Drug Information Journal, Vol. 32, pp. 1063–1072, 1998
Printed in the USA.

0092-8615/98

THE PHENYTOIN TRIAL IS A CASE STUDY OF 'INDIVIDUAL' BIOEQUIVALENCE*

R. C. SHUMAKER, PHD
ALPHARMA, Baltimore, Maryland
C. M. METZLER, PHD
Nutwood Associates, Goshen, Indiana

Proposals are being made for changes in bioequivalence criteria, the so-called 'population' and 'individual' bioequivalence methodologies. The proposed criteria require assumptions that have not yet been validated, including the equality of intra-subject variances. More data are needed about the different sources of variation. This type of information was provided by a single dose, two-formulation, four-period trial of phenytoin that used a replicated design. There were no formulation effects, thus each subject provided four observations from which intra-subject variances could be estimated; the individual intra-subject variances were equal within each formulation and were uncorrelated with individual means. This was true on both original and log-transformed scales. Observations of area under the time by concentration curve and maximum observed concentration were analyzed as the total study and within each of the two replications. Bioequivalence inferences were the same for all three analyses, indicating that for this drug 'individual' bioequivalence is not likely to be needed.

Key Words: Phenytoin; Individual bioequivalence

INTRODUCTION

STARTING IN 1990, THE sufficiency of 'average' bioequivalence has been questioned (1,2) and proposals have been made for new bioequivalence criteria that would ensure both 'prescribability' and 'switchability' (2,3,4). A recent issue of the *Journal of Biopharmaceutical Statistics* was largely devoted to this subject, including a report of the Food and Drug Administration Working Group (5). These new criteria have been labeled 'individual' and 'population' bioequivalence. Most of these criteria are based on a function of population parameters of the form

$$\frac{(\mu_T - \mu_R)^2 + C_1\sigma_D^2 + C_2(\sigma_{WT}^2 - \sigma_{WR}^2) \leq \Theta}{\sigma_{WR}^2} \quad (1)$$

where μ_T, μ_R are the test and reference population means,
σ_D^2 is the subject x product interaction,
σ_{WT}^2, σ_{WR}^2 are the within-subject variances of test and reference, and
Θ is the bioequivalence criterion determined by regulatory agencies.

Reprint address: Robert C. Shumaker, ALPHARMA, 333 Cassell Drive, Suite 3500, Baltimore, MD 21224.

*Presented as a poster at the American Society for Clinical Pharmacology and Therapeutics Meeting, March 5–7, 1997, San Diego, California.

This is equation (2) on page 8 of Reference 5, with the exception of the C_1 and C_2 weights which do not appear there. This equation, without the weights, may be written in words as

$$\frac{\text{squared difference of means + interaction + squared different of variances}}{\text{within-subject variance of reference}} \leq K. \quad (2)$$

The use of this criterion requires trial designs which observe each formulation at least twice in at least some of the subjects. These are the so-called 'replicate' designs, which have many advantages over 2×2 designs even for the conventional 'average' bioequivalence studies. Some of these advantages are: less confounding of main effects with interactions, better estimates of variances, and better estimates of carry-over effects.

There are three unresolved issues which should have at least partial resolution before the proposed new bioequivalence criteria are put in place as regulations:

1. Is there a need to change the bioequivalence criteria? There does not seem to be any clinical evidence that the use of 'average' bioequivalence criteria has harmed any patients in the past 25 years. The rationale for changing is, at this point, based on hypothesized possible harm rather than on any real data,
2. What are appropriate values for the weights C_1 and C_2, and for Θ? Is there a scientific knowledge base to inform the choice, or will they be political choices? (As this paper was being revised there is evidence that the weights will no longer be proposed), and
3. The criterion in Equation (1) is a function of population parameters and thus unknown. In practice, sample values would need to be used. How would the criterion perform with sample values substituted for the parameters?

The first issue is a medical issue, and outside the scope of this discussion. Little is known about issues two and three, but there is common agreement that larger studies will be required to obtain acceptable performance (3). To resolve issues two and three more experience with replicate designs is needed, that is, more data. There have been many bioequivalence studies done with these designs but the data are not in the public domain. Thus, one such study is reported for the insight it provides, and in the hope that it will encourage others to report more studies with replicate designs.

A BIOEQUIVALENCE STUDY

A single dose (125 mg), two-formulation, four-period, bioequivalence trial of phenytoin compared the test product, ALPHA-RMA lot PB6198, with the reference product, Parke-Davis Dilantin-125® Lot 14605L. The study used the replicated design:

R	T	T	R
T	R	R	T

where R is the reference product and T is the test product. This design can be considered two replications:

Replicate 1		Replicate 2
RT	and	TR
TR		RT.

The AUC and CMAX parameters observed in this study are listed in the Appendix.

Study Objectives

The objective of this randomized, single-dose, four-way crossover study was to compare the oral bioavailability of the test 125 mg/5 mL phenytoin oral suspension formulation to an equivalent does of the commercially available reference drug in a fasted test population of 26 healthy adult male volunteers.

Methods

The protocol and volunteer consent form were approved by an institutional review board. A total of 26 healthy adult male volunteers were entered into this study. All 26 subjects completed the study in its entirety. The subjects ranged in age from 21 to 47 years (mean = 32.5 years). The subjects' weights ranged from 142 to 207 pounds (mean = 175.8 pounds). The subjects' heights varied from 65 to 74 inches (mean = 69.7 inches). All subjects were within 10% of their desirable height/weight ratio according to the 1983 Metropolitan Insurance Table.

Study Results

Mean concentrations for each product and each replication are shown in Figure 1; the four means are so nearly the same that the lines are almost indistinguishable. As shown in Table 1 of sample statistics, area under the time by concentration curve from dosing to infinity (AUCI) showed neither product not replication differences; the differences in CMAX were small.

STATISTICAL ANALYSIS

Statistical Methods

Analyses of variances were computed for AUCI and CMAX on both the original and log-transformed scales; these ANOVAs were computed for the total trial and for each replication. The ANOVAs were computed with SAS PROC MIXED using the following model statements:

```
PROC MIXED METHOD = REML
DATA = DATA1;

CLASS SUB SEQ PERIOD FORM;

MODEL (PK parameter) = SEQ
SUB(SEQ) PERIOD FORM
FORM*SUB(SEQ) / CHISQ;

RANDOM SUB(SEQ) / TYPE = CS
SUBJECT = SUB;

LSMEANS FORM / ADJUST = BON
PDIFF ALPHA = 0.1;
```

Those familiar with SAS PROC MIXED will recognize that these statements define the statistical model used. In the usual convention of bioequivalence studies, FORM is formulation; SUB is subject; and SEQ is sequence.

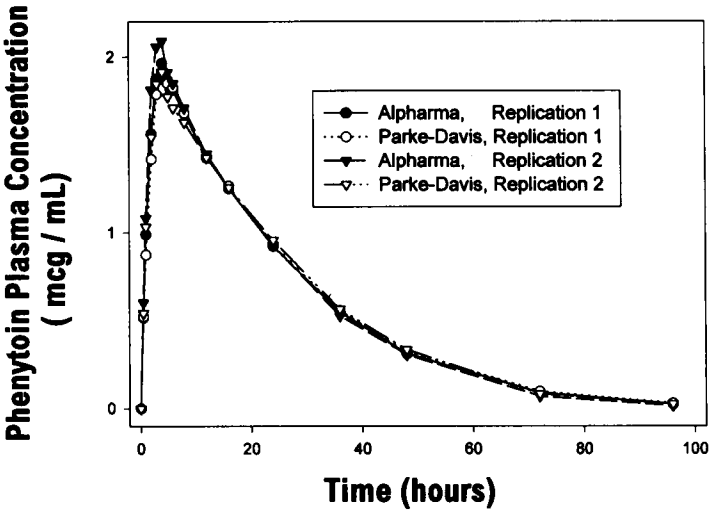

FIGURE 1. Mean phenytoin concentrations by product and replication.

TABLE 1
Sample Statistics

	Product Means		Standard deviation		CV(%)	
Product, Replication	AUCI	CMAX	AUCI	CMAX	AUCI	CMAX
Reference, Replication 1	55.72	1.994	18.52	0.399	33.2	20.0
Reference, Replication 2	55.54	1.948	15.47	0.350	27.9	18.0
Test, Replication 1	55.09	2.077	16.68	0.426	30.3	20.5
Test, Replication 2	55.02	2.182	16.86	0.409	30.6	18.8

For all analyses the pharmacokinetic parameters were AUCI, ln(AUCI), CMAX, and ln(CMAX). Conventional 90% confidence intervals for deciding bioequivalence were computed with the above model for the total trial and for each replication. For each subject the four residuals from the subject's mean AUC were plotted. A mixed model ANOVA with sequence and subject effects was used to estimate pooled between- and within-subject variances; the model without sequence effects was used to estimate the variances within each formulation.

Statistical Results

For all ANOVAs the F-statistic for the formation by subject interaction term was less than 1.00; thus the interaction term is not reported in the tables. The analyses are summarized in Table 2, and the confidence interval estimates in Table 3. The estimates of within- and between-subject variances are reported in Table 4. Since the variance estimates in the two formulations are not independent (same subjects) the F-test for equality of variances may not be appropriate. But both "F-ratios" for within-subject variances in Table 4 are less than 1.6, suggesting equality.

The graphs and tables indicate that:

1. There was no formulation difference in AUCI in this study and only a small difference in CMAX. The statistical results can be anticipated by seeing the closeness of the mean curves in Figure 1. Figures 2 and 3, which plot test and reference AUCI and CMAX by replication, show that the AUCI are almost identical, while the slight shift of CMAX above the diagonal line of

TABLE 2
Summary of Analyses of Variance

		p-levels			
Replication	Variable	Sequence	Formulation	Period	Error CV%
1	AUCI	0.152	0.356	0.002	4.34
	CMAX	0.497	0.168	0.032	10.36
2	AUCI	0.120	0.550	0.173	5.49
	CMAX	0.229	0.003	0.245	12.41
BOTH	AUCI	0.134	0.365	0.025	5.45
	CMAX	0.326	0.002	0.079	11.80
1	LOGAUCI	0.175	0.777	0.002	1.22
	LOGCMAX	0.478	0.176	0.036	14.29
2	LOGAUCI	0.116	0.316	0.144	1.40
	LOGCMAX	0.306	0.002	0.322	16.96
BOTH	LOGAUCI	0.142	0.385	0.026	1.40
	LOGCMAX	0.358	0.002	0.120	16.46

TABLE 3
Summary of Confidence Interval Estimates

Replication	BE Parameter	Means			Confidence Intervals	
		Reference	Test	Test-Reference	90%	% Reference mean
1	AUCI	55.717	55.090	−0.627	(−1.77, 0.51)	(−3.21%, 0.93%)
	CMAX	1.994	2.077	0.083	(−0.02, 0.18)	(−0.81%, 8.81%)
2	AUCI	55.532	55.022	−0.510	(−1.95, 0.93)	(−3.54%, 1.69%)
	CMAX	1.948	2.182	0.235	(0.11, 0.36)	(5.18%, 16.3%)
BOTH	AUCI	55.620	55.056	−0.568	(−1.61, 0.47)	(−2.93%, 0.86%)
	CMAX	1.971	2.130	0.159	(0.08, 0.24)	(3.65%, 11.3%)
1	LOGAUCI	3.973	3.969	−0.001	(−0.03, 0.02)	(−2.63%, 1.93%)
	LOGCMAX	0.672	0.712	0.038	(−0.01, 0.08)	(−0.87%, 8.88%)
2	LOGAUCI	3.983	3.674	−0.309	(−0.04, 0.01)	(−4.15%, 1.07%)
	LOGCMAX	0.651	0.784	0.113	(0.06, 0.17)	(5.77%, 18.5%)
BOTH	LOGAUCI	3.978	3.698	−0.010	(−0.03, 0.01)	(−2.83%, 0.90%)
	LOGCMAX	0.662	0.737	0.076	(0.04, 0.11)	(3.66%, 12.2%)

equality indicates that the test CMAX were a little larger. Figure 3 also indicates that the within-subject variance of CMAX was greater than that of AUCI. (In Figures 2 and 3 variation along the diagonal line is between-subject variation, and variation perpendicular to the diagonal represents within-subject variation.),

2. The variance in the data is so small that the study was overpowered, that is, it was much larger than necessary. This is reflected in the very short confidence intervals for AUCI in Table 3. It is also reflected by the fact that although the ANOVAs for CMAX show some statistically significant differences, all of the confidence intervals are well within the criterion for bioequivalence,
3. There was no difference in within-subject variances between the two products, and

TABLE 4
Estimation of Variances (by PROC VARCOMP)

Within Formulations

Formulation = Reference

scale	between	within	within/between
original	260.06	9.062	3.48%
log	0.07370	0.00299	4.06%

Formulation = Test

scale	between	within	within/between	within Test/ Reference
original	259.71	13.400	5.16%	1.479
log	0.07115	0.00454	6.38%	1.518

Total study

scale	sequence	between	within	within/between
original	28.551	261.27	9.889	3.78%
log	0.00739	0.07278	0.00342	4.70%

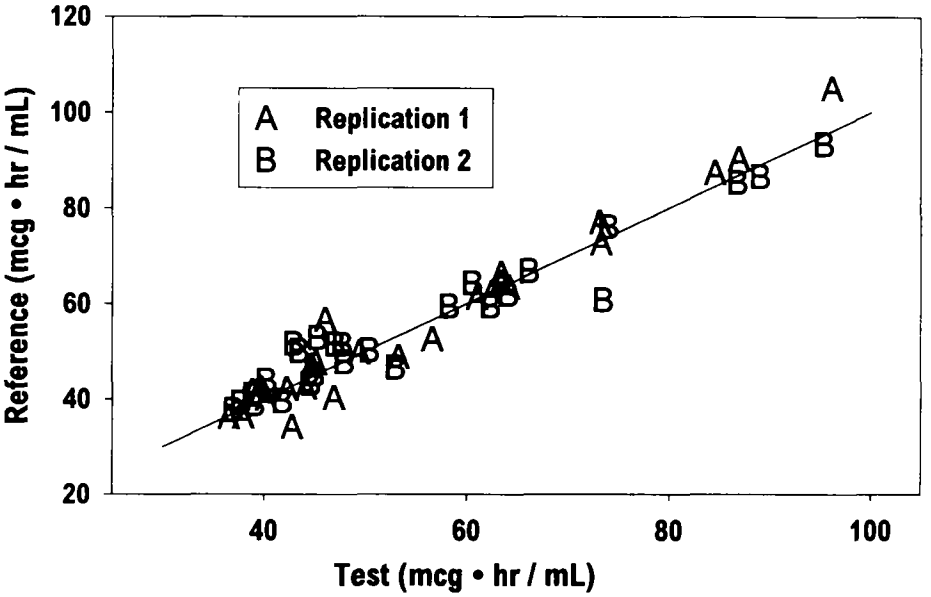

FIGURE 2. Reference versus test AUCI, by replication.

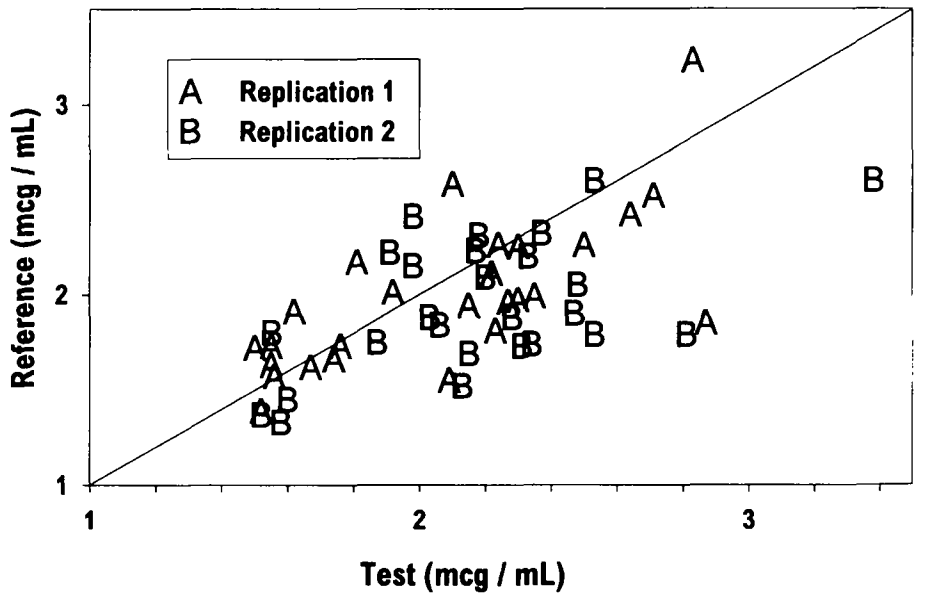

FIGURE 3. Reference versus test CMAX, by replication.

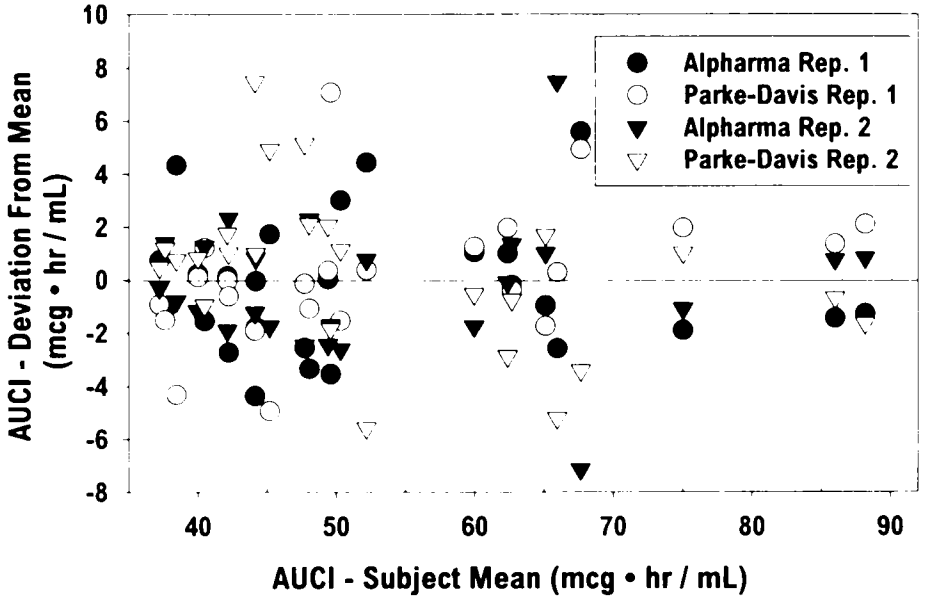

FIGURE 4. Deviation from subject mean AUCI versus subject mean, by subject (original scale).

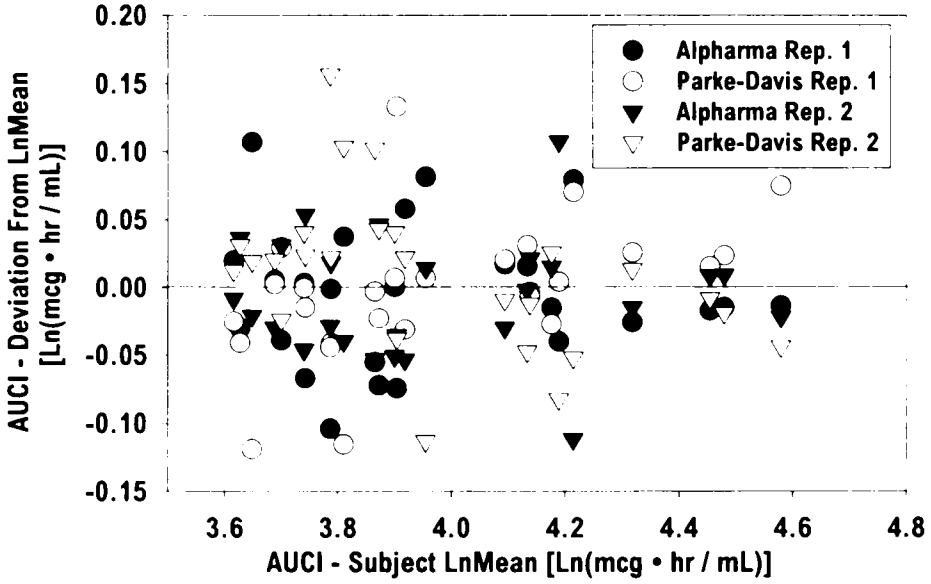

FIGURE 5. Deviation from subject mean AUCI versus subject mean, by subject (log scale).

there is no indication that the subjects do not all have the same error distribution. Figures 4 and 5 show no indication of trend in the deviations of each subject from the subject mean. Table 4 indicates that the variances are not different in the two formulations.

4. No differences in deciding bioequivalence were made from the replicated study that would not have been made had only one of the replicates been run. Table 3 shows that with either original or log-scaled AUC and CMAX, and in either replication, the decision is the same: The test product is equivalent to the reference product, and
5. As is the case for many drugs, the between-subject variance is much larger than the within-subject variance; Table 4 shows that with one exception the within-subject variance is less than 5% of the between-subject variance.

Since there was so little difference between the two products, and the estimate of the interaction term is zero, the proposed "individual" bioequivalence criteria were not computed.

CONCLUSIONS

Analysis of this replicated bioequivalence study of phenytoin shows that for this drug a test formulation can be produced that is equal to the reference formulation in both average bioavailability and variance of bioavailability. The ALPHARMA phenytoin suspension is bioequivalent to the Parke-Davis phenytoin suspension. The study also shows that there is no indication that subjects have different within-subject variances and thus there is no need for "individual" bioequivalence criteria to be applied.

REFERENCES

1. Anderson S, Hauck WW. Consideration of individual bioequivalence. *J Pharmacokin Biopharm.* 1990; 18(3):259–273.
2. Hauck WW, Anderson S. Measuring switchability and prescribability: When is average bioequivalence sufficient? *J Pharmacokin Biopharm.* 1994;22(6):551–564.
3. Schall R, Williams RL. Towards a practical strategy for assessing individual bioequivalence. *J Pharmacokin Biopharm.* 1996;24(1):133–149.
4. Vuorinen J, Turunen J. A simple three-step procedure for parametric and nonparametric assessment of bioequivalence. *Drug Inf J.* 1997;31(1):167–180.
5. Chen L. Individual bioequivalence—A regulatory update. *J Biopharm Statist.* 1997;7(1):5–11.

APPENDIX
Data Listing

SUB	SEQ	PERIOD	PROD	REP	AUCI	CMAX
1	BAAB	1	REF	1	36.270	1.63
1	BAAB	2	TEST	1	37.945	1.55
1	BAAB	3	TEST	2	36.893	2.20
1	BAAB	4	REF	2	37.663	2.09
2	BAAB	1	REF	1	49.748	2.26
2	BAAB	2	TEST	1	49.422	2.50
2	BAAB	3	TEST	2	46.957	1.98
2	BAAB	4	REF	2	51.451	2.41
3	ABBA	1	TEST	1	73.241	1.50
3	ABBA	2	REF	1	72.581	1.72
3	ABBA	3	REF	2	64.232	1.37
3	ABBA	4	TEST	2	60.498	1.52
4	ABBA	1	TEST	1	46.055	2.64
4	ABBA	2	REF	1	56.667	2.42
4	ABBA	3	REF	2	47.760	2.20
4	ABBA	4	TEST	2	47.877	2.33
5	BAAB	1	REF	1	52.544	1.97
5	BAAB	2	TEST	1	56.609	2.30
5	BAAB	3	TEST	2	52.930	2.28
5	BAAB	4	REF	2	46.601	1.88
6	ABBA	1	TEST	1	38.908	1.62
6	ABBA	2	REF	1	41.667	1.91
6	ABBA	3	REF	2	39.495	1.69
6	ABBA	4	TEST	2	41.725	2.15
7	ABBA	1	TEST	1	86.878	2.83
7	ABBA	2	REF	1	90.262	3.23
7	ABBA	3	REF	2	86.512	2.60
7	ABBA	4	TEST	2	88.976	3.38
8	BAAB	1	REF	1	34.117	1.85
8	BAAB	2	TEST	1	42.762	2.87
8	BAAB	3	TEST	2	37.629	2.18
8	BAAB	4	REF	2	39.175	2.31
9	BAAB	1	REF	1	48.768	2.17
9	BAAB	2	TEST	1	53.306	1.81
9	BAAB	3	TEST	2	47.685	1.91
9	BAAB	4	REF	2	51.427	2.22
10	ABBA	1	TEST	1	73.128	1.74
10	ABBA	2	REF	1	76.991	1.66
10	ABBA	3	REF	2	76.034	2.23
10	ABBA	4	TEST	2	73.926	2.17
11	ABBA	1	TEST	1	44.711	2.30
11	ABBA	2	REF	1	46.971	2.25
11	ABBA	3	REF	2	50.170	2.60
11	ABBA	4	TEST	2	50.340	2.53
12	BAAB	1	REF	1	87.344	1.99
12	BAAB	2	TEST	1	84.546	2.35
12	BAAB	3	TEST	2	86.733	2.31
12	BAAB	4	REF	2	85.301	1.73
13	ABBA	1	TEST	1	64.142	1.92
13	ABBA	2	REF	1	63.369	2.01
13	ABBA	3	REF	2	66.799	1.84
13	ABBA	4	TEST	2	66.111	2.06
14	BAAB	1	REF	1	62.276	1.81
14	BAAB	2	TEST	1	62.455	2.23

14	BAAB	3	TEST	2	64.007	1.98
14	BAAB	4	REF	2	61.904	2.15
15	BAAB	1	REF	1	40.069	1.57
15	BAAB	2	TEST	1	40.196	1.56
15	BAAB	3	TEST	2	38.802	1.55
15	BAAB	4	REF	2	40.800	1.80
16	ABBA	1	TEST	1	63.364	2.10
16	ABBA	2	REF	1	64.339	2.58
16	ABBA	3	REF	2	59.492	2.32
16	ABBA	4	TEST	2	62.267	2.37
17	BAAB	1	REF	1	42.097	1.39
17	BAAB	2	TEST	1	42.258	1.52
17	BAAB	3	TEST	2	40.217	1.58
17	BAAB	4	REF	2	43.867	1.33
18	ABBA	1	TEST	1	63.379	2.22
18	ABBA	2	REF	1	66.237	2.11
18	ABBA	3	REF	2	60.755	1.79
18	ABBA	4	TEST	2	73.428	2.81
19	BAAB	1	REF	1	61.217	2.52
19	BAAB	2	TEST	1	61.003	2.71
19	BAAB	3	TEST	2	58.207	2.47
19	BAAB	4	REF	2	59.418	1.90
20	BAAB	1	REF	1	42.384	1.94
20	BAAB	2	TEST	1	44.115	2.15
20	BAAB	3	TEST	2	44.978	2.34
20	BAAB	4	REF	2	45.157	1.74
21	ABBA	1	TEST	1	36.508	2.27
21	ABBA	2	REF	1	36.106	1.96
21	ABBA	3	REF	2	38.804	1.88
21	ABBA	4	TEST	2	39.017	2.03
22	ABBA	1	TEST	1	96.140	2.24
22	ABBA	2	REF	1	105.022	2.26
22	ABBA	3	REF	2	93.275	1.79
22	ABBA	4	TEST	2	95.308	2.53
23	ABBA	1	TEST	1	39.469	1.67
23	ABBA	2	REF	1	41.583	1.62
23	ABBA	3	REF	2	43.212	2.05
23	ABBA	4	TEST	2	44.520	2.48
24	BAAB	1	REF	1	40.251	1.55
24	BAAB	2	TEST	1	46.900	2.09
24	BAAB	3	TEST	2	43.421	1.87
24	BAAB	4	REF	2	50.101	1.75
25	BAAB	1	REF	1	47.567	1.73
25	BAAB	2	TEST	1	45.156	1.55
25	BAAB	3	TEST	2	45.245	2.13
25	BAAB	4	REF	2	52.852	1.52
26	ABBA	1	TEST	1	39.759	1.76
26	ABBA	2	REF	1	42.206	1.73
26	ABBA	3	REF	2	51.582	1.45
26	ABBA	4	TEST	2	42.872	1.60