ABSTRACT

Objective: In spite of increases in short-term kidney transplant survival rates and reductions in acute rejection rates, increasing long-term graft survival rates remains a major challenge. The objective here was to project long-term graft- and survival-related outcomes occurring among renal transplant recipients based on short-term outcomes including acute rejection and estimated glomerular filtration rates observed in randomized trials. Methods: We developed a two-phase decision model including a trial phase and a Markov state transition phase to project long-term outcomes over the lifetimes of hypothetical renal graft recipients who survived the trial period with a functioning graft. Health states included functioning graft stratified by level of renal function, failed graft, functioning regraft, and death. Transitions between health states were predicted using statistical models that accounted for renal function, acute rejection, and new-onset diabetes after transplant and for donor and recipient predictors of long-term graft and patient survival. Models were estimated using data from 38,015 renal transplant recipients from the United States Renal Data System. The model was populated with data from a 3-year, randomized phase III trial comparing belatacept to cyclosporine. Results: The decision model was well calibrated with data from the United States Renal Data System. Long-term extrapolation of Belatacept Evaluation of Nephroprotection and Efficacy as Firstline Immunosuppression Trial was projected to yield a 1.9-year increase in time alive with a functioning graft and a 1.2 life-year increase over a 20-year time horizon. Conclusions: This is the first long-term follow-up model of renal transplant patients to be based on renal function, acute rejection, and new-onset diabetes. It is a useful tool for undertaking comparative effectiveness and cost-effectiveness studies of immunosuppressive medications.

Keywords: decision model, end stage renal disease, modeling, renal transplantation.

Projecting Long-Term Graft and Patient Survival after Transplantation

Adrian R. Levy, PhD, Andrew H. Briggs, DPhil, Karissa Johnston, PhD, J. Ross MacLean, MD, Yong Yuan, PhD, Gilbert J. L’Italien, PhD, Anupama Kalsekar, PhD, Mark A. Schnitzler, PhD

1Dalhousie University, Halifax, NS, Canada; 2Oxford Outcomes Ltd., Vancouver, BC, Canada; 3University of Glasgow, Glasgow, UK; 4Bristol-Myers Squibb, Princeton, NJ, USA; 5Yale University School of Medicine, New Haven, CT, USA; 6Saint Louis University, St Louis, MO, USA

Introduction

Rapid increases in the incidence of end-stage renal disease and aging of the population in industrialized countries are leading to growing numbers of individuals requiring lifelong renal replacement therapy and a greater call on limited health care resources for this condition. In the United States, the incidence of end-stage renal disease doubled between 1998 and 2008, from 183 to 351 per million population [1]. Renal transplantation offers substantial benefits over long-term dialysis [2] and is the treatment of choice for end-stage renal disease [3]. The number of persons with end-stage renal disease being placed on US wait lists for kidney transplantation continues to grow each year, with approximately 99,250 candidates registered as of December 2013 [4].

The success of renal transplantation has arisen in large measure as a result of the efficacy of immunosuppressive medications. Early trials in kidney transplantation from the 1970s demonstrated that azathioprine and prednisone were efficacious for immunosuppression, leading to improvements in the end points of graft and patient survival [5,6]. The introduction of cyclosporine in 1978 led to the use of acute rejection as the primary trial end point [7-14] because graft failure and death became too rare to design realistically sized trials within reasonable periods of observation.

In spite of reductions in acute rejection [15], long-term renal graft and patient survival have not improved, and transplant researchers are shifting focus to other surrogates as end points [16]. For instance, renal functioning at 1 year posttransplant has been shown to be associated with long-term graft and patient survival [17-20] and new-onset diabetes is known to be a major complication after kidney transplant [21].

Most published decision models of immunosuppressive medications in kidney transplant are based primarily on associations between acute rejection and graft and patient survival. Although episodes of acute rejection can have deleterious consequences to the patient, be costly to treat, and increase the risk of graft failure, acute rejection alone is not a reliable predictor of long-term outcomes [22]. The specific objective of this study is to...
develop a decision model that incorporates the observed post-
transplant distributions of renal function, acute rejection, and
new-onset diabetes at the end of follow-up in randomized trials
to estimate the effect of immunosuppressant therapy on graft
and patient survival among renal transplant recipients in the
United States.

**Methods**

**Model Structure and Outputs**

The decision model includes two phases (Fig. 1). Phase 1, the
“trial period,” incorporates renal functioning (categorical esti-
mated glomerular filtration rates [eGFRs] measured in mL/min/
1.73 m²) and the probabilities of experiencing unintended and
undesirable outcomes at the end of follow-up of a randomized
trial: new-onset diabetes, acute rejection, graft failure, and death.
It is assumed that individuals with an eGFR of less than 15 mL/
min/1.73 m² are in the graft failure state.

Phase 2, the “extrapolation period,” incorporates a Markov
model to reflect 20-year follow-up of hypothetical individuals
surviving the trial period with a functioning graft. Markov models
are used widely in the health economic modeling of disease and
represent a reasonable compromise between simplicity on the
one hand, which aids transparency and understanding of the
model, and flexibility on the other, which allows key aspects of
the disease course and treatment pathways to be captured [23].
Starting with the distribution of functioning graft health states,
subjects experience declining eGFR over time [24–26], progressing
in 1-year cycles. At the end of each cycle, subjects can 1) remain
in the same state, 2) experience graft failure and return to
hemodialysis, or 3) move to the absorbing death state. Patients
are categorized at the end of the trial period into one of four
categories of renal functioning defined by the US National Kidney
Foundation [27]: eGFR greater than or equal to 60, eGFR greater
than or equal to 45 and less than 60, eGFR greater than or equal
to 30 and less than 45, and eGFR greater than or equal to 15 and
less than 30. While in the functioning graft health states, eGFR is
assumed to decline linearly until graft failure occurs at an eGFR of
less than 15 mL/min/1.73m². Following graft failure, individuals
remain in the hemodialysis state until death or regraft. A regraft
health state is included because the third most common cause
for being placed on a wait list in the United States is a previously
failed transplant [28]. Following a regraft, individuals reenter into
an undifferentiated functioning graft state and the time to graft
failure or death is based on an exponential distribution.

The Markov model is run separately for each functioning graft
health state, and the outcome measures are weighted by the
observational eGFR distribution from the trial period and
summed to obtain results. This allows flexibility because the
model can be used to project results from other studies by
incorporating information on relevant parameters from those
studies. The number of life-years spent in each graft functioning
health state is calculated once a subject enters the graft failure
state by allocating life-years assuming that eGFR declines linearly
over time starting at the entry health state and transitioning
through subsequent health states stopping at an eGFR of 15. The
time spent in each functioning graft health state is weighted by a
utility to obtain quality-adjusted life-years.

Outcomes output by model include cumulative proportions
alive with a functioning graft, time alive with a functioning graft,

**Model structure**

![Fig. 1 – Schematic of the decision model for extrapolating long-term outcomes after renal transplantation. AR, acute rejection; eGFR, estimated glomerular filtration rate; NODM, new-onset diabetes mellitus. Functioning graft categories refer to categories of renal functioning defined by the US National Kidney Foundation [27]: 2 = eGFR greater than or equal to 60; 3a = eGFR greater than or equal to 45 and less than 60; 3b = eGFR greater than or equal to 30 and less than 45; and 4 = eGFR greater than or equal to 15 and less than 30.](image-url)
life-years, and quality-adjusted life-years. The model is programmed in Microsoft EXCEL using visual basic macros and is fully parameterized, allowing different values of model parameters and one-way and probabilistic sensitivity analyses [29]. Model verification included testing for internal consistency using extensive debugging and testing extreme conditions and calibration against the source data (i.e., United States Renal Data System [USRDS]) [30].

Data Sources
Information for the model is derived from three sources. 1) All trial period outcomes from the Belatacept Evaluation of Nephroprotection and Efficacy as Firstline Immunosuppression Trial (BENEFIT) [31] were reported at 3 years (see Appendix Table, Table 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.01.001) except new-onset diabetes and the distributions in patients 3a and 3b, which were not reported in the publication. New-onset diabetes incidence was taken from the 2-year published results [32], while eGFR stage 3 was assumed to be evenly distributed between patients 3a and 3b. 2) Data for processes used to derive transition probabilities in the Markov model as well as donor and recipient characteristics of transplant patient populations not available from the phase III trial published data are from the USRDS. The USRDS is a national data system available for researchers and is made up of multiple component registries of organ donors and recipients in the United States [33]. The historical development, structure, and limitations of the data for research purposes are well described [34,35]. The USRDS makes available data on the occurrence, clinical characteristics, treatment, mortality, and survival rate of transplant recipients. This study included all kidney transplant recipients listed in the USRDS of deceased donor, single-organ transplants between 1995 and 2003, and who had Medicare as the primary payer for their transplant care. Medicare was the insurer for at least 66% of adult patients included in the USRDS between April 1995 and December 2004. The models are based on 38,015 recipients who had functioning grafts and on whom estimated eGFR could be calculated on day 366 posttransplant. Measurements of eGFR are reported annually. The list of candidate demographic and clinical variables available for use come from a published model on long-term outcomes in kidney transplant [36]. 3) Utilities for each eGFR category were from a study of 386 kidney transplant recipients aged 18 to 74 years treated in two Midwestern outpatient clinics who completed the Health Utilities Index Mark III questionnaire [37]. The utilities are as follows: 0.70 for eGFR greater than or equal to 60, 0.63 for eGFR greater than or equal to 45 and less than 60, 0.63 for eGFR greater than or equal to 30 and less than 45, 0.62 for eGFR greater than or equal to 15 and less than 30, and 0.59 for eGFR less than 15 [38]. The utility with a functioning regraft was assumed to be the arithmetic mean utility of functioning eGFR categories [38].

Renal functioning was characterized using eGFR because it is a more accurate measure than serum creatinine [27]. The eGFR was computed according to the abbreviated Modification of Diet in Renal Disease equation [39] as follows: eGFR (mL/min/1.73 m²) = 186 × (Serum Creatinine mg/dL) − 1.154 × Age − 0.203 × (1.212, if African-American) × (0.742, if Female). The abbreviated Modification of Diet in Renal Disease equation has been validated as a measure of renal functioning among kidney transplant recipients [40].

Transition Probabilities
The transition probabilities in the extrapolation period are derived from a series of processes that are modeled using different types of parametric regression models fit using USRDS data (Appendix Table in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.01.001) [41]. For those subjects entering the extrapolation period, the annual probabilities of experiencing graft failure or death, respectively, are estimated using Weibull models. These models are parametric regression models, which allow extrapolation beyond the period of observation and allow the likelihood of a transition to change with time under observation [42]. Once the graft fails, the time spent in the dialysis health state is calculated using an exponential model because the Markov assumption does not allow the transition probability of this health state to depend on the timing of graft failure [43]. The USRDS data did not include information describing new-onset diabetes, and the effect of new-onset diabetes on graft failure and death, respectively, was incorporated into the model via published hazard ratios [21].

Target Population
The target population is US Medicare-covered kidney transplant recipients. The statistical models are based on a historical cohort of patients receiving organs from deceased donors and treated predominantly with calcineurin-inhibitor-based immunosuppression [44]. To account for the observation that eGFR decline was slower in 2003 than in 1995, a variable for calendar year was included in the published models.

Treatments
In this case, the treatments included were cyclosporine (any formulation) and belatacept (Nulojix). The two immunosuppressants were compared directly in BENEFIT, a 3-year, randomized, active-controlled, parallel-group, multicenter phase III study conducted at 100 centers worldwide, enrolling standard criteria donors [31,45]. Belatacept received approval from the US Food and Drug Administration on June 15, 2011, to prevent acute rejection in adult patients who have had a kidney transplant.

Sensitivity Analysis
Stochastic distributions were applied to uncertain parameters, allowing for results to be varied in a probabilistic sensitivity analysis. The variability was based on observed variability in trial outcomes [31,45] and regression equation coefficients [44].

Results
The Appendix Table in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.01.001 shows the risk equations in the underlying Markov model predicting graft and patient survival modeled in the extrapolation phase using the data from the USRDS. In addition to regression coefficients, hazard ratios associated with new-onset diabetes were 1.46 for death-censored graft failure and 1.87 for death [21]. The probabilities describing the transitions from an initial functioning graft to graft failure or death were dependent on time and stratified by eGFR at 1 year posttransplant and are displayed in Figure 2A,B. The probabilities over time for graft failure and death were inversely related to eGFR category, although minimal differences in survival probabilities were observed across the two highest eGFR categories. The regression equations for death were censored by graft failure, and as such do not represent the probability of overall mortality but rather the probability of mortality conditional on having a functioning graft. As shown in Figure 2A, the long-term probability of having a functioning graft decreased sharply after 20 years, and the subsequent mortality probabilities were relevant only to the minority of individuals remaining in the functioning graft state.

Transition probabilities estimated via exponential equations can be summarized by a single time-invariant probability: the
the annual probability of transitioning from dialysis to retransplantation was 0.04; the annual probability of dying while on dialysis was 0.12; the annual probability of experiencing graft failure with a functioning regraft was 0.10; and the annual probability of dying with a functioning regraft was 0.06.

Among graft recipients in BENEFIT, demographic and baseline characteristics were well balanced between groups and comparable to US population [46]. The mean age was 43 years, and 70% were men. At 3 years posttransplant, a higher proportion of patients in the cyclosporine group had died or suffered graft failure than in the belatacept group, while a higher proportion of patients in the belatacept group had acute rejection, most of which occurred in the first 6 months posttransplant (Table 1). The mean eGFR was 21.4 mL/min/1.73 m² higher in the belatacept group, yielding a distribution of eGFR categories shifted to the higher functioning categories.

The calibration of the decision model was done by comparing the model-predicted survival to Kaplan-Meier survival curves fit to USRDS data over the first 5 years posttransplant (Fig. 3). Because of the large number of individuals included within the USRDS (n = 38,015), the 95% CIs for Kaplan-Meier USRDS-predicted survival display little variability, with upper and lower bounds within 2% of each other throughout the 5-year period. For the first 3 years, model-predicted survival is within 0.5% of the upper bound of the 95% CI and in the final 2 years, model-predicted survival falls between the 95% CI bounds. The existing, relatively minor, discrepancies are likely due to deviations in continuous covariates from the normal distribution.

Using the transition probabilities derived from the USRDS equations, the long-term extrapolation phase of the model indicated that the difference in the proportions of recipients suffering graft failure and having a regraft was lower in the belatacept group at all points in time (data not shown). The differences at 3 years in eGFR are projected to avert graft failure (57% cumulative graft failures projected for the belatacept group relative to 65% for the cyclosporine group over 20 years) and yield 2.5 incremental years alive with a functioning graft (95% empirical CI 1.3–3.4), a 10.1% increase in patients surviving over a 20-year time horizon, and a corresponding incremental 1.5 life-years (95% CI 0.9–1.9) and 1.2 quality-adjusted life-years (95% CI 0.8–1.6). Of the time alive with a functioning graft, a greater proportion of time, for example, 46% (5.2 years of 11.4 with a functioning graft) for belatacept versus 22% (2.0 years of 8.9 with a functioning graft) for cyclosporine, was projected to be spent with renal functioning of 60 mL or more.

At every point in time, the cumulative proportions of individuals with a functioning graft (Fig. 4 and Table 2) and which
survive (see Appendix Figure in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.01.001) are higher among persons receiving belatacept than cyclosporine.

Discussion

In the coming years, novel immunosuppressive medications may offer the possibility of preserving renal function [45]. Characterizing the therapeutic value of these medications, in terms of either comparative effectiveness or incremental cost-effectiveness, will require long-term projections of graft and patient survival. The model described in this article permits the lifetime extrapolation of long-term outcomes given any distribution of eGFR categories, acute rejection, and new-onset diabetes. The model incorporates important demographic and clinical characteristics of recipients during the early posttransplant trial period that can affect graft- and survival-related outcomes. It can be extended in a straightforward fashion to include longer trial durations and include other immunosuppressive medications including existing agents such as tacrolimus and sirolimus, or novel agents under development. Sufficient information is provided to allow other investigators to recreate the model using standard software (e.g., Microsoft EXCEL and Treeage).

Extensive verification processes demonstrate that inputs and outputs of this model are consistent with known facts regarding renal transplant. Given the high precision of the Kaplan-Meier analysis (Fig. 3), the level of consistency between the two sets of curves provides considerable assurance that the model is well calibrated. In the future, the model’s validity may be tested with longer-term follow-up of the trial populations. The probabilistic structure of the model is strength insofar as it allows statistical comparisons of projections to observed findings [30].

A published two-stage literature-based model has been used to estimate the economic impact among kidney transplant recipients with lower renal function by one measure of renal function, serum creatinine [47]. The current model is an important step forward in that: it uses renal functioning expressed as eGFR; transition probabilities are based on primary data analysis using outcomes data on actual transplant recipients; and important predictors of graft and patient survival, such as acute rejection in the trial period and new-onset diabetes, are included. More generally, predicting long-term outcomes from intermediate end points is often of interest to clinicians and researchers. Except in rare circumstances such as those provided by the Framingham studies, models that incorporate both trial data and long-term follow-up are required [48]. The current model applies those techniques using high-quality data from the USRDS registry. This application serves as an example for decision models in other areas of surgery.

Important limitations exist regarding the model, including those related to the structure and data inputs. There are at least five structural features of the model that may have an important impact on the results and interpretations. First, the underlying premise of the model is that early survival of patients with higher renal functioning predicts late survival. While current evidence indicates that this premise is correct [17–20], longer follow-up of observational cohorts will provide a better characterization of the

### Table 1 – Clinical outcomes and distributions of eGFR (mL/min/1.73m²) at 4 y posttransplantation among graft recipients randomized to receive cyclosporine or belatacept (less intensive regimen) in BENEFIT [47].

<table>
<thead>
<tr>
<th></th>
<th>Cyclosporine (n = 221)</th>
<th>Belatacept (n = 226)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipients surviving with functioning graft and GFR ≥ 15 (%)</td>
<td>73.7</td>
<td>82.6</td>
</tr>
<tr>
<td>Death (%)</td>
<td>6.8</td>
<td>4.4</td>
</tr>
<tr>
<td>Graft loss (%)</td>
<td>19.5</td>
<td>12.3</td>
</tr>
<tr>
<td>Acute rejection (%)</td>
<td>10.0</td>
<td>17.0</td>
</tr>
<tr>
<td>New-onset diabetes (%)</td>
<td>11.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Mean eGFR ± SD</td>
<td>44.4 ± 23.6</td>
<td>65.8 ± 27.0</td>
</tr>
<tr>
<td>Distribution of eGFR at 3 y posttransplant (%)</td>
<td>29.0</td>
<td>71.0</td>
</tr>
<tr>
<td>≥60</td>
<td>25.0</td>
<td>10.0</td>
</tr>
<tr>
<td>≥45 and &lt;60</td>
<td>25.0</td>
<td>10.0</td>
</tr>
<tr>
<td>≥30 and &lt;45</td>
<td>6.0</td>
<td>0.0</td>
</tr>
<tr>
<td>≥15 and &lt;30</td>
<td>15.0</td>
<td>9.0</td>
</tr>
<tr>
<td>&lt;15</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

BENEFIT, Belatacept Evaluation of Nephroprotection and Efficacy as Firstline Immunosuppression Trial; eGFR, estimated glomerular filtration rate.
relationship. Second, the model includes benefits observed only in the trial period. If the separation in GFR between the treatment groups continues to increase beyond that which was observed in the 3-year follow-up period in BENEFIT [31], the long-term estimates of clinical benefit projected here may represent the lower limit of incremental benefit for new treatments. Third, posttransplant eGFR is assumed to reach its zenith at 3 years and then decline equally in both groups. In months 3 to 36 of BENEFIT III, the mean calculated eGFR increased in the belatacept arms by +1.0 mL/min/1.73 m²/y (more intensive regimen) and +1.2 mL/ min/1.73 m²/y (less intensive regimen) versus a decline of −2.0 mL/min/1.73 m²/y in the cyclosporine arm [31]. Fourth, in the functioning graft health state, subjects are allowed to stay only in the same eGFR category or a lower one. The results are likely to be robust to minor violations of this assumption because subjects are grouped into eGFR categories. Fourth, the variables included in the predictive models are limited to those available in the USRDS. Because new-onset diabetes following transplant is known to have a large impact on long-term outcomes [49], it was included using other published information [21]. There may be other potential confounders, however, which are not included.

Conclusions

This is the first long-term follow-up model of renal transplant patients to be based on graft function, acute rejection, and new-onset diabetes and can incorporate data from clinical trials of newly approved treatments in renal transplant. It is a useful tool for undertaking comparative effectiveness studies of new immunosuppressive medications. By incorporating resource utilization and costs, it can also be used to carry out methodologically sound economic evaluations in renal transplant.

Acknowledgments

We are grateful to the following individuals for their contributions: Patricia Corey-Lisle, Sarah Goring, Sandra Joshua-Gotlib, Bertrand Kasiske, Jon Snyder, and Vickie Tuomari.

At the time this study was carried out, A.R.L. and A.H.B. were shareholders and directors, and K.J. was an employee, of Oxford Outcomes Ltd. R.M., Y.Y., A.K., and G.J.L. are employees of Bristol Myers Squibb.

The data reported here have been supplied by the United States Renal Data System. The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the United States government.

Source of financial support: This study was funded by Bristol-Myers Squibb.

Supplemental Materials

Supplemental materials accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2014.01.001, or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).
REFERENCES


