Mispredicting and Misremembering: Patients With Renal Failure Overestimate Improvements in Quality of Life After a Kidney Transplant

Dylan Smith
University of Michigan

Christopher Jepson
University of Pennsylvania

Harold Feldman
University of Pennsylvania

George Loewenstein
Carnegie Mellon University

Aleksandra Jankovich
University of Michigan

Peter Ubel
University of Michigan

Objective: People tend to overestimate the impact that future events will have on their quality of life. In the case of a medical treatment like kidney transplant, this should result in biased predictions—overestimates of how much the transplant will benefit quality of life. The authors surveyed kidney transplant patients, both before and after transplant, to test whether they would overestimate the benefits of a successful transplant for their quality of life. Design: The authors interviewed 307 patients on a waiting list for cadaveric renal or renal-pancreatic transplant, and 195 patients one year after a successful transplant. A sub sample of patients were interviewed both before and after transplant. Main Outcome Measures: The survey included measures of quality of life, both in terms of an overall estimate (0–100), and across sub domains, including health, employment, and travel. Results: Cross-sectional results suggested that overall quality of life improved after transplant, but the predictions of pretransplant patients overestimated the magnitude of the improvement (p < .01). In addition, patients predicted large improvements in specific life domains that did not change. These results were confirmed in longitudinal, prospective analyses. Additional analyses showed that posttransplant patients recalled their pretransplant quality of life to be much lower than what they had reported at the time. Conclusion: Consistent with an impact bias, patients substantially overestimated the benefits of a successful kidney transplant, both in terms of predictions of life after treatment, and in their memories of QOL before the transplant.

Keywords: affective forecasting, impact bias, quality of life, renal transplant

When confronted with a choice about whether to undergo a treatment, patients naturally rely on judgments about how the treatment will affect their overall quality of life (QOL). These judgments are based on predictions about how the changes in their health status caused by the treatment will affect various aspects of their lives. Little is known about whether such predictions are accurate, or whether patients systematically mispredict the benefits of treatments for their subsequent QOL. However, based on prior research, there is reason to suspect that patients will overestimate the benefits of a given treatment on their QOL. Research in psychology demonstrates an “impact bias”—a tendency to overestimate the influence that events will have on one’s QOL (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). Consistent with this possibility, numerous studies have indicated that people tend to believe that a disability will have a greater impact on QOL than is reported by disabled individuals (Riis et al., 2005; Smith, Sherriff, Damschroder, Loewenstein, & Ubel, 2006; Ubel, Loewenstein, Schwarz, & Smith, 2005). For example, members of the general public often think that a disability will result in much lower quality of life than what is reported by patients who actually have the disability (Smith et al., 2006). These community members may be engaging in a misprediction because of impact bias—they expect an overly large decrement in their QOL after onset of a disability. Applying this logic to a positive event, such as a successful medical treatment, impact bias might cause people to predict a larger improvement in QOL after a treatment than what actually occurs.
In a similar vein, after a successful treatment, patients’ appraisals of the treatment’s effectiveness are partly based on recollections of their QOL before the treatment, and an estimate of how much QOL has improved. Again, these recollections could be biased. For example, research has shown that memories of emotional well-being are often not reliable. Instead, when asked to remember an emotional reaction to some event, people often rely on a theory about how the event should have affected their emotions—a phenomenon called “theory based recall bias” (Ross, 1989; Wilson, Meyers, & Gilbert, 2003). Consistent with this idea, a prospective study of 22 renal transplant patients showed that recollections of QOL before transplant were substantially lower than what the patients had reported at the time (Adang, Kootstra, Engel, van Hoooff, & Merckelbach, 1998).

Like mispredictions because of impact bias, a tendency to believe that life before treatment was worse than it actually was would result in overestimates of a treatment’s effectiveness. This would have important implications for patient decisions, as previous experiences with treatments likely affect future decisions in some health contexts. For example, when deciding whether to undergo a second knee replacement, patients may rely on their estimate of the benefits of the first replacement. In addition, some research designs rely on retrospective appraisals to document the value of treatments in improving QOL (Ahmen, Mayo, Wood-Dauphinee, Hanley, & Cohen, 2005). To the extent that recollections are systematically biased, these studies may overestimate the effectiveness of treatments for QOL.

In the current investigation, we asked patients currently on a waiting list for a kidney transplant to report their current overall QOL, and also to predict their QOL after receiving a successful transplant. We asked for both an overall QOL estimate, and for estimates in a number of specific life domains, such as health, employment, and personal relationships. In a cross-sectional analysis, we compared these predictions to the current QOL of a group of successful transplant recipients.

In addition, we asked posttransplant patients to remember what their QOL was like before their transplant, when they were on the wait list. By comparing their remembered QOL to the actual QOL reported by pretransplant patients, we were able to test the possibility that recalled QOL could be biased, again resulting in overestimation of the positive effects of transplant.

We also performed longitudinal comparisons using a subsample of participants. After some of the wait-listed patients had received a transplant, we again asked them for their current QOL. This procedure allowed us to make within-subject comparisons of the actual improvement in perceived QOL, compared to the predictions and recollections of the patients, and thus test the possibility that patients would overestimate the positive effects of transplant.

We tested three hypotheses. First, based on previous findings, we predicted that reported health and overall QOL would increase after transplant (Park et al., 1996). Second, we predicted that patients’ forecasts would overestimate the amount of improvement in QOL. Finally, we predicted that, after transplant, patients would remember their pretransplant QOL as being worse than it actually was (Adang et al., 1998).

Participants

We invited 460 patients from the waiting list for cadaveric renal or renal-pancreatic transplant at the Hospital of the University of Pennsylvania. A total of 77 were determined to be ineligible for various reasons. Most of these (n = 41) had already received their transplant by the time we contacted them; other reasons for ineligibility included death, hospitalization, or a language barrier. Of the remaining 383 individuals, 307 (80.2%) completed an interview.

Procedures

Interviews were conducted by a professional survey firm using a computer-assisted telephone interview technique. The interview began with questions about kidney failure history; how long the patient had been waiting for a transplant and how they were being treated for their kidney problems. Respondents were next asked to rate the quality of their current lives using a 0 (worst imaginable) to 100 (best imaginable) scale (QOL rating).

The QOL rating was followed by the SF-12, a brief measure of physical and mental health functioning (Ware, Kosinski, & Keller, 1996). Patients were then asked a set of three questions, developed for this study, asking about three aspects of functioning that are known to be influenced by kidney failure:

How many days have you spent overnight in the hospital in the past year?

In the past year, how many days did you spend traveling outside of town, where you traveled far enough to stay overnight?

How many hours a week are you currently working as a paid employee?

After completing the above items, pretransplant participants were asked to answer the 0-to-100 QOL item again, but this time with respect to what they thought their life would be like one year after a successful transplant. In addition, they were asked to predict four of the SF-12 items (general health, limited in moderate activities, calm and peaceful, and downhearted and blue), and the three items about functioning (hospitalization, travel, employment).

Posttransplant Interviews

To obtain posttransplant data, we combined our study with a separate study on posttransplant outcomes that was being conducted at the same medical center (this other study focused on adherence to taking posttransplant medications). We added our measures to the end of the interviews, and these were administered to consenting participants. In addition to answering the questions described above about their current QOL, posttransplant participants were also asked to answer the QOL question (0–100) with respect to what their life had been like 1 year before their transplant.

Patients were enrolled in this study soon after transplant (most within a few weeks, up to 2 months). Baseline measures—including our measures—were collected at this time, and again at 6 and 12 months after transplant (Weng et al., 2005).
This posttransplant adherence study was conducted over a period of several years. Eventually, our measures were administered to a total of 429 patients. Three hundred and twelve patients completed the measures at 6 months, and 210 at 12 months. Reasons that patients did not complete these latter measures included death \((n = 26)\), transplant failure and other health problems \((n = 23)\), voluntary withdrawal \((n = 15)\), and a number of technical factors related to the administration of the posttransplant study \((n = 170)\). These factors included problems monitoring patient medication adherence and the early termination of the study because of lack of funding.

Successful transplant recipients were aged 48 years on average. Forty-one percent were female, 33% non-White, 91% had completed high school, and 51% had some college education.

**Longitudinal analyses.** Because the posttransplant interviews were conducted over a period of several years, eventually a number of the patients we interviewed before their transplant happened to be enrolled in the posttransplant study \((n = 75)\). A subset of these patients remained in the study long enough to provide 12 month data \((n = 33)\). As with the cross-sectional data, participant loss was due partly to factors such as transplant failure and illness \((n = 4)\), and voluntary withdrawal \((n = 7)\), but also to factors related to the administration of the posttransplant study, especially the early termination \((n = 8)\) and other problems such as adherence measurement \((n = 21)\).

Thus, the overlap in time frame (and location) of the two separate studies created a circumstance wherein we could perform within-subjects analyses on a small subsample of patients, directly comparing their expectations of QOL after transplant with what they actually experienced—and comparing their recollections of QOL before the transplant with what they reported at the time.

### Results

**Cross-Sectional Analyses**

We first compared the current QOL \((0–100)\) reported by pretransplant patients to their predicted QOL 12 months after a successful transplant, using a within-subjects \(t\) test. As shown in the first row of Table 1, pretransplant patients predicted a posttransplant QOL that was substantially higher than their current QOL \((within-subjects \(t(297) = 22.11, p < .001; d = 1.28)\). QOL was indeed higher after transplant \((between-subjects \(t(509) = 8.23, p < .001; d = .68)\). However, the predictions of pretransplant QOL patients were also significantly higher than the QOL reported by posttransplant patients \((between-subjects \(t(510) = 9.72, p < .001; d = .76)\).

Table 1 also displays the results for the various specific QOL domains. The pattern observed for the overall QOL estimates was mirrored in some cases, with the exceptions of “travel,” “hours worked,” “hospital days,” and “family life.” For all of these variables, pretransplant patients predicted a substantial improvement, but in fact, the level was not different before versus after transplant. These items, therefore, showed an even greater bias than the others, such that patients predicted improvement where no observable improvement occurred.

For only one variable, “limited in social activity” (reverse coded), did patients correctly predict an improvement in functioning—that is, the predicted posttransplant value was higher and so was the actual posttransplant value, by a similar amount.

In summary, the results of the cross-sectional analyses are consistent with the idea that patients overestimate the impact of successful kidney transplant on QOL. However, these analyses must be considered preliminary, as their cross-sectional nature complicates our ability to interpret the results. Fortunately, as outlined in the methods, we were able to conduct longitudinal analyses in a group of participants who provided both pretransplant and posttransplant data.

**Longitudinal Analyses**

Because the group of patients who provided both pretransplant data and 12 month posttransplant data is a subset of the overall sample, we began by conducting attrition analyses. Specifically, we wanted to see if this group differed from the rest of the sample in terms of their baseline (pretransplant) QOL, or in their predictions for posttransplant QOL. We compared the 33 patients who

### Table 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pretransplant QOL ((n = 307))</th>
<th>Predicted posttransplant QOL ((n = 307))</th>
<th>QOL at 12 months ((n = 210))</th>
<th>Effect size (Cohen’s (d)) for predicted versus actual post score</th>
</tr>
</thead>
<tbody>
<tr>
<td>QOL ((0–100))</td>
<td>64.1 (20.5)(^1)</td>
<td>90.4 (12.5)(^2)</td>
<td>78.1 (16.2)(^3)</td>
<td>.76</td>
</tr>
<tr>
<td>Travel (days per year)</td>
<td>7.3 (12.6)(^1)</td>
<td>27.0 (38.5)(^2)</td>
<td>9.3 (17.7)(^1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Work (hours per week)</td>
<td>11.9 (18.1)(^1)</td>
<td>29.5 (18.2)(^2)</td>
<td>12.3 (18.5)(^1)</td>
<td>.93</td>
</tr>
<tr>
<td>Hospital (days per year)</td>
<td>9.6 (18.8)(^1)</td>
<td>1.6 (4.2)(^2)</td>
<td>9.2 (17.2)(^1)</td>
<td>.44</td>
</tr>
<tr>
<td>Rate current health ((1–5))</td>
<td>2.7 (1.0)(^1)</td>
<td>4.0 (0.8)(^2)</td>
<td>3.4 (0.9)(^1)</td>
<td>.67</td>
</tr>
<tr>
<td>Limit on mod. activity(^*) (1–3)</td>
<td>2.1 (0.8)(^1)</td>
<td>2.6 (0.6)(^2)</td>
<td>2.3 (0.7)(^1)</td>
<td>.43</td>
</tr>
<tr>
<td>Had a lot of energy ((1–6))</td>
<td>3.1 (1.4)(^1)</td>
<td>5.1 (0.9)(^2)</td>
<td>3.8 (1.4)(^1)</td>
<td>.93</td>
</tr>
<tr>
<td>Limited social activity(^*) (1–5)</td>
<td>3.6 (1.2)(^1)</td>
<td>4.4 (0.9)(^2)</td>
<td>4.3 (1.0)(^2)</td>
<td>.10</td>
</tr>
</tbody>
</table>

* These items are coded such that higher values indicate better functioning.

Note. Standard deviations in parentheses. Mean values not sharing a common superscript are significantly different from one another at \(p < .05\).
provided 12 month data to the rest of the baseline sample (n = 274). We found no differences in terms of QOL at baseline (63.9 for noncompleters vs. 65.9 for completers, p > .50; d = .10), or in predictions of QOL at 12 months (90.3 vs. 91.4 for completers, p > .50, d = .07). Thus, we found no evidence that the participants who provided 12 month data were different from the larger pretransplant sample on our primary variables of interest.

Table 2 makes the same comparisons as Table 1, but this time all of the comparisons are within-subject. The results are highly consistent with what we observed in the larger cross-sectional analyses. QOL increased after transplant (within-subjects t(32) = 4.40, p < .001; d = .82), but patient predictions of QOL significantly overestimated this increase (within-subjects t(32) = 3.12, p < .01; d = .79). Pretransplant predictions on more specific domains were also biased high, and some, especially days of travel and days in hospital, were substantially more biased than the general QOL ratings.

In Table 3, we see that after a successful transplant, patients underrecalled their pretransplant QOL (for 12 months after surgery, within-subjects t(32) = 3.27, p < .01; d = .87). Interestingly, this tendency increased as time increased from transplant, even as their current QOL also increased. As a result, by 12 months after transplant, patients recalled a total improvement in their QOL of 35 points, approximately double the observed improvement of 17 points.

**Discussion**

In examining judgments about QOL before and after a successful renal transplant, we observed patterns that were consistent with an impact bias. Patients experienced an increase in their QOL after a transplant, but they substantially overestimated the magnitude of this increase, both in terms of predictions made before the transplant, and in terms of remembered QOL after the transplant. This is the first study, to our knowledge, to demonstrate that predictions about a treatment that improves health are subject to impact bias. The prediction bias revealed by our study reinforces the conclusion of prior studies that non patients’ estimates of the quality of life associated with medical conditions should not form the basis of public policy (Lacey et al., 2006; Smith et al., 2006; Ubel, Loewenstein, Schwarz et al., 2005). However, we do not think it can easily account for all our current results, especially because some of the largest mispredictions occurred on the subjective, 0–100 QOL scale differently before versus after transplant, rescaling the values to fit their circumstances. This would tend to limit the ability of the scale to quantify real improvements in QOL. In this case, there might not be any over prediction; instead, we may be underestimating the magnitude of the real improvement in QOL. We and others have argued elsewhere that scale recalibration is an important issue for measuring health-related QOL, demanding further study (Schwartz & Sprangers, 1999; Ubel, Loewenstein, Schwarz et al., 2005). However, we do not think it can easily account for all our current results, especially because some of the largest mispredictions occurred on scales using objective units that would be resistant to rescaling (e.g., hours at work, days traveling).

Table 2

<table>
<thead>
<tr>
<th>Longitudinal Comparisons: Predicted and Actual Mean QOL Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual pretransplant QOL</strong></td>
</tr>
<tr>
<td>QOL (0–100)</td>
</tr>
<tr>
<td>Travel (days per year)</td>
</tr>
<tr>
<td>Work (hours per week)</td>
</tr>
<tr>
<td>Hospital (days per year)</td>
</tr>
<tr>
<td>Rate current health (1–5)</td>
</tr>
<tr>
<td>Limit on mod. Activity (1–3)</td>
</tr>
<tr>
<td>Had a lot of energy (1–6)</td>
</tr>
<tr>
<td>Limited social activity (1–5)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations in parentheses. Mean values not sharing a common superscript are significantly different from one another at p < .05. *These items are coded such that higher values indicate better functioning.
Overfocusing or Overgeneralizing?

Some researchers have suggested that mispredictions about how some event will affect QOL are because of “focusing illusion,” wherein people are overly focused on the area of life directly influenced by the event, and fail to consider that other important domains will be unaffected (Schkade & Kahneman, 1998; Ubel, Loewenstein, & Jepson, 2005). However, our data suggest that at least when considering an event that will improve health, people may overgeneralize the improvement that would occur in health status to other areas. Indeed, we found that some of the most optimistic predictions for improvements were in areas that did not change at all after transplant, such as employment and travel. It would be interesting to see if this “spreading” effect also occurs when making predictions about negative events, such as the onset of a disability. In such cases, do people merely focus too much on the disability itself, or do they believe that the disability will negatively affect areas of life that will actually be relatively unaffected?

Methodological Considerations

As mentioned in the results section, interpretation of the cross sectional analyses is complicated by the fact that there are substantial differences between people who are waiting for a kidney and people who have received a successful transplant. Successful transplant recipients tend to be younger and healthier, with fewer co-morbid conditions. This is no accident—these factors predict success, and the decision to match available kidneys with potential recipients on a transplant wait lists takes such factors into consideration, precisely for this reason. In our sample, for example, transplant patients who received a successful transplant spent less than half as many days in the hospital in the year before the pretransplant survey, compared to the rest of the pretransplant sample (4 vs. 11 days, p < .01). Nevertheless, we do not believe that this creates major problems in interpreting our results. If the posttransplant group is healthier, on average, than the pretransplant group, then their posttransplant QOL scores would, if anything, be higher than would be expected if everyone received a transplant. Yet, we found that the posttransplant QOL scores were lower than the predictions of pretransplant patients. Therefore, it is plausible that we are actually underestimating the amount of misprediction, by comparing pretransplant patients to relatively healthier posttransplant patients.

But there may some other reason that the pretransplant patients, as a group, were happier and had higher expectations than those who actually went on to get a transplant. To address this possibility, we turn to the sub sample of participants who provided prospectively longitudinal data. Here, we can directly compare the variables of interest (current QOL and predictions about QOL) for a subset of pretransplant patients who received a successful transplant to the rest of the pretransplant sample. We found no differences, comparing the 75 who received a transplant and provided some posttransplant data to the rest of the pretransplant sample (likewise with the 33 who gave us data one year after transplant, all p’s > .5). Thus, we found no evidence that patients who got a successful transplant were different in terms of their overall pretransplant QOL, or in predictions about QOL after transplant.

Last but not least, we have the longitudinal analyses themselves, which are not subject to demographic or psychological confounding, and which perfectly replicate the cross sectional analyses. Of course, this is a relatively small sample, but the hypothesized effects are all significant, and, as described previously, this sub sample does not differ from the full sample on the key variables of interest, so there do not appear to be problems with either statistical power or generalizability.

Conclusion

Overall, the findings provide new support for the generality of the impact bias. In addition, we found evidence for an impact bias not only in patients’ predictions about subjective moods, where such biases have been previously demonstrated, but also for objective features of people’s lives, such as how much they will work and travel.

References


