Published in final edited form as:
JAMA. 2012 January 11; 307(2): 182–192. doi:10.1001/jama.2011.1966.

Prognostic Indices for Older Adults: A Systematic Review

Lindsey C. Yourman, MD[1], **Sei J. Lee, MD, MAS**[1,2], **Mara A. Schonberg, MD, MPH**[3], **Eric W. Widera**[1,2], and **Alexander K. Smith, MD, MS, MPH**[1,2]

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

[2]San Francisco Veterans Affairs Medical Center, San Francisco, CA

[3]Beth Israel Deaconess Medical Center, Department of Medicine, Division of General Medicine and Primary Care, Boston, MA

Abstract

Context—To better target services to those who may benefit, many guidelines recommend incorporating life expectancy into clinical decisions.

Objectives—We conducted a systematic review to help physicians assess the quality and limitations of prognostic indices for mortality in older adults.

Data Sources—We searched MEDLINE, EMBASE, Cochrane, and Google Scholar through November 2011.

Study Selection—We included indices if they were validated and predicted absolute risk of mortality in patients whose average age was ≥ 60. We excluded indices that estimated ICU, disease-specific, or in-hospital mortality.

Data Extraction—For each prognostic index, we extracted data on clinical setting, potential for bias, generalizability, and accuracy.

Results—We reviewed 21,593 titles to identify 16 indices that predict risk of mortality from 6-months to 5 years for older adults in a variety of clinical settings: the community (six indices), the nursing home (two indices), and the hospital (eight indices). At least 1 measure of transportability was tested for all but 3 indices. By our measures, no study was free from potential bias. While 13 indices had c-statistics ≥ 0.70, none of the indices had c-statistics ≥ 0.90. Only two indices were independently validated by investigators who were not involved in the index's development.

Conclusion—We identified several indices for predicting overall mortality in different patient groups; future studies need to independently test their accuracy in heterogeneous populations and their ability to improve clinical outcomes before their widespread use can be recommended.

INTRODUCTION

Failure to consider prognosis in the context of clinical decision-making can lead to poor care. Hospice is underutilized for patients with non-malignant yet life-threatening diseases.[1] Healthy older patients with good prognosis have low rates of cancer screening.[2] Older adults with advanced dementia or metastatic cancer are screened for slow growing cancers that are unlikely to ever cause them symptoms, but may lead to distress from false positive results,

Corresponding Author: Alexander K. Smith, MD, MPH, MS, UCSF Division of Geriatrics, 4150 Clement St. (181G), San Francisco, Ca 94121, Voice: 415-221-4810 x 4684, Fax: 415-750-6641, aksmith@ucsf.edu.

Financial Disclosures: None.

invasive work-ups, and treatments.[3,4] In recognition of these phenomena, guidelines increasingly incorporate life expectancy as a central factor in weighing the benefits and the burdens of tests and treatments (see Table 1). Prognostic indices offer a potential role for moving beyond arbitrary age-based cutoffs in clinical decision-making for older adults.[2] However, little is known about the quality of prognostic indices for older adults, limiting their clinical use.

We performed a systematic review to describe the quality and limitations of validated non-disease specific prognostic indices that predict absolute risk of all-cause mortality in older adults. Recognizing that older adults are more likely to have more than one chronic illness than younger adults, we focused on non-disease specific indices.

METHODS

We used broad Medical Subject Heading Terms (e.g. mortality, prognosis, aged) to search MEDLINE, EMBASE, Cochrane, and Google scholar through November 2011 for English-language validated prognostic indices that predict absolute risk of all-cause mortality in patients whose average age was ≥60. Authors of included studies and experts in the field were contacted and asked for additional published and unpublished sources. We excluded indices that estimated ICU, in-hospital, or disease-specific mortality. Two investigators (L.C.Y. and A.K.S) independently applied these inclusion and exclusion criteria to select prognostic indices, and independently abstract their data. Disagreements were resolved by consensus or, if necessary, the involvement of a third investigator (S.J.L).

There are no accepted criteria to assess the quality of prognostic indices. Therefore, we adapted criteria from previous work published by experts in medicine and epidemiology.[5–12] We abstracted data on the quality of prognostic indices, including information on potential bias, generalizability, and accuracy (see Table 2). For discrimination, we considered c-statistics in the range of 0.5–0.59 to indicate poor, 0.6–0.69 to indicate moderate, 0.7–0.79 to indicate good, 0.8–0.89 to indicate very good, and >0.9 to indicate excellent discrimination.[13] For calibration, we considered ≥10% points difference between predicted and observed mortality evidence of poor calibration, and <10% points difference evidence that the model was well calibrated. To further assess the potential limitations of these indices in clinical practice, we tracked studies that predicted > 50% mortality, since 50% mortality represents the median residual lifespan. We report 95% CI on measures of discrimination and calibration where available.

RESULTS

One investigator title-screened 21,593 studies to identify 4,120 potentially relevant abstracts (eFigure 1). After excluding studies with participants average age <60 years old, studies that predicted only relative risk, or indices that predicted only disease-specific, in-hospital or ICU mortality, there were 341 studies published between January, 1987 and November, 2011. After review of the full text of these studies, 317 studies were excluded, leaving 24 studies (eFigure 1).[14–37] Three of these studies present updated versions of an index,[20,21,27] and 5 provide additional validation for an index,[28,29,34–36] resulting in a total of 16 unique indices.

All indices were developed using secondary analysis of existing datasets of participants from the United States (eleven indices)[14,15,18,20–23,30–33] and Western Europe (four indices)[16,17,24,25]. The most common final predictors of mortality included functional status and comorbidities (each only absent in <5 indices). Three indices tested only reproducibility and did not evaluate any form of transportability (split sample validation only: Di Bari 2010

and Drame 2008; bootstrapping only: Teno 2000)[16,17,33]. Only a single form of transportability was tested for 4 indices (Mazzaglia 2007, geographic; Carey 2004, geographic; Lee 2006 geographic; Levine 2007, historical)[15,22–24]. For 4 indices, the investigators who developed the index tested the transportability of their index in a separate validation study (Schonberg 2009 and 2011; Pilotto 2008 and Sancarlo 2010; Porock 2005 and Kruse 2010; Fischer 2006 and Youngwerth 2011)[18,25,26,29,30,34–36]. Two indices were additionally validated by an investigator not involved in the index's development (Flacker 2003 and Kruse 2010; Walter 2001 and Rozzini 2001)[20,28,32,34]. The highest risk group was at > 50% risk of mortality for 10 indices (Carey 2008, Lee 2006, Schonberg 2009 and 2011, Porock 2005 and 2010, Flacker 2003, di Bari 2010, Inouye 2003, Teno 2000, Walter 2001, and Drame 2008)[14,16,17,20–22,26,27,30,33,35].

None of the examined indices had a c-statistic ≥0.9; three indices had c-statistics between 0.8 and 0.89 suggesting very good discrimination (Lee 2006, Fischer 2006, Inouye 2003 development cohort)[18,21,22]; 10 indices had c-statistics between 0.7 and 0.79 suggesting good discrimination (Gagne 2011, Mazzaglia 2007, Carey 2004, Schonberg 2009, Porock 2005, Flacker 2003, Pilotto 2008, Teno 2000, Walter 2001, and Drame 2008)[15,17,20,24–26,30–33]; and 3 indices had c-statistics between 0.6 and 0.69 suggesting moderate discrimination (Carey 2008, di Bari 2010, and Levine 2007)[14,16,23]. Indices were generally well calibrated across risk groups (Table 3). Two indices reported a > 10% difference between predicted and observed mortality (Flacker 2003, Inouye 2003)[20,21].

Below we present a descriptive summary of each index by setting. Results of data abstraction regarding potential bias, generalizability, and accuracy shown in Tables 3 and 4.

Community-Dwelling Older Adults

Our review identified six indices for community-dwelling older adults. Indices estimated mortality risk from one (Gagne 2011)[31] to five years (Schonberg 2009)[30]. Schonberg's highest risk group at 9-year follow-up predicts 92% (95% CI 86–96) mortality.[35]

Gagne (2011)[31] developed a mortality risk score to predict 1-year mortality by combining conditions in the Romano[38] implementation of the Charlson Index[39] and the van Walraven[40] implementation of the Elixhauser system.[41] The sample was a secondary analysis of Medicare enrollees age>65 who in 2004 participated in a pharmacy assistance contract for low-income seniors who did not qualify for Medicaid prescription drug coverage in Pennsylvania (development cohort, n=120,679) and New Jersey (validation cohort, n=123,855). The model has good discrimination and was well calibrated (Table 3). Reclassification measures compare the model favorably against the Ramono/Charlson and van Walraven/Elixhauser indices.

The Mazzaglia (2007)[24] 15-month index is a 7-item questionnaire for primary care physicians that was developed in 2,470 primary care patients >65 residing in northwestern Florence, Italy, and validated in a sample of 2,926 similar patients residing in southwestern, Florence. The model was well calibrated and had good discrimination, but it predicted the narrowest range of mortality of any examined index (0–10% risk).

Carey (2004)[15] developed a 2-year index for community-dwelling elders from a sample of 4,516 adults age >70 from the eastern, western, and central US, who had been interviewed in the Asset and Health Dynamics Among the Oldest Old (AHEAD) study in 1993. Carey et al subsequently validated the index in 2,877 similar interviewees from the southern US. The index had good discrimination, was well calibrated across all three risk levels, but predicted only a narrow range of mortality (5–36% risk).

The Carey (2008)[14] Index for 3-year mortality was developed in functionally impaired nursing home-eligible community dwelling older adults >55 for the years 1988–1996 living in the western US (n=2,232) and enrolled in Program of All-Inclusive Care for the Elderly (PACE), a senior daycare program providing multidisciplinary services. Validation was conducted in PACE participants from the Eastern and Midwestern US (n=1,667). The index was well-calibrated but showed only moderate discrimination. Accuracy was similar for 1 year mortality.

Lee (2006)[22] developed a 4-year mortality index in community-dwelling adults age >50 from eastern, western, and central United States who were interviewed in the Health & Retirement Survey of 1998 (81% participation rate, n=11,701). To test geographic transportability, the index was validated in interviewees from the southern United States (n=8,009). The Lee index was well-calibrated and showed very good discrimination.

The Schonberg (2009)[30] index to predict 5-year mortality was developed from a nationally representative sample of adults >65 (n=16,077) who responded to the 1997–2000 National Health Interview Survey (NHIS) (74% participation rate), and was well calibrated and had good discrimination in a random sample of n=8,038 adults drawn from the same data source. Schonberg (2011)[35] then further validated the index in respondents to the 2001–2004 NHIS (n=22,057, 25% age>80, 57% female, 12% dependent in at least one AIDL, 18% diabetes, 15% cancer) and found no change in discrimination (c-statistic, 0.75). The Kaplan-Meier method demonstrated widening separation between risk groups out to 9 years.

Nursing Home Residents

Two indices were developed for the nursing home, both using the Minimum Data Set (MDS), a clinical and administrative dataset that is federally required of all US nursing homes. The MDS Mortality Rating Index by Porock (2005)[26] to estimate 6-month mortality in nursing home patients was developed using data from all Missouri long-term care residents in 1999. Study authors later created a simplified version of this model, using the same dataset (Porock 2010)[27]. The Flacker-R (Flacker 1998 and 2003)[19,20] long-stay index for 1-year mortality was developed and validated from the MDS using a split sample of nursing home residents >65 residing longer than one year in Medicare-certified nursing homes within New York (N=63,077). Both indices demonstrated very good discrimination and were well calibrated across a wide range of mortality risk levels, except the Flacker-R for the highest risk group (20% difference).

Kruse (2010)[34] prospectively validated indices by Porock and Flacker in a small prospective single nursing home study in 2007 (n=130, mean age 83, 61% female, 24% dementia, 23% congestive heart failure [CHF]). For the Porock index, discriminatory ability was lower in the validation study by Kruse (c-statistic, 0.59, 95% CI 0.46–0.72) than in Porock's original derivation study (c-statistic, 0.75) or using the simplified score (c-statistic, 0.76). For the Flacker-R index, discriminatory ability was the same in both Flacker's original derivation study (c-statistic, 0.71) and the external validation by Kruse (c-statistic, 0.72, 95% CI 0.62–0.81).

Hospitalized Older Adults

We identified eight indices that estimate mortality risk for hospitalized older adults. Seven indices estimate 1-year mortality. Five were intended for use in the emergency department or on hospital admission (Di Bari 2010, Fischer 2006, Inuye 2003, Pilotto 2008, and Teno 2000)[16,18,21,25,33], and three following hospital discharge (Levine 2007, Walter 2001, and Drame 2008)[17,23,32].

The Di Bari (2010)[16] "Silver Code," a 1-year index for emergency triage of elders age>75, was developed and validated using administrative records of patients admitted to the hospital via the emergency department from Florence, Italy in 2005 (n=10,913). They achieved a 91% linkage across 4 administrative datasets (demographics, hospitalizations, prescription medications, and mortality). Random split sample validation was conducted on half the cohort. The index was well calibrated and discriminatory ability was moderate.

Fischer (2006)[18] conducted a retrospective chart review to develop a 1-year index for hospitalized elders using four pre-specified predictors called the "CARING criteria," collected at admission. Fischer's sample included patients admitted to the medical service of a VA hospital in a 4 month period in 1999 (n=873). Participants admitted in the first two months of the study period were included in the development cohort, the remainder in the validation cohort. The model had very good discrimination and a reported error rate of 0.26 in the validation cohort. Youngwerth (2011)[36] later prospectively tested the external validity of the CARING criteria in a younger, gender-balanced sample from a University hospital in 2005 (n=427, average age 54, 50% female). No c-statistic is reported for the external validation.

The Burden of Illness Score for Elderly Persons by Inouye (2003)[21] updates previous indices developed by the same group[37,42] by adding functional and laboratory data to diagnoses from administrative data to estimate 1-year mortality. Participants were drawn from a prospective study of elders age ≥70 hospitalized at Yale New Haven Hospital 1989–1991 (n=525). The study was validated in a sample of 1,246 participants from 27 Connecticut hospitals age ≥65 with a principle discharge diagnosis of pneumonia 1995–1996. The investigators demonstrated improvement in the c-statistic with the addition of laboratory and functional/cognitive measures to administrative data (validation c-statistic, administrative alone 0.59; all measures, 0.77). The model was well calibrated at the extremes, but was less accurate in middle risk groups (see Table 4).

Pilotto (2008)[25] used information from the standardized Geriatrics Assessment, performed at admission, to develop a 1-year prognostic index for hospitalized elders age ≥65 in a sample of 838 consecutively admitted elders to the Geriatrics Unit of an Italian hospital in 2004, validating in 857 participants from 2005. They subsequently tested the model's accuracy at 1-year and 1-month in participants from the same hospital 2005–2007 (n=4,088),.[29] The model was well calibrated and demonstrated good discrimination in the larger validation study (c-statistic, 0.71, 95% CI 0.70–0.74), and performance was similar at 1-month (c-statistic, 0.76, 95% CI 0.73–0.79).

Teno (2000)[33] developed a nomogram to predict 1 and 2-year mortality based on medicine and ICU patients >80 enrolled in the Hospitalized Elder Longitudinal Project (HELP) from five different hospitals across the US 1993–1994 (n=1266). Teno tested the reproducibility of the index in 150 random samples from the original 1266 patients. The Teno nomogram is convenient in that it predicts multiple endpoints from a single score. The index includes the APACHE III scale, which requires arterial blood gas measurement.

Levine (2007)[23] developed a 1-year prognostic model for hospitalized elders following discharge using data from a cohort of patients admitted to hospitalist and non-hospitalist physicians at the University of Chicago Hospitals 7/1997–6/1999 (development cohort, n=2,739) and 7/1999–6/2001 (validation cohort, n=3,643). The index had moderate discriminatory ability and was well calibrated.

Walter (2001)[32] developed a 1-year index for elders following hospital discharge using secondary data from a study of elders age > 70 hospitalized between 1993–1997 at the University of Hospitals Cleveland (development cohort, n=1,495) and the Akron City

Hospital (validation cohort, n=1,427). The model demonstrated good discrimination and was well calibrated across risk groups. Rozzini[28] subsequently externally validated the index's performance predicting 6-month mortality in a retrospective analysis of 840 consecutively admitted participants to a hospital in Italy, and found monotonic increases in mortality for each predicted risk level (observed 4%, 10%, 25%, and 46% 6-month mortality).

The Drame (2008)[17] index for 2-year mortality was developed in hospitalized adults age >75 based on secondary data obtained in the Emergency Department as part of the SAFES study in France (n=870). It showed good calibration and discrimination in a split sample validation of 436 older adults.

COMMENT

Our review identified 16 validated non-disease specific prognostic indices for older adults. Studies were abstracted for information about index quality, including potential for bias, generalizability, and accuracy.

We highlight criteria for evaluating prognostic indices and identified several high-quality prognostic indices. Unfortunately, although these indices hold the promise of improving the targeting of interventions in older adults, there is insufficient evidence at this time to recommend the widespread use of prognostic indices in clinical practice. Only two indices were validated by investigators not involved in the studies development, and no index had been prospectively tested and found to be accurate in a large diverse sample. Confidence intervals were not presented for either measures of discrimination or calibration for fourteen indices. By our measures, no study was completely free from potential sources of bias. Testing of transportability was limited, raising concerns about overfitting and underfitting. These factors limit the clinician's ability to assess the accuracy of these indices across patient groups that differ according to severity of illness, methodology of data collection, geographic location, and time.

Even if quality barriers are overcome, important limitations remain. Several indices require collection of information that may not be routinely assessed in elderly patients, such as Activities of Daily Living. Many of these indices rely on clinical information from administrative datasets, and the accuracy of ICD-9 codes has been called into question.[43] Thus, indices by Gagne, Inouye, and Levine may be better suited to risk adjustment than clinical use. Moreover, coding algorithms are subject to change. The MDS has been updated to a new version (3.0) since the development of indices for nursing home patients, and some variables in indices by Porock and Flacker have been changed or are no longer present.[44] Finally, PubMed has no single Medical Subject Heading term for prognostic index, making it difficult for a busy clinician to locate these studies.

Ultimately, an index will be judged not only on it's accuracy across diverse settings, but on it's clinical impact. Studies that demonstrate impact on prognostic estimates, clinician behavior, and patient outcomes have a higher level of evidence for use in clinical decision making (e.g. Ottowa Ankle rules).[12] We are aware of only two small studies that tested the impact of these indices on clinical outcomes.[23,45] The highest level of evidence, however, would come from large prospective trials that randomize clinicians to using the index or not, evaluating the impact of the index on prognostic estimates, clinical decision making, and patient outcomes. Such large randomized trials have not been performed.

None of the c-statistics for the included indices were higher than 0.90, suggesting unexplained variation in mortality. However, discriminatory ability of these indices is consistent with other indices that commonly drive clinical decisions, such as: the CHADS2 index to help determine warfarin therapy (c-statistic 0.68–0.72)[46]; the Framingham risk

score to help determine lipid therapy (c-statistic 0.63–0.83)[47]; and the TIMI risk score to help determine invasive therapy for unstable angina (c-statistic 0.65)[48].

There may be a limited role for the highest quality indices in the right settings. If patient characteristics align closely with those of the development or validation cohorts, clinicians may find prognostic information useful to help inform, though not replace, their clinical judgment. Prediction rules have been shown to outperform clinicians in terms of prognostication,[49,50] whereas human prediction on its own is fraught with bias.[51] The indices we identified were developed from heterogeneous groups of patients. Applying this information to the individual patient, therefore, requires a nuanced use of the index. Patients are likely to have conditions that are not included in the index (e.g. Parkinson's). The clinician must account for these conditions and decide whether their impact is adequately accounted for by the indices' predictors.

Indices are most likely to be clinically useful when they predict a wide range of mortality. Clinical decisions are most likely to be influenced by either very low or very high mortality risk. While 10 indices predicted >50% mortality, only 3 predicted > 80% risk in the highest risk group. Mid-range probabilities may still useful in clinical decisions in which life-expectancy plays a role, allowing patient preferences to drive the physician's recommendation. The following case illustrates this issue.

Case

Ms. A is a 75-year-old clinic patient who has been hospitalized twice in the past year for COPD, and has a history of diabetes and difficulty walking a quarter mile. She has not been previously screened for colon cancer. The USPSTF recommends that individual factors should determine the decision to screen or not screen patients ages 75–85, patients must live at least 7 years to benefit from screening, and the net benefits in this age group are small.[52] Using indices developed for community dwelling elders, Ms. A has a 54–67% mortality risk at 4 years (Lee index) and 75% at 9 years (Schonberg index). Should Ms. A. have colorectal cancer screening?

In this case, the prognostic information may be helpful as her physician discusses the possibility of colon cancer screening in relation to other health priorities, such as maintaining mobility. Since her median life expectancy is less than 4 years, Ms A will probably not live long enough to benefit from screening. And if screening is difficult for her, there is enough uncertainty in her likelihood of benefit that she probably should focus on other priorities. However, if she feels strongly about wanting to be screened, the estimates are not strong enough on their own to refute that decision.

Limitations

We have refrained from explicitly ranking or categorizing the quality of these indices, recognizing that no agreed upon scientifically-developed system for rating index quality currently exists. Some will argue that minimizing risk for potential bias is of critical importance, while others argue that an index should be judged on its ability to perform accurately across diverse settings. Our review excluded indices that estimated only relative risk or had not been validated, and future research may find that some of these indices are generalizable and accurate. Our ability to assess publication bias was limited by our small sample size.

CONCLUSION

While neither a clinician nor an index can predict with absolute certainty how long an older adult will live, validated prognostic indices might improve the accuracy of the prognostic

assumptions that influence our clinical decisions. However, further research is needed before general prognostic indices for elders can be recommended for routine use. Future research should focus on prospectively testing the validity of these indices across diverse clinical settings, and analyzing their impact on clinical decision making and patient outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support: Dr. Yourman was supported by an NIA T32 pre-doctoral fellowship position (5T32AG000212). Dr. Smith was supported by a career development grant from the National Center for Research Resources UCSF-CTSI (UL1 RR024131).

Role of the Sponsor: The funding organizations had no role in the design and conduct of the study, in the collection, analysis, and interpretation of the data, or in the preparation, review, or approval of the manuscript.

We gratefully acknowledge Gloria Won, MLIS, Librarian, H.M. Fishbon Memorial Library, University of California, San Francisco for her help with searching EMBASE.

References

1. Casarett D, Karlawish J, Morales K, Crowley R, Mirsch T, Asch DA. Improving the use of hospice services in nursing homes: a randomized controlled trial. JAMA. Jul 13; 2005 294(2):211–217. [PubMed: 16014595]
2. Walter LC, Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. JAMA. Jun 6; 2001 285(21):2750–2756. [PubMed: 11386931]
3. Mehta KM, Fung KZ, Kistler CE, Chang A, Walter LC. Impact of cognitive impairment on screening mammography use in older US women. Am J Public Health. Oct; 2010 100(10):1917–1923. [PubMed: 20075325]
4. Sima CS, Panageas KS, Schrag D. Cancer screening among patients with advanced cancer. JAMA. Oct 13; 2010 304(14):1584–1591. [PubMed: 20940384]
5. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. Mar 16; 1999 130(6):515–524. [PubMed: 10075620]
6. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. Jan; 2010 21(1):128–138. [PubMed: 20010215]
7. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. JAMA. 1997; 277(6):488–494. [PubMed: 9020274]
8. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. JAMA. 1994; 272(3): 234–237. [PubMed: 8022043]
9. Braitman LE, Davidoff F. Predicting clinical states in individual patients. Ann Intern Med. 1996; 125(5):406–412. [PubMed: 8702092]
10. Randolph, A.; Cook, DJ.; Guyatt, G. Prognosis. In: GG; DR; MOM; DJC, editors. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. New York, NY: McGraw-Hill; 2008.
11. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med. Mar 21; 2006 144(6):427–437. [PubMed: 16549855]
12. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature:XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. JAMA. Jul 5; 2000 284(1):79–84. [PubMed: 10872017]
13. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. Apr; 1982 143(1):29–36. [PubMed: 7063747]

14. Carey EC, Covinsky KE, Lui LY, Eng C, Sands LP, Walter LC. Prediction of mortality in community-living frail elderly people with long-term care needs. J Am Geriatr Soc. Jan; 2008 56(1):68–75. [PubMed: 18031487]
15. Carey EC, Walter LC, Lindquist K, Covinsky KE. Development and validation of a functional morbidity index to predict mortality in community-dwelling elders. J Gen Intern Med. Oct; 2004 19(10):1027–1033. [PubMed: 15482555]
16. Di Bari M, Balzi D, Roberts AT, et al. Prognostic stratification of older persons based on simple administrative data: development and validation of the “Silver Code,” to be used in emergency department triage. J Gerontol A Biol Sci Med Sci. 2010; 65(2):159–164. [PubMed: 19349591]
17. Drame M, Novella JL, Lang PO, et al. Derivation and validation of a mortality-risk index from a cohort of frail elderly patients hospitalised in medical wards via emergencies: the SAFES study. Eur J Epidemiol. 2008; 23(12):783–791. [PubMed: 18941907]
18. Fischer SM, Gozansky WS, Sauaia A, Min SJ, Kutner JS, Kramer A. A practical tool to identify patients who may benefit from a palliative approach: the CARING criteria. J Pain Symptom Manage. 2006; 31(4):285–292. [PubMed: 16632076]
19. Flacker JM, Kiely DK. A practical approach to identifying mortality-related factors in established long-term care residents. J Am Geriatr Soc. 1998; 46(8):1012–1015. [PubMed: 9706893]
20. Flacker JM, Kiely DK. Mortality-related factors and 1-year survival in nursing home residents. J Am Geriatr Soc. 2003; 51(2):213–221. [PubMed: 12558718]
21. Inouye SK, Bogardus ST Jr, Vitagliano G, et al. Burden of illness score for elderly persons: risk adjustment incorporating the cumulative impact of diseases, physiologic abnormalities, and functional impairments. Med Care. 2003; 41(1):70–83. [PubMed: 12544545]
22. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. Jama. Feb 15; 2006 295(7):801–808. [PubMed: 16478903]
23. Levine SK, Sachs GA, Jin L, Meltzer D. A prognostic model for 1-year mortality in older adults after hospital discharge. Am J Med. 2007; 120(5):455–460. [PubMed: 17466658]
24. Mazzaglia G, Roti L, Corsini G, et al. Screening of older community-dwelling people at risk for death and hospitalization: the Assistenza Socio-Sanitaria in Italia project. J Am Geriatr Soc. 2007; 55(12):1955–1960. [PubMed: 17944891]
25. Pilotto A, Ferrucci L, Franceschi M, et al. Development and validation of a multidimensional prognostic index for one-year mortality from comprehensive geriatric assessment in hospitalized older patients. Rejuvenation Res. 2008; 11(1):151–161. [PubMed: 18173367]
26. Porock D, Oliver DP, Zweig S, et al. Predicting death in the nursing home: development and validation of the 6-month Minimum Data Set mortality risk index. J Gerontol A Biol Sci Med Sci. 2005; 60(4):491–498. [PubMed: 15933390]
27. Porock D, Parker-Oliver D, Petroski GF, Rantz M. The MDS Mortality Risk Index: The evolution of a method for predicting 6-month mortality in nursing home residents. BMC Res Notes. 2010; 3:200. [PubMed: 20637076]
28. Rozzini R, Sabatini T, Trabucchi M. Prediction of 6-month mortality among older hospitalized adults. JAMA. 2001; 286(11):1315–1316. [PubMed: 11560535]
29. Sancarlo D, D’Onofrio G, Franceschi M, et al. Validation of a modified-multidimensional prognostic index (m-MPI) including the mini nutritional assessment short-form (MNA-SF) for the prediction of one-year mortality in hospitalized elderly patients. The Journal of Nutrition, Health & Aging. 2010:1–5.
30. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. J Gen Intern Med. 2009; 24(10):1115–1122. [PubMed: 19649678]
31. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. Jul; 2011 64(7):749–759. [PubMed: 21208778]
32. Walter LC, Brand RJ, Counsell SR, et al. Development and validation of a prognostic index for 1-year mortality in older adults after hospitalization. Jama. Jun 20; 2001 285(23):2987–2994. [PubMed: 11410097]

33. Teno JM, Harrell FE Jr, Knaus W, et al. Prediction of survival for older hospitalized patients: the HELP survival model. Hospitalized Elderly Longitudinal Project. J Am Geriatr Soc. May; 2000 48(5 Suppl):S16–24. [PubMed: 10809452]
34. Kruse RL, Parker Oliver D, Mehr DR, Petroski GF, Swenson DL, Zweig SC. Using mortality risk scores for long-term prognosis of nursing home residents: caution is recommended. J Gerontol A Biol Sci Med Sci. 2010; 65(11):1235–1241. [PubMed: 20639529]
35. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. External Validation of an Index to Predict Up to 9-Year Mortality of Community-Dwelling Adults Aged 65 and Older. J Am Geriatr Soc. Aug; 2011 59(8):1444–1451. [PubMed: 21797837]
36. Youngwerth J, Fischer SM, S-J M, Kutner JS. Who would benefit from a palliative care consult in the hospital? CARING about prognosis (CAP) (313-C). J Pain Symptom Manage. 2011; 41(1): 192–193.
37. Desai MM, Bogardus ST Jr, Williams CS, Vitagliano G, Inouye SK. Development and validation of a risk-adjustment index for older patients: The high-risk diagnoses for the elderly scale. Journal of the American Geriatrics Society. 2002; 50(3):474–481. [PubMed: 11943043]
38. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. J Clin Epidemiol. Oct; 1993 46(10):1075–1079. discussion 1081–1090. [PubMed: 8410092]
39. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987; 40(5):373–383. [PubMed: 3558716]
40. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care. Jun; 2009 47(6):626–633. [PubMed: 19433995]
41. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care. Jan; 1998 36(1):8–27. [PubMed: 9431328]
42. Inouye SK, Wagner DR, Acampora D, et al. A predictive index for functional decline in hospitalized elderly medical patients. J Gen Intern Med. Dec; 1993 8(12):645–652. [PubMed: 8120679]
43. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. N Engl J Med. Feb 11; 1988 318(6): 352–355. [PubMed: 3123929]
44. [Accessed November 15, 2011, 2011] MDS 3.0 Technical Information. 2011. http://www.cms.gov/NursingHomeQualityInits/30_NHQIMDS30TechnicalInformation.asp#TopOfPage
45. Oliver DP, Bickel-Swenson D, Zweig S, Kruse R, Mehr D. Experience with implementation of a quality improvement project for the care of nursing home residents. J Nurs Care Qual. Apr-Jun; 2009 24(2):100–104. [PubMed: 19287246]
46. Rietbrock S, Heeley E, Plumb J, van Staa T. Chronic atrial fibrillation: Incidence, prevalence, and prediction of stroke using the Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, and prior Stroke or transient ischemic attack (CHADS2) risk stratification scheme. Am Heart J. 2008; 156(1):57–64. [PubMed: 18585497]
47. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA. Jul 11; 2001 286(2):180–187. [PubMed: 11448281]
48. Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. JAMA. 2000; 284(7):835–842. [PubMed: 10938172]
49. Van Zee KJ, Manasseh DM, Bevilacqua JL, et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. Ann Surg Oncol. Dec; 2003 10(10):1140–1151. [PubMed: 14654469]
50. Ross PL, Gerigk C, Gonen M, et al. Comparisons of nomograms and urologists' predictions in prostate cancer. Semin Urol Oncol. May; 2002 20(2):82–88. [PubMed: 12012293]

51. Einhorn, HJ.; Hogarth, RM. Behavioral decision theory: process of judgement and choice. In: Bell, DE.; Raiffa, H.; Tversky, A., editors. Decision-Making: Descriptive, Normative, and Prescriptive Interactions. Cambridge, MA: Cambridge University Press; 1988. p. 113-146.
52. U.S. Preventive Services Task Force. [Accessed November 15, 2011, 2011] Screening for Colorectal Cancer. 2008. http://www.uspreventiveservicestaskforce.org/uspstf08/colocancer/colors.htm
53. Holmes HM, Hayley DC, Alexander GC, Sachs GA. Reconsidering medication appropriateness for patients late in life. Arch Intern Med. 2006; 166(6):605–609. [PubMed: 16567597]
54. Vollrath AM, Sinclair C, Hallenbeck J. Discontinuing cardiovascular medications at the end of life: lipid-lowering agents. J Palliat Med. Aug; 2005 8(4):876–881. [PubMed: 16128666]
55. Fillinger M. Who should we operate on and how do we decide: predicting rupture and survival in patients with aortic aneurysm. Semin Vasc Surg. 2007; 20(2):121–127. [PubMed: 17580250]
56. Conway KP, Byrne J, Townsend M, Lane IF. Prognosis of patients turned down for conventional abdominal aortic aneurysm repair in the endovascular and sonographic era: Szilagyi revisited? J Vasc Surg. 2001; 33(4):752–757. [PubMed: 11296328]
57. Lederle FA, Johnson GR, Wilson SE, et al. Rupture rate of large abdominal aortic aneurysms in patients refusing or unfit for elective repair. JAMA. 2002; 287(22):2968–2972. [PubMed: 12052126]
58. Mohan IV, Harris PL. When not to operate for abdominal aortic aneurysms. Semin Interv Cardiol. 2000; 5(1):15–19. [PubMed: 10875219]
59. Brown AF, Mangione CM, Saliba D, Sarkisian CA. Guidelines for improving the care of the older person with diabetes mellitus. J Am Geriatr Soc. 2003; 51(5 Suppl Guidelines):S265–280. [PubMed: 12694461]
60. Beckett NS, Peters R, Fletcher AE, et al. Treatment of hypertension in patients 80 years of age or older. N Engl J Med. 2008; 358(18):1887–1898. [PubMed: 18378519]
61. Walter LC, Lewis CL, Barton MB. Screening for colorectal, breast, and cervical cancer in the elderly: a review of the evidence. Am J Med. Oct; 2005 118(10):1078–1086. [PubMed: 16194635]
62. Walter LC, Lindquist K, Nugent S, et al. Impact of age and comorbidity on colorectal cancer screening among older veterans. Ann Intern Med. Apr 7; 2009 150(7):465–473. [PubMed: 19349631]
63. U.S. Preventive Services Task Force. Colon cancer screening (USPSTF recommendation). J Am Geriatr Soc. 2000; 48(3):333–335. [PubMed: 10733063]
64. Warner E. Clinical practice. Breast-cancer screening. N Engl J Med. Sep 15; 2011 365(11):1025–1032. [PubMed: 21916640]
65. American Geriatrics Society Clinical Practice Committee. Breast cancer screening in older women. J Am Geriatr Soc. 2000; 48(7):842–844. [PubMed: 10894327]
66. U.S. Preventive Services Task Force. [Accessed November 15, 2011, 2011] Screening for Breast Cancer. 2009. http://www.uspreventiveservicestaskforce.org/uspstf09/breastcancer/brcanrs.htm
67. Rahimtoola SH. Choice of prosthetic heart valve in adults an update. J Am Coll Cardiol. 2010; 55(22):2413–2426. [PubMed: 20510209]
68. Huang ES, Zhang Q, Gandra N, Chin MH, Meltzer DO. The effect of comorbid illness and functional status on the expected benefits of intensive glucose control in older patients with type 2 diabetes: a decision analysis. Ann Intern Med. 2008; 149(1):11–19. [PubMed: 18591633]
69. Skyler JS, Bergenstal R, Bonow RO, et al. Intensive glycemic control and the prevention of cardiovascular events: implications of the ACCORD, ADVANCE, and VA Diabetes Trials: a position statement of the American Diabetes Association and a Scientific Statement of the American College of Cardiology Foundation and the American Heart Association. J Am Coll Cardiol. 2009; 53(3):298–304. [PubMed: 19147051]
70. Walter LC, Bertenthal D, Lindquist K, Konety BR. PSA screening among elderly men with limited life expectancies. JAMA. Nov 15; 2006 296(19):2336–2342. [PubMed: 17105796]
71. Wolf AM, Wender RC, Etzioni RB, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. CA Cancer J Clin. 60(2):70–98. [PubMed: 20200110]

72. Akhoundi FH, Ghorbani A, Soltani A, Meysamie A. Favorable functional outcomes in acute ischemic stroke patients with subclinical hypothyroidism. Neurology. Jul 26; 2011 77(4):349–354. [PubMed: 21715712]

73. Buchholz TA. Radiation therapy for early-stage breast cancer after breast-conserving surgery. N Engl J Med. 2009; 360(1):63–70. [PubMed: 19118305]

74. Geiger AM, Thwin SS, Lash TL, et al. Recurrences and second primary breast cancers in older women with initial early-stage disease. Cancer. 2007; 109(5):966–974. [PubMed: 17243096]

TABLE 1

Example Clinical Decisions Influenced by Life Expectancy[a]

LIFE EXPECTANCY	EXAMPLE CLINICAL DECISIONS	GUIDELINES
Short-Term (<2 years)		
<6 months	consider discontinuation of statins[53,54]	None
<6 months	consider referral to hospice	Medicare Regulations
<1–2 years	consider non-operative management of asymptomatic abdominal aortic aneurysm[55–58]	None
Mid-Term (2–3 years)		
<2–3 years	consider that blood pressure/lipid control in Diabetes Mellitus unlikely to prevent *macrovascular* complications	California Healthcare Foundation and AGS[59]
<2–3 years	consider that lowering blood pressure <140/80 unlikely to improve cardiovascular outcomes[53,60]	None
Long-term (>3 years)		
a) <5 years or b) <7 years	consider discontinuation of colon cancer screening[61,62]	**a.** AGS[63] **b.** USPSTF[52]
a) <5 years or b) "limited"	consider discontinuation of breast cancer screening[61,64]	**a.** AGS[65] **b.** USPSTF[66]
< 5 years	consider that stented bioprosthetic heart valve may be preferable to metallic valve[67]	None
<5 years	consider limited benefit to lowering HbA1C therapeutic target to <8%[53]	California Healthcare Foundation and AGS[59]
<8 years	consider that tight glycemic control in Diabetes Mellitus is unlikely to prevent *microvascular* complications[53,68,69]	California Healthcare Foundation and AGS[59]
<10 years	consider discontinuation of prostate cancer screening[70]	ACS[71] and AUA[72]
<15 years	consider that irradiation therapy to ipsilateral breast may not have mortality benefit if life expectancy <15 years (for patients with T1,T2 ER+ breast cancer status post breast-conserving surgery and hormonal therapy) [73,74]	None

[a]Prognosis is only one of many important factors to consider for these clinical decisions.

Abbreviations: AGS – American Geriatrics Society; USPSTF – US Preventive Services Task Force; ACS – American Cancer Society; AUA – American Urological Association.

TABLE 2

Factors to Consider When Evaluating the Quality of Prognostic Indices[a]

TERMS	EXPLANATIONS	MEASUREMENTS/EXAMPLES
BIAS	**Systematic variation (non-random error) in the development or validation of a prognostic index**	-Example: 13% of participants in the Flacker[20] development cohort were lost to follow up (unknown mortality at 1 year), and may have systematically differed.
ACCURACY	**The degree to which predicted outcomes match observed outcomes**	
calibration	How close each level of prediction is to what is observed for that risk group	-Compares predicted vs. observed mortality rate -Hosmer-Lemeshow[b]
discrimination	How well those who die are separated from those who do not die	-c-statistic[c]
GENERALIZABILITY	**Ability of a prognostic index to provide accurate predictions in a new sample of patients**	
reproducibility	The index is accurate in patients who were not included in the development cohort but who are from the same underlying population. A measure of overfitting (matching the predictive model to random noise in the data).	-Data-resampling (also called bootstrapping)[d]
transportability	The index is accurate in patients drawn from a different but related population or in data collected by using methods that differ from those used in development. A measure of both overfitting and underfitting (the omission of important predictors of mortality).	-Non-randomly split-sample[e] or independent validation
i. Methodologic	Accuracy is maintained when the index is tested in data collected using different methods. Independent validation tests the accuracy of the index by investigators not involved in the development of the index	-Example: Porock[26] developed index and Kruse[34] independently validated it
ii. Historical	Accuracy is maintained when the index is tested in data from a different calendar time	-Example: Inouye development sample was from 1989–1991, validation sample was from 1995–1996[21]
iii. Geographic	Accuracy is maintained when the index is tested in data from different locations	-Example: Lee developed in eastern/western/central US, validated in southern US[22]
iv. Spectrum	Accuracy is maintained in a patient sample that is, on average, more or less advanced in disease process or that has a somewhat different disease process or trajectory	-Example: Walter developed in tertiary care hospital and validated in community hospital[32]
v. Follow-up interval	Accuracy is maintained when the index is tested over a longer or shorter period	-Example: Pilotto[25] developed for 1-year and San Carlo[29] validated for 1-month mortality

[a] Table 2 Adapted from Justice et al. (Ann Intern Med. 1999), Hayden (Ann Intern Med 2006), McGinn (JAMA 2000), and Steyerberg et al. (Epidemiology 2009)[6]

[b] Higher values closest to 1 indicate better fit

[c] Higher values closest to 1 indicate better discrimination

[d] Develop the index in the entire data set, then validate it in multiple bootstrap samples generated from the original sample with replacement.

[e] Develop the index in one part of the data, and validate it in another portion that differs on some major variable. Non-randomly split samples measure an index's transportability better than randomly split samples

TABLE 3

Summary of 16 Validated General Prognostic Indices for Older Adults

AUTHOR	INDEX DESCRIPTION	GENERALIZABILITY[a]		ACCURACY		
		Characteristics of Development Cohort Participants	Characteristics of Validation Cohort Participants	Discrimination (95% CI where reported)	Calibration – Predicted and Observed Mortality for Low, Intermediate, and High Risk Groups over Specified Time Period	
					Predicted Mortality (95% CI where reported)	Observed Mortality (95% CI where reported)
Community						
Gagne 2011[31]	1-year index for low income community dwelling elders	n=120,679 average age 80 83% female 29% hospitalized/last year 9% nursing home residents median 18 distinct ICD-9 diagnoses 9% 1-year mortality	n=123,855 average age 79 77% female 27% hospitalized/last year 9% nursing home residents median 12 distinct ICD-9 diagnoses 8% 1-year mortality	validation c=0.79 (0.79–0.79)	<7% 7–17% >17%	3% 12% 29%
Mazzaglia 2007[24]	15-month index for community-dwelling elders	n=2,470 mean age 75 56% female 56% 5% 15-month mortality	n=2,926 mean age 75 59% female 4% 15-month mortality	derivation c=0.75 (0.72–0.78) validation c=0.75 (0.73–0.78)	0% (0.04–1.1) 1% (0.4–3.6) 1% (0.4–2.3) 10% (7.9–11.5)	0% (0.03–1.1) 1% (0.1–2.1) 1% (0.2–1.1) 8% (6.7–9.8)
Carey 2004[15]	2-year index for community dwelling elders	n=4,516 mean age 78 61% female 84% white 13% dependent in ≥1 ADL 28% difficulty with stairs 13% diabetes 14% cancer 31% heart disease 10% mortality	n=2,877 mean age 78 61% female 73% white 17% dependent in ≥1 ADL 41% difficulty with stairs 14% diabetes 13% cancer 32% heart disease 12% 2-year mortality	derivation c=0.76 validation c=0.74	3% 11% 34%	5% 12% 36%
Carey 2008[14]	3-year index for nursing-home eligible community- dwelling elders	n=2,232 mean age 79 68% female 62% difficulty bathing on own 23% diabetes 23% coronary artery disease 37% 3-year mortality	n=1,667 mean age 79 76% female 72% difficulty bathing on own 27% diabetes 27% coronary artery disease 36% 3-year mortality	derivation c=0.66 validation c=0.69	21% 36% 54%	18% 35% 55%

JAMA. Author manuscript; available in PMC 2013 October 08.

AUTHOR	INDEX DESCRIPTION	GENERALIZABILITY[a]		ACCURACY		
		Characteristics of Development Cohort Participants	Characteristics of Validation Cohort Participants	Discrimination (95% CI where reported)	Calibration – Predicted and Observed Mortality for Low, Intermediate, and High Risk Groups over Specified Time Period	
					Predicted Mortality (95% CI where reported)	Observed Mortality (95% CI where reported)
Lee 2006[22]	Lee 4-year index for community dwelling elders	n=11,701 mean age 67 57% female 81% white 15% diabetes 12% cancer 17% coronary artery disease 12% 4-year mortality	n=8009 mean age 67 57% female 71% white 16% diabetes 11% cancer 19% coronary artery disease 13% 4-year mortality	derivation c=0.84 validation c=0.82	<5% 4–9% 12–19% 22–24% 43–48% 54–67%	<5% 6–9% 15–20% 20–28% 44–45% 59–64%
Schonberg 2009[30]	5-year index for community dwelling elders	n=16,077 27% age>80 60% female 85% white 18% dependent in at least 1 ADL or AIDL 15% diabetes 15% cancer 11% coronary artery disease 17% 5-year mortality	n=8,038 validation cohort participants reported as “similar” to development participants	validation c=0.75	2% (1–4) 8% (6–9) 25% (23–28) 47% (32–42) 71% (65–77)	3% (1–6) 8% (6–10) 29% (25–33) 49% (43–55) 62% (54–70)
Nursing Home						
Porock 2005[26]	6-month index for nursing home patients	n=32,599 51% age >85 74% female 92% white 26% 6-month mortality	n=10,991 50% age>85 73% female 92% white 26% 6-month mortality	development c=0.75	9% 23% 43% 62% 81%	10% 23% 43% 58% 82%
Flacker 2003[20]	1-year index for long stay (>1 yr) nursing home patients	n=22,749 49% age>84 74% female 83% white 9% cancer 26% heart disease 56% dementia 21% mortality	N=40,328 46%>84 73% female 82% white 9% cancer 24% heart disease 54% dementia 22% mortality	derivation c=0.71	8% 13% 31% 52% 76% 80%	9% 13% 31% 59% 79% 100%
Hospital						
di Bari 2010[16]	1-year index for emergency department triage of hospitalized elders	n=5,457 71% age ≥ 80 55% female 6% cardiovascular disease 2% respiratory disease	n=5,456 Characteristics reported in development cohort are for all participants, random split	development c=0.66 validation c=0.64	20% 28% 41% 52%	21% 29% 40% 50%

AUTHOR	INDEX DESCRIPTION	GENERALIZABILITY[a]		ACCURACY		
		Characteristics of Development Cohort Participants	Characteristics of Validation Cohort Participants	Discrimination (95% CI where reported)	Calibration – Predicted and Observed Mortality for Low, Intermediate, and High Risk Groups over Specified Time Period	
					Predicted Mortality (95% CI where reported)	Observed Mortality (95% CI where reported)
		50% ≥ 5 medications 34% 1-year mortality	sample validation not reported separately			
Fischer 2006[18]	1-year index for hospitalized elders on admission	n=435 mean age 63 2% female 23% cancer 36% ≥ 2 hospitalizations in past year 26% 1-year mortality	n=438 Characteristics reported in development cohort are for all participants, historical split sample validation not reported separately	development c=0.82[d]	<18% 18–48% >49%	NR
Inouye 2003[21]	1-year index for hospitalized elders on admission	n=525 mean age 79 56% female 91% white 7% nursing home resident 11% pneumonia 24% albumin ≤ 3.5 27% creatinine > 1.5 29% 1-year mortality	n=1,246 average age 81 52% female 94% white 32% nursing home resident 100% pneumonia 49% albumin ≤ 3.5 20% creatinine > 1.5 39% 1-year mortality	development c=0.83 validation c=0.77	8% 24% 51% 74%	5% 17% 33% 61%
Pilotto 2008[25]	1-year index for hospitalized elders on admission	n=838 mean age 79 55% female mean 4/6 functional ADL mean 3 errors on SPMSQ mean 4 medications 18% 1-year mortality	n=857 mean age 78 53% female mean 4/6 functional ADL mean 3 errors on SPMSQ mean 4 medications 17% 1-year mortality	development c=0.75 (0.71–0.80)	8% 21% 43%	6% 23% 45%
Teno 2000[33]	1 year index for hospitalized elders on admission	n=1,266 median age 85 25% cancer 9% congestive heart failure 61% female 40% 1-year mortality	Validation performed using resampling from the development cohort	derivation c=0.73 validation c=0.74	22% 58% 66% 82% 86% 93%	26% 57% 72% 80% 89% 95%
Levine 2007[23]	1-year index for hospitalized elders following discharge	n=2,739 mean age 78 63% female 15% discharged to SNF 23% CHF 27% COPD 8% metastatic cancer	n=3,643 mean age 78 65% female 17% discharged to SNF 24% CHF 24% COPD 4% metastatic cancer	derivation c=0.67 validation c=0.65	14% (11–16) 18% (15–21 32% (28–36)	14% (12–16) 24% (22–27) 30% (26–33)

AUTHOR	INDEX DESCRIPTION	GENERALIZABILITY[a]		ACCURACY		
		Characteristics of Development Cohort Participants	Characteristics of Validation Cohort Participants	Discrimination (95% CI where reported)	Calibration – Predicted and Observed Mortality for Low, Intermediate, and High Risk Groups over Specified Time Period	
					Predicted Mortality (95% CI where reported)	Observed Mortality (95% CI where reported)
		26% 1-year mortality	26% 1-year mortality		46% (42–50)	42% (38–45)
Walter 2001[32]	1 year index for hospitalized elders following discharge	n=1,495 mean age 81 67% female 60% white 40% black 27% CHF 30% discharged to a nursing home or SNF 27% dependent in all 5 ADL 10% albumin < 3.0 40% creatinine ≥ 1.5 33% 1-year mortality	n=1,427 mean age 79 61% female 88% white 12% black 29% CHF 14% discharged to a nursing home or SNF 15% dependent in 5 ADL 19% albumin < 3.0 20% creatinine ≥ 1.5 28% 1-year mortality	derivation c=0.75 validation c=0.79	13% (10–16) 20% (16–24) 37% (33–41) 68% (63–73)	4% (2–6) 19% (15–23) 34% (29–39) 64% (58–70)
Drame 2008[17]	2-year index for hospitalized elders following discharge	n=870 mean age 85 64% female 60% dependent in 1 ADL 15% Charlson co-morbidity score ≥ 3 44% 2-year mortality	n=436 mean age 85 64% female 61% dependent in 1 ADL 19% Charlson co-morbidity score ≥ 3 44% 2-year mortality	derivation c=0.72 (0.68–0.75) validation c=0.71 (0.66–0.76)	21% (15–26) 50% (45–54) 62% (59–71)	22% (14–29) 49% (42–55) 65% (55–76)

Abbreviations: NR= not reported; SNF=skilled nursing facility; SPMSQ=Short Portable Mental Status Questionnaire; CHF=congestive heart failure; COPD=chronic obstructive pulmonary disease.

[a]Descriptive information on age, sex, race, morbidity, and mortality is reported where available.

Table 4

Potential Sources of Bias for 16 Validated General Prognostic Indices

INDEX	Adequate description of sample, (Participation rate)[a]	Clearly defined, reproducible prognostic variables[b]	Blinded measurement of potential prognostic variables and mortality[c]	Completeness of potential predictors[d]	Completeness of mortality outcome[e]	Model building is conceptually based, stability of model tested[f]
Community						
Gagne 2011[31]	Partly Race/ethnicity not described (participation not optional in this administrative dataset)	Partly ICD-9 codes have limited reproducibility	Yes	NR	NR	Partly Stability of model not tested
Mazzaglia 2007[24]	Partly Race/ethnicity not described (Italian sample), participants not compared with non-participants	Partly “Inadequacy of income” not well described	Yes	NR	99%	Yes
Carey 2004[15]	Partly No comparison of respondents to non-respondents	Yes	Yes	99.3%	NR	Yes
Carey 2008[14]	Yes (participation not optional in this administrative dataset)	Yes	Yes	92%	NR	No Not conceptually based; stability of model not tested
Lee 2006[22]	Partly Participants not compared with non-participants (81% participation rate)	Yes	Yes	NR	98%	Yes
Schonberg 2009[30]	Partly Participants not compared with non-participants (74% participation rate)	Yes	Yes	95%	97%	Yes
Nursing Home						
Porock 2005[26]	Partly Comorbities not described (participation not optional in this administrative dataset)	Yes	Yes	NR[g]	>99% linkage to Missouri death certificates	Yes
Flacker 2003[20]	Yes (participation not optional in this administrative dataset)	Yes	Yes	NR	87%	Yes
Hospital						
di Bari 2010[16]	Partly Race/ethnicity not described (Italian sample); “admitted for medical reasons” not clear; (participation not optional in this administrative dataset)	Partly Admission to “day hospital” not clearly defined	Yes	91% linkage across 4 datasets, 0% after linkage	91% linkage across 4 datasets, including mortality	No Not conceptually based; stability not tested
Fischer 2006[18]	Yes	Yes	Partly	NR	98%	Partly

INDEX	Adequate description of sample, (Participation rate)[a]	Clearly defined, reproducible prognostic variables[b]	Blinded measurement of potential prognostic variables and mortality[c]	Completeness of potential predictors[d]	Completeness of mortality outcome[e]	Model building is conceptually based, stability of model tested[f]
			For validation, 10% blinded chart review with 100% agreement			Final predictors for model selected a priori; stability not tested
Inouye 2003[21]	Partly Participants not compared with non-participants (86% participation rate)	Partly ICD-9 codes have limited reproducibility	Yes	>99% for all predictors	100%	Yes
Pilotto 2008[25]	Partly Race/ethnicity not described (Italian sample), participants not compared with non-participants (80% participation rate)	Yes	Yes	90%	82%	Yes
Teno 2000[33]	Partly Race/ethnicity and participation rate not described	Yes	Yes	81%	100%	Yes
Levine 2007[23]	Partly Participation rate not reported	Partly ICD-9 codes have limited reproducibility	Yes	NR	>99%	No Not conceptually based; stability not tested
Walter 2001[32]	Partly Participation rate not described.	Yes	Yes	96%	100%	Yes
Drame 2008[17]	Partly Race/ethnicity not described (French sample) (87% participation rate)	Yes	Yes	NR	92%	No Not conceptually based; stability of model not tested

[a] Sample description: study and source populations clearly defined and study sample clearly described (age, sex, race/ethnicity, comorbidities, baseline mortality rates), enrollment procedures clear and, unless administrative data, comparison of participants and non-participants (Yes/Partly/No/Unsure). Participation rates provided for studies requiring consent.

[b] Prognostic variables defined: clear, reproducible measures (Yes/Partly/No/Unsure). ICD-9 codes rated partly due to concerns about reproducibility.[43]

[c] Blinding: developers of the prognostic index were blinded to the measurement of potential prognostic variables and mortality outcomes (Yes/Partly/No/Unsure). Secondary analyses of existing data categorized as yes.

[d] Completeness of Predictors: % sample with complete predictors

[e] Completeness of Mortality: % sample with complete follow up or % successful linkage to vital statistics records (e.g. National Death Index)

[f] Model building: Selection of potential predictors is conceptually based, and stability of model by varying assumptions and/or modeling techniques is tested (Yes/Partly/No/Unsure)

Abbreviations: NR=not reported; SS#= social security number; DOB=date of birth