A study on serum calcium and phosphate in dialysis patients via a mixed effects latent Markov model with two outcomes

Alessio Farcomeni, Gian Luca Di Tanna, Annarita Vestri, Sara Viviani

Department of Experimental Medicine
Sapienza - University of Rome

Corresponding Author:
Dr. Alessio Farcomeni
alessio.farcomeni@uniroma1.it

Summary
We propose a latent Markov model to simoultaneously model two longitudinal binary outcomes. The model does not assume normality of the random effects and allows the subject-specific parameters to evolve over time. Its usefulness is demonstrated on an original study on the effects of parathyroidectomy on targeting recommended levels of serum calcium and phosphate in dialysis patients with secondary hyperparathyroidism. We illustrate the clinical implications of our empirical analysis and the methodological issues related to modeling multivariate longitudinal categorical outcomes in general.

KEY WORDS: *Dialysis, latent Markov model, multivariate responses, longitudinal data, nephrology, random effects.*

1. Introduction

The recently introduced K/DOQI guidelines (1) indicate safe ranges for serum levels of different nutrients. These ranges have affected the management of mineral metabolism in uremic patients, who are more likely to fail targeting serum calcium (Ca) and serum phosphate (P), see for instance (2).
Despite drug terapies, many patients suffer from unresponsive secondary hyperparathyroidism, which requires surgery. Surgery is expected to be curative. Our research question is how much surgery is able to allow patients accomplish the recommended K/DOQI ranges. A recent study (3) demonstrated that the K/DOQI ranges are expected to be targeted only in the short term, thereby questioning the efficacy of surgery for unresponsive secondary hyperparathyroidism.
Along these lines, in this study we aim at evaluating if patients submitted to parathyroidectomy actually target optimal serum levels of Ca and P. As a consequence, we need to analyze data in which two outcomes were measured simoultaneosly over different time occasions.

When observing more than one outcome simoultaneously, a common approach, especially in longitudinal studies, is to fit a model separately on each one. We argue in this work that this practice may be awed at least in two respects: (i) it does model dependence among outcomes, and direct effects one outcome can have on another and (ii) it is inefficient since it does not take into account possible dependence among the outcomes.
Another common practice is to model longitudinal data using mixed effects models in which the random intercepts are time-constant and assumed to arise from a Normal distribution. While we believe these two latter assumptions may be sensible in many applications, we argue they may be sometimes restrictive, leading to biased regression estimates and hence possibly awed conclusions. We will demonstrate both assumptions are restrictive for the data at hand, concluding that the researcher should routinelly check the normality and time-constant assumptions for the random intercepts, and possibly use models which bypass them.
In order to provide efficient, informative and less biased estimates we introduce a multivariate mixed-ef-

fects model based on a latent Markov heterogeneity structure. The subject-specific parameters change over time and there is no normality assumption. The model is a special case of a class of models recently proposed in (4), and will be tailored to fitting simoultaneously the two binary outcomes (within/outside ranges for each nutrient).

The same data is analyzed by (3), but two separate models are used there.

The scope of this paper is then threefold: firstly, we outline a latent Markov model which can be used to fit multivariate marginal longitudinal logit models on two binary outcomes, allowing for non-informative dropout which is ubiquitous in longitudinal studies. Secondly, we develop an original application on a study on levels of serum calcium and phosphate in dialysis patients who underwent PTX for unresponsive secondary hyperparathyroidism. Finally, we use the application to demonstrate the need of using multivariate models when more than one measured variable should be deemed as an outcome; and the need of checking the common assumptions of normality and time-constancy of random effects, both of which will be rejected for our data.

The rest of the paper is as follows: in the next section we describe the motivating example and briefly argument why the common approaches may not be suitable for the data at hand. We describe our proposed model in Section 3 and detail issues about random effects in Section 3.1. Section 4 shows the results on K/DOQI data and compares our model with more common approaches. We conclude in Section 5 with a brief discussion.

2. Data

Recently issued K/DOQI ranges recommend the following levels for serum calcium and phosphate:

8.5 mg/dl ≤	Ca	≤ 9.5mg/dl
3.5 mg/dl ≤	P	≤ 5.5mg/dl

Between 2000 and 2005 a total of 77 dialysis patients, referred to our hospital from several dialysis units, received a parathyroid surgery. For every patient the existence of very high PTH levels (>800 pg/ml) together with hypercalcemia and hyperphosphatemia or a condition of unresponsiveness to current medical management were the criteria for surgical intervention.

Three different types of parathyroidectomy were performed: total (all isolated glands) in 36 patients; subtotal (three glands and 5/6th of the fourth) in 8, and total plus forearm auto-transplantation in 33. After surgery patients were summoned on a regular basis in order to check the biochemical and clinical outcomes.

We recorded biochemical assays one month, and one, three and five years after surgery. Daily adjustments in the therapy of secondary hyperparathyroidism, according to the levels of Ca, P and PTH, were left to the single nephrologist responsible for each patient.

Data were firstly analyzed by (3).

We can define the two binary response variables as:

- *Calcium*: equal to 1 if serum calcium is within K/DOQI ranges, 0 otherwise
- *Phosphate*: equal to 1 if serum phosphate is within K/DOQI ranges, 0 otherwise

We only have recorded two covariates, measuring sex and age of the patient at time of PTX. We had 38 males, 29 females, aged 54 ± 12 years.

Other covariates were not included into the present analysis because of clear insignificance at the outset.

In Figure 1 we show boxplots, one for each time occasion, of observed levels of Ca, with red lines indicating the K/DOQI thresholds. Figure 2 shows the same for P.

From the gures it can be seen that Ca levels tend to grow after PTX, with more than 50% subjects above the upper threshold after 1, 3 and 5 years. For P the medians stay within the threshold after surgery, even if this does not happen for the third quartile. Further, many subjects are outside the thresholds at all time points and there are few subjects who fail to target the recommended ranges because they are below the lower thresholds.

Hypoparathyroidism is a well known risk associated with PTX.

Follow-up data are available for all patients after one month, while those checked after 1, 3 and 5 years are respectively 72, 47 and 27. Some patients have not yet reached all the temporal end-points (n = 24), while others have been censored due to renal transplantation (n = 5) or death (n = 9). Five cases are lost in the follow-up for unknown reasons.

It could be argued that drop-out is informative at least for the n = 14 censored cases. We have for the time

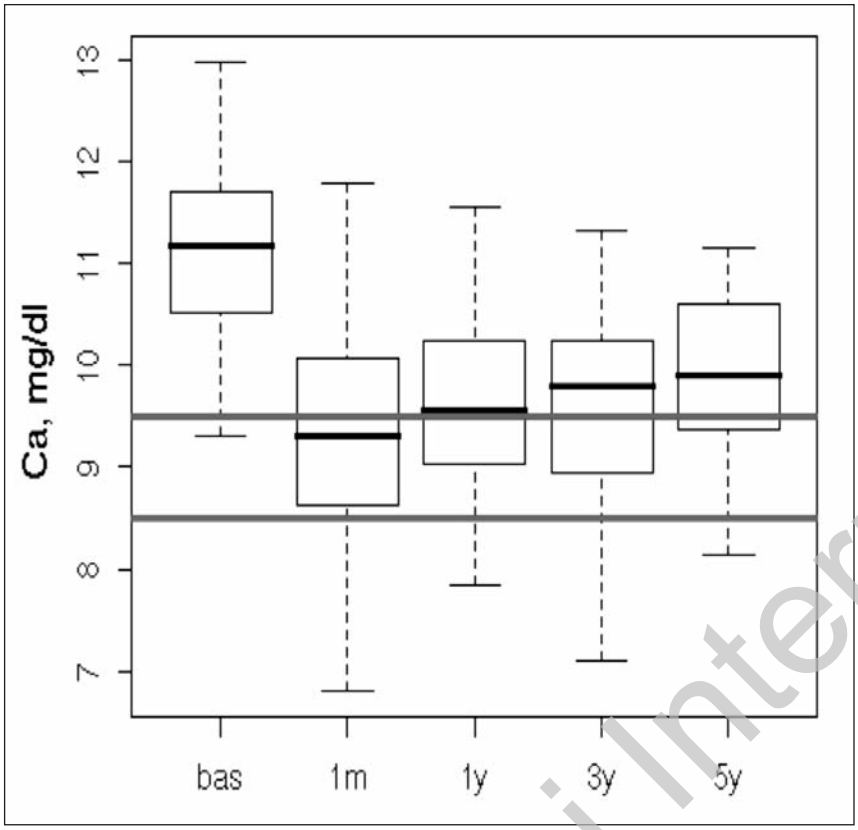

Figure 1. Boxplot of observed Serum Calcium levels after PTX. The red lines indicate the K/DOQI thresholds.

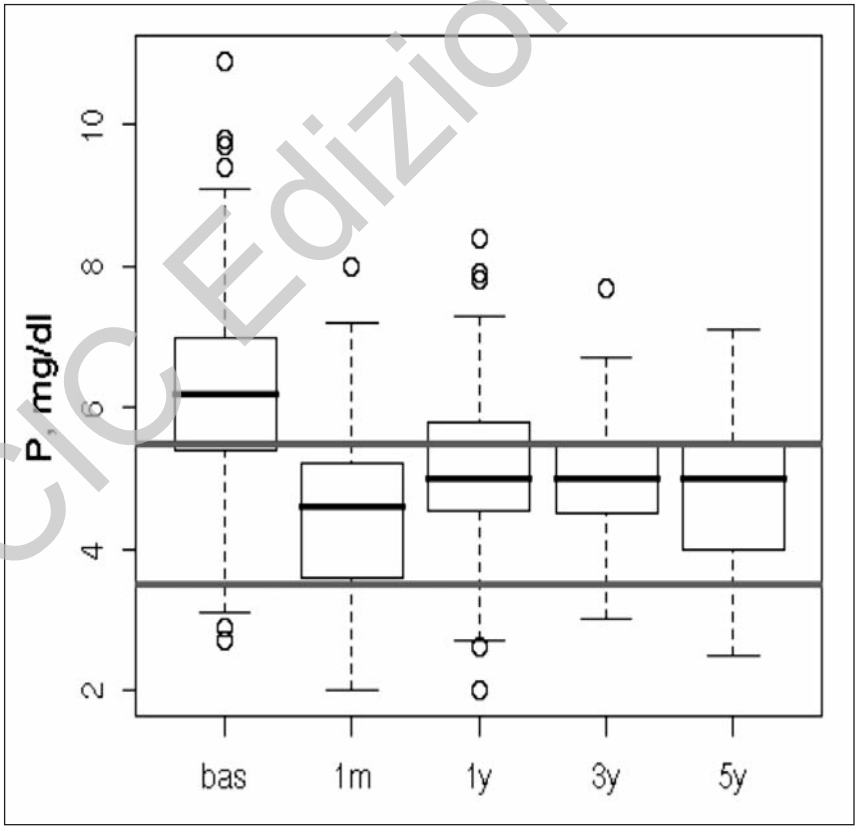

Figure 2. Boxplot of observed Serum Phosphate levels after PTX. The red lines indicate the K/DOQI thresholds.

being excluded a significant effect of these drop-outs by fitting models with a categorical variable indicating reason of drop-out, both the univariate and multivariate analyses; and also including a general dummy variable for possibly informative drop-outs. In all cases we recorded very large p-values (in all cases, $p > 0.8$), indicating a possibly negligible effect.

3. Methods

To formalize, let n denote the number of subjects. Each subject has been observed for a number of times, say T_i, $i = 1,\ldots, n$. For subject i at time t we have two binary outcomes $\boldsymbol{Y}_{1it}$ and $\boldsymbol{Y}_{2it}$. The vectors $\boldsymbol{Y}_{1i}$ and $\boldsymbol{Y}_{2i}$ contain the repeated measures of the categorial outcomes for subject j.

A very common modeling approach would fit a logistic model with random effects separately on each outcome. Formally, the Y_{it} outcome would be modeled as

$$\log \frac{P(Y_{it} = 1|\boldsymbol{X}_{it})}{P(Y_{it} = 0|\boldsymbol{X}_{it})} = \alpha_i + \boldsymbol{\beta}'\boldsymbol{X}_{it}, \qquad (1)$$

where β is a vector of logistic regression coefficients, $\boldsymbol{X}_{it}$ is a vector of possibly time-dependent covariates for the i-th subject at time t, and $\alpha_i \sim N(0, \sigma^2)$, $i = 1,\ldots, n$, $t = 1,\ldots,T_i$. For a review of such models, see for instance (5).

There are different approaches in the literature for simoultaneously fitting two binary outcomes, which are mostly generalizations of model (1). For instance, see (6) or (7).

We prefer to use a different approach since we question the normality assumption for the random effects, and the assumption that they are time fixed. We prefer to allow for subject *and* time specific intercepts α_{it}. See (8) for a thorough discussion of this topic. The model we propose is based on two logits and one log-odds ratio:

$$\log \frac{p(y_{1it} = 1|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})}{p(y_{1it} = 0|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})} = \alpha_{1it} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_1 + \boldsymbol{y}'_{i,t-1}\boldsymbol{\gamma}_1$$

$$\log \frac{p(y_{2it} = 1|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})}{p(y_{2it} = 0|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})} = \alpha_{2it} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_2 + \boldsymbol{y}'_{i,t-1}\boldsymbol{\gamma}_2$$

$$\log \frac{p(y_{1it} = 1, y_{2it} = 1|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})}{p(y_{1it} = 1, y_{2it} = 0|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})} +$$

$$+ \log \frac{p(y_{1it} = 0, y_{2it} = 0|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})}{p(y_{1it} = 0, y_{2it} = 1|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, \boldsymbol{y}_{i,t-1})} = \phi$$

We denote with α_{hit} a subject and time specific random intercept and with β_h a vector of regression parameters for the h-th response variable, $h = 1; 2$. The log-odds ratio is denoted with ϕ.

We further introduce two vectors of regression parameters Υ_h which are used to estimate the direct effect, adjusted for the covariates, of experimenting the event at the previous time on the probability of experimenting the event at current time. This is known as *state dependence*.

A full account of the model and a description in full generality (i.e., for arbitrary number of arbitrary categorical outcomes) can be found in (4). The main difference with (4) is that they use a closed panel assumption that $T_i = T\ \forall i$.

We stress that, even if our notation suggests so, not necessarily the same covariates must be used on each marginal logit. This is a particularly useful feature of our marginal logit parameterization since it can be expected in general applications that different covariates have significant effects on different outcomes, even if they may be related. Refer to (4) for further discussion.

Finally, it is straightforward to check that modelling the outcomes with two marginal logits and the log-odds does not impose any restriction on the joint distribution.

3.1 Random effects

The random intercepts α_{hit} are commonly assumed to arise from a normal distribution, possibly with an autoregressive structure to model change over time. In this work we prefer to use a Markov process, which avoids the questionable normality assumption and provides more information at the interpretation stage. Hence, assuming there are k latent states, the random effects model can be summarized with the initial and transition probabilities $p(\boldsymbol{\alpha}_{i1} = \boldsymbol{\xi}_c) = \boldsymbol{\lambda}_c(\boldsymbol{y}_{i0}, \boldsymbol{x}_{i0})$ and $p(\boldsymbol{\alpha}_{it} = \boldsymbol{\xi}_d|\boldsymbol{\alpha}_{i,t-1} = \boldsymbol{\xi}_c) = \boldsymbol{\pi}_{cd}$, where we use a logit parameterization to allow the initial probabilities to depend on a baseline and covariates.

For each i and t, the random parameter vector $\boldsymbol{\alpha}_{it}$ is assumed to arise from a *discrete* distribution with k support points, which are denoted by $\boldsymbol{\xi}_c$, $c = 1,\ldots,k$. The k support points can be thought of being latent intercepts, which are estimated together with the other model parameters.

For fixed k, the model is then fully described by the regression parameters, the k latent intercepts and the mixing distributions $P(\boldsymbol{\alpha}_{it} = \boldsymbol{\xi}_c)$, $c = 1,\ldots,k$.

In order to allow the intercepts to evolve over time, i.e., to allow subjects to move from a latent state (identi ed by ξ_c) to another, we assume a latent Markov structure. The Markov assumption is that α_{it} is independent of α_{iv},
$v < t - 1$, conditionally on $\alpha_{i,t-1}$. This assumption is seldom restrictive and is used to limit the number of free parameters.
Summarizing, for each i, the random parameter vectors $\{\alpha_{i1},\ldots, \alpha_{iT}\}$ are assumed to follow unobservable *first-order Markov chains* with

- initial probabilities $P(\alpha_{i1} = \xi_c) = \lambda_c(y_{i0}, x_{i0})$, $c = 1,\ldots,k$
- transition probabilities $P(\alpha_{it} = \xi_c | \alpha_{i,t-1} = \xi_d) = \pi_{cd}$; $c, d = 1,\ldots,k$

The initial probabilities are allowed to depend on baseline outcomes and covariates through a logit parameterization. Parameters for this further logit model and latent transition matrix $\Pi = \{\pi_{cd}, c, d = 1,\ldots,k\}$ are estimated together with β, γ, ϕ and ξ.
We give no details here on model fitting.
Refer to (4) for a description of an Expectation-Maximization (EM) algorithm for deriving the maximum likelihood estimates for the model.
At the E-step, we make use of forward and backward recursions adapted from the hidden Markov literature, see also (9).
At the M-step, we mostly set up ad-hoc Fisher-scoring iterations.

4. Results

We fit our model to the data about serum calcium (Y_1) and phosphate (Y_2).
Levels at baseline are used to model initial probabilities of the latent trait.
In order to select the number of states we use the Akaike Information Criterion (AIC) (10). We repeatedly fit the model for different values of k, and select the model minimizing AIC. We end up selecting a model with $k = 2$ latent states. Table 1 shows the estimated β, γ and ϕ parameters, an asterisk indicating significance at 5% level.
The strong negative trend has been observed also by (3), and has shed doubts over parathyroidectomy as a therapeutic tool for secondary hyperparathyroidism.
The state dependence estimators reveal that, after adjusting for possible confounders, both calcium and phosphate are likely to stay within the ranges once they have been targeted. Most importantly, it is better to target calcium (for instance, through subscription of vitamin D) since this significantly raises the probability that phosphate is targeted in the near future. State dependence of calcium on phosphate had not been unveiled in (3), and may help refining clinical treatment of hyperparathyroidism.

Table 1. Regression estimates for the proposed model fit on *K/DOQI* data. An asterisk indicates significance at 5% level.

Effect	logit fertility	logit employment	log-odds ratio
intercept	α_{1it}	α_{2it}	-0.055
t = 1y	-0.97	-1.06	-
t = 3y	-0.63	-1.75*	-
t = 5y	-1.08	-2.13*	-
age/100	0.01	-0.41	-
sex(Male)	-0.52	-0.31	-
lagged calcium	1.20*	0.59*	-
lagged phosphate	0.23	1.00*	-

We conclude this section demonstrating the need, at least for the data at hand, of allowing for time-varying subject-specific parameters and for relaxing the normality assumption. To do so, we fit a model with Normal random effects, and then a model with time-constant subject specific parameters.
The latter model is fit by constraining the hidden transition matrix to be diagonal, and corresponds to a latent class model. We perform a likelihood ratio tests, using results from (4) to evaluate significance. Both models are rejected with $p < 0.001$, indicating that time-constant random effects may be restrictive and that a Gaussian distribution may not adequately model the random effects.
Finally, in order to convince the reader about necessity of simoultaneous modeling of joint outcomes, we compare the predictive performance of the simoultaneous and of the separate models. To do so, we use 10-fold cross-validation. We randomly choose a test set of 10 subjects, fit the model on the remaining subjects, and use the fitted model to predict the outcomes of the 10 subjects. We repeat the operation one hundred times, each time randomly choosing a set of 10 subjects for validation.
The result of this cross-validation experiment are as follows: when one uses the joint model, the cross-validated

misclassification estimate for the joint outcomes is 0.11, with a 95% confidence interval (0.06-0.20). When separate models are fit on each of the two outcomes, the misclassification estimate raises to 0.39, with a 95% confidence interval (0.28-0.45). This happens even if the parameter for log-odds is not significant.

5. Discussion

We have illustrated a latent Markov model for two longitudinal binary out-comes. Matlab code for fitting the model is available from the web page `http://afarcome.interfree.it/codemarglong.zip`. The proposed model allows the subject specific parameters to change over time. The common assumption that the subject-specific parameters are time-constant is particularly at risk when the time horizon is large, like in the proposed motivating example. In that case, it is very likely that unobserved factors contributing to overdispersion have changed over time. The proposed model also relaxes the normality assumption for the random effects. It could be argued that normality assumption is bypassed at the price of assuming a discrete latent distribution, but it is well known that any smooth enough continuous density can be well approximated by a finite mixture of masses. Simulations in (4) hint in a more general framework that when the random effects are actually Normal, there is a very small bias in approximating the Normal distribution by a finite mixture. On the other hand, it is natural to expect that when the random effects are far from being Normal, a large bias can be expected in assuming normality.

The rejection of normality and the need for time varying random effects for the data at hand issue a warning to applied statisticians. Both assumptions should be checked when analyzing longitudinal binary data. A possible alternative if any of the two fails is to use our proposed model. The model can be easily adapted to the case of time-constant random effects by constraining the latent transition matrix to be diagonal, thereby only bypassing the normality assumption.

With respect to our application, we can confirm that in the long term (3-5 years) a tendency toward a loss of control of serum levels of P and

Ca exists. This phenomenon of disease recurrence, e.g. (11) and references therein, underlines the importance of medical surveillance and treatment. After carefully re-analyzing data from (3) with a more suitable simoultaneous model for the two outcomes of main interest, we can confirm that parathyroid surgery may not represent an optimal therapeutic tool for targeting the recommended ranges for Ca and P.

Acknowledgements: This study was partially supported by EIEF research grant “Advances in non-linear panel models with socio-economic applications”.

References

1. National Kidney Foundation. K/DOQI clinical practice guidelines: Bone metabolism and disease in chronic kidney disease. American Journal of Kidney Diseases 2003; 42:S1–S201.
2. Young E, Akiba T, Albert J, McCarthy J, Kerr P, Mendelssohn D, Jadoul M. Magnitude and impact of abnormal mineral metabolism in hemodialysis patients in the dialysis outcomes and practice patterns study (DOPPS). American Journal of Kidney Diseases 2004; 44:34–38.
3. Mazzaferro S, Pasquali M, Farcomeni A, Vestri A, Filippini A, Romani AM, Barresi G, Pugliese F. Parathyroidectomy as a therapeutic tool for targeting the recommended NKF-K/DOQI ranges for serum calcium, phosphate and parathyroid hormone in dialysis patients. Nephrology Dialysis Transplantation 2008; 23:2319–2323.
4. Bartolucci F, Farcomeni A. A multivariate extension of the dynamic logit model for longitudinal data based on a latent Markov heterogeneity structure. Journal of the American Statistical Association 2009; 104:816–831.
5. Pendergast J, Gange SJ, Lindstrom M, Newton MA, Palta M, Fisher MR. A survey of methods for analyzing clustered binary response data. International Statistical Review 1996; 64:89–118.
6. Ten Have TR, Morabia A. Mixed effects models with bivariate and univariate association parameters for longitudinal bivariate binary response data. Biometrics 1999; 55:85–93.
7. Todem D, Kim K, Lesaffre E. Latent-variable models for longitudinal data with bivariate ordinal outcomes. Statistics in Medicine 2007; 26:1034–1054.
8. Hsiao C. Analysis of Panel Data. Cambridge University Press: New York, 2005.
9. MacDonald IL, Zucchini W. Hidden Markov and other Models for Discrete-Valued Time Series. Chapman and Hall: London, 1997.
10. Akaike H. Information theory as an extension of the maximum likelihood principle. Second International symposium on information theory, Petrov BN, F C (eds.), Akademiai Kiado: Budapest, 1973; 267–281.
11. Tominaga Y, Katayama A, Sato T, Matsuoka S, Goto N, Haba T, Hibi Y, Numano M, Ichimori T, Uchida K. Re-operation is frequently required when parathyroid glands remain after initial parathyroidectomy for advanced secondary hyperparathyroidism in uraemic patients. Nephrology Dialisis and Transplantation 2003; 18:65–70.