Recursive estimation method for predicting residual bladder urine volumes to improve accuracy of timed urine collections

David Afshartous and Richard A. Preston, Division of Clinical Pharmacology, Department of Medicine, Miller School of Medicine, University of Miami, Miami, FL 33136, USA

Clinical research studies often collect data via repeated measurements of collected urine. Unfortunately, the accuracy of timed urine collections is limited by the presence of a residual volume of urine remaining in the bladder following each timed void due to incomplete emptying of the bladder. This residual urine volume adds significant imprecision to the urine collection method, rendering an important and fundamental clinical research tool inaccurate. We present an unbiased method to estimate the residual bladder volumes via a mathematical model of the bladder process. Regardless of the substance of primary interest, the model leverages conservation of mass and conservation of concentration principles towards a substance of secondary interest in order to solve a system of recursive equations, resulting in our Recursive Residual Estimation method to predict the residual volumes at each time point. We verify the model on simulated patients and also investigate the sensitivity of the model to initial value specification.

1 Introduction

Clinical research studies often employ timed urine collections to collect physiological data. Such measurements are collected over fixed time intervals and analysed as a function of time. Timed urine collections are inexpensive, non-invasive, relatively simple to conduct and can provide valuable data on kidney function and the clearance of various electrolytes and drugs by the kidney. For example, studies of drug elimination by the kidney require timed urine collections in order to calculate a drug’s clearance by the kidney. Renal elimination of potassium, a critical physiological process, can be quantitatively studied by administering a standard potassium load followed by hourly determination of urinary potassium. Unfortunately, the utility and widespread use of timed urine collections is limited by the inherent imprecision of the measurement, largely due to a residual volume of urine remaining in the bladder following each timed void. In many research study participants, the bladder is not fully emptied upon voiding, and a residual volume of urine remains in the bladder after each urine collection. The existence of a residual volume of urine at each collection point presents difficulties when attempting to estimate the actual amount of a substance cleared by the kidney during the specified time interval. For example, the potassium collected during a 1 h urine collection may not represent the actual potassium introduced into the bladder.
during that hour. Furthermore, the residual bladder volume may vary from one timed period to the next, and may actually increase when urine volumes are large and urine is collected frequently at short intervals. Thus, we seek to model this residual bladder process.

A model that can predict the residual bladder volume would allow for the correction of the parameters of interest and would increase the accuracy, and therefore the utility of a valuable clinical research tool, the timed urine collection. The goal of this investigation is to estimate these residual volumes such that they may be used to adjust the measurements of interest from the collected urine. The model incorporates quantities such as the new urine volume introduced into the bladder each time period, the urine volume voided each time period, the residual urine volume remaining after voiding and their corresponding concentrations. Regardless of the substance of primary interest, the model will take advantage of conservation of mass and conservation of concentration principles towards a substance of secondary interest. For example, under steady state conditions, the body produces a relatively constant mass of creatinine per unit time. In addition, the concentration of a substance in the voided urine specimen is the same as that in the bladder residual volume. We will show that the incorporation of these principles facilitates solving a recursive set of equations for the residual volumes.

The outline of this article is as follows. In Section 2 we present a model of the bladder process and solve a recursive set of equations for the residual volumes. In Section 3 we investigate the robustness of the model via a sensitivity analysis with simulated data. In Section 4 we discuss the implementation of the model in practice and in Section 5 we conclude with a brief summary and directions for future research.

2 Bladder model

Figure 1 illustrates the bladder process. At the start of a given time interval, say \( t = 1 \), there exists a residual volume \( r_1 \) in the bladder that was not emptied during voiding. During the next time interval, the volume \( N_2 \) is introduced into the bladder. Thus, exactly at time \( t = 2 \) (right before voiding), the total volume in the bladder is \( B_2 = N_2 + r_1 \). At time \( t = 2 + \epsilon \), a volume of urine \( C_2 \) is collected. We differentiate between time \( t = 2 \) and \( t = 2 + \epsilon \) in order to better differentiate between the volumes within and outside of the bladder. Specifically, we may also write \( B_2 = C_2 + r_2 \), i.e. the total volume at the end of the interval (just before voiding) is equal to the sum of the volume that will be voided plus the volume that will be leftover. Thus, the general formula is:

\[
B_t = N_t + r_{t-1} = C_t + r_t,
\]

or, in terms of the residual volume:

\[
r_t = B_t - C_t = (N_t + r_{t-1}) - C_t.
\]

Since we always have the urine volume collected, we can solve for the residual volumes \( r_t \) if we know the total bladder volume at the end of each time interval \( B_t \) (or equivalently the new volume).
Conservation of mass and conservation of concentration principles may be employed to form a useful expression for $B_t$. By incorporating these principles towards a substance of secondary interest (creatinine), the objective of estimating the residual volumes may be achieved; this allows one to accurately describe the behaviour of the substance of primary interest. Specifically, the mass of creatinine produced each interval is constant, say equal to $\gamma$. Thus, the mass that exists in $B_{t+1}$ is equal to the mass leftover in $r_t$ plus the new mass $\gamma$ introduced during the time period. Formally,

\[
\text{Mass}_{B_{t+1}} = \text{Mass}_{r_t} + \gamma. \tag{3}
\]

Given that mass is concentration (of creatinine) multiplied by volume, we have:

\[
\text{Conc}_{B_{t+1}} \times B_{t+1} = (\text{Conc}_{r_t} \times r_t) + \gamma. \tag{4}
\]

Solving for $B_{t+1}$, we have:

\[
B_{t+1} = \frac{(\text{Conc}_{r_t} \times r_t) + \gamma}{\text{Conc}_{B_{t+1}}} = \frac{(\text{Conc}_{C_t} \times r_t) + \gamma}{\text{Conc}_{C_{t+1}}}. \tag{5}
\]

The last equality follows from the fact that the concentration of the collected urine is equal to that of the bladder volume just before collection. In terms of the new volume, we have:

\[
N_{t+1} = \frac{(\text{Conc}_{C_t} \times r_t) + \gamma}{\text{Conc}_{C_{t+1}}} - r_t. \tag{6}
\]

Substituting Equation (5) into Equation (2) provides a single recursive equation that requires initial values for $r_1$ and $\gamma$:

\[
r_t = \frac{(\text{Conc}_{C_{t-1}} \times r_{t-1}) + \gamma}{\text{Conc}_{C_t}} - C_t. \tag{7}
\]

As with any Markovian system, one can express $r_t$ for $t \geq 2$ in terms of $r_1$ by substituting...
4 **D Afshartous and RA Preston**

accordingly. After some algebra, we obtain:

\[ r_t = \left( \frac{\text{Conc}_{C_1}}{\text{Conc}_{C_t}} \right) r_1 + \left( \frac{t - 1}{\text{Conc}_{C_t}} \right) \gamma - \left[ \sum_{i=2, t+2}^{i=t-1} \frac{\text{Conc}_{C_i} \times C_i}{\text{Conc}_{C_t}} \right] + C_t \].

(8)

Thus, the general form of the recursion is linear and may be expressed as \( r_t = a_{r1} r_1 + a_{r2} \gamma + a_{r3} \). Inspecting Equation (8), we observe that these coefficients can be computed recursively as well, providing a general algorithm to compute the residual volumes from initial values for \( r_1 \) and \( \gamma \) for each patient:

**Algorithm 2.1** (RRE – Recursive residual estimation method).

1. Specify an initial value for \( \gamma \).
2. Specify an initial value for \( r_1 \).
3. Calculate the coefficients for \( r_2 \) and then calculate \( r_2 \):
   (a) \( a_{21} = \frac{\text{Conc}_{C_1}}{\text{Conc}_{C_2}} \)
   (b) \( a_{22} = \frac{1}{\text{Conc}_{C_2}} \)
   (c) \( a_{23} = -C_2 \)
   (d) \( r_2 = a_{21} r_1 + a_{22} \gamma + a_{23} \)
4. Recursively calculate the coefficients for \( r_{t+1} \) and then calculate \( r_{t+1}, t \geq 2 \):
   (a) \( a_{(t+1)1} = a_{t1} \frac{\text{Conc}_{C_{t+1}}}{\text{Conc}_{C_{t+1}}} \)
   (b) \( a_{(t+1)2} = a_{t2} \frac{\text{Conc}_{C_{t+1}}}{\text{Conc}_{C_{t+1}}} + \frac{1}{\text{Conc}_{C_{t+1}}} \)
   (c) \( a_{(t+1)3} = a_{t3} \frac{\text{Conc}_{C_{t+1}}}{\text{Conc}_{C_{t+1}}} - C_{t+1} \)
   (d) \( r_{t+1} = a_{(t+1)1} r_1 + a_{(t+1)2} \gamma + a_{(t+1)3} \)

To be sure, one could attempt to extend the basic model by introducing a varying propensity to empty the bladder to predict correlations of residuals. For instance, it may be that for some individuals, a large residual is always immediately followed by a very small residual since the large residual stimulates a higher likelihood of emptying the bladder during the next period, whereas, for other individuals the stimulus takes longer. However, the current model captures this behaviour without the additional complexity, i.e. regardless of an individual’s bladder emptying propensity, the only measurements needed to fully capture the residual behaviour are the collected measurements and the \( r_1 \) and \( \gamma \) parameters.

3 **Results – sensitivity analysis**

As mentioned earlier, the purpose of the residual estimation method is to adjust the potential bias of collected measurements. This would be useful in any clinical trial where key measurements are collected via voiding. Unfortunately, since the actual residual volumes are unobserved, we cannot directly assess the accuracy of the residual estimation method without actually placing a catheter directly into the bladder. Bladder
Recursive estimation method for predicting residual bladder urine volumes

Table 1  Simulated patient

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hour</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voided volume (cc)</td>
<td></td>
<td>390</td>
<td>270</td>
<td>380</td>
<td>250</td>
<td>310</td>
<td>370</td>
</tr>
<tr>
<td>Actual residual (cc)</td>
<td></td>
<td>20</td>
<td>50</td>
<td>70</td>
<td>20</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Voided concentration (mg/cc)</td>
<td></td>
<td>.10</td>
<td>.13</td>
<td>.11</td>
<td>.18</td>
<td>.14</td>
<td>.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New mass</th>
<th>New vol</th>
</tr>
</thead>
<tbody>
<tr>
<td>41.5 mg</td>
<td>400 cc</td>
</tr>
</tbody>
</table>

| Conc 1 = 41.6 + 1 | Conc 2 = 41.6 + 20 (Conc 1) | Conc 3 = 41.6 + 50 (Conc 2) | Conc 4 = 41.6 + 70 (Conc 3) | Conc 5 = 41.6 + 20 (Conc 4) | Conc 6 = 41.6 + 10 (Conc 5) |
| 390 + 20         | 270 + 50                       | 380 + 70                      | 250 + 20                       | 310 + 10                       | 370 + 40                       |
| = 0.1039 mg/cc   | = 0.1364 mg/cc                 | = 0.1076 mg/cc                | = 0.1820 mg/cc                | = 0.1414 mg/cc                | = 0.1049 mg/cc                |

Figure 2  Mean absolute deviation for various initial values.

catheterisation is invasive, can potentially introduce infection, and therefore would not be acceptable for safety and ethical reasons for the purpose of directly verifying this model. Thus, in order to verify our model, we simulated several patients under a typical repeated measurements clinical trial design, where six hourly measurements are taken for each patient, measuring variables such as voided volume and creatinine concentration. Table 1 illustrates one of these patients.

This patient was simulated according to an actual $\gamma = 41.6$ mg and a first residual $r_1 = 20$ cc.† The value of $\gamma$ was chosen by assuming a daily creatinine production of 1000 mg. Note that we cannot simulate the values for the separate variables independently since they are dependent. Thus, we employed a stepwise process emulating Figure 1 in order to simulate the actual values. The simulation requires the specification of the volume entering the bladder during each hour and both the residual volume and its associated concentration carried into the baseline hour. Figure 2 illustrates the details of this process. To avoid overly crowding the diagram, only the time points corresponding to immediately after voiding are shown; this corresponds to the $t + \epsilon$ notion of Figure 1.

†These values were chosen to reflect typical values for a ‘normal’ patient; we also investigated other values of $r_1$ and $\gamma$ and obtained similar results.
The simulated data illustrate the potential bias in the observed measurements. Indeed, if say changes in hourly creatinine clearance are of interest, the observed measurements can be severely misleading. Consider the change from hour 1 to hour 2. Based on the observed measurements, there is a decrease in creatinine handling from 39 mg (390 cc × .10 mg/cc) to 35.1 mg (270 cc × .13 mg/cc). However, by design the creatine production of this patient is constant each hour; thus, a 0% change would be represented as a 10% decrease! Such discrepancies are commonplace whenever there is variation in the residual and voided volumes. Such large variations can in many cases render data from timed urine collections inaccurate. Thus, an inexpensive, non-invasive and valuable research tool is made much less accurate and therefore potentially unusable. This underscores the need for accurate residual estimation, regardless of the substance of interest.‡

The Recursive Residual Estimation method and all subsequent coding and analysis were implemented in the statistical software R.17 As expected, the method predicts residuals exactly if we correctly specify $\gamma$ and $r_1$. To be sure, in actual practice $\gamma$ and $r_1$ will be unknown and will have to be estimated. Thus, we would like to assess the robustness of the model to departures from the real $\gamma$ and $r_1$ values; recall that these are the only two values that need to be specified to estimate all the residual volumes from the collected data.

In order to assess the aforementioned model robustness, we investigated many scenario combinations where initial values for $\gamma$ and $r_1$ were both under-specified and over-specified. We examined incremental changes in $r_1$ values of 20% from the actual value and incremental changes in $\gamma$ of 5% from the actual value, both below and above the real value. In each scenario, model robustness was assessed via the mean absolute deviation of the real residuals from the estimated residuals. For the simulated design above, note that although there are six hourly measurements, we only estimated the last five residuals since the first must be specified as an initial value. Table 2 contains the results for the simulated patient described earlier. Figure 3 illustrates these results via an interpolated smooth surface to further facilitate pattern recognition.

The zero value for the mean absolute deviation in the middle of Table 2 corresponds to the ideal case where we exactly specify the initial values for $\gamma$ and $r_1$. The main message from these results, dramatically displayed in Figure 2, is that the model is fairly robust with respect to $r_1$ estimation but not very robust with respect to $\gamma$ estimation. In addition, note that the matrix of results in Table 2 is symmetric. This seems unusual at first, as there is no clinical reason to expect the same robustness when we specify say $\gamma = 49.92$ and $r_1 = 4$ versus $\gamma = 33.28$ and $r_1 = 36$, e.g. when the actual values are $\gamma = 41.6$ and $r_1 = 20$. Further insight, however, is revealed if we consider the general case and focus on the estimation of the second residual $r_2$. Consider the following two

\[ \text{New Vol}_{t} = C_t + r_t - r_{t-1}. \]

However, the new concentration is not equal to the concentration of the collected volume ($C_t$) since the collected volume ($C_t$) is affected by the concentration of the previous residual. After some algebra, utilising the fact that the collected concentration is a weighted sum of the new volume concentration and the previous residual concentration, we obtain the desired concentration as:

\[ \text{New Vol Conc}_{t} = \left( \frac{C_t - r_{t-1} \text{Conc}_{t-1}}{\text{New Vol}_{t}} \right) \left( \frac{C_t}{\text{New Vol}_{t}} \right). \]

---

‡Given estimates of the residuals, one may estimate the amount cleared of any substance for a given time period by multiplying the new volume introduced during the time period by its concentration. The new volume is simply the collected volume, adjusted accordingly for the current and previous residual: $\text{New Vol}_{t} = C_t + r_t - r_{t-1}$. However, the new concentration is not equal to the concentration of the collected volume ($C_t$) since the collected volume ($C_t$) is affected by the concentration of the previous residual. After some algebra, utilising the fact that the collected concentration is a weighted sum of the new volume concentration and the previous residual concentration, we obtain the desired concentration as: $\text{New Vol Conc}_{t} = \left( \frac{C_t - r_{t-1} \text{Conc}_{t-1}}{\text{New Vol}_{t}} \right) \left( \frac{C_t}{\text{New Vol}_{t}} \right)$. 

---

6 D Afshartous and RA Preston
Recursive estimation method for predicting residual bladder urine volumes

Table 2  M.A.D. – Mean absolute deviation

<table>
<thead>
<tr>
<th>γ</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>24</th>
<th>28</th>
<th>32</th>
<th>36</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.2</td>
<td>262.2</td>
<td>259.0</td>
<td>255.8</td>
<td>252.6</td>
<td>249.3</td>
<td>246.1</td>
<td>242.9</td>
<td>239.7</td>
<td>236.5</td>
<td>233.3</td>
<td>230.0</td>
</tr>
<tr>
<td>33.2</td>
<td>213.0</td>
<td>209.8</td>
<td>206.5</td>
<td>203.3</td>
<td>200.1</td>
<td>196.9</td>
<td>193.7</td>
<td>190.5</td>
<td>187.2</td>
<td>184.0</td>
<td>180.8</td>
</tr>
<tr>
<td>35.3</td>
<td>163.7</td>
<td>160.5</td>
<td>157.3</td>
<td>154.1</td>
<td>150.9</td>
<td>147.7</td>
<td>144.4</td>
<td>141.2</td>
<td>138.0</td>
<td>134.8</td>
<td>131.6</td>
</tr>
<tr>
<td>37.4</td>
<td>114.5</td>
<td>111.3</td>
<td>108.1</td>
<td>104.9</td>
<td>101.6</td>
<td>98.4</td>
<td>95.2</td>
<td>92.0</td>
<td>88.8</td>
<td>85.5</td>
<td>82.3</td>
</tr>
<tr>
<td>39.5</td>
<td>65.3</td>
<td>62.1</td>
<td>58.8</td>
<td>55.6</td>
<td>52.4</td>
<td>49.2</td>
<td>46.0</td>
<td>42.8</td>
<td>39.5</td>
<td>36.3</td>
<td>33.1</td>
</tr>
<tr>
<td>41.6</td>
<td>16.0</td>
<td>12.8</td>
<td>9.6</td>
<td>6.4</td>
<td>3.2</td>
<td>0.0</td>
<td>3.2</td>
<td>6.4</td>
<td>9.6</td>
<td>12.8</td>
<td>16.0</td>
</tr>
<tr>
<td>43.6</td>
<td>33.1</td>
<td>36.3</td>
<td>39.5</td>
<td>42.8</td>
<td>46.0</td>
<td>49.2</td>
<td>52.4</td>
<td>55.6</td>
<td>58.8</td>
<td>62.1</td>
<td>65.3</td>
</tr>
<tr>
<td>45.7</td>
<td>82.3</td>
<td>85.5</td>
<td>88.8</td>
<td>92.0</td>
<td>95.2</td>
<td>98.4</td>
<td>101.6</td>
<td>104.9</td>
<td>108.1</td>
<td>111.3</td>
<td>114.5</td>
</tr>
<tr>
<td>47.8</td>
<td>131.6</td>
<td>134.8</td>
<td>138.0</td>
<td>141.2</td>
<td>144.4</td>
<td>147.7</td>
<td>150.9</td>
<td>154.1</td>
<td>157.3</td>
<td>160.5</td>
<td>163.7</td>
</tr>
<tr>
<td>49.9</td>
<td>180.8</td>
<td>184.0</td>
<td>187.2</td>
<td>190.5</td>
<td>193.7</td>
<td>196.9</td>
<td>200.1</td>
<td>203.3</td>
<td>206.5</td>
<td>209.8</td>
<td>213.0</td>
</tr>
<tr>
<td>52.0</td>
<td>230.0</td>
<td>233.3</td>
<td>236.5</td>
<td>239.7</td>
<td>242.9</td>
<td>246.1</td>
<td>249.3</td>
<td>252.6</td>
<td>255.8</td>
<td>259.0</td>
<td>262.2</td>
</tr>
</tbody>
</table>

general cases: (1) over-specify $r_1$ by $\delta$ and under-specify $\gamma$ by $\theta$ and (2) under-specify $r_1$ by $\delta$ and over-specify $\gamma$ by $\theta$. The estimation of $\hat{r}_2$ for the two cases then proceeds as follows, with the actual $r_2$ also provided for completeness:

$$\hat{r}_2 = a_{21}(r_1 + \delta) + a_{22}(\gamma - \theta) + a_{23}$$

$$\hat{\hat{r}}_2 = a_{21}(r_1 - \delta) + a_{22}(\gamma + \theta) + a_{23}$$

$$r_2 = a_{21}r_1 + a_{22}\gamma + a_{23}$$

In both cases, the absolute deviation is equal to $|a_{21}\delta - a_{22}\theta|$. Indeed, it readily follows that we will have similar symmetry for all $r_i$, hence explaining the equal values for the mean absolute deviation for these two scenarios. Furthermore, we note that if the condition $a_{21}\delta - a_{22}\theta > 0$ is satisfied, the first scenario leads to over-predicting $r_1$ whereas the second scenario leads to under-predicting $r_1$.

3.1 A Bayesian extension

The model may be extended to a Bayesian framework by applying formal distributions to the key parameters $r_1$ and $\gamma$. For instance, assume that $r_1 \sim N(\mu_{r_1}, \sigma_{r_1}^2)$ and $\gamma \sim N(\mu_\gamma, \sigma_\gamma^2)$. The analysis is simplified by the fact that the coefficients in the main recursive residual formula do not directly depend on either of these parameters, i.e. $r_t = a_{t1}r_1 + a_{t2}\gamma + a_{t3}$. As $r_t$ is now simply a linear combination of normal distributions, $r_t$ is also normally distributed:

$$r_t \sim N(a_{t1}\mu_{r_1} + a_{t2}\mu_\gamma + a_{t3}, a_{t1}^2\sigma_{r_1}^2 + a_{t2}^2\sigma_\gamma^2).$$  

Thus, for any given $r_t$, uncertainty is conveyed via the uncertainty in $r_1$ and $\gamma$, whereas the previous analysis specified a single value for these key parameters and thus no uncertainty in $r_t$. One may relate the two methods by considering the previous analysis as a special case of the Bayesian approach with the variances of
r₁ and γ set equal to 0. Although, it is informative to examine the derived distribution for r₁, the ultimate goal of adjusting observed electrolyte measurements collected via voiding is facilitated by using the quantiles of this distribution to obtain intervals for r₁. When setting the means of the distributions for r₁ and γ to the previously supplied single values, the centre of these intervals will be equal the previous point estimates. Regarding implementation, two additional computations must be added to Algorithm 2.1 at step 3(d) and 4(d): in addition to the current calculation (which may be viewed as inputting the mean values of the prior distributions for r₁ and γ), we perform the same computation to calculate ‘upper-residuals’ with (r₁, γ) = (μᵣ₁ + σᵣ₁, μᵣ + σᵣ) and ‘lower residuals’ with (r₁, γ) = (μᵣ₁ − σᵣ₁, μᵣ − σᵣ). This yields two additional sets of predicted residuals for each patient to add uncertainty around the original predictions, thereby formalising the uncertainty in specifying r₁ and γ.

However, we note that the widths of these intervals are not constant across the time points, as the coefficients in the recursive formula are time dependent. For example, for the patient considered in Figure 3, if we specify r₁ ∼ N(20, 5²) and γ ∼ N(41.6, 1.2²) instead of simply r₁ = 20 and γ = 41.6, the interval widths are (12, 27, 22, 37, 62), where the centre of the intervals are the true residuals since we have centred the input.
Recursive estimation method for predicting residual bladder urine volumes

Parameter prior distributions around the known values. Although, in practice the prior distributions will not be centred around the true values as these are unknown, the same procedure applies. With respect to the mean absolute deviations in Figure 2 that cover a range of mis-specifications (for the mean of the prior distribution), the Bayesian approach corresponds to creating an upper and lower confidence surface around the existing MAD surface. Furthermore, the upper MAD surface at a given point does not necessarily correspond to the upper residuals, i.e. those calculated using the upper quantiles of the prior distributions for $(r_1, \gamma)$. Rather, somewhat intuitively, when $\gamma$ is under-specified, the upper residuals yield a lower MAD, and vice versa when $\gamma$ is over-specified; the specification of $r_1$ is dominated by that of $\gamma$ for the values considered, viz., even if $\gamma$ is only slightly under-specified and $r_1$ is grossly over-specified, the upper residuals yield a lower MAD. Graphically, the surfaces corresponding to the upper and lower residuals reverse their ordering according to whether $\gamma$ is over/under specified. Moreover, the upper and lower residuals would also yield the same type of symmetry noted earlier in Table 2.

4 Discussion

The Recursive Residual Estimation Method allows one to solve for the unknown residual volumes. However, in order to be implemented, one must specify the values of $r_1$ and $\gamma$ for each patient (or the mean and standard deviations of their distributions in the Bayesian framework). Mathematically, one may use the recursive equation to determine the feasible domain from which to choose these initial values, i.e. values such that the method is guaranteed to produce positive values for the estimated residuals. Thus, the challenge is to find the feasible range of $r_1$ and $\gamma$ that lead to a positive solution for the residual volumes. One could experiment with different starting values, but then this would have to be repeated for each patient. Instead, one may view each residual equation $r_t = a_{11}r_1 + a_{12}\gamma + a_{13}$ in terms of the constraint $r_t \geq 0$. As the coefficients $a_{ti}$ do not depend on $r_1$ or $\gamma$, we may solve for them independently and graph each constraint $r_t \geq 0$ as a function of $\gamma$ and $r_1$. For each additional measurement interval we have an additional constraint and hence feasible region, and the intersection of all the feasible regions provides feasible starting values for which we are assured a solution of positive residual volumes. For example, Figure 4 provides the constraint lines for the initial values for the simulated data. Note that the true $\gamma$ is very close to the region where the lines intersect, while the true $r_1$ is less restricted. We noticed a similar pattern for other simulated patients. Initially, it seems that visual inspection of these feasible regions is informative with respect to $\gamma$ specification, a very important result given the previous sensitivity results. Thus, along with clinical information, these constraint diagrams could be used for parameter specification with actual patients.

Although, the feasible domain is quite large, this can be narrowed down further to a 1D search space if we are confident in our estimate of $\gamma$ (see below). Based on the previous simulation results, the specification of $\gamma$ is more important with respect to the accurate estimation of the residual volumes. An approximate value of $r_1$ may be specified based on patient clinical data and important covariates such as age group.
To supplement the feasible region method discussed above, we may consider more active methods for the estimation of $\gamma$. In any given experiment, we will be able to calculate the mass collected at each time period from the collected volume and concentration measurements. Suppose we estimate $\gamma$ via the average of the creatinine masses collected each time period. Does this provide a good approximation to $\gamma$? Suppose the creatinine concentration in the residual volume from the previous period $r_{t-1}$ equals the concentration in new volume $N_t$. Let $\gamma^*_t$ equal the creatinine mass collected at time $t$ and let $\gamma_t = \gamma$ equal the creatinine mass produced in the new volume, which will remain constant every time period. We will have $\gamma^*_t = \gamma$ when $r_{t-1} = 0$ and $r_t = 0$, i.e. there was no residual to begin with and everything emptied. Moreover, based on our initial assumption regarding concentration, we will have $\gamma^*_t < \gamma$ when $C_t < N_t$, i.e. when the volume collected is less than new volume produced. (As the creatinine in new volume mixes evenly with leftover volume, we only need to examine volumes to determine if less creatinine mass is collected.) The condition $C_t < N_t$ is equivalent to an increase in the residual volume. Similarly, $\gamma^*_t > \gamma$ when $C_t > N_t$, i.e. when the volume collected is greater than new volume produced. This is equivalent to a decrease in the residual volume.

Although, a patient may exhibit trends in residual behaviour, the increase/decrease of the residual volume should be symmetric, i.e. for any given time point, the likelihood of a residual increase should be similar to that of a decrease, ceteris paribus. Hence, the probability that $\gamma^*_t$ is less (greater) than $\gamma$ should be equal to 0.5, and thus averaging them should estimate $\gamma$ unbiasedly. In summary, our proposed method to estimate $\gamma$ via
Recursive estimation method for predicting residual bladder urine volumes

averaging the collected creatinine masses is unbiased if (1) the creatinine concentration in the residual volume from the previous period \( r_{t-1} \) equals the concentration in new volume \( N_t \); or, less strictly, the creatinine concentration in the new volume is not systematically higher/lower than that remaining from the previous time period, and (2) the residual volume is not systematically increasing or decreasing. If aspects of a study design lead to a violation of the second assumption, one could adjust the model accordingly.

In an actual clinical trial, measurement error in either the creatinine concentration or the voided volume could lead to overestimating the residual volumes. Even small measurement error could have a possible ‘snowball’ effect that would be carried forward at each iteration, see Equation (8). Thus, in order to be confident in the estimated residuals one must insure that the measurement of creatinine and voided volumes are relatively accurate. Fortunately, these measurements are easily performed and quite accurate.

In addition, in an actual clinical trial patients may exhibit multiple voids during each time interval, instead of a single void at the end of interval. These multiple voids are usually added together to produce the total volume voided for the interval. Fortunately, this does not affect the accuracy of our method. Indeed, assuming that the multiple void volumes mix properly, it is relatively easy to show that one can act as if this total volume was produced at one time point instead of multiple time points. The lost information of the residuals during the multiple voids within the time interval are irrelevant; the only pieces of information needed are the last residual and the total volume collected and corresponding concentration.

5 Conclusion

The Recursive Residual Estimation Method represents an unbiased method to estimate the residual bladder volumes via a mathematical model of the bladder process. The knowledge of such residual volumes is extremely important with respect to understanding the true handling of various electrolytes and drugs by the kidney. Regardless of the substance of primary interest, the model leverages conservation of mass and conservation of concentration principles towards a substance of secondary interest in order to solve a system of recursive equations. The model only requires the specification of two initial values: the creatinine mass introduced per unit time into the bladder \( \gamma \), and the initial residual volume \( r_1 \). The simulated data indicate that the robustness of the model is highly dependent on the accurate specification of \( \gamma \), while not so much so with respect to \( r_1 \). In future research we plan to employ the model in the context of a potassium handling study and adjust the collected measurements accordingly.

Acknowledgements

We thank Hernan Awad for some very helpful discussions of this article, and an anonymous reviewer for many detailed and helpful comments that have greatly improved the article.
12 D Afshartous and RA Preston

References


