A comparative study of some approaches for constructing tolerance limits

by Max Petzold

Department of Statistics, School of Economics and Commercial Law, Göteborg University, Box 660, SE-405 30 GÖTEBORG, Sweden.
Email: Max.Petzold@statistics.gu.se

Key words and phrases: conservative; dose finding study; non-parametric; order statistics; parametric; population pharmacokinetics; small sample; tolerance limits.

Abstract

In a dose finding study the aim is to come up with a safe and efficient drug administration. By comparing the obtained tolerance limits with predetermined desired concentration limits, i.e. the therapeutic window, one may be able to adjust the drug dosage. For drugs with a therapeutic window situated at high concentrations close to toxic levels, one has to balance between attaining a high proportion of the population at efficient high levels on one hand and the risks of an overdose on the other hand. In such cases it is important to use the proper estimation approach for the upper tolerance limit. Here a conservative estimation approach intended for a drug with potentially adverse side effects with minor overdoses is compared with an approach intended for a drug with harmless side effects. It is shown that the conservative approach can be considerably less efficient when used in the latter case, a disadvantage that is rarely discussed when proposing conservative estimators.

The properties of the two approaches are evaluated using well-known parametric and non-parametric estimators in a simulation study for small and moderate sample sizes.

1 Introduction

The question of how much drug should be used and how it should be administered for a given therapeutic purpose can not readily be answered. Basically two different methods have been used to answer the question in a population: empirical and population kinetic. The latter is to be discussed in this paper. The kinetic method is based on the hypothesis that therapeutic and toxic responses are related to the amount of drug in the body or to the plasma drug concentration. By studying pharmacokinetic data following a single dose, the tolerance limits of the concentrations in a population may be obtained. The tolerance limits can then in turn be used to modify the dosage regimen according to the known relationships among concentration levels, therapeutic response, and toxic effects.

The idea of tolerance limits and tolerance intervals has to be explained in order to get the right interpretation. Historically, *ordinary tolerance intervals* arose in response to engineers' concern with mass-production processes. For example, industry-wide specifications might dictate that any component which measures less than v_1 or greater than v_2, i.e. the *ordinary tolerance limits*, is to be considered as unsatisfactory. The variability in the interval $[v_1, v_2]$ is tolerated due to design considerations or economical reasons etc. The probability that the measurement X_i of a single component i will fall in the interval indicates how successful the production process is. As a check on the true distribution of the production process one may take sample data and estimate the *statistical tolerance interval*, which is to be compared with the ordinary tolerance interval $[v_1, v_2]$. The *statistical tolerance limits* provide information about where the population sampled is likely to be concentrated, and from this, the production process can be calibrated to get a high proportion of the population within the desired interval $[v_1, v_2]$. (Statistical) tolerance limits and tolerance intervals are frequently used today in both production and research. The terms tolerance limit and percentile may be used interchangeably depending on the area of research.

In this paper a principal question of how to obtain the upper tolerance limit for a particular drug is discussed. In section 2 a conservative estimation approach based on probability restrictions and an approach based on close-

ness of the estimators are introduced. In each approach one parametric and one non-parametric estimator are considered. The parametric estimators are based on the normal distribution and the non-parametric estimators are based on continuous data. The original drug concentration is assumed to have the lognormal distribution at each time point, but all results are based on the normal distributed logarithmic scale. All the estimators are well-known but arbitrary for the actual question. In section 3 some consequences of the two approaches when used in a dose finding study are described. The results of the simulation studies are presented in section 4 followed by a discussion in section 5. The simulation model and parameter settings can be found in the Appendix.

2 Two approaches for obtaining the tolerance limits

In this paper two approaches for obtaining the upper tolerance limit are compared: the *conservative approach* and the *closeness approach*. Note that the terminology for tolerance intervals has not been standardized and the proposed names may be interpreted differently in other papers, see Mee [7] for a discussion. In the conservative approach, let a single measurement be represented by a random variable X_i, and let w be the upper tolerance limit. Then one requires that

$$P_{X_i}\left(-\infty \leq X_i \leq w\right) \geq \beta \tag{1}$$

If $\mathbf{X} = (X_1, ..., X_n)$ represents a sample of n measurements, then the estimator of w can be written $\widehat{w} = \widehat{w}\left(\mathbf{X}\right)$. Over all possible samples one may require that

$$P_{\mathbf{X}}[P_{X_i}\left(-\infty \leq X_i \leq \widehat{w}\right) \geq \beta] \geq \gamma \tag{2}$$

The expression (2) defines a one-sided *statistical β-content tolerance interval* at the confidence level γ. The interpretation of (2) is that at least $100 \times \beta\%$ of the population is covered by the tolerance interval with a probability of

at least γ. Equality is obtained in the continuous case. The upper limit w of the β-content tolerance interval is interpreted as a conservative tolerance limit, i.e. the limit indicates what extreme value that covers at least $100 \times \beta\%$ of the population with confidence γ. Expression (1) and (2) may easily be exemplified by assuming X_i to have the exponential distribution with mean and variance λ^{-1}. From (1) we have $1 - \exp(-\lambda w) = \beta$ which defines $w = -\lambda^{-1} \ln(1-\beta)$. Here one may take $\widehat{w} = -\overline{X} \ln(1-\beta)$ which in (2) gives $P[1 - \exp(-\lambda \widehat{w}(\mathbf{X})) \geq \beta] = \gamma$, i.e. the probability over all samples should be γ. Important results in the area are due to Wilks [19], Wald [16] and Tukey [14]. In the simulation study some estimators were to be chosen. Since the principle difference between the two approaches is independent of the estimators, some well-known but arbitrary estimators were chosen. In the parametric case, under normality assumption, the conservative tolerance limit was estimated by $\widehat{w}(\mathbf{X}) = \widehat{\mu} + K\widehat{\sigma}$. The constant K is determined by $P_{\mathbf{X}}[T_{n-1}(\sqrt{n} z_{1-\beta}) \leq \sqrt{n} K] \geq \gamma$ where $T_f(\delta)$ is a non-central t variable with f degrees of freedom and non-centrality parameter δ, see Kotz&Johnson [5] for a detailed description. In the non-parametric case, Wilks' [18] conservative estimator based on the beta distributed coverage between succeeding order statistics from a continuous distribution was used. The tolerance limit is estimated by the upper order statistic of the β-content tolerance interval. The limiting order statistic is chosen with a step function of n, and the confidence level will attain its minimal level γ at the shifts of order statistics. This estimator has been proposed by e.g. Gillespie&Srinivasan [2] to obtain a normal range for screening purpose in clinical medicine, Walsh [17] to compare two samples and Nickens [8] for analyzing safety data in population pharmacokinetics. As the coverage and the confidence both tend to 1, increasingly larger sample sizes are required to achieve informative limits. The minimum required sample size for a one-sided tolerance interval would e.g. increase from 59 to 459 when both β and γ increase from 0.95 to 0.99. Sample size tables for one-sided and two-sided tolerance intervals are given by Somerville [12]. In all simulations in the present paper $\gamma = 0.95$ was used.

In the closeness approach, one requires

$$E_{\mathbf{X}}[P_{X_i}(-\infty \leq X_i \leq \widehat{w})] \approx \beta \tag{3}$$

i.e. closeness in coverage over all samples. The approximative requirement in expression (3) makes it possible to choose the estimator with minimal variance $V[\widehat{w}]$ and with small bias $E[\widehat{w} - w]$. Here, under normality assumption, the parametric estimator $\widehat{w}(\mathbf{X}) = \widehat{\mu} + \sqrt{(n+1)/n}\widehat{\sigma}t_{n-1;1-\beta}$ proposed by Wilks [18] was used. In the non-parametric case, let j and g be the integer and the decimal part of $n \times \beta$, respectively. Then the $100 \times \beta\%$ tolerance limit was estimated using the quantity $\widehat{w}(\mathbf{X}) = \left(X_{(j)} + X_{(j+1)}\right)/2$ if $g = 0$, and $\widehat{w}(\mathbf{X}) = X_{(j+1)}$ if $g > 0$. This is one out of five sample percentile definitions included in the statistical software packages SAS(r) and SPSS(r), see [11] and [13] respectively. Sample percentile definitions are frequently used to establish reference intervals in determining abnormality, see e.g. Ooi et al [9].

3 Consequences of the two approaches in a dose finding study

In a dose finding study, the obtained tolerance limits are compared with the limits of the *therapeutic window*, i.e. the desired concentration range where the chance of successful therapy is high. The information relating concentration to response can be obtained at three levels: through in vitro experiments, animal studies and investigations in human. Once the distribution of therapeutic and toxic effects at different concentration levels for a drug is obtained, it is possible to obtain a *utility curve* by weighting and adding the probabilities of the different therapeutic and toxic effects. The probabilities of adverse side effects are weighted with large negative values, and vice versa for the probability of desired effects, e.g. recovery**.** The utility curve has an optimum concentration at which therapeutic success is most likely, and there is a range of concentrations in which the chances of successful therapy is high. This is the *therapeutic window.* Precise limits are not definable, particularly considering the subjective nature of the utility curve,

but approximate limits will be determined by the location of the maximal efficacy and the slope of the curve. The shape of the curve may differ for different populations and even between objects, which must be taken into consideration in a study. For a further introduction to therapeutic response and toxicity, see Rowland&Tozer [10]. The utility curves and the therapeutic window for the two types of drugs discussed in this paper are shown in Figure 1. Note that the upper limit of the therapeutic window is close to toxic concentration levels and the different shapes of the two utility curves.

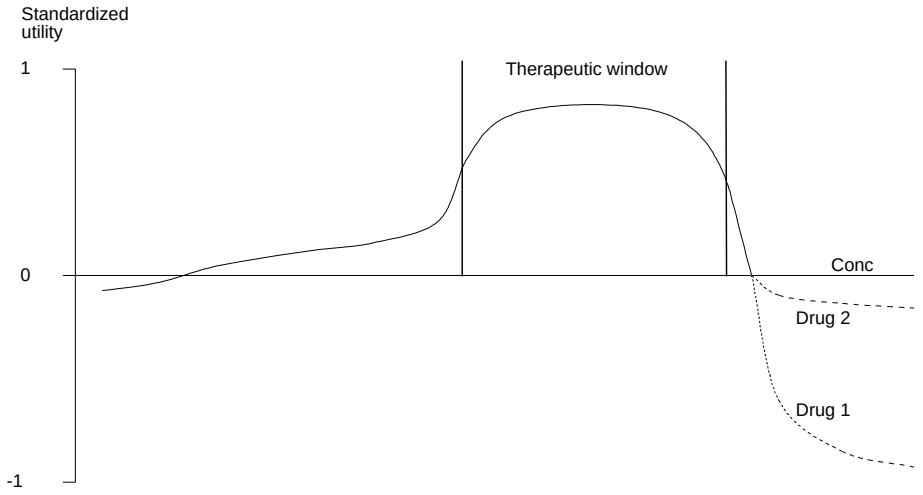

Figure 1: Utility functions of the concentration for Drug 1 and Drug 2. At positive values, the expected outcome is a successful therapy, e.g. well-being and recovery, and vice versa for negative values. The therapeutic window is the desired drug concentration interval. The weights are standardized for an outcome in the interval $[-1, 1]$.

The appropriate approach must be chosen depending on the utility function of the drug at hand. For Drug 1, the conservative approach is appropriate. When calibrating the dosage regimen for Drug 1, it is important not to underestimate the upper tolerance limit. The reason for this is that an obtained estimate lower than the upper limit of the therapeutic window indicates that the dosage should be increased. When based on an underestimate, the increased dosage will have the result that a larger than desired proportion of the population will receive unacceptably high concentrations. At these levels of concentration the risks of adverse side effects rapidly increase for Drug 1. Even a minor overdose may result in death. Using the conservative approach,

one may control for these risks. The coverage and the confidence level are chosen depending on the shape of the utility curve. However, requiring high coverages and confidence levels the dosage regimen tends to be low. This may result in a non-optimal dosage regimen where a larger than desired proportion of the population receives inefficient concentrations to the left of the therapeutic window. This is the disadvantage of the conservative approach, which may be well motivated by the side effects. With Drug 2, the expected outcome of an overdose is harmless, e.g. discomfort, and the non-optimal dosage regimen attained by the conservative approach is no longer justifiable by the side effects. In this case, the closeness approach may be more appropriate. The closeness criterion makes it possible to calibrate the dosage regimen to attain a large proportion of the population within the therapeutic window.

4 Results

Irrespective of which estimator is being used (parametric or non-parametric), the distributions of the estimates for the two approaches will principally differ in the same way. For the conservative approach the distribution will be skewed to the right due to the confidence criterion in expression (2). The skewness will increase with an increased confidence level. However, for the closeness approach the distribution of the estimates is less skewed. This is illustrated in Figure 2 where the obtained distributions of $\widehat{w}$ when estimating the 95% tolerance limit using the two non-parametric estimators are plotted. The true tolerance limit is marked with a straight vertical line at 10.67. In this particular simulation, a proportion of only 4.9% were underestimated using the conservative estimator due to the confidence criterion $\gamma \geq 95\%$. Using the closeness estimator, a proportion of 43% was underestimated. Notable is the difference in variance, which for the closeness approach was 39% of the variance for the conservative approach.

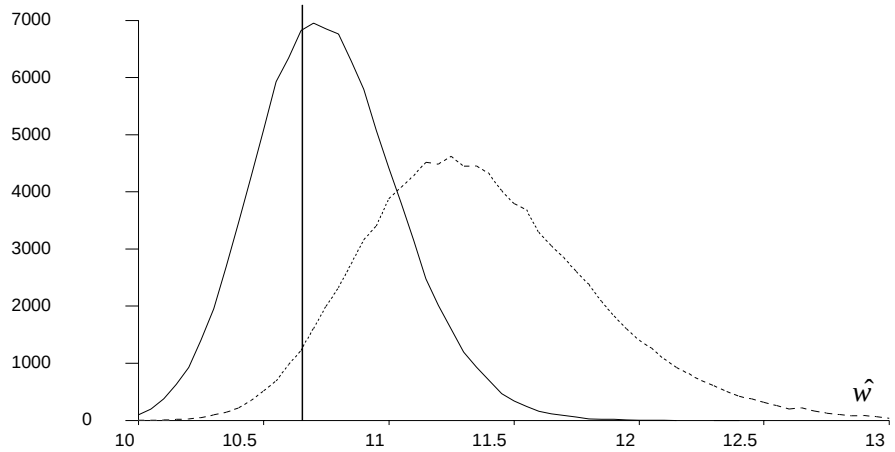

Figure 2: The distributions of the non-parametric $\widehat{w}$ when estimating the 95% tolerance limit, marked with a vertical line at 10.67,when $n = 59$. For the closeness approach, solid line, the distribution of $\widehat{w}$ is almost centered around the true limit but is highly skewed to the right for the conservative approach, dashed line. The number of replicates was 100.000.

4.1 Proportion of underestimates, bias and variance

The differences between the two approaches will principally be reflected by the proportion of underestimates and the bias when estimating the upper tolerance limit. The choice of confidence level for the conservative approach will to a large extent determine these differences with a maximum as the confidence is limiting to 1. Also the variance will be influenced in the same way by the confidence level. As could be seen from Figure 2, the differences may be considerable for the non-parametric estimators. A simulation study was carried out to illustrate how the two approaches differ for some β tolerance limits and some sample sizes using both parametric and non-parametric estimators. The sample sizes chosen are the minimal sample sizes where expression (2) still holds true for Wilks' non-parametric estimator. For these sample sizes the upper conservative limit is estimated with $X_{(n)}$. The results are given as $X/Y/Z$ in Table 1 and Table 2 where X is the proportion of underestimates and Y and Z are the bias and the variance, respectively, given as percentages of the corresponding values for the conservative approach. In Table 1, for example, it can be seen that using the closeness approach the bias and variance are 9.0% and 79% respectively compared to the use of the conservative approach for estimating the 95% tolerance limit when $n = 59$.

As expected, the bias and the variance are considerably smaller for the closeness approach than for the conservative approach. Using the closeness approach, it will be possible to obtain a relative stable estimate near the true tolerance limit. This would be preferable for drugs like Drug 2. For the two parametric estimators an increased sample size will monotonically lower the bias and the variance. However, the bias will decrease faster for the closeness approach than for the conservative approach, implying an increasing difference in bias. The opposite relation is true for the variance.

For the non-parametric estimators there will be a somewhat different relation. The bias and variance of the two non-parametric estimators will not decrease monotonically with an increasing sample size. Using order statistics as estimators the bias and variance will be determined by the positive relation between the expected value of an order statistic and the sample size. Both the non-parametric estimators use step functions of n for including and excluding order statistics. However, increasing the sample size, the bias will monotonically increase up to the next shift of order statistic, e.g. $X_{(n)}$ to $X_{(n-1)}$, where there is a decrease in one "jump" to a new lower level. This process will continue and the bias will asymptotically be zero. The variance also will not be a monotonically decreasing function of n.

Table 1: The performances of the conservative (CON) and the closeness (CLO) approach using the parametric estimators are compared. The percentage values $X/Y/Z$ are interpreted as the proportion of underestimates / the bias compared to the bias of CON / the variance compared to the variance of CON.

n	β :	97.5%	95%	90%	80%
119	CON	5.0/100/100	5.0/100/100	5.0/100/100	5.0/100/100
	CLO	44/8.6/84	46/6.5/85	47/4.5/86	48/2.6/89
59	CON	5.0/100/100	5.0/100/100	5.0/100/100	5.0/100/100
	CLO	42/12/79	44/9.0/79	46/6.2/81	48/3.5/84
29	CON	5.0/100/100	5.0/100/100	5.0/100/100	5.0/100/100
	CLO	38/16/71	41/12/71	44/8.4/73	46/5.0/76
14	CON	5.0/100/100	5.0/100/100	5.0/100/100	5.0/100/100
	CLO	33/22/61	37/17/60	41/11/61	45/6.5/65

Table 2: The performances of the conservative (CON) and the closeness (CLO) approach using the non-parametric estimators are compared. The percentage values $X/Y/Z$ are interpreted as the proportion of underestimates / the bias compared to the bias of CON / the variance compared to the variance of CON.

n	β :	97.5%	95%	90%	80%
119	CON	4.9/100/100	1.6/100/100	4.0/100/100	4.2/100/100
	CLO	43/10/37	45/5.5/42	47/5.1/71	48/2.9/84
59	CON		4.8/100/100	1.5/100/100	3.5/100/100
	CLO		43/9.5/39	45/4.9/46	47/3.8/75
29	CON			4.7/100/100	1.3/100/100
	CLO			44/8.6/42	46/3.8/51
14	CON				4.4/100/100
	CLO				45/6.9/47

4.2 Attained efficiency

Until now, focus has been on safety and the upper concentration limit of the therapeutic window. To achieve an efficient dosage regimen, i.e. a large proportion of the population within the therapeutic window, one also has to consider the lower limit. Calibrating a dosage regimen in a repeated study, one will successively increase the dosage from a low level until the obtained $100 \times \beta\%$ tolerance limit equals the upper limit of the therapeutic window. Assuming a constant variance of the concentration, this equality will be attained at different dosage levels for the two approaches. As can be seen from the comparisons of bias in the previous subsection and from Figure 2, the conservative approach is expected to reach an agreement at lower dosage levels than the closeness approach. Consequently, a higher proportion of the population will attain the inefficient concentration levels to the left of the therapeutic window for the conservative approach than for the closeness approach. This is the price of safety, which may be well motivated for Drug 1. For Drug 2, the risks of an overdose are negligible and it is then possible to use the closeness approach that does not control the risks of an overdose. The choice of approach depends now only on the expected efficiency which in some situations may be considerably higher for the closeness approach.

A simulation study of the difference in attained efficiency was performed for the parametric estimators, defining efficiency as the proportion of the population within the therapeutic window. The true 95:th percentile of the expression (5) in the Appendix was used as the fixed upper limit UL_{95} of the therapeutic window, and the lower limit LL_p was chosen as some p:th percentiles where $p < 95$. For each sample replicate the sample variance $\widehat{\sigma}^2$ was calculated. Utilizing the fact that the estimators have the same structure

$$\begin{aligned} \widehat{w}_{clo} &= \widehat{\mu} + g_{clo}\left(\widehat{\sigma}\right) \\ \widehat{w}_{con} &= \widehat{\mu} + g_{con}\left(\widehat{\sigma}\right) \end{aligned}$$

a dose finding study was imitated by calculating

$$\begin{aligned} \widehat{\mu}_{clo} &= UL - g_{clo}\left(\widehat{\sigma}\right) \\ \widehat{\mu}_{con} &= UL - g_{con}\left(\widehat{\sigma}\right) \end{aligned}$$

The parameters μ_{clo} and μ_{con} correspond to the attained logarithmic concentrations where the estimated tolerance limit equals the upper limit of the therapeutic window in a real dose finding study. Assuming constant variance σ^2 from expression (5) the difference in efficiency D was calculated for each replicate as:

$$\widehat{D} = P\left(LL_p \leq \widetilde{X}_{clo} \leq UL_{95}\right) - P\left(LL_p \leq \widetilde{X}_{con} \leq UL_{95}\right) \tag{4}$$

where $\widetilde{X}_{clo} \sim N\left(\widehat{\mu}_{clo}; \sigma^2\right)$ and $\widetilde{X}_{con} \sim N\left(\widehat{\mu}_{con}; \sigma^2\right)$. These distributions agree with the logarithmic concentration distributions that would be obtained in a dose finding study.

As can be seen in Table 3, the closeness approach will on the average attain a larger proportion within the window. When increasing UL or decreasing the sample sizes, the differences will increase. The standard deviation seems to have a more complex dependence on UL_p and the sample size, but is not further analyzed. For the difference there will be an optimal LL_p depending on the underlying distribution. In this case the differences will decrease for increasing $p > 40$.

Table 3: The mean of $\widehat{D}$ when using the parametric estimators. The standard deviation is given within brackets.

$n \backslash p$	5	10	30
14	0.14 (0.12)	0.19 (0.11)	0.24 (0.069)
29	0.058 (0.056)	0.098 (0.056)	0.16 (0.043)
59	0.025 (0.022)	0.052 (0.026)	0.10 (0.023)
119	0.011(0.001)	0.030(0.012)	0.067(0.012)

For the non-parametric estimators the simulation study had to be done in a different way. Using the same therapeutic window as before, a successively increased dosage was simulated letting the mean value of X_i in expression (5) successively increase from a low level until the estimated tolerance limit on the average was equal to UL_{95}. This was done independently for the two approaches, and the obtained means $\widehat{\mu}_{clo}$ and $\widehat{\mu}_{con}$ were used as in the

parametric case to calculate the difference in efficiency. From Table 4 it can be seen that the difference is considerable also for relative large sample sizes when using the non-parametric estimators. The difference was found to have the same dependences on UL, LL_p and the sample size as in the parametric case.

Table 4: The mean of $\widehat{D}$ when using the parametric estimators.

$n\backslash p$	5	10	30
119	0.054	0.10	0.17
59	0.072	0.12	0.20

It is now obvious that a proper usage of the two approaches is important. To avoid overdoses of Drug 1, one has to use the conservative approach, but this approach would often be too inefficient to be motivated for Drug 2.

5 Discussion

The importance of using the appropriate efficiency and safety approach in dose finding studies was discussed. It was shown that an appropriate approach for a drug with adverse side effects with minor overdoses might be inappropriate for a drug with harmless side effects. The performance of two common approaches, the closeness and the conservative approaches, were compared for some relevant tolerance limits. Using the conservative approach, one will only risk that a certain proportion of the population attains overdoses. The drawback is that the resulting dosage tends to be low, with the consequence that an undesired large proportion of the population attains inefficient concentrations. In many papers though conservative tolerance limits are proposed without a discussion of the consequences of an inappropriate use, see e.g. the discussions in Nickens [8] and Holst&Christensen [3]. Here, efficiency was defined as the proportion of the population within the therapeutic window. A more detailed study taking the exact form of the utility curve into account would be valuable.

A simulation study showed that the estimates of the true tolerance limits obtained with the closeness approach attain relative low bias and variance.

Such properties offer opportunities to adjust the dosage to obtain a large proportion of the population within the therapeutic window, i.e. where the chance of successful therapy is high. With this approach the confidence criterion in expression (2) may not be fulfilled, and consequently it is more appropriate in dose finding studies for drugs with harmless side effects with minor overdoses. However, for diseases with very harsh outcomes, e.g. AIDS, this approach may still be motivated ethically, even if there is a high probability of adverse side effects with minor overdoses. Here the proportion within the therapeutic window is the most important property. In the opposite situation, when treating a harmless disease, no drug with any side effects may be ethical.

Although this paper does not primarily discuss the estimators, some general comments may be done. Within both approaches there are several possible estimators. In applied research the underlying data generating process is never or seldom known and distributional assumptions may be questioned. The Food and Drug Administration [1] states that distributional assumptions should be done with restraint, which motivates the consideration of non-parametric estimators in this paper. There may be drawbacks with these simple estimators, e.g. relatively high and not monotonically decreasing variance and bias. In this paper the only assumption made for the non-parametric estimators was continuous data. When making more assumptions about the data, there are of course many possibilities to perform better. One possibility to make improvements is to introduce restrictions about the shape of the data, e.g. strictly increasing or decreasing functions. Such minor restrictions open a wide field of more efficient non-parametric methods. Also under distributional assumptions the estimates may be improved by using the longitudinal structure of the data. Cross sectional analysis is still dominating in pharmacokinetics, and an important task is to introduce longitudinal models.

Acknowledgment

I wish to thank my supervisor Ph. D. Robert Jonsson for suggesting the problem and for his devoted guidance throughout this work.

Appendix: The simulation model

The following model was used in the simulations when estimating the studied tolerance limits:

$$X_i = A_i + B_i \cdot t + U_i \qquad i = 1, 2, ..., n \tag{5}$$

where

- X_i the logarithm of the drug concentration for person i.
- A_i level factor specific to person i (metabolism, weight, sex, age).
- B_i change factor specific to person i (metabolism, renal clearance).
- t the time value
- U_i error term person i.

Letting A_i and B_i have the normal distribution with the parameter settings from Table 5, $cov\,(A_i, B_i) = 0.01$ and $t = 1$, then X_i will have the normal distribution with $\mu = 9$ and $\sigma^2 = 1.03$. The software SAS(r) 6.12 IML was used for the simulations with 200.000 replicates at each sample size giving a stable outcome. Further, the computer clock was used as random seed, and for the conservative approach the confidence level γ was set to 95%.

The original concentration scale $\exp(X_i)$ will have the lognormal distribution which is a common assumption for drug concentrations. The multiplicative model $\exp(X_i)$ has been used in pharmacokinetics when estimating *area under curve*, see Jonsson [4], Mandallaz&Mau [6] and Vuorinen&Turunen [15].

Table 5: Parameter settings for the simulations.

	Mean	Variance
A_i	10	0.5
B_i	-1	0.01
U_i	0	0.5

References

[1] Food and Drug Administration (1997). Draft Guideline on Statistical Principles for Clinical Trials. FDA Federal Register vol. 62 no. 90, 25712-25726.

[2] Gillespie, J. A. & Srinivasan, R. (1994). Distribution-Free Simultaneous Outer and Inner Tolerance Limits For Screening and Descriptive Purposes. Communications in Statistics-Theory and Methods 23(5), 1357-1364.

[3] Holst, E. & Christensen, J. M. (1992). Intervals For the Description of the Biological Level of a Trace-Element in a Reference Population. Statistician 41(2), 233-242.

[4] Jonsson, R. (1998). Comparison of trapezoidal and regression-based methods for estimating the area under curve using exponential time-response data. Working paper. Department of Statistics, Göteborg University.

[5] Kotz, S. & Johnson, N. L., Eds. (1988). Tolerance regions, statistical. Vol. 9. Encyclopedia of statistical sciences. 9 vol. + Suppl. vols. New York: Wiley.

[6] Mandallaz, D. and Mau, J. (1981). Comparison of different methods for decision-making in bioequivalence assessment. Biometrics 37, 213-222.

[7] Mee, R. W. (1989). Normal-Distribution Tolerance Limits For Stratified Random Samples. Technometrics 31(1), 99-105.

[8] Nickens, D. J. (1998). Using tolerance limits to evaluate laboratory data. Drug Information Journal 32(1), 261-269.

[9] Ooi, D. S., Innanen, V. T., Wang, D., Chong, G. L., Donnelly, J. G., Arseneault, J. J., Pronovost, C. & Wells, G. (1998). Establishing reference intervals for DPC's free testosterone radioimmunoassay. Clinical Biochemistry 31(1), 15-21.

[10] Rowland, M. & Tozer, T. N. (1995). Clinical pharmacokinetics : concepts and applications. 3 edit, Williams & Wilkins, London.

[11] SAS(r) (1990). SAS(r) Procedures Guide. SAS Institute Inc., 625-626.

[12] Somerville, P. N. (1958). Tables for obtaining non-parametric tolerance limits. The Annals of Mathematical Statistics 29, 599-601.

[13] SPSS(r) (1997). SPSS(r) Base 7.5 Syntax Reference Guide. SPSS Inc., 318-319.

[14] Tukey, J. W. (1947). Non-parametric estimation II: Statistically equivalent blocks and tolerance regions - the continuous case. The Annals of Mathematical Statistics 18, 529-539.

[15] Vuorinen, J. and Turunen, J. (1996). A three-step procedure for assessing bioequivalence in the general mixed model framework. Statistics in medicine 15, 2635-2655.

[16] Wald, A. (1943). An extension of Wilks´method for setting tolerance limits. The Annals of Mathematical Statistics 14, 45-55.

[17] Walsh, J. E. (1962). Handbook of nonparametric statistics, Van Nostrand, Princeton.

[18] Wilks, S. S. (1941). Determination of sample sizes for setting tolerance limits. The Annals of Mathematical Statistics 12, 91-96.

[19] Wilks, S. S. (1962). Mathematical statistics. Wiley publications in statistics, Wiley, New York.