Monitoring Kidney Transplant Patients*

A. F. M. SMITH†, M. WEST‡, K. GORDON§, M. S. KNAPP$ and I. M. G. TRIMBLE§

Department of Mathematics, University of Nottingham, Nottingham†;
Department of Statistics, University of Warwick‡; Renal Unit, City Hospital, Nottingham§

1 Introduction

In the first two months following a kidney transplant, there are almost invariably "rejection episodes," when the patient's body reacts negatively to the presence of the transplanted organ, resulting in a deterioration in functioning of the latter. Clinicians are therefore concerned to find some method of monitoring patients in order to detect sudden changes in the performance of the new kidney. Renal function is indicated by the glomerular filtration rate (GFR), which measures the rate at which various substances are cleared through the kidney, but GFR is not itself directly observable and so clinicians must attempt to infer something about it on the basis of blood or urine concentrations of chemicals for which kidney physiology suggests a simple form of relationship with GFR.

This particular study is based on plasma creatinine measurements (Knapp et al., 1977). In a normally functioning kidney, GFR is constant and creatinine is excreted at a constant rate. In a recently transplanted kidney, typically fluctuating between periods of deterioration and improvement in functioning, GFR and plasma creatinine are inversely related. If GFR decreases, the observed concentration of creatinine will increase, and vice versa. In theory, therefore, by monitoring an observed plasma creatinine series for sudden changes it should be possible to infer the occurrence of significant underlying biological changes.

In practice, messages in observed plasma creatinine series are obscured by a combination of factors which together add considerable "noise" to the underlying process. In part, these derive from biological variation and errors arising in the collection, measurement and processing of the data. In addition, however, there is more than one type of abrupt, discontinuous change in pattern which might occur and it is important to be able to distinguish these different forms of change. For this reason, a statistically based, automatic monitoring procedure requires considerable care in modelling the underlying system and observation processes. In particular, it requires the incorporation of substantial prior information based on physiological and clinical expertise.

In section 2, we outline the model building process leading to a representation of the evolution of the plasma creatinine series in steady periods (either of improvement or deterioration). This model is extended, in section 3, to incorporate the various forms of discontinuity which may occur in the process. Section 4 summarizes the implementation of the monitoring procedure and section 5 outlines the way in which this procedure is used as a decision aid for clinicians.

The emphasis in this paper is on the modelling process. Detailed mathematical developments are given in Smith and West (1983), detailed discussion of the clinical background and

validation of the procedure is given in Trimble et al. (1983). Some earlier, related, analysis is given in Smith and Cook (1980).

2 Modelling steady evolution

Part of a sequence of plasma creatinine measurements from a patient who had undergone transplantation 10 days before is shown in Figure 1.

![Graph showing plasma creatinine values over time](image)

**Fig. 1.** Observed plasma creatinine values before transformation.

As a first step in the process of introducing prior information, we recall that basic kidney physiology requires an inverse relationship between GFR and creatinine, so that information about the former is more directly contained in the reciprocal creatinine values. A plot of these, derived from the values in Figure 1, is shown in Figure 2. The resulting pattern is still rather noisy, but there is now a suggestion that a "change in direction" has occurred in the latter part of the series.

However, there is another, common-sense, piece of information which has not yet been incorporated. When a chemical is measured in concentration the measured value will reflect both the absolute amount of chemical present and the amount of fluid in which it is carried. This latter is potentially important in the case of patients with a recent transplant, since their body fluid is particularly prone to rapid and significant changes. Using changes in body weight as a proxy indicator for changes in body fluid (see Knapp et al., 1977), weight corrected reciprocal creatinine values can be plotted. Comparison of the three plots in Figures 1 and 2 shows the importance of working on a transformed scale which takes into account available prior information (a point emphasized by Gore, 1981).

The pattern which now emerges clearly indicates a rejection episode around day 24 (a fact confirmed by other clinical tests). Moreover, and most importantly from the point of view of developing a model, actual weight corrected reciprocal creatinine values can reasonably
be regarded as evolving, during steady periods (of improvement or deterioration), as straight lines (allowing for the fact that Figure 2 is based on observed values, which have overlaid an element of noise on to the pattern of evolution of the actual values). Switches from improvement to deterioration, or vice versa, are then evidenced by sudden switches in the slope of the lines (from positive to negative, or vice versa). For the moment, however, we shall concentrate on modelling the steady, straight-line segments.

We believe that an appropriate time-series modelling philosophy for medical monitoring needs to separate clearly two different components of the model: first we need a component which models the way in which the underlying system actually evolves; secondly, we need a component describing the way in which the system is measured or observed. The plots in Figures 1 and 2 derive from observations on the system and thus result from a combination of the system and observation processes.

If, in our case, we assume deterministic straight-line evolution of the system in steady periods, the evolution of the system between arbitrary times \( \tau_0, \tau_1, \tau_2, \ldots \), is as shown in Figure 3. The \( \mu \) values denote the actual levels of weight corrected reciprocal creatinine at the given time points and the \( \beta \) values are the incremental changes during the successive time intervals, whose lengths are given by the \( r \) values in terms of some basic unit time interval. In general, the evolution of the steady system between times \( \tau_{t-1} \) and \( \tau_t = \tau_{t-1} + r_t \) is described by

\[
\begin{align*}
    \mu_t &= \mu_{t-1} + \beta_t \\
    \beta_t &= (r_t/r_{t-1}) \beta_{t-1}
\end{align*}
\]

Fig. 2. Reciprocal observed creatinine values.
The model for the observation process is more complicated — even assuming no sudden discontinuities — since it is $\mu t^{-1}$ that is observed directly, and with a multiplicative error. Moreover, there are rounding errors present in the equipment used to measure creatinine. Despite these complications, a detailed mathematical investigation of the orders of magnitude of the various error terms involved (see Smith and West, 1983) shows that, to a satisfactory order of approximation, we can assume the simple measurement model

$$y_t = \mu_t + v_t$$  \hspace{1cm} (2)

where $y_t$ is observed, weight corrected, reciprocal creatinine, and $v_t \sim N(0, c^2\mu_t^2)$, $c$ being the coefficient of variation of the creatinine analyser used (with $c \approx 0.04$ being a typical value). Equations (1) and (2) together constitute the system and observation process models for steady evolution and, as we shall see in the next section, provide a convenient starting point for the extension to incorporate various possible discontinuities.

3 Modelling discontinuities

There are three forms of sudden discontinuity which can occur in the observed series. The first of these discontinuities relates to the observation process, (2), and occurs in the form of an "outlier": an inaccurate laboratory measurement resulting in an aberrant looking data point.

The other two discontinuities both relate to the system process, (1). One occurs when a patient receives dialysis, which results in a sudden change in the reciprocal creatinine level $\mu_t$ and thus affects the first system equation. The other corresponds to the onset of a rejection episode, whose effect is to cause a sudden jump in the incremental change $\beta_t$, thus disturbing the second system equation.

Reconsidering (1) and (2) in the light of Figure 4, it turns out that a suitable framework
Fig. 4. Three types of discontinuity.

for monitoring kidney transplant patients is provided by the four-state multi-process Kalman filter, based on an underlying linear growth model (see Harrison and Stevens, 1976; Stoodley and Mirnia, 1979).

Allowing for unequally spaced observations, we express the full model in the form

\[
\begin{align*}
\gamma_t &= \mu_t + \nu_t \\
\mu_t &= \mu_{t-1} + \beta_t + \gamma_t \\
\beta_t &= (r_t/r_{t-1}) \beta_{t-1} + \delta_t
\end{align*}
\]

where

\[
\nu_t \sim N(0, c^2 \mu_t^2 K_v), \quad \gamma_t \sim N(0, c^2 K_y), \quad \delta_t \sim N(0, c^2 r_t^2 K_v)
\]

and the \( K \)'s are selected in the following way, with \( j = 1, 2, 3, 4 \), indexing the four states; steady evolution, change in level, change in slope, outlier:

\[
\begin{array}{cccc}
1 & 1 & 0 & 0 \\
1 & 1 & 90 & 0 \\
1 & 0 & 60 & 0 \\
100 & 0 & 0 & 0 \\
\end{array}
\]

These variance multipliers – which, in effect, tell the monitoring procedure that discontinuities will manifest themselves as elements having bigger variances than anticipated in the steady evolution state – were chosen on the basis of extensive empirical trials. For convenience, we shall write \( K^{(j)}(t) = r_t^2 K^{(j)} \).

The model described by (3), (4) and (5) can be rewritten in the form

\[
\begin{align*}
\gamma_t &= F \theta_t + \nu_t \\
\theta_t &= G_t \theta_{t-1} + w_t
\end{align*}
\]

where

\[
F = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad G_t = \begin{bmatrix} 1 & (r_t/r_{t-1}) \\
0 & (r_t/r_{t-1}) \end{bmatrix}
\]

\[
\theta_t^T = [\mu_t, \beta_t], \quad w_t^T = [\gamma_t, \delta_t]
\]
and \( v_t, w_t \) are normally distributed, with zero means and variances given by
\[
\text{Var} \left( v_t \Big| \mu_t, M_t^{(j)} \right) = C^2 \mu_t^2 K_v^{(j)} \\
\text{Var} \left( w_t \Big| M_t^{(j)} \right) = C^2 K_w^{(j)}(t)
\]
with
\[
K_w^{(j)}(t) = \begin{bmatrix} K_{v}^{(j)}(t) + K_{w}^{(j)}(t) \\ K_{v}^{(j)}(t) \\ K_{w}^{(j)}(t) \end{bmatrix}
\]
and \( M_t^{(j)} \) denoting the assumption that the \( t \)th observation is of the system in state \( j \).

4 Implementation

We shall write \( D_t \) to denote the observed sequence of observations on a given patient, up to and including the \( t \)th; \( D_0 \) will denote the situation before any observations have been made on the particular patient.

Among the main quantities of interest after observing the \( t \)th observation are
\[
1 - P_r(M_t^{(1)} \big| D_t), \quad P_r(M_{t-1}^{(3)} \big| D_t), \quad P_r(M_{t-2}^{(6)} \big| D_t)
\]
representing, respectively, the probabilities of "instability now", "slope change one-step back", "slope change two-steps back", all in the light of observations up to the current time (and, for the latter two values, corresponding to times where \( E(\beta_{t-1} \big| D_t) \) or \( E(\beta_{t-2} \big| D_t) \) are negative, thus indicating deterioration in kidney function).

The model structure defined by (6)–(9) leads to straightforward recursive procedures, which enable these and other quantities to be calculated straightforwardly. Particular prior inputs required are \( P_r(M_t^{(j)} \big| D_0) \) – assessed on the basis of retrospective study of previous cases – and a prior specification for \( \theta_0 \), assumed to be of the form \( N(m_0, C^2 C_0) \). The values used for the state probabilities were
\[
\{ P_r(M_t^{(j)} \big| D_0), j = 1, 2, 3, 4 \} = (0.85, 0.06, 0.07, 0.02)
\]
reflecting part experience of about 2 per cent outliers, 6 per cent of observations affected by dialysis and 7 per cent corresponding to changes in slope. It is important to note that the procedure is designed for implementation without current knowledge of the occurrence or otherwise of dialysis. If such information were available, the values given above could be overridden, for the time point in question, by the form \( (0, 1, 0, 0) \).

The values specified for \( m_0 \) and \( C_0 \) are derived from clinical knowledge of typical initial creatinine levels and changes immediately following transplantation. The values used were
\[
m_0 = \begin{pmatrix} 0.001 \\ 0 \end{pmatrix}, \quad C_0 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}
\]

Detailed derivations of the required recursive forms, both for \( C \) known and unknown, are given in Smith and West (1983). The latter only refers explicitly to the case of equally spaced observations, but the more general model defined by (6)–(9) is analysed in precisely the same way.

A summary output of the analysis from a typical patient is shown in Figure 5.

5 Approximate decision analysis

In this section we summarize a crude form of decision analysis based on the quantity
Figure 5. Warning probabilities.

$P_r(M_{t-1}(^3) | D_t)$: in other words, using the one-step back probability of a slope change in cases where $E(\beta_{t-1} | D_t) < 0$.

The clinician wishes to have a decision rule of the form

"take action if $P_r(M_{t-1}(^3) | D_t) > \alpha$",

for some suitable value of $\alpha$. The question is, how to choose $\alpha$?

The underlying decision tree is as shown in Figure 6.

A proper analysis of this problem would require the full specification of the utilities of the four consequences. However, if we assume that $|d|$ is very much less than $\min \{ |a|, |b|, |c| \}$ and that $\alpha$ is chosen (whatever happens) such that $P_r$ (No action | Rejection) is very
small, the optimization problem reduces to: choose \( \alpha \) to maximize

\[ a \cdot Pr(\text{Action} \mid \text{Rejection}) + b \cdot Pr(\text{Action} \mid \text{No rejection}). \]

In order to solve this problem, we took about 30 previous series (with an average of 65 days in each) and ran the probabilistic analysis, based on \( Pr(M_t-1(5) \mid D_t) \) and the simple decision rule, with a set of specified cut-off levels \( 0 = \alpha_1 < \alpha_2 < \ldots < \alpha_m = 1 \). For each \( \alpha \), we noted the set of time points, \( P(\alpha) \), for which \( Pr(M_t-1(5) \mid D_t) \) exceeded \( \alpha \).

Following this, the observed series (but not the probabilistic analyses) were given to a group of consultant physicians, who also had access to full patient records and the results of other clinical tests. The physicians then agreed a category, "probable or definition rejection" versus "no rejection", for each data point in each series (see, Trimble, 1980, for full details).

Taking the clinicians' category as the true one, it was then possible to derive, from the results of the probabilistic analysis, estimates, for each of the \( \alpha \) levels, of

\[ T(\alpha) = Pr(\text{Action} \mid \text{Rejection}) \]
\[ F(\alpha) = Pr(\text{Action} \mid \text{No rejection}) \]

Returning to the approximate decision analysis, if we set \( a = K, b = -1 \) (thus expressing \( a \) on a scale where a false alarm is regarded as one negative unit), the expected utility of a decision rule with cut-off level \( \alpha \) is given by

\[ U(\alpha) = KT(\alpha) - F(\alpha) \]

Assuming \( U(\alpha) \) to be differentiable, the optimal cut-off level \( \alpha^* \) satisfies

\[ \frac{K}{\partial \alpha} \frac{\partial T(\alpha)}{\partial \alpha} - \frac{1}{\partial \alpha} = 0 \]

or, since

\[ \frac{\partial T(\alpha)}{\partial \alpha} = \frac{\partial T(\alpha)}{\partial F(\alpha)} \frac{\partial F(\alpha)}{\partial \alpha} \]

assuming \( \partial F(\alpha)/\partial \alpha \neq 0 \), we have

\[ \frac{\partial T}{\partial F} = K^{-1} \quad \text{at} \quad \alpha = \alpha^* \]

We therefore plot the pairs of values \([T(\alpha_i), F(\alpha_i)]\) obtained from the retrospective exercise for the set of levels \( \alpha_1, \alpha_2, \ldots, \) etc., and use these to roughly sketch \( T \) as a function of \( F \). We can then interpolate to find an approximate value for \( \alpha^* \). For example, a value \( K=3 \) gave a cut-off value of \( \alpha^* \approx 0.2 \).
Detailed discussion of this form of implementation and its validation is given in Trimble (1980). We can summarize his findings by saying that extensive retrospective and prospective use of the Kalman filter procedure has proved highly satisfactory to clinicians, both as a source of genuine guidance for the analysis of very noisy series and as a convenient "automatic pilot" for series which could, in principle, have been examined by eye, but at the cost of considerable effort on the part of the clinician, which could thereby be released for other purposes.

References