Statistical Papers 45, 51-66 (2004)

Statistical Papers

A cautionary note on the analysis of randomized block designs with a few missing values

Devan V. Mehrotra

Merck Research Labs., UM-A102, 785 Jolly Rd., Bldg. C,
Blue Bell, PA 19422, USA

Received: December 6, 2001; revised version: June 28, 2002

The randomized block design is routinely employed in the social and biopharmaceutical sciences. With no missing values, analysis of variance (AOV) can be used to analyze such experiments. However, if some data are missing, the AOV formulae are no longer applicable, and iterative methods such as restricted maximum likelihood (REML) are recommended, assuming block effects are treated as random. Despite the well-known advantages of REML, methods like AOV based on complete cases (blocks) only (CC-AOV) continue to be used by researchers, particularly in situations where routinely only a few missing values are encountered. Reasons for this appear to include a natural proclivity for non-iterative, summary-statistic-based methods, and a presumption that CC-AOV is only trivially less efficient than REML with only a few missing values (say $\leq 10\%$). The purpose of this note is two-fold. First, to caution that CC-AOV can be considerably less powerful than REML even with only a few missing values. Second, to offer a summary-statistic-based, pairwise-available-case-estimation (PACE) alternative to CC-AOV. PACE, which is identical to AOV (and REML) with no missing values, outperforms CC-AOV in terms of statistical power. However, it is recommended in lieu of REML *only* if software to implement the latter is unavailable, or the use of a "transparent" formula-based approach is deemed necessary. An example using real data is provided for illustration.

Key words: analysis of variance; linear mixed model; restricted maximum likelihood; Satterthwaite approximation.

1. Introduction

Randomized block experiments are routinely conducted in the social and biopharmaceutical sciences, and it is not uncommon for some data to be missing in such studies. To be clear, missing data refer to those values that were planned to be made available for analysis but could not be obtained due to factors beyond experimental control. For example, the data in Table I (from May and Johnson, 1995) are from an experiment in which blood was drawn from 40 subjects (blocks), and each sample was divided into four equal parts. Coagulation times were recorded after each part was randomly assigned to be treated with one of four experimental drugs. Note that 19 of the 40 subjects have at least one missing value; reasons for missingness were not provided by the authors. We will assume throughout this paper that missing values, such as those in Table I, are missing completely at random, in the sense formalized by Little and Rubin (1987, chapter 1). We will also assume that block

Table 1. Coagulation times for four experimental drugs (T1-T4)

Subject	T1	T2	T3	T4
1	1.24	2.11	1.19	1.63
2	.	1.50	1.67	2.24
3	1.02	1.18	1.50	1.63
4	1.48	1.84	1.85	2.07
5	.	1.58	1.66	1.42
6	1.14	1.47	1.78	1.88
7	1.67	1.99	1.88	2.07
8	1.06	.	1.57	.
9	.	1.72	2.08	2.00
10	1.28	1.72	1.82	1.41
11	1.42	1.73	1.80	2.25
12	1.24	1.34	1.39	1.98
13	1.51	.	1.58	1.95
14	.	1.47	2.11	1.65
15	1.44	1.71	.	2.09
16	1.89	2.42	1.81	2.14
17	1.30	1.91	1.29	2.02
18	1.63	1.79	2.12	1.87
19	1.67	2.14	2.09	2.09
20	1.36	1.61	1.60	1.82
21	1.35	2.04	1.61	.
22	0.86	1.37	.	2.02
23	.	1.88	1.92	2.00
24	1.47	1.52	1.97	.
25	1.69	1.57	.	2.08
26	1.63	1.39	1.78	.
27	1.13	1.44	1.27	1.41
28	.	1.78	1.65	1.75
29	1.28	1.72	1.47	1.93
30	1.22	0.74	1.51	1.35
31	.	.	1.34	1.59
32	1.43	2.46	2.08	2.24
33	1.29	1.64	.	1.69
34	.	1.07	1.36	1.37
35	1.11	1.25	1.45	1.82
36	1.50	.	1.38	1.71
37	1.47	.	.	1.79
38	1.84	1.97	1.88	2.29
39	1.83	2.35	2.07	1.98
40	1.58	1.92	1.85	1.61

Data abstracted from May and Johnson (1995); "." = missing value.

effects are random (instead of fixed), implying that responses within a block are expected to be correlated, hence requiring the use of a mixed effects model.

With no missing values, the analysis is straightforward due to the existence of a summary-statistic-based approach, namely analysis of variance (AOV). However, if some data are missing, the intuitive appeal and simplicity of this traditional approach are *presumed* lost because the usual AOV formulae for inference and estimation are no longer applicable. Accordingly, the missing data problem is commonly tackled using non-summary-statistic-based methods such as restricted maximum likelihood (REML, Patterson and Thompson, 1971; Corbeil and Searle, 1976). These iterative methods have gained in popularity because they have attractive theoretical properties, and are easily implemented using commercial software packages such as SAS PROC MIXED.

It is well-known that REML makes excellent use of all the available data by cleverly combining interblock and intrablock information to estimate treatment differences (e.g., Hocking, 1985, chapter 9). This note was motivated by the authors observation that, despite the well-known advantages of REML, methods like AOV based on complete cases (blocks) only (CC-AOV) continue to be used by some researchers, especially in situations where routinely only a few missing values (if any) are encountered, such as in preclinical or phase I clinical trials. Reasons for using CC-AOV include a natural proclivity for non-iterative, summary-statistic-based methods, a presumption that CC-AOV is only trivially less efficient than REML with only a few missing values (say $\leq$ 10%), and/or lack of access to software such as SAS PROC MIXED.

The purpose of this note is two-fold. First, to illustrate that REML is much more powerful than CC-AOV even with only a few missing values. Second, to offer a summary-statistic-based, pairwise-available-case-estimation (PACE) alternative to CC-AOV. Section 2 contains a review of the standard analysis with no missing values. In Section 3, we describe the summary-statistic-based PACE method, including the test statistic for comparing treatments, a derivation of its approximate null distribution, and formulae for variance component estimation. Analysis of the data in Table I based on CC-AOV, PACE and REML is provided in Section 4. Simulation results comparing PACE with CC-AOV and REML are described in Section 5, followed by concluding remarks in Section 6.

2. Standard Analysis With No Missing Values

The following traditional mixed effects model is assumed:

$$y_{ij} = \mu_i + \beta_j + \varepsilon_{ij},\ i \leq a,\ j \leq n, \tag{2.1}$$

where y_{ij} is the response for treatment i in block j, μ_i is the population mean for treatment i, and β_j and ε_{ij} are random block and residual effects, assumed to be independently distributed normal variates with zero means and variances ϕ_1 and ϕ_0, respectively. ϕ_1 and ϕ_0 are commonly referred to as the interblock and intrablock variance components, respectively. Implicit in (2.1) is the following compound symmetry covariance assumption:

$$\begin{aligned} Cov(y_{ij}, y_{kl}) &= \phi_1 \qquad i \neq k, j = l \\ &= \phi_1 + \phi_0 \quad i = k, j = l \\ &= 0 \qquad j \neq l\,. \end{aligned} \tag{2.2}$$

From (2.2), it can be inferred that responses between blocks are assumed to be uncorrelated, but responses within each block share a common intrablock correlation of $\rho = \dfrac{\phi_1}{\phi_1 + \phi_0}$.

For now, let us assume that there are no missing values. Throughout this paper, a bar indicates a mean and a dot replaces the subscript over which the mean is taken. Let $\bar{d}_{ik(.)}$ and V_{ik} denote, respectively, the sample mean and variance of the normalized pairwise contrast $d_{ik(j)} = \dfrac{(y_{ij} - y_{kj})}{\sqrt{2}}$, $i<k$, and let C_{ik} denote the sample covariance between observations for treatments i and k. Hence, $V_{ik} = (n-1)^{-1}\sum_j \left(d_{ik(j)} - \bar{d}_{ik(.)}\right)^2$ and $C_{ik} = (n-1)^{-1}\sum_j (y_{ij} - \bar{y}_{i.})(y_{kj} - \bar{y}_{k.})$. It is easily seen that under (2.2), V_{ik} and C_{ik} are unbiased estimators of ϕ_0 and ϕ_1, respectively. Hence, an intuitively appealing approach for estimating the variance components is to take an average of all the pairwise contrast variances, V_{ik}, and all the pairwise covariances C_{ik}, for estimating ϕ_0 and ϕ_1, respectively. Hocking (1985, chapter 10) made the astute observation that the AOV estimators of the variance components, obtained by equating the expected block and residual mean squares to their corresponding observed values, are in fact identical to the intuitive estimators based on averages of sample variances and covariances. In other words, the AOV estimators can be expressed as

$$\hat{\phi}_1^{AOV} = \frac{\sum_{i<k}(n-1)C_{ik}}{\sum_{i<k}(n-1)} \tag{2.3}$$

and

$$\hat{\phi}_0^{AOV} = \frac{\sum_{i<k}(n-1)V_{ik}}{\sum_{i<k}(n-1)}. \tag{2.4}$$

Even though the (n-1) terms cancel out in (2.3) and (2.4), they have been retained for ease of presentation later.

Next, if $100(1-\alpha)\%$ confidence intervals need to be constructed for an individual treatment mean, or for a difference between a pair of means, the following well-known formulae are available:

$$\bar{y}_{i.} \pm t_f^{\alpha/2}\sqrt{\frac{\hat{\phi}_0^{AOV} + \hat{\phi}_1^{AOV}}{n}} \tag{2.5}$$

and

$$\left(\bar{y}_{i.}-\bar{y}_{k.}\right)\pm t^{\alpha/2}_{(a-1)(n-1)}\sqrt{\frac{2\hat{\phi}_0^{AOV}}{n}}. \tag{2.6}$$

Above, $t_{\nu}^{\alpha/2}$ denotes the $100(1-\alpha/2)^{\text{th}}$ percentile of a central t distribution with ν degrees of freedom. Under (2.2), the degrees of freedom for use in (2.5) are usually estimated using a Satterthwaite (1941) approximation, leading to $f=\frac{f_2(1+f_1)^2(1+\lambda)^2}{f_1\{1+\lambda(1+f_1)\}^2+f_1^2}$, where $f_1=(a-1)$, $f_2=(a-1)(n-1)$ and $\lambda=\frac{\hat{\phi}_1^{AOV}}{\hat{\phi}_0^{AOV}}$. Note the intuitively appealing result that $f\to a(n-1)$ as $\lambda\to 0$, and $f\to(n-1)$ as $\lambda\to\infty$. Why is this intuitively appealing? To answer this, it is helpful to point out that $n^{-1}\left(\hat{\phi}_0^{AOV}+\hat{\phi}_1^{AOV}\right)$ is mathematically identical to $n^{-1}\left(a^{-1}\sum_i V_i\right)$, where $V_i=(n-1)^{-1}\sum_j\left(y_{ij}-\bar{y}_{i.}\right)^2$ is the observed sample variance for treatment i with (n-1) degrees of freedom. However, under (2.2), the a treatment variances are correlated, and it can be shown that the correlation between any two observed variances is proportional to ρ^2. Hence, when $\rho\approx 0$, the a variances effectively contribute $(n-1)$ degrees of freedom each, resulting in $f\approx a(n-1)$. Conversely, when $\rho\approx 1$, there is effectively only a single independent variance, resulting in $f\approx(n-1)$. In general, $(n-1)\le f\le a(n-1)$.

Finally, the traditional AOV statistic for testing $H_0:\mu_1=\mu_2=\ldots=\mu_a$ is

$$F_{cal}^{AOV}=\frac{\frac{n\sum_i\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2}{(a-1)}}{\frac{\sum_i\sum_j\left(y_{ij}-\bar{y}_{i.}-\bar{y}_{j}+\bar{y}_{..}\right)^2}{(a-1)(n-1)}} \tag{2.7}$$

Under H_0, $F_{cal}^{AOV}\sim F\left(a-1,(a-1)(n-1)\right)$, where $F(\nu_1,\nu_2)$ denotes a central F distribution with ν_1 and ν_2 degrees of freedom. An important observation, one that leads us to the natural extension of the AOV statistic for dealing with missing values later, is to note that the statistic in (2.7) can be written in the more convenient form

$$F_{cal}^{AOV}=\frac{\frac{2\sum_{i<k}\left(\bar{d}_{ik(.)}\sqrt{n}\right)^2}{a(a-1)}}{\frac{\sum_{i<k}(n-1)V_{ik}}{\sum_{i<k}(n-1)}}. \tag{2.8}$$

In AOV parlance, the numerator and denominator in (2.8) represent the average mean square and the average variance, respectively, of all normalized pairwise treatment contrasts.

It should be noted that with no missing values, AOV and REML are mathematically identical as long as ϕ_1 is not constrained to be non-negative (Hocking, 1985, chapter 10). In summary, the analysis is straightforward when there are no missing values, with formulae (2.3) through (2.8) readily available for hypothesis testing and estimation. In the next section, we show how the AOV formulae can be extended in a natural way to deal with missing values.

3. PACE Analysis With Missing Values

To allow for missing values, additional notation is required. Let n_{ik} denote the number of blocks in which there is a response available for both treatments i and k. Also, let $\bar{d}^P_{ik(.)}, V^P_{ik}$ and C^P_{ik} $(i<k)$ denote the analogs of $\bar{d}_{ik(.)}, V_{ik}$ and C_{ik}, respectively, calculated using pairwise-available-cases only. In other words, the means, variances and covariances are obtained separately for each (i,k) treatment pair using only those blocks that contain a response for both treatment i and treatment k. Accordingly, if all the blocks contain responses for treatments i and k, then $\bar{d}^P_{ik(.)} = \bar{d}_{ik(.)}, V^P_{ik} = V_{ik}$ and $C^P_{ik} = C_{ik}$.

Based on the pairwise-available-case-estimation (PACE) approach, a natural extension of the AOV method is as follows. First, variance component estimates can be obtained by using

$$\hat{\phi}_1^{PACE} = \frac{\sum_{i<k}(n_{ik}-1)C^P_{ik}}{\sum_{i<k}(n_{ik}-1)} \tag{3.1}$$

and

$$\hat{\phi}_0^{PACE} = \frac{\sum_{i<k}(n_{ik}-1)V^P_{ik}}{\sum_{i<k}(n_{ik}-1)} \; . \tag{3.2}$$

Note that we are using weighted averages above, with weights $(n_{ik}-1)$ being inversely proportional to the sampling variances of V^P_{ik} and C^P_{ik} in (3.1) and (3.2), respectively, conditioned on the observed n_{ik}.

Next, for confidence intervals related to a single mean or for the difference between two means, we replace n under the square root sign in (2.5) and (2.6) with $n_{i,obs}$ and n_{ik}, respectively, and use the PACE variance component estimates based on (3.1) and (3.2). Here, $n_{i,obs}$ is the number of observed (i.e., non-missing) responses for treatment i. In addition, for the degrees of freedom in (2.5) and (2.6), we replace f_1 and f_2 in the expression for f given earlier with $\tilde{f}_1$ and $\tilde{f}_2$ given later in (3.7) and (3.8), respectively.

Finally, the PACE extension of the AOV statistic for testing the overall null hypothesis of equal treatment means is

$$F_{cal}^{PACE} = \frac{\dfrac{2\sum\limits_{i<k}\left(\bar{d}_{ik(.)}^{P}\sqrt{n_{ik}}\right)^2}{a(a-1)}}{\dfrac{\sum\limits_{i<k}(n_{ik}-1)V_{ik}^{P}}{\sum\limits_{i<k}(n_{ik}-1)}} . \tag{3.3}$$

Note that the PACE approach is based entirely on intrablock treatment differences, and, unlike REML, makes no attempt at recovering interblock information. However, it uses information in the incomplete blocks that CC-AOV completely ignores.

Null Distribution of F_{cal}^{PACE}

The PACE statistic in (3.3) is a ratio of two quadratic forms that are independent with no missing data, and approximately independent if data are missing completely at random. Under H_o, the distributions of these quadratic forms can be approximated as follows:

$$Q_1 = \frac{2\sum\limits_{i<k}\left(\bar{d}_{ik(.)}^{P}\sqrt{n_{ik}}\right)^2}{a(a-1)} \sim a_1\chi^2_{\tilde{f}_1} \tag{3.4}$$

and

$$Q_2 = \frac{\sum\limits_{i<k}(n_{ik}-1)V_{ik}^{P}}{\sum\limits_{i<k}(n_{ik}-1)} \sim a_2\chi^2_{\tilde{f}_2} , \tag{3.5}$$

where χ^2_ν denotes a central chi-square distribution with ν degrees of freedom, and $a_1, a_2, \tilde{f}_1$ and $\tilde{f}_2$ are constants that can be determined using the Satterthwaite (1941) approximation, i.e., by solving the simultaneous equations $E(Q_i) = a_i\tilde{f}_i$ and $Var(Q_i) = 2a_i^2\tilde{f}_i$, for $i = 1, 2$. Once $a_1, a_2, \tilde{f}_1$ and $\tilde{f}_2$ are obtained, using Theorem 6.1 in Box (1954), the null distribution of the PACE statistic can be approximated as

$$F_{cal}^{PACE} \dot{\sim} \left(\frac{a_1\tilde{f}_1}{a_2\tilde{f}_2}\right)F(\tilde{f}_1, \tilde{f}_2) . \tag{3.6}$$

Simple algebra (see Appendix) reveals that $a_1\tilde{f}_1 = a_2\tilde{f}_2$, and that

$$\tilde{f}_1 \equiv \frac{\left[\dfrac{a(a-1)}{2}\right]^2}{\dfrac{a(a-1)}{2} + \sum\limits_{P}\dfrac{n_{ik}n_{i'k'}}{2m^2_{ik,i'k'}}} \tag{3.7}$$

and

$$\tilde{f}_2 \cong \frac{\left[\sum_{i<k}(n_{ik}-1)\right]^2}{\sum_{i<k}(n_{ik}-1)+\sum_{P}\frac{(n_{ik}-1)(n_{i^*k^*}-1)}{2(m_{ik,i^*k^*}-1)}}. \qquad (3.8)$$

Above, m_{ik,i^*k^*} = harmonic mean of n_{ik} and $n_{i^*k^*}$, and the set P is given by

$$P=\left\{(i,i^*,k,k^*)\,\middle|\,\{i=i^*<k<k^*\}\cup\{i<i^*=k<k^*\}\cup\{i<i^*<k=k^*\}\right\}. \qquad (3.9)$$

For example, when $a=4$, the table below illustrates which (i,k) and (i^*,k^*) pairings would (✓), or would not (x) be included in P.

		(i^*,k^*)					
		(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)
	(1,2)	x	✓	✓	✓	✓	X
	(1,3)	x	x	✓	✓	x	✓
(i,k)	(1,4)	x	x	x	x	✓	✓
	(2,3)	x	x	x	x	✓	✓
	(2,4)	x	x	x	x	x	✓
	(3,4)	x	x	x	x	x	x

Since $a_1\tilde{f}_1 = a_2\tilde{f}_2$, it follows from (3.6) that the null distribution of the PACE statistic in (3.3) is approximately $F(\tilde{f}_1,\tilde{f}_2)$, where $\tilde{f}_1$ and $\tilde{f}_2$ are given by (3.7) and (3.8), respectively. Note that, since the cardinality of P is $a(a\text{-}1)(a\text{-}2)/2$, when there are no missing values, $\tilde{f}_1 = f_1 = (a-1)$ and $\tilde{f}_2 = f_2 = (a-1)(n-1)$, as expected, for in that case F_{cal}^{PACE} is mathematically identical to the AOV statistic in (2.7) and (2.8).

4. Illustrative Example

A detailed analysis of the coagulation time data in Table I is provided to help readers check their implementation of the PACE method, and to caution against CC-AOV.

Table II displays the standardized contrast scores and their summary statistics for all pairwise contrasts. We have: a=4, n_{12}=28, n_{13}=27, n_{14}=28, n_{23}=31, n_{24}=32, n_{34}=31, $m_{12,13}$=27.5, $m_{12,14}$=28.0, $m_{12,23}$=29.4, $m_{12,24}$=29.9, $m_{13,14}$=27.5, $m_{13,23}$=28.9, $m_{13,34}$=28.9, $m_{14,24}$=29.9, $m_{14,34}$=29.4, $m_{23,24}$=31.5, $m_{23,34}$=31.0, and $m_{24,34}$=31.5. Next, for the means, variances, and covariances, we have $\bar{d}_{12(.)}^{P} = -.224$, $\bar{d}_{13(.)}^{P} = -.191$, $\bar{d}_{14(.)}^{P} = -.336$, $\bar{d}_{23(.)}^{P} = -.010$, $\bar{d}_{24(.)}^{P} = -.120$, $\bar{d}_{34(.)}^{P} = -.108$, $V_{12}^{P} = .048$, $V_{13}^{P} = .023$, $V_{14}^{P} = .035$, $V_{23}^{P} = .069$, $V_{24}^{P} = .055$, $V_{34}^{P} = .044$, $C_{12}^{P} = .062$, $C_{13}^{P} = .043$, $C_{14}^{P} = .031$, $C_{23}^{P} = .045$, $C_{24}^{P} = .056$, and $C_{34}^{P} = .035$. Hence, (3.1) and (3.2) yield $\hat{\phi}_1^{PACE} = .046$ and $\hat{\phi}_0^{PACE} = .046$, respectively, and using (3.3) gives $F_{cal}^{PACE} = 22.918$. Finally, using (3.7) and (3.8), we get $\tilde{f}_1 = 2.997$ and $\tilde{f}_2 = 85.503$. Consequently, the p-value for the

Table II. Normalized pairwise contrasts and summary statistics for coagulation data

Subject	$d_{12}=\frac{T1-T2}{\sqrt{2}}$	$d_{13}=\frac{T1-T3}{\sqrt{2}}$	$d_{14}=\frac{T1-T4}{\sqrt{2}}$	$d_{23}=\frac{T2-T3}{\sqrt{2}}$	$d_{24}=\frac{T2-T4}{\sqrt{2}}$	$d_{34}=\frac{T3-T4}{\sqrt{2}}$
1	-.615	.035	-.276	.651	.339	-.311
2	.	.	.	-.120	-.523	-.403
3	-.113	-.339	-.431	-.226	-.318	-.092
4	-.255	-.262	-.417	-.007	-.163	-.156
5	.	.	.	-.057	.113	.170
6	-.233	-.453	-.523	-.219	-.290	-.071
7	-.226	-.148	-.283	.078	-.057	-.134
8	.	-.361	.	.	.	.
9	.	.	.	-.255	-.198	.057
10	-.311	-.382	-.092	-.071	.219	.290
11	-.219	-.269	-.587	-.049	-.368	-.318
12	-.071	-.106	-.523	-.035	-.453	-.417
13	.	-.049	-.311	.	.	-.262
14	.	.	.	-.453	-.127	.325
15	-.191	.	-.460	.	-.269	.
16	-.375	.057	-.177	.431	.198	-.233
17	-.431	.007	-.509	.438	-.078	-.516
18	-.113	-.346	-.170	-.233	-.057	.177
19	-.332	-.297	-.297	.035	.035	.000
20	-.177	-.170	-.325	.007	-.148	-.156
21	-.488	-.184	.	.304	.	.
22	-.361	.	-.820	.	-.460	.
23	.	.	.	-.028	-.085	-.057
24	-.035	-.354	.	-.318	.	.
25	.085	.	-.276	.	-.361	.
26	.170	-.106	.	-.276	.	.
27	-.219	-.099	-.198	.120	.021	-.099
28	.	.	.	.092	.021	-.071
29	-.311	-.134	-.460	.177	-.148	-.325
30	.339	-.205	-.092	-.544	-.431	.113
31	.	.	.	.	.	-.177
32	-.728	-.460	-.573	.269	.156	-.113
33	-.247	.	-.283	.	-.035	.
34	.	.	.	-.205	-.212	-.007
35	-.099	-.240	-.502	-.141	-.403	-.262
36	.	.085	-.148	.	.	-.233
37	.	.	-.226	.	.	.
38	-.092	-.028	-.318	.064	-.226	-.290
39	-.368	-.170	-.106	.198	.262	.064
40	-.240	-.191	-.021	.049	.219	.170
n_{ik}	28	27	28	31	32	31
Mean	-.224	-.191	-.336	-.010	-.120	-.108
Variance	.048	.023	.035	.069	.055	.044

overall treatment comparison is 5.7×10^{-11}. The p-values for the pairwise comparisons are obtained in a similar manner.

Table III contains an overall summary of the PACE results, along with those for CC-AOV and REML, the latter obtained using SAS PROC MIXED (version 6.12, using the REPEATED statement with TYPE=CS and the DDFM=SATTERTH option in the model statement to estimate denominator degrees of freedom for the fixed effects). For these data, the three methods yield generally similar conclusions. The single exception is for the Treatment 2 versus Treatment 4 comparison, for which the CC-AOV approach (which uses data from only 21 of the 40 subjects!) yields a two-tailed p-value of .0895, while the PACE and REML approaches yield highly significant p-values of .0023 and .0012, respectively.

Table III. Summary of results for coagulation data

	CC-AOV	PACE	REML
Interblock ($\hat{\phi}_1$)	.046	.046	.040
Intrablock ($\hat{\phi}_0$)	.046	.046	.046
	Two-tailed p-values (numerator d.f., denominator d.f.)		
Overall	< .0001 (3, 60)	< .0001 (3.0, 85.5)	< .0001 (3, 99.2)
T1 vs. T2	< .0001 (1, 60)	< .0001 (1, 85.5)	< .0001 (1, 99.6)
T1 vs. T3	< .0001 (1, 60)	< .0001 (1, 85.5)	< .0001 (1, 100)
T1 vs. T4	< .0001 (1, 60)	< .0001 (1, 85.5)	< .0001 (1, 99.8)
T2 vs. T3	.3161 (1, 60)	.7867 (1, 85.5)	.8361 (1, 98.7)
T2 vs. T4	.0895 (1, 60)	.0023 (1, 85.5)	.0012 (1, 98.4)
T3 vs. T4	.0082 (1, 60)	.0066 (1, 85.5)	.0024 (1, 98.9)

5. Simulation Study

5.1 Description

The inferential properties of the PACE method were investigated via a simulation study designed to mimic conditions typically encountered in trials that motivated this research. The following four factors were varied: number of treatments (a=3, 4 and 5), number of blocks (n=10, 20 and 30), the percentage of missing values (5%, 10% and 15%), and the interblock variance component (ϕ_1=1/9, 1 and 9). For each combination of these factors, normally distributed data were generated according to model (2.1), with $\phi_0 = 1$. Note that values of 1/9, 1 and 9 for ϕ_1 correspond to intrablock correlations, $\rho = \phi_1(\phi_1 + \phi_0)^{-1}$, of 0.1, 0.5 and 0.9, respectively.

The overall treatment comparison, as well as the pairwise comparison of μ_1 to μ_2, were studied. To investigate test size, all treatment means were set equal to zero, and to study power, μ_1 and μ_2 were set equal to 0.3 and 1, respectively, with all other mean(s) continuing to be zero. Empirical test sizes and powers for PACE, CC-AOV and REML were determined after 10,000 simulations. In each case, the nominal test size was 5%.

5.2 Results

For brevity, only the results for $a = 3$ and 5, and $\rho = 0.1$ and $\rho = 0.9$ are discussed; other results were similar, and are available from the author. Empirical test sizes are displayed in Table IV. As expected, none of the three methods have inflated test sizes, either for the overall comparison or pairwise comparisons. Of note, the test size for PACE tends to get smaller with more missing data.

Table V contains the empirical power results for all methods. Once again, the results are as expected, with $\text{power}_{\text{CC-AOV}} < \text{power}_{\text{PACE}} < \text{power}_{\text{REML}}$ when there are missing values. As the percentage of missing values increases, all methods steadily lose power, with the rate of loss being fastest for CC-AOV and slowest for REML.

Since PACE ignores interblock information, it is expected to have smaller power than REML. Note that, with only 5% missing data, the power of REML is slightly bigger than that of PACE, but can be substantially bigger than that of CC-AOV! This is one of the key messages of this note. For example, with a=5, n=30, $\rho = 0.9$ and only 5% missing data, for the pairwise comparison, the powers for CC-AOV, PACE and REML are 65%, 72% and 74%, respectively. When the percentage of missing data is increased to 10%, the correspondence power difference between CC-AOV and REML is even more pronounced, for both the pairwise comparison (CC-AOV = 53%, REML = 71%) as well as the overall comparison (CC-AOV = 80%, REML = 95%). It is clear that even when the anticipated percentage of missing data is small ($\leq$ 10%), using CC-AOV can result in a substantial loss of power compared with REML.

Note that as ρ increases from 0.1 to 0.9, power differences between REML and PACE become monotonically smaller. The reason for this is that as ρ increases, i.e., as ϕ_1 increases relative to ϕ_0, there is increased uncertainty about whether interblock differences represent true treatment differences estimated via incomplete blocks, or whether they are merely an artifact of the large interblock variability. Due to increased confounding between block and treatment effects, there is a poorer "recovery" of interblock information for REML, and, hence, REML comes closer to PACE, with the bulk of the weight moving away from the interblock to the intrablock estimates of treatment differences.

Table IV. Empirical Test Size (nominal α=5%)

			3 treatments					
			Overall Comparison			Pairwise Comparison		
ρ	N	Missing	CC-AOV	PACE	REML	CC-AOV	PACE	REML
0.1	10	5%	5.2	5.1	5.5	5.3	5.5	5.4
		10%	4.7	4.6	5.3	5.0	4.9	5.4
		15%	5.0	4.0	5.4	5.0	4.9	5.4
	20	5%	4.9	4.6	5.1	4.9	4.9	4.9
		10%	5.1	4.8	5.2	5.3	5.4	5.3
		15%	4.8	4.3	4.8	5.0	5.0	4.6
	30	5%	5.1	4.9	5.0	5.1	5.1	5.2
		10%	5.3	4.7	5.2	5.0	5.0	5.1
		15%	5.0	4.2	5.2	5.2	4.8	5.3
0.9	10	5%	4.9	4.7	5.0	4.9	4.9	4.9
		10%	5.5	5.1	5.4	5.1	4.6	5.0
		15%	5.1	4.5	5.1	5.2	4.7	5.3
	20	5%	5.1	4.8	4.9	5.4	5.2	5.2
		10%	5.0	4.7	5.1	4.8	4.8	5.0
		15%	5.0	4.7	5.3	5.0	5.1	4.9
	30	5%	4.9	4.6	4.8	4.6	4.8	4.8
		10%	5.0	4.5	4.9	5.0	5.1	5.1
		15%	5.2	5.0	5.5	5.2	5.0	5.2
			5 treatments					
0.1	10	5%	4.9	4.5	5.0	5.0	5.0	5.1
		10%	5.1	4.2	5.2	5.0	4.7	4.9
		15%	5.0	4.0	5.2	5.1	5.0	5.3
		15%	5.0	4.0	5.2	5.1	5.0	5.3
	20	5%	5.3	4.8	5.4	5.0	5.3	5.2
		10%	5.1	4.6	5.1	5.2	5.0	5.1
		15%	5.1	4.3	5.2	4.9	5.0	4.7
	30	5%	4.8	4.3	4.8	5.2	5.0	5.1
		10%	5.0	4.4	5.2	4.5	4.8	5.1
		15%	5.1	4.0	5.1	4.7	5.0	5.2
0.9	10	5%	4.6	4.7	5.0	4.8	4.8	4.8
		10%	4.9	4.4	5.2	4.7	4.7	5.0
		15%	5.0	3.9	5.1	5.1	4.7	4.9
	20	5%	5.3	4.6	4.9	4.9	4.7	4.7
		10%	4.9	4.2	5.0	5.1	4.9	4.9
		15%	5.4	3.9	5.0	5.2	5.1	5.4
	30	5%	5.1	4.7	5.1	5.0	4.8	4.8
		10%	5.2	4.5	5.3	5.0	5.0	5.2
		15%	5.0	3.8	5.0	5.1	4.8	4.8

Table V. Empirical Power (%)

			Overall Comparison			Pairwise Comparison		
ρ	N	Missing	CC-AOV	PACE	REML	CC-AOV	PACE	REML
			3 treatments					
0.1	10	5%	38	40	43	27	28	30
		10%	32	36	41	23	25	27
		15%	27	32	39	20	23	27
	20	5%	73	76	78	50	52	55
		10%	64	70	75	44	48	52
		15%	55	65	73	38	44	51
	30	5%	90	92	93	69	71	73
		10%	84	89	92	62	67	71
		15%	76	84	89	54	60	67
0.9	10	5%	39	41	43	28	29	30
		10%	33	36	39	24	27	28
		15%	27	32	35	19	23	25
	20	5%	72	76	77	50	53	54
		10%	64	71	73	44	49	51
		15%	56	65	69	39	44	48
	30	5%	89	91	92	69	71	72
		10%	84	88	90	61	67	69
		15%	76	85	87	55	62	66
			5 treatments					
0.1	10	5%	39	46	49	26	31	32
		10%	29	41	46	22	29	31
		15%	22	36	44	16	25	28
	20	5%	75	83	85	49	55	57
		10%	60	78	82	37	49	53
		15%	45	72	79	29	45	51
	30	5%	92	96	97	65	72	75
		10%	81	94	96	53	67	72
		15%	66	91	94	42	62	69
0.9	10	5%	39	47	49	26	30	31
		10%	29	41	45	21	27	29
		15%	21	35	41	16	25	28
	20	5%	75	83	84	48	55	56
		10%	60	79	81	39	50	53
		15%	45	72	77	29	44	49
	30	5%	92	97	97	65	72	74
		10%	80	94	95	53	67	71
		15%	67	91	93	43	62	68

6. Concluding Remarks

This note has focused on the analysis of data from randomized block experiments where routinely only a few missing values (if any) are encountered. REML, which continues to be increasingly popular for dealing with missing data, is a powerful method that makes efficient use of all the available data. However, our experience suggests that (unfortunately) some researchers shy away from using REML because they view it as a "black-box" approach that lacks the simplicity and intuitive appeal of a summary-statistic-based approach like AOV. Accordingly, a natural extension of the traditional AOV method has been proposed to deal with the design imbalance caused by missing values. The proposed method (PACE) is a significant improvement over the naïve (but not uncommon) CC-AOV approach. Moreover, when only a few data are missing (5%), PACE is almost as powerful as REML, but with many missing values REML is significantly more powerful.

In conclusion, our recommendation is to stay away from CC-AOV and adopt REML as the method of choice even when routinely only a few missing values are expected. PACE could be used an alternative to REML, but *only* if software to implement the latter is unavailable, or the use of a "transparent" formula-based approach is deemed necessary.

APPENDIX

We need to solve $E(Q_i) = a_i \tilde{f}_i$ and $Var(Q_i) = 2a_i^2 \tilde{f}_i$, for i = 1, 2, where Q_1 and Q_2 are defined in (3.4) and (3.5), respectively. First, note that:

$$\begin{aligned} Cov(d_{ik(j)}, d_{i'k'(j)}) &= \frac{\phi_0}{2} \qquad i = i^* < k < k^* \\ &= \frac{-\phi_0}{2} \qquad i < i^* = k < k^* \\ &= \frac{\phi_0}{2} \qquad i < i^* < k = k^*. \end{aligned}$$

Next, the following results are easily established under the null hypothesis of equal means:

$$\frac{n_{ik}(\bar{d}^P_{ik(.)})^2}{\phi_0} \sim \chi^2_1 \text{ and } \frac{(n_{ik}-1)V^P_{ik}}{\varphi_0} \sim \chi^2_{n_{ik}-1}.$$

From this, it follows that:

$$E[(\bar{d}^P_{ik(.)})^2] = \frac{\phi_0}{n_{ik}} \text{ and } Var[(\bar{d}^P_{ik(.)})^2] = \frac{2\phi_0^2}{n_{ik}^2},$$

and

$$E[V_{ik}] = \phi_0 \quad \text{and} \quad Var[V_{ik}] = \frac{2\phi_0^2}{n_{ik}-1},$$

and, hence, that $E[Q_1] = a_1\tilde{f}_1 = E[Q_2] = a_2\tilde{f}_2 = \phi_0$.

The only results that remain to be determined are $Cov[(\bar{d}^P_{ik(.)})^2, (\bar{d}^P_{i'k'(.)})^2]$ and $Cov(V^P_{ik}, V^P_{i'k'})$. Suppose there were no missing values, i.e., $n_{ik}=n$, for all $i<k$. Let

$\vec{d}_{ik}^{\,T} = (d_{ik(1)}, d_{ik(2)}, ..., d_{ik(n)})$ denote the 1 x n vector of normalized contrast scores comparing treatment i to treatment k. Note that $\bar{d}^2_{ik(.)}$ and V_{ik} can be written as quadratic forms in $\vec{d}_{ik}$. Specifically,

$$\bar{d}^2_{ik(.)} = \vec{d}_{ik}^{\,T}\left(\frac{1}{n^2}\vec{J}_n\vec{J}_n^{\,T}\right)\vec{d}_{ik} \quad \text{and} \quad V_{ik} = \vec{d}_{ik}^{\,T}\left(\frac{1}{n-1}\left\{I_n - \frac{1}{n}\vec{J}_n\vec{J}_n^{\,T}\right\}\right)\vec{d}_{ik},$$

where $\vec{J}_n$ is an n x 1 vector of ones and I_n is the identity matrix of order n. Hence, using a well-known result on the covariance between two quadratic forms in normal variates (Searle, 1971, pg. 66), it is quickly established that, with no missing values,

$$Cov(\bar{d}^2_{ik(.)}, \bar{d}^2_{i'k'(.)}) = \frac{\phi_0^2}{2n^2} \quad \text{and} \quad Cov(V_{ik}, V_{i'k'}) = \frac{\phi_0^2}{2(n-1)}.$$

The above covariances are as shown provided the subscripts belong to the set P given by (3.9); otherwise the covariances are zero. With missing values, there are no closed-form expressions for the above covariances. In light of the missing completely at random (MCAR) assumption, the following approximations are used:

$$Cov[(\bar{d}^P_{ik(.)})^2, (\bar{d}^P_{i'k'(.)})^2] \cong \frac{\phi_0^2}{2m^2_{ik,i'k'}} \quad \text{and} \quad Cov(V^P_{ik}, V^P_{i'k'}) \cong \frac{\phi_0^2}{2(m_{ik,i'k'}-1)},$$

where $m_{ik,i'k'}$ = harmonic mean of n_{ik} and $n_{i'k'}$; the arithmetic mean is not used because its use does not guarantee $\tilde{f}_1 \le (a-1)$.

The final expressions for $\tilde{f}_1$ and $\tilde{f}_2$ in (3.7) and (3.8), respectively, are now easily obtained using the intermediary results described above.

ACKNOWLEDGMENTS

Helpful comments by an anonymous referee are gratefully acknowledged.

REFERENCES

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Annals of Mathematical Statistics*, 25, 290-302.

Corbeil, R. R. and Searle, S. R. (1976). Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model. *Technometrics*, 18, 31-38.

Hocking, R. R. (1985). *The Analysis of Linear Models*, Monterey, CA:Brooks-Cole.

Little, R. J. A. and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, John Wiley & Sons, Inc.

May, W. L. and Johnson, W. D. (1995). Some Applications of the Analysis of Multivariate Normal Data With Missing Observations. *Journal of Biopharmaceutical Statistics*, 5, 215-228.

Patterson, H. D. and Thompson, R. (1971). Recovery of interblock information when block sizes are unequal. *Biometrika*, 58, 545-554.

Satterthwaite, F. F. (1941). Synthesis of variance. *Psychometrika*, 6, 309-316.

Searle, S. R. (1971). *Linear Models*, New York: John Wiley.