Projecting the Number of Patients with End-Stage Renal Disease in the United States to the Year 2015

David T. Gilbertson,* Jiannong Liu,* Jay L. Xue,*† Thomas A. Louis,*‡ Craig A. Solid,* James P. Ebben,* and Allan J. Collins*†

*United States Renal Data System Coordinating Center and †University of Minnesota, Minneapolis, Minnesota; and ‡Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland

The size of the prevalent ESRD population in the United States increased dramatically during the 1990s, from 196,000 in 1991 to 382,000 in 2000. Incidence also increased considerably during the same period, from 53,000 to 93,000 per year. If previous trends in ESRD incidence and prevalence continue, then current levels of health care resources that are devoted to the care of these patients will eventually be unable to meet the demand. This study discusses a Markov model developed to predict ESRD incidence, prevalence, and mortality to the year 2015 and incorporating expected changes in age/race distributions, diabetes prevalence, ESRD incidence, and probability of death. The model predicted that by 2015 there will be 136,166 incident ESRD patients per year (lower/upper limits 110,989 to 164,550), 712,290 prevalent patients (595,046 to 842,761), and 107,760 ESRD deaths annually (96,068 to 118,220). Incidence and prevalence counts are expected to increase by 44 and 85%, respectively, from 2000 to 2015 and incidence and prevalence rates per million population by 32 and 70%, respectively. The financial and human resources that will be needed to care for these patients in 2015 will be considerably greater than in 2005.

As reported by the United States Renal Data System (USRDS), the size of the prevalent ESRD population in the United States nearly doubled during the 1990s, increasing from 196,000 in 1991 to 382,000 in 2000. The number of incident cases also increased considerably during the same period, from 53,000 to 93,000 per year (1). A recent USRDS projection of the growth in the ESRD population through 2010 showed a near doubling of this population to 650,000 individuals, with 520,000 prevalent on dialysis (2). If these trends in ESRD incidence and prevalence continue, then the level of health care resources that currently are allocated to the care of these patients will eventually be unable to meet the demand. Predicting ESRD incidence and prevalence is important so that policy makers and health care providers can plan appropriately.

Several factors will influence the future numbers of patients with ESRD, including the aging of the baby boomers, changing racial distributions, increasing prevalence of diabetes, increasing probability of surviving to ESRD before dying, and changing life expectancy of patients with ESRD. To predict ESRD incidence, prevalence, and mortality to the year 2015, we developed a Markov model, incorporating expected changes in demographic distributions, diabetes prevalence, ESRD incidence, and probability of death.

Materials and Methods

The Model

A discrete-time, annually incremented, nonstationary Markov model was developed to model the ESRD population from 1978 to 2000 and project to the year 2015. The model (Figure 1) consists of an incident block and a prevalent block. Patients who develop ESRD are assigned to the incident block; if they remain alive to the end of the calendar year, then they are moved to the prevalent block for the beginning of the next year. By creating these two blocks, we were able to estimate transition probabilities (probability of moving from one state in the model to another state) for each of the two groups separately. This ability was especially important for cause of renal failure because the effect of diabetes varies greatly between incident and prevalent patients, resulting in different transition probabilities.

As shown in Figure 1, within the incident and prevalent blocks, there are two states: "DM" and "No DM." The former represents individuals with diabetes as the cause of renal failure; the latter comprises those with renal failure for any other cause. Within each of these states, there are also substates (not shown in the figure) created by the cross-classification of seven age groups (0 to 18, 19 to 40, 41 to 64, 65 to 69, 70 to 74, 75 to 79, and ≥80 yr) and three race groups (white, black, and other). Hence, there are 21 substates (7 × 3) within each of the two states ("DM" and "No DM") within each of the two blocks (incident and prevalent), for a total of 84 possible cells.

Data Sources and Model Parameter Estimation

ESRD Incident Counts. Figure 2 displays the data sources and method used to estimate ESRD incident counts by age, race, and cause of renal failure for the years 1978 to 2015. Population estimates by age and race for 1981 to 2015 were obtained from the US Census Bureau (www.census.gov). These estimates were linearly extrapolated backward to obtain population estimates by age and race for 1978 to 1981. Diabetes prevalence estimates by age and race for 1981 to 1999 were obtained from Boyle *et al.* (3); these estimates were based on data from
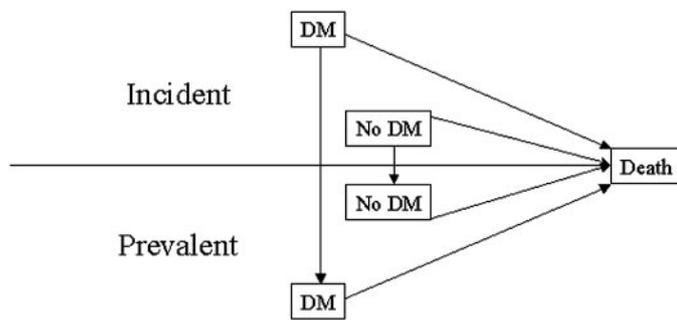
Figure 1. Model of ESRD. See Materials and Methods (The Model) for explanation. DM, diabetes mellitus.

the National Health Interview Survey (NHIS), which is performed yearly by the National Center for Health Statistics (an organizational component of the Centers for Disease Control and Prevention). The diabetes prevalence estimates by age and race were linearly extrapolated backward and forward to obtain estimates of diabetes prevalence for 1978 to 2015. The resulting estimates were multiplied by census estimates for each year to obtain estimates of the number of individuals with and without diabetes, by age and race, for 1978 to 2015. ESRD incident counts by age, race, and cause of renal failure for 1978 to 2000 were divided by diabetic and nondiabetic population counts during that same period to obtain ESRD incidence rates by age, race, and cause of renal failure (4). These rates were extrapolated forward to obtain ESRD incidence rates for all years from 1978 through 2015. Three methods of extrapolation were investigated: Fitting linear trend lines to the incidence rates and then extrapolating, fitting Gompertz lines and then extrapolating, and holding ESRD incidence rates at their year 2000 levels. The linear extrapolation was performed giving more weight to later years to reflect the flattening of the incidence rates observed for 1999 to 2001, resulting in a more conservative extrapolation to the year 2015. These ESRD incidence rates then were multiplied by expected diabetes and nondiabetes population counts for 1978 to 2015 to obtain expected incident counts. Because the rate of increase in the incidence of ESRD slowed somewhat during the 1990s, separate model runs were performed fitting separate lines using 1978 to 2000 *versus* 1990 to 2000.

Other Model Parameters.   Actual ESRD prevalence in 1978 was used in the model and was modeled from 1979 through 2015 (4). A relatively small percentage of ESRD patients are lost to follow-up each year for a variety of reasons, such as recovery of renal function, leaving the country, and nonreporting of death to the Social Security Administration. Actual lost-to-follow-up counts from 1978 through 2000 were used and were linearly extrapolated from 2001 to 2015. Some patients are not covered by Medicare (patients with Medicare as secondary payer) for a variety of reasons, such as Centers for Medicare and Medicaid Services coverage rules (the first 30 to 33 mo are not covered for patients with employer coverage) and health maintenance organization coverage. Patients with Medicare as a secondary payer were included because it was determined that for the current model, claims (which would be affected by coverage other than Medicare) are not needed, but death information (which would not be affected by coverage other than Medicare) is needed. For the prevalent population from 1978 through 2000, the 1-yr probability of death was calculated for each year for each of the 42 groups within it: Seven age groups × three race groups × two cause-of-renal-failure groups (diabetes *versus* any other cause). For the incident population from 1978 through 2000, the probability of death by year end was calculated for each year in each of its 42 groups. For both incident and prevalent transition probabilities of

death, we fit 42 exponential curves for 1978 through 2000 and extrapolated through 2015.

Model Implementation

The starting point for the model was January 1, 1978. Each "year" of the model was implemented by multiplying the number of individuals in each cell by the corresponding transition probability for each possible model transition. In each cell, the number of prevalent patients and number of deaths were recorded; the remaining patients were "aged" 1 yr, and new incident patients were added. This process was repeated for each year through 2015, using actual incident counts for each year through 2000 and extrapolated incident counts through 2015.

Two implementations of the model were performed. The first, which used counts and transition probabilities based on data from 1978 through 1990 and extrapolation through 2000, was used to assess the model's ability to predict incidence, prevalence, and mortality during 1991 through 2000. The second implementation used counts and probabilities based on 1978 through 2000 to project 2001 through 2015. Because projected numbers did not agree exactly with actual numbers for 2000, incident counts, prevalent counts, and numbers of deaths were recalibrated to actual counts in the year 2000 to project to 2015.

Two sets of upper and lower limits were generated. The first set of limits was generated as follows: Upper limits were obtained by increasing ESRD incidence and diabetes prevalence rates beyond the year 2000, decreasing the probability of death among patients with ESRD, and increasing the underlying age/race US population numbers in a structured manner. The changes to ESRD incidence, diabetes prevalence, and probability of death were made such that, by the year 2015, ESRD incidence and diabetes prevalence were 10% higher and probability of death was 10% lower than the values used for the point estimates in 2015. For the underlying population age/race structure, the US Census Bureau's "Highest Series" was used. Lower limits were obtained similarly by decreasing ESRD incidence and diabetes prevalence, increasing the probability of death, and using the Census Bureau's "Lowest Series."

The second set of upper and lower limits, which are more statistical in nature and can be considered actual 95% confidence limits, was generated by putting sampling distributions around model parameters and running the model 1000 times, each time sampling each model parameter from the appropriate distribution. This resulted in 1000 estimates of incidence, prevalence, and death counts; the 2.5th and 97.5th percentiles of these estimates were used as 95% confidence limits. Counts were generated using a Poisson distribution, and proportions were generated using a Beta distribution. This second set of limits addressed statistical uncertainty in terms of estimation of model parameters, whereas the first set addressed broader uncertainty regarding model structure and general assumptions about future change in the primary model parameters.

Results
Model Validation

Figure 3 shows actual and predicted ESRD incidence, prevalence, and mortality counts from 1978 to 2000, using information from 1978 to 1990. The model underpredicted the number of incident cases by 5%, prevalent cases by 3%, and deaths by 10%. This indicates that there were changes in diabetes prevalence, ESRD incidence, age/race distributions, or probability of death beyond our extrapolation of the 1980s. Regarding diabetes prevalence, NHIS data show a relatively flat prevalence during the 1980s but a slightly increasing prevalence during the
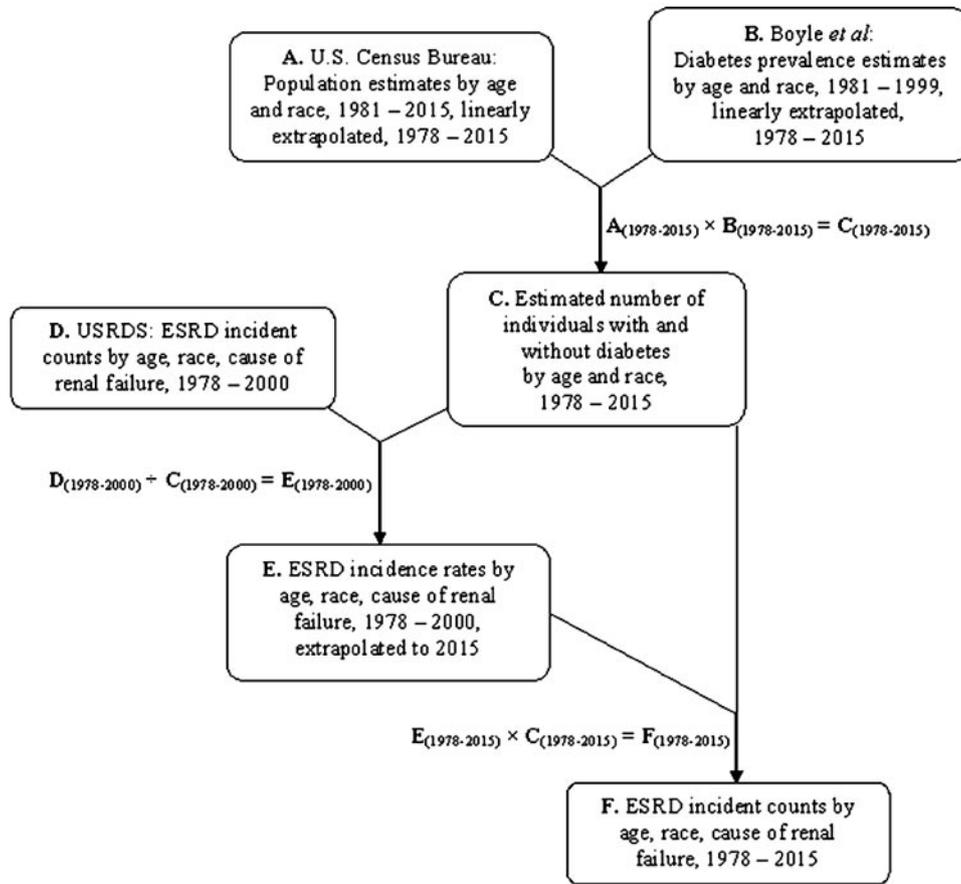
Figure 2. Data sources and method used to estimate ESRD incident counts, 1978 to 2015. USRDS, United States Renal Data System.
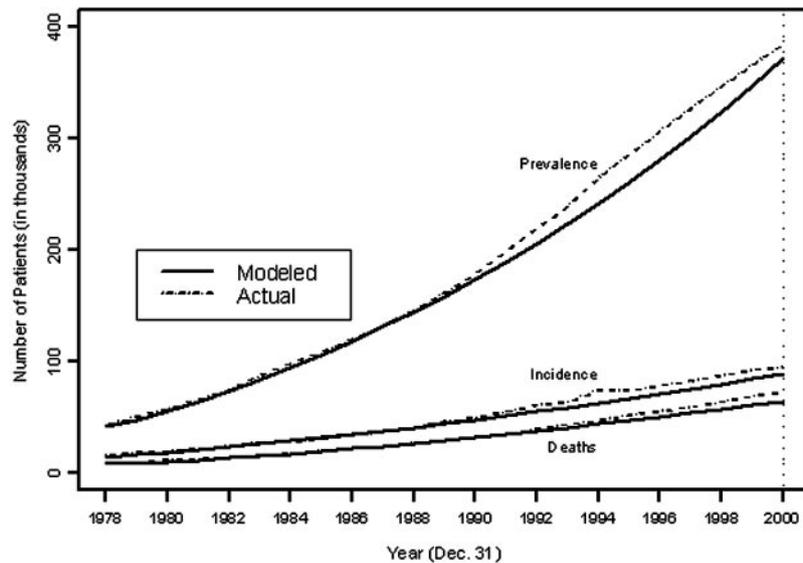

Figure 3. Model validation: Actual and predicted ESRD incidence, prevalence, and mortality, using information from 1978 to 1990.

1990s (3). This led to underprediction of diabetes prevalence in the current model during the 1990s because the model used a linear extrapolation of the 1980s. The prediction of ESRD incidence was affected in a similar way. Incidence rates increased at a slightly higher rate during the 1990s than during the 1980s

(1). Therefore, extrapolation of ESRD incidence using only the 1980s incidence rates underestimated 1990s incidence. With use of diabetes prevalence and ESRD incidence from the 1990s as well as from the 1980s, incidence and prevalence predicted for 2000 were within 0.5% of the actual values.

Prediction

Figures 4 and 5 show ESRD incidence and prevalence estimates to the year 2015. The model predicted that by 2015, annual incidence will be 136,166 patients per year and annual prevalence will be 712,290, with 107,760 deaths per year occurring among patients with ESRD. The upper and lower limits obtained by changing the rates of increase or decrease of diabetes prevalence, ESRD incidence, and probability of death showed that incidence could range from 110,989 to 164,550; prevalence from 595,046 to 842,761; and deaths from 96,068 to 118,220. Ninety-five percent confidence limits that were obtained by iteratively sampling model parameters from appropriate distributions resulted in considerably tighter bounds: 135,200 to 137,200 for incidence, 708,400 to 716,000 for prevalence, and 106,900 to 108,600 for deaths.

Rates per Million Population and Growth Rates

This article focuses on incident and prevalent counts, as opposed to rates, because it is primarily counts of people that will influence planning by policy makers and health care providers. However, by dividing overall actual and projected counts by census numbers, actual and projected, we can examine rates. From 2000 to 2015, the model predicted an increase in the incidence, from 343 per million in 2000 to 453 per million in 2015, a 32% increase. For prevalence, it predicted a 70% increase, from 1396 to 2371 per million.

Table 1 displays the average annual growth rate by decade. The greatest growth occurred during the 1980s, with less growth occurring during the 1990s. The model predicted a further decline in the growth rate through 2015. The pattern observed (high, early growth rates followed by a leveling off) probably reflects early increases in acceptance to therapy and expanding access to dialysis, given the growth of dialysis providers and units (4).

Discussion

Although the number of incident ESRD cases has leveled off during the past few years (1), our Markov model predicts that this trend will not continue, as a number of factors are expected to lead to further increases in the future. Currently, patients with chronic kidney disease (CKD) are five to 10 times more likely to die than to reach ESRD (4,5). With continued progress in the treatment of CKD and cardiovascular disease (the main cause of death among patients with CKD), more patients will reach ESRD before dying, further contributing to the increasing incidence of ESRD. The aging of the baby boomers (born between 1946 and 1964) will also contribute significantly to expected increases in incidence and prevalence, as will increasing prevalence of diabetes and obesity and improved care of patients with ESRD (the latter leading to increased life expectancy for those who receive renal replacement therapy) (6,7).

Using a stepwise autoregressive method with exponential smoothing, Xue *et al.* (2) estimated the number of incident and prevalent patients with ESRD through the year 2010. Our model predicts 9600 fewer incident cases in 2010 (119,562 *versus* 129,200) and 60,000 fewer prevalent cases (591,023 *versus* 651,330). This can be explained partially by the method of extrapolating ESRD incidence in the current model. Because the *USRDS 2005 Annual Data Report* shows a continuing of the trend in ESRD incidence rates shown in earlier Annual Data Reports, that is, a leveling off in virtually all age/race/disease groups, we held ESRD incidence rates constant at the 2000 levels for projecting to 2015 (as opposed to using either linear or Gompertz extrapolation) (8). The autoregressive model extrapolated information from the 1980s and early 1990s, during a time of consistently increasing ESRD incidence rates.

Other researchers have used various techniques to predict future numbers of ESRD patients in the United States and elsewhere (9–17). Port (14) showed projections of patients with ESRD per million population from 1992 through 2000, assuming either continued exponential increase as seen during the 1980s or a more level increase that is not exponential. The actual rate reported by the *USRDS 2003 Annual Data Report* in 2000 is 332 per million population, which is between Port's two esti-
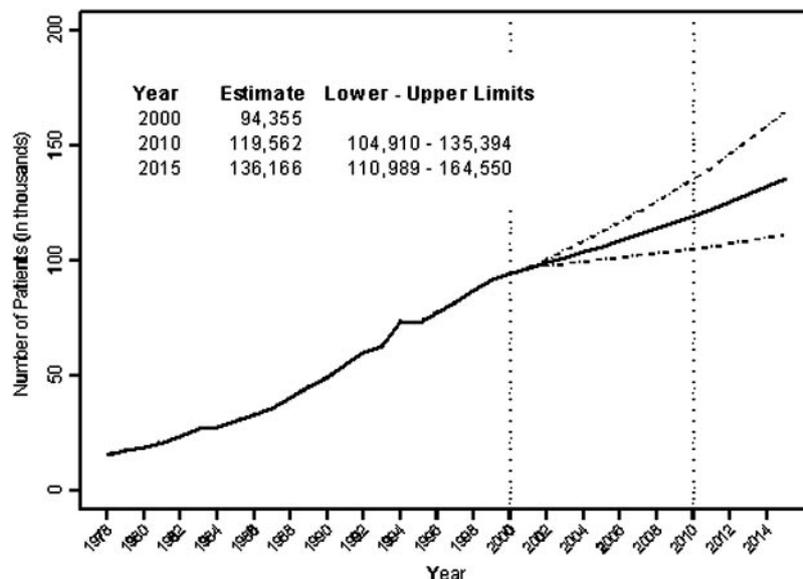

Figure 4. Incidence estimates (solid line) and upper/lower limits (dashed lines) for 1978 to 2015.
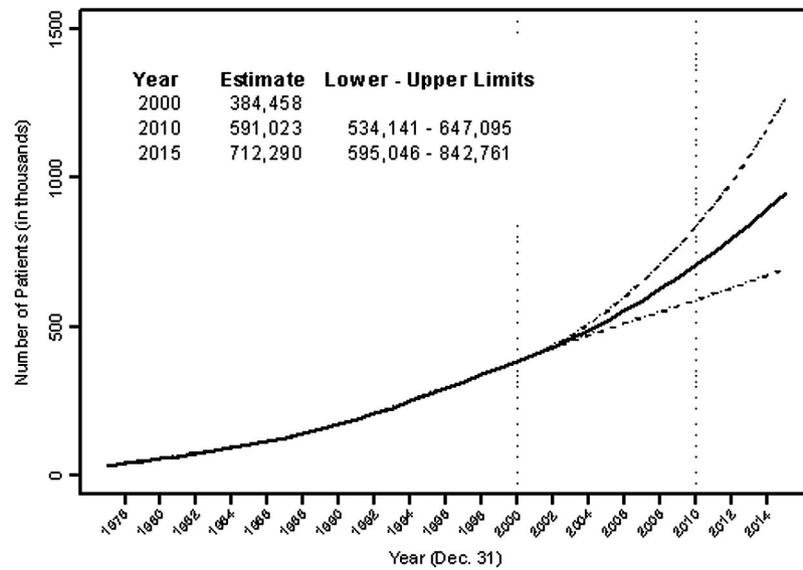
Figure 5. Prevalence estimates (solid line) and upper/lower limits (dashed lines) for 1978 to 2015.

Table 1. Average percentage annual growth rate by decade

Decade	Incidence	Prevalence
1980 to 1989	10.0	12.0
1990 to 1999	7.4	8.8
2000 to 2009	2.5	4.6
2010 to 2015	2.6	3.8

mates. These results, taken together with results from our model, point to continued slowing of ESRD growth rates but continued increases in sheer numbers of patients, leading to a much larger ESRD population in 2015 than exists today.

Even though our model is relatively complex, consisting of 84 cells created by groupings of age, race, cause of renal failure, and incident *versus* prevalent cases, it is still a relatively basic model of ESRD at the population level. In particular, although incident and prevalent cases are separated, after incident cases reach year end, they become part of the prevalent group, and transition probabilities for prevalent patients are independent of vintage. This suggests that the survival experience for patients who are on dialysis for 1 yr is similar to that for patients who are on dialysis for 5 or 10 yr. Because patients with increasing vintage generally have higher death rates (5), the probabilities of death in the model for prevalent patients may be an underestimate, which could lead to overprediction of the expected numbers of patients with ESRD in the future. Also, cause of renal failure is grouped into diabetic *versus* all other. Despite these simplifications, results from the current model give estimates for incidence and prevalence in 2010 that are within 10% of results from the previous autoregressive model.

Validation of the current model, in which information from the 1980s was used to predict the 1990s, showed that the numbers of predicted *versus* actual cases for the year 2000 differed by 3% for prevalence and 5% for incidence. Two primary factors seem to explain these discrepancies: Changes in diabetes prevalence and changes in ESRD incidence from the 1980s to the 1990s. Data from the NHIS show little change in diabetes prevalence during the 1980s (18). However, data from the NHIS and from the Behavioral Risk Factor Surveillance System show substantial increases in diabetes prevalence during the 1990s (19). The validation method of using data from the 1980s to predict the 1990s therefore extrapolates a flat diabetes trend throughout the 1990s, instead of an increasing prevalence, resulting in fewer predicted ESRD cases and underprediction of ESRD incidence, prevalence, and death counts. In terms of changing ESRD incidence rates, the *USRDS 2003 Annual Data Report* shows that ESRD incidence increased at a slightly higher rate during the 1990s than during the 1980s (1). Therefore, extrapolation of ESRD incidence using the 1980s data underestimated 1990s incidence. When the model was run for validation using 1990s diabetes prevalence estimates and ESRD incidence rates, agreement between model and actual incidence, prevalence, and death counts was much closer, within 0.5%.

The two methods that were used for computing upper and lower limits produced different results. The more statistically based 95% confidence limits were much tighter than the limits that were not statistically based. The 95% confidence limits were tighter because the model parameters were generally estimated from large numbers of individuals, so the resulting parameters were well estimated, and, therefore, these model-based confidence limits were narrow. The wider, nonstatistically based limits were intended to indicate what would happen if our assumptions about the model parameters were significantly different from what actually occurred. The structure of our model is adequate for describing the ESRD population currently, but it may be inadequate in the future.

Like the autoregressive model, our model makes predictions that are based on information that is currently available. Predicted growth in the ESRD program may be significantly al-

tered by increased attention to prevention. Rising national attention to the epidemic of obesity may slow the increasing rates of diabetes, which in turn would lead to fewer ESRD cases in the future. New treatments for diabetes also may slow the increase. Detection programs and more use of renoprotective therapies, as outlined in the Kidney Disease Outcomes Quality Initiative guidelines, are heightening provider awareness of CKD (20), which in turn may lead to lower ESRD incidence rates in the future as well. Continued improvements in efficiency of care, use of more cost-effective treatments, consideration of palliative as opposed to interventive care in some cases, and better management of patients who initiate dialysis may also contribute to altered growth in the program, thereby decreasing the burden on the health care system. However, even if rates of diabetes prevalence and ESRD incidence were 10% lower than expected in 2015, the probability of death were 10% higher, and the "lowest" census projections are used, the predicted number of prevalent cases in 2015 still would be >595,000. The financial and human resources that will be needed to care for these patients in 2015 will be considerably greater than in 2005, regardless of the exact number of cases.

Given these findings, the public health implications are considerable. Much more attention may be needed to manage the size of the pool of individuals who subsequently proceed to ESRD. It seems that a 20% or greater improvement in outcomes will be needed to lower the expected growth of the ESRD program for both Medicare and private health plans. Earlier identification of patients with CKD and improved treatment of patients with ESRD both are likely to reduce the death rate. What is not clear is the extent to which the rate of progression to ESRD will be affected by changes "upstream" of ESRD. A reduction in the CKD death rate may increase the rate at which patients with CKD progress to ESRD; in this case, the dialysis and transplant populations will grow even faster than predicted with our model. Conversely, a reduction in the ESRD rate may exceed a reduction in the CKD death rate; in this case, the size of the ESRD population will diminish. Prospective data collection is needed to address such pressing epidemiologic questions in CKD and ESRD; the insights to be gained will play no small role in shaping public health policy and health care planning related to the treatment of kidney disease.

Acknowledgments

References

1. US Renal Data System: *USRDS 2003 Annual Data Report*, Bethesda, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 2003
2. Xue JL, Ma JZ, Louis TA, Collins AJ: Forecast of the number of patients with end-stage renal disease in the United States to the year 2010. *J Am Soc Nephrol* 12: 2753–2758, 2001
3. Boyle JP, Honeycutt AA, Narayan KM, Hoerger TJ, Geiss LS, Chen H, Thompson TJ: Projection of diabetes burden through 2050: Impact of changing demography and disease prevalence in the US. *Diabetes Care* 24: 1936–1940, 2001
4. US Renal Data System: *USRDS 2002 Annual Data Report*, Bethesda, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 2002
5. US Renal Data System: *USRDS 2004 Annual Data Report*, Bethesda, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 2004
6. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, Marks JS: Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA* 289: 76–79, 2003
7. Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, Koplan JP: The continuing epidemics of obesity and diabetes in the United States. *JAMA* 286: 1195–1200, 2001
8. US Renal Data System: *USRDS 2005 Annual Data Report*, Bethesda, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 2005
9. Schaubel DE, Morrison HI, Desmeules M, Parsons DA, Fenton SS: End-stage renal disease in Canada: Prevalence projections to 2005. *CMAJ* 160: 1557–1563, 1999
10. Schaubel DE, Morrison HI, Desmeules M, Parsons DA, Fenton SS: End-stage renal disease projections for Canada to 2005 using Poisson and Markov models. *Int J Epidemiol* 27: 274–281, 1998
11. Motohashi Y, Nishi S: Prediction of end-stage renal disease patient population in Japan by system dynamics model. *Int J Epidemiol* 20: 1032–1036, 1991
12. Molzahn M: Future evolution of the ESRD patient population—A perspective for the year 2000. *Nephrol Dial Transplant* 11[Suppl 8]: 59–62, 1996
13. Ruwaard D, Hoogenveen RT, Verkleij H, Kromhout D, Casparie AF, van der Veen EA: Forecasting the number of diabetic patients in The Netherlands in 2005. *Am J Public Health* 83: 989–995, 1993
14. Port FK: End-stage renal disease: Magnitude of the problem, prognosis of future trends and possible solutions. *Kidney Int* 50[Suppl]: S3–S6, 1995
15. Vestergaard P, Lokkegaard H: Predicting future trends in the number of patients on renal replacement therapy in Denmark. *Nephrol Dial Transplant* 12: 2117–2123, 1997
16. Davies R, Roderick P: Predicting the future demand for renal replacement therapy in England using simulation modelling. *Nephrol Dial Transplant* 12: 2512–2516, 1997
17. Roderick P, Davies R, Jones C, Feest T, Smith S, Farrington K: Simulation model of renal replacement therapy: Predicting future demand in England. *Nephrol Dial Transplant* 19: 692–701, 2004
18. Mokdad AH, Ford ES, Bowman BA, Nelson DE, Engelgau MM, Vinicor F, Marks JS: Diabetes trends in the US: 1990–1998. *Diabetes Care* 23: 1278–1283, 2000
19. National Diabetes Surveillance System. Available: http://www.cdc.gov/diabetes/statistics. Accessed November 12, 2004
20. K/DOQI clinical practice guidelines for chronic kidney disease: Evaluation, classification, and stratification. Kidney Disease Outcome Quality Initiative. *Am J Kidney Dis* 39: S1–S246, 2002

Correction

Pertosa *et al.*: Coagulation Cascade Activation Causes CC Chemokine Receptor-2 Gene Expression and Mononuclear Cell Activation in Hemodialysis Patients. *J Am Soc Nephrol* 16: 2477–2486, 2005.

The authors regretfully report two labeling errors in the figures of this article. In Figure 2B, the cellulose acetate (CA) and ethylen-vinyl-alcohol (EVAL) labels should be reversed (see below). Also, in Figure 5 the black bars indicate CA, not AC. Please see the corrected figures below.
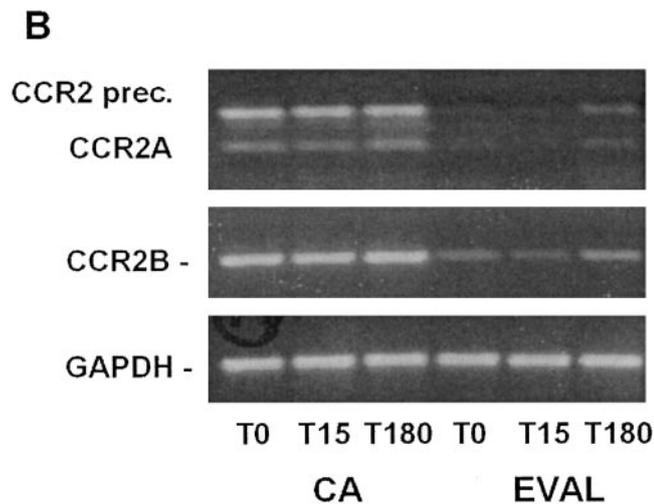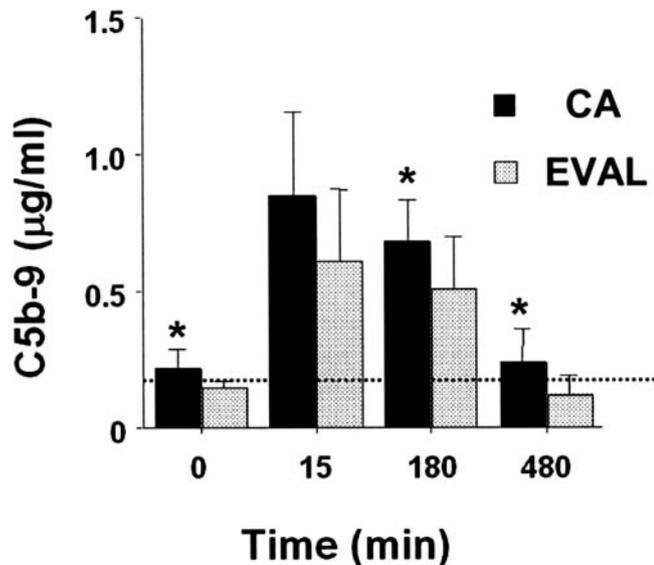

Figure 2B.

Figure 5.

Correction

Bernier *et al.*: Pharmacologic Chaperones as a Potential Treatment for X-Linked Nephrogenic Diabetes Insipidus. *J Am Soc Nephrol* 17: 232–243, 2006.

In this article, the authors regretfully report an error in a plasma concentration datum. The second sentence of the first full paragraph in the right-hand column on page 239 should read as follows (bold text indicates corrected information):

A 300-mg dose of the F1 formulation was reported in the Clinical Investigator Brochure to increase plasma concentration to 18.5 ng**/ml** (**30 nanomolar**, MW of SR49059 is 620.5) 3 h after the administration of SR49059 to normal volunteers.

Correction

Gilbertson *et al.*: Projecting the Number of Patients with End-Stage Renal Disease in the United States to the Year 2015. *J Am Soc Nephrol* 16: 3736–3741, 2005.

In this article, the authors regretfully report an error in the representation of data as seen in Figure 5. The data discussed in the text is correct, as are the numbers in the figure, but the *y*-axis and the lines in the graph are not correct. Please see the corrected figure below.
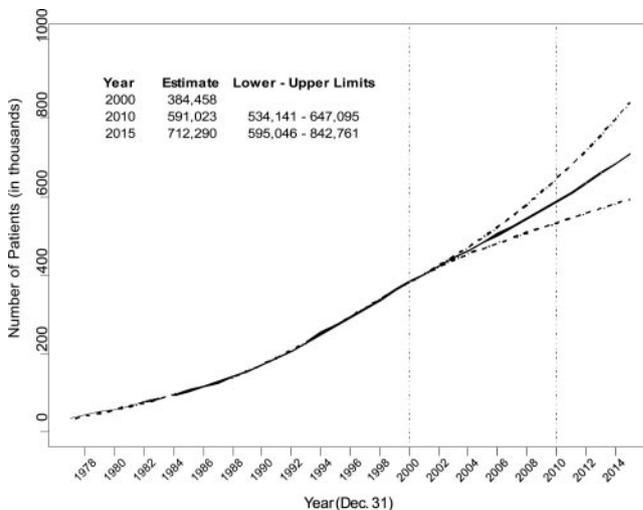

Figure 5. Prevalence estimates (solid line) and upper/lower limits (dashed lines) for 1978 to 2015.