



HAL
open science

Statistical aspects of design and validation of microtiter plate based linear and non-linear parallel in-vitro bioassay

Theo Dingermann, Hanne Zimmermann, Ludwig A. Hothorn, Daniel Gerhard

► To cite this version:

Theo Dingermann, Hanne Zimmermann, Ludwig A. Hothorn, Daniel Gerhard. Statistical aspects of design and validation of microtiter plate based linear and non-linear parallel in-vitro bioassay. *Biotechnology Journal*, 2009, 5 (1), pp.62. 10.1002/biot.200900146 . hal-00527653

HAL Id: hal-00527653

<https://hal.science/hal-00527653v1>

Submitted on 20 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Statistical aspects of design and validation of microtiter plate based linear and non-linear parallel in-vitro bioassay

Journal:	<i>Biotechnology Journal</i>
Manuscript ID:	biot.200900146.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	20-Jul-2009
Complete List of Authors:	Dingermann, Theo; Institut für Pharmazeutische Biologie, Biozentrum Zimmermann, Hanne; Roche Diagnostics GmbH, TG-QW Hothorn, Ludwig; Leibniz University Hanover, Institute of Biostatistics Gerhard, Daniel; Leibniz University Hanover, Institute of Biostatistics
Keywords:	Assay validation, non-linear PLA, linear PLA, mixed model approach, variance component analysis



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Research Article ((6103 words))

Statistical aspects of design and validation of microtitre-plate-based linear and non-linear parallel in-vitro bioassays

H. Zimmermann¹, D. Gerhard², T. Dingermann³, L.A. Hothorn²,

¹ Roche Diagnostics GmbH, 68305 Mannheim, Germany

² Leibniz University Hanover, Institute of Biostatistics, 30419 Hanover, Germany

³ Goethe University Frankfurt, Institute of Pharmaceutical Biology/ZAFES, 60438 Frankfurt/Main, Germany

Keywords

Assay validation, variance component analysis, mixed model approach, linear PLA, non-linear PLA

Corresponding author

Prof. Theo Dingermann, Goethe University Frankfurt, Institute of Pharmaceutical Biology/ZAFES, Max-von-Laue-Straße 9, 60438 Frankfurt/Main

Email: dingermann@em.uni-frankfurt.de, phone: +49 69 798-29650, fax: +49 69 798-29662

Abbreviations

EC₅₀ / ED₅₀: Effective concentration / effective dose resulting in 50% of the maximum response. Depending on the literature quoted, both terms are used in this paper, but have essentially the same meaning.

RFU: Relative Fluorescence Units

For Peer Review

Abstract

Assay validation was performed using four consecutive experiments with the related statistical evaluation. A cell-based assay on microtitre plates measured repeatedly within one day and on consecutive days was chosen as the model. The following problems were addressed: i) choosing an appropriate design on a plate to avoid heterogeneities, ii) quantification of all sources of variability, iii) selecting between linear and non-linear parallel line assays. A mixed model was used with the random factors: rows, columns and plates and fixed effect factors with either linear or non-linear parallel line models.

1 Introduction

Therapeutic proteins are nowadays produced using recombinant gene technology. The determination of their activity is commonly not feasible by physico-chemical methods and, therefore, in-vitro bioassays are used. In these, the relative potency of a test sample is estimated in comparison with an international or internal standard of known and clearly defined activity. The design and statistical evaluation of such assays are described in the US and European Pharmacopoeias [1, 2]. Commonly, a sigmoid dose-response curve is obtained with an approximately linear range around the ED_{50} . A non-linear, parallel model using all data or a linear, parallel model using a fraction of the data in the linear range can therefore be used as alternative evaluation methods.

In-vitro bioassays are commonly carried out on replicated microtitre plates. Several statistical problems may occur in the assay validation and evaluation. These are:

- i) Spatial heterogeneities caused by different positions within and between the plates that interfere with the measurements of the responses to different samples and dosages, e.g. differences between rows, columns, plates, days of measurements and the positions of the standards and samples,
- ii) Selection of an optimum design for either a non-linear or linear model utilising the available wells per plate. The objective is to guarantee an

1
2
3
4 unbiased and stable estimate of the relative potency with a narrow
5 confidence interval. Further, real data problems might also occur, such as
6 comparison of two different measurement devices and outlier detection.
7
8
9

10
11 The first problem was addressed by the variance component analyses. On
12 the basis of these results optimised linear and non-linear assay designs
13 were defined and compared.
14
15
16

17 18 **2 Materials and Methods**

19 20 **2.1 The bioassay**

21
22 An antibody based cytotoxicity assay was chosen as a model for a cell-
23 based bioassay and was used in all the experiments performed. It is an in-
24 vitro bioassay which measures the biological response of the cell line WIL
25 2-S to the antibody. The cell suspension (1×10^6 cells/ml), the growth
26 medium (RPMI1640), the dilutions of standard and unknown and the
27 controls were dispensed on microtitre plates. Incubation was performed at
28 37°C , 95% rH and 5% CO_2 for 15 – 26 hours. After incubation the viable
29 cell count was measured by determining the metabolic reduction of Alamar
30 Blue®. This dye exhibits fluorescence in response to cellular enzymatic
31 activity. The intensity of the fluorescence is measured with a photometer in
32 relative fluorescence units (RFU).
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

2.2 Microtitre plates

Standard commercial 96-well microtitre plates were used. As can be seen from Figure 1, the plates are divided into 8 rows (A - H) and twelve columns (1 - 12).

2.3 Variance component analysis

Variance component analysis is a statistical approach to characterise the variance components for a single continuous measure, e.g. optical density (RFU), for the possible sources of variation (factors), such as between days, between plates, between rows, between columns and the residual error. Commonly, a hierarchical design is used in variance component analysis, i.e. the factors are nested hierarchically, e.g. wells within a plate, replicated plates within days. A single plate in a microtitre-plate assay is characterised by a spatial row-column design (see Figure 1) which is not hierarchical and therefore a specific random effect model should be used. In particular, in a microtitre-plate assay both fixed effect factors (such as treatment with the test compound and the chosen fixed levels of standard, sample 1, sample 2) or potency (with fixed chosen levels of 70, 80, 100, 125 and 140%) and random effect factors (such days or plates) are included. A corpus of literature exists on the definition of fixed and random effect factors. Here it is clear that the three test levels were well chosen and are fixed, whereas the various days and various plates can be regarded as random selection from many days and many plates, not just

1
2
3
4 Monday, Tuesday, Wednesday, etc. For such a design, we used a mixed
5 model approach that included both fixed effect factors and random effect
6 factors, whereby the variance component estimates for the random effects
7 are of particular interest for the characterisation of the assay variability.
8
9 The objectives were i) to rank the sources of variation from seriously
10 important to negligible, ii) to reduce the impact of the serious factors, e.g.
11 by restricting their levels, or, if reduction was technically not feasible, iii) to
12 guarantee the control and stability of these serious factors.
13
14

15
16 In some publications on assay validation, the variance components are
17 inadequately characterised by simple estimation of the coefficient of
18 variation (CV) [3] or fixed effect analysis of variance (ANOVA) [4]. For
19 example, Morelock (2005) [5] characterised the variability for an IL-1RA in
20 vitro assay by comparison of coefficients of variation (CVs) only in a
21 complex validation study on 96 well microtitre plates using the IC₅₀
22 (estimated from 12 non-equidistant concentration on a plate using a Hill
23 model) for several days, runs, plates and rows.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 For the estimation of the EC₅₀ (effective dose) alone in microtitre-plate
41 assays used for screening for drug discovery, Rodrigues et al. (2007) [6]
42 compared the usual intra-plate design with a proposed inter-plate assay.
43 Systematic effects across each dimension of the plates were assessed, as
44 well as inter-plate effects. They detected apparent spatial trends between
45 the columns and rows that were considered to be insignificant.
46
47
48
49
50
51
52
53

1
2
3
4 Nevertheless, slight differences in the resulting potency values were
5 observed which were discussed and could as well be due to other
6 reasons, such as dilution issues. Unlike these findings, we found relevant
7 spatial differences within the plates. This may be explained by the fact that
8 we used a sensitive testing model and the most relevant concentrations of
9 the analyte, including the EC₅₀, in our experiments. In contrast, the
10 findings relating to spatial heterogeneities reported by Rodrigues et al. [6]
11 are mainly based on comparisons of the maximum and minimum controls.
12 Whereas Rodrigues [6] does not specify in detail the statistical
13 methodology, we have mainly applied variance component analysis as a
14 basic tool and combined it, where appropriate, with hierarchical models of
15 variance analysis. The results of this mixed model are both conclusive in
16 terms of statistics and valuable in terms of guidance for developing and
17 assessing optimised assay designs.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35 Several publications on assay validation using hierarchical variance
36 components exist, e.g. [7], for the random factors: two days, two analysts,
37 two HPLC devices and two columns. Variance component analysis has
38 already been described in [8], where it was used to assess a bioassay
39 based on luminescent bacteria with the sources “between batches”,
40 “between vials”, “between tubes” variance components for a hierarchical
41 design “batch > vial > tube”. In a pilot study, the authors estimated the
42 variance components and presented them as a percentage of the total
43
44
45
46
47
48
49
50
51
52

1
2
3
4 variability. Hereby the “between vials” variability dominated with 71%. Five
5 possible factors influencing the bioassay were investigated by means of a
6 factorial design, whereby the activation temperature was strictly controlled
7 in a second study. The total variability was reduced by 85%.
8
9

10
11
12
13
14 This technique was used in section 3.1. However, the focus of this paper
15 is on the use of mixed linear and non-linear models for both dose-
16 response models in a parallel-line assay (PLA) and the estimation of the
17 variance contributions of the random effects, i.e. we are not just
18 characterising the impact of the random factors, we are also modelling the
19 dose-response relationship – either linearly according to a parallel-line
20 assay (PLA) or according to a non-linear parallel-line assay.
21
22
23
24
25
26
27
28
29
30

31 For the estimation of an unbiased and precise potency level and its
32 confidence interval, an important issue is to choose between a linear and
33 a non-linear PLA. First of all, statistical approaches and related algorithms
34 are available, even as public domain software [9]. For ideal data, there is
35 no most powerful approach with relation to the number of concentrations,
36 number of replicates per concentration, variance, slope of the dose-
37 response relationship, etc. [10]. In practice, the selection of the linear part
38 for a PLA is a challenge, because the true range of the underlying sigmoid
39 dose-response curve is unknown and an un-centred choice results in non-
40 linearity and even non-parallelism. However, a linear PLA needs only two
41
42
43
44
45
46
47
48
49
50
51
52
53

1
2
3
4 concentrations (or three to check for non-linearity) and therefore many
5 more replicates can be used on a microtitre plate. From this perspective, a
6 non-linear PLA is much more robust. However it needs a design with more
7 concentrations and five instead of only two concentrations have to be
8 assessed. Therefore, in addition to the statistical simulation study [10], we
9 compare the two designs using two experiments reported in sections 3.2
10 and 3.3.
11
12
13
14
15
16
17
18

19 **3 Experiments and Results**

20 **3.1 Analysis of spatial heterogeneities and trends on microtitre** 21 **plates**

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Spatial heterogeneities and trends are known to occur within and between microtitre plates that will bias the results of the bioassays with respect to the EC_{50} and potency estimates. This problem is usually overcome by omitting extreme positions on the plates, i.e. the margin positions of rows A and H and the columns 1 and 12 (see Figure 1). Furthermore, replicates of the dosages of the standards and samples are measured that are distributed on the plates in such a way that preferably the replicates of each treatment but not their means are affected by the spatial heterogeneities caused by different positions within and between the plates.

In order to investigate the influence of any spatial heterogeneity on the growth of the cells, all wells of all plates were loaded with a constant

1
2
3
4 number of 1×10^6 cells/ml and a uniform concentration of $1 \mu\text{g/ml}$
5 antibody. The concentration of antibody falls in the range of the EC_{50}
6 concentrations observed. Therefore, the resulting responses can be
7 reasonably expected to be in the central and steep part of the dose-
8 response curve. The plates were incubated with Alamar Blue® and
9 measured photometrically in accordance with the usual procedures.

10
11 A total of $N = 9$ plates was prepared on three days, three plates per day.
12 As described above, all wells were uniformly loaded with medium, cells
13 and $1 \mu\text{g/ml}$ of the antibody. In addition, the plates were read with
14 photometers from two different manufacturers. This also allows analysis of
15 device-specific influences on the measurement.

16
17 Heat maps were chosen as spatial descriptive statistics, (see Figure 3)
18 which is a convenient way of displaying the complete raw data, in which
19 the relative fluorescence units (RFUs) are represented by different grey
20 scales. Dark grey squares represent low values, whereas light grey
21 regions represent measurements of higher relative fluorescence units.
22 Significant spatial effects are seen predominantly at the borders of the
23 plates, which therefore justifies the omission of the marginal rows and
24 columns. Other spatial effects are apparently distributed randomly over the
25 plates and should not therefore interfere with the chosen assay design.
26 The question arises whether this observation in a descriptive plot can be
27 quantified by means of quantitative estimation of the underlying variance
28 components using a mixed model.

1
2
3
4 The MIXED procedure in the SAS System 9.1 was used for the estimation
5 of variance components because asymptotic confidence intervals are
6 available. A random-component model was fitted using the Mivque
7 method with the hierarchical random factors: days, plates (within days),
8 rows (within plates, within days) and columns (within plates, within days).
9 Figure 4a shows the percentages of the variances in a pie-chart for the
10 complete raw RFU data.
11
12
13
14
15
16
17

18
19 Most of the observed variance is explained by the differences between
20 days (35%), which is not surprising because it is not possible to use the
21 same cell suspension on different days and, therefore, the variance
22 between days also reflects the variance caused by different physiological
23 conditions, growth stages and cell count between days. The variance
24 component for the plates within days is also quite significant with 20%.
25 This may reflect unavoidable differences in the incubation conditions
26 resulting, for example, from different position in the incubators. The
27 variance component for columns within plates and dates was found to be
28 16%, which is also a significant value. The question is whether one of
29 these serious variability-determining factors can be reduced. After some
30 experimenting, we were able to reduce the variance components of the
31 factor columns within plates and dates and rows within plates and dates to
32 only 8% or 1 % of the complete variance respectively, (see Figure 4b) by
33 changing the design on the plate by omitting the marginal columns and
34 rows, i.e. a plate design using rows B - G and columns 2 – 11 only.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

1
2
3
4 Further omission of the adjacent columns/rows, i.e. using only rows B - F
5 and columns 2 – 10, reduced the percentage influence to 1% and 0 %
6 respectively, (see Figure 4c). This design (see Figure 2) was used for the
7 validation of the linear parallel line assay described in section 3.3.
8
9

10
11
12 Therefore, the question arises whether the design with triplicate standard
13 samples in rows {3, 6, 9}, with triplicate test A samples in rows {2, 5, 8},
14 and with triplicate test B samples in rows {4, 7, 10} (see Figure 2) is
15 appropriate. Figure 4c shows instead of the variability between columns
16 on a plate, the variability between the columns corresponding to three
17 specific sample locations. Also, most of the variability of the experiment
18 can be explained by the different days and plates. The variability between
19 columns of the treatment locations is negligibly small. Therefore, one can
20 argue that, although there is no complete randomisation, the spatial
21 effects on a plate may not influence the inference between the different
22 treatments. To analyse possible heterogeneities between these row
23 positions, a pseudo-PLA between the three can be determined. Figure 5
24 shows linear curves with near-to-zero slopes without any heterogeneity
25 between standard, which is represented by dark grey lines and circles (O),
26 test A, represented by medium grey lines and triangles (Δ) and test B,
27 represented by light grey lines and plus signs (+). This design for a
28 microtitre plate was therefore used in the following.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

Deleted: b

3.2 Evaluation of a design for both linear and non-linear PLA

In this experiment we used a design including nine concentrations (10, 3, 2, 1.33, 0.89, 0.59, 0.40, 0.26, 0.08 $\mu\text{g/ml}$) for both standard and samples which were triplicated on a single microtitre plate. Three independent “known” samples per day were used to cover the possible potency range from 80 to 125% of the standard. This complete set was investigated on 6 days, i.e. a total of 18 plates were used. The statistical analysis of the variance components was performed with a linear mixed model to characterise the linear PLA and with a non-linear mixed model to characterise the non-linear PLA. Figure 6 shows the scatter plots for the raw RFU data of the 18 plates, where the standard is represented by a circle (O), the 80% sample by a triangle (Δ) and the 125% sample by a plus sign (+). Reproducible sigmoid dose-response curves for the 18 plates are shown.

3.2.1 Linear PLA

We first analysed the data by means of a linear PLA. To determine the variance components, the relationship between RFU and the logarithmic concentration in $\log(\mu\text{g/ml})$ was described with a hierarchical linear model. The differences between the treatments (80% and 125% of the concentration of the control) and the slope of the linear model were determined as fixed factors. To reflect the experimental design, a random effects structure was assumed with treatments nested in the plates, which

1
2
3
4
5 in turn were nested in a factor accounting for the different dates when the
6 experiment was performed. In addition a random linear slope was
7 assumed for all of the random factors. The predictions by means of this
8 model are presented in Figure 7, together with the raw data. It can easily
9 be seen that a linear model for a linear PLA was inappropriate for this wide
10 range of concentrations. This is supported by diagnostic graphics such as
11 QQ-Plots or plotting the residual versus the predicted values in Figures 8a
12 and b.
13
14
15
16
17
18
19

20 Table 1 presents the estimated variance components for the random
21 factors in this linear mixed model. The dominant effects are the intercept
22 heterogeneities between the three treatments within the 18 plates.
23 Irrespective of these interfering effects, the estimated potencies are near
24 the target values and their confidence intervals clearly contain the target
25 values, see Table 2.
26
27
28
29
30
31
32
33

34 **3.2.2 Non linear 4-PL**

35
36 Secondly we analysed the data by a non-linear 4-PL model. The
37 relationship between RFU and the concentration in $\mu\text{g/ml}$ was described
38 by a hierarchical four parametric log-logistic model (4-PL). The model fits
39 are shown in Figure 9.
40
41
42
43

44 The model fit and the data correspond well over these 18 plates. The
45 differences between the treatments (80% and 125% of the concentration
46 of the control) were determined as fixed factors. To reflect the
47 experimental design, a random effect structure was assumed with
48
49
50
51
52

1
2
3
4 treatments nested in the plates, which in turn were nested in a factor
5 accounting for the different dates when the experiment was performed.
6
7 For these random effect factors, variance components were determined,
8
9 (see Table 3) showing a particularly high variability of the upper
10 asymptotes for both the factors date and plate. This is quite obvious
11 because the upper asymptotes stand for uninfluenced growth of the cells
12 due to a negligible concentration of antibody and, therefore, reflect the
13 biological variability of the cells.
14
15
16
17
18
19

20 The estimated potencies meet the target values quite well but, when
21 compared with the linear model, the confidence intervals are much wider,
22 thus indicating lower precision of the non linear PLA as performed.
23
24
25
26
27

28 **3.2.3 Conclusions**

29
30 At first sight the combined design for either a linear or a non-linear PLA
31 does not seem to be a really appropriate solution. Nevertheless, a
32 comparison of the two models that is not biased by different quality of data
33 can be made from a single set of data. This comparison should include
34 selection of the data used, e.g. limiting the linear PLA to the linear portion
35 of the curve and/or modifications in the methods of evaluation, e.g. to
36 reduce the variability of the asymptotes by taking account of the biological
37 background. On the basis of the unmodified data, with the reservations
38 made, both models seem to be equivalent in terms of accuracy, but efforts
39 are necessary to improve the precision of the non-linear model.
40
41
42
43
44
45
46
47
48
49
50

3.3 Performance characteristics of an optimised linear parallel line design

In this experiment the biological activity was determined by means of an optimised linear parallel line model using measurements from the linear part of the dose response curve only. Five dose levels were used, namely 2, 1.33, 0.89, 0.59 and 0.4 µg/ml. The objective was to evaluate the results in terms of accuracy and precision of the linear PLA and later comparison with the non-linear parallel line model.

Three compounds were tested for each treatment, the standard (S) and two samples (A and B) with unknown activity. The whole set of replicates for the standard and the unknown samples were placed on one plate following an assay design that had been verified in pre-validation studies. In this validation study, the samples A and B were identical, the only difference being their positions on the plate. The samples were derived from the standard solution which was diluted to target concentrations of 70%, 80%, 100%, 125% and 140% of the standard. This strategy allows assessment of accuracy in terms of percent recovery. Five plates, one for each target concentration, were loaded on one day. This procedure was repeated on three days resulting in a total of 15 plates.

Figure 10 shows the predictions of the model for the 15 individual plates (plates 1, 2 and 3 for the 5 concentrations). As expected, it can be seen that the dark grey curve for the standard (O) overlaps with the medium and light grey lines for sample A (Δ) and B (+) for 100%, whereas the

1
2
3
4 standard curve is above the sample curves for activities larger than 100%
5 and below for activities less than 100%. The medium and light grey lines
6 of samples A (Δ) and B (+) are both similar with respect to slopes and
7 intercepts. Already here in Figure 10 we see heterogeneity between the
8 plates 1, 2 and 3 according to all intercepts of the three lines, i.e. plate-
9 specific different levels of RFU can be observed. The relative potency
10 estimates and their two-sided 95% confidence intervals are summarised in
11 Table 5.
12
13
14
15
16
17
18
19
20

21
22 The reproducibility of the relative potencies (Table 5) is excellent, i.e. the
23 two-sided 95% confidence intervals include the true activity in each case.
24
25 Moreover, the relative potency for the comparison between the samples A
26 and B is close to 1. The estimated percentage bias is extremely small. A
27 plot of the residual versus the predicted values for the parallel line model
28 is given in Figure 11. A clearly non-linear pattern is evident, which
29 suggests that the optimum linear range was not chosen in this particular
30 experiment. In practice this can be done using different dose levels,
31 different dose increments, smaller dose range or less dose levels (down to
32 three or even two).
33
34
35
36
37
38
39
40
41
42

43 As a statistical evaluation, a mixed model including the fixed factors
44 treatment (with levels 'Standard', 'Sample A', 'Sample B'), potency (with
45 levels 70, 80, 100, 125 and 140% target potencies) and its interaction, and
46 the random effects plate (with 3 levels), potency was nested inside plate
47
48
49
50

1
2
3
4 as the factor plate denotes also the day the experiment was performed,
5
6 whereas for all of the random effects a random slope was also assumed.
7
8

9 The estimated variance components for the intercepts and slopes between
10 plates and between potencies are presented in Table 6. The dominant
11 heterogeneity, with nearly 39% explaining the variability, was detected for
12 the plate-specific intercepts, as we already guessed from Figure 10. This
13 means for later routine bioassays: a potency estimate from a single plate
14 is over-optimistically precise. To achieve an unbiased potency estimate,
15 several plates should be used.
16
17
18
19
20
21
22
23

24 **3.4 Performance characteristics of an optimised non-linear parallel** 25 **line design** 26 27

28 An alternative approach is to estimate potency by means of the non-linear
29 4-PL model. Therefore, a special design that included six dosages (10,
30 2.8, 1.6, 0.8, 0.4 and 0.08 $\mu\text{g/ml}$) with 3 replicates for the standard and
31 each unknown sample was used in this experiment. The remaining set-up
32 was analogous to section 3.3. One standard and two samples (unknowns
33 A and B) were prepared on each plate. The samples A and B were
34 identical, the only difference being their positions on the plate. The
35 samples were derived from the standard solution, which was diluted to
36 target concentrations of 70%, 80%, 100%, 125% and 140% of the
37 standard. Five plates, one for each target concentration, were loaded on
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

1
2
3
4 one day. This procedure was repeated on two separate days thus giving a
5
6 total of 15 plates.
7

8
9 Figure 12 shows the model fit for the 15 individual plates. Analogously to
10
11 the linear PLA we see the dark grey curve for the standard (O) overlaps
12
13 with the medium and light grey lines of samples A (Δ) and B (+) for 100%,
14
15 whereas the standard curve is above the sample curves for activities
16
17 greater than 100% and below for activities less than 100%. The lines of
18
19 samples A (Δ) and B (+) are similar over the entire range of
20
21 concentrations. Also we assume heterogeneity between plates 1, 2 and 3
22
23 according to all upper asymptotes, i.e. generally plate-specific different
24
25 levels of RFU were measured.
26

27
28 The relative potency estimates and their two-sided 95% confidence
29
30 intervals are summarised in Table 7. As compared with the linear PLA, the
31
32 reproducibility of the relative potencies is limited, in particular for the
33
34 samples with the extreme targets 0.70 and 1.40. However, the two-sided
35
36 95% confidence intervals include the true activity in each case. Moreover,
37
38 the width of the confidence intervals is much greater than that in the linear
39
40 PLA, reflecting the determination of the non-linear models over
41
42 heterogeneous plates. The bias was as large, with a tendency to
43
44 underestimate the true potencies of less than 100% and an overestimation
45
46 of the true potencies greater than 100%.
47
48
49
50
51
52
53

1
2
3
4 In a non-linear mixed model assuming a four parameter log-logistic dose-
5 response model, the three fixed effects (determining the four parameters
6 of the logistic model for each factor level): treatments (with levels:
7 'Standard', 'Sample A', 'Sample B'), potency (with levels 70, 80, 100, 125,
8 and 140%) and its interaction and the random effects: plate (with 3 levels)
9 and relative potency (nested inside the plate), the variance components
10 were estimated for the upper and lower asymptote, the EC50 value and
11 the slope. The estimated variance components for the four model
12 parameters and the random factors are shown in Table 8. The dominant
13 heterogeneity is the upper asymptote, which is visualised in Figure 12,
14 analogously to the heterogeneities in the intercepts in section 3.3.
15
16
17
18
19
20
21
22
23
24
25
26
27

28 **3.5 Method comparison**

29
30 In the experiments in section 3.13.1, the RFUs for the same concentration
31 in all wells were measured using two different fluorescence reader –
32 Mithras LB 940 and Varioskan Flash Multimode reader. The two
33 measurements are on rather different scales: Mithras LB 940 between
34 70,000 and 100,000 and Varioskan Flash between 300 and 500. The two
35 approaches for model comparison were therefore performed: linear
36 regression analysis and variance component analysis for the ratio of the
37 levels of individual wells. Figure 13 shows the scatter plot with the fit.
38
39 There is no specific structure, either between the different days or the
40 different plates. The Pearson correlation coefficient was calculated to be
41 0.967 with two-sided confidence interval [0.962, 0.971], i.e. at the level of
42
43
44
45
46
47
48
49
50
51
52
53

1
2
3
4 individual wells, the two measurements are comparable. No substantial
5
6 plate or day effect was found either for the elementary correlation
7
8 coefficients. Table 9 shows overlapping confidence intervals for plate/day
9
10 specific correlation coefficients, i.e. there is no difference between the two
11
12 measurement approaches in any condition.
13
14

15 16 **4 Concluding Remarks**

17
18 Assay validation was determined in a series of experiments with their
19
20 related statistical evaluation. The principle design was a cell-based assay
21
22 on a microtitre plate, repeatedly measured within one day and on some
23
24 consecutive days. The basic questions to be answered were: i) what is an
25
26 appropriate design on a plate to avoid heterogeneities, ii) what sources of
27
28 variability are serious and can they be eliminated, or at least restricted, iii)
29
30 is the evaluation appropriate by means of a linear or a non-linear parallel line
31
32 assay? The mixed model with random factors rows, columns, plates, and
33
34 the fixed effect factors of either a linear or even a non-linear parallel line
35
36 model (PLA) was used as standard statistical technique.
37

38
39 In a first experiment possible heterogeneities on the plate were
40
41 determined, whereby an optimum design for a single plate was chosen. In
42
43 the second experiment a compromise design for both linear and non-linear
44
45 PLA was used to compare performance parameters for both strategies
46
47 from one set of data. In the third and fourth experiments, optimum designs
48
49 for either a linear or a non-linear PLA were investigated. The linear PLA
50
51
52

1
2
3
4 gives the most precise and unbiased potency estimations and narrow
5 widths of the confidence intervals. However, there are further possibilities
6 for improvement of the design of the linear PLA by choosing the linear
7 range of the sigmoid dose-response curve better, e.g. using only four or
8 even three concentrations and/or using smaller dose increments. The non-
9 linear PLA was apparently inferior to the linear PLA in terms of
10 reproducibility and precision. Moreover, it showed a bias of overestimating
11 target potencies greater than 100% and underestimating potencies smaller
12 than 100%. Bearing in mind the results of the variance component
13 analysis, the reason for this surprising finding may be the variability of the
14 upper asymptote, which also influences the calculation of the ED50. In our
15 experiments the dose-response curves for the standard and the samples
16 were calculated individually for each treatment. Depending on the nature
17 of the experiments performed and supported by validation experiments, it
18 may be wise to calculate common asymptotes for a set of dose response
19 curves and focus the evaluation on the data points between the ED80 and
20 the ED20 of the individual dose response curves.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 It is important to notice the – not surprising – dominant variability between
42 plates within a day and especially between consecutive days. Clearly this
43 is technically unavoidable. However, relevant estimations of potencies
44 (and their confidence limits) should not be based on one plate alone, but
45
46
47
48
49
50
51
52
53

1
2
3
4 precisely on such repeated plates, in order to avoid pseudo-precise
5
6 results.
7
8

9
10 Finally, we compared two means of measurements of the relative
11
12 fluorescence units and found comparability.
13
14

17 5 References

18
19
20 [1] USP 32 – NF 27, 2008, <111> Design and Analysis of Biological assays
21

22
23
24 [2] European Pharmacopoeia 6.0. 2008; 5.3 Statistical Analysis of Results of
25
26 Biological Assays and Tests
27

28
29
30 [3] Jacobson L, Middleton B, Holmgren J: An optimised automated assay for
31
32 determination of metabolic stability using hepatocytes: assay validation, variance
33
34 component analysis and in vivo relevance. Assay and Drug Development
35
36 Technology (2007) 5, 3, 403-415.
37

38
39 [4] European Pharmacopoeia, Chapter 3.2.2 "Analysis of variance", 2005, p. 696.
40
41

42
43 [5] Morelock, M.M., Hunter E.A. et al. (2005) Statistics of assay validation in high-
44
45 throughput cell imaging of nuclear factor xB nuclear translocation. Assay and
46
47 Drug Development Technology. Vol. 3, No. 5.
48
49

1
2
3
4 [6] Tsang PKS, Larew, JSA, Larew, LA: Statistical approaches to determine
5 analytical variability and specifications: applications of experimental design and
6 variance component analysis. J of Pharmaceutical and Biomedical Analysis
7
8 16(1998), 1125-1141.
9
10

11
12
13 [7] Ren S, Frymier PD: Reducing bioassay variability by identifying source of
14 variation and controlling key parameters in assay protocol. Chemosphere
15
16 57(2004) 81-90.
17
18

19
20
21 [8] Rodrigues, D.J., Lyons, R. et al (2007) A three-stage experimental strategy to
22 evaluate and validate an interpolate IC50 format. Assay and Drug Development
23
24 Technology. Vol. 5, No. 6.
25
26

27
28 [9] Ritz, C. and Streibig, JC (2009). Non-linear regression with R. useR! Springer,
29
30 New York,
31
32

33
34 [10] Gerhard, D. and Hothorn, L.A. (2009) A simulation study for linear and non-
35
36 linear PLA to estimate relative potency. Internal Report.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

Figure legends

Figure 1: 96 well microtitre plates

Figure 2: Positions of the treatments in the linear parallel line design

S₁ - S₅: designated positions of the dosages 1 to 5 of the standard

A₁ - A₅ / B₁ - B₅: designated positions of the dosages 1 to 5 of the samples A and B

M = designated positions of the blanks (growth medium)

(In order to detect spatial differences that might interfere with the assay design, all wells were uniformly loaded)

Figure 3: Heat maps for the measured RFU values

Figure 4a: Percentage of variance components of the RFU distribution on plates

– complete data

Figure 4b: Percentage of variance components of the RFU distribution on plates

– new design omitting marginal rows and columns (rows B - G and columns 2 – 11)

Figure 4c: Percentage of variance components of the RFU distribution on plates

– new design omitting marginal rows and columns (rows B – F and columns 2 – 10)

Figure 5: Pseudo linear PLA for detecting possible heterogeneities

Figure 6: Raw data scatter plots (80%: medium grey, Δ; 100%: dark grey, O;

125%: light grey, +)

Figure 7: Predictions of the hierarchical linear model separately for each of the

18 plates (80%: medium grey, Δ; 100%: dark grey, O; 125%: light grey, +)

Figure 8a, b: QQ- and residual plot for the linear mixed model

1
2
3
4 **Figure 9:** Predictions of the hierarchical 4 parameter log-logistic model

5
6 separately for each of the 18 plates (80%: medium grey, Δ ; 100%: dark grey, O;

7
8 125%: light grey, +)

9
10 **Figure 10:** Model fit for the optimized linear parallel line model

11
12 **Figure 11:** Residual plot for the parallel line model

13
14 **Figure 12:** Model fit for the non linear parallel line model

15
16 **Figure 13:** Scatter plot of Mithras LB940 and Varioskan Flash multimode reader

17
18 RFU measures (different greyscales identify different days, different symbols

19 identify different plates)

Table 1: Estimated variance components for the random effects structure of the mixed effect linear model

Factor	Source	$\sqrt{Variance}$
Date	Intercept	0
	Slope	0
Plate in date	Intercept	3820.42
	Slope	1378.28
Treatment in plate in date	Intercept	4607.49
	Slope	996.26
Residual		17186.42

Table 2: Estimated potencies for the comparison of the standard with the treatment with 80% of the active compound and comparison of the control with the 125% treatment.

Target potency	EC ₅₀ Comparison	Estimated rel. potency	Lower CI	Upper CI
0.8	Standard / Sample	0.76	0.70	0.83
1	Standard / Sample	1.02	0.94	1.11
1.25	Standard / Sample	1.25	1.15	1.36

Table 3: Estimated variance components for the random effects for the 4 PL model

Factor	Source	$\sqrt{\text{Variance}}$
Date	Upper Asymptote	10532.51
	Lower Asymptote	1472.71
	Slope	0.10
Plate in date	Upper Asymptote	12527.26
	Lower Asymptote	605.58
	Slope	0.04
Treatment in plate in date	Upper Asymptote	18.40
	Lower Asymptote	0.06
	Slope	0.00
Residual		4704.72

Table 4: Potency estimation by the 4-PL model

Target potency	EC50 Comparison	Estimated rel.	Lower	Upper
0.8	Standard / Sample A	0.85	0.59	1.47
1	Standard / Sample A	1.02	0.74	1.52
1.25	Standard / Sample A	1.25	0.78	2.53

For Peer Review

Table 5: Potency estimates for the optimised linear parallel line assays

Target potency	EC50 Comparison	Estimated			Estimated % bias
		rel. potency	Lower CI	Upper CI	
0.7	Standard / Sample A	0.71	0.67	0.74	-1.4
0.7	Standard / Sample B	0.72	0.69	0.75	-2.9
1	Sample A / Sample B	1.02	0.98	1.06	-2.0
0.8	Standard / Sample A	0.80	0.76	0.83	0.0
0.8	Standard / Sample B	0.81	0.78	0.85	-1.3
1	Sample A / Sample B	1.02	0.98	1.06	-2.0
1	Standard / Sample A	1.00	0.97	1.04	0.0
1	Standard / Sample B	0.99	0.95	1.03	1.0
1	Sample A / Sample B	0.99	0.95	1.03	1.0
1.25	Standard / Sample A	1.20	1.15	1.25	4.0
1.25	Standard / Sample B	1.20	1.16	1.25	4.0
1	Sample A / Sample B	1.01	0.97	1.04	-1.0
1.4	Standard / Sample A	1.35	1.29	1.42	3.6
1.4	Standard / Sample B	1.36	1.30	1.43	2.9
1	Sample A / Sample B	1.01	0.97	1.05	-1.0

Table 6: Variance component estimates for the estimated relative potencies

$\sqrt{\text{Variance}}$	Source		% Variance component
	Plate:potency (Intercept)	3266.1	8.77
	Slope	6646.5	17.85
	Plate (Intercept)	14353.8	38.54
	Slope	7233.9	19.42
	Residual	5744.2	15.42

For Peer Review

Table 7: Potency estimates of the optimised non linear parallel line assays

Target potency	EC50 Comparison	Estimated			Estimated % bias
		rel. potency	Lower CI	Upper CI	
0.70	Standard / Sample A	0.52	0.37	0.74	25.7
0.70	Standard / Sample B	0.54	0.38	0.75	22.9
1.00	Sample A / Sample B	1.02	0.74	1.42	-2.0
0.80	Standard / Sample A	0.74	0.50	1.08	7.5
0.80	Standard / Sample B	0.85	0.56	1.29	-6.2
1.00	Sample A / Sample B	1.05	0.62	1.77	-5.0
1.00	Standard / Sample A	1.08	0.78	1.50	-8.0
1.00	Standard / Sample B	0.95	0.70	1.30	5.0
1.00	Sample A / Sample B	0.91	0.63	1.33	9.0
1.25	Standard / Sample A	1.24	0.91	1.70	0.8
1.25	Standard / Sample B	1.12	0.83	1.52	10.4
1.00	Sample A / Sample B	1.23	0.78	1.94	-23.0
1.40	Standard / Sample A	1.72	1.09	2.72	-22.9
1.40	Standard / Sample B	1.91	1.22	2.99	-36.4
1.00	Sample A / Sample B	1.14	0.62	2.09	-14.0

Table 8: Variance component estimates for the estimated relative potencies and their confidence intervals

Parameter	$\sqrt{\text{Variance}}$	
	Plate	Potency within plate
Upper Asymptote	26660.27	11161.04
Lower Asymptote	1970.10	719.29
EC50	0.06	0.06
Slope	0.14	0.03
Residual error		3865.63

Table 9: Day- and plate-specific correlation coefficients between Mithras and Varioscan RFU measures

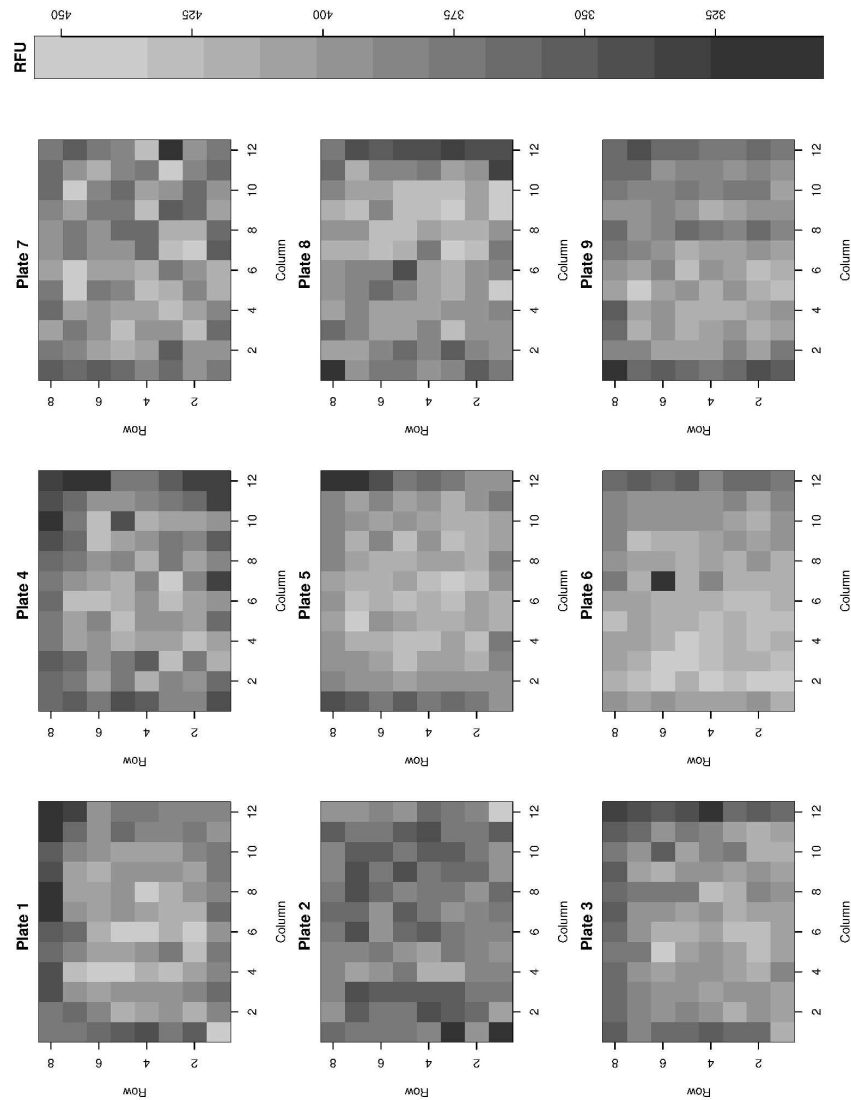
Day	Plate	Coefficient of correlation	Lower limit	Upper limit
1	1	0.952	0.929	0.968
1	2	0.962	0.943	0.974
1	3	0.926	0.891	0.950
2	4	0.941	0.912	0.960
2	5	0.972	0.959	0.981
2	6	0.949	0.924	0.966
3	7	0.958	0.938	0.972
3	8	0.931	0.898	0.954
3	9	0.964	0.947	0.976

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

		Columns											
		1	2	3	4	5	6	7	8	9	10	11	12
Rows	A	○	○	○	○	○	○	○	○	○	○	○	○
	B	○	○	○	○	○	○	○	○	○	○	○	○
	C	○	○	○	○	○	○	○	○	○	○	○	○
	D	○	○	○	○	○	○	○	○	○	○	○	○
	E	○	○	○	○	○	○	○	○	○	○	○	○
	F	○	○	○	○	○	○	○	○	○	○	○	○
	G	○	○	○	○	○	○	○	○	○	○	○	○
	H	○	○	○	○	○	○	○	○	○	○	○	○

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

		Columns											
		1	2	3	4	5	6	7	8	9	10	11	12
Rows	A	M	M	M	M	M	M	M	M	M	M	M	M
	B	M	A₁	S₁	B₁	A₁	S₁	B₁	A₁	S₁	B₁	M	M
	C	M	A₂	S₂	B₂	A₂	S₂	B₂	A₂	S₂	B₂	M	M
	D	M	A₃	S₃	B₃	A₃	S₃	B₃	A₃	S₃	B₃	M	M
	E	M	A₄	S₄	B₄	A₄	S₄	B₄	A₄	S₄	B₄	M	M
	F	M	A₅	S₅	B₅	A₅	S₅	B₅	A₅	S₅	B₅	M	M
	G	M	M	M	M	M	M	M	M	M	M	M	M
	H	M	M	M	M	M	M	M	M	M	M	M	M

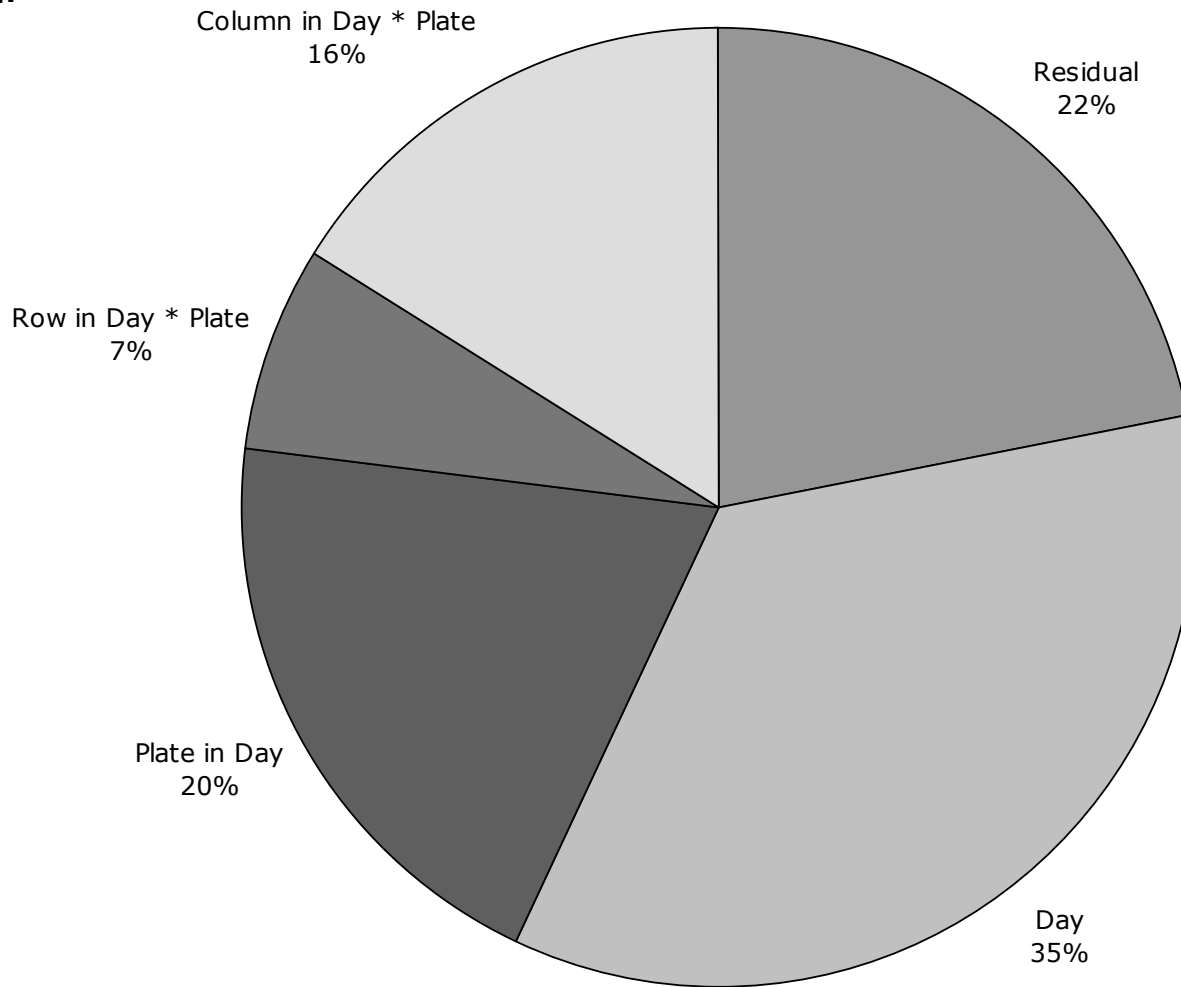


Heat maps for the measured RFU values
215x279mm (600 x 600 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

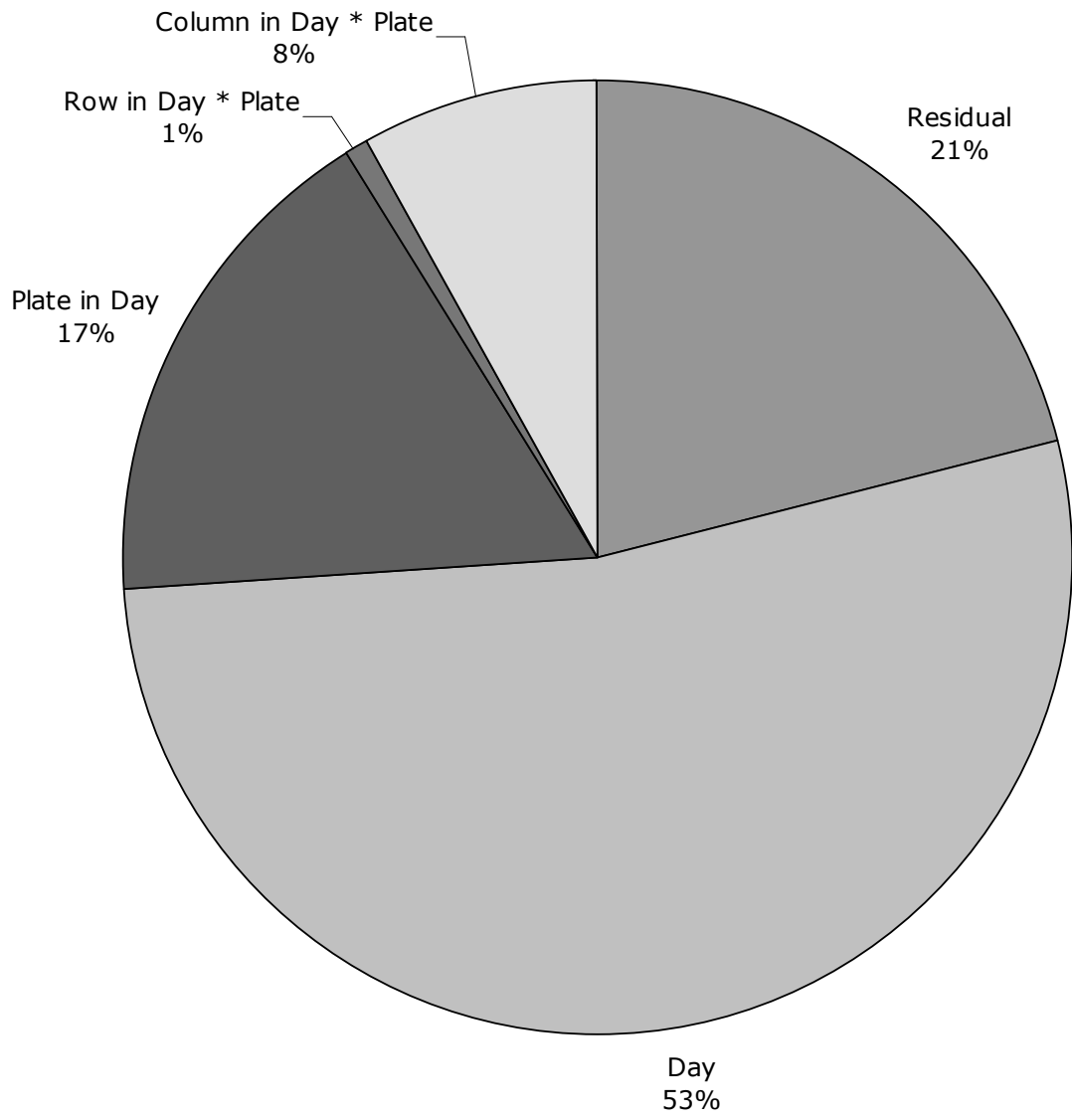
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Figure 4a:



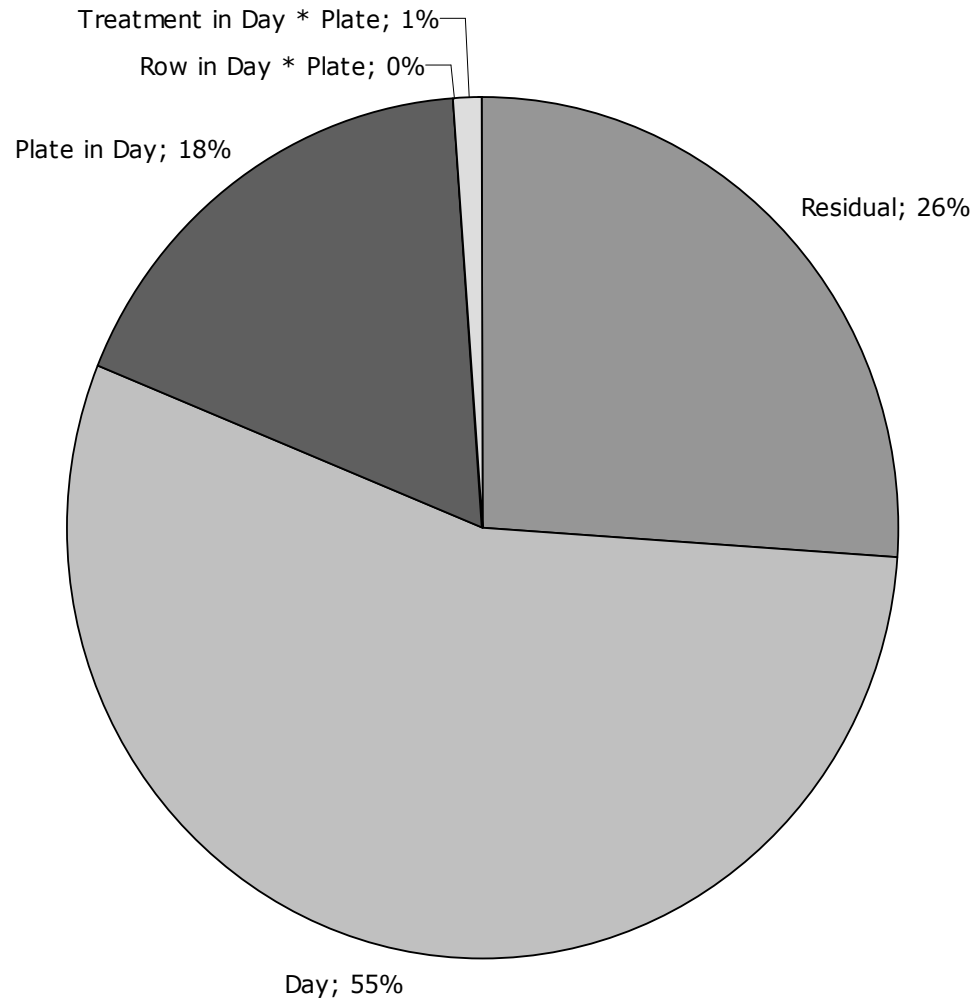
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Figure 4b:



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Figure 4c:



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

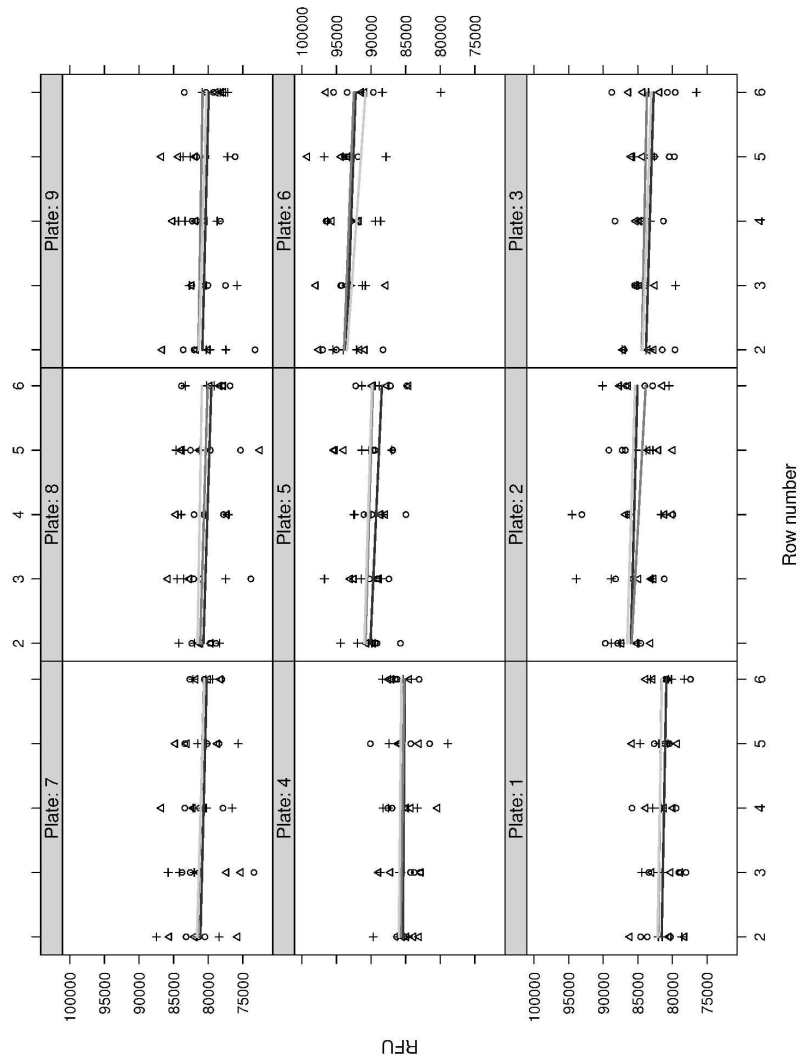


Figure 5: Pseudo linear PLA for detecting possible heterogeneities
215x279mm (600 x 600 DPI)

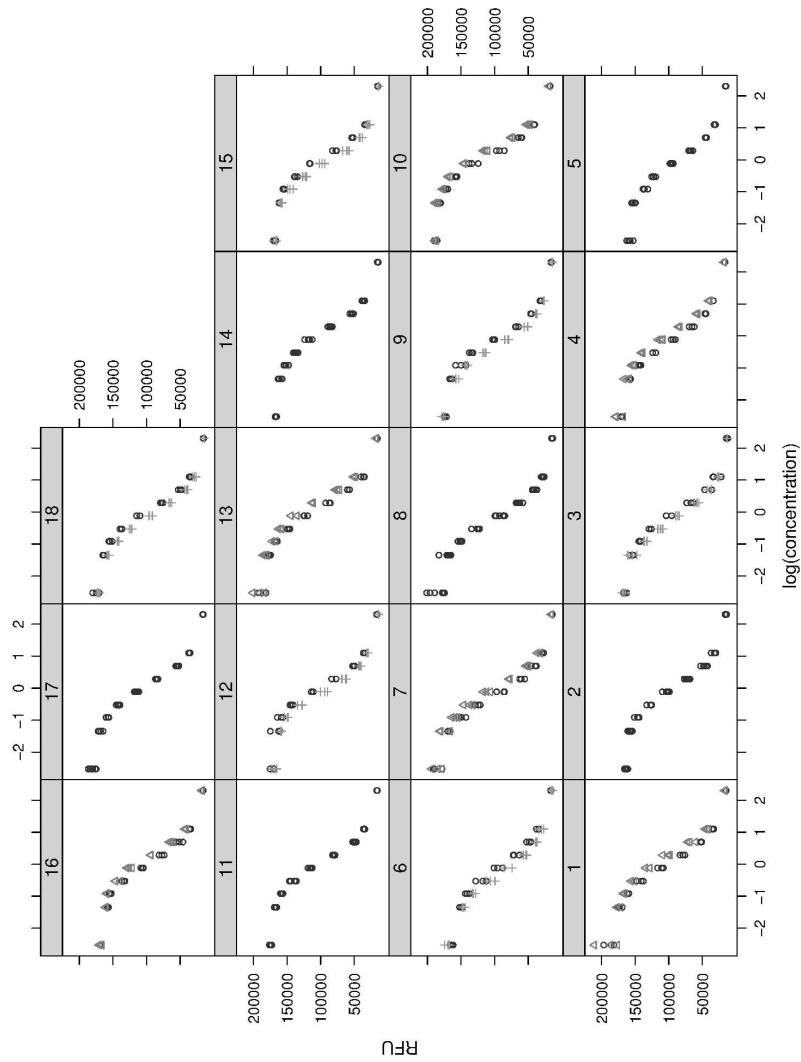


Figure 6: Raw data scatter plots (80%: medium grey, Δ; 100%: dark grey, O; 125%: light grey, +)
215x279mm (600 x 600 DPI)

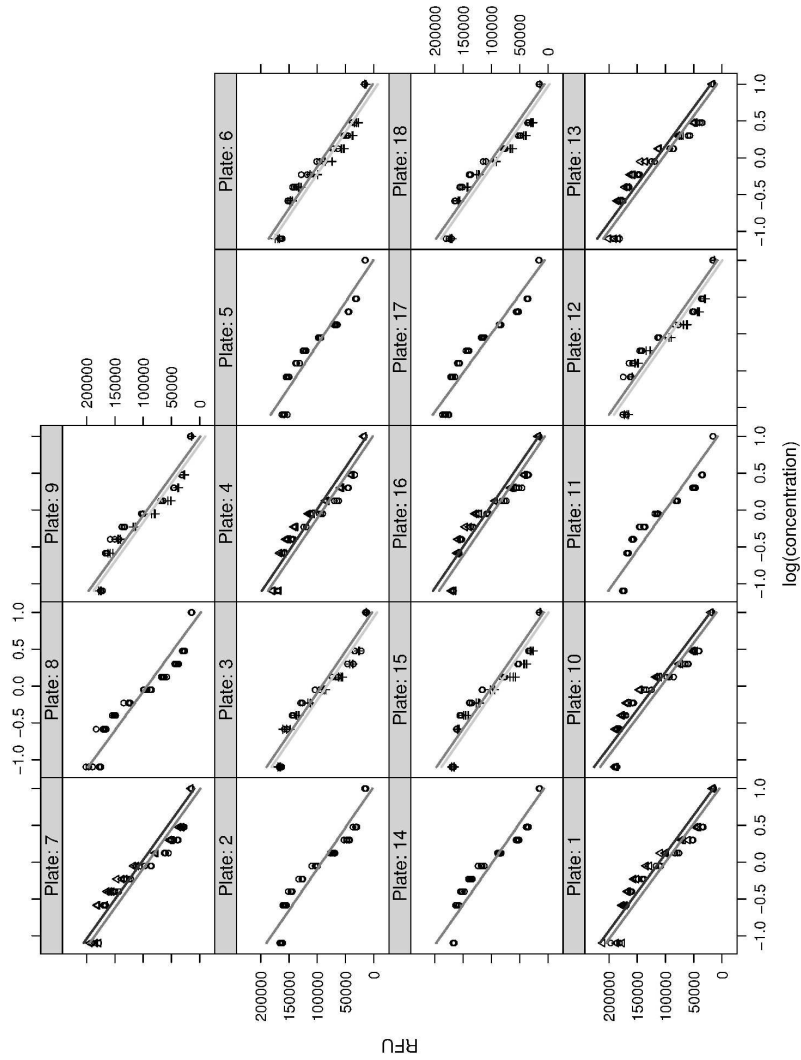


Figure 7: Predictions of the hierarchical linear model separately for each of the 18 plates (80%: medium grey, Δ ; 100%: dark grey, O ; 125%: light grey, $+$) 215x279mm (600 x 600 DPI)

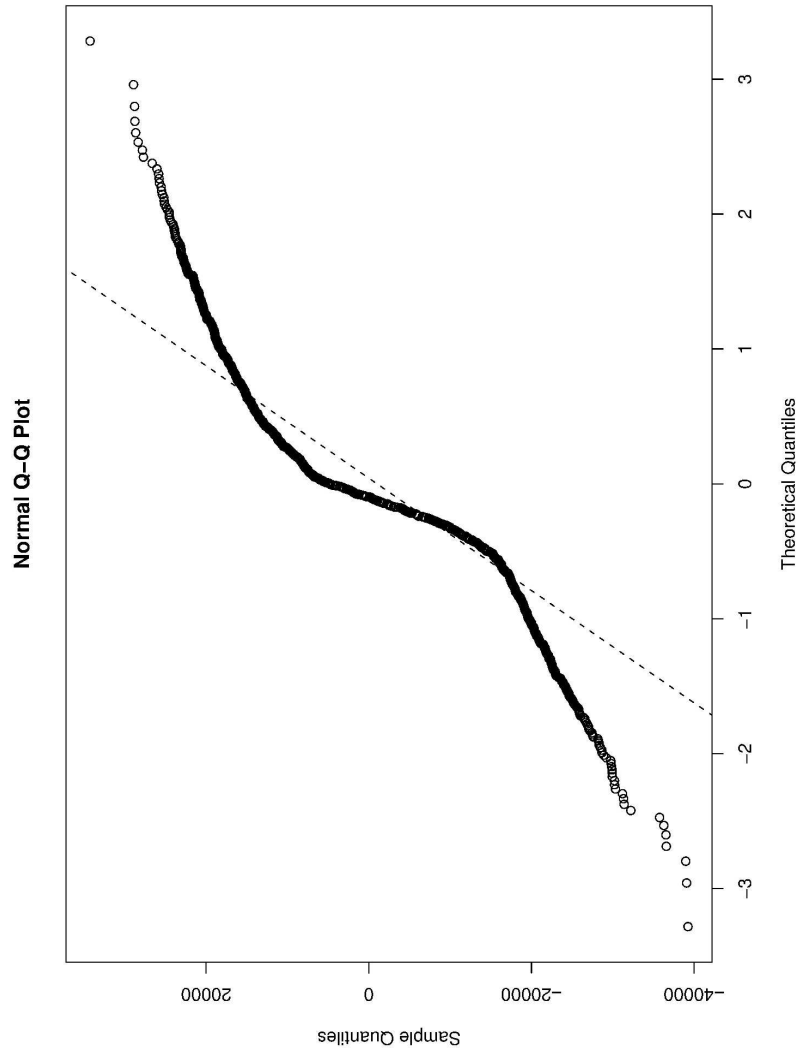


Figure 8a: QQ- and residual plot for the linear mixed model
215x279mm (600 x 600 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

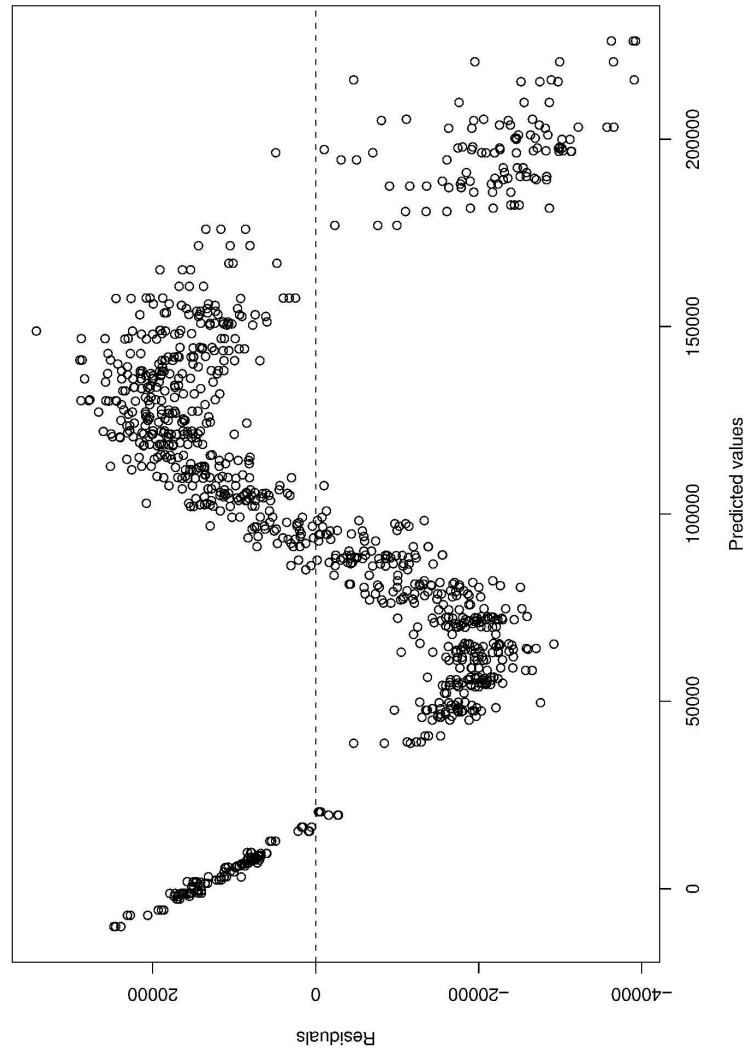


Figure 8b: QQ- and residual plot for the linear mixed model
215x279mm (600 x 600 DPI)

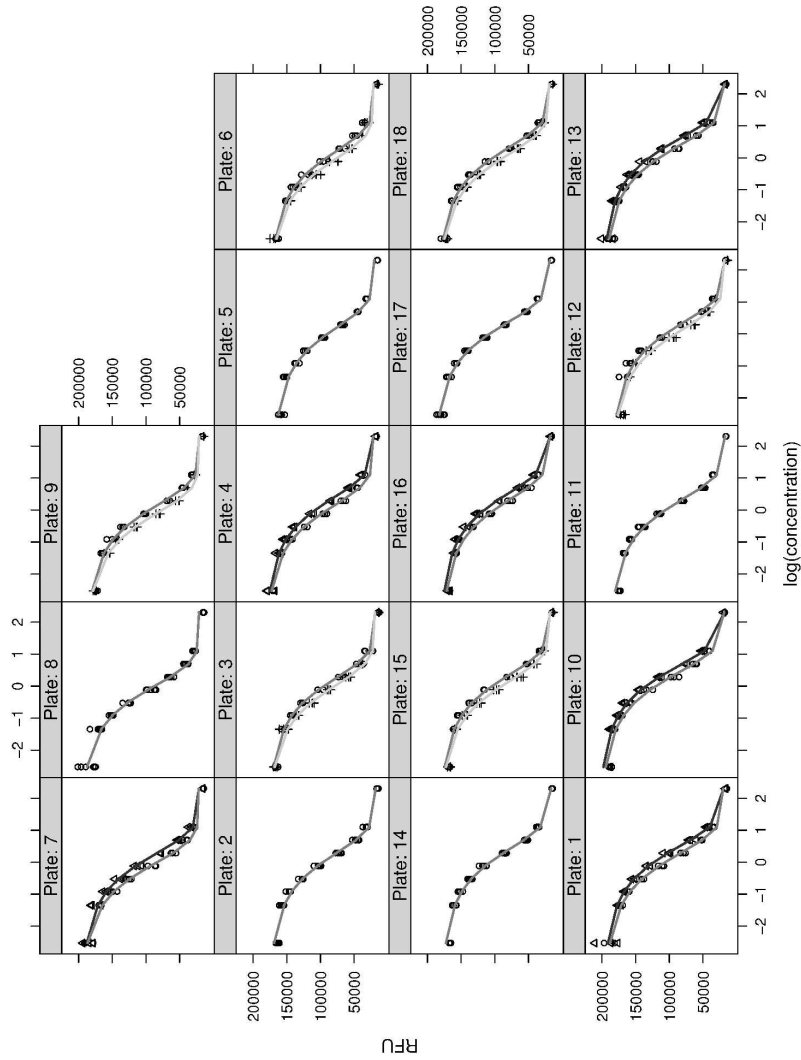


Figure 9: Predictions of the hierarchical 4 parameter log-logistic model separately for each of the 18 plates (80%: medium grey, Δ ; 100%: dark grey, O ; 125%: light grey, $+$) 215x279mm (600 x 600 DPI)

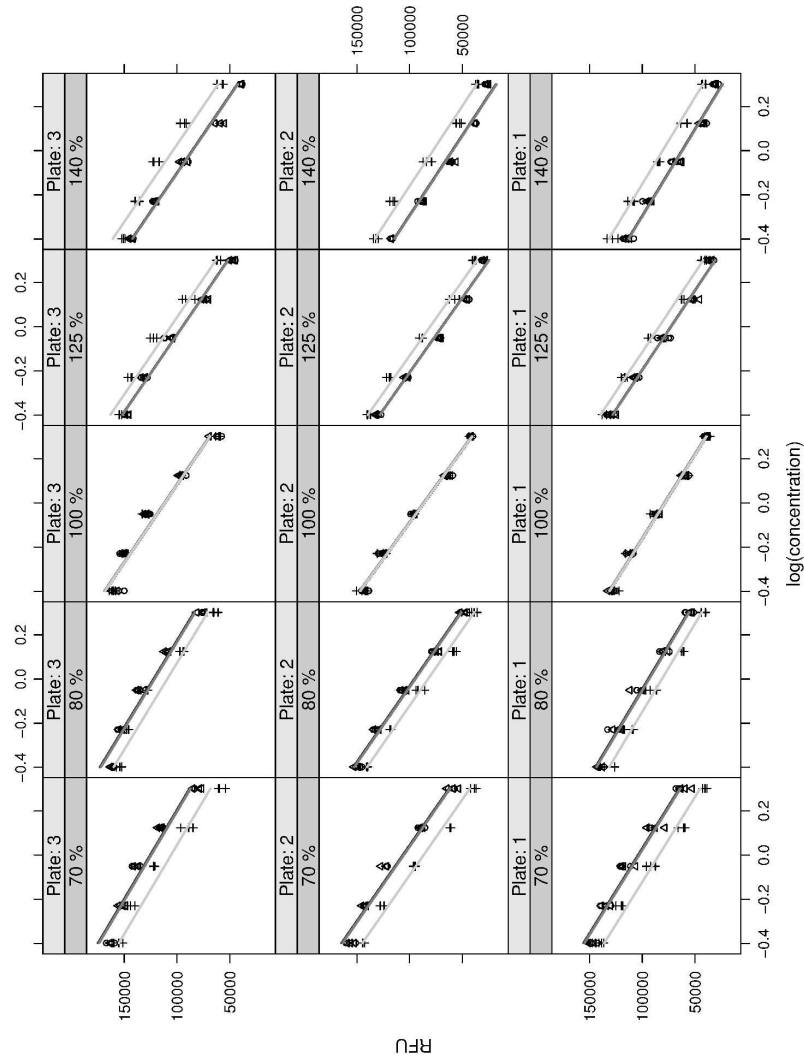


Figure 10: Model fit for the optimized linear parallel line model
215x279mm (600 x 600 DPI)

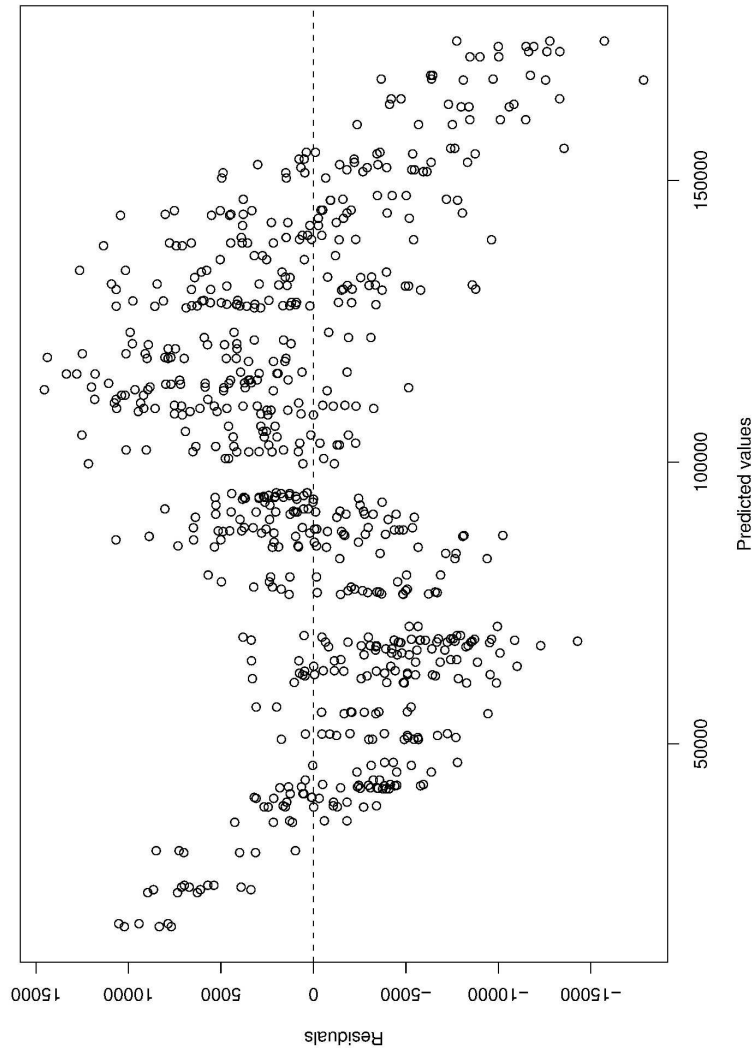


Figure 11: Residual plot for the parallel line model
215x279mm (600 x 600 DPI)

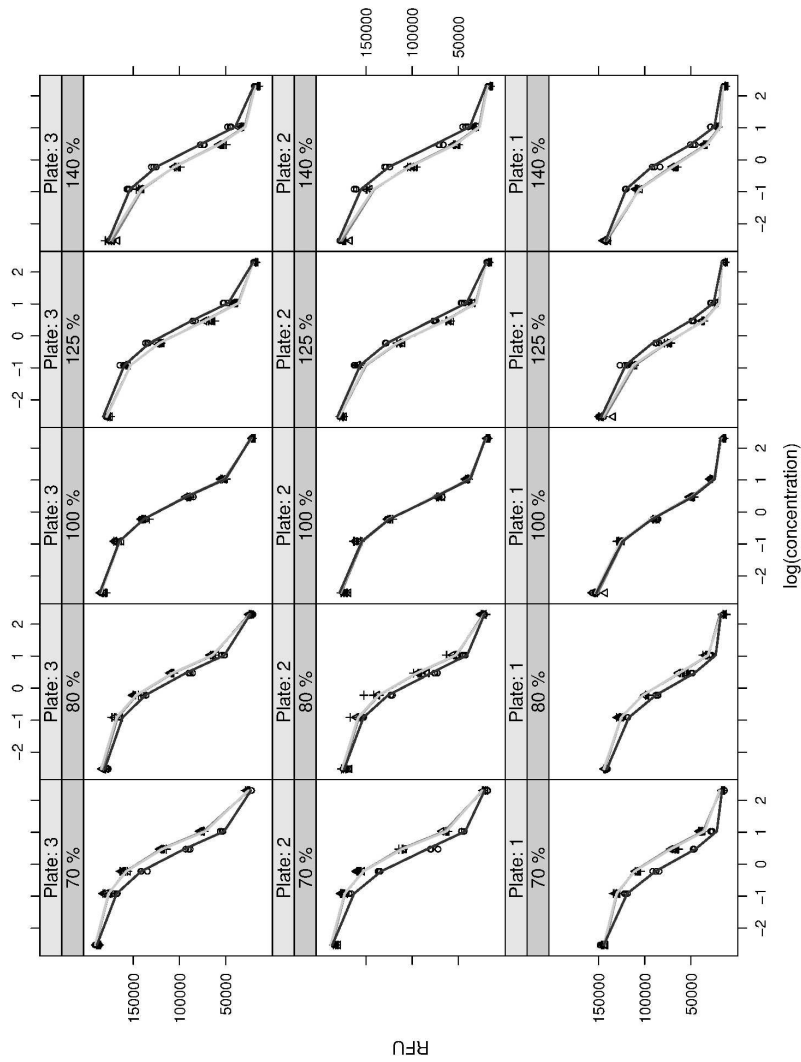


Figure 12: Model fit for the non linear parallel line model
215x279mm (600 x 600 DPI)

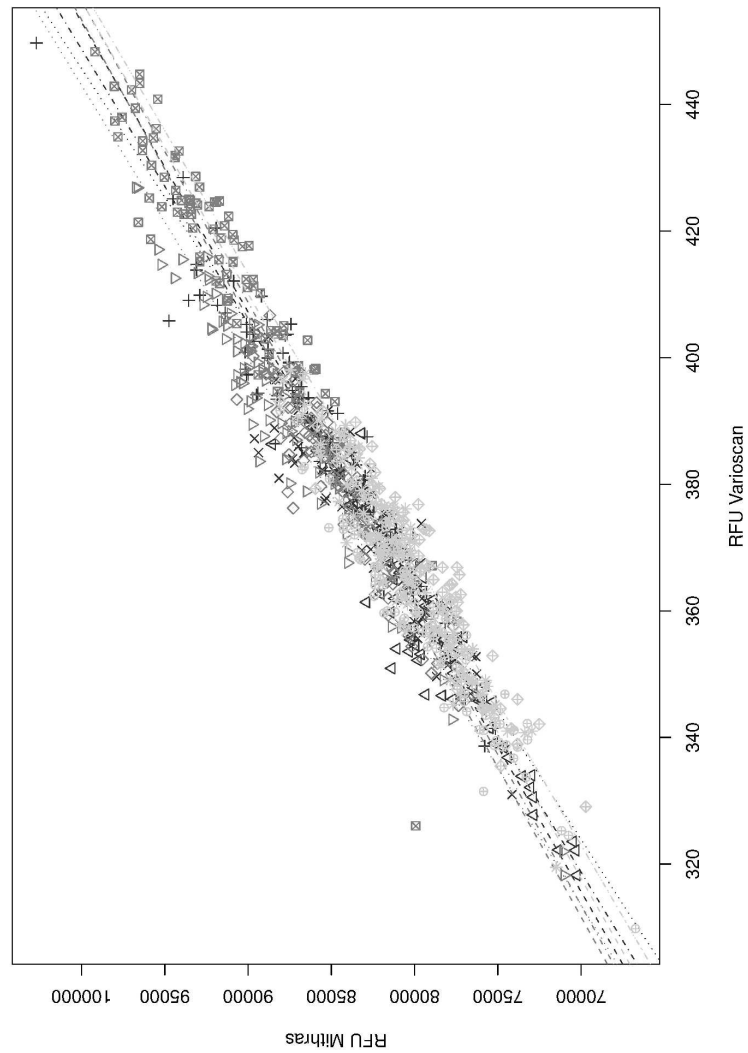


Figure 13: Scatter plot of Mithras LB940 and Varioskan Flash multimode reader RFU measures (different greyscales identify different days, different symbols identify different plates) 215x279mm (600 x 600 DPI)