Parametric and Nonparametric Population Methods:

Their Comparative Performance in Analysing a Clinical Data Set

and Two Monte Carlo Simulation Studies

Aida Bustad BA[1], Dimiter Terziivanov MD[2], Robert Leary PhD[3], Ruediger Port MD[4], Alan Schumitzky PhD[1], and Roger Jelliffe MD[1]

1. Laboratory of Applied Pharmacokinetics, USC Keck School of Medicine, Los Angeles CA, USA (bustad@usc.edu, jelliffe@usc.edu)
2. Clinic of Clinical Pharmacology and Pharmacokinetics, University Hospital "St. I. Rilsky", Sofia, Bulgaria (terziiv@yahoo.com)
3. San Diego Supercomputer Center, San Diego CA. Present address: Pharsight Corp, Cary, NC, bleary@Pharsight.com
4. German Cancer Research Center, Heidelberg, Germany.

Running title: Parametric and Nonparametric Population Approaches

ACKNOWLEDGEMENTS**:** Supported by US Government grants RR11526 and GM65619.

Corresponding Author: Roger Jelliffe, M.D., Laboratory of Applied Pharmacokinetics, Division of Geriatric and General Internal Medicine, USC Keck School of Medicine, 2250 Alcazar Street, Room CSC-134-B, Los Angeles, CA, 90033. Tel=1-323-442-1300, fax=1-323-442-1302, email=jelliffe@usc.edu. Many thanks also to Dr. Irina Bondareva, of the Institute for Physical-Chemical Medicine, Moscow, Russia, for her very thoughtful help and critiques in the preparation of this manuscript.

LEGENDS FOR FIGURES

Figure 1. Left, smoothed true simulated population joint density of K and V. The unimodal V is on the axis from the bottom corner to the left corner, and the bimodal K is on the axis from the bottom to the right corner. Right, the smoothed empirical density of 20 exactly known simulated subjects randomly sampled from the true density.

Figure 2. Graph of the IT2B marginal frequency of population parameter KA. The plot is divided, for display purposes only, into 100 cells over the range from 1.19 to 1.47 hr^{-1} (horizontal axis). The frequency of the patient parameter values in each cell is shown on the vertical. See text for discussion.

Figure 3. Graph of the IT2B marginal frequency of population parameter KS1. The plot is divided, for display purposes only, into 100 cells over the range from 0.0019 to 0.0041 hr^{-1} per unit of creatinine clearance (horizontal axis). The frequency of the patient parameter values in each cell is shown on the vertical. See text for discussion.

Figure 4. Graph of the IT2B marginal frequency of population parameter VS1. The plot is divided, for display purposes only, into 100 cells over the range from 0.15 to 0.32 L/kg (horizontal axis). The frequency of the patient parameter values in each cell is shown on the vertical. See text for discussion.

Figure 5. Joint marginal density of KS1 and VS1 obtained with the FOCE IT2B, left, and with NPAG, right. Note the very high, and probably erroneous, correlation between the parameters seen with IT2B, and the much more realistic correlations seen with NPAG. NPEM results (not shown) were similar to NPAG.

Figure 6. Scattergram of IT2B relationship between estimated serum concentrations (ug/ml, horizontal) and measured ones (vertical), based on median population parameter values.

Figure 7. Scattergram of IT2B relationship between estimated serum concentrations (ug/ml, horizontal) and measured ones (vertical), based on each subject's own maximum aposteriori probability (MAP) Bayesian posterior parameter values, where each subject predicts only his/her own measured concentrations.

Figure 8. Graph of the NPAG marginal density of population parameter KA. The plot is divided, for display purposes only, into 100 cells over the range from 0.0 to 6.0 hr^{-1} (horizontal axis). The estimated probability of the parameter values in each cell is shown on the vertical. See text for discussion.

Figure 9. Graph of the NPAG marginal density of population parameter KS1. The plot is divided, for display purposes only, into 100 cells over the range from 0.0 to 0.008 hr^{-1} per unit of creatinine clearance (horizontal axis). The estimated probability of the parameter values in each cell is shown on the vertical. The distribution is more skewed to the left than that from IT2B, shown in Figure 3. See text for discussion.

Figure 10. Graph of the NPAG marginal density of population parameter VS1. The plot is divided, for display purposes only, into 100 cells over the range from 0.0 to 0.60 L/kg (horizontal axis). The estimated probability of the parameter values in each cell is shown on the vertical. See text for discussion.

Figure 11. Scatterplot of predicted (horizontal) and measured (vertical) serum concentrations (ug/ml) based on the population median parameter values, obtained using the NPAG program. NPEM results (not shown) were similar to NPAG.

Figure 12. Scatterplot of predicted (horizontal) and measured (vertical) serum concentrations (ug/ml) based on the median parameter values of each subject's Bayesian posterior joint probability density, predicting only that subject's own data, using the NPAG program. NPEM results (not shown) were similar to NPAG.

Figure 13. Plot of the 800 Monte Carlo simulated subjects (stars) whose parameter values (V, horizontal, and K, vertical) were exactly known prior to the analysis. The circles represent the 70 support points into which this population was resolved by the NPAG program. Note that the means, standard deviations, and the correlation of V and K were well captured by the NPAG program.

Figure 14. Consistency of estimators of the mean of V. The true mean of V is 1.1. NPAG and PEM (top and middle) are consistent. Their estimates approach the true value as more subjects are analyzed. FOCE IT2B (bottom) is not consistent. Results actually stray from the true values as more subjects are studied.

Figure 15. Consistency of estimators of mean of K. True value of mean K is 1.0. NPAG and PEM (middle and bottom) are consistent. Again, the estimates approach the true value as more subjects are studied. FOCE IT2B (bottom) is not consistent. Results stray from the true values as more subjects are studied.

Figure 16. Consistency of estimators of SD of K. True SD is 0.25. NPAG and PEM (top and middle) are consistent. Results approach the true value as more subjects are analyzed. FOCE IT2B (bottom) is not consistent. Results drift away at least 25% from the true value as more subjects are studied.

Figure 17. Consistency of estimators of correlation coefficient between V and K. The true value is 0.0. NPAG and PEM (middle and bottom) are consistent. FOCE IT2B (top) is not consistent, drifting away from the true value, and is severely biased.

Figure 18. Asymptotic convergence rate of IT2B (top) is much less than that of NPAG and PEM (essentially superimposed at bottom). For NPAG and PEM, the estimated SD decreases by half as 4 times the number of subjects are analyzed, as predicted by theory. For the FOCE IT2B,

top, however, fully 16 times the number of subjects are required to decrease the estimated SD by half.

Figure 19. As convergence proceeds, the true log-likelihood (vertical) increases monotonically with PEM (top), but with the FOCE IT2B approximation (bottom), the true likelihood reaches a high point and then decreases, while the FOCE approximation (not shown) is what actually increases monotonically with IT2B as the true log-likelihood drifts downward.

Figure 20. Distribution of 50 NPAG estimates of the variance of V (top) and of 50 NONMEM FO estimates (bottom). NPAG parameter estimates are much more precise.

Figure 21. (Axes as in Figure 1). Left: the best parametric representation of the joint population density using the assumption of a joint Gaussian distribution, as might be seen with PEM, for example. The actual distribution cannot be seen at this initial step, as other information such as covariates or the collection of individual MAP Bayesian posterior values, for example, is required to suggest to the user to anticipate a bimodal, trimodal, or some other specifically assumed mixture distribution. The actual most likely distribution may well not be seen, because of the assumptions made about the shape of the mixture distribution. Right, smoothed results from NPEM. The actual bimodal distribution is easily recognized by NPEM at this initial step, without any further information. See text for discussion.

ABSTRACT

Introduction. This study examined parametric and nonparametric population modeling methods in three different analyses. The first was of a real, though small, clinical data set from 17 patients receiving intramuscular Amikacin. The second was of a Monte Carlo simulation study in which the populations ranged from 25 to 800 subjects, the model parameter distributions were Gaussian, and all the simulated subjects' parameter values were exactly known prior to the analysis. The third analysis was again of a Monte Carlo study in which the exactly known population sample consisted of a unimodal Gaussian distribution for the volume of distribution V, but a bimodal one for the elimination rate constant K, simulating rapid and slow eliminators of a drug.

Methods. For the clinical data set, the parametric iterative 2-stage Bayesian (IT2B) approach, with the First Order, Conditional Expectation (FOCE) approximate calculation of the conditional likelihoods was used [1], and the nonparametric expectation maximization (NPEM) and nonparametric adaptive grid (NPAG) approaches, both of which use exact computations of the likelihood [1].

For the first Monte Carlo simulation study, the above programs were also used. A 1-compartment model with unimodal Gaussian parameters V and K was employed, with a simulated IV bolus dose and two simulated serum concentrations per subject In addition, a newer parametric EM (PEM) program, with a Faure low discrepancy computation of the conditional likelihoods [2], and also NONMEM, both the first order (FO) and the FOCE versions, were used.

For the second Monte Carlo study, a 1 compartment model with an intravenous (IV) bolus dose was again used, with five simulated serum samples obtained from early to late after a dose. A unimodal distribution for V but a bimodal one for K were chosen, to simulate two subpopulations, of “fast” and “slow” metabolizers of a drug. NPEM results were compared that of a unimodal parametric joint density having the true population parameter means and covariance.

Results. For the clinical data set, the interindividual parameter percent coefficients of variation (CV%) were least with IT2B, suggesting less diversity in the population parameter distributions. However, the exact likelihood of the results was also least with IT2B, and was 14 logs greater with NPEM and NPAG, which found a greater and more likely diversity in the population studied.

For the first Monte Carlo data set, NPAG and PEM, both using accurate computations, showed statistical consistency, in accordance with theory [3,4,16]. Consistency means that the more subjects studied, the closer the estimated parameter values approach the true ones. NONMEM FOCE and NONMEM FO, and the IT2B FOCE methods do not have this guarantee [3,4]. Results obtained by IT2B FOCE, for example, often strayed visibly from the true values as more subjects were studied.

Further, with respect to statistical efficiency (precision of parameter estimates), NPAG and PEM had good efficiency and precise parameter estimates, while precision suffered with NONMEM FOCE and IT2B FOCE, and severely so with NONMEM FO.

For the second Monte Carlo data set, NPEM closely approximated the true bimodal population joint density, while an exact parametric representation of an assumed joint unimodal density having the true population means, SD’s, and correlation gave a totally different picture.

Conclusions. The smaller population interindividual CV% estimates with IT2B on the clinical data set are probably the result of assuming Gaussian parameter distributions, and/or of using the FOCE approximation. NPEM and NPAG, having no constraints on the shape of the population parameter distributions, and which compute the likelihood exactly and estimate parameter values with greater precision, detected the more likely greater diversity in the parameter values in the population studied.

In the first Monte Carlo study, NPAG and PEM had more precise parameter estimates than either IT2B FOCE or NONMEM FOCE, and very much more precise estimates than NONMEM FO. In the second Monte Carlo study, NPEM easily detected the bimodal parameter distribution at this initial step, without requiring any further information.

Population modeling methods using exact or accurate computations have more precise parameter estimation, better stochastic convergence properties, and are, very importantly, statistically consistent. Nonparametric methods are better than parametric ones at analyzing populations having unanticipated non-Gaussian or multimodal parameter distributions.

1. INTRODUCTION

Population pharmacokinetic (PK) and pharmacodynamic (PD) modeling is done to describe experience with a drug in a collection of patients in a way that will be useful for understanding the drug's behavior, and especially for subsequent use as a Bayesian prior in the treatment of similar patients who must receive that drug in the future. The present report describes a comparison of parametric and nonparametric population modeling methods in three situations: first, to analyze a real, though small, clinical data set of patients receiving

intramuscular amikacin; second, to analyze a Monte Carlo simulation study of a population in which each subject's parameter values were simulated, and therefore were exactly known prior to the analysis, as can never be done with real clinical data; third, to analyze a second Monte Carlo simulated population having both rapid and slow eliminators of a drug.

The parametric iterative 2-stage Bayesian (IT2B) approach, employing the widely used first-order, conditional expectation (FOCE) approximation to compute the conditional likelihoods of the results given the data, was the parametric method used to analyze the clinical data set. This was compared with the nonparametric expectation maximization (NPEM) method [10,11,17] and the nonparametric adaptive grid (NPAG) method [2]. Further, NONMEM, both its first order (FO) and first order, conditional expectation (FOCE) versions [7-9], and a new parametric expectation-maximization (PEM) method [17] were also used for the first Monte Carlo study. NPEM was used for the second Monte Carlo analysis.

2. A BRIEF REVIEW OF PARAMETRIC APPROACHES TO POPULATION PHARMACOKINETIC MODELING

A variety of parametric population modeling methods exist [5-9]. They assume that the pharmacokinetic parameters are either normally or lognormally distributed in the population studied, and that the population parameter distributions are therefore fully described by the estimated parameter means, standard deviations (SD's), and the correlations (and covariances) between them, as the means, SD's and covariances are the definitive parameters in the equations describing the shape of Gaussian distributions.

2.1 THE PARAMETRIC ITERATIVE TWO-STAGE BAYESIAN (IT2B) METHOD

The IT2B population modeling method, used here from the USC*PACK collection of software [1] begins by setting up an initial estimate of the mean values for each parameter in the pharmacokinetic structural model, and their standard deviations (SD's). These are used as Bayesian priors. In the first stage, the method examines each patient's data and obtains each patient's maximum a posteriori probability (MAP) Bayesian posterior parameter values [12]. Then, in the second stage, the new summary parameter means and SD's are found. The IT2B approach then uses these revised summary population parameter means and SD's as a new initial Bayesian prior, and does another MAP Bayesian analysis of the data. It again first obtains each patient's new MAP Bayesian values, and then again summarizes them as new population parameter means and SD's. This two-stage process continues iteratively, and ends when a convergence criterion is reached. The IT2B method described here uses the First Order Conditional Expectation (FOCE) approximation to calculate the conditional means and covariances of the population parameters, given the population raw data [1,17].

2.2 Weighting the Data: Determining the Assay Error Polynomial

It is useful, in both parametric and nonparametric analyses, to assign a measure of credibility to each data point to be fitted or analyzed. In the IT2B program, and also in the NPEM and NPAG programs of the USC*PACK collection [1], one is encouraged, first of all, to determine the error pattern of the assay quite specifically, by measuring several assay samples in at least quadruplicate (a blank, a low sample, a medium one, a high one, and a very high one, for example, that cover the full working range of the assay), and to find the mean and SD for

each of these samples [13,14]. One can then express the overall relationship between serum concentration and assay SD as

$$SD = A_0 + A_1C + A_2C^2 + A_3C^3 \quad (1)$$

where SD is the assay SD, A_0 through A_3 are the coefficients of the polynomial, C is the measured concentration, C^2 is the concentration squared, and C^3 is the concentration cubed. In this way, each assay data point can be given a weight in the modeling process appropriate to the precision with which it was measured, and thus to its credibility, according to its Fisher information [13-15].

2.3 Weighting the Data: Estimating the Remaining Environmental Error

In addition, a further parameter, gamma, can then yield an estimate of the overall contribution of the remaining environmental sources of intra-individual variability, such as the errors in preparation and administration of the doses, errors in recording the times when doses were given and serum samples drawn, the mis-specification of the pharmacokinetic model used, and any unsuspected changes in parameter values of the subjects during the period of the data analysis. Gamma was used in the USC*PACK IT2B program [1] as a multiplier of the above assay error polynomial, as shown below. It is now also implemented in the NPAG population modeling program, but this was not yet done at the time of this analysis.

$$\text{Total error SD} = \text{Gamma}(A_0 + A_1C + A_2C^2 + A_3C^3) \quad (2)$$

If the value of gamma is 1.0, it suggests that there is no other source of variability than the assay error pattern itself. Gamma is usually greater than 1.0, but may sometimes be less.

Gamma is an overall reflection of all the other sources of intraindividual variability besides the assay error. In this way, one can get an impression of how much of the total noise is due to the assay SD, and how much is due to the remaining environmental noise. This is useful, as it gives separate information both about the precision of the assay and about the precision of the environment in which the study was done.

2.4 THE PARAMETRIC EXPECTATION – MAXIMIZATION (PEM) METHOD.

This parametric method [2,17] is exactly similar to the IT2B approach described above, except that the FOCE approximate calculation of the conditional means and covariances is replaced by a Faure low discrepancy integrator, which gives an accurate computation.

3.0 A BRIEF REVIEW OF NONPARAMETRIC (NP) POPULATION MODELING METHODS

If the correct structural PK/PD model of the drug could somehow be exactly known, and if each individual subject's pharmacokinetic parameter values in a given population could also somehow be directly observed and exactly known (a clinical impossibility), and if, for example, we were examining two typical model parameters such as volume of distribution (V), and elimination rate constant (K), the true population distribution of these parameter values would simply be the collection of each individual subject’s exactly known values of V and K. For such a two-parameter model, for example, the graphic result would be a scattergram of the collection of those points, one point for each subject studied. All genetically determined subpopulations would also be truly known (perhaps not explicitly recognized or classified yet), but nevertheless specifically located and quantified. If this distribution were summarized simply as means, SD’s,

and correlations, important information about the distribution, which is often genetically polymorphic, would be lost. However, in this “ideal” model, no statistical statements can be made about the reproducibility of the parameter values, such as confidence limits of the parameter distributions or their means, SD’s, etc.

Lindsay [19] and Mallet [16] were the first to show that the optimal maximum likelihood solution to the population modeling problem is a discrete (not continuous), spiky (not smooth) probability distribution in which no preconceived parametric assumptions (such as Gaussian, lognormal, or other) are made about its shape. The nonparametric maximum likelihood (NPML) estimate of the population model joint parameter distribution, whatever its shape or distribution turns out to be, is resolved into at most N discrete points for the N patients in the population studied. Each such discrete support point is a collection of estimated single parameter values, one for each model parameter such as V, K, etc., along with an estimate of the probability associated with each such set or combination of values. The probabilities of all the support points in the population sum to 1.0. Parameter means, SD’s, and covariances are easily obtained as well. The only assumption made about the shape of the discrete parameter distributions is that, for each model parameter, the shape, whatever it is, is the same for all subjects in the population [10,16]. The NPML method of Mallet [16] can function with only one sample per patient, if required. Just as with the so-called “ideal” model, however, no statistical statements about confidence limits can be made for the NP models. However, the various percentile values of the specific NP parameter distributions are easily obtained.

A nonparametric expectation-maximization (NPEM) method was developed by Schumitzky [10,11]. It is an iterative EM method, but is nonparametric. Like NPML, it also can

function with only one sample per patient if needed. Like NPML, it also does not make any parametric assumptions about the shape of the joint probability distribution. It also computes the entire discrete joint distribution of population support points. In contrast to NPML, though, NPEM obtains a continuous (although very spiky) distribution, which finally becomes discrete in the limit. With each iteration, NPEM examines the patient data and develops a more and more spiky (and more likely) joint population parameter distribution. In the limit, the spikes become discrete support points, up to one for each subject studied, each of which contains a set of parameter values, each set of which has a certain probability, just as with the NPML method.

Both the NPML and the NPEM programs approach the unattainable “ideal” population model described earlier. Both NPML and NPEM have been shown to converge to essentially the same results [20]. Both NPML and NPEM are proven under suitable hypotheses to have the desirable property of mathematical consistency [3,4,21]. This important property of consistency means that the greater the number of subjects that are studied in a population, the closer the results obtained approach the true population parameter values. If a method is not statistically consistent, it is not proven to have that behavior, and there is no guarantee that studying any more subjects will actually yield any better or more trustworthy results. In fact, as shown later, results can actually get worse as more patients are studied.

Most currently available parametric maximum likelihood population modeling methods do not have the proven property of consistency, as they use approximations such as the FO (first order) or FOCE (first order, conditional expectation) software to compute the conditional likelihoods, which then are not exact. Many workers using current parametric population modeling methods such as NONMEM, for example, often seem not to have reported the actual

value of the likelihood of the results given the raw data and the error model used. Instead, they have used indices of “goodness of fit”, the confidence limits of the parameter values, and the bias and precision of predictions of subsequent serum concentrations. However, the recent class of Bayesian population analysis methods [33] has the property of consistency

The nonparametric methods compute the likelihood exactly, simply by summing the discrete support points rather than having to integrate a continuous function. They give percentile estimates of the dispersion of the parameter distributions, but not confidence limits. This is an important difference between the parametric and nonparametric methods.

However, the likelihood of the IT2B and the NPEM collections of support points can be computed and compared directly, under exactly similar conditions, as the final IT2B collection of MAP Bayesian posterior support points can be analyzed in the IT2B program in the USC*PACK collection just as if it had come from NPEM or NPAG, and its final likelihood can be computed exactly [1]. This feature was implemented in the USC*PACK IT2B software specifically to permit such comparisons.

Since it is never possible to know any patient's parameter values exactly in real life, we must infer them or estimate them from the data of the doses of the drug given and the responses (usually serum concentrations) found. We study a sample of patients requiring therapy with a drug (the most relevant population sample) by giving the drug, measuring serum concentrations and/or other responses, and estimating the population model parameter distributions. On the other hand, using Monte Carlo simulation, we may study a sample of simulated subjects in which each subject’s parameter values are exactly known prior to the

analysis, so that we can then see just how well the method in question works to discover the true simulated population parameter distributions.

Many patient populations are made up of genetically determined clusters or subpopulations such as fast and slow metabolizers of a drug. The relative proportions of fast and slow subjects may vary from one population (Caucasian people, for example) to another (Asian people, for example [18]). Describing a real distribution of subpopulations (whatever it is) optimally may be very difficult parametrically, assuming normal, lognormal, or multimodal distributions. One must somehow anticipate and assume in advance, based on other information, that the population parameter distribution will be bimodal or multimodal, for example.

The word nonparametric (NP) is also not to be confused with noncompartmental modeling approaches based on statistical moments, which also are sometimes called nonparametric. The NP approach described in the present report always has a specific structural model, with its specific model differential equations, and their model parameters (volumes, rate constants, clearances, etc.). It also has a specific error model like the assay and environmental error polynomial described above.

3.1 NEWER DEVELOPMENTS IN NONPARAMETRIC POPULATION MODELING

A significant improvement in NP modeling was then made by Leary [2]. The original NPEM strategy, based on a large fixed grid covering the parameter space, was computationally quite intensive. Leary [2] showed that the likelihood of the results obtained correlated strongly with the number of grid points used in the computations. The quality of the NPEM result (the log-likelihood) thus depends on the number of grid points used in the analysis. A powerful

machine and much computer time were often needed to obtain precise results with the original NPEM, using many grid points to get good resolution over the ranges of the parameters.

Leary then developed a new nonparametric “adaptive grid” (NPAG) procedure, which combined with an interior point rather than an EM algorithm (as suggested to Leary by Burke [31]), made significant advances in the quality, speed and memory requirements for an NP population analysis [2]. The NPAG method thus begins with a smaller and coarser grid. After this is initially solved and the support points found, the grid is refined by adding perturbations (extra grid points, about 10 for each previous solution support point), near them. Using this new grid, the problem is solved again. Once again, new grid points are placed near the previous solution points. This process then continues iteratively, using decreasing perturbations, gradually using a finer and finer adaptive resolution of the grid, until a convergence criterion is met.

NPAG made a significant improvement in quality of the results, with far less overall computational time and effort. For example, analyzing an 8 subject, 5 parameter problem with the original NPEM, on the 1152 processor IBM Blue Horizon computer at the San Diego Supercomputer Center (at the time the fastest non-classified computer in the world), took 2037 processor-hours, used 164 million grid points and 10000 Mbytes of memory. It achieved a likelihood of -433.1. In contrast, NPAG, running only on a single processor 833 MHz Dell PC, used only 5000 grid points, only 6 Mbytes of memory, took only 1.7 processor hours. It achieved a likelihood of -433.0. NPAG has greatly reduced the computational time and memory requirements compared to NPEM, and many population modeling tasks can now be done on a single processor notebook PC which used to require a large parallel mainframe machine.

4.0 METHODS: THE CLINICAL DATA SET

This retrospective clinical data set was obtained from patients all of whom received intramuscular (IM) Amikacin therapy. It presented a common problem in population analysis, permitting a useful comparison of the methods described here. The data set was also a reasonable one for understanding the IM rate of uptake of the drug somewhat better (although certainly not perfectly) than with other studies where it is usually given intravenously. This data set was obtained by one of us (DT) [22] from 17 adult patients with urinary tract infections who received intramuscular Amikacin, 1000 mg every 24 hours for 5 or 6 days. For each patient, two clusters of 4 (or sometimes 5) serum concentrations were measured, one cluster on the first day and the other on the 5th or 6th day, with serum samples taken at 1.0, 3.0, (sometimes at 5.0), and at 7.0 and 23.5 hours after the dose. Maximum measured serum concentrations ranged up to 48 ug/ml, usually at 1 hr after the dose, and trough concentrations were less than 2.0 ug/ml. Creatinine clearance (CCr) was estimated from data of age, gender, serum creatinine, height and weight [23]. Of the patients, 6 were male and 11 female, age ranged from 52 to 74 years, height ranged from 61 to 73 inches, weight ranged from 53 to 105 kg, and CCr ranged from 38 to 87 ml/min/1,73M^2. The patients were all clinically stable from day to day, and their renal function also was stable. All patients had eradication of the urinary pathogen after 3 days of their 5-6 day course of therapy, and no relapse seen with one year of follow-up. No nephrotoxicity or ototoxicity was seen.

Initial parameter estimates for the IT2B analysis were set at (mean ±SD) 0.1 ± 0.1 hr^{-1} for the elimination rate constant Kel, 25 ± 25 L for the apparent volume of distribution V, and

1.5 ± 1.5 hr^{-1} for the absorptive rate constant KA. Bioavailability of IM Amikacin was assumed to be 100%.

Initial analyses with IT2B used a basic model with parameters KA, the absorptive rate constant, KE, the elimination rate constant, and total apparent volume of distribution V. However, the initial results showed strong correlations between V and the covariate body weight, and between KE and the covariate CCr. These results were also found with NPEM and NPAG. Because of this, the model was reparameterized as KS1, with respect to the covariate CCR, and VS1, with respect to the covariate body weight. Thus KE (hr^{-1}) = KI (hr^{-1}) + KS1 (hr^{-1} per unit of CCr) x CCr (ml/min/1.73 M^2) and V (L) = VS1(L/kg) x body weight (kg). KI was held fixed at 0.0069325 (hr^{-1}), appropriate to a half-time of 100 hrs in an anuric patient.

The same data of the 17 patients receiving intramuscular Amikacin were also analyzed using both the NPEM and the NPAG software [22]. As above, the parameters were KA, the absorptive rate constant from the intramuscular injection site, VS, the volume of distribution in L/kg, and KS, the increment of elimination rate constant per unit of creatinine clearance. Initial ranges for these parameters were set at 0 to 6.0 hr^{-1} for KA, 0 to 0.6 l/kg for VS, and 0.0000001 to 0.008 hr^{-1} per unit of CCr for KS1. Gamma was set at 3.2158, the value previously obtained with the IT2B analysis. Again, KI was held fixed at 0.0069325 hr^{-1}.

4.1. THE AMIKACIN ASSAY

Amikacin concentrations in serum were assayed microbiologically within 48 h by an agar well diffusion method employing medium 1 according to USP XXI, using square plates (28 by 28 cm) with 49 wells and *Bacillus subtilis* ATCC 3399. The agar was inoculated with a suspension

adjusted by an optic standard to a density of approximately 10^9 CFU/ml. The final inoculum was 2 x 10^7 CFU/ml. Each assay was performed in triplicate. In preliminary in vitro studies, antibiotic standards were prepared both in pH 7.4 buffer and in serum. No difference in the sizes of zones of inhibition induced by a single antibiotic concentration was found between buffer and serum. The standards (ranging from 1 to 20 μg/ml in pH 7.4 buffer) and sera were assayed without dilution. The plates were read after incubation at 37 °C for 18 h. The detection limit of the microbiological method was 1 μg/ml and intra- and inter-day variations were below 5%. The explicit error polynomial of this assay was found to be

$$\text{Assay SD} = 0.12834 + 0.045645C, \qquad (3)$$

where C is the serum concentration.

4.2 METHODS: THE FIRST MONTE CARLO SIMULATION STUDY

Since it is not possible ever to know the parameter values exactly in clinical studies, a Monte Carlo simulation study was done in which each simulated subject's parameter values were exactly known prior to the analysis. The initial objectives were to examine and compare the statistical consistency, efficiency, and asymptotic convergence rate of NPAG with that of IT2B (using the FOCE approximation), and also with a new parametric EM (PEM) method developed initially by Schumitzky [17,21], but now using the more recent Faure low discrepancy sequence integration method [2], which yields accurate computations of the required integrations [21].

A simulated population in which the parameter distributions were in fact truly Gaussian was studied. This favored the parametric population modeling program. A one-compartment, two-parameter model was used, with parameters apparent volume of distribution V and elimination rate constant K. The true mean V was 1.1, with standard deviation 0.25. The true mean K was 1.0, also with SD = 0.25. The true correlation coefficient between the two parameters was set at three different values in three different simulated scenarios: -0.6, 0.0, and +0.6. A single unit intravenous bolus dose was given to a simulated “average” patient at time zero, and a sparse data set of two simulated “measured” serum concentrations were used, an early one after the dose, somewhat similar to a “peak” sample, and a later sample similar to a “trough”, each with a 10 % coefficient of variation. Simulated populations ranging in size from 25 to 800 subjects were studied.

4.3 METHODS: THE SECOND MONTE CARLO STUDY, OF A NONGAUSSIAN DISTRIBUTION.

Another Monte Carlo simulation study [10] examined results from an NPEM analysis when the true simulated population parameter distribution was not Gaussian, but bimodal. Again, a 1 compartment model with parameters V and K was used, an intravenous bolus dose was given, and five simulated serum concentrations were obtained, each one obtained at approximately equal time intervals from early to late after the dose. In the population, V and K were assumed to be independent. V had a mean of 2.0, with SD 0.2. K, however, was bimodal, a mixture of half slow eliminators with mean 0.5 with SD 0.05, and half fast eliminators with mean 1.5 with SD 0.15 units. The bimodal (for K) population joint parameter density is shown in Figure 1, left. Twenty simulated subjects were randomly sampled from the above population.

Their empirical, exactly known parameter values are shown, smoothed, in Figure 1, right. The task of population modeling now is to discover this true sample joint parameter density.

Figure 1 here

5.0 RESULTS: CLINICAL DATA SET - THE PARAMETRIC IT2B ANALYSIS.

The IT2B program converged on the 918th iteration. The mean values found for the population parameters KA, KS1, and VS1 were 1.349 hr^{-1}, 0.00326 hr^{-1}(per unit of CCr), and 0.2579 L/kg respectively. The median values were 1.352 hr^{-1}, 0.00327hr^{-1}(per unit of CCr), and 0.2591 L/kg respectively. The population parameter standard deviations were 0.062 hr^{-1}, 0.000485 hr^{-1}(per unit of CCr), and 0.0350 L/kg respectively, yielding population parameter coefficients of variation of 4.55, 14.83, and 13.86 percent respectively. Gamma was found to be 3.2158, showing that the SD of the environmental noise was probably about 3.2 times that of the assay SD, or conversely, that the assay SD was about 1/3 of the total of the assay and the environmental noise SD.

The individual subjects' IT2B distributions of KA, KS1, and VS1 are shown in Figures 2 through 4. While the distributions of KA and VS1 appeared to be fairly Gaussian, that of KS1 was slightly skewed to the left. The joint IT2B distribution of KS1 and VS1 is shown in Figure 5, left, which shows an extremely high positive correlation between the two parameters, consistent with their estimated correlation coefficient of +0.991. The correlation coefficient between KA and

KS1 (not shown) was similar, (+0.924), and that between KA and VS1 (not shown) was also very high (+0.950). These are probably spuriously high correlations found with IT2B, reflecting overparameterization of this problem for the IT2B method. Such very high correlations were not found with IT2B when gamma was held fixed at 1.0 and was not estimated.

Figures 2-5 here

Figures 6 and 7 are scattergrams of the IT2B estimated versus measured serum concentrations. Figure 6 shows the estimates based on the population parameter medians and the doses each subject received. In contrast, Figure 7 shows the estimates made using each subject's individual MAP Bayesian posterior parameter values (based on the population parameter medians and SD's as the Bayesian prior) to predict only his/her own measured serum concentrations. The improved estimates in Figure 7 are due to the removal of the population inter-individual variability, as perceived by the IT2B program. The remaining scatter is due to the intra-individual variability resulting from the assay error and the other sources of noise in the environment. The results, and the relatively low value of gamma (only 3.2), suggest that the clinical environment in the patients studied was probably managed with reasonable precision and care.

Figures 6-7 here

The final exact log-likelihood (not just the FOCE approximation) can also be computed exactly by IT2B, if desired [1], This permits a direct comparison of the results of the exact likelihood values obtained with the three methods [1]. The exact final log-likelihood with IT2B was -389.548 (see Table I).

Table 1 here

5.1 RESULTS: CLINICAL DATA SET - THE NONPARAMETRIC POPULATION ANALYSES

The results are summarized in Table I, where they are also compared with the previous results from the IT2B program. When comparing the results from IT2B, NPEM, and NPAG, at first glance there seems to be little difference between them. The mean and median parameter values were all quite similar. On closer inspection, though, the population interindividual parameter percent coefficients of variation (CV%) were always least with the IT2B program. This suggests that the population parameters were more narrowly distributed. However, the exact log-likelihood of the results was also clearly least with the IT2B parametric program [1], as shown in Table I. The log-likelihood was 14 logs greater with NPEM than with IT2B, and slightly greater still with NPAG. Similar differences in log-likelihood between a method using an approximate method to compute the likelihood versus one using an accurate one are also seen in Figure 19, further on, between the accurate parametric method PEM, and the approximate FOCE parametric method IT2B.

Since the likelihood was greater with the exact likelihood nonparametric methods, the smaller population parameter interindividual CV% values found with IT2B are probably due to its constraining assumption that the population parameters must have Gaussian distributions, and/or due to the FOCE approximate calculation used by the IT2B method.

Both NPEM and NPAG found a more likely greater diversity in the population parameter distributions studied. The specific marginal distributions of KA, KS1 and VS1 obtained with NPAG are shown in Figures 8 through 10. They should be compared with Figures 2 through 4 respectively, which were obtained with the IT2B program. Note that KS1 is skewed to the left.

Figures 8-10 here

An analysis of the scatterplots of estimated versus measured serum concentrations done by NPAG is shown in Figures 11 and 12, (and also those of the NPEM program, not shown), reveals that R^2, the coefficient of the determination (the fraction of the variance explained by the regression relationship between the estimated and the measured data), was slightly greater with the exact likelihood nonparametric programs. As shown in Table II, the mean error was the least, and the mean squared error was also the least, with NPAG, showing that the scatterplots actually were somewhat more correlated, probably less biased, and more precisely estimated by the two nonparametric modeling methods.

Table II here

Figs 11-12 here

5.2 RESULTS: THE FIRST MONTE CARLO SIMULATION STUDY

Figure 13 shows the exactly known parameter values of the 800 simulated subjects in the population studied (stars). As shown, this population was well resolved into 70 estimated support points (circles) by NPAG. The size of the circle represents the estimated probability of each support point.

Figure 13 here

Figure 14 shows that the accurate likelihood NPAG and PEM programs have consistent behavior. As the number of subjects examined in the simulated population increased from 25 to 800, the estimated mean apparent volume of distribution V approached closer and closer to the true value, while the value estimated by the IT2B program with the FOCE approximation actually drifted visibly away from the true value as more subjects were studied.

Figure 14 here

Figure 15 shows that the same was true for the mean elimination rate constant K. The estimates of K with NPAG and PEM approached the true value as the number of subjects in the population increased, while the IT2B FOCE estimates were again not consistent, and again drifted away from the true value as more subjects were studied.

Figure 15 here

Figure 16 shows the same behavior for the estimates of the SD of K. NPAG and PEM had consistent behavior while the FOCE IT2B estimate again drifted away. Similar behavior was found for the estimation of the SD of V (not shown).

Figure 16 here

Figure 17 shows that the estimation of the correlation coefficient between V and K with NPAG and PEM was consistent, approaching the true value of 0.0 more and more closely as the number of subjects increased, while FOCE IT2B was severely biased, starting at about +0.5 instead of the true value, and then increasing further up to about +0.65. Similar quite biased and incorrect behavior (not shown) was also seen with the FOCE IT2B when the true simulated correlation coefficient was -0.6, and also when it was +0.6. Since gamma was not estimated here, the high correlations may have been due to the rather large assumed assay error of a 10% CV. Both the results here and those described earlier when estimating gamma in the clinical data set with IT2B may have been due to the overall error model used, and (in the present Monte Carlo simulation) to the sparse data of only two samples per subject. Further study of this problem is warranted.

Figure 17 here

Furthermore, with respect to asymptotic convergence, as shown in Figure 18, bottom, in order to decrease the estimated SD of a parameter by half, 4 times as many subjects were required by the accurate NPAG and PEM, consistent with asymptotic theory. In contrast, fully 16 times as many subjects were required by the approximate FOCE IT2B program, top, to obtain the same reduction in SD.

Figure 18 here

Figure 19 shows that the true log-likelihood increased monotonically with PEM, but not with IT2B, where the true likelihood reached an early high, and then decreased to a lower value. This difference in likelihoods is similar to that seen between NPAG and IT2B in the analysis of the clinical data set and shown in Table I. This behavior in likelihood is very similar to the behavior shown in Figures 14 through 17, in which the parameter values similarly reached somewhat of a “best point” with a few subjects, and then got worse as more subjects were studied. The behavior of the likelihood here and the parameter estimates in Figures 14-17 may well be related.

Figure 19 here

Figure 20 shows a frequency histogram comparing 50 NPAG estimated values of the variance of the estimated volume of distribution (V), at top, with 50 similar estimates made by

the NONMEM program using the FO approximation, at bottom. This was done on a simulated population of 200 subjects, for 50 separate analyses. Very much greater precision in parameter estimation is seen with NPAG.

Figure 20 here

In summary, NPAG and PEM, which have exact or accurate computations, had consistent behavior, good efficiency with precise parameter estimates, and good asymptotic convergence. In contrast, the IT2B program, which used the FOCE approximation, was not consistent, with a small (1 – 2%) bias for the mean parameter values, a larger (20 – 30%) bias for the SD's, and a severe bias for the correlation coefficients, as described above. Further, the NPAG and PEM programs had much better asymptotic convergence, close to theoretical.

Especially disturbing in this study was the loss of statistical efficiency and precision of parameter estimation found with the FOCE approximation. This simulation study was recently extended by one of us (RP) to include the FO and FOCE approximations specifically as implemented in the parametric population modeling program NONMEM, version V.1.1 [30], The first order approximation in NONMEM FO had biases as high as 50% in estimates of variances, and statistical efficiencies (for the estimation of the mean volume of distribution, for example, as shown here in Table III) less than 2% of those of the accurate PEM and NPAG methods for 800 subjects, with a relative error over 100! NONMEM FOCE was a modest improvement over its IT2B FOCE counterpart. However, NONMEM FOCE still had significantly compromised

statistical efficiency (29%, with relative error 3.45), less than half that of the accurate methods (61% for NPAG with relative error 1.63, and 75% for PEM, with relative error 1.33) as shown in Table III. When the parameter distribution is truly known to be Gaussian, the efficiency and relative error of PEM are somewhat better than with NPAG, as predicted by theory.

Table III here

The PEM bias results shown in Table III were recently confirmed in an independent blind comparison of ten Nonlinear Mixed Effects parametric population modeling methods conducted by INSERM [34]. The methods included the NONMEM FOCE, Splus FOCE, and SAS FO approximate algorithms as well as other 'accurate likelihood' methods based on either Monte Carlo or Gauss-Hermite numerical quadrature. When analysing one hundred test problems generated by simulation from a one compartment first order absorption PK model with three random effects related to absorption, elimination, and volume of distribution, PEM exhibited the least bias among all methods.

5.3 RESULTS: THE SECOND MONTE CARLO SIMULATION STUDY

Since the differences between the accurate (PEM, NPEM, and NPAG) and the approximate (FOCE IT2B and NONMEM) methods were relatively small (in this example) with regard to the mean parameter values, although larger for the interindividual SD's, and only grossly different for the correlations (see above), many workers have continued to use such approximate methods. However, when one encounters unsuspected non-Gaussian parameter

distributions (difficult to see with parametric methods when having to assume various candidate multimodal distributions), things can be quite different. The results of the NPEM analysis of this Monte Carlo simulated bimodal population [10] gave the joint density shown in Figure 21, right. In the initial analysis, it easily discovered the two populations of simulated subjects without any further information, and was similar in shape to the empirical joint density shown in Figure 1, right.

Figure 21 here

In contrast, Figure 21, left, shows a joint unimodal Gaussian distribution having the same true means and covariances as that of the original population shown in Figure 1. It is similar to what would be obtained by PEM, and more accurate than what would be obtained by IT2B or NONMEM, if V and K had been assumed to be unimodal, which is why it was used for this comparison. This result gives a totally different and incorrect perception of the true density shown in Figure 1. These results illustrate the difficulties encountered using parametric modeling approaches when the parameter distributions are not truly Gaussian, and one has no other advance information such as covariates or the collection of MAP Bayesian posterior parameter values, for example, to permit one to assume a bimodal, trimodal, or some other mixture distribution, which will still always be constrained by the specific parametric assumptions used.

6. DISCUSSION:

The IT2B method of population modeling is based on the widely used strategy of MAP Bayesian individualization of pharmacokinetic models [12]. Like any parametric method, it perceives population parameter distributions only in terms of single point estimates of its means, modes, medians, variances, and correlations, and it computes the conditional likelihoods only in an approximate (FOCE) way.

A significant weakness of many, but not all, current parametric maximum likelihood methods is that they usually have lacked the desirable property of statistical consistency. With the parametric methods using the FO and FOCE approximations of the likelihoods, as shown earlier, there is no guarantee that the greater the number of subjects studied in a population, the closer the results obtained will actually approach more closely to the true parameter values. However, as shown here, the newer parametric PEM method using the Faure low discrepancy integration method, was in fact consistent, efficient, and had good convergence properties. When the parameter distributions were truly Gaussian, as in the first Monte Carlo simulation, PEM was also the most efficient and precise, as predicted by theory. That new method, though, was not available to analyze the clinical data set.

Parametric methods make parametric assumptions about the shape of the parameter distributions. Even when multimodal assumptions are used, they usually do not take into account the actual most likely shape of the distributions, as the nonparametric methods can do. Further, they give only single point summary parameter estimates such as the mean, median, or mode, and the variances and covariances of each of the overall parameter distributions, and not the full distributions themselves. In contrast, the nonparametric methods clearly, and without

any other information, can detect any number of totally unanticipated subpopulations of patients, such as those with unusually large clearances of caffeine, for example, who may have altered CYP1A2 activity [32].

In addition, when one uses a parametric population model as the Bayesian prior to develop an initial drug dosage regimen for a patient to achieve a desired target goal at a desired target time, the MAP Bayesian regimen is based only on the point estimates of the central tendency (usually the mean or median) of each parameter distribution, and is then simply assumed to hit the desired target exactly. With such approaches, it is not possible to estimate in advance the degree to which the regimen may fail to hit the target, as there is only one single model, with each parameter consisting of only a single point summary estimate. The full shape of the parameter distribution, with its recognized or unrecognized subpopulations and outliers, is not considered. Such action (the dosage regimen) therefore cannot be specifically designed to achieve target goal(s) optimally, with maximal precision. When the distribution is a mixture of Gaussians, there is also a difficult problem ensuring that the regimen chosen is truly the most precise in hitting the desired target goals.

In summary, PEM, NPEM, and NPAG, which have exact or accurate computations, have consistent behavior, good statistical efficiency and precise parameter estimation, and good asymptotic convergence. In contrast, the IT2B program, which uses the FOCE approximation, suffered a loss of consistency, with a small (1 – 2%) bias for the mean parameter values, a moderate (20 – 30%) bias for the SD's, and severe bias for the correlation coefficients. The NONMEM program also, with the FOCE approximation, suffered in precision of its parameter estimates. NONMEM FO suffered severely. The NPAG program was statistically quite efficient,

much more so than the FOCE NONMEM or IT2B, and had much better asymptotic convergence, close to theoretical [2].

A weakness of nonparametric approaches is that there is currently no method to obtain confidence limits for these oddly shaped nonparametric parameter distributions. To get confidence intervals for the entire nonparametric parameter distributions, one must use much more computationally intensive methods such as profile likelihood or bootstrap methods to obtain them. Work to implement these approaches is in progress.

6. CONCLUSIONS

Population modeling approaches which compute the likelihood function and/or conditional likelihoods exactly or accurately have the desirable properties of statistical consistency, good statistical efficiency (precise parameter estimates), and good asymptotic convergence [2,21]. This is true of the nonparametric NPEM and NPAG programs, and also of the PEM parametric approach. Most currently available parametric methods such as IT2B and NONMEM, however, do not have these desirable proven properties [2-4], as the computations are done only as an FO or FOCE approximation.

A significant clinical benefit of the NP population modeling approaches is that the multiple support points, with their multiple sets of parameter values, provide multiple predictions of future serum concentrations and other responses from any given future dosage regimen, compared to only one from a parametric model. Because of this, the NP models provide a tool to circumvent the problems presented by the separation principle [24] and can not only calculate and but also can specifically optimize in advance the precision with which any dosage regimen is predicted to hit a desired target goal at a desired time. Nonparametric population models thus

permit “multiple model” design of dosage regimens to optimize a specific performance criterion [25-29], such as minimizing a weighted least-squares criterion in achieving a desired clinically selected target goal, thus performing an optimized clinical simulated trial with each dosage regimen developed. It is useful to consider all the above information when selecting software for population PK/PD modeling.

The following approach currently appears to be useful. It seems good, first, to carefully determine the assay error polynomial. Then, a parametric method such as IT2B can be useful to help to estimate and explore the probable ranges of the population parameter values. Gamma can be computed either here of with the NPAG program. Then the NPAG program can be used to get the final definitive entire parameter distributions. These can then be used with multiple model dosage design [26-29] for optimal clinical planning, monitoring, and adjustment of a patient’s drug regimen.

REFERENCES

1. Jelliffe R, Schumitzky A, Van Guilder M, and Jiang F: User Manual for Version 10.7 of the USC*PACK Collection of PC Programs, USC Laboratory of Applied Pharmacokinetics, USC School of Medicine, Los Angeles, CA, December 1, 1995. This software is available from the corresponding author by license from the University of Southern California for a nominal donation

2. Leary R, Jelliffe R, Schumitzky A, and Van Guilder M: A Unified Parametric / Nonparametric Approach to Population PK/PD Modeling. Presented at the Annual Meeting of the Population Approach Group in Europe, Paris, France, June 6-7, 2002.

3. Spieler G and Schumitzky A: Asymptotic Properties of Extended Least Squares Estimates with Application to Population Pharmacokinetics. Proceedings of the American Statistical Society, Biopharmaceutical Section, 1993, pp. 177-182.

4. Vonesh E, Chinchilli V. Linear and nonlinear models for analysis of repeated measurements. New York: Marcel Dekker, 1997.

5. Variability in Drug Therapy: Description, Estimation, and Control. Ed by Rowland M, Sheiner L, and Steimer JL. Raven Press, New York, 1985.

6. Davidian M, Giltinan D. Nonlinear models for repeated measurement data. New York: Chapman and Hall, 1995.

7. Beal S, and Sheiner L: NONMEM User's Guide I. Users Basic Guide. Division of Clinical Pharmacology, University of California, San Francisco, 1979.

8. Sheiner L: The population Approach to Pharmacokinetic Data Analysis: Rationale and Standard Data Analysis Methods. Drug Metab. Rev. 1984; 15: 153-171.

9. Beal S: Population Pharmacokinetic Data and Parameter Estimation Based on their First Two Statistical Moments. Drug Metab. Rev. 1984; 15: 173-193.

10. Schumitzky A: The Nonparametric Maximum Likelihood Approach to Pharmacokinetic Population Analysis. Proceedings of the 1993 Western Simulation Multiconference - Simulation for Health Care. Society for Computer Simulation, 1993, pp 95-100. (Also available as Technical Report 92-3, Laboratory of Applied Pharmacokinetics, USC School of Medicine, 1992).

11. Schumitzky A: Nonparametric EM Algorithms for Estimating Prior Distributions. App. Math. and Computation. 1991; 45: 143-157.

12. Sheiner L, Beal S, Rosenberg B, and Marathe V: Forecasting Individual Pharmacokinetics. Clin. Pharmacol. Therap. 1979; 26: 294-305.

13. Jelliffe R: Explicit Determination of laboratory assay error patterns: a useful aid in therapeutic drug monitoring. No. DM 89-4 (DM56). Drug. Monit. Toxicol. 1989; 10: (4) pp.1-6.

14. Jelliffe R, Schumitzky A, Van Guilder M, Liu M, Hu L, Maire P, Gomis P, Barbaut X, and Tahani B: Individualizing Drug Dosage Regimens: Roles of Population

Pharmacokinetic and Dynamic Models, Bayesian Fitting, and Adaptive Control. Therap. Drug Monit. 1993; 15: 380-393.

15. De Groot M: Probability and Statistics, 2nd edition, 1986, reprinted 1989, Addison-Wesley, Reading MA, pp. 420-423.

16. Mallet A: A Maximum Likelihood Estimation Method for Random Coefficient Regression Models. Biometrika. 1986; 73: 645-656.

17. Schumitzky A: EM Algorithms and Two Stage Methods in Pharmacokinetic Population Analysis. In: Advanced Methods of Pharmacokinetic and Pharmacodynamic Systems Analysis II. D. Z.D'Argenio, ed., Plenum Press, New York, 1995, pp. 145-160.

18. Bertilsson L: Geographic/Interracial Differences in Polymorphic Drug Oxidation. Clin. Pharmacokinet. 1995; 29: 192-209.

19. Lindsay B: The Geometry of Mixture Likelihoods: A General Theory. Ann. Statist. 1983; 11: 86-94.

20. Maire P, Barbaut X, Girard P, Mallet A, Jelliffe R, and Berod T: Preliminary results of three methods for population pharmacokinetic analysis (NONMEM, NPML, NPEM) of

amikacin in geriatric and general medicine patients. Int. J. Biomed. Comput., 1994; 36: 139-141.

21. Leary R, Jelliffe R, Schumitzky A, and Van Guilder M: Improved Computational Methods for Statistically Consistent and Efficient PK/PD Population Analysis. Presented at the Annual Meeting of the Population Approach Group in Europe, Verona, Italy, June 12-13, 2003.

22. Bustad A, Jelliffe R, and Terziivanov D: A comparison of Parametric and Nonparametric Methods of Population Pharmacokinetic Modeling. Presented at the Annual Meeting of the American Society for Clinical Pharmacology and Therapeutics, Atlanta, GA, March 26, 2002.

23. Jelliffe R: Estimation of Creatinine Clearance in Patients with Unstable Renal Function, without a Urine Specimen. Am. J. Nephrol. 2002; 22: 320-324.

24. Bertsekas D: Dynamic Programming: deterministic and stochastic models. Englewood Cliffs (NJ): Prentice-Hall, 1987; pp.144-146.

25. Tarıght N, Mentre F, Mallet A, and Jouvent R: Nonparametric Estimation of Population Characteristics of the Kinetics of Lithium from Observational and Experimental Data:

Individualization of Chronic Dosing Regimen Using a New Bayesian Approach. Therap. Drug Monit. 1994; 16: 258-269.

26. Bayard D, Jelliffe R, Schumitzky A, Milman M, and Van Guilder M: Precision Drug Dosage Regimens using Multiple Model Adaptive Control: Theory and Application to Simulated Vancomycin Therapy. in Selected Topics in Mathematical Physics, Professor R. Vasudevan Memorial Volume, ed. by Sridhar R, Srinavasa Rao K, and Vasudevan Lakshminarayanan, Allied Publishers Inc., Madras, 1995, pp. 407-426.

27. Jerling M: Population Kinetics of Antidepressant and Neuroleptic Drugs. Studies of Therapeutic Drug Monitoring data to Evaluate Kinetic Variability, Drug Interactions, Nonlinear Kinetics, and the Influence of Genetic Factors. Ph. D. Thesis, Division of Clinical Pharmacology, Department of Medical Laboratory Sciences and Technology, Karolinska Institute at Huddinge University Hospital, Stockholm, Sweden, 1995; pp 28-29.

28. Jelliffe R, Schumitzky A, Bayard D, Milman M, Van Guilder M, Wang X, Jiang F, Barbaut X, and Maire P: Model-Based, Goal-Oriented, Individualised Drug Therapy: Linkage of Population Modelling, New “Multiple Model” Dosage Design, Bayesian Feedback and Individualised Target Goals. Clin. Pharmacokinet. 1998; 34: 57-77.

29. Bayard D, Milman M, and Schumitzky A: Design of Dosage Regimens: A Multiple Model Stochastic Approach. Int. J. Biomed. Comput. 1994; 36: 103-115.

30. Beal L, and Sheiner L. (eds.): NONMEM version V.1.1 Users Guides, University of California, San Francisco, CA, 1999.

31. Burke J: personal communication to R Leary.

32. Terziivanov D, Bozhinova K, Dimitrova V, and Atanasova I: Nonparametric Expectation Maximisation (NPEM) Population Analysis of Caffeine Disposition from Sparse Data in Adult Caucasians: Systemic Caffeine Clearance as a Biomarker for Cytochrome P4501A2 Activity. Clin. Pharmacokinet. 42: 1393-1409, 2003.

33. Bennet JE, Racine-Poon A, Wakefield JC. Markov chain Monte Carlo for nonlinear hierarchical models. In: *Markov Chain Monte Carlo in Practice*. Gilks WR, Richardson S, Spiegelhalter DJ, editors. London: Chapman and Hall: 339-357. 1996.

34. Girard P and Mentre F: A comparison of estimation methods in nonlinear mixed effects models using a blind analysis. Population Approach Group Europe annual meeting, Pamplona, Spain, 16-17 June 2005, http://www.page-meeting.org/

TABLES

		METHOD		
		IT2B	NPEM	NPAG
Mean				
	KA	1.349	1.408	1.380
	VS1	0.258	0.259	0.258
	KS1	0.003358	0.003371	0.003375
Median/CV%				
	KA	1.352/4.55	1.363/20.42	1.333/21.24
	VS1	0.2591/13.86	0.249/17.44	0.254/17.38
	KS1	0.003273/14.83	0.003371/15.53	0.003283/15.76
Log – Likelihood		-389.548	-374.790	-374.326

Table I. Clinical Study Results. Parameter values (mean, median, percent coefficient of variation, CV%), and the exact final log likelihood obtained with the IT2B, the NPEM, and the NPAG programs. KA = absorptive rate constant (hr^{-1}), VS1, apparent central volume of distribution (L/kg), KS1, increment of elimination rate constant (hr^{-1} per unit of creatinine clearance). CV% is less with IT2B, but so was the log likelihood, which was better (less negative here) with NPEM and NPAG.

	IT2B	NPEM	NPAG
R^2 =	0.814	0.879	0.880
ME =	-0.575	-0.751	0.169
MSE =	48.69	29.01	29.70

Table II. Clinical Study Results. Analysis of estimated versus measured serum concentrations based on population median parameter values, using the IT2B, NPEM, and NPAG programs. R^2: square of the correlation coefficient, ME: mean error, MSE: mean squared error.

:

Estimator	Relative efficiency	Relative error
DIRECT OBSERVATION	**100.0 %**	**1.00**
PEM	**75.4%**	**1.33**
NPAG	**61.4%**	**1.63**
NONMEM FOCE	**29.0%**	**3.45**
IT2B FOCE	**25.3%**	**3.95**
NONMEM FO	**0.9%**	**111.11**

Table III. Results of First Monte Carlo Simulation. Comparison of the relative statistical efficiency and relative error in parameter estimation for the mean value of V, the apparent volume of distribution, of the PEM, NPAG, NONMEM FOCE, IT2B FOCE, and NONMEM FO population Modeling Methods.

FIGURES

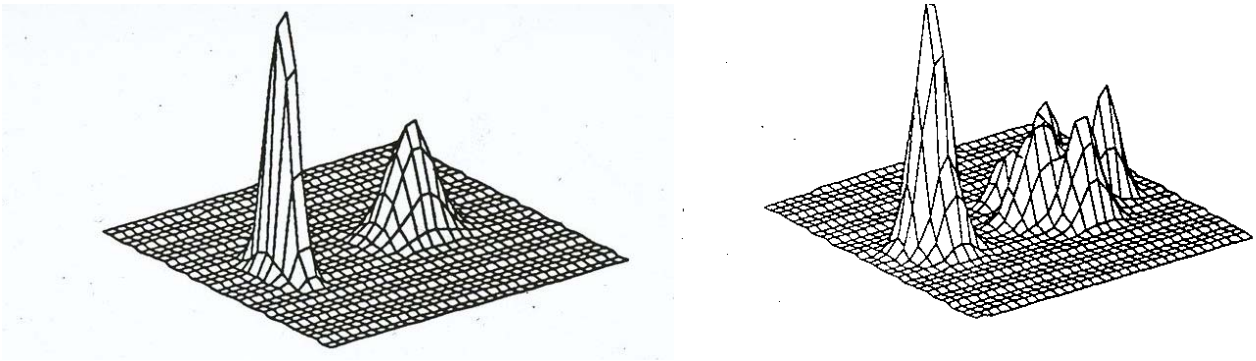

Figure 1.

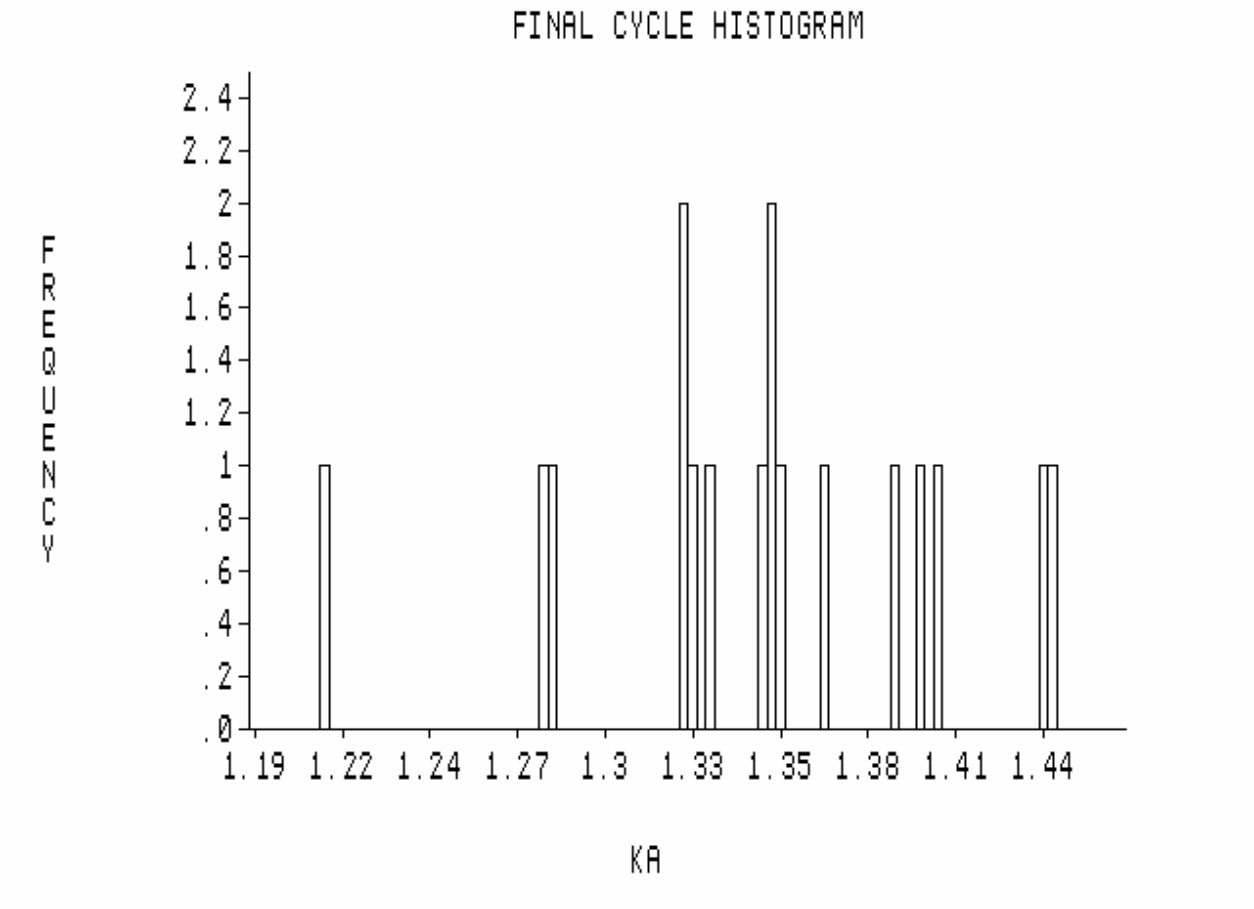
FINAL CYCLE HISTOGRAM
FREQUENCY
2.4
2.2
2
1.8
1.6
1.4
1.2
1
.8
.6
.4
.2
.0
1.19
1.22
1.24
1.27
1.3
1.33
1.35
1.38
1.41
1.44
KA

Figure 2.

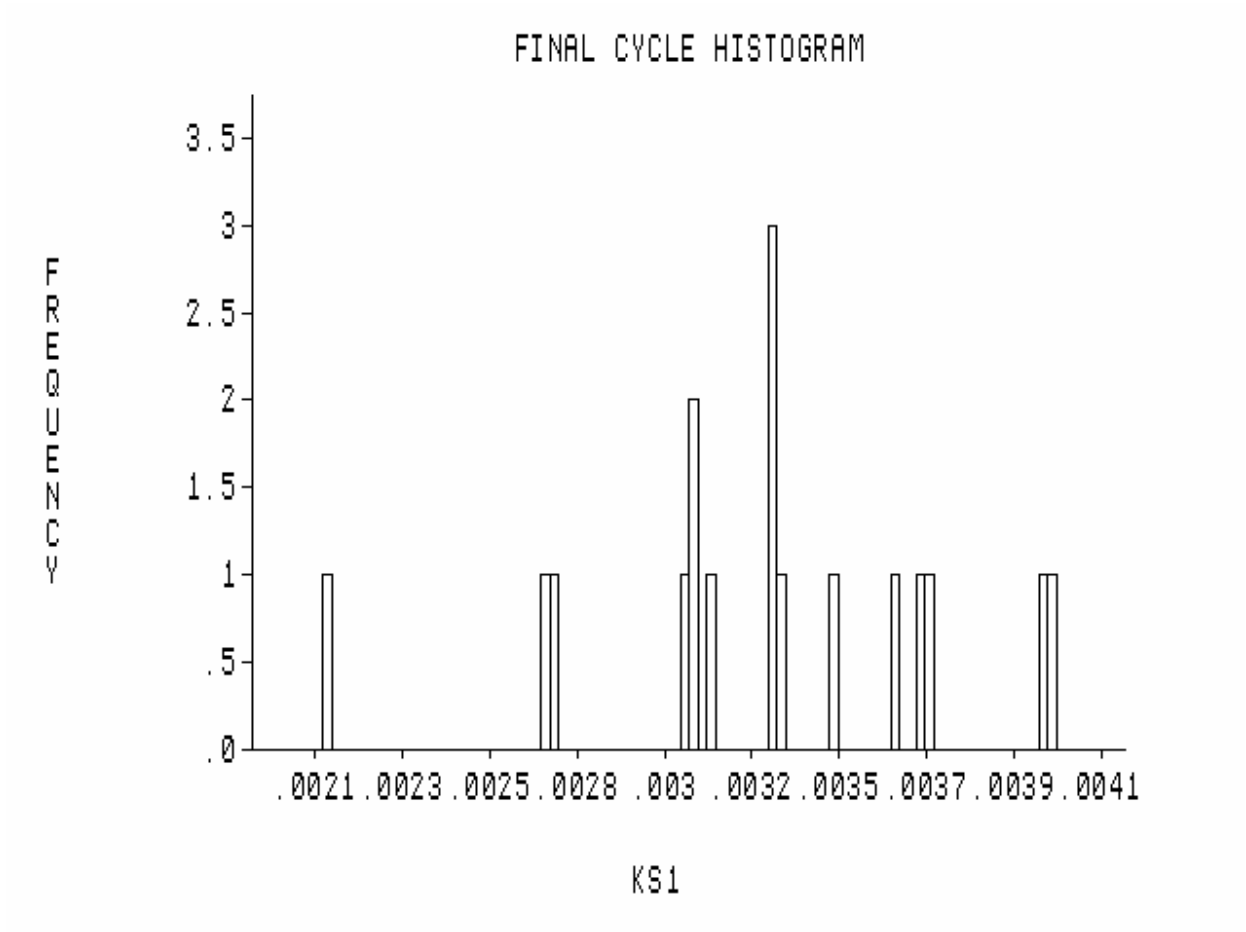
FINAL CYCLE HISTOGRAM
FREQUENCY
3.5
3
2.5
2
1.5
1
.5
.0
.0021 .0023 .0025 .0028 .003 .0032 .0035 .0037 .0039 .0041
KS1

Figure 3.

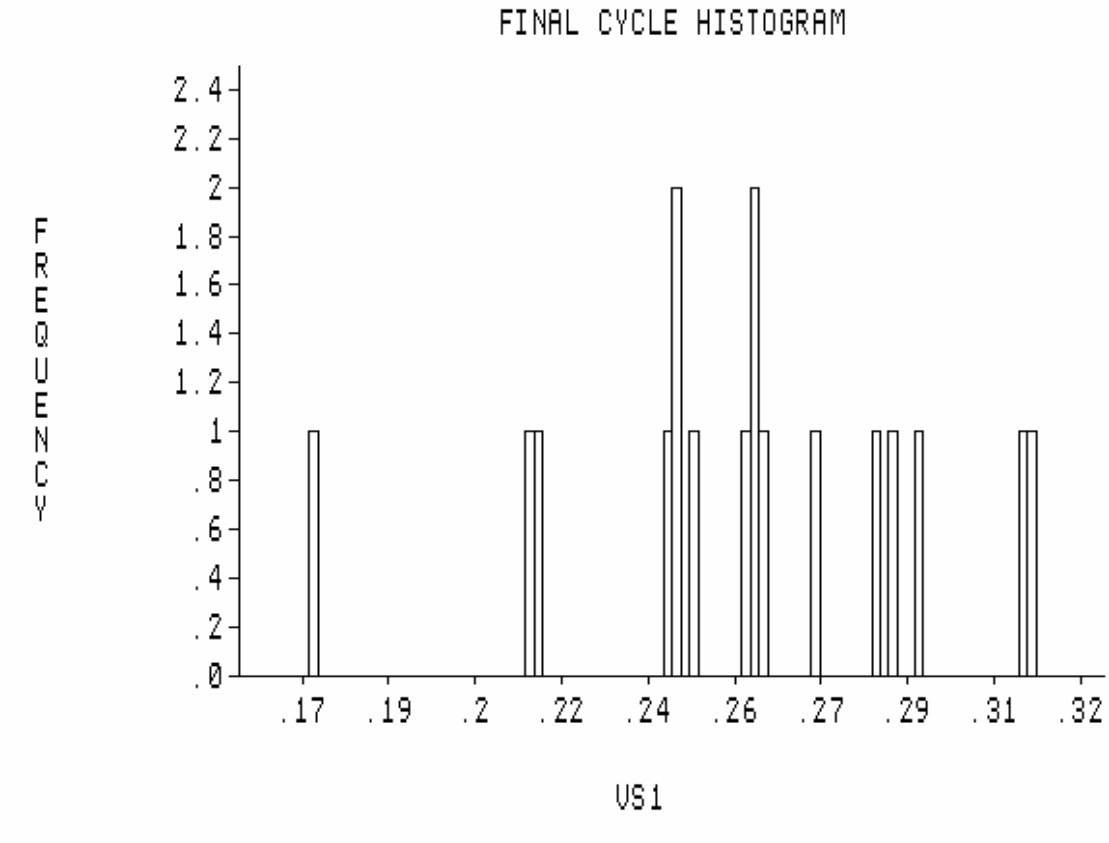
FINAL CYCLE HISTOGRAM
FREQUENCY
2.4
2.2
2
1.8
1.6
1.4
1.2
1
.8
.6
.4
.2
.0
.17
.19
.2
.22
.24
.26
.27
.29
.31
.32
VS1

Figure 4.

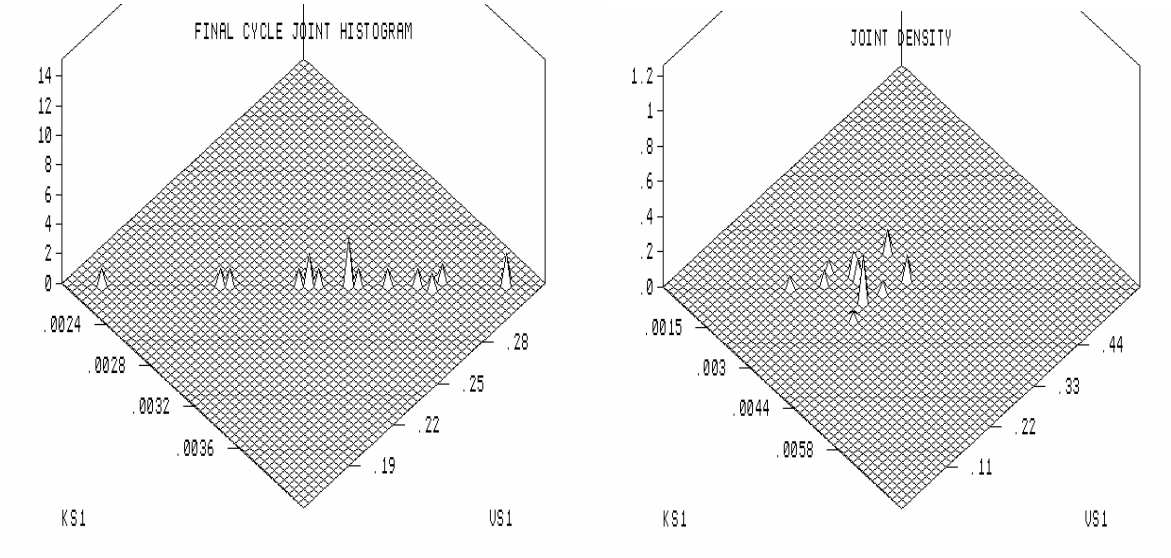
FINAL CYCLE JOINT HISTOGRAM
14
12
10
8
6
4
2
0
.0024
.0028
.0032
.0036
KS1
.19
.22
.25
.28
VS1
JOINT DENSITY
1.2
1
.8
.6
.4
.2
.0
.0015
.003
.0044
.0058
KS1
.11
.22
.33
.44
VS1

Figure 5.

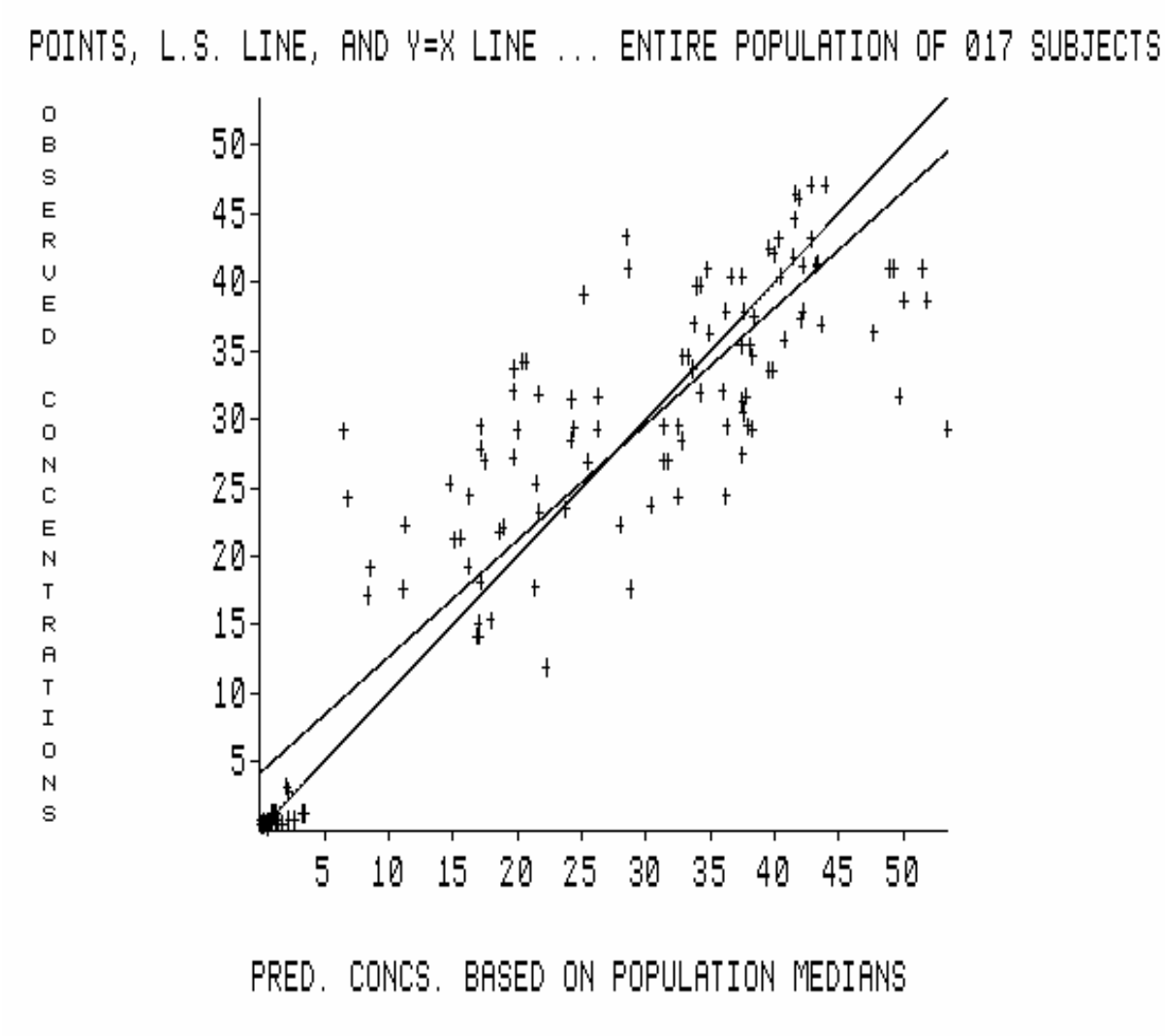
POINTS, L.S. LINE, AND Y=X LINE ... ENTIRE POPULATION OF 017 SUBJECTS
OBSERVED CONCENTRATIONS
50
45
40
35
30
25
20
15
10
5
5 10 15 20 25 30 35 40 45 50
PRED. CONCS. BASED ON POPULATION MEDIANS

Figure 6

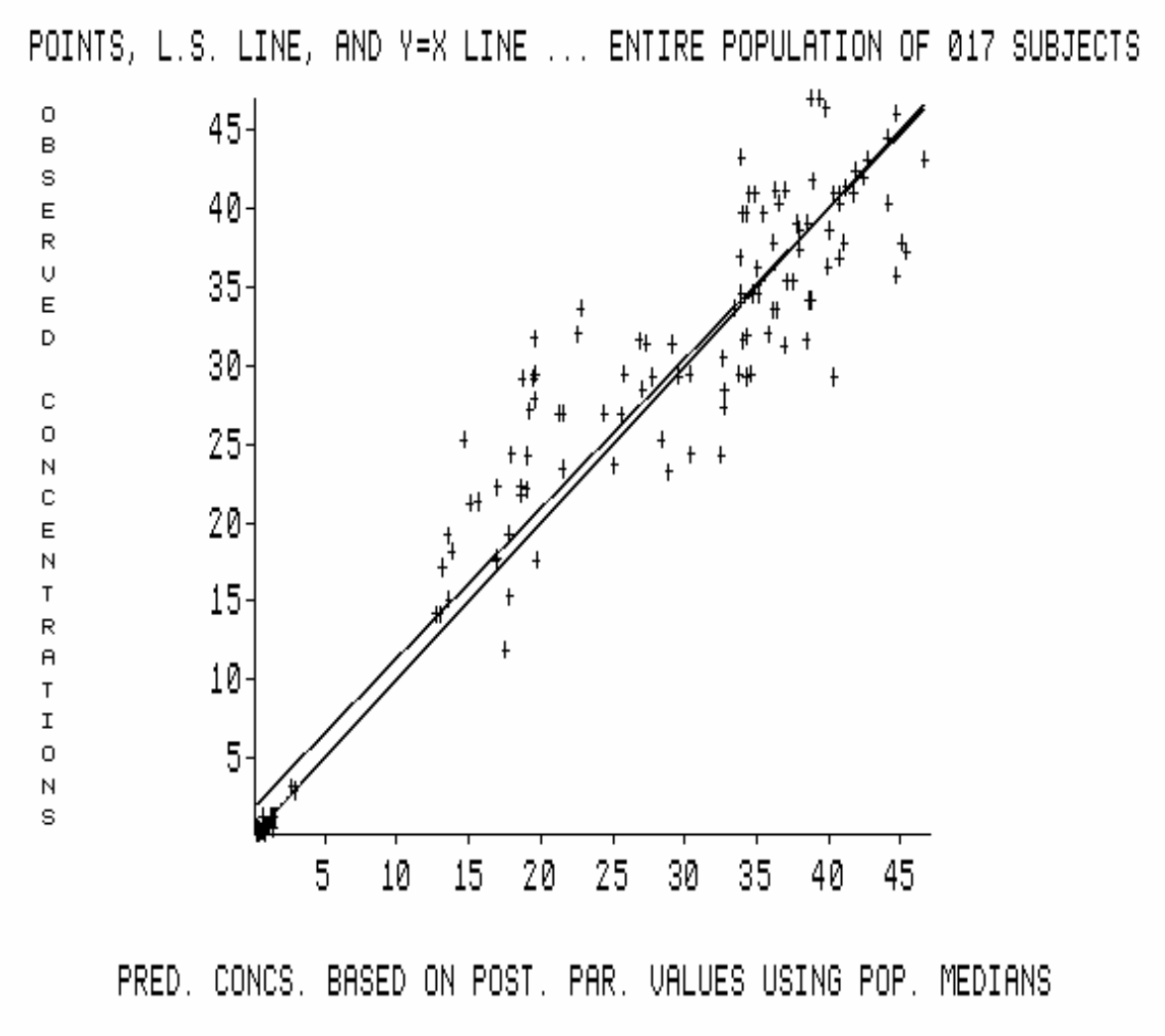
POINTS, L.S. LINE, AND Y=X LINE ... ENTIRE POPULATION OF 017 SUBJECTS
OBSERVED CONCENTRATIONS
45
40
35
30
25
20
15
10
5
5 10 15 20 25 30 35 40 45
PRED. CONCS. BASED ON POST. PAR. VALUES USING POP. MEDIANS

Figure 7

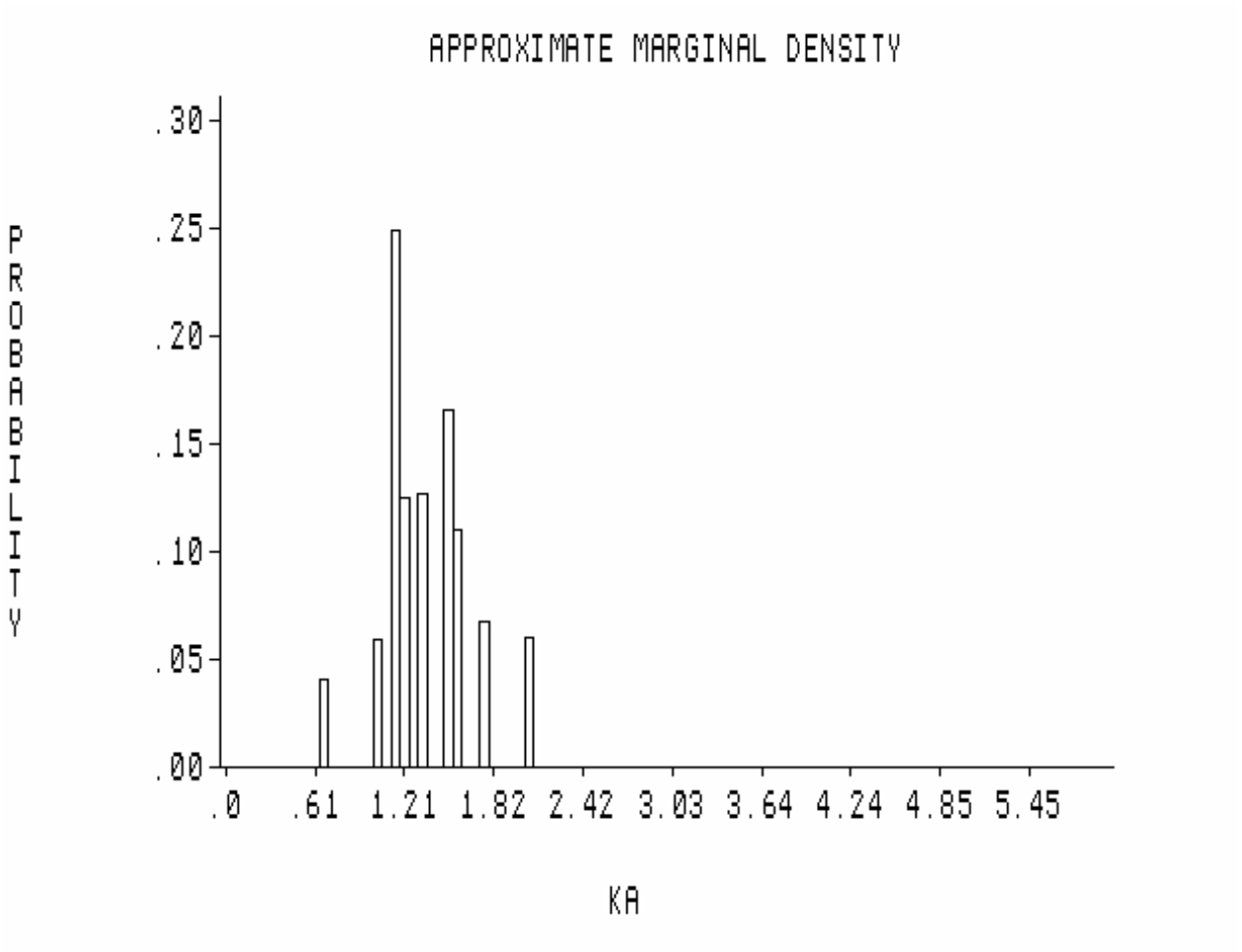
APPROXIMATE MARGINAL DENSITY
PROBABILITY
.30
.25
.20
.15
.10
.05
.00
.0
.61
1.21
1.82
2.42
3.03
3.64
4.24
4.85
5.45
KA

Figure 8.

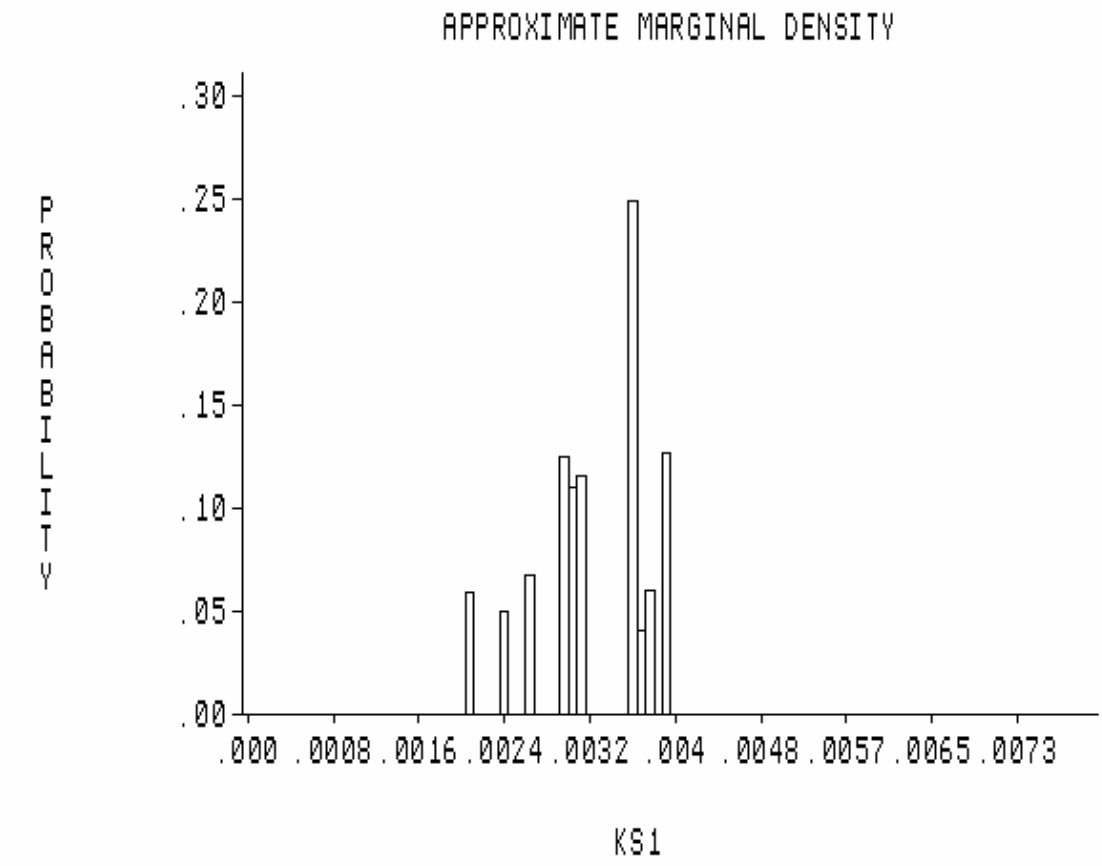
APPROXIMATE MARGINAL DENSITY
PROBABILITY
.30
.25
.20
.15
.10
.05
.00
.000 .0008 .0016 .0024 .0032 .004 .0048 .0057 .0065 .0073
KS1

Figure 9.

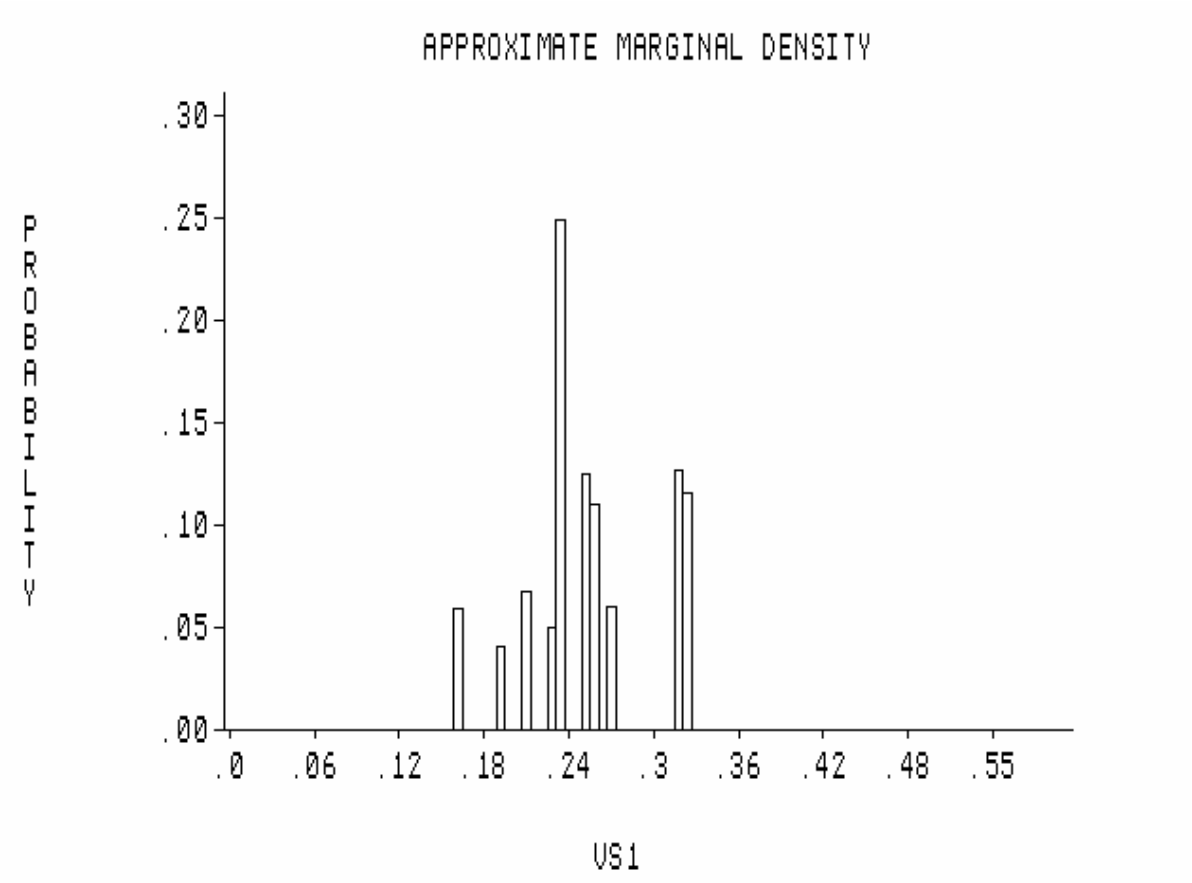
APPROXIMATE MARGINAL DENSITY
PROBABILITY
.30
.25
.20
.15
.10
.05
.00
.0
.06
.12
.18
.24
.3
.36
.42
.48
.55
VS1

Figure 10.

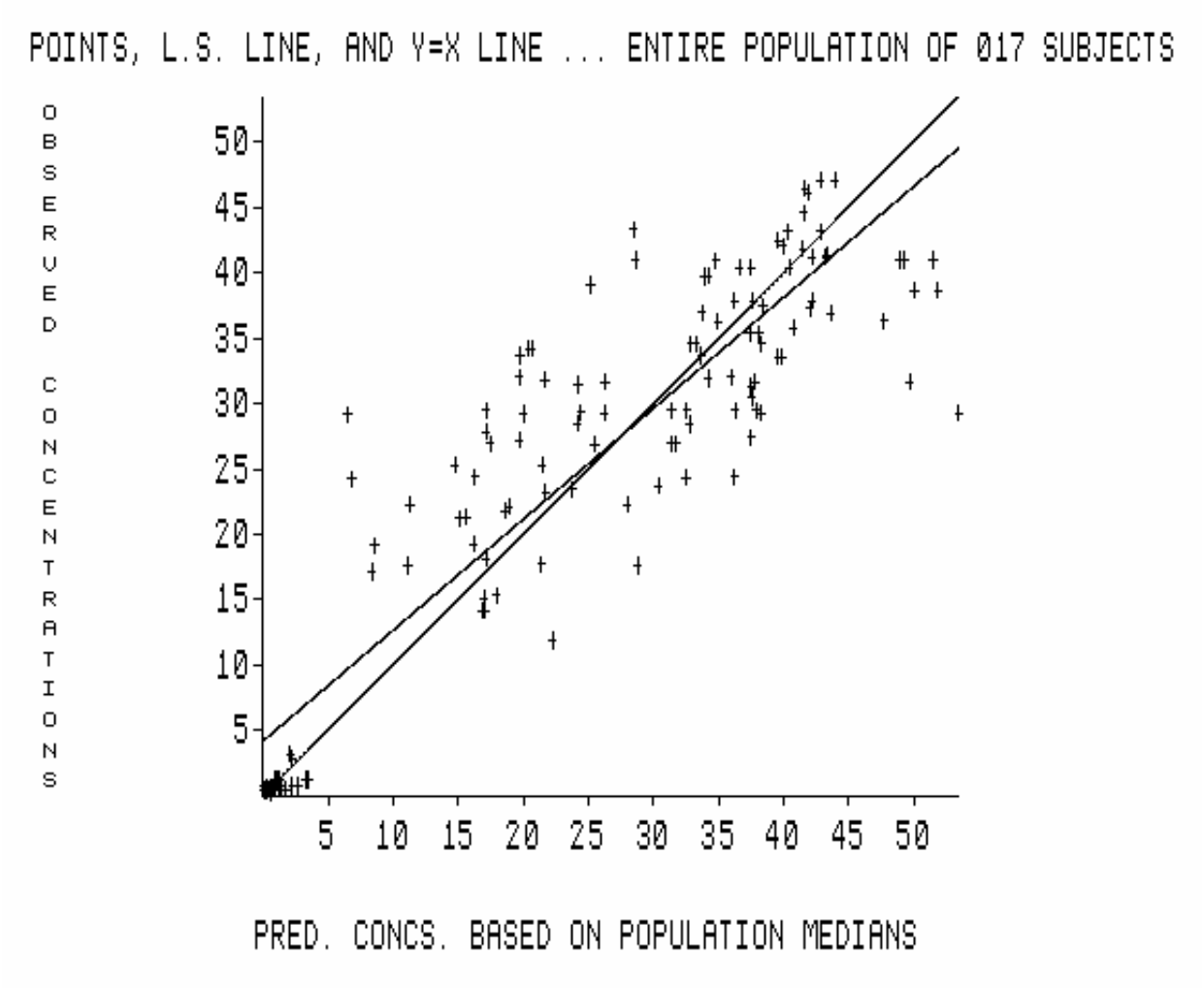
POINTS, L.S. LINE, AND Y=X LINE ... ENTIRE POPULATION OF 017 SUBJECTS
OBSERVED CONCENTRATIONS
50
45
40
35
30
25
20
15
10
5
5 10 15 20 25 30 35 40 45 50
PRED. CONCS. BASED ON POPULATION MEDIANS

Figure 11.

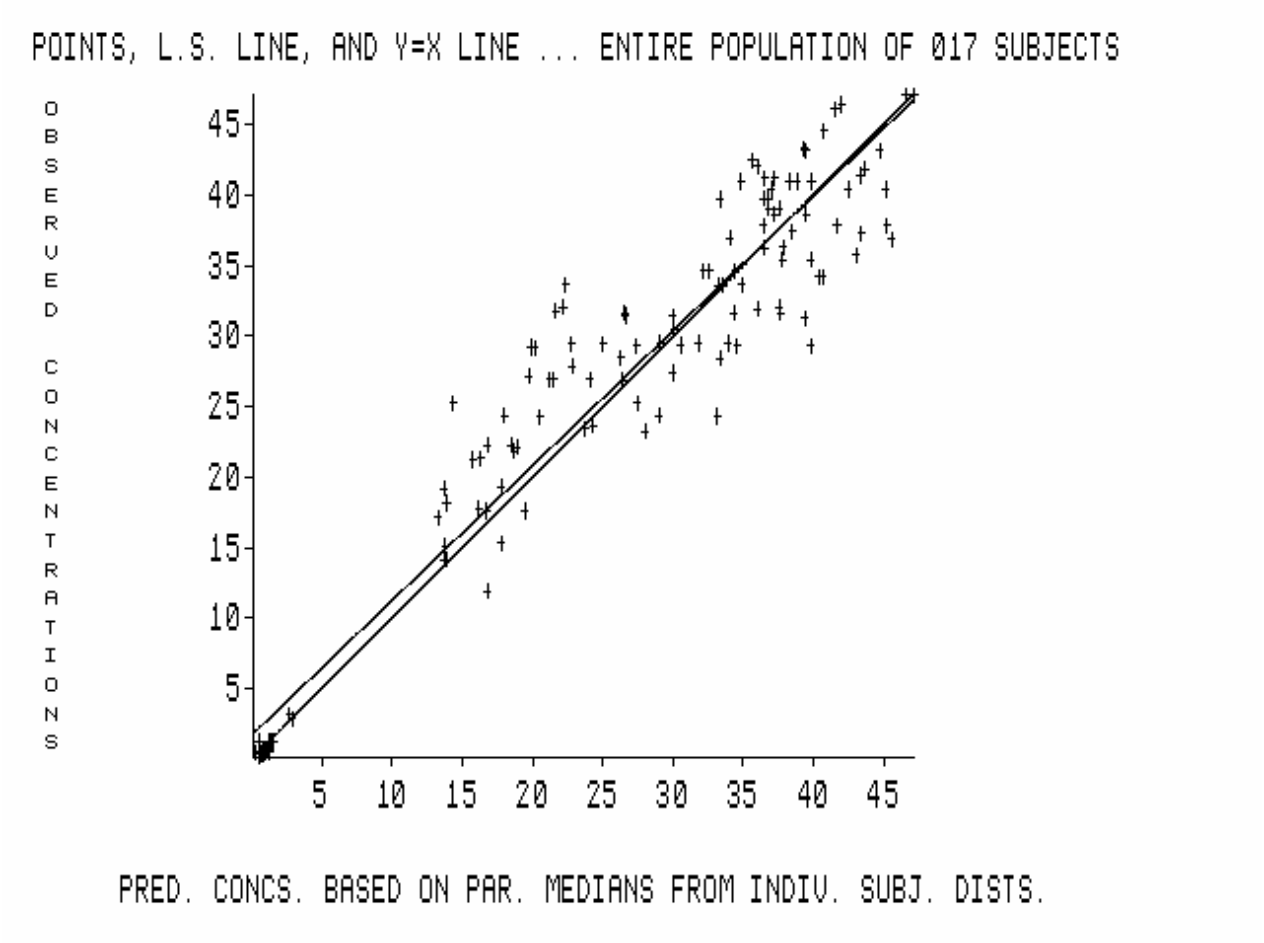
POINTS, L.S. LINE, AND Y=X LINE ... ENTIRE POPULATION OF 017 SUBJECTS
OBSERVED CONCENTRATIONS
45
40
35
30
25
20
15
10
5
5 10 15 20 25 30 35 40 45
PRED. CONCS. BASED ON PAR. MEDIANS FROM INDIV. SUBJ. DISTS.

Figure 12.

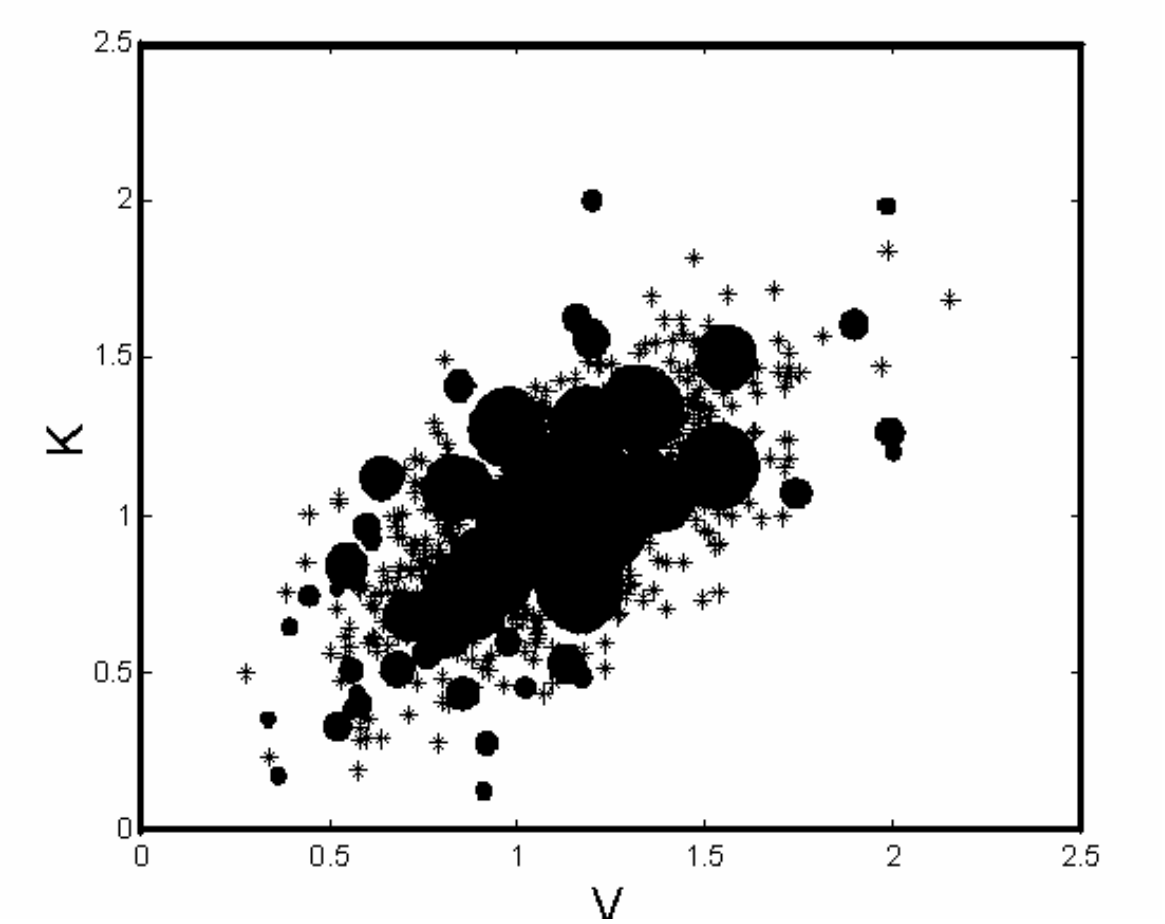

Figure 13

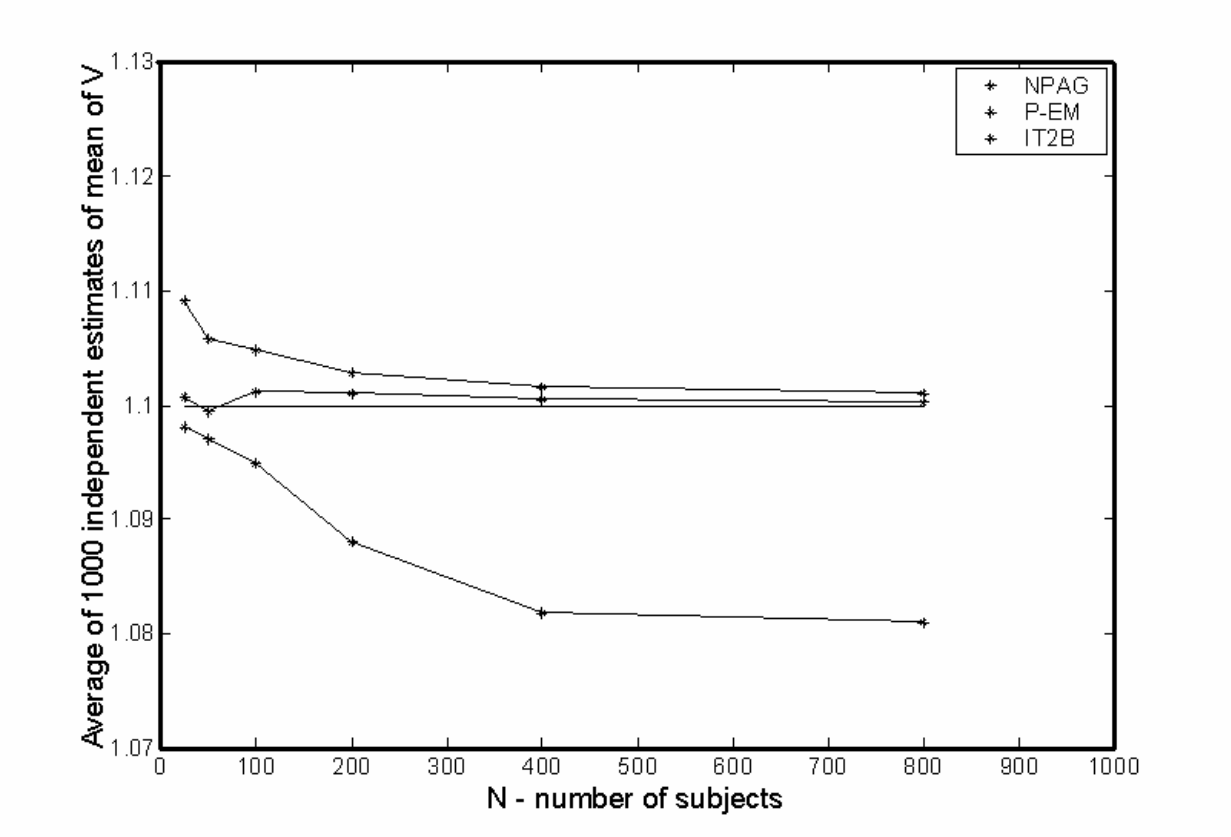
NPAG
P-EM
IT2B
Average of 1000 independent estimates of mean of V
N - number of subjects
1.13
1.12
1.11
1.1
1.09
1.08
1.07
0
100
200
300
400
500
600
700
800
900
1000

Figure 14.

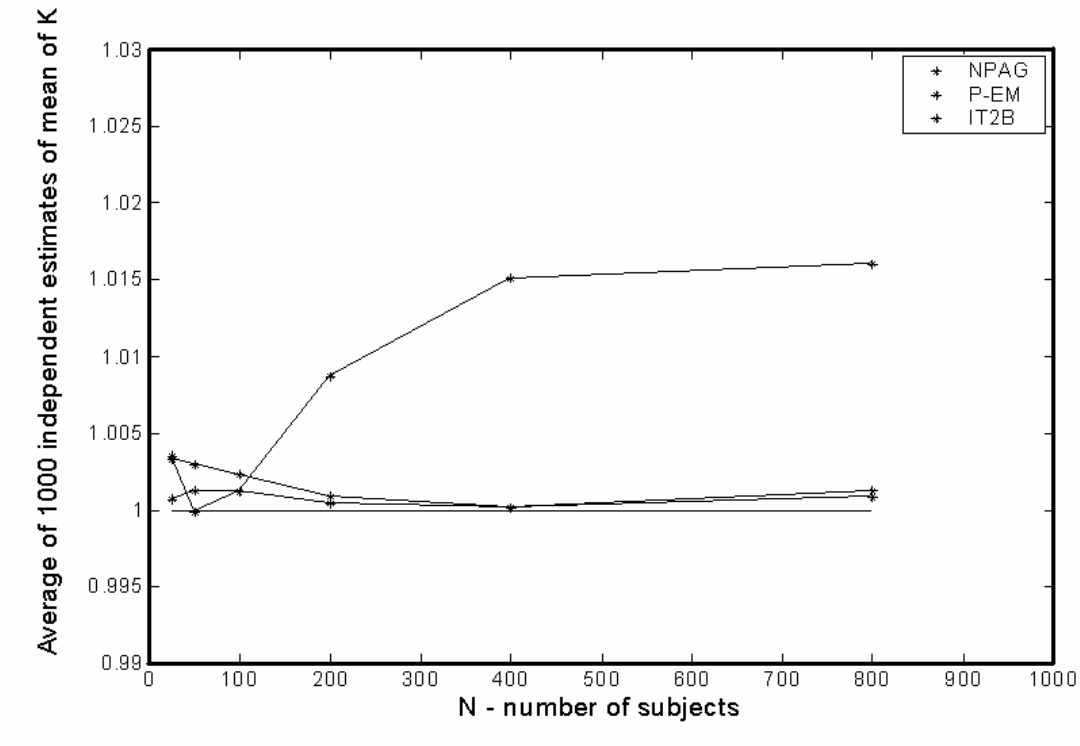
Average of 1000 independent estimates of mean of K
N - number of subjects
NPAG
P-EM
IT2B
1.03
1.025
1.02
1.015
1.01
1.005
1
0.995
0.99
0
100
200
300
400
500
600
700
800
900
1000

Figure 15

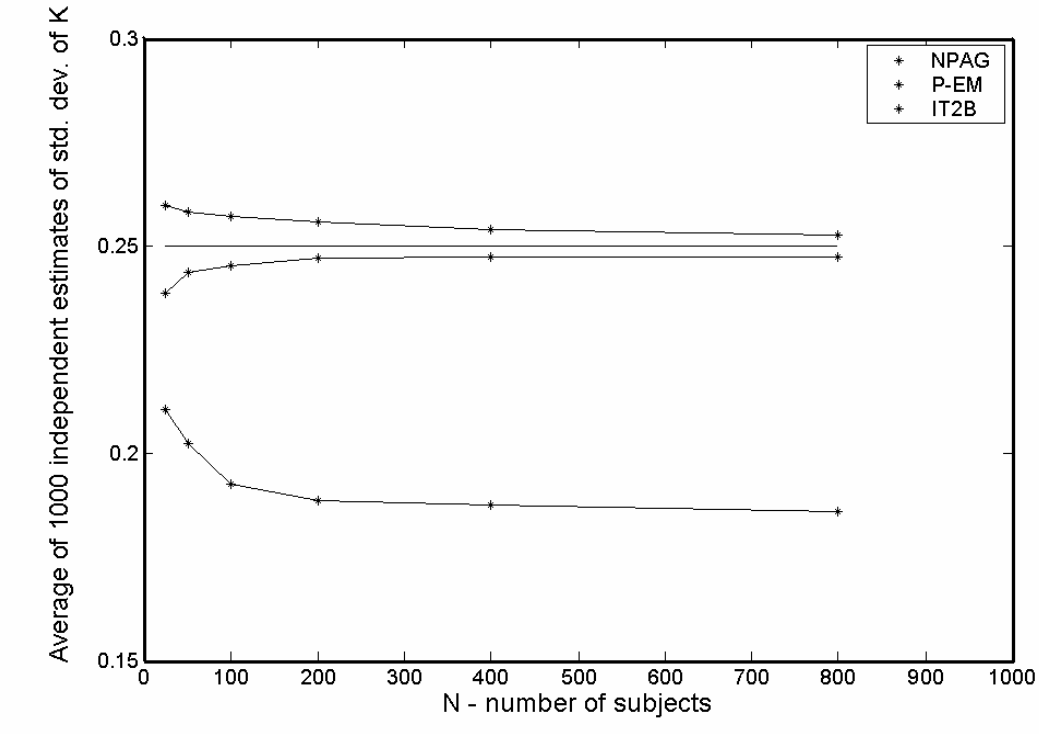
Average of 1000 independent estimates of std. dev. of K
0.3
0.25
0.2
0.15
0
100
200
300
400
500
600
700
800
900
1000
N - number of subjects
NPAG
P-EM
IT2B

Figure 16

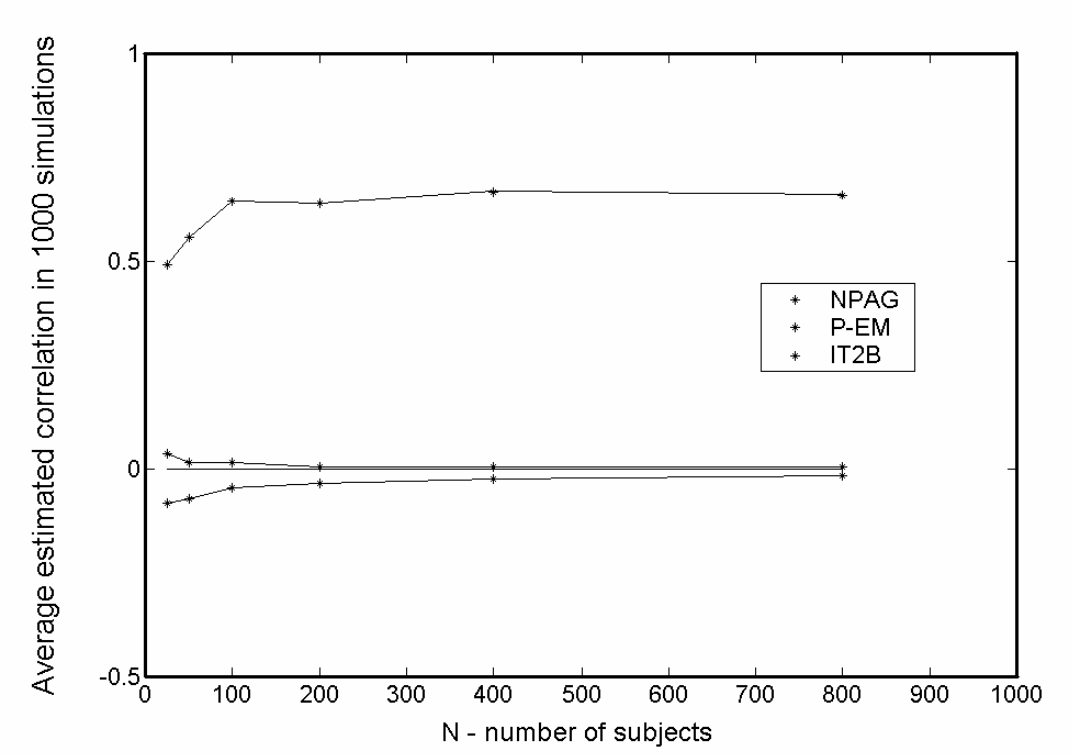
Average estimated correlation in 1000 simulations
1
0.5
0
-0.5
0
100
200
300
400
500
600
700
800
900
1000
N - number of subjects
NPAG
P-EM
IT2B

Figure 17

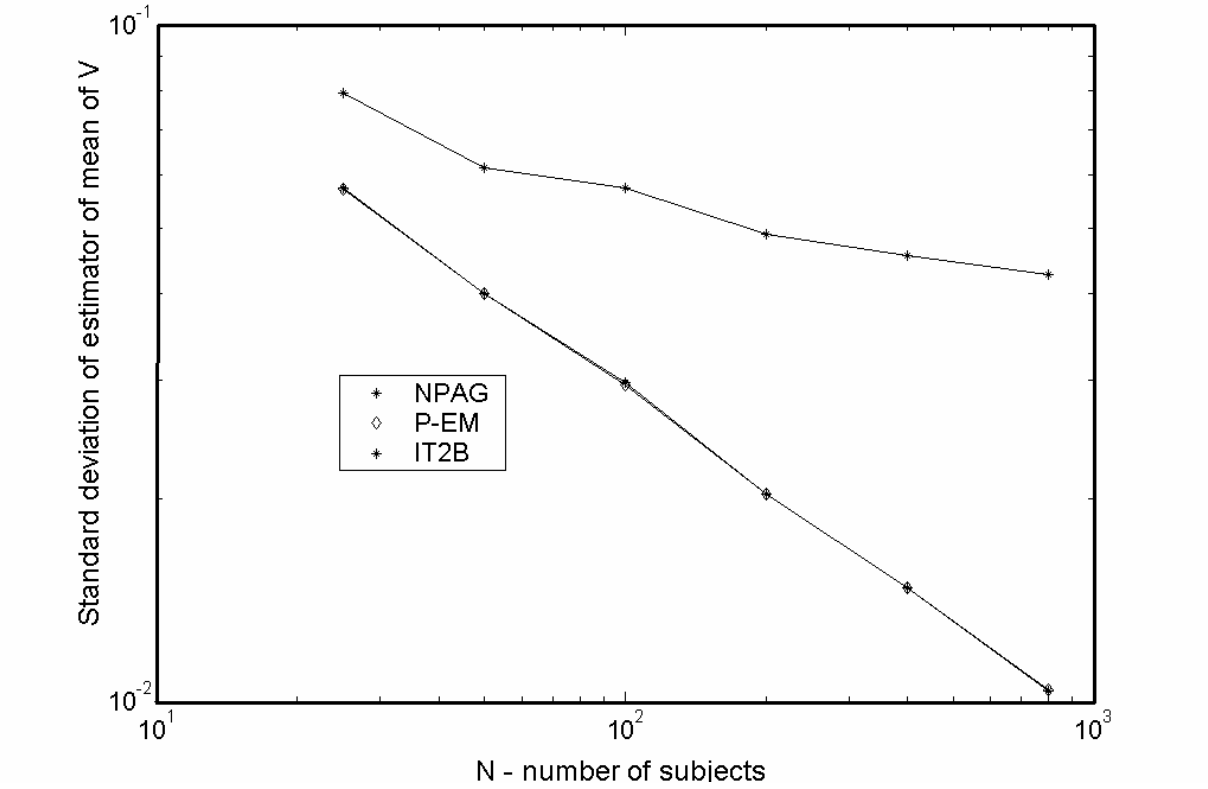
Standard deviation of estimator of mean of V
N - number of subjects
NPAG
P-EM
IT2B
10^-1
10^-2
10^1
10^2
10^3

Figure 18

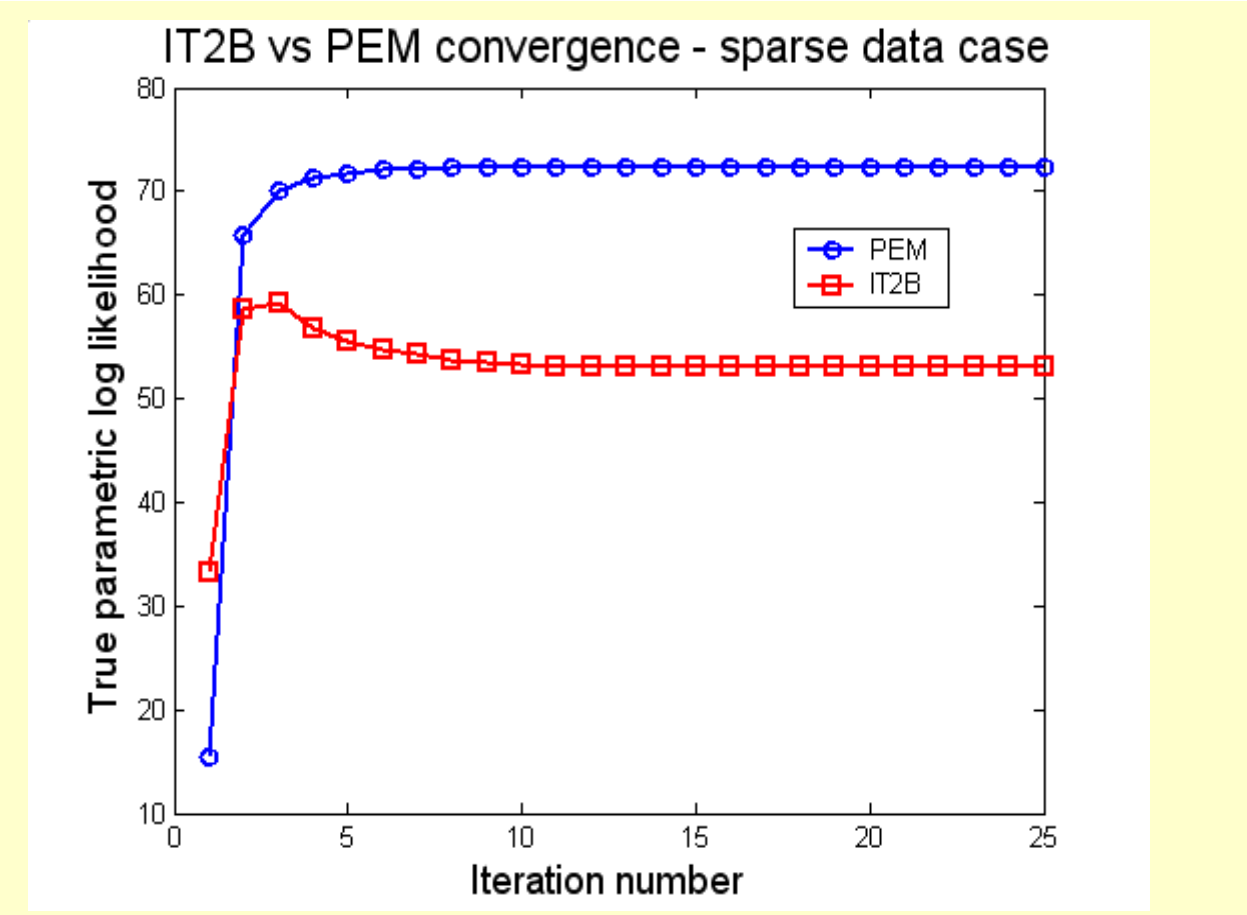
IT2B vs PEM convergence - sparse data case
True parametric log likelihood
Iteration number
PEM
IT2B
80
70
60
50
40
30
20
10
0
5
10
15
20
25

Figure 19

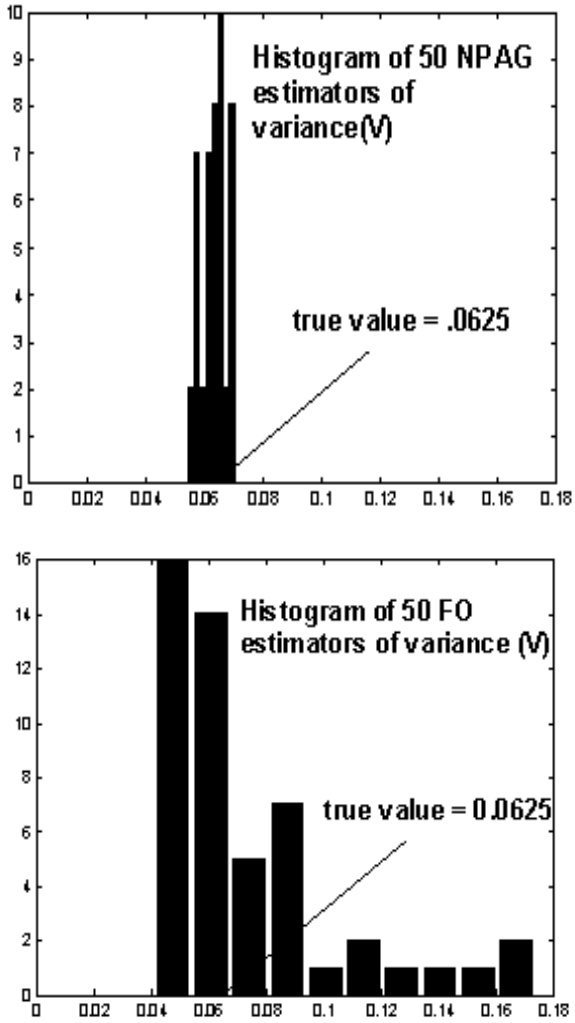
Histogram of 50 NPAG estimators of variance(V)
true value = .0625
Histogram of 50 FO estimators of variance (V)
true value = 0.0625

Figure 20

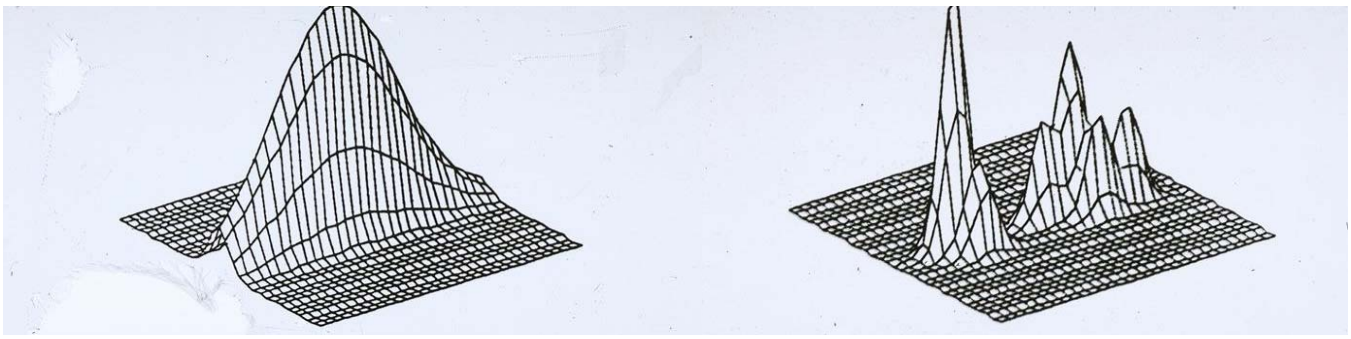

Figure 21.