see commentary on page 83

# An improved comorbidity index for outcome analyses among dialysis patients

Jiannong Liu[1], Zhi Huang[1,2], David T. Gilbertson[1], Robert N. Foley[1,3] and Allan J. Collins[1,3]

[1]United States Renal Data System, Minneapolis, Minnesota, USA; [2]Clinical Statistics, GPRD, Abbott Laboratories, Abbott Park, Illinois, USA and [3]Department of Medicine, University of Minnesota, Minneapolis, Minnesota, USA

**Since comorbid conditions are highly prevalent among patients with end-stage renal disease, indexes measuring them have been widely used to describe the comorbidity burden and to predict outcomes as well as adjust for their roles as confounders. The current comorbidity indexes, however, were developed for general populations or on small patient cohorts. In this study we developed a new index for mortality analyses of dialysis patients based on the 2000 US incident dialysis population, and validated this using the 1999 and 2001 incident and 2000 prevalent dialysis patient populations. Numerical weights were assigned to the comorbid conditions of atherosclerotic heart disease, congestive heart failure, cerebrovascular accident/transient ischemic attack, peripheral vascular disease, dysrhythmia, other cardiac diseases, chronic obstructive pulmonary disease, gastrointestinal bleeding, liver disease, cancer, and diabetes. A patient's comorbidity score was the sum of the weights corresponding to the individual conditions present and could be used as a continuous variable in analyses. Our index performance was almost identical to the individual comorbid conditions regarding model fit, predictive ability, and effect on inference, and it outperformed the widely used Charlson Comorbidity Index.**

*Kidney International* (2010) **77,** 141–151; doi:10.1038/ki.2009.413; published online 11 November 2009

KEYWORDS: comorbidity score; dialysis patients; outcome analysis; validation

**Correspondence:** *Jiannong Liu, United States Renal Data System Coordinating Center, 914 South 8th Street, Suite S-253, Minneapolis, Minnesota 55404, USA. E-mail: jliu@usrds.org*

**Received 2 January 2009; revised 25 August 2009; accepted 1 September 2009; published online 11 November 2009**

Comorbid conditions are highly prevalent among end-stage renal disease (ESRD) patients, and comorbidity indexes have been widely used for describing comorbidity burden, predicting outcomes, and adjusting as a confounder in analyses involving ESRD patients.[1–4] A comorbidity index can give a single-value summary for several comorbid conditions, thereby simplifying the comparison. A comorbidity index can also reduce the dimension of model-based analysis. Too many comorbid conditions and their correlations may distort the information an analysis yields. Large numbers of variables and their correlations may also make the parameter estimation inefficient and the result difficult to interpret.[5] Reducing the dimension of the analysis and therefore reducing the correlations among variables is necessary to produce reliable and meaningful results, especially when the sample size is small.

Several comorbidity indexes have been used for analysis of ESRD patients. The Charlson Comorbidity Index (CCI)[6] is the most widely used. It was developed for mortality analysis based on 604 patients admitted to the medical services at New York Hospital-Cornell Medical Center during a 1-month period in 1984, and was validated based on 685 women with histologically proven primary carcinoma of the breast, who received their first treatment at Yale New Haven Hospital between 1 January 1962 and 31 December 1969. Khan *et al.*[7] proposed a comorbidity index for survival analysis based on 375 ESRD patients, and Davies *et al.*[8] used a different comorbidity index for analyses of 97 continuous ambulatory peritoneal dialysis patients. Van Manen *et al.*[9] compared these three indexes and showed that the CCI performs slightly better than the other two, based on *c*-statistic, a model predictive ability statistic.[10] Fried *et al.*[3] compared the CCI with the Davies comorbidity index based on 415 incident peritoneal dialysis patients. Results showed that CCI was a better predictor for mortality, but the Davies index was a better predictor for hospitalizations.

These comorbidity indexes were developed for general populations or on small samples. The effects on survival of the comorbid conditions included in the CCI are different for the general population than for ESRD patients.[11] Whether the CCI conditions can accurately describe the comorbidity burden for ESRD patients is also questionable. The Khan index did not specify which conditions should be included, and mixed chronological age with comorbid conditions

without showing clear evidence for doing so. In the Davies index, all comorbid conditions were assigned the same weight no matter how different their effects on outcome, and the definitions of the conditions were unclear.

In addition, these indexes were not formally validated or were validated based only on predictive ability or significance as a predictor. The validation of an index should be goal-driven, against a gold standard. To be used to predict mortality, a comorbidity index should have the same, or very similar, predictive ability for mortality as the individual comorbid conditions represented by the index. To be used as an adjustor in survival analyses, a comorbidity index should make the same inferences made when using individual comorbid conditions. An index validated for mortality prediction should not be used for medical cost prediction, unless it was also validated for doing so.

Increasing numbers of analyses are done using administrative data. In addition to observational studies, some clinical trials[12,13] and 'quasi-clinical trials'[14] are conducted based at least partially on the Centers for Medicare & Medicaid (CMS) ESRD database, which is the largest administrative database for ESRD patients in the United States. For those studies, most information on comorbid conditions was derived from the CMS ESRD database. A comorbidity index developed for analyses based on administrative data would be useful. Thus, we propose a new comorbidity index, including and excluding ESRD primary cause, for mortality analyses of dialysis patients, using administrative data, based on the comorbid conditions used by the United States Renal Data System (USRDS).[15] The index was developed using the 2000 US incident dialysis population and validated using the 1999 and 2001 US incident dialysis populations and the 2000 US prevalent dialysis population. The validation was based on model fit, model predictive ability, index predictive ability, and effect on inference. The new index was also checked to see if it can be used for hospitalization and medical cost analysis. Because the CCI performs better than the other indexes for mortality analyses among dialysis patients,[3,9] the new index was compared with it.

## RESULTS

### Description of data

A total of 102,134 incident and 142,517 prevalent dialysis patients were included in this study (Table 1). Mean follow-up time was 2.3 years per patient for incident and 2.5 years per patient for prevalent patients. Percentages of patients with atherosclerotic heart disease (ASHD), congestive heart failure (CHF), cerebrovascular accident/transient ischemic attack (CVA/TIA), chronic obstructive pulmonary disease (COPD), dysrhythmia, diabetes and liver disease increased over the incident year. Compared with the incident cohort, prevalent patients were younger, fewer were white, fewer had diabetes as primary ESRD cause, and fewer had comorbid conditions, possibly because of older patients with more comorbidity dying earlier than younger, healthier patients. The death rate was 26.36, 26.21, 25.59, and 24.55 per 100 patient-years for the 1999, 2000, 2001 incident cohorts and the 2000 prevalent cohort, respectively. It decreased slightly over time for incident patients.

**Table 1 | Patient characteristics: 1999–2001 US incident dialysis patients and 2000 prevalent dialysis patients**

| | Cohort[a] | | | |
|---|---|---|---|---|
| Characteristics | 1999 Incident *n*=33,166 | 2000 Incident *n*=33,077 | 2001 Incident *n*=35,891 | 2000 Prevalent *n*=142,517 |
| *Age* | | | | |
| Mean | 65.6 | 65.0 | 66.0 | 61.0 |
| Median | 69.0 | 68.0 | 69.0 | 63.0 |
| s.d. | 14.7 | 15.0 | 15.0 | 16.0 |
| *Age group, years* | | | | |
| 0–19 | 0.3 | 0.4 | 0.4 | 0.4 |
| 20–29 | 1.9 | 2.1 | 1.9 | 3.0 |
| 30–39 | 4.6 | 4.6 | 4.4 | 7.5 |
| 40–49 | 8.4 | 9.0 | 8.2 | 13.7 |
| 50–59 | 13.6 | 13.7 | 13.5 | 18.2 |
| 60–64 | 8.3 | 8.9 | 8.9 | 10.1 |
| 65–69 | 15.4 | 14.5 | 15.0 | 12.7 |
| 70–79 | 33.0 | 32.1 | 32.1 | 24.9 |
| ⩾80 | 14.6 | 14.8 | 15.6 | 9.7 |
| *Sex* | | | | |
| Women | 48.3 | 48.1 | 48.0 | 48.3 |
| Men | 51.7 | 51.9 | 52.0 | 51.7 |
| *Race* | | | | |
| White | 64.0 | 64.1 | 65.4 | 52.9 |
| African American | 30.7 | 30.8 | 29.8 | 41.7 |
| Native American | 1.6 | 1.6 | 1.4 | 1.6 |
| Asian | 2.8 | 2.7 | 2.6 | 3.1 |
| Other | 0.8 | 0.9 | 0.8 | 0.7 |
| *ESRD primary cause* | | | | |
| Diabetes | 46.9 | 47.5 | 48.4 | 38.6 |
| Hypertension | 29.8 | 29.7 | 29.4 | 29.9 |
| Glomerulonephritis/ cystic kidney disease | 10.4 | 9.6 | 9.2 | 17.4 |
| Other | 12.9 | 13.2 | 13.1 | 14.1 |
| *Comorbid conditions* | | | | |
| ASHD | 51.5 | 52.2 | 53.8 | 41.2 |
| CHF | 54.3 | 55.0 | 55.5 | 44.3 |
| CVA/TIA | 24.2 | 25.1 | 25.6 | 18.4 |
| PVD | 44.4 | 44.5 | 45.6 | 38.0 |
| Other cardiac[b] | 35.3 | 34.9 | 36.9 | 33.1 |
| COPD | 20.5 | 21.2 | 22.2 | 16.1 |
| GI | 10.8 | 10.7 | 10.7 | 9.9 |
| Liver disease | 5.3 | 6.8 | 7.1 | 6.8 |
| Dysrhythmia | 30.6 | 31.7 | 32.3 | 26.2 |
| Cancer | 12.2 | 12.2 | 12.8 | 9.4 |
| Diabetes | 63.6 | 65.3 | 67.2 | 53.4 |

Abbreviations: ASHD, atherosclerotic heart disease; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; CVA/TIA, cerebrovascular accident/transient ischemic attack; ESRD, end-stage renal disease; GI, gastrointestinal bleeding; PVD, peripheral vascular disease.
[a]Values are percents unless otherwise specified.
[b]Includes pericarditis, endocarditis, myocarditis, other complications of heart disease, heart transplant, heart valve replacement, and cardiac devices.

### Calculation of comorbidity score

The coefficient estimates and their *P*-values for all variables from the Cox proportional regression model for the 2000

incident dialysis cohort with mortality as end point are displayed in Table 2. Coefficients for all comorbid conditions are positive, meaning that all conditions were associated with shorter survival and are significant predictors. Weights were assigned to each comorbid condition as follows: a weight of 1 was assigned to ASHD and diabetes; 2 to CVA/TIA, PVD, COPD, gastro-intestinal bleeding (GI), dysrhythmia, other cardiac disease, liver disease, and cancer; and 3 to CHF. The comorbidity score for each patient was defined as the sum of the weights based on presence or absence of the conditions. For example, for a patient with diabetes and CHF, and no other comorbid conditions, the comorbidity score would be 1 (diabetes) + 3 (CHF) = 4. For a patient with ASHD, PVD, GI bleeding, and CHF, the comorbidity score would be 1 (ASHD) + 2 (PVD) + 2 (GI) + 3 (CHF) = 8. The comorbidity score distribution is displayed in Table 3 and Figure 1 (1999 cohort only; results were similar for the other cohorts). As with individual comorbid conditions,

**Table 3 | Comorbidity score distribution for the four cohorts: 1999, 2000, 2001 incident dialysis patients and 2000 prevalent dialysis patients**

| | Cohort | | | |
|---|---|---|---|---|
| | 1999 Incident | 2000 Incident | 2001 Incident | 2000 Prevalent |
| *Score* | | | | |
| Mean | 6.4 | 6.6 | 6.7 | 5.4 |
| Median | 6.0 | 6.0 | 7.0 | 5.0 |
| s.d. | 4.1 | 4.1 | 4.1 | 4.1 |
| Minimum | 0.0 | 0.0 | 0.0 | 0.0 |
| Maximum | 19.0 | 21.0 | 21.0 | 21.0 |
| *Score groups, %* | | | | |
| 0 | 6.47 | 6.48 | 5.65 | 12.74 |
| 1–2 | 13.26 | 12.83 | 12.42 | 16.53 |
| 3–4 | 16.31 | 15.44 | 15.22 | 17.27 |
| 5–7 | 24.74 | 25.05 | 24.50 | 23.07 |
| 8–10 | 20.41 | 20.13 | 21.27 | 16.72 |
| 11–15 | 17.79 | 18.97 | 19.62 | 12.96 |
| >15 | 1.02 | 1.08 | 1.32 | 0.72 |

**Table 2 | Parameter estimates from a Cox model for survival analysis using the 2000 US incident dialysis cohort**

| Variable | Coefficient estimates | s.e. | Relative risk (95% CI) | *P*-value | Weight[a] |
|---|---|---|---|---|---|
| *Age group, years* | | | | | |
| 0–19 | −1.1560 | 0.2248 | 0.315 (0.203–0.489) | <0.0001 | |
| 20–29 | −1.5569 | 0.0979 | 0.211 (0.174–0.255) | <0.0001 | |
| 30–39 | −1.1940 | 0.0535 | 0.303 (0.273–0.336) | <0.0001 | |
| 40–49 | −0.9646 | 0.0352 | 0.381 (0.356–0.408) | <0.0001 | |
| 50–59 | −0.8238 | 0.0280 | 0.439 (0.415–0.464) | <0.0001 | |
| 60–64 | −0.6663 | 0.0303 | 0.514 (0.484–0.545) | <0.0001 | |
| 65–69 | −0.5770 | 0.0248 | 0.562 (0.535–0.590) | <0.0001 | |
| 70–79 | −0.3712 | 0.0196 | 0.690 (0.664–0.717) | <0.0001 | |
| ⩾80[b] | 0.0000 | 0.0000 | 1.000 | NA | |
| *Sex* | | | | | |
| Men[b] | 0.0000 | 0.0000 | 1.000 | NA | |
| Women | −0.0137 | 0.0144 | 0.986 (0.959–1.015) | 0.3421 | |
| *Race* | | | | | |
| White[b] | 0.0000 | 0.0000 | 1.000 | NA | |
| Asian/Pacific Islander | −0.3242 | 0.0482 | 0.723 (0.658–0.795) | <0.0001 | |
| African American | −0.2141 | 0.0170 | 0.807 (0.781–0.835) | <0.0001 | |
| Native American | −0.2098 | 0.0616 | 0.811 (0.719–0.915) | 0.0007 | |
| Other | −0.1945 | 0.0778 | 0.823 (0.707–0.959) | 0.0124 | |
| *ESRD primary cause* | | | | | |
| Diabetes | 0.2900 | 0.0322 | 1.336 (1.255–1.423) | <0.0001 | 3 |
| Hypertension | 0.1862 | 0.0303 | 1.205 (1.135–1.278) | <0.0001 | 2 |
| GN/cystic kidney disease[b] | 0.0000 | 0.0000 | 1.000 | NA | 0 |
| Other | 0.3463 | 0.0334 | 1.414 (1.324–1.510) | <0.0001 | 3 |
| *Comorbid conditions* | | | | | |
| ASHD | 0.0616 | 0.0165 | 1.064 (1.030–1.098) | 0.0002 | 1 |
| CHF | 0.2950 | 0.0164 | 1.343 (1.301–1.387) | <0.0001 | 3 |
| CVA/TIA | 0.1881 | 0.0159 | 1.207 (1.170–1.245) | <0.0001 | 2 |
| PVD | 0.1990 | 0.0149 | 1.220 (1.185–1.256) | <0.0001 | 2 |
| Other cardiac | 0.1559 | 0.0154 | 1.169 (1.134–1.205) | <0.0001 | 2 |
| COPD | 0.2333 | 0.0167 | 1.263 (1.222–1.305) | <0.0001 | 2 |
| GI | 0.1794 | 0.0216 | 1.196 (1.147–1.248) | <0.0001 | 2 |
| Liver disease | 0.1885 | 0.0270 | 1.207 (1.145–1.273) | <0.0001 | 2 |
| Dysrhythmia | 0.2090 | 0.0158 | 1.232 (1.195–1.271) | <0.0001 | 2 |
| Cancer | 0.2020 | 0.0205 | 1.224 (1.176–1.274) | <0.0001 | 2 |
| Diabetes | 0.1250 | 0.0205 | 1.133 (1.088–1.180) | <0.0001 | 1 |

Abbreviations: ASHD, atherosclerotic heart disease; CI, confidence interval; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; CVA/TIA, cerebrovascular accident/transient ischemic attack; ESRD, end-stage renal disease; GI, gastrointestinal bleeding; GN, glomerulonephritis; PVD, peripheral vascular disease.
[a]Weight assigned to ESRD primary cause and comorbid conditions: 0 for coefficient <0.05, 1 for 0.05 to <0.15, 2 for 0.15 to <0.25, and 3 for ⩾0.25.
[b]Reference group.

comorbidity scores increased over time for the incident cohorts ($P<0.0001$), and scores were lower for prevalent than for incident patients ($P<0.0001$).

Figure 2 presents survival curves for the patient groups with comorbidity scores in the intervals $\leqslant 3$, 4–6, 7–9, and $\geqslant 10$ for each cohort. Patients in lower-score groups survived longer. Log rank test shows that the four curves are different ($P<0.0001$ for all cohorts). The relative risks for the 4–6, 7–9, and $\geqslant 10$ groups relative to $\leqslant 3$ group are 1.533, 2.117, and 3.138, respectively, for the 1999 cohort, and similar for the other cohorts. These numbers indicate that the comorbidity score is an important predictor for survival.

We also defined the index including comorbid conditions and ESRD primary cause based on the same model. We assigned a weight to each ESRD primary cause in the same way as for comorbid conditions. A weight of 2 was therefore assigned to hypertension, and 3 to diabetes and other causes (Table 2). Glomerulonephritis/cystic kidney disease was used as the reference in the Cox model, and assigned a weight of 0. Scores for each patient are calculated in the same way as in the index without ESRD cause. For example, the comorbidity score for a patient with diabetes and CHF as comorbid conditions and diabetes as primary ESRD cause would be 1 (diabetes as comorbid condition) + 3 (CHF) + 3 (diabetes as primary ESRD cause) = 7.

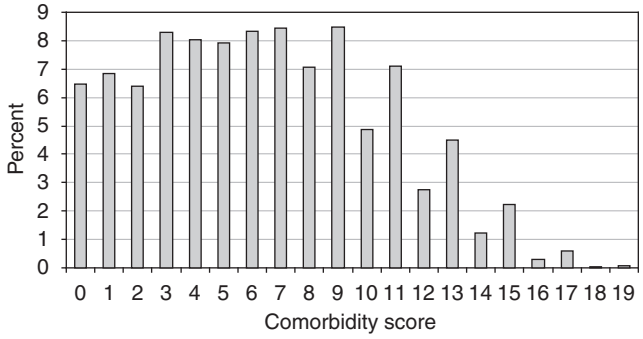


**Figure 1 | Percent of patients with each comorbidity score, 1999 cohort.**

### Comorbidity index validation

Table 4 displays the model fit statistics, predictive ability statistics, and parameter estimates for all variables in the model not used for the indexes, from a model with 11 individual comorbid conditions and models with two indexes for the 1999 incident cohort. A positive percent change value means that the corresponding statistic has a higher value from the comorbidity score model than from the individual comorbid condition model. Otherwise, the statistic has a lower value from the comorbidity score model. Although some changes are statistically significant because of the large sample size, Table 4 shows that the model with the new comorbidity index without ESRD primary cause was almost identical in value to the model with 11 individual comorbid conditions regarding model fit, model predictive ability, comorbidity predictive ability, and relative risk estimates for all variables. Except for diabetes as ESRD primary cause and age groups 0–19, 50–59, and 60–64 years, changes in relative risk estimates across models for all variables are less than 1%. The importance of the magnitude of change is subjective, but

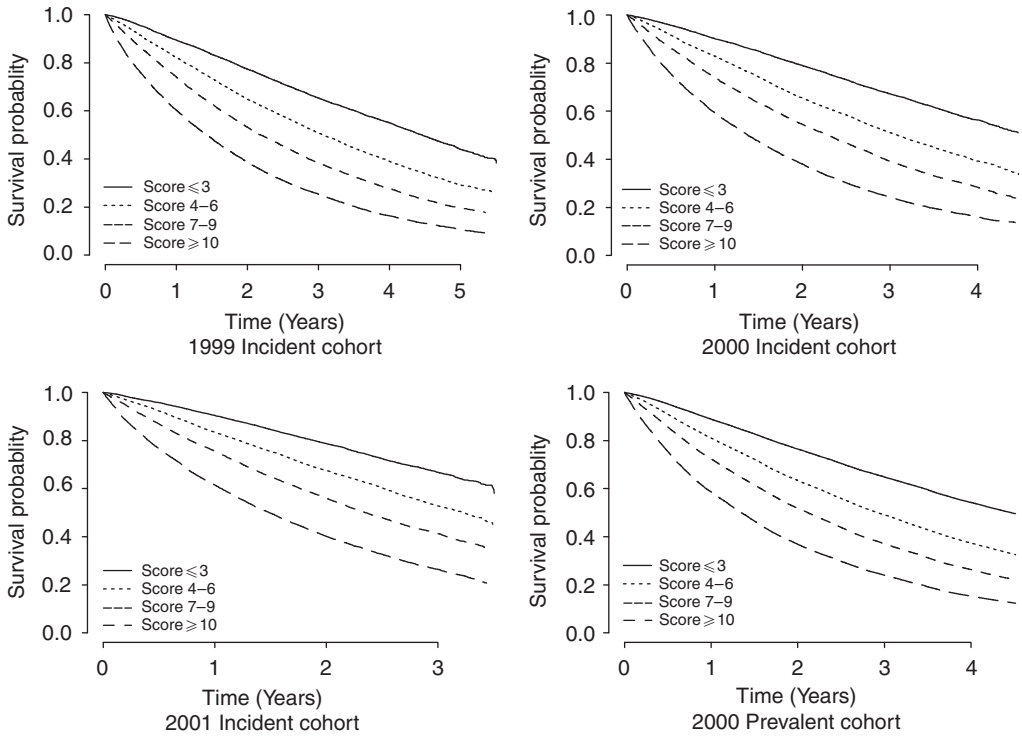


**Figure 2 | Survival curves for patient groups with comorbidity scores in the intervals $\leqslant 3$, 4–6, 7–9, and $\geqslant 10$ for each cohort.**

**Table 4 | Comorbidity index validation: comparing the index with the individual comorbid conditions as a set in a Cox proportional hazard model, 1999 incident dialysis patients, a validation cohort**

| Parameter | 11 Individual conditions | Score excluding ESRD cause | Percent change (*P*-value)[a] | Score including ESRD cause | Percent change (*P*-value)[a] |
|---|---|---|---|---|---|
| *Model fit statistics* | | | | | |
| −2 LOG L | 438141 | 438223 | 0.02 | 438236 | 0.02 |
| AIC | 438195 | 438257 | 0.01 | 438264 | 0.02 |
| SBC | 438412 | 438394 | 0.00 | 438376 | −0.01 |
| *Model predictive ability* | | | | | |
| *c*-Statistic | 0.6695 | 0.6687 | −0.11 (<0.001) | 0.6693 | −0.02 (<0.001) |
| *Predictive ability of comorbid conditions* | | | | | |
| Marginal $R^2$ | 0.1119 | 0.1072 | −4.44 | 0.1096 | −2.09 |
| Partial $R^2$ | 0.0543 | 0.0526 | −3.17 | 0.0623 | 14.81 |
| *Relative risks for other variables* | | | | | |
| Age group, years | | | | | |
| 0–19 | 0.220 | 0.215 | −1.95 (0.033) | 0.212 | −3.60 (0.010) |
| 20–29 | 0.196 | 0.195 | −0.71 (0.633) | 0.194 | −0.94 (0.210) |
| 30–39 | 0.285 | 0.286 | 0.14 (0.377) | 0.288 | 0.93 (0.123) |
| 40–49 | 0.329 | 0.331 | 0.43 (0.123) | 0.335 | 1.68 (0.007) |
| 50–59 | 0.404 | 0.409 | 1.11 (<0.001) | 0.416 | 2.96 (<0.001) |
| 60–64 | 0.478 | 0.483 | 1.10 (0.010) | 0.492 | 3.01 (<0.001) |
| 65–69 | 0.545 | 0.550 | 0.89 (<0.001) | 0.558 | 2.40 (<0.001) |
| 70–79 | 0.681 | 0.685 | 0.60 (0.013) | 0.692 | 1.52 (<0.001) |
| ⩾80[b] | 1.000 | 1.000 | NA | 1.000 | NA |
| Sex | | | | | |
| Men[b] | 1.000 | 1.000 | NA | 1.000 | NA |
| Women | 1.001 | 1.000 | −0.06 (0.463) | 1.006 | 0.47 (0.040) |
| Race | | | | | |
| White[b] | 1.000 | 1.000 | NA | 1.000 | NA |
| Asian/Pacific islander | 0.702 | 0.704 | 0.25 (0.797) | 0.707 | 0.70 (0.140) |
| African American | 0.786 | 0.790 | 0.41 (0.007) | 0.792 | 0.75 (0.777) |
| Native American | 0.864 | 0.864 | −0.07 (0.413) | 0.875 | 1.24 (<0.001) |
| Other | 1.145 | 1.142 | −0.25 (0.530) | 1.152 | 0.57 (0.287) |
| ESRD primary cause | | | | | |
| Diabetes | 1.292 | 1.361 | 5.27 (<0.001) | | |
| Hypertension | 1.193 | 1.199 | 0.52 (0.133) | | |
| Glomerulonephritis/cystic renal disease[b] | 1.000 | 1.000 | NA | | |
| Other | 1.267 | 1.275 | 0.64 (0.034) | | |

Abbreviations: AIC, Akaike's information criterion; ESRD, end-stage renal disease; SBC, Schwartz Bayesian criterion.
[a]Percent change of parameter estimate from the model with comorbidity score relative to estimate from model with individual comorbid conditions. *P*-value from the bootstrap method.
[b]Reference group.

we consider less than 5% an acceptable change. Findings are similar for the 2001 incident cohort and 2000 prevalent cohort (data not shown). Standard errors for all parameter estimates were also compared across models (data not shown). They are almost identical. The comorbidity index including ESRD primary cause also works well, but not as well as the version excluding it. The new index also works well for hospitalization and medical cost analyses (Table 5). The percent change in relative risk and relative cost are less than two for most variables. The differences in $R^2$ are small, but the percent changes are large because of the small values of $R^2$.

Regarding model fit (−2 log likelihood, AIC [Akaike's information criterion], and SBC [Schwartz Bayesian criterion]), model predictive ability (*c*-statistic), and the predictive ability of comorbid conditions (marginal and partial $R^2$), the 11 comorbid conditions in the new comorbidity index and the 17 comorbid conditions in the CCI are almost identical (Table 6, 1999 incident cohort). Regarding predictive ability and the effect on relative risk estimate, the new index works as well as the 11 individual comorbid conditions (Table 6). The model fit and the predictive ability of the model with CCI, which includes CCI and the other variables, is similar to the model with the 17 individual comorbid conditions. However, the $R^2$, which measure the predictive ability of CCI alone, are much lower than the $R^2$ for the 17 individual comorbid conditions (marginal $R^2$ 0.0642 versus 0.0981, partial $R^2$ 0.0381 versus 0.0549). Also, the effect of CCI on relative risk estimates for variables not in the index is high (see the 'Percent Change' column for 'Estimates of relative

**Table 5 | Parameters from the model compared with individual comorbid conditions and with comorbidity score for hospitalization and medical cost, 2000 incident dialysis patients**

| Parameter | First hospitalization | | | Multiple hospitalizations | | |
|---|---|---|---|---|---|---|
| | Model with individual comorbidities | Model with comorbidity score | Percent change (*P*-value)[a] | Model with individual comorbidities | Model with comorbidity score | Percent change (*P*-value)[a] |
| $R^2$ | | | | | | |
| Marginal | 0.0628 | 0.0563 | −10.32 | | | |
| Partial | 0.0546 | 0.0491 | −10.07 | | | |
| *Estimates for variables not in the index*[b] | | | | | | |
| Age group, years | | | | | | |
| 0–19 | 1.772 | 1.795 | 1.32 | 1.894 | 1.919 | 1.32 |
| 20–29 | 1.081 | 1.094 | 1.20 | 1.361 | 1.381 | 1.51 |
| 30–39 | 1.072 | 1.080 | 0.80 | 1.295 | 1.315 | 1.52 |
| 40–49 | 0.970 | 0.981 | 1.16 | 1.231 | 1.255 | 1.98 |
| 50–59 | 0.958 | 0.964 | 0.62 | 1.029 | 1.044 | 1.45 |
| 60–64 | 0.882 | 0.891 | 1.03 | 0.942 | 0.954 | 1.25 |
| 65–69 | 0.880 | 0.891 | 1.27 | 0.947 | 0.961 | 1.49 |
| 70–79 | 0.897 | 0.902 | 0.61 | 0.927 | 0.934 | 0.71 |
| ⩾80[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Sex | | | | | | |
| Men[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Women | 1.153 | 1.160 | 0.67 | 1.138 | 1.144 | 0.57 |
| Race | | | | | | |
| White[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Asian/Pacific Islander | 0.855 | 0.860 | 0.51 | 0.854 | 0.856 | 0.28 |
| African American | 0.992 | 0.990 | −0.16 | 1.055 | 1.051 | −0.37 |
| Native American | 0.984 | 0.981 | −0.28 | 0.977 | 0.969 | −0.78 |
| Other | 1.034 | 1.050 | 1.54 | 1.100 | 1.102 | 0.21 |
| ESRD primary cause | | | | | | |
| Diabetes | 1.108 | 1.150 | 3.84 | 1.184 | 1.208 | 2.00 |
| Hypertension | 1.026 | 1.032 | 0.62 | 1.080 | 1.084 | 0.40 |
| Glomerulonephritis/ cystic renal disease[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Other | 1.134 | 1.132 | −0.18 | 1.143 | 1.144 | 0.04 |
| | Hospitalization days | | | Medical cost | | |
| $R^2$ | | | | | | |
| Marginal | 0.0300 | 0.0128 | −57.39 | 0.1033 | 0.0903 | −12.56 |
| Partial | 0.0198 | 0.0159 | −19.94 | 0.0745 | 0.0686 | −7.93 |
| *Estimates for variables not in model*[b] | | | | | | |
| Age group, years | 1.000 | 1.000 | 0.00 | 1.000 | 1.000 | 0.00 |
| 0–19 | 1.780 | 1.796 | 0.92 | 0.896 | 0.904 | 0.82 |
| 20–29 | 1.048 | 1.066 | 1.68 | 0.880 | 0.891 | 1.28 |
| 30–39 | 1.059 | 1.077 | 1.67 | 0.910 | 0.921 | 1.23 |
| 40–49 | 1.098 | 1.121 | 2.09 | 0.940 | 0.951 | 1.21 |
| 50–59 | 1.002 | 1.018 | 1.62 | 0.956 | 0.963 | 0.75 |
| 60–64 | 0.949 | 0.964 | 1.53 | 0.958 | 0.965 | 0.73 |
| 65–69 | 0.960 | 0.974 | 1.53 | 0.948 | 0.955 | 0.70 |
| 70–79 | 0.938 | 0.946 | 0.82 | 0.970 | 0.973 | 0.37 |
| ⩾80[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Sex | 1.000 | 1.000 | 0.00 | 1.000 | 1.000 | 0.00 |
| Men[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Women | 1.161 | 1.170 | 0.72 | 1.073 | 1.076 | 0.29 |
| Race | | | | | | |
| White[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Asian/Pacific Islander | 0.900 | 0.904 | 0.45 | 1.026 | 1.031 | 0.49 |
| African American | 1.082 | 1.082 | −0.04 | 1.100 | 1.107 | 0.58 |

Table 5 continued on the following page

**Table 5 | Continued**

| Parameter | First hospitalization | | | Multiple hospitalizations | | |
|---|---|---|---|---|---|---|
| | Model with individual comorbidities | Model with comorbidity score | Percent change (*P*-value)[a] | Model with individual comorbidities | Model with comorbidity score | Percent change (*P*-value)[a] |
| Native American | 0.921 | 0.919 | −0.25 | 0.957 | 0.959 | 0.19 |
| Other | 1.071 | 1.078 | 0.59 | 1.037 | 1.043 | 0.56 |
| ESRD primary cause | | | | | | |
| Diabetes | 1.231 | 1.303 | 5.80 | 1.109 | 1.150 | 3.69 |
| Hypertension | 1.060 | 1.070 | 0.87 | 1.012 | 1.014 | 0.27 |
| Glomerulonephritis/ cystic renal disease[c] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Other | 1.188 | 1.193 | 0.43 | 1.020 | 1.027 | 0.72 |

Abbreviations: ESRD, end-stage renal disease.
[a]Percent change of parameter estimate from the model with comorbidity score relative to estimate from model with individual comorbid conditions.
[b]The parameters corresponding to these variables are relative risk for hospitalization and times of increment for medical cost.
[c]Reference group.

**Table 6 | Comparison between new comorbidity score and Charlson Comorbidity Index, 1999 incident dialysis patients, a validation cohort**

| | Model | | | Model | | |
|---|---|---|---|---|---|---|
| | 11 Individual comorbid conditions | New index score | Percent change (*P*-value)[a] | 17 Individual comorbid conditions | Charlson score | Percent change (*P*-value)[a] |
| *Model fit statistics* | | | | | | |
| −2 LOG L | 438186 | 438247 | 0.01 | 438024 | 438804 | 0.18 |
| AIC | 438240 | 438281 | 0.01 | 438090 | 438838 | 0.17 |
| SBC | 438456 | 438418 | −0.01 | 438355 | 438975 | 0.14 |
| *Model predictive ability* | | | | | | |
| *c*-Statistic | 0.6691 | 0.6685 | −0.09 | 0.6709 | 0.6623 | −1.30 |
| *Predictive ability of comorbid conditions* | | | | | | |
| Marginal $R^2$ | 0.1007 | 0.0957 | −5.22 | 0.0981 | 0.0642 | −52.85 |
| Partial $R^2$ | 0.0524 | 0.0511 | −2.56 | 0.0549 | 0.0381 | −44.00 |
| *Estimates of relative risks for variables not in the Index* | | | | | | |
| Age group, years | | | | | | |
| 0–19 | 0.201 | 0.200 | −0.65 | 0.186 | 0.177 | −4.69 |
| 20–29 | 0.182 | 0.181 | −0.76 | 0.178 | 0.166 | −6.95 |
| 30–39 | 0.269 | 0.269 | −0.14 | 0.266 | 0.244 | −8.43 |
| 40–49 | 0.316 | 0.317 | 0.31 | 0.312 | 0.289 | −7.30 |
| 50–59 | 0.398 | 0.400 | 0.63 | 0.393 | 0.369 | −5.92 |
| 60–64 | 0.474 | 0.478 | 0.68 | 0.468 | 0.443 | −5.34 |
| 65–69 | 0.543 | 0.547 | 0.69 | 0.541 | 0.521 | −3.73 |
| 70–79 | 0.682 | 0.684 | 0.30 | 0.678 | 0.663 | −2.18 |
| ⩾80[b] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Sex | | | | | | |
| Men[b] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Women | 0.991 | 0.991 | 0.05 | 0.994 | 0.992 | −0.26 |
| Race | | | | | | |
| White[b] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Asian/Pacific Islander | 0.691 | 0.690 | −0.15 | 0.699 | 0.684 | −2.14 |
| African American | 0.776 | 0.776 | −0.07 | 0.765 | 0.765 | −0.03 |
| Native American | 0.867 | 0.863 | −0.44 | 0.861 | 0.831 | −3.49 |
| Other | 1.138 | 1.136 | −0.12 | 1.134 | 1.116 | −1.64 |
| ESRD primary cause | | | | | | |
| Diabetes | 1.367 | 1.409 | 3.08 | 1.323 | 1.223 | −7.59 |
| Hypertension | 1.224 | 1.231 | 0.57 | 1.215 | 1.242 | 2.27 |
| Glomerulonephritis/ cystic renal disease[b] | 1.000 | 1.000 | NA | 1.000 | 1.000 | NA |
| Other | 1.275 | 1.280 | 0.39 | 1.262 | 1.258 | −0.31 |

Abbreviations: AIC, Akaike's information criterion; ESRD, end-stage renal disease; SBC, Schwartz Bayesian criterion.
[a]Percent change of parameter estimate from the model with comorbidity score relative to estimate from model with individual comorbidities.
[b]Reference group.

risks for variables not in the Index' in Table 6). Results for the 2000 and 2001 incident and 2000 prevalent cohorts are almost identical.

## DISCUSSION

A useful comorbidity index should be simple and validated for specific use as a substitute for individual comorbid conditions. In this study, a new comorbidity index, including and excluding ESRD primary cause, was developed for survival analysis of dialysis patients and validated as a predictor and an adjuster for replacing the individual comorbid conditions. The new index performs well in analyses for mortality, hospitalization, and medical costs in both incident and prevalent cohorts. It also outperforms the CCI in both predictive ability and inference. This comorbidity index can be used for the comparison of comorbidity burden and comorbidity burden change over a fixed period, reflecting quality of care, among dialysis facilities; standard mortality ratio calculation; and outcome analysis that requires adjustment for comorbid conditions. Use of the index will make the comparisons easier, the computations simpler, and the estimates of effects more stable.

The ESRD Medical Evidence Report (form CMS-2728) lists eight categories of primary ESRD causes: diabetes, hypertension, glomerulonephritis, cystic renal disease, other urologic disease, unknown cause, other cause, and missing. We began by using these eight categories. Because the effect of cystic renal disease is similar to that of glomerulonephritis, we combined these two categories into one. For the same reason, we combined 'other urologic disease,' 'unknown cause,' and 'other cause' into a category called 'other.' Because the missing-cause group was too small for a significant comparison with the other causes, we grouped it with 'other.'

The effect of diabetes looks smaller than expected because diabetes was the primary cause of ESRD for most of the diabetic patients, and diabetes as ESRD primary cause, which has a high score, is also in the model. Thus, most diabetes patients actually receive a high score, 4 (3 from diabetes as primary cause of ESRD and 1 from diabetes as a comorbid condition), for diabetes.

Age was not included in our comorbidity index because the effects of age and comorbid conditions were at very different levels (see coefficient estimates in Table 2), and this difference makes assigning definite scores difficult. Also, age affects mortality, hospitalization, and medical costs differently. An index including age developed for mortality would be difficult to generalize to hospitalization and medical cost analyses. But, because age, comorbid conditions, and ESRD primary cause are highly correlated, age and ESRD primary cause should be included in the model when the comorbidity index excluding ESRD primary cause is used, and age should be included when the index including ESRD primary cause is used.

Patients who returned to dialysis after failed transplants were included in the validation using the 2000 prevalent cohort. Neither adding previous transplant as an indicator in the model nor excluding those patients from the model changes the validation result. This might be because these patients were not considerably different from the other patients, or because this is a small subgroup of the prevalent dialysis cohort (about 8%), or both.

Another comorbidity index, the Index of Coexistent Disease (ICED), also works better than the CCI in analyses of ESRD patients.[16] The ICED includes a patient physical functioning component, which might improve its performance. However, the ICED is difficult to use.[17,18] The score given is based not only on the presence or absence of the specific diseases and physical functions, but also on the severity of diseases or physical impairment. Assigning scores at the start of a study requires trained personnel. This is impractical for many studies, especially for studies based on administrative data. A useful comorbidity index should work well, be easy to use, and thus become widely used.

The new comorbidity index proposed in this study was designated for analyses using administrative data. The accuracy of the comorbid conditions documented in the data may affect the performance of the index. We provide the International Classifications of Diseases, Ninth Edition, Clinical Modification (ICD-9 CM) codes and Current Procedural Terminology codes used for defining comorbid conditions from claims in the Appendix. If different codes are used for comorbid condition definitions, the scores for the conditions defined in this study may not be applicable. Information on the comorbid conditions included in this index was collected from both the Medical Evidence Report and claims over a 6-month entry period. The index cannot be used for studies that evaluate outcomes in the entry period. The index was developed based on the national cohort of Medicare dialysis patients and is not validated for non-Medicare patients. It should be used with caution when non-Medicare patients are included in the study. Both hemodialysis and peritoneal dialysis cohorts were included in score development and evaluation. Separate validations were performed for hemodialysis and peritoneal dialysis patients, with good results. However, because of the small sample size of the peritoneal dialysis cohort, validation may not be reliable, and the score should be applied with caution to peritoneal dialysis patients.

## MATERIALS AND METHODS

### Study population and data source

This study included all US patients who developed ESRD in 1999, 2000, and 2001; survived at least 9 months; used dialysis as renal replacement therapy; and had Medicare as primary payer with Part A and Part B coverage at day 91 of ESRD. The first 90 days were excluded because Medicare information for that period was incomplete for many patients; months 4–9 were used to define patient baseline characteristics. Patients who underwent transplant, were lost to follow-up, or changed primary payer during the baseline period were excluded. Patients who were HIV positive, had AIDS during the baseline period, or died of AIDS in the follow-up period were also excluded. Patients were followed from the first day after the baseline period until the date of death, transplant, loss to

follow-up, change of primary payer, or end of 2005. The 2000 incident cohort was used as the training population and the 1999 and 2001 incident cohorts, separately, as the testing populations. The 2000 point prevalent dialysis patients (on dialysis on 1 January 2000) who survived at least 6 months in 2000 and had Medicare as primary payer with Part A and Part B coverage on 1 January 2000, were also used as a testing population. The first 6 months (1 January 2000, to 30 June 2000) were used to define patient baseline characteristics, and the same exclusion criteria were applied.

Data were from the CMS Renal Management Information System (REMIS) and CMS Standard Analytical Files. The REMIS database includes information from the CMS Medicare Enrollment Database, the United Network for Organ Sharing transplant database, the ESRD Medical Evidence Report (form CMS-2728), and the ESRD Death Notification (form CMS-2746). The Standard Analytical Files include Medicare Part A institutional claims and Part B physician/supplier claims. Patient demographic information, including age, sex, race (white, African American, Native American, Asian, and other); ESRD primary cause (diabetes, hypertension, glomerulonephritis/cystic renal disease, and other); and date of death were obtained from the REMIS database. Patient comorbid conditions, hospitalization dates, and medical costs were derived from CMS claim files.

### Comorbidity and outcomes

The new comorbidity index was defined by 11 comorbid conditions used by the USRDS,[15] most of which are listed in the Medical Evidence Report. The conditions are diabetes, ASHD, CHF, PVD, CVA/TIA, dysrhythmia, other cardiac diseases (including pericarditis, endocarditis, myocarditis, other complications of heart disease, heart transplant, heart valve replacement, and cardiac devices), cancer, liver disease, GI bleeding, and COPD. Liver disease and GI bleeding are not included on the Medical Evidence Report. A patient is defined as having a comorbid condition if a code for it can be found on the Medical Evidence Report or in claims during the 6-month baseline period. Comorbid conditions from claims were defined using ICD-9 CM codes. The ICD-9 CM diagnosis codes and V codes defining the conditions were designated by the USRDS,[15] and are listed in the Appendix. The method used to define comorbid conditions from claims has been previously described.[19] Claims over 6 months were used for defining comorbid conditions because the Medical Evidence Report may underestimate patient comorbid conditions and the most recent claims supply supplemental and updated information, especially for prevalent patients.

The outcome used to develop and test the index was death. We also examined whether the index could be used for hospitalization (first hospitalization, multiple hospitalizations, and hospital days) and medical cost analyses. Data on death were obtained from the ESRD Death Notification (form CMS-2746). Hospitalization data in the follow-up period were derived from the Medicare Part A inpatient claims. Per-patient per-month Medicare allowable costs in the follow-up period were calculated from both Part A and Part B claims. Institutional Medicare allowable expenditures include Medicare payment, coinsurance, deductibles, and any payment provided by a payer other than Medicare. Dollar amounts during the follow-up period were assembled and were linearly prorated for claims that spanned the beginning or end of the follow-up period. Expenditures in dollars were summed for each patient and divided by the number of follow-up months to arrive at the per-patient per-month expenditure.

### Development and validation of the index

A Cox proportional regression model was fitted for the 2000 incident dialysis cohort with time to death as the response variable and age, race, sex, primary ESRD cause, and the 11 comorbid conditions as covariates. Categorical age was used with nine age groups (ages 0–19, 20–29, 30–39, 40–49, 50–59, 60–64, 65–69, 70–79, and $\geqslant$80 years). A Cox proportional regression model can be expressed as

$$\lambda(\mathrm{t}|x_1x_2,\ldots,x_k) = \lambda_0(t)\exp(\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k),$$

where $x_1$, $x_2$, …, and $x_k$ are covariates; $\lambda(t|x_1, x_2, \ldots, x_k)$ is the hazard at time $t$ given $x_1$, $x_2$, …, $x_k$; $\lambda_0(t)$ is the baseline hazard at time t corresponding to $x_1 = 0$, $x_2 = 0$, …, $x_k = 0$; $\beta_1$, $\beta_2$,…, and $\beta_k$, are coefficients of $x_1$, $x_2$, …, and $x_k$, and $\exp(\beta_1)$, $\exp(\beta_2)$, …, and $\exp(\beta_k)$ are the relative risk of $x_1$, $x_2$, …, and $x_k$. Another expression of a Cox proportional regression model is

$$\mathrm{S}(\mathrm{t}|x_1,x_2,\ldots,x_k) = [\mathrm{S}_0(\mathrm{t})]\exp(\beta_1x_1+\beta_2x_2+\ldots+\beta_\mathrm{k}x_\mathrm{k}),$$

where $S(t|x_1, x_2, \ldots, x_k)$ is the survival probability at time t given $x_1$, $x_2$, …, $x_k$ and $S_0(t)$ is the baseline survival probability at time t corresponding to $x_1 = 0$, $x_2 = 0$, …, $x_k = 0$.

After the coefficients $\beta_1$, $\beta_2$,…, and $\beta_k$, were estimated using the 2000 incident cohort, an integer weight was assigned to each comorbid condition as follows: 0 for coefficient $<0.05$, 1 for 0.05 to $<0.15$, 2 for 0.15 to $<0.25$, and 3 for $\geqslant 0.25$. To maintain the linear relationship of comorbidity score and log hazard in the Cox model, coefficients, not relative risks, were used to assign weights. The comorbidity score for each patient was the sum of the weights based on the presence or absence of the conditions. Specifically, weights are $w_1$, $w_2$, …, $w_{11}$, for the 11 conditions, respectively; for a patient with conditions 1 and 5 only, the comorbidity score would be $w_1 + w_5$. For a patient with conditions 3, 5, 8, and 9, the comorbidity score would be $w_3 + w_5 + w_8 + w_9$. The comorbidity score is used as a continuous covariate in models for outcome analyses.

The index was validated using the 1999 and 2001 US incident dialysis cohorts and the 2000 US prevalent dialysis cohort, separately. Model fit statistics ($-2$ times log likelihood, AIC and SBC), model predictive ability statistic (*c*-statistic for time-to-event data), and the parameter estimates for the variables not included in the score and their s.e. were compared between the model using comorbidity score and the model using individual comorbid conditions. The $-2$ times log likelihood can be used to compare model fit. A smaller value means a higher probability of the observed values being observed. Therefore, the smaller the $-2$ times log likelihood, the better the model. The *c*-statistic is used to measure to what extent predicted values from the model are concordant with observed values, that is, whether a pair of observed values occurs in the same order as their predicted values. It was used to measure the model predictive ability; the higher the *c*-statistic, the better the model. The predictive abilities of comorbidity scores and the individual comorbid conditions as a set were also compared using the Kent-O'Quigley $R^2$ as an indicator.[20] Similar to the $R^2$ for a linear regression model, the Kent-O'Quigley $R^2$ can be used to check how much of the data 'variation' can be explained by specific variables. The differences in *c*-statistic and estimates of variables not in the index between the model using individual comorbid conditions and the model using indexes were tested using the bootstrap method. Checking usability of the index for hospitalization and medical cost analyses was done similarly using the 2000 incident data. Cox proportional regression models, Andersen–Gill models, Poisson regression models, and linear regression models were used for first hospitalization, multiple hospitalizations, hospital

days, and medical cost analyses, respectively. The cost analysis was based on logged scale of per-patient per-month expenditures because of the skewness of expenditure data.

A comorbidity index including ESRD primary cause was also developed for survival analysis based on the same model, using the 2000 incident dialysis cohort, and validated using the 1999 and 2001 US incident dialysis cohorts and 2000 prevalent dialysis cohort in the same way as for the index without ESRD primary cause.

The new comorbidity index cannot be compared with the CCI directly because many of the 17 comorbid conditions the CCI uses cannot be derived from the Medical Evidence Report (we excluded renal disease because all patients had renal disease). The 17 comorbid conditions were defined from claims only. To make it comparable, we redefined our 11 comorbid conditions using claims only; however, this might lead to underestimates because the 6-month period might not be long enough to yield claims for all comorbid conditions. We then repeated the procedures described above and compared the results with results from models using the CCI. The 17 comorbid conditions used for the CCI were derived from patient Medicare Part A and Part B claims using the ICD-9 CM diagnosis codes given by Deyo.[21] We compared the changes in the statistics used for validating the new comorbidity index from the 11-individual-condition model to the new comorbidity index model and from the 17-individual-condition model to the CCI model for all four cohorts (1999, 2000, 2001 incident cohorts and 2000 prevalent cohort). SAS version 9.1 (SAS Institute, Cary, NC, USA) was used to perform all analyses in this study.

## DISCLOSURE

The data reported here have been supplied by the United States Renal Data System. This study was performed as a deliverable under Contract No. HHSN267200715002C (National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland). The authors have no conflict of interest with its subject matter.

## ACKNOWLEDGMENTS

We thank United States Renal Data System colleagues Beth Forrest, Shane Nygaard, and Nan Booth, for regulatory assistance, article preparation, and article editing, respectively. The data reported here have been supplied by the United States Renal Data System. This study was performed as a deliverable under Contract No. HHSN267200715002C (National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland).

## REFERENCES

1. Oniscu GC, Forsythe JL. Which index is optimal for assessment of baseline comorbidities in kidney transplant recipients? *Nat Clin Pract Nephrol* 2006; **2**: 74–75.
2. Beddhu S, Bruns FJ, Saul M *et al.* A simple comorbidity scale predicts clinical outcomes and costs in dialysis patients. *Am J Med* 2000; **108**: 609–613.
3. Fried L, Bernardini J, Piraino B. Charlson comorbidity index as a predictor of outcomes in incident peritoneal dialysis patients. *Am J Kidney Dis* 2001; **37**: 337–342.
4. Collins AJ, Weinhandl E, Snyder JJ *et al.* Comparison and survival of hemodialysis and peritoneal dialysis in the elderly. *Semin Dial* 2002; **15**: 98–102.
5. Wall MM, Li R. Comparison of multiple regression to two latent variable techniques for estimation and prediction. *Stat Med* 2003; **22**: 3671–3685.
6. Charlson ME, Pompei P, Ales KL *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; **40**: 373–383.
7. Khan I, Catto G, Edward N *et al.* Influence of coexisting disease on survival on renal-replacement therapy. *Lancet* 1993; **341**: 415–418.
8. Davies SJ, Russell L, Bryan J *et al.* Comorbidity, urea kinetics, and appetite in continuous ambulatory peritoneal dialysis patients: their interrelationship and prediction of survival. *Am J Kidney Dis* 1995; **26**: 353–361.
9. van Manen JG, Korevaar JC, Dekker FW *et al.* How to adjust for comorbidity in survival studies in ESRD patients: a comparison of different indices. *Am J Kidney Dis* 2002; **40**: 82–89.
10. Harrell F. *Regression Modeling Strategies With Applications To Linear Models, Logistic Regression, and Survival Analysis*. In Third Printing ed Springer: New York, 2001.
11. Hemmelgarn BR, Manns BJ, Quan H *et al.* Adapting the Charlson comorbidity index for use in patients with ESRD. *Am J Kidney Dis* 2003; **42**: 125–132.
12. St Peter WL, Liu J, Weinhandl ED *et al.* Linking centers for medicare & medicaid services data with prospective DCOR trial data: methods and data comparison results. *Hemodial Int* 2008; **12**: 480–491.
13. St Peter WL, Liu J, Weinhandl ED *et al.* A comparison of sevelamer and calcium-based phosphate binders on mortality, hospitalization, and morbidity in hemodialysis: a secondary analysis of the dialysis clinical outcomes revisited (DCOR) Randomized Trial Using Claims Data. *Am J Kidney Dis* 2008; **51**: 445–454.
14. Jaber BL, Finkelstein FO, Glickman JD *et al.* Scope and Design of the Following Rehabilitation, Economics and Everyday-Dialysis Outcome Measurements (FREEDOM) Study. *Am J Kidney Dis* 2008; **53**: 310–320. (in press).
15. US Renal Data System. *USRDS 2005 Annual Data Report: Atlas of end-stage renal disease in the United States*. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases: Bethesda, MD, 2005.
16. Miskulin DC, Martin AA, Brown R *et al.* Predicting 1 year mortality in an outpatient haemodialysis population: a comparison of comorbidity instruments. *Nephrol Dial Transplant* 2004; **19**: 413–420.
17. Athienites NV, Miskulin DC, Fernandez G *et al.* Comorbidity assessment in hemodialysis and peritoneal dialysis using the index of coexistent disease. *Semin Dial* 2000; **13**: 320–326.
18. Imamura K, McKinnon M, Middleton R *et al.* Reliability of a comorbidity measure: the Index of Co-Existent Disease (ICED). *J Clin Epidemiol* 1997; **50**: 1011–1016.
19. Hebert PL, Geiss LS, Tierney EF *et al.* Identifying persons with diabetes using Medicare claims data. *Am J Med Qual* 1999; **14**: 270–277.
20. Kent J, O'Quigley J. Measures of dependence for censored survival data. *Biometrika* 1988; **75**: 525–534.
21. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992; **45**: 613–619.

**Appendix**

**International classification of diseases, ninth revision, clinical modification codes (ICD-9-CM) used for defining the 13 comorbid conditions for the new comorbidity index**

| Comorbid condition | ICD-9-CM diagnosis codes | ICD-9-CM V codes |
|---|---|---|
| Atherosclerotic heart disease | 410–414 | V45.81; V45.82 |
| Congestive heart failure | 398.91; 422; 425; 428; 402.X1; 404.x1; 404.x3 | V42.1 |
| Cerebrovascular accident/ transient ischemic attack | 430–438 | |
| Peripheral vascular disease | 40–444; 447; 451–453; 557 | |
| Other cardiac | 420–421; 423–424; 429; 785.0–785.3 | V42.2;V43.3 |
| Chronic obstructive pulmonary disease | 491–494; 496; 510 | |
| Gastrointestinal bleeding | 456.0–456.2; 530.7; 531–534; 569.84; 569.85; 578 | |
| Liver disease | 570; 571; 572.1; 572.4; 573.1–573.3 | V42.7 |
| Dysrhythmia | 426–427 | V45.0; V53.3 |
| Cancer | 140–172; 174–208; 230–231; 233–234 | |
| Diabetes | 250; 357.2; 362.0x; 366.41 | |