Development of a Novel Frailty Index to Predict Mortality in Patients with End-Stage Liver Disease

Jennifer C. Lai, MD, MBA[a], **Kenneth E. Covinsky, MD, MPH**[b], **Jennifer L. Dodge, MPH**[c], **W. John Boscardin, PhD**[b,d], **Dorry L. Segev, MD, PhD**[e], **John P. Roberts, MD**[c], and **Sandy Feng, MD, PhD**[c]

[a]Department of Medicine, Division of Gastroenterology and Hepatology, University of California-San Francisco, San Francisco, CA

[b]Department of Medicine, Division of Geriatrics, University of California-San Francisco, San Francisco, CA

[c]Department of Surgery, Division of Transplant Surgery, University of California-San Francisco, San Francisco, CA

[d]Department of Epidemiology and Biostatistics, University of California-San Francisco, San Francisco, CA

[e]Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD

Abstract

Background—Cirrhosis is characterized by muscle wasting, malnutrition, and functional decline that confer excess mortality not well quantified by the MELDNa score. We aimed to develop a frailty index to capture these extrahepatic complications of cirrhosis *and* enhance mortality prediction in cirrhotics.

Methods—Consecutive outpatients listed for LT at a single transplant center *without MELD exceptions* were assessed with candidate frailty measures. Best subset selection analyses with Cox regression identified subsets of frailty measures that predicted waitlist mortality (=death or delisting due to sickness). We selected the Frailty Index by balancing statistical accuracy with clinical utility. The net reclassification index (NRI) evaluated the %patients correctly reclassified by adding the Frailty Index to MELDNa.

Corresponding Author: Jennifer C. Lai, 513 Parnassus Avenue, UCSF Box 0538, San Francisco, CA 94143. Jennifer.lai@ucsf.edu; fax: 415-476-0659.

Results—Included were 536 cirrhotics with median MELDNa of 18. 107(20%) died/were delisted. The final Frailty Index consisted of: grip strength, chair stands, and balance. The ability of MELDNa and the Frailty Index to correctly rank patients according to their 3-mo waitlist mortality risk (i.e., C-statistic) was 0.80 and 0.76, respectively, but 0.82 for MELDNa + Frailty Index together. Compared with MELDNa alone, MELDNa + Frailty Index correctly re-classified 16% of deaths/delistings (p=0.005) and 3% of non-deaths/delistings (p=0.17) with a total NRI of 19% (p<0.001). Compared to those with robust Frailty Index scores (<20%ile), cirrhotics with poor Frailty Index Scores (>80%ile) were more impaired by gait speed, IADL difficulty, exhaustion, and low physical activity [p<0.001 for each].

Conclusions—Our Frailty Index for cirrhotics, comprised of 3 performance-based metrics, has construct validity for the concept of frailty and improves risk prediction of waitlist mortality over MELDNa alone.

INTRODUCTION

Cirrhosis is the terminal complication of a multitude of chronic liver conditions that lead to progressive liver failure and, ultimately, death. The most commonly used metric to determine prognosis for cirrhotics is the Model for End-Stage Liver Disease (MELDNa) score. Comprised of a logarithmic combination of serum total bilirubin, creatinine, international normalized ratio (INR) for prothrombin time, and sodium the MELDNa score can accurately predict 90-day mortality for most cirrhotics.[1] What these blood tests fail to capture, however, are the effects of muscle wasting, under-nutrition, and functional decline that are nearly universal findings in decompensated cirrhotics and contribute to excess mortality in this population.[2]

We have sought to objectively measure "physical frailty" – a term that we believe embodies these extra-hepatic manifestations of cirrhosis – in patients with end-stage liver disease and quantify its impact on health-related outcomes. In the field of geriatrics, frailty has been commonly defined as a distinct biological condition of increased vulnerability to health stressors[3] and operationalized using a number of instruments such as the Fried Frailty Index[3] and Short Physical Performance Battery.[4] The individual components of these instruments (e.g., hand grip strength, depression scale, short gait speed) have the strong advantage of being easy to administer at baseline and longitudinally at the bedside, even in patients with cirrhosis.[2,5] However, these measures were originally developed in cohorts of community-dwelling adults over age 65 years without known liver disease. As a result, their true discriminative ability for mortality in cirrhotics is unknown. Furthermore, the components of the frail phenotype may differ between patients with and without cirrhosis. For example, sarcopenia may play a greater role in physical frailty in a patient with end-stage liver disease compared to an older adult with normal hepatic synthetic function. On the other hand, decreased cardiopulmonary reserve may contribute more to the frail phenotype in the 90 year old adult without liver disease than in a frail 60 year old cirrhotic.

Therefore, in this study, we aimed to develop a continuous and quantitative index of physical frailty based on readily available instruments that predicts mortality for patients with cirrhosis.

METHODS

This study included patients enrolled in the Functional Assessment in Liver Transplantation (FrAILT) Study from March 2012 until February 2016. The FrAILT Study, initiated in July 2012, is an ongoing study of adults (18 years) with cirrhosis who are listed for liver transplantation at the University of California, San Francisco (UCSF) and are seen in the outpatient UCSF Transplant Hepatology clinic. To ensure an adequate number of events during follow-up, we prioritized consecutive recruitment of all patients with a laboratory MELD score 12. In September 2013, we relaxed our recruitment criteria to include *all* patients 60 years and older, regardless of laboratory MELD score, given the conceptual association between frailty and advancing age. Patients were excluded if they were listed with MELD exception points, as these patients have a trajectory to liver transplantation that is independent of hepatic decompensation. While national liver allocation changed in January 2016 to be based on MELDNa rather than MELD, we maintained study inclusion criteria as MELD for consistency. Also excluded were those with severe hepatic encephalopathy (n=8), as defined by the time to complete a Numbers Connection Test[6] of >120 seconds, as this may impair the patient's ability to provide informed consent and complete tests of physical function. Of those who met inclusion criteria, 97% enrolled in the FrAILT Study.[2,5] Four subjects who refused to complete all study procedures (i.e., assessments of frailty) were excluded from the analyses.

At enrollment, all patients underwent the tests of physical frailty that have been commonly utilized in the geriatric literature (Table 1). These measures included four performance-based tests (gait speed, grip strength, chair stands, and balance) and five self-reported tests (unintentional weight loss, exhaustion, physical activity, activities of daily living (ADL), and instrumental ADLs. All assessments were performed by one of two study personnel specifically trained at administering these study procedures in the same order and same manner for each study subject. On the same day as the clinic visit, the patient's hepatologist was asked to subjectively rate his or her patient's health using the following question:

> "We are interested in your general impression about your patient's overall health, as compared to other patients with underlying liver disease. How would you rate this patient's overall health today? Excellent (0), very good (1), good (2), fair (3), poor (4), or very poor (5)".

This rating was collected solely for the purposes of providing information regarding construct validity of the frailty measures. We have previously demonstrated that this subjective clinician assessment can identify liver transplant candidates at high risk for waitlist mortality.[7]

At the time of enrollment, demographic data were extracted from the clinic visit note from the same day as the physical frailty testing. Patients were classified as having hypertension or diabetes if listed in the past medical history or taking a medication to manage

hypertension or diabetes. Ascites was ascertained from the physical examination or mention of ascites in the management plan. Laboratory data within 3 months of the frailty assessment were collected from the electronic health record. Candidate prognostic indicators were recorded in blind with respect to the primary endpoint. All patients were followed prospectively until their terminal waitlist event (e.g., death/delisting, liver transplantation) or, for those who had not experienced a terminal waitlist event, until February 2016. "Delisting for being too sick for liver transplant" was decided by consensus among the liver transplant team members if there was concern that an individual would not achieve acceptable outcomes after liver transplantation due to medical co-morbidities or current medical acuity. This decision was made independently of the frailty assessments performed for the study, as the results from the study assessments were not made available to the clinical care team. Outcomes (e.g., death, delisting, transplant) were ascertained quarterly from UNet[SM], the official online database for the United Network for Organ Sharing (UNOS). Per UNOS requirements, outcomes must be recorded into UNet[SM] within 24 hours of the outcome and therefore, is a reliable source of information about the patients' current waitlist status.

Statistical analysis

The primary outcome was time from first study assessment to waitlist mortality, defined as death prior to liver transplantation or delisting for being too sick for transplant. For the analyses, patients who underwent living donor liver transplant were censored at the time of transplant as living donor surgery interrupts the natural trajectory of liver disease on the waitlist. Patients delisted for reasons other than being too sick (e.g., substance abuse, non-adherence) were censored at the time of waitlist removal.

Probability of waitlist mortality at 3, 6, and 12 months and 95% confidence intervals (95% CI) were estimated using the Kaplan-Meier method and calculated as 1-survival. Univariable Cox regression assessed the association between co-variables and waitlist mortality. The proportional hazards assumption was assessed by reviewing the survival function plots. Linearity of continuous variables was tested by adding quadratic terms to the model and Loess plots were used to visualize the functional form of the variables. Given the inherent sex differences in grip strength[8] and walk speed[9] norms, we converted grip strength and walk speed into sex-specific z-scores for use in the prediction modeling:

$$\text{Adjusted test for a woman} = \frac{(\text{individual's test value} - \text{average female test value})}{\text{standard deviation of female test value}}$$

$$\text{Adjusted test for a man} = \frac{(\text{individual's test value} - \text{average male test value})}{\text{standard deviation of male test value}}$$

The resulting z-score centers grip strength and walk speed around the gender-specific means and scales the variables to standard deviation units away from each gender mean. In other words, women and men with a mean score for their respective genders will have the same score of 0, while women and men with a grip strength value that is 1 standard deviation above their gender mean will have the same score of 1.

We then proceeded to develop a frailty index intended for use in adult patients with cirrhosis awaiting liver transplantation in the outpatient setting. We used a combination of stepwise selection, Akaike information criteria (AIC) reduction, and best subset selection, developed by Shtatland, *et al*,[10] to identify prediction models for waitlist mortality consisting of multiple permutations of the physical frailty instruments. First, stepwise selection using Cox regression (PROC PHREG, SELECTION=STEPWISE) identified the full stepwise sequence of potential models from the null model to the full model containing all explanatory variables, for a total of 10 candidate models. We calculated AIC for the 10 candidate models and identified the optimal number of explanatory variables (3 to 4 variables) to include in the frailty index, indicated by the number of covariables in the model with the lowest AIC. To assess a broad spectrum of candidate models that were both predictive and feasible to perform in the clinical setting, we applied best subset selection (PROC PHREG, SELECTION=SCORE) to identify the 5 best models within and just beyond the optimal covariable range. Best subset selection uses the branch-and-bound algorithm[11] to rank models (i.e, combinations of variables) in order of global score chi-square statistic. Given the optimal range of 3 to 4 explanatory variables, we assessed models with 2 covariables (most clinically feasible) to 6 covariables (potentially most predictive). In other words, we identified and evaluated the 5 best combinations of 2-variable models, 3-variable models, 4-variable models, 5-variable models, and 6-variable models, capturing a total of 25 candidate models. We selected the two "best" models by balancing model accuracy with simplicity, using a combination of: 1) AIC to measure the relative quality of the models, where a lower AIC indicates higher model quality, 2) concordance statistic [C-statistic with bootstrap 95% confidence intervals (CI)], which evaluates the absolute discrimination of the model, and 3) ease of use in clinical practice, with an underlying assumption that a model including more variables is less clinically useful than a more parsimonious model. The final frailty index was calculated as:

$$\text{Frailty index} = (\text{coefficient} * \text{test} \#1) + (\text{coefficient} * \text{test} \#2) + (\text{coefficient} * \text{test} \#3) + 6$$

where the coefficients are the parameter estimates directly estimated in the Cox regression model containing the components of the frailty index. To convert the index to a positive, rather than negative, scale, a constant was added to the equation to start the scale at a value of 1.

We selected this approach – best subset selection combined with stepwise selection and AIC reduction – because this methodology is asymptotically equivalent to the commonly-used internal validation methods, cross-validation and bootstrap. This was critical as an external validation cohort, the preferred validation method, is not yet available. The disadvantage is that best subset selection methodology is not available for competing risk regression.[12] However, in our cohort, the estimates of waitlist mortality, our primary outcome, were similar at 3-, 6-, and 12-months (supplemental table 2).

To evaluate the impact of the selected frailty index on improving prediction of MELDNa, we compared the proportion of patients whose risk of waitlist mortality at 3 months and 12 months, estimated from the Cox model, was correctly reclassified using MELDNa versus

MELDNa *plus* the frailty index using the net reclassification index (NRI). The NRI utilized a risk of waitlist mortality stratified by risks categories based on the cohort event rate and clinical relevance[13] at 3-months (<5%, 5–10%, and 10%) and 12-months (<10%, 10–15%, and 10%).

Statistical analyses were performed using SAS (v9.4, Cary, NC) and Stata (v14, College Station, TX). The Institutional Review Board at the University of California, San Francisco approved this study. This manuscript adheres to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement (Supplementary Table 3).[14]

RESULTS

Baseline characteristics, physical frailty scores, associations with waitlist mortality

Baseline characteristics of the 536 patients enrolled in the FrAILT Study are shown in Table 2. Median age of this cohort was 58 years and 41% were women. Over half (58%) self-identified as non-Hispanic White, 26% as Hispanic White, 3% Black, 5% Asian. Median height, weight, and body mass index were 170 cm, 82 kg, and 28 kg/m^2, respectively. The etiology of cirrhosis was alcoholic liver disease in 22%, chronic hepatitis C in 40%, non-alcoholic steatohepatitis in 14%, autoimmune/cholestatic liver diseases in 14%, and "other" in 10%. The prevalence of hypertension was 39% and diabetes was 28%.

Standard metrics of liver disease severity were assessed upon study enrollment (Table 2). Median MELD score was 15, MELDNa was 18, and albumin was 3.0 g/dL. Four percent of the cohort had end-stage renal disease requiring hemodialysis. Ascites was present in a total of 34% of the cohort, and hepatic encephalopathy, defined as a numbers connection test time 60 seconds, was present in 20%. Median Child-Pugh score was 8.

Table 3 presents a summary of the measures of physical frailty in the study cohort, along with available normative data for individuals in the general population of older adults without end-stage liver disease. With respect to performance-based characteristics: median walk speed for the FrAILT cohort was 1.2 meters per second, grip strength was 28 kg, number of chair stands per seconds was 0.4, and balance was 30 seconds. For the self-reported measures, 46% reported unintentional weight loss, 52% met criteria for exhaustion, and median kilocalories expended per week was 126. The median number of ADLs and IADLs that patients reported being able to perform independently was 6 and 8, respectively.

Median (interquartile range) follow-up time was 11 (4–22) months. By the end of follow-up, 107 (20%) died or were delisted for being too sick for transplant and 128 (24%) underwent deceased donor liver transplant. The probability of waitlist mortality at 3, 6, and 12 months was 5% (95% CI: 3–7%), 9% (95% CI: 6–12%), and 15% (95% CI: 12–19%), respectively. All baseline liver disease severity parameters were associated with wait-list mortality in univariable Cox regression: MELD/MELDNa score, serum total bilirubin, creatinine, INR, serum albumin, dialysis, ascites, hepatic encephalopathy, and Child Pugh score (Table 2). Eight of nine measures of physical frailty were significantly associated with wait-list

mortality in univariable Cox regression; the only variable that did not show an association was self-reporting of unintentional weight loss (Table 2).

Developing the frailty index for patients with cirrhosis

In Table 4, we display seven candidate models for the frailty index, along with their AIC values and C-statistics. As a point of reference, in our cohort, the AIC and C-statistic (including all follow-up time) for MELDNa was 1126 and 0.70, respectively. Model #2, which included grip strength, chair stands, and balance, best balanced accuracy (AIC 1112, C-statistic 0.72) with ease of use in clinical practice. Substituting walk speed for grip strength (Model #3), in order to avoid the use of the hand dynamometer, resulted in a substantially higher AIC (1117), indicating lower model quality. Adding variables to Model #2 (Models #4 through 7) did not meaningfully improve either the AIC (1112-1113) or C-statistic (0.72-0.73).

Our final frailty index was calculated as:

$$(-0.330 * \text{gender}-\text{adjusted grip strength})$$
$$+ (-2.529 * \text{number of chair stands per second})$$
$$+ (-0.040 * \text{balance time}) + 6$$

Construct validity of the frailty index for patients with cirrhosis

Figure 1 shows the distribution of frailty index scores for the cohort. Higher frailty index scores indicate a higher degree of frailty. Compared to those with a frailty index score in the bottom 20%ile ("robust"), patients with cirrhosis with a frailty index score in the top 80%ile ("frail") were more impaired by median walk speed (0.8 versus 1.5 m/sec; p<0.001), ADLs (5 versus 6; p<0.001), IADLs (5 versus 8; p<0.001), low physical activity (31 versus 0%; p<0.001), % meeting criteria for exhaustion (78 versus 32%; p<0.001). Furthermore, subjective clinician assessment scores (on a scale of 0 to 5 where 0 is robust) were significantly worse among patients who were frail versus not frail by the frailty index (3 versus 2; p<0.001).

Prognostic value of the frailty index for patients with cirrhosis

We then evaluated the ability of the frailty index to enhance risk prediction for waitlist mortality at 3-months over MELDNa alone. The ability of the MELDNa and the frailty index to correctly rank patients according to their 3-month risk of waitlist mortality (i.e., C-statistic) was 0.80 and 0.76, respectively. The combination of MELDNa *and* the frailty index together resulted in a C-statistic of 0.82 for 3-month waitlist mortality prediction. Compared with MELDNa alone, MELDNa + the frailty index correctly re-classified 16% of deaths/delistings (p=0.005) and 3% of non-deaths/non-delistings (p=0.17) for a total net re-classification index of 19% (p<0.001). The number of patients reclassified is shown in Table 5. Patients who died/were delisted who were correctly re-classified by MELDNa + the frailty index were older (62 versus 54 years) and substantially more likely to be obese (59 versus 41%), have hypertension (44 versus 31%), diabetes (59 versus 24%), end-stage renal disease (26 versus 6%), or hepatic encephalopathy (89 versus 46%) [Table 6].

Finally, given the potential clinical utility of the frailty index in predicting outcomes over a longer term on the waitlist, we evaluated the prognostic performance of the frailty index *at 12 months*. The C-statistics for prediction of 12-month risk of waitlist mortality was 0.73 for MELDNa and 0.73 for the frailty index alone, and the combination of MELDNa *and* the frailty index together resulted in a C-statistic of 0.77 at 12-months. Compared to MELDNa alone, the addition of the frailty index to MELDNa alone correctly reclassified 2% of deaths/delistings (p=0.72) and 17% of non-deaths/non-delistings (p<0.001) for a total net-reclassification index of 19% (p=0.002).

DISCUSSION

The MELD score has proven, time and time again, to predict mortality in patients with cirrhosis.[1,15] However, clinicians have long recognized that, at a given MELD score, the clinical manifestations of patients with end-stage liver disease are highly heterogeneous. For example, at a MELD score of 15, one patient may have well-controlled ascites and be working full-time, while another patient may have refractory ascites, sarcopenia, and barely be able to stand on his own. On paper, these two patients have the same (relatively low) predicted probability of death based on their MELD score – and therefore, the same (low) priority for liver transplantation – but any clinician will tell you that the latter patient clearly carries a higher risk of death. Yet at the current time, we lack the tools to objectively capture this risk.

Here, we present a novel liver-specific frailty index that objectively captures this frail phenotype in patients with end-stage liver disease in the outpatient setting. Consisting of three simple, performance-based tests of physical frailty – grip strength, chair stands, and balance testing – this index can feasibly be measured in the clinic setting at baseline and longitudinally. Furthermore, our analyses demonstrate that a frailty index can enhance risk prediction for waitlist mortality in cirrhotics awaiting liver transplantation over MELDNa alone. This was particularly true for those who were older, obese, or had hepatic encephalopathy or medical co-morbidities, risk factors that we would not expect to be adequately captured by the total bilirubin, creatinine, INR, or sodium. For this reason, we intentionally did not adjust for these patient-related related factors as we believe that they contribute directly to the frail phenotype rather than act independently of it. The strength of a frailty index is that it allows us to measure the *combined effect* of all of these factors on one's physiologic reserve. We have made this frailty index calculator available at: http:// liverfrailtyindex.ucsf.edu.

It is important to note the enduring tradition of frailty instruments in the literature. All of the tests that we administered in our study have repeatedly demonstrated prognostic utility in geriatrics populations (for whom they were originally established)[3,4,9,16] – and more recently in non-geriatrics[17,18], surgical,[19,20] and non-liver transplant populations.[21–24] This enduring history of frailty provides strong support for the broad applicability of these instruments to diverse populations. That an index consisting of a combination of these tests enhances mortality risk prediction in patients with end-stage liver disease awaiting liver transplantation builds upon this robust foundation of frailty in the literature.

How do we envision a frailty index to be incorporated into clinical practice? We do not intend for the frailty index to replace the MELD score. Given its strong prognostic value and sole reliance on common laboratory tests that are readily available, MELD plays an indispensable role in the national liver transplant allocation system. Rather, the frailty index should serve to *complement* the MELD in our clinical decision-making, facilitate discussions with patients. As an example of the impact that the frailty index might have on clinical decision-making and management, we offer predicted probabilities of death for four hypothetical outpatients classified by MELDNa scores and frailty status based on our model (Figure 2). Cirrhotic patients who display the frail phenotype experience a risk of waitlist mortality that exceeds that predicted by MELDNa alone. These frail patients should aggressively seek means to shorten their wait-time (e.g., living donor liver transplant, listing at centers with lower transplant MELDNa scores, or acceptance of higher risk donor livers). For those who are not eligible for liver transplant, frailty justifies initiation of palliative/ hospice care and provides the catalyst for developing end-stage liver disease management programs targeting the most vulnerable. For all cirrhotics, a firm "diagnosis" of frailty – which is potentially reversible through exercise and adequate nutrition – represents an opportunity for patients to gain control of their disease. There is nothing they can do about their MELDNa score, but the prospect of changing their frailty trajectory can offer them hope.

We acknowledge the following limitations to this study. We did not internally validate these results using a split-sample cohort, but rather utilized a combination of stepwise selection, AIC reduction, and best subset selection on the entire cohort to derive the frailty index. This was a deliberate choice, as studies have demonstrated that this combination of selection methods is equivalent to cross-validation and bootstrapping, and is preferred to internal validation.[10,25] *External* validation, however, remains the gold standard to assess the true validity of this index when applied to other patient samples, and is our next step. While reproducibility of these measures has not been evaluated the cirrhotic population, test-retest reliability is high in non-cirrhotic patients: 0.85 for grip strength,[26] 0.73–0.78 for chair stands,[27] and 0.55–0.75 for balance.[27] Nor did we evaluate inter- and intra-observer variability of the subjective hepatologist's rating of the patient's health, which would be important to further support the construct validity of this index. In addition, this cohort did not include patients in the inpatient setting, limiting the utility of this frailty index to outpatients. We believe that the construct of frailty can play its greatest role for candidates in the outpatient setting, allowing for sufficient time to engage in pre-habilitation, to seek faster paths to transplant (e.g., living donor liver transplant, transplant at a center with a lower transplant MELDNa score), or initiate hospice care services.

Also potentially limiting the generalizability of this index is the fact that we did not enroll patients <60 years if their MELD scores were <12. However, the intended use of this index is for patients with cirrhosis awaiting liver transplantation, as physical frailty likely disproportionately impacts this population due to the cumulative effects of *decompensated* cirrhosis. While this may introduce spectrum bias with possible overestimation of the predictive accuracy of the frailty index, we believe that the index was developed on the population for whom it would have the greatest clinical relevance. We excluded patients with severe hepatic encephalopathy due to concerns about their ability to follow frailty test

directions, which also could have introduced further spectrum bias in our results. However, this exclusion criterion applied to only eight patients (reflecting the outpatient nature of our cohort), so we do not believe that exclusion of these patients significantly impacts our findings. On the other hand, we included 21 patients with refractory ascites who also may have had greater difficulty with performing the frailty tests, but in this case, we felt that their lower scores would truly reflect their insufficient muscle mass and poor nutritional status, rather than reflect an inability to follow the frailty test instructions. Lastly, we did not evaluate inter-clinician variability of the decision to delist a patient for being "too sick for transplant", an outcome not defined by strict criteria. However, as described in a multi-center study evaluating center-level transplant decision-making, these decisions are reasonably consistent and based on consensus that for medical reasons they would have poor outcomes despite receiving a transplant.[28]

The strength of this study is that it includes a large number of patients who experienced a high number of events over a relatively short period of time, but the data are only from one center. However, this single-center index is the critical first-step to synchronize our nation's efforts at capturing the concept of frailty in clinical practice. Several centers have already reported on their use of frailty, as measured by a variety of frailty instruments originally developed in older adults without cirrhosis such as the six-minute walk test,[29] the Braden scale,[30] and the Fried Frailty Index.[31] Beyond the *published* literature, we know of other centers who have incorporated the timed-up-and-go test or hand grip strength into their evaluations of liver transplant candidates (personal communication). Our frailty index, developed specifically for patients with end-stage liver disease, allows us to converge on a relatively parsimonious number of individual tests of physical frailty. Only then will we be positioned to conduct the multi-center studies needed to better understand the impact of frailty on our patients – not just before transplant – but perhaps even more importantly, after.

In conclusion, we have developed a simple, easy-to-administer frailty index to predict mortality in patients with end-stage liver disease. Our index lays the foundation for future studies to investigate the relationship between frailty and outcomes after liver transplant, longitudinal changes in frailty over time, as well as develop interventions designed to reverse frailty and improve both survival and quality of life in patients with cirrhosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

List of abbreviations

AIC   Aikaike Information Criterion

| **C-statistic**	concordance-statistic
HCC	hepatocellular carcinoma
HCV	chronic hepatitis C
IQR	interquartile range
MELDNa	Model for End-Stage Liver Disease Sodium

References

1. Kamath P. A model to predict survival in patients with end-stage liver disease. Hepatology. 2001; 33(2):464–470. DOI: 10.1053/jhep.2001.22172 [PubMed: 11172350]

2. Lai JC, Feng S, Terrault NA, Lizaola B, Hayssen H, Covinsky K. Frailty Predicts Waitlist Mortality in Liver Transplant Candidates. American journal of transplantation. 2014; 14(8):1870–1879. DOI: 10.1111/ajt.12762 [PubMed: 24935609]

3. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001; 56(3):146–156.

4. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med. 1995; 332(9):556–561. DOI: 10.1056/NEJM199503023320902 [PubMed: 7838189]

5. Lai JC, Dodge JL, Sen S, Covinsky K, Feng S. Functional Decline in Patients with Cirrhosis Awaiting Liver Transplantation: Results from the Functional Assessment in Liver Transplantation (FrAILT) Study. Hepatology. 2016; 63:574–580. DOI: 10.1002/hep.28316 [PubMed: 26517301]

6. Weissenborn K, Rückert N, Hecker H, Manns MP. The number connection tests A and B: interindividual variability and use for the assessment of early hepatic encephalopathy. J Hepatol. 1998; 28(4):646–653. [PubMed: 9566834]

7. Lai JC, Covinsky KE, Hayssen H, et al. Clinician assessments of health status predict mortality in patients with end-stage liver disease awaiting liver transplantation. Liver Int. 2015; 35(9):2167–2173. DOI: 10.1111/liv.12792 [PubMed: 25644788]

8. Mathiowetz V, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: normative data for adults. Archives of Physical Medicine and Rehabilitation. 1985; 66(2):69–74. [PubMed: 3970660]

9. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA: The Journal of the American Medical Association. 2011; 305(1):50–58. DOI: 10.1001/jama.2010.1923 [PubMed: 21205966]

10. Shtatland, ES., Ken, K., Cain, EM. Statistics and Data Analysis. Vol. 30. SUGI; 2005. Model building in PROC PHREG with automatic variable selection and information criteria; p. 206-230.

11. Furnival GM, Wilson RW Jr. Regressions by Leaps and Bounds. Technometrics. 1974; 16(4):499–511.

12. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999; 94(446):496–509.

13. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Statist Med. 2010; 30(1):11–21. DOI: 10.1002/sim.4085

14. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med. 2015; 162(1):55–11. DOI: 10.7326/M14-0697 [PubMed: 25560714]

15. Wiesner R, Edwards E, Freeman R, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. Gastroenterology. 2003; 124(1):91–96. DOI: 10.1053/gast.2003.50016 [PubMed: 12512033]

16. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. Journal of Gerontology. 1994; 49(2):M85–M94. [PubMed: 8126356]

17. Leong DP, Teo KK, Rangarajan S, et al. Prognostic value of grip strength: findings from the Prospective Urban Rural Epidemiology (PURE) study. The Lancet. 2015; 386(9990):266–273. DOI: 10.1016/S0140-6736(14)62000-6

18. Pavasini R, Guralnik J, Brown JC, et al. Short Physical Performance Battery and all-cause mortality: systematic review and meta-analysis. BMC Medicine. 2016; 14(1):1–9. DOI: 10.1186/s12916-016-0763-7 [PubMed: 26728489]

19. Makary MA, Segev DL, Pronovost PJ, et al. Frailty as a predictor of surgical outcomes in older patients. J Am Coll Surg. 2010; 210(6):901–908. DOI: 10.1016/j.jamcollsurg.2010.01.028 [PubMed: 20510798]

20. Lin H-S, Watts JN, Peel NM, Hubbard RE. Frailty and post-operative outcomes in older surgical patients: a systematic review. BMC Geriatrics. 2016; 16(1):1–12. DOI: 10.1186/s12877-016-0329-8 [PubMed: 26729190]

21. McAdams-DeMarco MA, Law A, King E, et al. Frailty and Mortality in Kidney Transplant Recipients. American journal of transplantation. 2014; 15(1):149–154. DOI: 10.1111/ajt.12992 [PubMed: 25359393]

22. Singer JP, Katz PP, Dean MY, et al. Frailty Is Common in Lung Transplant Candidates and Associated with Poorer Health-Related Quality of Life. Journal of Heart and Lung Transplantation. 2013; 32(Supplement):S43.doi: 10.1016/j.healun.2013.01.912

23. Jha SR, Hannu MK, Chang S, et al. The Prevalence and Prognostic Significance of Frailty in Patients With Advanced Heart Failure Referred for Heart Transplantation. Transplantation. 2016; 100(2):429–436. DOI: 10.1097/TP.0000000000000991 [PubMed: 26516676]

24. Dunn MA, Josbeno DA, Tevar AD, et al. Frailty as tested by gait speed is an independent risk factor for cirrhosis complications that require hospitalization. Am J Gastroenterol.

25. Steyerberg EW, Harrell FE, Borsbroom GJ, Eijkemans MR, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology. 2001; (54):774–781. [PubMed: 11470385]

26. Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. J Hand Surg Am. 1984; 9(2):222–226. [PubMed: 6715829]

27. Freire AN, Guerra RO, Alvarado B, Guralnik JM, Zunzunegui MV. Validity and Reliability of the Short Physical Performance Battery in Two Diverse Older Adult Populations in Quebec and Brazil. Journal of Aging and Health. 2012; 24(5):863–878. DOI: 10.1177/0898264312438551 [PubMed: 22422762]

28. Volk ML, Biggins SW, Huang MA, Argo CK, Fontana RJ, Anspach RR. Decision making in liver transplant selection committees: a multicenter study. Ann Intern Med. 2011; 155(8):503–508. DOI: 10.7326/0003-4819-155-8-201110180-00006 [PubMed: 22007044]

29. Carey EJ, Steidley DE, Aqel BA, et al. Six-minute walk distance predicts mortality in liver transplant candidates. Liver Transplantation. 2010; 16(12):1373–1378. DOI: 10.1002/lt.22167 [PubMed: 21117246]

30. Tapper EB, Finkelstein D, Mittleman MA, Piatkowski G, Lai M. Standard assessments of frailty are validated predictors of mortality in hospitalized patients with cirrhosis. Hepatology. 2015; 62(2):584–590. DOI: 10.1002/hep.27830 [PubMed: 25846824]

31. Cron DC, Friedman JF, Winder GS, et al. Depression and Frailty in Patients with End-Stage Liver Disease Referred for Transplant Evaluation. American journal of transplantation. Nov.2015 n/a–n/a. doi: 10.1111/ajt.13639

32. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. JAMA: The Journal of the American Medical Association. 1963; 185:914–919. [PubMed: 14044222]

33. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969; 9(3):179–186. [PubMed: 5349366]

34. Oberg T, Karsznia A, Oberg K. Basic gait parameters: reference data for normal subjects, 10–79 years of age. J Rehabil Res Dev. 1993; 30(2):210–223. [PubMed: 8035350]
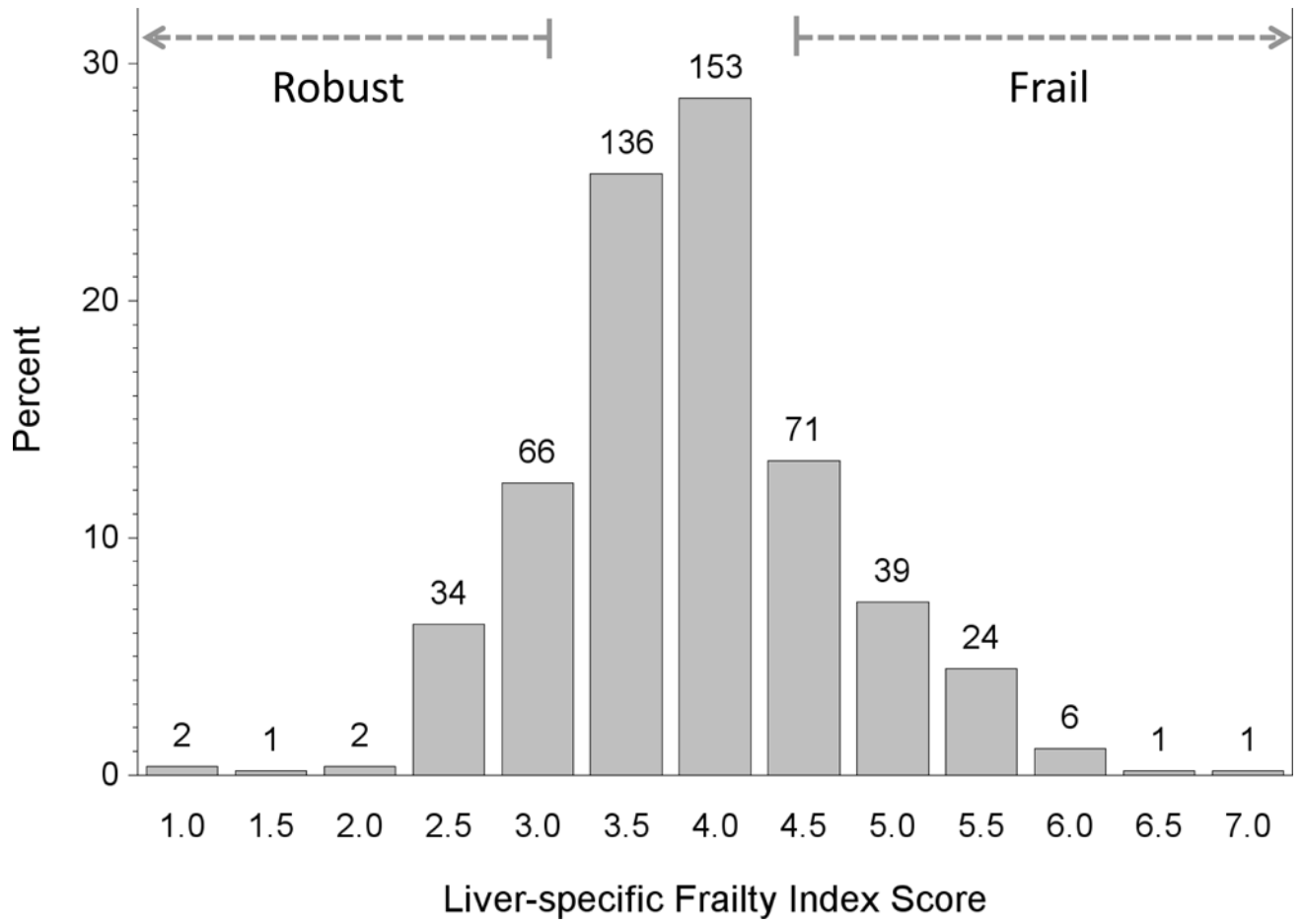
Figure 1.
Distribution of frailty index scores for 536 outpatient cirrhotics awaiting liver transplantation. Higher values indicate a higher degree of frailty.
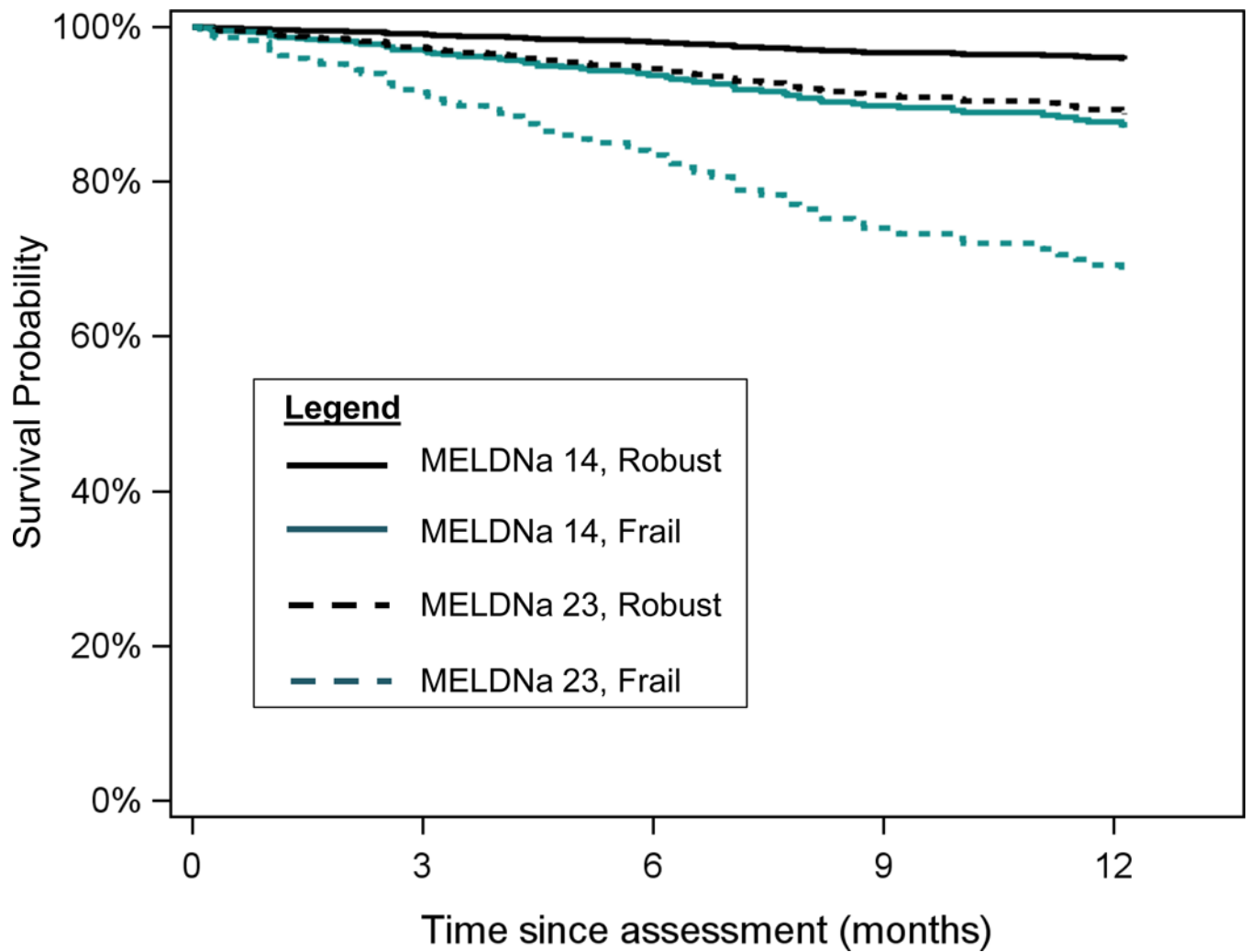
Figure 2.
Predicted probabilities of survival for four patients with cirrhosis on the liver transplant waitlist, classified by a combination of MELDNa and frailty index scores. MELDNa scores (14 and 23) and frailty index scores (3.2=robust and 4.5=frail) were selected because they represented the bottom 20%ile and top 80%ile values for the cohort.

Table 1

Tests administered in the FrAILT study.

Test	Self-reported or performance-based	Unit of measure	Description of the test
Gait speed[3]	Performance-based	Meters/second	Subjects were asked to walk eight feet as quickly as they could. Patients unable to walk at all (i.e., wheelchair bound) were assigned a gait speed of 0.01 meters/second.
Grip strength[3]	Performance-based	Kilogram	This was measured in each subject's dominant hand using a hand dynamometer (Jamar Hydraulic hand dynamometer). We averaged three trials.
Chair stands[4]	Performance-based	Number of chair stands per second	Number of chair stands completed in 30 seconds
Balance[4]	Performance-based	Seconds	Ability to balance in 3 positions (feet placed side-to-side, semi-tandem, and tandem) for 10 seconds each.
Unintentional weight loss[3]	Self-report	0 = no 1= yes	Self-reported unintentional weight loss of 10 pounds in the last year.
Exhaustion[3]	Self-report	0 = no 1 = yes	Identified by two self-reported questions from the Center for Epidmiologic Studies-Depression scale, as originally used in the Fried Frailty Index.[3]
Physical activity[3]	Self-report	Kilocalories per week	The number of kilocalories expended per week is assessed using the self-reported Minnesota Leisure Time Activity Scale.
Activities of Daily Living[32]	Self-report	Maximum 6 activities	Ability to perform basic activities of self-care without assistance (e.g., dressing, bathing, toileting)
Instrumental Activities of Daily Living[33]	Self-report	Maximum 8 activities	Ability to perform activities to live within society without assistance (e.g., shopping, preparing meals, administering medications, laundry)

Table 2

Baseline characteristics of the 536 patients enrolled in the FrAILT Study

Characteristic[*]		n=536	Univariable association with waitlist mortality[†]	
Demographics			Hazard ratio (95% CI)	p-value
Age, year		58 (50–63)	1.02 (1.00–1.04)	0.08
% female		41%	1.22 (0.83–1.80)	0.30
Race/Ethnicity	Non-Hispanic White	58%	Ref	Ref
	Black	3%	0.63 (0.15–2.59)	0.52
	Hispanic White	26%	1.09 (0.70–1.68)	0.71
	Asian	5%	1.92 (0.92–4.03)	0.08
	Other	8%	1.06 (0.49–2.33)	0.88
Height, cm		170 (163–178)	0.98 (0.97–1.00)	0.10
Weight, kg		82 (70–98)	1.00 (0.99–1.01)	0.36
BMI		28 (25–33)	1.00 (0.96–1.03)	0.83
Etiology of Liver Disease	Alcohol	22%	0.89 (0.53–1.49)	0.66
	Chronic hepatitis C	40%	Ref	Ref
	Non-alcoholic steatohepatitis	14%	1.00 (0.56–1.78)	0.99
	Cholestatic	14%	0.90 (0.48–1.70)	0.75
	Other	10%	1.06 (0.55–2.04)	0.86
Hypertension		39%	0.76 (0.51–1.13)	0.18
Diabetes		28%	1.23 (0.85–1.87)	0.26
Markers of liver disease severity				
MELD		15 (13–19)	1.15 (1.10–1.19)	<0.001
MELDNa		18 (15–23)	1.15 (1.11–1.20)	<0.001
Total bilirubin, mg/dL		2.6 (1.7–4.0)	1.05 (1.01–1.10)	0.02
Creatinine, mg/dL		1.0 (0.8–1.3)	1.03 (1.01–1.06)	0.002
INR		1.4 (1.3–1.6)	2.33 (1.71–3.18)	<0.001
Sodium, meQ/L		136 (133–139)	0.91 (0.87–0.95)	<0.001
Albumin, g/dL		3.0 (2.6–3.4)	0.56 (0.40–0.78)	0.001
Dialysis		4%	2.55 (1.33–4.89)	0.005
Ascites		34%	1.57 (1.06–2.32)	0.02
Hepatic encephalopathy		20%	2.04 (1.29–3.21)	0.002
Child Pugh Score		8 (7–9)	1.39 (1.24–1.55)	<0.001

[*] Median, interquartile range or n (%)

[†] Number of events = 107

[‡] Missing data: Sodium n=1, Child Pugh Score n=3, ascites n=3,

Table 3

Summary of physical frailty components.

| Measure[*] | Normative data in older adults, mean (SD) or range[†] | FrAILT Cohort n=536 | Univariable association with wait-list mortality[‡] | |
			Hazard ratio (95% CI)	p-value
Walk speed, m/sec	Men: 1.3 (1.2) m/s Women: 1.2 (1.7) m/s[34]	1.2 (1.0–1.4)	0.25 (0.16–0.39)	<0.001
Grip strength, kg	Men: 40 (8.3) kg Women: 24 (5.3) kg[8]	28 (21–37)	0.95 (0.93–0.97)	<0.001
Chair stands, number per second	<0.5 is impaired[16]	0.4 (0.3–0.5)	0.02 (0.01–0.07)	<0.001
Balance, seconds	<30 seconds is impaired[16]	30 (30–30)	0.91 (0.88–0.93)	<0.001
Unintentional weight loss	6%[3]	46%	1.43 (0.98–2.09)	0.07
Exhaustion	17%[3]	52%	1.77 (1.20–2.63)	0.004
Low physical activity	20%	13%	2.66 (1.69–4.20)	<0.001
Activities of Daily Living	Expected to be 6 (i.e., fully independent)	6 (5–6)	0.78 (0.69–0.89)	<0.001
Instrumental Activities of Daily Living	Expected to be 8 (i.e., fully independent)	8 (6–8)	0.81 (0.75–0.87)	<0.001

[*]
Median (interquartile range) or %

[†]Available normative data provided for the following age ranges: walk speed – 60–69 years; grip strength – 60–69 years; chair stands per 30 seconds – 70 years; balance – 70 years; unintentional weight loss – 65 years; exhaustion – 65 years

[‡]Number of events = 107

[§]Missing data: Physical activity n=24

Table 4

Derivation of a liver-specific frailty index: Candidate models incorporating various measures of physical frailty derived from best subsets selection with Cox regression. *

	Number of variables included						
	2	3	3	4	4	5	6
Model #	#1	#2	#3†	#4	#5	#6	#7
Grip strength, per kg, gender-adjusted	✓	✓		✓	✓	✓	✓
Chair stands, number per second	✓	✓	✓	✓	✓	✓	✓
Balance, per second		✓	✓	✓	✓	✓	✓
Walk speed, per m/sec, gender-adjusted			✓				
Exhaustion							
Physical activity				✓		✓	✓
Unintentional weight loss							
ADLs, per ADL					✓		✓
Independent IADL, per IADL						✓	✓
AIC	1113	1112	1117	1113	1113	1112	1113
C-statistic	0.72 (0.67–0.77)	0.72 (0.67–0.77)	0.71 (0.61–0.82)	0.72 (0.68–0.77)	0.72 (0.66–0.78)	0.72 (0.68–0.78)	0.73 (0.67–0.78)

*
As a point of reference, in our cohort, the AIC and C-statistic (including all follow-up time) for MELDNa was 1126 and 0.70, respectively. Higher AIC indicates relative lower model quality.

†
This model was not among the highest performing models by AIC using best subset selection but is included here to demonstrate performance characteristics of a performance-based model that does not require any specialized equipment for testing.

Table 5

Reclassification table comparing predicted 3-month risk of waitlist death estimated by MELDNa versus MELDNa and the frailty index, stratified by patients without (non-events) and with (events) waitlist death. Interior cells reflect the number of patients within the predicted risk categories. Pink cells indicate patients with an increased predicted risk of waitlist death estimated by MELDNa + frailty index versus MELDNa alone. Blue cells indicate patients with a decreased risk of death estimated by MELDNa + frailty index versus MELDNa alone. Yellow cells indicate no change in risk category when the frailty index is included in estimating risk.

Deaths/Delisting (n=107)				
MELDNa risk category	**MELDNa + Frailty Index risk category**			
	<5%	**5–<10%**	**10%+**	**Total**
<5%	50	7	4	61
5–<10%	8	5	16	29
10%+	0	2	15	17
Total	58	14	35	107

Non-deaths/Non-delistings (n=429)				
MELDNa risk category	**MELDNa + Frailty Index risk category**			
	<5%	**5–<10%**	**10%+**	**Total**
<5%	287	24	3	314
5–<10%	40	33	11	84
10%+	5	6	20	31
Total	332	63	34	429

Table 6

Select characteristics of patients with cirrhosis awaiting liver transplantation whose 3-month risk of waitlist mortality was *correctly* reclassified by MELDNa + Frailty Index compared to MELDNa alone.

Characteristic[*]		Correctly reclassified to death/delisting n=27/107 (25%)	Correctly reclassified to non-death/delisting n=51/429 (12%)
Frailty index[†]		4.1 (3.7–4.4)	2.4 (2.0–2.7)
MELDNa		23 (19–25)	23 (22–25)
Age, years		62 (53–68)	54 (46–60)
% female		44%	39%
BMI 30 kg/m²		59%	41%
Serum albumin, g/dL		2.9 (2.5–3.1)	3.0 (2.5–3.3)
Liver disease etiology	Alcohol	22%	29%
	HCV	41%	29%
	NASH	19%	16%
	Cholestatic	7%	14%
Hypertension		44%	31%
Diabetes		59%	24%
Dialysis		26%	6%
Ascites		44%	37%
Hepatic encephalopathy		93%	45%

[*] Median (interquartile range) or n (%)

[†] Higher frailty index scores indicate higher degree of frailty.