J Am Soc Nephrol 14: 478–485, 2003

Adjustment for Comorbidity in Studies on Health Status in ESRD Patients: Which Comorbidity Index to Use?

JEANNETTE G. VAN MANEN,* JOHANNA C. KOREVAAR,* FRIEDO W. DEKKER,† ELISABETH W. BOESCHOTEN,‡§ PATRICK M. M. BOSSUYT,* and RAYMOND T. KREDIET,‡ for the NECOSAD-Study Group

**Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; †Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; ‡Department of Nephrology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; and §Dianet Dialysis Centres, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands*

Abstract. Health status can be an important outcome in studies on patients with end-stage renal disease (ESRD). In these studies, adjustment for prognostic factors, such as comorbidity, often has to be made. None of the comorbidity indices that are commonly used in research on ESRD patients has been validated for studies on health status. This study evaluated three existing indices (Khan, Davies, and Charlson) and four indices specifically developed for use in studies on health status. In a large prospective multi-center study (NECOSAD-2), new ESRD patients were included ($n = 1041$). Comorbidity was assessed at the start of dialysis. Health status was assessed with the physical and mental component summary score of the SF-36 (PCS and MCS), the symptoms dimension of the KDQOL-SF, and the Karnofsky Scale. Patient data were randomly allocated to a modeling or a testing set. The new indices were developed in the modeling set. The three existing indices explained about the same percentage of variance in the PCS (7 to 8%), MCS (1 to 3%), symptoms (2 to 4%), and Karnofsky (10 to 12%). The new indices performed better than the existing indices in the modeling population (13% PCS, 10% MCS, 10% symptoms, 18% Karnofsky), but not in the testing population (8% PCS, 1% MCS, 3% symptoms, 8% Karnofsky). Individual comorbidities explained more variance in PCS (10 to 15%), MCS (1 to 7%), symptoms (6 to 11%), and Karnofsky (11 to 18%) than comorbidity indices. The Khan, Davies, and the Charlson indices will adjust to the same extent for the potential confounding effect of comorbidity in studies with health status as an outcome. Separate comorbidity diagnoses will adjust best for comorbidity.

Treatment of patients with end-stage renal disease (ESRD) aims not only at prolonging life, but also at achieving the greatest possible well being and functioning. For this reason, health status — sometimes also referred to as quality of life — is an important outcome in research on ESRD patients. Several studies have been carried out in which treatment modalities, drugs, times of initiation of dialysis, and rehabilitation program were evaluated with respect to their effect on health status.

To allow comparisons of health status among groups of patients, an adjustment for differences in prognostic factors has to be made. In many studies, results are adjusted for risk factors such as residual renal function or hematocrit level.

Comorbidity is an inevitable prognostic factor, but there is no agreement on how to adjust for it. The simplest proxy for comorbidity — also the most widely used — is age. Age seems a poor index of comorbidity when used in isolation; therefore, several attempts have been made to create comorbidity indices. So far, several such comorbidity indices have been developed. In some studies, the number of comorbidities is calculated. In other studies, more advanced ways to summarize comorbidity have been used by indices in which more severe comorbidities have higher weights than less severe comorbidities. Registries with limited information on comorbidity, such as the United States Renal Data System (1), only take the most common comorbidities into account. Other studies also look at less frequent comorbidities.

The preferred comorbidity index in studies that use health status as an outcome measure is one that most adequately reflects the impact of comorbidity on health status. Such an index will capture the potential confounding effect of comor-

Received June 11, 2002. Accepted October 4, 2002.

The NECOSAD-study group: Apperloo AJ, Barendregt JNM, Birnie RJ, Boekhout M, Boer WH, Bommel EFH van, Büller HR, Charro FTh de, Doorenbos CJ, Dorp WT van, Es A van, Fagel WJ, Feith GW, Franssen CFM, Frenken LAM, Geelen JACA van, Geerlings W, Gerlag PGG, Gorgels JPMC, Grave W, Huisman RM, Jager KJ, Jansen MAM, Jie K, Koning-Mulder WAH, Koolen MI, Kremer Hovinga TK, Lavrijssen ATJ, Mulder AW, Parlevliet KJ, Rosman JB, Saase JLCM van, Schonk MJM, Schuurmans MMJ, Stevens P, Tijssen JGP, Valentijn RM, Vastenburg G, Verburg CA, Verhagen CE, Verstappen VMC, Vincent HH, Vos P

Correspondence to Dr. Jeannette G. van Manen, AMC, Dept. Clinical Epidemiology & Biostatistics, P.O. Box 22700, 1100 DE Amsterdam, the Netherlands. Phone: 31-20-5663607; Fax: 31-20-6912683; E-mail: j.g.vanmanen@amc.uva.nl

1046-6673/1402-0478
Journal of the American Society of Nephrology

DOI: 10.1097/01.ASN.0000043902.30577.C9

bidity best. It can be argued that the sheer number of comorbidities, as used in the commonly used Davies index (2), is a crude measure for the prognostic impact of comorbidity on health status. Unlike other indices, the Davies index does not include age, because it was specifically designed to be used in conjunction with age as an independent covariate. Other comorbidity indices that are in use in studies on ESRD patients assign different weights to different comorbidities, such as the Khan (3) or the Charlson index (4), with the weights based on the impact of comorbid diseases on survival. However, the impact of comorbid diseases on survival may be rather different from their impact on health status.

None of the indices that are commonly used in research on ESRD patients has been validated for studies on health status. We therefore evaluated these existing indices with respect to their association with health status. Four important components of health status were used for this purpose: overall physical and mental health status, symptoms dimension of disease-specific health status, all three assessed by the patient, and a more objective measure, namely overall health status assessed by the patients' physician or nurse. We also investigated whether indices that include comorbidities differentially weighted by their impact on health status are more strongly associated with health status than existing indices. For this purpose, we created four new comorbidity indices. The impact of existing indices and new indices on health status were evaluated in a prospective cohort study on patients new on dialysis.

Materials and Methods

Patients

All new ESRD patients from 36 dialysis centers in the Netherlands were consecutively invited to participate in the Netherlands Cooperative Study on the Adequacy of Dialysis (NECOSAD-2), a prospective cohort study. All invited patients were 18 yr or older, and chronic dialysis had to be their first renal replacement therapy. The inclusion period was between January 1997 and November 2000.

Data Collection Procedures

Patient data were collected at the start of dialysis treatment. The following data were recorded: age, gender, primary kidney disease, comorbidity, serum albumin, residual renal function, and initial dialysis treatment. Primary kidney disease was classified according to the codes of the European Renal Association–European Dialysis and Transplantation Association (ERA-EDTA).

A plasma sample and 24-h urine sample were obtained simultaneously. Serum albumin, plasma creatinine, and plasma urea levels were determined. Urea and creatinine were analyzed in the urine sample. Renal function was expressed as residual GFR (rGFR), calculated as the mean of creatinine and urea clearance, corrected for body surface area, and as renal Kt/V_{urea} per week, calculated as urea clearance corrected for distribution volume (V) according to Watson *et al.* (5). Calculation of normalized protein equivalent of nitrogen appearance (nPNA) was performed according to Bergström *et al.* (6,7) normalized to actual body weight.

Comorbidity was defined in terms of presence or absence of nonrenal disease at the onset of chronic dialysis treatment. The patients' nephrologists completed a list of 15 different comorbid conditions, including an expression of their severity (see Table 2 for a detailed list). The primary kidney diseases diabetes mellitus (ERA-EDTA codes: 80, 81), renal vascular disease (ERA-EDTA codes: 70, 71, 72, 79), and systemic collagen diseases (ERA-EDTA codes: 73, 74, 82, 83, 84, 85, 86, 87, 88, 89) were considered comorbid diseases as well and were added up to diabetes mellitus and systemic collagen disease that were recorded in the list of 15 comorbid conditions.

Health status was assessed by the Medical Outcomes Study 36-item Short-Form Health Survey (SF-36), the Kidney Disease Quality of Life-Short Form (KDQOL-SF), and the Karnofsky Scale. The SF-36 is a generic measure of health status and is composed of 36 questions

Table 1. Patient characteristics of the modeling and testing population[a]

	Modeling Population $n = 515$	Testing Population $n = 526$
Age (yr)	59.5 (14.9)	58.4 (15.7)
Males (%)	62	60
Primary kidney disease (%)		
diabetes	16	14
glomerulonephritis	15	14
renal vascular disease	19	18
other	51	53
Serum albumin (g/dl)	3.6 (0.6)	3.7 (0.6)
Body mass index (kg/m^2)	24.6 (3.9)	24.9 (4.5)
rGFR (mL/min per 1.73 m^2)	4.9 (2.9)	4.9 (3.1)
Renal Kt/V_{urea} (/wk)	1.0 (0.6)	1.0 (0.7)
nPNA (g/kg per d)	0.93 (0.25)	0.95 (0.29)
Modality (%)		
hemodialysis	60	59
peritoneal dialysis	40	41
Physical summary scale SF-36	38.7 (9.7)	38.9 (9.3)
Mental summary scale SF-36	40.8 (11.9)	41.4 (12.1)

[a] Mean values (SD) are given for continuous variables.

Table 2. Prevalence of chronic diseases in the modeling population and association between diseases and PCS and MCS (beta coefficients) ($n = 515$)[a]

	Prevalence *n* (%)	Beta for PCS	Beta for MCS	Khan	Davies	Charlson
Diabetes mellitus	99 (19.5)			X	X	X
non-insulin–dependent	37 (7.3)	−2.0				
insulin-dependent <15 yr	20 (3.9)	−3.7				
insulin-dependent ≥15 yr	42 (8.3)	−6.1				
Renal vascular disease	96 (18.6)					
Stroke	37 (7.2)	−3.5	−6.0			X
Myocardial infarction	57 (11.1)		−3.7	X	X	X
Angina pectoris	55 (10.7)			X	X	
no pain	25 (4.9)	0.4	−3.7			
pain doing heavy work	15 (2.9)	−2.3	−6.0			
pain doing light work	12 (2.3)	−5.0	−0.5			
pain in rest	3 (0.6)	−2.9	−7.9			
Congestive heart failure	63 (12.3)			X	X	X
no dyspnea	14 (2.7)					
dyspnea doing heavy work	24 (4.7)					
dyspnea doing light work	16 (3.1)					
dyspnea in rest	9 (1.8)					
Other heart diseases	11 (2.1)					
Peripheral vascular disease	67 (13.0)				X	X
Malignancies	52 (10.1)				X	X
previous	37 (7.2)					
current	15 (2.9)					
Liver cirrhosis	3 (0.6)			X		X
Obstructive pulmonary disease	37 (7.2)	−5.4		X		X
Musculoskeletal disorders	40 (7.8)	−3.0	5.8			
Systemic collagen disease	43 (8.3)	−4.2		X		
Chronic gastric/intestinal diseases	26 (5.1)		−4.1			
Psychiatric disorders	16 (3.1)		−9.7			
Pulmonary fibrosis				X		
Visceral malignancies				X		
Other significant pathology					X	
Age (yr)				X[b]		
18 to 40	72 (12.1)	0	0			X
40 to 50	86 (14.5)	−2.2	−2.9			X
50 to 60	115 (19.4)	−1.4	−1.0			X
60 to 70	144 (24.2)	−2.4	−0.5			X
70 to 80	149 (25.1)	−4.6	−1.4			X
≥80	28 (4.7)	−5.4	1.8			X

[a] The beta coefficients were used to create two new indices. Comorbidities that are included in the Khan, Davies, and Charlson indices are marked with an X.

[b] Age in three classes.

and is widely used and validated (8). Two summary scales of the SF-36 can be computed. The Physical Component Summary (PCS) and Mental Component Summary (MCS) scales summarize the eight separate dimensions of the SF-36 (9). The eight dimensions have different weights for the PCS and the MCS: the PCS is predominantly related to physical functioning, role-functioning due to physical problems, bodily pain, and general health, whereas the MCS is predominantly related to social functioning, mental health, role-functioning due to emotional problems, and vitality. The scales have a mean of 50 and a SD of 10 in the general Dutch population. The KDQOL-SF is a disease-specific instrument that encompasses 44 items that focus on particular health-related concerns of individuals with kidney disease and on dialysis and that are organized into twelve domains (10). To limit the number of health status measures in the present study, the domain Symptoms was chosen to evaluate the indices, because this domain is mostly concerned with kidney/dialysis-related problems and it is the most continuous scale. The Karnofsky Scale is a clinician-assessed scale of functional status. It consists of ten levels, with scores ranging from 10 (moribund) to 100 (normal, without limitations) (11).

Analyses

The Khan, Davies, and Charlson indices were calculated for each patient. The Khan index is a combination of age and comorbidity, used to assign patients to one of three risk groups; low, medium, and high risk (3). Patients over 80 yr are always classified in the high-risk group. Patients between 70 and 80 yr are allocated to the medium or high-risk group. The Davies score is based on the presence or absence of seven comorbid conditions, also producing three risk groups (2). Age is not included in this index. Patients without comorbid conditions are classified as low risk. Patients with one or two comorbid diseases are regarded as medium-risk patients. Patients with three or more comorbid conditions are classified as high-risk patients. The Charlson index is based on weights for each comorbidity and age class (4). The weights express the associated risk of mortality. From 40 yr onwards, an increase in age results in one additional point per decade. The weights for the different comorbid conditions range from one to six points. The weights of all comorbidities that are present and the concurrent age score are summed to obtain a final score. To be able to compare the Charlson index with the Khan and Davies indices, the Charlson index was used to classify patients into three risk groups, based on tertile limits.

We also created four comorbidity indices that were based on the impact of several comorbid conditions on the two summary dimensions of the SF-36, the physical and the mental summary scale (PCS and MCS), on the Symptoms dimension of the KDQOL and on the Karnofsky Scale. These indices were created in a modeling group and subsequently evaluated in a testing group. Patients were randomly allocated to one of these two groups. To create these four indices, a linear regression analysis was performed in the modeling population with either the PCS, the MCS, Symptoms dimension of the KDQOL, or the Karnofsky Scale as dependent variable and comorbidities and age as independent variables. Age was categorized into six groups. Each single comorbid condition and, when applicable, its different severity stages were included in the regression analysis with a forward procedure. The four final multivariable models consisted of those comorbid conditions that were significantly associated with the PCS, the MCS, Symptoms dimension or Karnofsky Scale respectively. The regression coefficients (weights) of each of the four final models were then used to obtain four scores for each patient by adding the coefficients of the diseases that were present and by adding the coefficient of the appropriate age class.

The prognostic impact of the indices on health status was compared by looking at the percentage of variance they explained of the two summary scores of the SF-36. This was performed in two ways. First, a linear regression analysis was performed with the PCS, the MCS, the Symptoms dimension, or the Karnofsky Scale as the dependent variable and each index as the independent variable in the modeling population. Second, these analyses were repeated in the testing population to find out the extent to which the association among the four newly developed indices and health status as found in the modeling population would differ from the association of the index in an external population.

The indices take age into account differently. This makes it difficult to compare how these indices assess the prognostic impact of comorbidity on health status. We therefore evaluated them in a separate analysis. In this analysis, each index and age were included in a linear regression model as independent variables, and the PCS, the MCS, the Symptoms dimension, or the Karnofsky Scale were included as dependent variable. We compared the indices with respect to the percentage of variance that could be explained by the respective age-adjusted models.

Finally, health status was modeled as a function of single comorbid conditions, instead of comorbidities summarized in an index. First, we looked at three conditions that are commonly recorded: diabetes, myocardial infarction, and congestive heart failure. In addition, we examined all conditions recorded in our study, as well as the conditions included in the Khan, Davies, and Charlson indices (see Table 2).

Results

In total, 1041 ESRD patients were included at the start of dialysis in the NECOSAD-2 study. These patients were randomly allocated to a modeling or a testing group. The modeling group consisted of 515 patients. This group was used to create four comorbidity indices that are based on their association with the four main components of health status. The testing group consisted of 526 patients. The latter group was used to validate these and other commonly used summarizations of comorbidity.

Table 1 contains the general characteristics of both populations. These were highly similar. The mean age was 60 and 58 yr, 62% and 60% were male, about 15% had diabetes as primary kidney disease, mean serum albumin levels were 3.6 and 3.7 g/dl, residual renal function was 4.9 ml/min per 1.73 m^2, and about 60% were on hemodialysis. Figures on health status (SF-36) were highly similar in both groups.

Table 2 shows the prevalence of the comorbid conditions, and their association with the physical and mental components of the SF-36 (PCS and MCS) in the modeling group. Diabetes mellitus and renal vascular disease were the most frequent conditions, whereas liver cirrhosis and psychiatric disorders were least common. Renal vascular disease, malignancies, liver cirrhosis, congestive heart failure, other heart diseases, and peripheral vascular disease were not significantly associated with either the PCS or the MCS. The other comorbidities were differently associated with the PCS and the MCS. Diabetes, obstructive pulmonary disease, and systemic collagen disease were associated with the PCS but not with the MCS, whereas the opposite was true for myocardial infarction, chronic gastric/intestinal disease, and psychiatric disorders. Furthermore, musculoskeletal disorders were negatively associated with the PCS, whereas they were positively associated with the MCS. For the PCS, there were clear trends toward more physical impairment with increasing severity of the disease (diabetes and angina pectoris) and with increasing age. For example, patients with insulin-dependent diabetes that existed for 15 yr or more have a PCS score of about six points lower on average than patients without diabetes, whereas patients with non-insulin–dependent diabetes have a PCS score of two points lower than patients without diabetes.

The coefficients or weights shown in Table 2 were used to create two new indices, the index for the PCS and for the MCS. Table 3 shows the association of the existing comorbidity indices and the new indices with health status in the modeling and in the testing population by means of the percentage of explained variance in the PCS, MCS, Symptoms, and Karnofsky. In the modeling population, the proportion of PCS variance that the Davies, Khan, and Charlson indices explained

Table 3. Percentage of variance in health status explained by each comorbidity index (and age) in the modeling population ($n = 515$) and in the testing population ($n = 526$)

	PCS		MCS		Symptoms		Karnofsky	
	Index	Index + Age[b]	Index	Index + Age[b]	Index	Index + Age[b]	Index	Index + Age[b]
Modelling population								
age	4[a]	4[a]	<1	<1	2[a]	2[a]	6[a]	6[a]
number of comorbidities								
Davies index	7[a]	8[a]	3[a]	3[a]	4[a]	4[a]	10[a]	11[a]
survival-based indices								
Khan index	8[a]	8[a]	3[a]	4[a]	2[a]	3[a]	11[a]	12[a]
Charlson index	8[a]	8[a]	1[a]	2[a]	3[a]	3[a]	12[a]	12[a]
health status-based indices								
index for PCS	13[a]	13[a]	1	1	3[a]	3[a]	11[a]	12[a]
index for MCS	<1	4	10[a]	10[a]	1[a]	3[a]	2[a]	8[a]
index for Symptoms	4[a]	5[a]	1[a]	1	10[a]	10[a]	5[a]	8[a]
index for Karnofsky	7[a]	8[a]	1[a]	1[a]	3[a]	3[a]	18[a]	18[a]
Testing population								
age	5[a]	5[a]	<1	<1	<1	<1	6[a]	6[a]
number of comorbidities								
Davies index	8[a]	9[a]	1[a]	1[a]	6[a]	6[a]	8[a]	11[a]
survival-based indices								
Khan index	8[a]	8[a]	1	1	4[a]	4[a]	10[a]	11[a]
Charlson index	8[a]	9[a]	1	1	3[a]	5[a]	10[a]	10[a]
health status-based indices								
index for PCS	8[a]	9[a]	<1	<1	4[a]	4[a]	7[a]	9[a]
index for MCS	1	5	1[a]	1[a]	1	1	1[a]	7[a]
index for Symptoms	3[a]	5[a]	<1	<1	3[a]	3[a]	1[a]	6[a]
index for Karnofsky	5[a]	6[a]	<1	<1	1[a]	1	8[a]	9[a]

[a] $P < 0.05$.

[b] Some of the indices have incorporated age, and age was incorporated in these indices in different ways. Therefore, it is difficult to separate the effects of age from the effects of comorbidity diagnoses on health status. For this reason, the analyses were also performed with age next to the index included in the model.

was rather similar (7% to 8%). This was also true for the MCS (1% to 3%), Symptoms (2% to 4%), and Karnofsky (10% to 12%). However, the indices that were specifically developed for the association with each of the health status measures, explained more variance of the PCS (13%), the MCS (10%), Symptoms (10%), and Karnofsky (18%) than for any of the other indices.

The age-adjusted analyses showed that age accounts for a large part of the variance in the PCS (4%) that can be explained by the index and age together (8% to 13%). This was also true for Symptoms (2% for age; 3 to 10% for index+age) and Karnofsky (6% for age; 11 to 18% for index+age). Age explained less than 1% of the variance in the MCS. After adjusting the results for age, we still found that the two indices that were specifically created for each of the health status measures performed best (13% PCS, 10% MCS, 10% Symptoms, 18% Karnofsky). The Khan, Davies, and Charlson indices performed rather equally (8% PCS, 2% to 4% MCS, 3% to 4% Symptoms, 11% to 12% Karnofsky).

The indices for the health status measures were created in the modeling group; therefore, the above-mentioned figures on the performance of these indices that were also calculated in the modeling group may be overestimations of their performance. Therefore, the analyses were repeated in the testing population as well. Now, the new indices (8% PCS, 1% MCS, 3% Symptoms, 8% Karnofsky) performed similar as compared with the existing indices (8% PCS, 1% MCS, 3 to 6% Symptoms, 8 to 10% Karnofsky). Thus, the two new indices performed less in the testing population than in the modeling population. When adjustments for differences between the summarizations with respect to age were made, all indices still performed rather similarly (8% to 9% PCS, 1% MCS, 3 to 6% Symptoms, 9 to 11% Karnofsky).

Finally, health status was modeled as a function of individual comorbidities and age in the testing population (Table 4). Three common comorbidities, diabetes mellitus, myocardial infarction, and congestive heart failure, and age explained 10% of the variance of the PCS, 1% of the variance of the MCS, 6% of the variance of Symptoms, 11% of the variance of Karnofsky. This was rather similar compared with the percentage of

Table 4. Percentage of variance in health status explained by modeling health status as a function of single comorbidities and age. Testing population ($n = 526$)

	PCS	MCS	Symptoms	Karnofsky
Age	5[a]	0	0	6[a]
+ diabetes, MI, congestive heart failure	10[a]	1	6[a]	11[a]
+ comorbidities from Davies[b]	11[a]	2	7[a]	12[a]
+ comorbidities from Khan[b]	13[a]	2	9[a]	13[a]
+ comorbidities from Charlson[b]	12[a]	2	8[a]	12[a]
+ all comorbidities[b]	15[a]	7	11[a]	18[a]

[a] $P < 0.05$.
[b] See Table 2 for list of comorbidities.

variance explained by the separate comorbidities from the Davies, Khan, and Charlson indices (11% to 13% PCS, 2% MCS, 7% to 9% Symptoms, 12% to 13% Karnofsky). However, the largest part of the variance in the PCS (15%), MCS (7%), Symptoms (11%), and Karnofsky (18%) was explained by all comorbidities that were listed in Table 2. Furthermore, each set of individual comorbidities and age explained a larger part of the variance of the PCS (10% to 15%) than each comorbidity index and age (8% to 9%). For the MCS, Symptoms, and Karnofsky, only the total list of individual comorbidities and age explained a substantial larger part of the variance than each of the comorbidity indices and age.

Discussion

In this study, several comorbidity indices were compared with regard to their association with physical, mental, disease-specific, and overall health status in ESRD patients. The indices were evaluated this way because indices that are most strongly related to health status will capture the possible confounding effect of comorbidity best in evaluation studies with health status as an outcome. The Davies, Khan, and Charlson indices that are commonly used in studies on ESRD patients, but that are considered to have the disadvantages to be either crude summarizations or to be developed for use in survival studies, performed rather equally in our study. The indices that were specifically created in this study, and that included comorbidities weighted by their impact on health status, performed better than the other indices when compared in the population in which they were created. However, these indices did not perform better than the other indices when compared in an external testing population. Finally, health status was more strongly associated with individual comorbidities than with comorbidities summarized into an index.

Three studies in ESRD patients have shown that a modified Charlson index, the Friedman comorbidity index, and the Khan index are significantly associated with health status as assessed by the Karnofsky index, the Sickness Impact Profile, or the SF-36 (12–14). To our knowledge, however, there are no studies in which different comorbidity indices are systematically compared with regard to the extent to which they capture the influence of comorbidity on health status. In the development and evaluation of comorbidity indices, emphasis has always been on finding a comorbidity index that can adequately stratify patients according to their risk of death. Studies in ESRD patients showed that several types of indices were all predictive of survival (3,15–22). In studies in other populations than ESRD patients and with survival as the outcome of interest, it was shown that comorbidity indices provide only a modest improvement on age adjustment and that the performance of different indices is rather similar (23–26). In a recent study by our own group, we found that the Charlson, Khan, and Davies indices had similar prognostic value with respect to mortality in ESRD patients (27). We furthermore found that a comorbidity index that included severity gradings of several diseases did not perform better than the Charlson, Khan, and Davies indices. Because indices that reflect the impact of comorbidity on survival do not necessarily reflect the impact of comorbidity on health status in the same way, we also studied the effect on health status and found in the present study that all comorbidity indices performed rather equally with respect to their association with health status. Moreover, the Khan, Davies, and Charlson indices that are developed for use in studies on survival performed similarly compared with the indices we developed in the present study for use in studies on health status. Thus, it made no difference whether the indices included comorbidities weighted by their impact on health status or whether the indices simply counted the number of comorbidities or included comorbidities weighted by mortality risks.

One reason for finding that the comorbidity indices in the present study performed similarly in explaining health status may be that three of the measures for health status we used are overall scores or summary scores of separate domains of health status and may therefore not be very sensitive for differences in comorbidity. For this reason we also analyzed the association between the indices and the specific domains of the SF-36. However, we still found no large differences in the performance of the indices. For example, the Davies, Khan, and Charlson indices explained 14% to 18% of the variance in physical functioning and 1% to 2% of the variance in mental health (1% to 2%), which was about the same as the variance that was explained by the indices that were specifically created for their association with physical functioning (15%) or mental health (2%).

Another reason for not finding large differences between the indices and their association with health status may be that indices of any kind perform poorly because they summarize a complex construct in an overly simplistic way. The presence or the number of comorbidities is an equally rough and imprecise indication of comorbidity as apparently more advanced indices that include comorbidities with different weights. The weights that are derived in a modeling population to create an index may overfit the data, particularly given the relatively small number of patients with a specific comorbidity. Therefore, it is unlikely that the weights will exactly apply to an external testing population. However, this problem of deriving imprecise weights when creating an index seems inevitable because our modeling population that consisted of about 500 patients was already quite large and it will be difficult to carry out much larger studies.

In our study, we also modeled health status as a function of the separate comorbidities to overcome the problem of summarizing comorbidity on the basis of weights that may not be exactly applicable to the population in which it is used. By modeling the individual comorbidities, weights are derived for each comorbidity that will exactly apply to the population in which it is used. We found that the separate comorbidities explained a larger part of the variance (15% PCS, 7% MCS, 11% Symptoms, 18% Karnofsky) than the comorbidity indices (8% to 9% PCS, 1% MCS, 3% to 6% Symptoms, 9% to 11% Karnofsky). Thus, the residual variance that can be attributed to comorbidity increases when using an index. However, the use of an index increases statistical efficiency because fewer variables are included in the model. This is important for studies that are not very large. In each study, it should be evaluated whether an increase in residual confounding outweighs an increase in statistical efficiency when using an index.

The separate comorbidities explained a larger part of the variance in health status than any of the comorbidity indices. However, the percentage of variance that could be explained by the separate comorbidities was still rather low (maximal 18%). One reason for this may be that the separate comorbidity diagnoses are still imprecise measures for comorbidity. Another and possibly more important reason may be that demographic and clinical factors only partly determine health status. Several studies have shown that the percentage of variance in health status that can be explained by symptoms, primary kidney disease, residual renal function, nutritional status, employment, or education varies from 7% to 39% (28–31). Thus, a large part of the variance in health status cannot be explained by these factors. Moreover, these factors and other factors, such as coping, personality, stress or social support may be more important in determining health status than comorbidity, which may indicate the limitations of improving health status by focusing on comorbidity alone.

In this study, we were able to evaluate the most commonly used comorbidity indices. One other index that has been used in ESRD patients is the Index of Co-Existent Disease (ICED) (32,33). The ICED summarizes comorbidity by adding the peak score of disease severity for the comorbidities (by level of symptoms, treatability, and threat to life) to the peak score for the impact of the comorbidities on physical impairment. This index has been shown to predict survival, but no data exist on its association with health status. However, as the degree of physical impairment of the patient is rated for this index, the ICED seems equivalent with a health status instrument that measures physical impairment. Even when assuming that the ICED would be an appropriate comorbidity index to use, the ICED requires answering more than 100 questions, and it can take a trained person up to one hour per patient to complete, which is impractical for many studies.

An additional finding of our study was that the indices we created performed better in the sample in which they were created than in the testing population, which demonstrates the importance of validating an index in an external population. Although this phenomenon has been mentioned in the literature (34), indices are not always validated this way. However, if we had not validated the indices in an external population, we would have concluded that the indices we created performed better than other comorbidity indices. Nevertheless, this is not in accordance with our findings in the testing population.

The comorbidity indices we compared in the present study performed equally with respect to their association with health status. Therefore, these indices will adjust to about the same extent for the potential confounding effect of comorbidity in evaluation studies with health status as an outcome. However, adjusting for comorbidity by using a summarization of comorbidity is associated with more residual confounding compared with adjusting for comorbidity by separate comorbidity diagnoses. The latter approach is however not always feasible, because it can only be performed in rather large populations. In smaller or exploratory studies, it will be necessary or sufficiently to use a summarization or index of comorbidity. With respect to the comorbidity indices evaluated in this study, we cannot recommend one index above another. Therefore, the Khan, Davies, and Charlson indices that are commonly used in studies on ESRD patients may all be appropriate. On the basis of practical considerations, however, the Davies index may be preferred because the least variables have to be recorded to score this index.

Acknowledgments

This work was supported by grants from The Dutch Kidney Foundation (E.018) and the Dutch National Health Insurance Board (OG97/005). The nursing staff of the dialysis centers, who collected most of the data, is gratefully acknowledged for their assistance. The authors also wish to thank J. Aurisch, A. Houweling, C. Janssen, K. Meijs, B. Nijman, H. Schrijver, and L. Ten Brinke for their assistance in the logistics of this study.

References

1. The USRDS Dialysis Morbidity and Mortality Study: Wave 2. United States Renal Data System. *Am J Kidney Dis* 30: S67–S85, 1997
2. Davies SJ, Phillips L, Naish PF, Russell GI: Quantifying comorbidity in peritoneal dialsyis patients and its relationship to other predictors of survival. *Nephrol Dial Transplant* 17: 1085–1092, 2002

3. Khan IH, Catto GRD, Edward N, Fleming LW, Henderson IS, MacLeod AM: Influence of coexisting disease on survival on renal replacement therapy. *The Lancet* 341: 415–418, 1993
4. Charlson ME, Pompei P, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis* 40: 373–383, 1987
5. Watson PE, Watson ID, Batt RD: Total body water volumes for adult males and females estimated from simple anthropometric measurements. *Am J Clin Nutr* 33: 27–39, 1980
6. Bergström J, Furst P, Alvestrand A, Lindholm B: Protein and energy intake, nitrogen balance and nitrogen losses in patients treated with continuous ambulatory peritoneal dialysis. *Kidney Int* 44: 1048–1057, 1993
7. Bergström J, Heimburger O, Lindholm B: Calculation of the protein equivalent of total nitrogen appearance from urea appearance. Which formulas should be used? *Perit Dial Int* 18: 467–473, 1998
8. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E: Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 51: 1055–1068, 1998
9. Ware JE, Kosinski M, Keller SD: *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston MA, The Health Institute, 1994
10. Korevaar JC, Merkus MP, Jansen MA, Dekker FW, Boeschoten EW, Krediet RT: Validation of the KDQOL-SF: A dialysis-targeted health measure. *Qual Life Res* 11: 437–447, 2002
11. Edgell ET, Coons SJ, Carter WB, Kallich JD, Mapes D, Damush TM, Hays RD: A review of health-related quality-of-life measures used in end-stage renal disease. *Clin Ther* 18: 887–938, 1996
12. Rebollo P, Ortega F, Baltar JM, Badia X, Alvarez-Ude F, Diaz-Corte C, Naves M, Navascues RA, Urena A, Alvarez-Grande J: Health related quality of life (HRQOL) of kidney transplanted patients: Variables that influence it. *Clin Transplant* 14: 199–207, 2000
13. Moreno F, Lopez Gomez JM, Sanz-Guajardo D, Jofre R, Valderrabano F: Quality of life in dialysis patients. A spanish multicentre study. Spanish Cooperative Renal Patients Quality of Life Study Group. *Nephrol Dial Transplant* 11[Suppl 2]: 125–129, 1996
14. Khan IH, Garratt AM, Kumar A, Cody DJ, Catto GR, Edward N, MacLeod AM: Patients' perception of health on renal replacement therapy: Evaluation using a new instrument. *Nephrol Dial Transplant* 10: 684–689, 1995
15. Hutchinson TA, Thomas DC, MacGibbon B: Predicting survival in adults with end-stage renal disease: An age equivalence index. *Ann Intern Med* 96: 417–423, 1982
16. Pollak VE, Pesce A, Kant KS: Continuous quality improvement in chronic disease: A computerized medical record enables description of a severity index to evaluate outcomes in end-stage renal disease. *Am J Kidney Dis* 19: 514–522, 1992
17. Davies SJ, Russell L, Bryan J, Phillips L, Russell GI: Comorbidity, urea kinetics, and appetite in continuous ambulatory peritoneal dialysis patients: Their interrelationship and prediction of survival. *Am J Kidney Dis* 26: 353–361, 1995
18. Khan IH, Campbell MK, Cantarovich D, Catto GR, Delcroix C, Edward N, Fontenaille C, Fleming LW, Gerlag PG, van Hamersvelt HW, Henderson IS, Koene RA, Papadimitriou M, Ritz E, Russell IT, Stier E, Tsakiris D, MacLeod AM: Survival on renal replacement therapy in Europe: is there a "centre effect"? *Nephrol Dial Transplant* 11: 300–307, 1996
19. Khan IH, Campbell MK, Cantarovich D, Catto GR, Delcroix C, Edward N, Fontenaille C, van Hamersvelt HW, Henderson IS, Koene RA, Papadimitriou M, Ritz E, Ramsay C, Tsakiris D, MacLeod AM: Comparing outcomes in renal replacement therapy: How should we correct for case mix? *Am J Kidney Dis* 31: 473–478, 1998
20. Khan IH: Comorbidity: The major challenge for survival and quality of life in end-stage renal disease. *Nephrol Dial Transplant* 13: 76–79, 1998
21. Beddhu S, Bruns FJ, Saul M, Seddon P, Zeidel ML: A simple comorbidity scale predicts clinical outcomes and costs in dialysis patients. *Am J Med* 108: 609–613, 2000
22. Fried L, Bernardini J, Piraino B: Charlson comorbidity index as a predictor of outcomes in incident peritoneal dialysis patients. *Am J Kidney Dis* 37: 337–342, 2001
23. Schneeweiss S, Maclure M: Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol* 29: 891–898, 2000
24. Gabriel SE, Crowson CS, O'Fallon WM: A comparison of two comorbidity instruments in arthritis. *J Clin Epidemiol* 52: 1137–1142, 1999
25. Krousel-Wood MA, Abdoh A, Re R: Comparing comorbid-illness indices assessing outcome variation: The case of prostatectomy. *J Gen Intern Med* 11: 32–38, 1996
26. van Doorn C, Bogardus ST, Williams CS, Concato J, Towle VR, Inouye SK: Risk adjustment for older hospitalized persons: A comparison of two methods of data collection for the Charlson index. *J Clin Epidemiol* 54: 694–701, 2001
27. van Manen JG, Korevaar JC, Dekker FW, Boeschoten EW, Bossuyt PMM, Krediet RT: How to adjust for comorbidity in survival studies in ESRD patients: A comparison of different indices. *Am J Kidney Dis* 40: 82–89, 2002
28. Moreno F, Lopez Gomez JM, Sanz-Guajardo D, Jofre R, Valderrabano F: Quality of life in dialysis patients. A spanish multicentre study. Spanish Cooperative Renal Patients Quality of Life Study Group. *Nephrol Dial Transplant* 11: 125–129, 1996
29. Harris LE, Luft FC, Rudy DW, Tierney WM: Clinical correlates of functional status in patients with chronic renal insufficiency. *Am J Kidney Dis* 21: 161–166, 1993
30. Merkus MP, Jager KJ, Dekker FW, De Haan RJ, Boeschoten EW, Krediet RT: Physical symptoms and quality of life in patients on chronic dialysis: Results of The Netherlands Cooperative Study on Adequacy of Dialysis (NECOSAD). *Nephrol Dial Transplant* 14: 1163–1170, 1999
31. Merkus MP, Jager KJ, Dekker FW, Boeschoten EW, Stevens P, Krediet RT: Quality of life in patients on chronic dialysis: Self-assessment 3 months after the start of treatment. The Necosad Study Group. *Am J Kidney Dis* 29: 584–592, 1997
32. Niccolucci A, Cubasso D, Labbrozzi D, Mari E, Impicciatore P, Procaccini DA, Forcella M, Stella I, Querques M, Pappani A: Effect of coexistent diseases on survival in patients undergoing dialysis. *ASAIO J* 38: m291–m295, 1992
33. Athienitis NV, Miskulin DC, Fernandez G, Bunnapradist S, Simon G, Landa M, Schmid CH, Greenfield S, Levey AS, Meyer KB: Comorbidity assessment in hemodialysis and peritoneal dialysis using the Index of Coexistent Disease. *Semin Dial* 13: 320–326, 2000
34. Charlson ME, Ales KL, Simon R, MacKenzie CR: Why predictive indexes perform less well in validation studies. Is it magic or methods? *Arch Intern Med* 147: 2155–2161, 1987