <12 years education 0 = ≥12 years education
Body-mass index	Measured height and weight, kg/m^2
Fried frailty	Fried frailty count (0-5)[5]
Chronic conditions	Number of chronic conditions (range 0-9)[a]
Probable dementia	An existing diagnosis of dementia or Alzheimer's disease, a score ≥2 on the AD8 Dementia Screening Interview, or a score ≤1.5 standard deviations below the mean in ≥2 of 4 cognitive domains (memory, orientation, executive function, and retrieval of information)[6]
Pre-ICU disability	Dependence in 7 functional activities (activities of daily living and mobility activity[b] operationalized as disabled (1) or nondisabled (0) at baseline
Social isolation	Score of ≥4 (range, 0-6) on a previously validated measure of social isolation[c,7]
Hearing impairment	Based on the response to 3 questions[d]
Vision impairment	Difficulty reading newspaper print, recognizing a person across the street, or seeing a television across the room (with the use of corrective lenses, if applicable)
Depressive symptoms	Based on responses to the PHQ-2[8] (range, 0-6)

Hospitalizations within the 6 months prior to admission	0, 1, or ≥2 hospitalizations in the 6 months prior to the ICU admission (drawn from claims data)
Hospital length of stay	In days, truncated at an upper limit of 16 (range, 0 through ≥16)
Mechanical ventilation	ICD-9 code 96.7x ICD-10 codes 5A1935Z, 5A1945Z, and 5A1955Z

Abbreviations: PHQ-2, Patient Health Questionnaire-2; ICD-9, International Classification of Diseases.

[a] Of a possible 9: hypertension, myocardial infarction, heart disease, stroke, diabetes mellitus, arthritis, osteoporosis, chronic lung disease, and cancer (other than minor skin cancers)

[b] The 7 functional activities were dressing, getting cleaned up, toileting, eating, getting out of bed, getting around inside, and going outside. Observations in which the participant was maximally disabled (in 7 of 7 activities) were excluded to avoid floor effects, so the range of the pre-ICU variable was 0-6.

[c] 1 point is assigned for each of the following 6 items: not married or living with a partner, do not talk to family about important things, do not talk to friends about important things, do not visit family or friends in person, do not attend religious services in the past month, do not attend clubs or other recreational activities

[d] Hearing impairment = yes was operationalized as:

In the last month have you used a hearing device or a hearing aid = 1 (Yes) or 7 (Deaf)

OR Are you able to hear well enough to have a conversation with the TV or Radio playing = 2 (NO)

OR Are you able to hear well enough to have a conversation in a quiet room = 2 (NO)

Supplemental Table S4: Full regression results for the model

Variable	Coefficient	Odds Ratio	95% CI
Intercept	-5.20	N/A	N/A
Age	0.05	1.05	1.02, 1.08
Any Pre-ICU disability	-0.36	0.70	0.51, 0.92
Probable dementia	0.46	1.58	0.76, 3.28
Fried Frailty score	0.15	1.16	0.92, 1.46
Number of prior hospitalizations	0.71	2.03	1.49, 2.77
Vision impairment	1.36	3.90	1.76, 8.62
Depressive symptoms	0.10	1.10	0.90, 1.34
Hospital length of stay	0.08	1.09	1.04, 1.14

The final multivariable consisted of the following predictors: age in years, pre-admission disability (0/1), probable dementia (0/1), Fried frailty phenotype (range, 0-5), number of hospitalizations in the 6 months prior to admission truncated at an upper limit of 2 (0, 1, ≥2), vision impairment (0/1), depressive symptoms (0-6), and hospital length of stay in days truncated at an upper limit of 16 (range, 0 through ≥16).

Supplemental Methods S3. Calculation of probabilities from model coefficients

One can take the Beta coefficients for the global intercept and the variables from eTable 2 to calculate the predicted probability for a given individual. Calculate the linear predictor by taking the sum of the model intercept and the products of the coefficients of the predictors by their respective predictor values as follows:

Linear Predictor =

(-5.21 +
0.05*age in years +
0.46*probable dementia +
0.15*FriedScore +
0.08*length of stay(upper truncation at 16) +
0.71*number of prior hospitalizations(upper truncation at 2) +
-0.36*AnyPreICUdisability +
0.10*depressive symptoms +
1.36*vision impairment)

The predicted probability is calculated as:

Phat = 1/ (1 + exponent^(-Linear Predictor))

Supplemental References S1

[1] SAS Institute Inc. 2013. Base SAS ® 9.4 Procedures Guide. Cary, N.C.: SAS Institute, Inc.

[2] White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med. 2011;30: 377-399.

[3] Raftery AE, Hoeting JA, Volinsky CT, Painter I, Yeung KY. R package “BMA,” version 3.18.14. 2019.

[4] Schwarz G. Estimating the Dimension of a Model. Ann Stat. 1978;6: 461-464.

[5] Fried LP, Tangen CM, Walston J*, et al.* Frailty in older adults: Evidence for a phenotype. J Gerontol A Biol Sci. 2001;56: M146-M156.

[6] Kasper JD, Freedman VA, Spillman BC. *Classification of Persons by Dementia Status in the National Health and Aging Trends Study. Technical Paper #5.* Baltimore: Johns Hopkins University School of Public Health, 2013.

[7] Pohl JS, Cochrane BB, Schepp KG, Woods NF. Measuring Social Isolation in the National Health and Aging Trends Study. Res Gerontol Nurs. 2017;10: 277-287.

[8] Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the patient health questionnaire (PHQ): A diagnostic meta-analysis. J Gen Intern Med. 2007;22: 1596-1602.

Supplemental Checklist S1. TRIPOD Checklist

Section/Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	4
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	6-7
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	7
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	7-8
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	7-10
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	7-10
	5b	Describe eligibility criteria for participants.	7, 9-10
	5c	Give details of treatments received, if relevant.	N/A
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	8-9
	6b	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	9 eTable 1
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	8	Explain how the study size was arrived at.	9-10
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	10, S3
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	10-11, S1
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for validation.	10-11, S1, eT2
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	11
Risk groups	11	Provide details on how risk groups were created, if done.	11
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Fig 1, 9-10, 12
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	12, Table 1
Model development	14a	Specify the number of participants and outcome events in each analysis.	12
	14b	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	eT2, S2
	15b	Explain how to the use the prediction model.	S2, 12-15
Model performance	16	Report performance measures (with CIs) for the prediction model.	12, Figure 2
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	16
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence.	13-16
Implications	20	Discuss the potential clinical use of the model and implications for future research.	13-16
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Suppl.
Funding	22	Give the source of funding and the role of the funders for the present study.	2, 17

We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.