Taylor & Francis
healthsciences

Application of the Monte Carlo Method for the Assessment of Long-term Success in Keratoprosthesis Surgery

W. HITZL* and G. GRABNER

Eye Clinic of the County Hospital Salzburg, St. Johanns-Spital Landeskliniken Salzburg, Muellner Hauptstrasse 48, 5020 Salzburg, Austria

The comparison of different methods of keratoprosthesis (KP) regarding their long-term success, as far as visual acuity is concerned, is difficult: this is the case both as a standardized reporting method agreed upon by all research groups has not been reported and far less accepted, and as the quality of life for the patient not only depends on the level of visual acuity, but also quite significantly on the "survival time" of the implant. Therefore, an analysis of a single series of patients with Osteo–Odonto–Keratoprosthesis (OOKP) was performed. Statistical analysis methods used by others in similar groups of surgical procedures have included descriptive statistics, survival analysis and ANOVA. These methods comprised comparisons of empirical densities or distribution functions and empirical survival curves. It is the objective of this paper to provide an inductive statistical method to avoid the problems with descriptive techniques and survival analysis. This statistical model meets four important standards: (1) the efficiency of a surgical technique can be assessed within an arbitrary time interval by a new index (VAT-index), (2) possible autocorrelations of the data are taken into consideration and (3) the efficiency is not only stated by a point estimator, but also 95% point-wise confidence limits are computed based on the Monte Carlo method, and finally, (4) the efficiency of a specific method is illustrated by line and range plots for quick illustration and can also be used for the comparison of different other surgical techniques such as refractive techniques, glaucoma and retinal surgery.

Keywords: Monte Carlo; Non-parametric model; Autocorrelation; VAT-index; Keratoprosthesis; OOKP

INTRODUCTION

Keratoprosthesis (KP) is a general term in ophthalmic surgery for any type of device that replaces the cornea of the human eye. These KPs are used whenever the chances of success (e.g. restoration of useful sight) for a conventional corneal transplant are minimal or non-existent, such as after severe lye burns or severe anterior segment diseases. Corneal blindness designates a state of vision loss that is caused by changes in the transparent anterior segment of the eye and that significantly precludes useful vision. In these instances often the retinal function (e.g. the activity of the deeper structures of the eye) can be sufficient for good vision, but cannot be used. KP is a type of ocular surgery that is performed at a comparatively infrequent rate when seen on a global scale. The reason for this is threefold: (a) the rarity of the cases that will profit from this type of procedure. It will only be considered in "only eyes" or in cases where both eyes have lost sight due to corneal blindness and not in patients where there is still a useful vision in at least one eye. (b) The complexity, cost and duration of the surgery. This is the cause for the comparatively small series of cases that can be accumulated and are reported by different research centers. At the most there are 30–40 surgeons worldwide to perform this type of surgery on a regular basis and the "KPro-study group" meetings are attended by approximately 100–120 participants every two to three years. (c) The ambiguities of evaluation of the results of these very complex surgical procedures on a long-term basis.

Several different types of KP implants are currently in use at comparatively few surgical centers worldwide for the treatment of otherwise intractable corneal blindness. The comparison of these differing surgical methods concerning their long-term success is an important issue among ophthalmologists, as progress in this field has been quite slow over the last decades and clinical stability

---

*Corresponding author. Tel.: +43-676-5233860. Fax: +43-662-44823703. E-mail: w.hitzl@lks.at

as well as visual outcomes over the "long-term" period
(>1 year) still leave much to be desired with most of the
currently available methods in these very severe cases.

The surgeons working in this field are, therefore,
interested in the clarification of several aspects related to
the efficiency of these comparatively few methods that are
currently in use, e.g. (1) "How well is an 'average patient'
able to see at a given time, postoperatively?", (2) "How
stable will the visual acuity of the patient be over time?",
(3) "Will the 'average' patient run the risk to possibly lose
his vision completely, when longer periods of time are
considered?" and (4) "If different surgical techniques are
compared, which one is better or worse at a given point in
time/within a given period?" Previous authors have used
different statistical methods to answer these questions.
The studies of Caselli *et al.* (1999), Pintucci *et al.* (1999)
and Liu and Pagliarini (1999) computed the empirical
density functions of their samples, however, no confidence
intervals of the corresponding percentages of the achieved
visual acuity are stated in their results. Hence, these results
only reflect descriptive aspects of their data and no
inferences concerning the quality of a surgical technique
as such can be made. The Kaplan–Meier method and the
log-rank test were used by Toledo *et al.* (1999) to test
different survival functions and Taloni *et al.* (1999) used
ANOVA to test contrast sensitivity, perimetric thresholds
and color vision.

From a statistical point of view, the above assessments
of the efficiency of a surgical method are unsatisfactory
and this is the case for several reasons. The most important
problem with descriptive methods is that the analysis does
not end with "statistically stable" results, i.e. should the
study be repeated several times over, it is very likely that
considerably different results would be obtained. This
problem was tackled by Toledo *et al.* (1999) and Taloni
et al. (1999), however, in both cited studies, serious
statistical problems still remain. The study of Toledo *et al.*
(1999) used the log-rank test for a comparison of survival
curves. As the failure criterion (terminal event), a final
visual acuity worse than preoperative or lower than 0.05
was chosen. The first disadvantage of this approach is the
use of the log-rank test. This test compares several
survival curves and is therefore not able to make
inferences across several survival curves within a *specific
time interval*. Compared with a method that can assess the
efficiency within a given time interval, this method
implies a loss of information. The second drawback is that
one does not receive any information about the expected
visual acuity (or another visual functional parameter) at a
given time point. However, this is of crucial interest for
the eye-specialist when comparing different methods.
Another approach was used in the study of Taloni *et al.*
(1999), which applies the ANOVA model for a series of
33 OOKP patients, in which the time after surgery ranged
from 1 to 16 years. The ANOVA model requires
measurements are made according to the necessities of
the clinical follow ups, which are made on different times
for different patients.

It is the aim of this paper to describe a method, based on
the statistical model of Rice and Silverman (1991) and in
conjunction with the Monte Carlo method (bootstrap), to
cope with the above-mentioned problems. The bootstrap is
a powerful computer-based method for statistical analysis.
It allows scientists to explore data and draw valid
statistical inferences without worrying about mathemati-
cal formulas and derivations. In its non-parametric form,
the bootstrap provides standard errors and confidence
intervals without the usual normal-theory assumptions.
The bootstrap method applied to the sample with 16
longitudinal sequences consists of resampling with
replacement from the sample, i.e. a random sample of
size 16 is drawn from the original sample and then the
VAT-index is computed. Again, another random sample of
size 16 (from the original sample) is drawn and again, the
VAT-index based on the new values is computed. This type
of simulation is done for $B = 1000$ times, where B is the
so-called bootstrap sample size. Based on these simu-
lations, more accurate confidence intervals can be
computed.

In order to assess the efficiency of a type of surgery as
objectively as possible, this paper suggests a minimum of
four standards that should be met.

First of all, the surgeon should be able to evaluate the
efficiency of a surgical method within a time interval
that has been set arbitrarily. This would allow
partitioning the postoperative time into *clinically
relevant* time intervals (e.g. the first 30 days, the first
6 months, the first year, the time period from the first to
the second year, etc.) and thereafter to compute the
efficiency for each method for exactly these time
intervals. This is of special interest for the comparison
of early postoperative phases and the long-term results
of different types of surgeries.

The second standard focuses on a possible existing
autocorrelation structure of the observations. This is
likely to happen, since the same patient is measured
over time (longitudinal study). For this reason, usual
regression methods with independent residual terms
should not be applied and models that take the
autocorrelation structure into consideration should be
preferred.

The third requirement concerns the proper statistical
assessment of the efficiency of a method, which should
not only be stated in terms of a point estimator, but also
should be expressed with 95% confidence intervals. The
substantial advantage of such a confidence interval is
that it allows to have a certain level of confidence or
reliability in the estimate.

Finally, the *fourth requirement* aims at the feasibility of
clear and comprehensive illustrations and the compara-
bility of the results with other types of KP surgery.

FIGURE 1   OOKP according to Strampelli, as modified by Falcinelly *et al.* The central optical cylinder, surrounded by healthy buccal mucosa that covers the anterior surface of the eye, is clearly seen.

This study was in fact undertaken to be able to better compare the visual results of the different study groups on a long-term basis and will improve the understanding of new developments in this field. We also feel that this new technique will help to better evaluate the visual results in a variety of different surgical fields such as refractive surgery, glaucoma surgery and vitrectomy studies.

METHODS

In order to describe and illustrate the proposed method, this model is applied to a total of 16 patients of the Eye Clinic of the County Hospital in Salzburg, Austria. Each patient underwent Osteo–Odonto–Keratoprosthesis-surgery (OOKP), a complex technique that uses the patients own tooth-bone-periosteal-complex as a lamina for the support of the optical cylinder. This optical cylinder is central part of the OOKP-implant that consists of a transparent material to allow a formation of a clear image at the surface of the retina and is fixed to the dental-bone-periosteal lamina by means of special glue. After
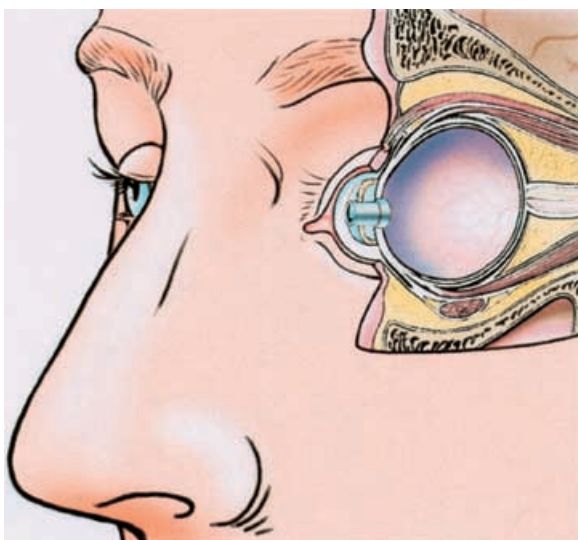

FIGURE 2   Schematic representation of the OOKP on the anterior surface of the eye of the patient, courtesy Prim. Univ.-Doz. Dr Chr. Krenkel.
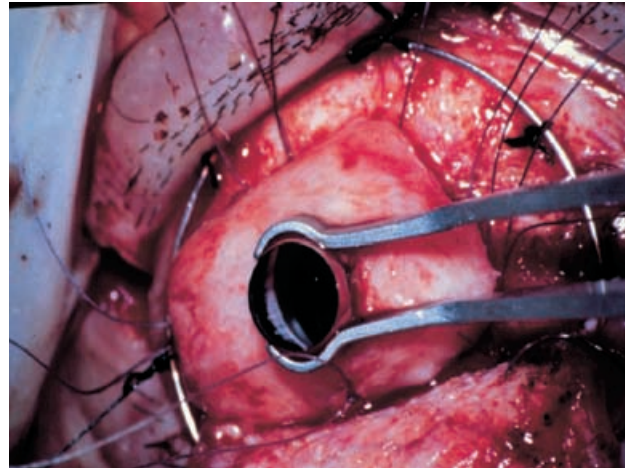

FIGURE 3   Fixation of the lamina is achieved by interrupted sutures, as shown. The optical cylinder is grasped with the help of a special forceps and fixed on the surface of the cornea by the assistant during this step of the surgery.

the preparation of the lamina in the operating room under sterile conditions it is implanted for a duration of two to three months in a skin pouch of the lower lid under the contralateral eye. At this time point it is explanted, it has also been covered by a newly formed periosteum and can be positioned and fixed on the corneal surface of the eye to be treated with the help of interrupted sutures following a trephination for the posterior part of the optical cylinder. The whole eye is then covered with the initially transplanted buccal mucosal graft, again with a central trephination for the anterior part of the optical cylinder. The globe is then pressurized with air and visual acuity can be regained within a fairly short period. The technique has been described in detail elsewhere (Strampelli, 1963; Aquavella *et al.*, 1982; Roper-Hall, 1991; Liu and Pagliarini, 1999, Figs. 1–4).

From each subject, the best corrected visual acuity was measured over a time period from 1994 to 2001,
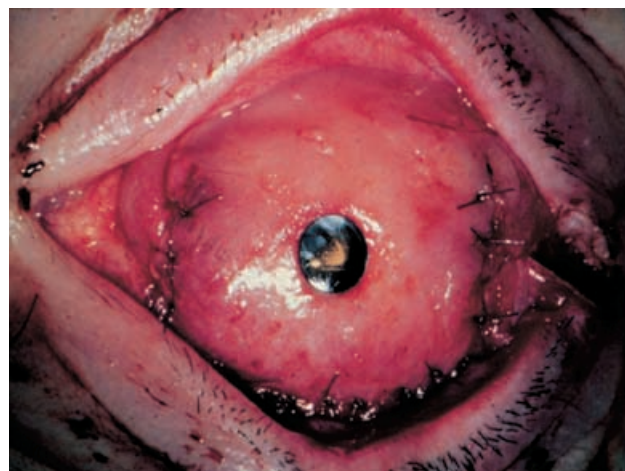

FIGURE 4   The lamina is finally covered by the buccal mucosa that was used to cover the anterior surface of the globe at the end of the initial (first step) of the procedure, at the time when the lamina was prepared for the implantation in the skin pouch.

with the maximum time of follow up being 2700 days. The best corrected visual acuity is the best possible achievable vision, regardless of what device or devices are necessary to help to get the best corrected visual acuity and ranges from 0 (no vision) to 1.6 (excellent vision). In five cases, the implant had to be removed after different periods of time. Due to different mechanisms, that are not clear in all cases (e.g. extremely rapid resorption of the lamina in children), the tooth-bone-periosteal-complex became unstable and the eyes had to be closed by other means, such as a penetrating keratoplasty (transplantation of a human donor cornea (in parts) onto a recipients eye to restore vision following different diseases that are amendable by this type of conventional surgery), to salvage the globe. From this time point on, the best corrected visual acuity was set to 0 until the end of the observation time (2700 days) with an increment of 100 days.

The best corrected visual acuity is estimated at any chosen time point using a weighted and smoothed average of the observations for all the patients (kernel estimation, Rice and Silverman, 1991). A window is centered at a given time t and the estimate of the mean response at time t is the weighted average of all points which are within this window. As a weighting function, the Gaussian kernel is used, which gives more weight to the observations close to t. The bandwidth of the kernel is the width of the window at time t. There is a continuum of curves, one for each possible window width. The wider the window, the smoother the resulting curve. To obtain an estimate of the smooth curve at every time, the window is moved from the extreme left across the data to the extreme right, calculating the weighted average of the points within the window at every time.

Confidence limits are derived for the estimates taking account the autocorrelations for a given patient and an index (VAT-index: Visual Acuity by Time index) is suggested that should meet the four standards described above. This index estimates the efficiency of KP surgery within a specified period and can be considered to "slide" from the far right (day 0) to the very left one (day 2700). These confidence limits are based on a bootstrap method (BCa method: bias corrected and accelerated) and achieve second order accuracy (Hall, 1992). This method is described in greater detail by Hall (1992), Efron and Tibshirani (1993) and Shao and Tu (1995).

Non-parametric Modeling of Mean Response

The $m = 16$ patients were measured repeatedly through time, whereby data were collected prospectively, following subjects forward in time. These measurements on the different patients are not made at a common set of time points. For this reason, the following notation is given. Let $(y_{i,1}, \ldots, y_i, n_i) \in \mathbb{R}^{n_i}$ be the vector of $n_i \in \mathbb{N}$ measurements on the ith patient ($\forall i \in 1, \ldots, m$) and $(t_{i,1}, \ldots, t_i, n_i) \in \mathbb{R}^{n_i}$ the corresponding set of times

($t_{i,j} \geq 0, \forall j \in \{1, \ldots, n_i\}$) at which the best corrected visual acuities were measured. In order to highlight the average change of visual acuity over time, the model

$$y_{ij} = \mu(t_{ij}) + \epsilon_i(t_{ij}),$$
$$\forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n_i\} \tag{1}$$

described by Diggle *et al.* (1994) is used throughout this paper, where $\{\epsilon_i(t), t \in \mathbb{R}\}, \forall i \in \{1, \ldots, m\}$ are independent realizations of a stationary random process, $\{\epsilon(t), t \in \mathbb{R}\}$ with variance $\sigma^2 > 0$, and the mean response $\mu(t)$, is a smooth function of t, i.e. $\mu'(t)$ is continuous. In order to find a non-parametric estimate of $\mu(t)$, Diggle *et al.* (1994) use the symmetric, non-negative kernel function (Gaussian kernel),

$$K(u) = \exp\left(\frac{-u^2}{2}\right) \tag{2}$$

and define the weights

$$\overset{*}{w}_{ij}(t) = h(t)^{-1} K\{(t_{ij} - t)/h(t)\} \tag{3}$$

with the function $h(t)$. This function describes the bandwidth of time-window and is estimated by the data. To find a reasonable function $h(t)$, Diggle *et al.* (1994) suggest to treat the function $h(t)$ as a constant h and then to adapt the kernel estimator with a new function $h(t)$. A description how this can be done is given later on. Diggle *et al.* (1994) uses standardized weights

$$w_{ij}(t) = \overset{*}{w}_{ij}(t)\left\{\sum_{u=1}^{m}\sum_{v=1}^{n_i}\overset{*}{w}_{uv}(t)\right\}^{-1}, \tag{4}$$

so that

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} w_{ij}(t) = 1, \quad \forall t \in \mathbb{R}.$$

The non-parametric estimate of $\mu(t)$ is

$$\hat{\mu}(t) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} w_{ij}(t) y_{ij}. \tag{5}$$

A cross-validatory prescription for choosing a constant h is given by Rice and Silverman (1991), which makes no assumption about the underlying correlation structure. For a given h, let $\hat{\mu}^{(i)}(t)$ be the estimate of $\mu(t)$ obtained by Eq. (5), but omitting the ith subject. The suggested prescription of Rice and Silverman (1991) chooses the constant h so as to minimize

$$S(h) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\{y_{ij} - \hat{\mu}^{(i)}(t_{ij})\}^2 \tag{6}$$

With

$$w_i(t) = \sum_{j=1}^{n_i} w_{ij}(t),$$

the expression $\hat{\mu}^{(i)}(t)$ can be expressed as

$$\hat{\mu}^{(i)}(t_{ij}) = \hat{\mu}(t_{ij})$$

$$-\frac{w_i(t_{ij})}{1 - w_i(t_{ij})}\left\{\sum_{v=1}^{n_i} \frac{w_{iv}(t_{ij})y_{iv}}{w_i(t_{ij})} - \hat{\mu}(t_{ij})\right\}. \quad (7)$$

Using to compute $S(h)$ avoids the need for explicit computation of the m leave-one-out estimates, $\hat{\mu}^{(i)}(t)$.

VAT-Index

In order to describe the efficiency of a surgical technique with unknown expectation of the best corrected visual acuity $\mu(t)$ at time t and to meet the four standards defined above, the index

$$\mathrm{VAT}(\mu, T_1, T_2) := \frac{1}{T_2 - T_1}\int_{T_1}^{T_2}\mu(t)\,\mathrm{d}t \quad (8)$$

is considered. This index can be interpreted as the average achieved best corrected visual acuity of a surgical technique within the time $[T_1, T_2]$ and can be computed for every time interval $[T_1, T_2]$ of arbitrary length. If one assumes that $\mu(t)$ is continuous—this is reasonable in the case of vision—then $\mathrm{VAT}(\mu, T_1, T_1)$ is equal to $\mu(T_1)$ for any time point T_1.

Confidence Limits Based on the BCa-method

In the non-parametric model described by Diggle *et al.* (1994), no explicit confidence limits for $\mu(t)$ are given. However, confidence limits for $\mu(t)$ and for the VAT-indices are required, but no explicit formulas are known to the authors to compute such confidence limits. The bootstrap method based on the BCa method, described by Hall (1992), Efron and Tibshirani (1993) and Shao and Tu (1995) was applied to the unknown mean response $\mu(t)$ and to the VAT-indices as test statistics Θ.

In a first step, $B = 1000$ independent bootstrap replications $\{\hat{\Theta}^*(b), b \in 1, \ldots, B\}$ of the test statistic Θ were computed. The mean $\hat{\Theta}$ was estimated based on the original data and the corresponding bias-correction $\hat{z}_0$ was computed, whereby

$$\hat{z}_0 = \Phi^{-1}\left(\frac{|\{b \in \{1,\ldots,B\}, \hat{\Theta}^*(b) < \hat{\Theta}\}|}{B}\right). \quad (9)$$

and Φ^{-1} is the inverse distribution function of the standard normal distribution. In order to compute the acceleration parameter $\hat{a}$, jackknife values $\{\hat{\Theta}_i, i \in 1, \ldots, m\}$ based on the original sample with the ith patient

deleted and

$$\hat{\Theta}_{(.)} = \sum_{i=1}^{m} \hat{\Theta}_i / m$$

was defined. The acceleration parameter is given by

$$\hat{a} = \frac{\sum_{i=1}^{m}(\hat{\Theta}_{(.)} - \hat{\Theta}_{(i)})^3}{6\left\{\sum_{i=1}^{m}(\hat{\Theta}_{(.)} - \hat{\Theta}_{(i)})^2\right\}^{3/2}} \quad (10)$$

The lower 0.025-quantile of the bootstrap distribution is given by

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + \Phi^{-1}(0.025)}{1 - \hat{a} + \Phi^{-1}(0.025)}\right) \quad (11)$$

and the upper 0.975-quantile of the bootstrap distribution is given by

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + \Phi^{-1}(0.9725)}{1 - \hat{a} + \Phi^{-1}(0.9725)}\right) \quad (12)$$

The confidence limits for θ are given by the α_1 and α_2 quantiles of the empirical bootstrap distribution.

Numerical Computations

The confidence limits for the mean of the best corrected visual acuity were computed at $0,2,4,\ldots,100$ and $100,125,150,\ldots,2700$ days. The confidence limits for the VAT-indices were computed for intervals of length comprising 60, 180 and 360 days. In addition, the VAT-indices for the periods ranging from the time of the surgery to the first, second, third, fourth, fifth, sixth and seventh year were computed, whereby for numerical integration a MATHEMATICA built-in third order polynomial as interpolation function with an increment of 10 days was used. All computations and illustrations were done with MATHEMATICA 3.0.1 (Wolfram, 1991).

RESULTS

Results Concerning the Model

All four requirements for the requested index are met by definition of the VAT-index. The index can be computed for any time interval of arbitrary length, including the case $[T_1, T_1]$, i.e. an interval with length 0, which means the best corrected visual acuity at the single time point T_1 after surgery. The model suggested by Rice and Silverman (1991) takes the autocorrelation into account, the computations of the bootstrap confidence limits ensure that the VAT-indices are estimated on the basis of interval estimators and not only on point estimators. The illustration of the VAT-index will be discussed later. The model was computed for values of h equal to 40, 60,
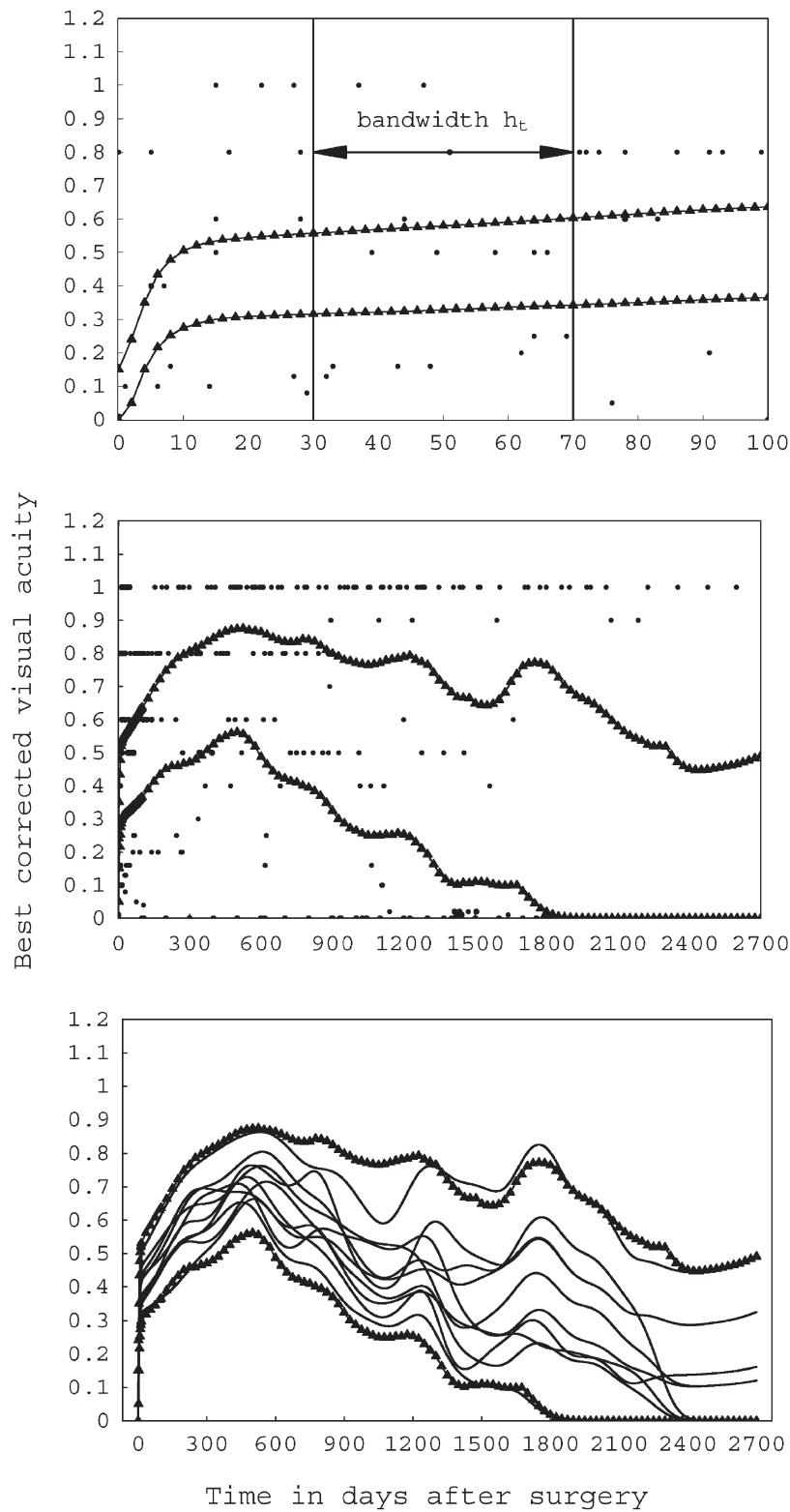
FIGURE 5   Plot of the non-parametric regression line $\hat{\mu}(t)$ with an illustration of the Gaussian kernel smoother (see text) and 95% pointwise confidence bands based on the BCa method for best visual acuity for OOKP patients within (a) [0,100] days and (b) [0,2700] days. The last figure (c) illustrates estimations for the non-parametric regression line generated by 10 bootstrap replicates.

80, 100 and 120 days and gave the following results for $S(h)$: 62.35, 62.15, 62.11, 62.18 and 62.40. A model with $h = 80$ described the data well, apart from the first 10 days, postoperatively. In addition, we computed the mean response curve and the corresponding confidence limits with $h = 40$ and $h = 120$ days, checked the results by eyes and could not found a substantial difference. Finally, the constant h was replaced by a function $h(t) = 80(1 - \exp(-t^2/100)) + 1$, to obtain a better approximation of $\mu(t)$ within the first 10 days.

Results Concerning OOKP

The initial phase after surgery (0,100 days) with corresponding 95% BCa confidence intervals is shown in Fig. 5a, the results for the total observation time ranging from 0 to 2700 days are given in Fig. 5b. The following results are reported based on the 95% confidence intervals for the expected best corrected visual acuity at a fixed time t. An illustration of 10 typical estimations of the mean response based on bootstrap replicates are given in Fig. 5c.

The results show that the OOKP patients start with an extremely low expectation of best corrected visual acuity of about (0, 0.15). This average is increasing rapidly over the first 10 days up to (0.27, 0.51), which means that there is a substantial visual improvement to be observed already within the initial time. After this period, there is a further increase of the best-corrected visual acuity beginning at day 10 until day 500. Within this time, the best-corrected visual acuity is improving from (0.27, 0.51) up to (0.56, 0.87). Beginning with day 500 until about day 2700, the best-corrected visual acuity is permanently decreasing until day 2700 with a final value covered by the 95% confidence interval (0, 0.49).

The results for the 95% lower and upper confidence limits of the VAT-indices with two-monthly, half-yearly, yearly steps and for the times ranging from the beginning of the surgery to the end of the first, second, third, fourth, fifth, sixth and seventh year were computed and shown in Fig. 6a–d, respectively.

For example, the VAT-index within the first two years ranges from 0.42 to 0.72, the VAT-index for the total observation time ranges from 0.25 to 0.7.

DISCUSSION

The assessment of a surgical procedure requires a statistical model that fulfills important prerequisites to provide the clinician with meaningful information, especially under complex circumstances such as in the context of evaluation of different KP results. This paper points out that (at least) four standards should be met. The major advantage of this model is the computation of confidence limits, since these bounds measure the precision of the estimated values, i.e. they reflect the amount of knowledge about the "true" value. Wider intervals indicate lower precision, as this is the case around 1800 days after surgery, whilst narrow intervals indicate higher precision (e.g. within the first 300 days). Consequently, a comparison of different surgical techniques can be carried out with much more care and precision. Clearly, an increase of the sample size would end up with smaller confidence intervals and hence would reflect higher precision.

The statistical methods applied by Caselli *et al.* (1999), Pintucci *et al.* (1999) and Liu and Pagliarini (1999) are inadequate from a statistical point of view, since they use
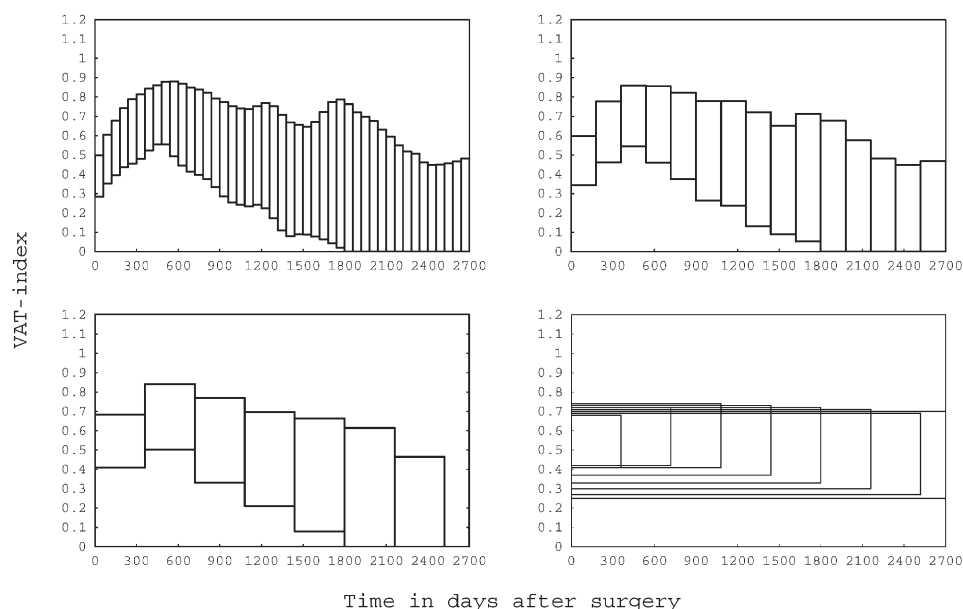

FIGURE 6 Illustration of different VAT-Indices with 95% confidence limits based on the BCa method for (a) 60 days (b) 180 days, (c) 360 days and (d) for the intervals [0,360], [0,720], [0,1080], [0,1440], [0,1800], [0,2160], [0,2520] and [0,2700] for the assessment of the long-term best corrected visual acuity in keratoprosthesis surgery.

descriptive methods. Strictly speaking, the conclusions of their results refer only to their own data in one study. It still remains unclear, what happens, if these studies are repeated.

The paper of Toledo *et al.* (1999) applying survival analysis (log-rank test) is more careful in the evaluation of the data, as the standards 1 and 4 are met to some extent. However, the survival analysis does not give any detailed information about the expectation of the best corrected visual acuity neither at a given time point nor within a chosen period of time.

As has been emphasized by Neter *et al.* (1990) and Diggle *et al.* (1994), the medical statistician should not apply regression models with uncorrelated error terms with autocorrelated data. This is mainly due to the fact, that whenever the residual terms are positively auto-correlated, the use of ordinary least squares procedures has a number of important consequences, e.g. the estimated regression coefficients are still unbiased, but they no longer have the minimum variance property and may be quite inefficient (Neter *et al.*, 1990). The mean square error (MSE) may thereby seriously underestimate the variance of the error terms. The estimations of the standard deviations of the regression coefficients calculated by ordinary least squares procedures may seriously underestimate the true standard deviation of the estimated regression coefficient.

We, therefore, propose the statistical method detailed above for the comparisons of different surgical techniques in a variety of ophthalmic specialities, e.g. KP and refractive techniques, but also in glaucoma and retinal surgery.

References

Aquavella, J.V., Rao, G.N., Brown, A.C. and Harris, J.K. (1982) "Kerato-prosthesis. Results, complications, and management", *Trans. Am. Acad. Ophthalmol. Otolaryngol.* **89**(6), 655–660.

Caselli, M., Colliardo, P., Falcinelli, G. and Nebbioso, M. (1999) "Falcinelli's osteo-odonto-keratoprosthesis: long-term results", *Ann. Inst. Barraquer (Barc.)* **28**(Suppl.), 113–114.

Diggle, P., Liang, K. and Zeger, S. (1994) *Analysis of Longitudinal Data Oxford Statistical Science Series 13*, (Clarendon Press, Oxford), pp. 106–110.

Efron, B. and Tibshirani, R. (1993) *An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability 57* (Chapman & Hall, New York), pp. 184–188.

Hall, P. (1992) *The Bootstrap and the Edgeworth Expansion Springer Series in Statistics*, (Springer, New York), pp. 135–137.

Liu, C. and Pagliarini, S. (1999) "Independent survey of long term results of the Falcinelli osteo-odonto-keratoprosthesis (OOKP)", *Ann. Inst. Barraquer (Barc.)* **28**(Suppl.), 91–93.

Pintucci, S., Pintucci, F., Cecconi, M. and Gaiazza, S. (1999) "The Pintucci dacron tissue KP: long term results, postoperative care and revisions in dry eyes and in eyes with tear secretion", *Ann. Inst. Barraquer (Barc.)* **28**(Suppl.), 109–112.

Rice, J.A. and Silverman, B.W. (1991) "Estimating the mean and covariance structure nonparametrically when data are curves", *J. Royal Stat. Soc. B* **53**, 233–243.

Roper-Hall, M. (1991) "Why are keratoprostheses not given more attention?", *Eur. J. Implant. Refract. Surg.* **3**, 79–81.

Shao, J. and Tu, D. (1995) *The Jackknife and the Bootstrap* (Springer, New York), pp. 135–140.

Strampelli, B. (1963) "Keratoprosthesis with osteodontal tissue", *Am. J. Ophthalmol.* **89**, 1023–1039.

Taloni, M., Falsini, B., Caselli, M., Micozzi, I., Piccardi, M. and Falcinelli, G.C. (1999) "Assessment of central visual function in patients with osteodontokeratoprosthesis", *Ann. Inst. Barraquer (Barc.)* **28**(Suppl.), 133–134.

Toledo, J.A., Barraquer, J.T., Carreras, H., Torres, E. and Barraquer, J. (1999) "Osteo-odonto-keratoprosthesis. A 30 years retrospective study", *Ann. Inst. Barraquer (Barc.)* **28**(Suppl.), 95–100.

Wolfram, S. (1991) *Mathematica: A System for Doing Mathematics by Computer* (Addison Wesley, New York).