Kidney International, Vol. 57, Suppl. 74 (2000), pp. S-49–S-54

Projecting renal replacement therapy–specific end-stage renal disease prevalence using registry data

Douglas E. Schaubel, Howard I. Morrison, and Stanley S.A. Fenton

Laboratory Centre for Disease Control, Health Canada, Ottawa, Ontario, Canada; Faculty of Medicine, University of Toronto, and Division of Nephrology, Toronto General Hospital, University Health Network, Toronto, Ontario, Canada; and Department of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill, North Carolina, USA

Projecting renal replacement therapy–specific end-stage renal disease prevalence using registry data. End-stage renal disease incidence and prevalence are increasing in many countries worldwide. Projections of ESRD prevalence are useful for forecasting future resource requirements, and organ failure registry databases are valuable for the development of appropriate projection models. We outline one method of generating renal replacement therapy (RRT)–specific ESRD prevalence projections based on data obtained from the Canadian Organ Replacement Register (CORR). To illustrate the methods, we present national RRT-specific prevalence projections for Canada to the year 2005. Continued large increases in ESRD incidence and prevalence are projected, particularly among diabetics. As of December 31, 1996, there were 17,807 patients receiving RRT in Canada. This number is projected to climb to 32,952 by the end of 2005, for a relative increase of 85% (average relative increase of 5.8% per year). Registry data are a useful basis for future health care planning.

End-stage renal disease (ESRD) incidence and prevalence are increasing in Canada [1], the United States [2], Australia [3], and several European countries [4]. The considerable public health importance of ESRD derives from the high mortality rate and impaired quality of life among patients, and the high cost of renal replacement therapy (RRT) [5]. Projections of future numbers of patients receiving RRT are of interest to health care planners and providers to forecast equipment, facility, manpower and other resource requirements. We propose one method of projecting RRT-specific ESRD prevalence using registry data, which has been previously employed to generate prevalence projections for Canada at both the national [6] and provincial [7] levels.

METHODS

Data sources

End-stage renal disease patient-specific data were obtained from the Canadian Organ Replacement Register (CORR), a population-based, nation-wide organ failure registry [1] operated by the Canadian Institute for Health Information (CIHI). Demographic data (e.g., birth date, gender, race, province of residence) and baseline clinical data (e.g., primary renal diagnosis (PRD), predialysis comorbid conditions) are collected from patients by each renal center upon RRT initiation. Therapeutic history data (e.g., RRT assignments, dialytic modality switches, transplantations, graft failures) are submitted annually. Death data (i.e., date and cause of death) are reported along with the other follow-up information. Coverage by CORR is complete in that records from all 86 Canadian renal centres are submitted annually. As well, all follow-up time is included, including the first 90 days post-initiation of RRT.

Population data, including census counts, intercensal estimates, and population projections, were obtained from Statistics Canada [8].

Overview of projection model

Below, we provide a basic description of a model previously employed to project RRT-specific national prevalence up to the year 2005 [6]. Since the funding of renal centres is largely within the jurisdiction of the provincial governments, projections to 2005 by province were later attempted using very similar methods [7]. As an illustration of the model's application, we present national results based on the latter study, derived by summing across all provinces. Readers interested in more statistical details of the projection model are referred to the former paper [6].

At the time of the most recent projections, CORR data were available for all patients initiating RRT in Canada between January 1, 1981 and December 31, 1996. Patients were classified by age ($\leq$44, 45–64, $\geq$65), province, and diabetes status. Separate projections were produced by province for each age $\times$ diabetes cross-classification, to allow the parameters of the model to differ across subpopulations defined by these variables. It was expected *a priori* that incidence (i.e., new case rate per

Key words: demography, dialysis, end-stage renal disease, epidemiology, Markov model, Poisson, transplantation.

year) and/or prevalence (i.e., year-end number of patients receiving RRT) trends in each subpopulation would be different.

The number of patients receiving RRT at any time beyond the conclusion of the period of observation, December 31, 1996, is composed of: (i) survivors of incident cases occurring between 1995 and 2005, and (ii) survivors among "currently prevalent" patients (i.e., patients receiving RRT as of December 31, 1996). Thus, although (i) and (ii) are not important quantities by themselves prior to being summed to calculate prevalence, it is important to consider the two in isolation since, as will be discussed in the paragraphs that follow, these are computed separately. All statistical analyses were conducted using the statistical computing language SAS (v6.12, SAS Institute; Cary, NC).

Incidence projections

Annual incidence rates between 1997 and 2005 were estimated by extrapolating Poisson regression [9–11] models based on the 1981–1996 data. That is, Poisson regression models were fitted to the 1981–1996 data, and the coefficients from these models were used to estimate annual incidence rates for each year between 1997 and 2005. Thus, the model for each subpopulation defined by age and diabetes status, would have the form:

$$\log\{I(y)\} = \beta_0 + \beta_1 y + \beta_2 (y - 1988)^2 \quad \text{(Eq. 1)}$$

where I(y) is the incidence rate during year *y*. Subtracting 1988 in the quadratic term dampens the correlation between β_1 and β_2. Estimates of β_0, β_1, and β_2, were used to project incidence rates for future years via the following equation:

$$I(y) = \exp[\beta_0 + \beta_1 y + \beta_2 (y - 1988)^2],$$
$$\text{for } y = 1997, 1998, \ldots, 2005 \quad \text{(Eq. 2)}$$

Having projected the incidence rate, the annual number of incident cases was obtained by multiplying the projected rate by the projected population size. Models were fitted separately by province and diabetes status, and contained terms for age, year, and all indicated age × year interactions. Quadratic terms in year were entered as needed. Diabetic status-specific models permitted the age and year effects to be different for diabetics and non-diabetics, without the need for third order interaction terms (e.g., age × year × diabetes), which are difficult to interpret.

A second set of projections was made wherein it was assumed that incidence during 1997–2005 remained constant at the average across the 1994–96 period.

Prevalence projections

Having projected the number of incident cases during 1997–2005, and knowing the number of currently prevalent cases (i.e., those receiving RRT as of December 31, 1996), we project the follow-up experience for each of these two groups separately. A Markov model [12, 13] was employed for this purpose. We first describe the fundamental quantities of the Markov model, then describe how they are used to project patient follow-up.

Table 1. Markov model states

State	RRT
1	Hemodialysis
2	Peritoneal dialysis
3	Functioning transplant
4	Death

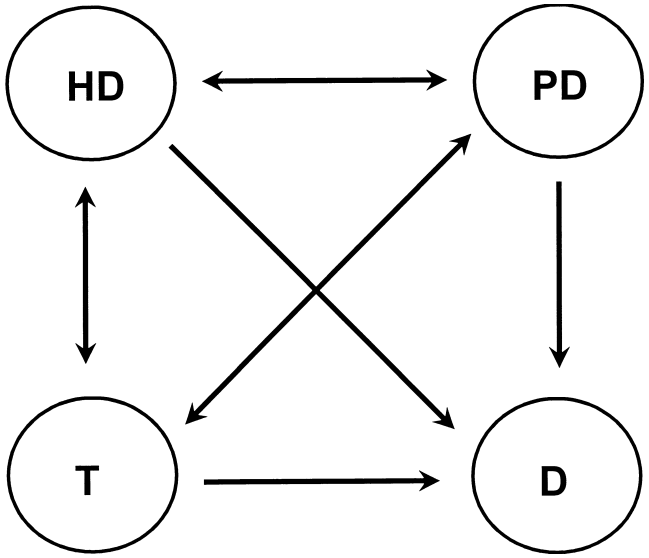

Fig. 1. Therapeutic "states" for the Markov model of patient follow-up. 1: Hemodialysis (HD), 2: peritoneal dialysis (PD), 3: functioning transplant (T), 4: death (D). Arrows denote possible state transitions.

A Markov model is, basically, a matrix of state transition probabilities. In our case, we have defined four states, as listed in Table 1: hemodialysis (HD), peritoneal dialysis (PD), functioning transplant, and death. In the formulae to follow, the numbering of the states is consistent with that listed in Table 1. Naturally, death is an "absorbing" state since transfer from it is impossible. Transitions among each of the other three states are possible, as are transitions from all RRT states to the death state. A state transition diagram is presented in Fig. 1.

A Markov model can be conceptualized as a generalization of the familiar survival model. Typically in survival analysis, there are 2 states: alive and dead. Since transition is possible in only one direction, a Markov model of this setting would reduce to one quantity: survival probability. Had only the total number of patients receiving RRT been of interest, as opposed to RRT-

specific prevalence, a survival model would have sufficed. Since the objective is RRT-specific projections, we must estimate the fraction of patients, among incident cases for each year, that will begin treatment on HD, PD and transplant. As well, we must estimate modality switching probabilities.

Modality assignment probabilities

For incident cases, it is necessary to estimate the fraction that will begin on each RRT. Set $q_j(t) = P(\text{state } j \text{ at time } t)$, for j = 1, 2, and 3, t = 0, 6, 12, 18, where t is measured in months of follow-up (i.e., months post-initiation of RRT), and where the numbering of the states is consistent with Table 1. Hence, t = 0 corresponds to initial RRT assignment. We estimate $q_j(0)$ by:

$$q_j(0) = c_j(0)/\{c_1(0) + c_2(0) + c_3(0)\} \quad \text{(Eq. 3)}$$

where $c_j(0)$ is the number of patients beginning in state j, based on the CORR data file previously described. Of course, $q_4(0) = 0$. Only 1990–1996 data were employed to estimate modality assignment probabilities for the 1997–2005 period, since future RRT assignment patterns are more likely to reflect those most recently experienced, and since data volume in the most recent 7 years was sufficient for such purposes. Consistent with the incidence models, each quantity in the follow-up models is estimated separately by province, age (≤44, 45–64, ≥65) and PRD (diabetic, non-diabetic). For clarity, notation identifying the particular subgroup have been suppressed. Thus, formulae which follow were applied separately to each province/age/diabetes stratum.

Transition matrices

It makes sense that state transition probabilities would not be constant over follow-up time (i.e., time post-initiation of RRT). For example, transplants among ESRD patients in Canada usually take place within the first three years of follow-up [1]. As well, switching dialytic modalities is much more frequent just after beginning on dialysis, and stabilizes thereafter. For these reasons, transition matrices were estimated separately for each 6-month follow-up window. We represent the probability of being in state *k* at time *t* + 6, conditional on being in state *j* at *t* months, by:

$$q_{jk}(t,t+6) = P(\text{state } k \text{ at time } t+6 \mid \text{state } j \text{ at time } t) \quad \text{(Eq. 4)}$$

where the "|" denotes "given" or "conditional on". We estimate these quantities by:

$$q_{jk}(t,t+6) = c_{jk}(t,t+6)/\{c_{j1}(t,t+6) + c_{j2}(t,t+6) + c_{j3}(t,t+6) + c_{j3}(t,t+6)\} \quad \text{(Eq. 5)}$$

where $c_{jk}(t,t+6)$ is the number of patients in state j at time t and state k at time t + 6, and $c_j(t)$ is the number of patients in state j at time t.

To illustrate the use of these quantities, we list some calculations. The number of patients with a functioning transplant for a given patient cohort after 12 months of follow-up could be computed, in terms of the RRT-specific patient numbers at time 6, as:

$$n_3(12) = n_1(6)q_{13}(6,12) + n_2(6)q_{23}(6,12) + n_3(6)q_{33}(6,12) \quad \text{(Eq. 6)}$$

Note the distinction between $n_j(t)$, the number of patients from a hypothetical cohort initiating RRT during a given calendar year, and $c_j(t)$ the actual number of patients on the CORR database (i.e., summed over all patient cohorts from 1981 to 1996, inclusive) upon which our analysis is based.

For notational convenience, we can write:

$$\boldsymbol{Q}(t,t+6) = \begin{vmatrix} q_{11}(t,t+6) & q_{12}(t,t+6) & q_{13}(t,t+6) & q_{14}(t,t+6) \\ q_{21}(t,t+6) & q_{22}(t,t+6) & q_{23}(t,t+6) & q_{24}(t,t+6) \\ q_{31}(t,t+6) & q_{32}(t,t+6) & q_{33}(t,t+6) & q_{34}(t,t+6) \\ q_{41}(t,t+6) & q_{42}(t,t+6) & q_{43}(t,t+6) & q_{44}(t,t+6) \end{vmatrix}$$

$$\mathbf{q}(t) = [q_1(t), q_2(t), q_3(t), q_4(t)], \text{ and}$$

$$\mathbf{n}(t) = [n_1(t), n_2(t), n_3(t), n_4(t)], \quad \text{(Eq. 7)}$$

which allows us to economically re-express quantities such as those in Eq. 6 as:

$$\mathbf{n}(t+6) = \mathbf{n}(t)\mathbf{Q}(t,t+6) \quad \text{(Eq. 8)}$$

We can further project a patient cohort through the follow-up period as follows:

$$\mathbf{n}(t+24) = \mathbf{n}(t)\mathbf{Q}(t,t+6)\mathbf{Q}(t+6,t+12) \times \mathbf{Q}(t+12,t+18) \times \mathbf{Q}(t+18,t+24) \quad \text{(Eq. 9)}$$

As an example, the following transition matrix is for Ontario non-diabetics aged 45–64, for follow-up interval 12–18 months post-initiation of RRT:

$$\boldsymbol{Q}(12,18) = \begin{vmatrix} 0.868 & 0.008 & 0.074 & 0.049 \\ 0.055 & 0.848 & 0.053 & 0.044 \\ 0.016 & 0 & 0.977 & 0.008 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad \text{(Eq. 10)}$$

Based on this matrix, for patients on HD one year post-initiation of RRT, the probability of remaining on HD at the end of month 18 equals 86.8%, while the probability of being on PD is 0.8%. The probability of having a functioning transplant at month 18 is 7.4%, and that of dying is 4.9%. Neglecting round-off error, the elements in each row sum to 1.

The following therapy assignment vector pertains to Ontario non-diabetics aged ≤44:

$$\mathbf{q}(0) = [0.584, 0.372, 0.044, 0] \qquad \text{(Eq. 11)}$$

indicating that the probability is 58.4% that such patients begin on HD, 37.2% for PD and 4.4% for an initial transplant. Note that although the 4th element of $\mathbf{q}(0)$ element is uninformative in that $q_4(0) = 0$ by definition, it is included so that the $\mathbf{q}(0)$ vector will conform with the $\mathbf{Q}$ matrices.

Projecting follow-up of incident cases

Now that the essential quantities of the projection model have been calculated, we outline how they are combined to generate prevalence projections. Denote the number of incident cases during year y, for a particular province/age/diabetes stratum by I_y, the product of the incidence rate and population size. If we can assume that ESRD incidence occurs uniformly across each calendar year, then, on average, cases will occur at the midpoint: 30/06/y. The distribution by modality of such incident cases, upon RRT initiation, is given by $I_y\mathbf{q}(0)$. After 6 months of follow-up, the distribution is given by $I_y\mathbf{q}(0)\mathbf{Q}(0,6)$; after 12 months: $I_y\mathbf{q}(0)\mathbf{Q}(0,6)\mathbf{Q}(6,12)$, and so on. As an example, consider the year 2003. Incident cases can occur any time between 01/01/2003 and 31/12/2003, and if we assume that they occur uniformly over that calendar time interval, then, on average, they will occur at 30/06/2003. Then, we project I_{2003} from 06/30/2003 to 31/12/2003 by $I_{2003}\mathbf{q}(0)\mathbf{Q}(0,6)$; from 31/12/2003 to 30/06/2004 by $I_{2003}\mathbf{q}(0)\mathbf{Q}(0,6)\mathbf{Q}(6,12)$, . . . , and from 30/06/2005 to 31/12/2005 by:

$$I_{2003}\mathbf{q}(0)\mathbf{Q}(0,6)\mathbf{Q}(6,12)\mathbf{Q}(12,18)\mathbf{Q}(18,24)\mathbf{Q}(24,30).$$

This process is repeated for all projected incident cohorts between 1997 and 2005. After summing the results across I_{1997} to I_{2005}, we have computed the contribution to (year-end) 2005 prevalence from projected incident ESRD cases.

Projecting follow-up of currently prevalent cases

Currently prevalent patients (i.e., those receiving RRT as of 31/12/1996) were categorized by RRT and length of follow-up (6 month intervals). For example, patients with 3–9 months of follow-up were classified, for projection purposes, to have been followed for exactly 6 months. Those with 9–15 months of follow-up were considered to be at month 12, and so on. Each prevalent cohort vector was then projected out to 31/12/2005 by multiplying by the appropriate set of transition matrices. To illustrate, denote the vector of patients prevalent on 12/31/1996 with 12 prior months of follow-up by: $\mathbf{p}(12) = [p_1(12), p_2(12), p_3(12), p_4(12)]$. We project this set of patients from 31/12/1996 to 30/06/1997 by: $\mathbf{p}(12)\mathbf{Q}(12,18)$; from 30/06/1997 to 31/12/1997 by: $\mathbf{p}(12)\mathbf{Q}(12,18)\mathbf{Q}(18,24)$, and so on, out to 31/12/2005. Repeating this procedure for $\mathbf{p}(0)$, $\mathbf{p}(6)$, $\mathbf{p}(12)$, $\mathbf{p}(18)$, . . . , and summing the results yields the contribution to year-end 2005 prevalence by currently prevalent patients.

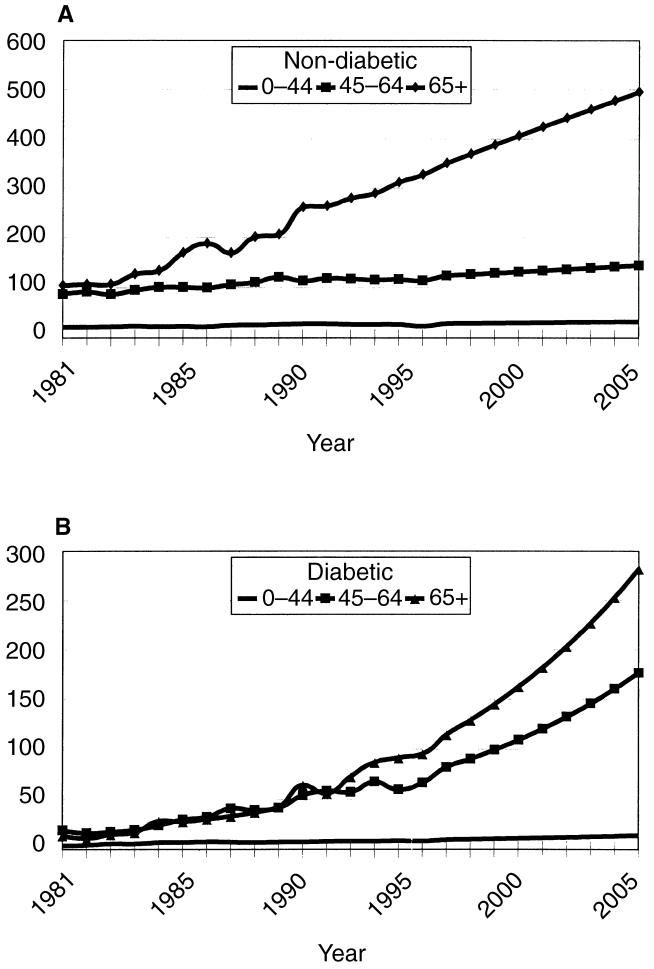

Fig. 2. Renal replacement therapy initiation rates per 10^6 population: Canada, 1981–96, and projections for 1997–2005 for non-diabetics (*A*) and diabetics (*B*). Projected incidence rates were extrapolated from Poisson regression models fitted to 1981–96 data.

RESULTS

Annual RRT initiation rates per 10^6 population in Canada by diabetes status are displayed in Fig. 2. The increase in incidence is projected to be highest for the ≥65 age group irrespective of diabetes status. Great increases are projected for the 45–64 age group only among diabetics. Incidence rates are projected to remain stable within in the ≤44 age group for both non-diabetics and diabetics.

Year-end RRT-specific ESRD prevalence (number of patients) during the 1981–2005 period is displayed in Fig. 3A. It is projected that 32,952 patients will be receiving RRT in Canada as of Dec. 31, 2005: 13,754 on HD, 6501 on PD, and 12,697 with a functioning transplant. The projected distribution by RRT for 2005 is virtually identi-

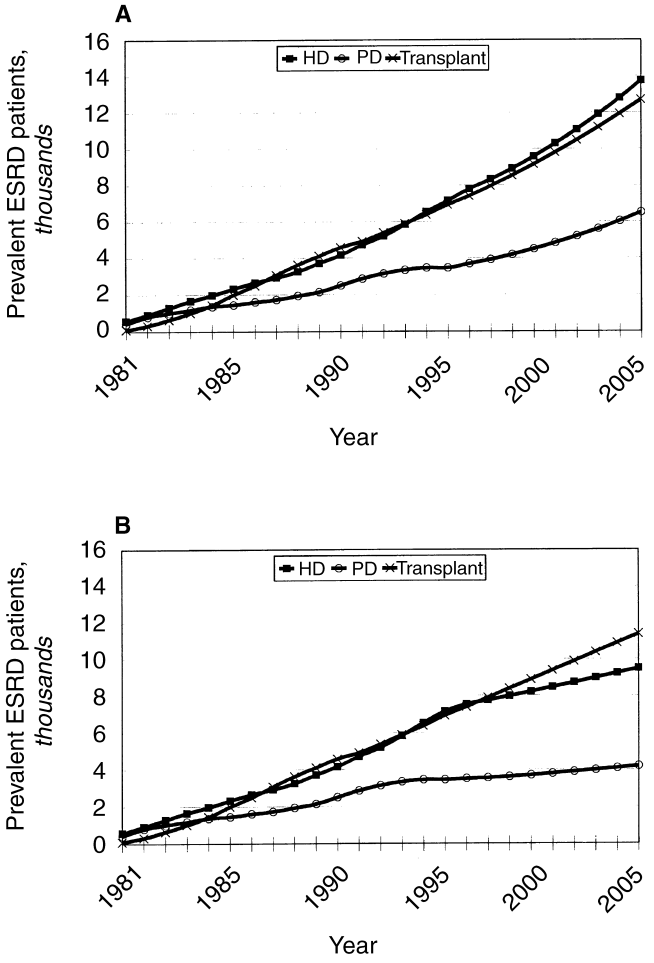

Fig. 3. End-stage renal disease prevalence (year-end number of registered patients): Canada, 1981–96, and projections for 1997–2005. The first set of prevalence projections are based on extrapolations of Poisson models fitted to the 1981–96 data (*A*). The second set assumes that incidence rates during 1997–2005 are constant and equal to the average rate experienced during 1994–96 (*B*). In both cases, a Markov model was used to project patient follow-up.

cal to that observed in 1996: 42% HD; 20% PD; 39% transplant. The average annual relative increase in prevalence projected for the 1997–2005 period is 5.8% nationally. Under the assumption of constant incidence rates during 1997–2005, it is projected that 25,065 patients in Canada will be receiving RRT in 2005, for an increase of 41% relative to 1996 national prevalence (Fig. 3B).

DISCUSSION

We present a method for projecting future numbers of patients receiving renal replacement therapy. The method was previously used to project RRT-specific ESRD prevalence in Canada nationally [6] and by province [7] up to the year 2005. Great increases in ESRD incidence and prevalence are projected, which should concern Canadian heath care planners and providers. RRT initiation rates are currently much higher in other developed countries such as the U.S., Germany and Japan [2–4] , with no obvious explanation why lower ESRD incidence should be expected in Canada. Further increases are projected for the U.S. [14].

The projection model presented is valid from a theoretical perspective [6]. Empirically, the method has proven accurate, when evaluated via data splitting, as 1981–1987 data accurately projected 1988–1994 RRT prevalence nationally [6], as did 1981–89 data for 1990–96 at the provincial level [7]. Limitations of our projection model have been discussed in detail previously [6, 7]. One noteworthy liability is that, although we account for lack of uniformity in transition probabilities over follow-up time, we have assumed constancy over calendar time. Given the supply of available donor organs has not nearly kept pace with the demand, particularly in Canada, it is likely that we over-estimate transplant probability for the 1997–2005 period. This would result in overestimation of the number of transplanted patients, underestimation of the number requiring dialysis. Another liability relates to the fact that we have data on RRT utilization, not true underlying ESRD incidence. Thus, the percentage of patients initiating RRT among the population with ESRD is not known, and is a quantity invaluable in assessing the extent to which the observed increase in RRT initiation rates could be expected to plateau. Note that, although RRT initiation rates have demonstrated no evidence of levelling off, even the supplementary "constant incidence" model yielded large increases in projected prevalence.

We present a method for projecting ESRD prevalence. Such projections are useful in forecasting future facility, equipment, manpower, and other resource requirements. Registry data, due to their volume, are invaluable for research purposes, including future health care planning, particularly when collected in the detail provided by CORR.

ACKNOWLEDGMENTS

The collection of data and the maintenance of the Canadian Organ Replacement Register (CORR) is made possible by the wholehearted collaboration of the 86 individual renal programs across Canada. The contribution of the current and past full-time staff assigned to the Register at the Canadian Institute for Health Information (formerly, Hospital Medical Records Institute) also has been essential to the success of the Register. The Canadian Society of Nephrology, The Canadian Transplant Society, and the Canadian Association of Nephrology, nurses and technicians and their constituent members also have made an essential contribution to the Register since its inception in 1981. CORR is funded 15% by the federal department of health, and 80% by the provincial and territorial departments of health based on population. Together, the Kidney Foundation of Canada and the health care industry provide approximately 5% of the funding.

Dr. Stanley S.A. Fenton, Division of Nephrology, Toronto General Hospital, 200 Elizabeth Street, Toronto, ON M5G 2C4, Canada.
E-mail: stanley.fenton@uhn.on.ca

REFERENCES

1. Canadian Organ Replacement Register, Canadian Institute for Health Information: *Annual Report 1998. Dialysis and Renal Transplantation.* Ottawa, Information, 1998
2. United States Renal Data System: *USRDS 1998 Annual Data Report.* Bethesda, MD, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 1998
3. Disney APS: *ANZDATA Report 1992.* Adelaide, Australia, ANZDATA Registry, 1992
4. Raine AEG, Margreiter R, Brunner FP, Ehrich JH, Geerlings W, Landais P, Lorait C, Mallick MP, Selwood NH, Tufveson G, Valderrbano F: Report on management of renal failure in Europe, XXII, 1991. *Nephrol Dial Transplant* 71 (Suppl. 2):7–35, 1992
5. Goeree R, Manalich J, Grootendorst P, Beecroft ML, Churchill DN: Cost analysis of dialysis treatments for end-stage renal disease. *Clin Invest Med* 18:455–464, 1996
6. Schaubel D, Morrison HI, Desmeules M, Parsons D, Fenton SSA: End-stage renal disease prevalence projections using Poisson and Markov models. *Int J Epidemiol* 27:274–281, 1998
7. Schaubel DE, Morrison HI, Desmeules M, Parsons DA, Fenton SSA: End-stage renal disease in Canada: Prevalence projections to 2005. *Can Med Assoc J* 1999:160 (11):1557–63
8. Statistics Canada: *Population Projections for Canada, Provinces and Territories: 1993–2016. Cat95–520.* Ottawa, Statistics Canada, 1995
9. Frome EL: The analysis of rates using Poisson regression models. *Biometrics* 39:665–674, 1983
10. Dobson AJ: *An Introduction to Generalized Linear Models.* New York, Chapman & Hall, 1990
11. Breslow NE: Extra-Poisson variation in log-linear models. *Appl Stat* 33:38–44, 1984
12. Ross SM: *Introduction to Probability Models.* London, Academic Press, 1989
13. Chiang CL: *Introduction to Stochastic Processes in Biostatistics.* New York, Wiley, 1968
14. Port FK: End-stage renal disease: magnitude of the problem, prognosis of future trends and possible solutions. *Kidney Int* 48 (Suppl. 50):S3–S6, 1995