Theor Appl Genet (1989) 78: 381–386

© Springer-Verlag 1989

Efficient design of crossbreeding experiments

E. P. Cunningham[1] **and J. Connolly**[2]

[1] Trinity College, Dublin 2, Ireland
[2] University College, Dublin 4, Ireland

Received: July 6, 1987; Accepted: March 17, 1989
Communicated by E. J. Eisen

Summary. The efficiencies of different experimental configurations for estimating additive (A) and heterotic (H) effects in purebred and crossbred populations derived from two parent breeds are examined. Allocation of resources over six groups is considered: the two parental breeds, the F_1, the F_2 and the backcrosses to the two parental breeds. Additive and heterotic effects are best estimated by allocating resources to the two parental groups and F_1 in the proportions 35%, 35%, 30% and for many practical situations, an experiment involving numbers in the region 250–300 is reasonable. If it is not possible to include all three groups, other combinations involving a subset of them and some of the other three groups can be used to give estimates of A and H. However, even the best of these alternatives requires over twice the resources to give the same precision as the optimal design and, further, these estimates may be correlated. Relatively modest reallocation of resources to the F_2 to estimate or test for an epistatic effect (E) leads to a minor reduction in the precision of estimates of A and H while giving reasonable precision for the estimate of E. The inclusion of maternal effects in the model greatly reduces the efficiency of estimation of A and H. Where one of the breeds is introduced through the sire line only, optimal allocation gives roughly equal replication to the pure lines and F_1 but about 63% of allocation is placed in equal amounts on the two backcrosses produced through crossing F_1 dams with pure sires of both breeds. The relevance of these results to the planning of livestock crossbreeding trials, particularly those involving a local and an exotic or imported breed, is discussed.

Key words: Heterosis – Crossbreeding – Epistasis – Design – Maternal effects

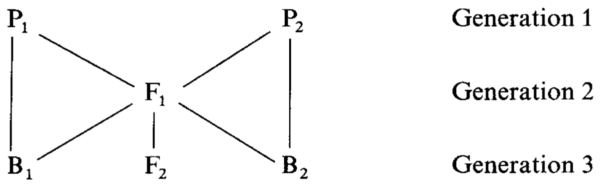
Introduction

Crossbreeding is widely used in animal and plant improvement programs. Its purpose may be to exploit additive and non-additive genetic differences between breeds (or strains) and/or to take advantage of complementarity between breeds for different traits. The crossbreeding strategy used may involve two or more breeds, and can have several forms: topcrossing or grading up, continuous hybrid production, rotational crossing, formation of new synthetic or composite populations.

The correct choice among these possible strategies depends on a proper assessment of the extent of additive (A) differences between the strains or breeds being considered. It also requires some knowledge of the scale of dominance effects which can give rise to heterosis (H) effects in their crosses. Finally, it requires some knowledge or assumptions about the extent of epistatic and maternal effects which can further modify the expected performance of crossbred generations (Cunningham 1987).

Well-planned trials at the early stages of such a crossbreeding program are necessary to provide an adequate information base on which to design the subsequent breeding strategy. Considerable care and investment is justified in the design and conduct of these trials, because of the scale, duration and economic impact of the breeding programs which follow.

The primary purpose of any such crossbreeding trial is the estimation of A and H with sufficient accuracy and precision for subsequent plans to be developed with reasonable confidence. If such trials are to be conducted within the first two generations of crossing between two breeds, they can involve any or all of six generation groups: the two parental breeds (P_1 and P_2), the first cross (F_1), the first-cross inter se (F_2), and the backcrosses (B_1 and B_2) to the two parental breeds.



Provided that the differences between the groups are not a reflection of environmental, time, location, nutritional or other non-genetic factors, A and H can be estimated from linear functions of these group means. In addition, it is possible to test for the presence of significant epistatic or maternal effects.

In this paper, we examine the efficiencies of different experimental configurations involving these six generation groups. Consideration is limited to trials involving two parental breeds or strains. For convenience, and also because it corresponds to a frequent situation in practice, one parental breed (P_1) is referred to as the local breed, and the other (P_2) as the exotic or imported breed.

Methodology

If the performances of the different groups are represented by P_1, P_2, F_1, B_1, F_2 and B_2 respectively, then a model in terms of A and H is

$$\begin{aligned} P_1 &= \mu && && + e_1 \\ P_2 &= \mu + A && && + e_2 \\ F_1 &= \mu + 0.5\,A && + H && + e_3 \\ B_1 &= \mu + 0.25\,A && + 0.5\,H && + e_4 \\ F_2 &= \mu + 0.5\,A && + 0.5\,H && + e_5 \\ B_2 &= \mu + 0.75\,A && + 0.5\,H && + e_6 \end{aligned} \tag{1}$$

The terms e_1 to e_6 are error elements with variances σ^2/n_i, for $i = 1 \ldots 6$, where n_i is replication in the i^{th} group and σ^2 is the between-animal variance, assumed equal in all six groups.

The individual variance σ^2 is a function of the between-sire (σ_s^2) and error or within-sire (σ_w^2) components of variance, and hence is sensitive to both the size of these variances and to the ratio of the numbers of sires (n_{si}) and progeny (n_i) per group. The relationship is

$$\sigma^2 = \sigma_s^2\,(n_i/n_{si}) + \sigma_w^2. \tag{2}$$

Our assumption of constant σ^2 across groups is satisfied if σ_w^2, the ratio σ_w^2/σ_s^2 (in effect the heritability) and the ratio n_i/n_{si} (the number of progeny per sire) are constant. It is not possible to pursue all combinations of deviations from these conditions. Furthermore, the question of optimum allocation between sires within a group, primarily a function of the heritability, is not pursued here. However, we show later how more complex error structures can be accommodated.

The assumption of equal heritability in all breeding groups requires some further comment. The parental groups, by definition, represent different populations, which may have quite different genetic histories and, therefore, differ in heritability for any particular trait. The sample of sires used in each group may be a more or less adequate representation of the genetic variance in the population from which they are drawn. In all breed and cross comparisons, the purpose is to draw inferences for the parent populations and their actual or potential crosses. It is, therefore, important that an adequate and representative sample of sires and dams is included. In the crossbred groups, covariances between individuals may contain heterotic as well as additive genetic components. The main focus in the present study is on cattle and sheep trials. In litter-bearing species such as pigs, a further level of complexity is added in that the covariances within and between litters may also be influenced by non-additive genetic effects. Finally, the samples of sires used in two or more groups may overlap, creating covariances between individuals in different groups.

A further complication in crossbreeding trials is that the confounding of time differences with genetic groups may be difficult to avoid. The F_1 is used to produce the F_2 and the backcrosses, and therefore cannot be contemporaneous with them unless a second generation of F_1's is produced. Furthermore, selection of sires and dams within the F_1 for the production of these three groups (B_1, B_2, F_2) can bias the comparisons. These questions are discussed in more detail by Dickerson (1969).

The additive breed difference, A, is presumed to depend largely on additive effects of genes. However, within-breed dominance and epistatic effects may also contribute to A. The heterosis effect, H, is defined as the difference between F_1 and the mean of P_1 and P_2. It is presumed to be mainly a function of dominance effects, and is therefore expected to be proportional to heterozygosity. In addition to the dominance effects, some epistatic effects may be confounded with the estimate of heterosis. For example, half of the parental additive by additive epistatic effects may contribute to H. Furthermore, to the extent that P_1 and P_2 differ in their gene content, new additive-by-additive effects can arise in the B_1, and these will also be confounded with the estimate of H. However, the balance of evidence suggests that these epistatic effects are of negligible importance in many instances (Cunningham 1980).

In order to obtain estimates of both A and H, a minimum of three of these groups is required in the trial. Starting with the local population (P_1), it is relatively easy to generate F_1 offspring using imported semen or males. It may also be possible to provide some animals of the exotic breed (P_2) for evaluation in the same environment. These three groups (both parents and F_1) give estimates of A and H as

$$\begin{aligned} \hat{A} &= P_2 - P_1 \\ \hat{H} &= F_1 - (P_1 + P_2)/2 \end{aligned} \tag{3}$$

With this set of three groups, what size of experiment is required to give an acceptable level of precision in the estimation of A and H? This is the basic question in the design of such trials. Assuming unbiased estimation, precision is best measured as the standard errors of the estimates of A or H.

In order to make the calculation of these standard errors possible, the allocation of replication over the three groups must first be decided. Suppose the total number of animals is $N = n_1 + n_2 + n_3$. The variances of the estimates of A and H are then

$$\begin{aligned} V(\hat{A}) &= \sigma^2(1/n_1 + 1/n_2) \\ V(\hat{H}) &= \sigma^2[1/n_3 + 0.25(1/n_1 + 1/n_2)] \end{aligned} \tag{4}$$

Since it is not possible to simultaneously minimise both of these variances, we have chosen to select the design that minimises their sum $V(\hat{A}) + V(\hat{H})$. This is a minimum when $n_1 = n_2 = 0.346\,N$ and $n_3 = 0.309\,N$.

The set discussed above includes only the first two generations (P_1, P_2 and F_1). In the next generation, three further groups are possible: F_2, B_1 and B_2. There can be difficulties in comparing these three groups to the parental and F_1 groups because the first and second generation are usually not contemporaneous. However, it may be possible to generate further samples of the parental and F_1 groups to give valid comparisons.

Assuming that problems of this nature can be overcome and that all three of these additional groups can be made available, do they contribute to the precision with which A and H are estimated? One way to respond to this question is to specify a fixed total number of animals in the trial, and to reallocate a certain proportion of them from the parental and F_1 groups to the F_2 and backcross groups. The effect on the standard errors of $\hat{A}$ and $\hat{H}$ can then be observed.

The main purpose to be served by adding groups beyond P_1, P_2 and F_1 is to provide an opportunity for testing the goodness of fit of the additive-heterosis model. Other than when maternal effects are of importance, any departure from this model is expected to be due to more complex gene interactions, collectively known as epistatic effects. Since these can occur in an infinite number of forms, attention is usually limited to the very simplest form, additive-by-additive gene effects. Following Dickerson (1969) we have chosen to define the epistatic effect (sometimes also called epistatic loss or recombination loss) as the expected value of

$$\hat{E} = F_2 - (1/4\,P_1 + 1/4\,P_2 + 1/2\,F_1). \tag{5}$$

This represents the deviation of the F_2 from the mean of the parental average and the F_1. Of the average additive-by-additive effects present in an average individual in the parental groups, the F_1 is expected to retain one-half, and the F_2 three-eighths. The epistatic effect, E, as defined here therefore represents minus three-eighths of the average additive-by-additive effects present in the parental groups.

With this definition of the epistatic effect, the performance of the F_2 and backcross groups is then represented by

$$\begin{aligned} B_1 &= \mu + 0.25\,A + 0.5\,H + 0.5\,E + e_4 \\ F_2 &= \mu + 0.5\,A \;\; + 0.5\,H + E \quad\;\; + e_5 \\ B_2 &= \mu + 0.75\,A + 0.5\,H + 0.5\,E + e_6 \end{aligned} \tag{6}$$

Minimisation of $V(\hat{A}) + V(\hat{H}) + V(\hat{E})$ gives values of $n_1 = n_2 = 0.260\,N$, $n_3 = 0.254\,N$ and $n_5 = 0.227\,N$, with $n_4 = n_6 = 0$. Thus, the only one of the three extra groups to be included would be F_2. Assuming that only these four groups are considered and that a certain proportion of resources are allocated to F_2, the increase in the standard errors of $\hat{A}$ and $\hat{H}$ relative to the optimum three-group allocation shows the loss incurred by reallocation. The gain from reallocation is that the epistatic effect can be estimated. The standard error of $\hat{E}$ relative to those of $\hat{A}$ and $\hat{H}$ for the various allocations gives some measure of the extent of the gain. The variances from which these may be calculated are shown below for the four-point design, with total animal numbers N allocated n_1, n_2, n_3 and n_5 to groups P_1, P_2, F_1 and F_2 respectively.

$$\begin{aligned} V(\hat{A}) &= \sigma^2(1/n_1 + 1/n_2) \\ V(\hat{H}) &= \sigma^2(1/n_3 + 1/4n_1 + 1/4n_2) \\ V(\hat{E}) &= \sigma^2(1/n_5 + 1/4n_3 + 1/16n_1 + 1/16n_2) \end{aligned} \tag{7}$$

Other formulations of the additive-by-additive epistatic effect are possible (Koch et al. 1985), but those reviewed there all result in identical analysis of variance to that chosen, since they all are coded values of each other.

In the design of such experiments, it is not always possible to choose the best combination of groups (P_1, P_2, F_1). For example, where P_2 is an exotic breed, it may not be possible to include it for practical or financial reasons. However, semen can be readily imported, so that F_1 progeny are usually easy to produce. From the F_1 generation, it is of course easy to produce F_2. Backcrosses to the exotic breed can be generated by further semen importations, while backcrosses to the local breed can be produced either by mating F_1 females to males of the local breed or vice versa.

Deciding the optimal allocation of numbers for various subsets of the six possible groups is not always algebraically tractable. If Eq. 1 are written in matrix notation as

$$Y = XB + E \tag{8}$$

with

$$\mathrm{Var}(Y) = V \tag{9}$$

where V is a diagonal matrix with typical element σ^2/n_i, then the least-squares estimates of B are

$$\hat{B} = (X'V^{-1}X)^{-1}X'V^{-1}Y \tag{10}$$

and the variance covariance matrix of these estimates is

$$V(\hat{B}) = (X'V^{-1}X)^{-1} \tag{11}$$

The sum of the variances $V(\hat{A}) + V(\hat{H})$ is now the sum of the second and third diagonal elements of $V(\hat{B})$. Selection of n_1 to n_6 to minimise the sum of these variances may be carried out using the minimisation algorithm of Nelder and Mead (1965), as programmed by O'Neill (1971). Where a point is excluded from consideration, the corresponding replication is constrained to be zero. The problem was structured to ensure non-negative solutions. One final point is that the estimates of A and H coming from these designs may be correlated, unlike those from the optimum three-point design which includes only the P_1, P_2 and F_1 groups.

The variance-covariance matrix V at (8) above has been defined as a diagonal matrix, implying that the separate generation groups are uncorrelated. This need not necessarily be the case. For example, the same sires might be used to generate the P_1, F_1 and B_1 groups. The structure of V will then be more complex. The necessary matrix has been developed for the most general case by Henderson (1985). It is not feasible to specify every possible covariance structure, and for every such structure the optimum allocations will be different. However, once the variance-covariance matrix V has been specified, the procedure outlined above can be followed to give an optimum allocation.

If maternal effects are considered important, they may be included in the model. The maternal effects on progeny phenotype are expressions of the genotype of the dam, but in terms of progeny performance they are really environmental effects. Model (12) below describes the phenotypic performance of nine genetic groups, where M is the maternal effect and MH maternal heterosis. Let superscripts s and d represent sire and dam respectively and let superscripts a and b represent two different ways of achieving a particular cross.

$$\begin{aligned} P_1 &= \mu + M \\ P_2 &= \mu + A \\ F_1^a &= P_1^s \times P_2^d = \mu + 0.5\,A \;\; + H \\ F_1^b &= P_1^d \times P_2^s = \mu + 0.5\,A \;\; + H + M \\ B_1^a &= P_1^s \times F_1^d = \mu + 0.25\,A + 0.5\,H + 0.5\,M + MH \\ B_1^b &= P_1^d \times F_1^s = \mu + 0.25\,A + 0.5\,H + M \\ B_2^a &= P_2^s \times F_1^d = \mu + 0.75\,A + 0.5\,H + 0.5\,M + MH \\ B_2^b &= P_2^d \times F_1^s = \mu + 0.75\,A + 0.5\,H \\ F_2 &= F_1 \times F_1 = \mu + 0.5\,A \;\; + 0.5\,H + 0.5\,M + MH \end{aligned} \tag{12}$$

Where P_2 is an exotic breed, introduced through the sire line only, the progeny equations will not include F_1^a or B_2^b. Other breeding systems use other subsets of these equations.

Results and discussion

As mentioned earlier, precision is best measured by the standard error of the estimate of A and H. If, e.g., the additive difference A is expected to be about 40% of the mid-parent mean, an estimate of A with a standard error equal to 10% of the mid-parent mean (i. e. one-quarter of the actual value estimated) might be regarded as adequate precision for the use of the estimate with confidence in the development of breeding plans. Similarly, if H were expected to be approximately 20% of mid-parent value, then a standard error of 5% (of mid-parent value) might be regarded as adequate precision. As the scale of the experiment goes up, the size of these standard errors of $\hat{A}$ and $\hat{H}$ comes down. It is then a matter of judgement as to what balance of precision versus scale is acceptable.

Table 1 shows the scale of experiment (with two parental groups and F_1) required for given levels of precision for the estimation of A and H. Traits will differ in their inherent variability, and this in turn will affect the relationship between precision and scale. Three levels of basic variability are therefore provided for: coefficients of variation [σ (Eq. 2) as percent of mid-parent mean] of 25%, 35% and 45%. To achieve given levels of the standard error of $\hat{A}$ or $\hat{H}$ (2.5%, 5%, 10%, 20%), the number of animals required in the trial is indicated. In all cases, optimal allocation of numbers to the three groups is assumed.

The following example illustrates the use of the table. If the main trait of interest has a coefficient of variation of 35%, and the standard errors of $\hat{A}$ and $\hat{H}$ are each required to be no greater than 5%, then the experiment should contain 284 animals to give this level of precision for the estimation of A, while 229 animals will achieve the desired precision in the estimation of H. Because the design is the same in all cases, H is always more precisely estimated than A (about 20% fewer animals being required to give the same precision).

In the case of this particular design (two parents and F_1) the optimum is achieved by allocating approximately 35% of the animals to the two parental groups and 31% to the F_1 group. In the example given above, therefore, the 284 animals in the experiment would be allocated 98 each to the two parental groups, and 88 to the F_1 group.

The allocation of some resources to the other three groups can be initially approached by equally distributing a certain proportion of the animals from the parental and F_1 groups over the F_2 and backcross groups. The results of this calculation are given in Table 2.

It can be seen that for fixed total experimental resources, the inclusion of these extra three groups in all cases reduces the precision of the estimates of A and H. If half of the animals are reallocated, the standard errors of the resulting estimates of A and H are increased by 27% and 36%, respectively.

Developing the other side of this question, the ability to test for goodness of fit if one or more of these three extra groups are included, the allocation of resources to F_2 to allow a test of epistasis is now considered. Various percentages of resources are allocated to F_2 and the remainder allocated to P_1, P_2 and F_1 in the ratio 0.346:0.346:0.308, the optimal ratios in the three-point design for two parents and F_1. Table 3 gives the size of the standard error for $\hat{A}$ or $\hat{H}$ for various reallocations to F_2 relative to that for the optimal three-point design (zero reallocation). Also given is the size of the standard error of the epistatic effect E relative to that of $\hat{A}$ and $\hat{H}$ for various reallocations.

Table 1. Numbers of animals required to give specified standard errors (SE) of $\hat{A}$ or $\hat{H}$ at different levels of variation. (SE and CV both measured as % of mid-parent mean)

Standard error of $\hat{A}$ or $\hat{H}$	Coefficient of variation					
	25%		35%		45%	
	$\hat{A}$	$\hat{H}$	$\hat{A}$	$\hat{H}$	$\hat{A}$	$\hat{H}$
2.5	579	468	1135	918	1876	1517
5	145	117	284	229	469	379
10	36	29	71	57	117	95
15	16	13	32	25	52	42
20	9	7	18	14	29	24

Table 2. The effect on the standard errors of $\hat{A}$ and $\hat{H}$ of reallocating resources from parental and F_1 groups to backcrosses and F_2 groups

Percent of animals reallocated from P_1, P_2 and F_1 to B_1, B_2 and F_2 groups	Relative size of standard errors of	
	$\hat{A}$	$\hat{H}$
0	100	100
10	104	105
20	109	110
30	114	117
40	120	125
50	127	136

Table 3. The effect on the standard errors of $\hat{A}$ or $\hat{H}$ of reallocating resources from parental and F_1 groups to the F_2 group. Also given is the size of the standard error for the epistatic effect relative to those for $\hat{A}$ and $\hat{H}$ for various reallocations

Percent of animals reallocated from P_1, P_2 and F_1 to the F_2 group	Standard error of $\hat{A}$ or $\hat{H}$ relative to optimal	Standard error of $\hat{E}$ relative to that of	
		$\hat{A}$	$\hat{H}$
0	100	inf	inf
5	103	187	207
10	105	133	147
15	108	109	121
20	112	95	105
25	115	85	94

Table 4. Comparison of the precision attainable with different combinations of P_1, P_2, F_1, B_1, B_2 and F_2 groups. Optimal allocation to groups minimises $V(\hat{A}) + V(\hat{H})$

Optimal percent of total animals allocated to groups						Relative size of standard errors of		Relative n. of animals for equal precision to optimal case	Correlation between estimates of $\hat{A}$ and $\hat{H}$
P_1	P_2	F_1	B_1	F_2	B_2	$\hat{A}$	$\hat{H}$		
35	35	31	–	–	–	100	100	100	0.00
37	23	–	41	–	–	111	191	231	0.17
29	29	–	–	41	–	109	188	223	0.00
23	37	–	–	–	41	111	191	231	−0.17
22	–	30	–	48	–	335	214	826	−0.80
22	–	35	–	–	43	170	145	254	−0.51
17	–	–	47	37	–	366	385	1405	−0.76
26	–	–	46	–	29	198	288	587	−0.68
13	–	–	–	49	38	360	558	2111	−0.87
–	–	19	38	43	–	370	254	1047	−0.38
–	–	31	35	–	35	200	200	400	0.00
–	–	19	–	43	38	370	254	1047	0.38
28	28	–	22	–	22	102	186	212	0.00
27	31	–	–	30	13	108	187	221	−0.05

Table 5. The effects of optimal allocation for the breeding structure including maternal effects and based on introduction of an exotic breed solely through the sire line

Optimal percent of total animals allocated to groups							Relative[1] size of SE of		Relative[2] no. of animals	Correlation between estimates of A and H
P_1	P_2	F_1^{b}	B_1^{a}	B_1^{b}	B_2^{a}	F_2	$\hat{A}$	$\hat{H}$		
14	11	11	31	0	32	0	210	217	454	−0.54

[1] Relative to 100 for the optimal design without maternal effects

[2] Relative number of animals for precision equal to that in three-point (P_1, P_2 and F_1) design

Quite small reallocation yields substantial benefits in terms of information on lack of fit. Even allocation of 20% of the resources to F_2 increases the size of the standard error of $\hat{A}$ (and $\hat{H}$) by only 12%, while giving a standard error of the epistatic effect of comparable size to that of the other two effects. However, the expected absolute values of $\hat{A}$ and $\hat{H}$ will typically be several times as large as the expected value of $\hat{E}$. Equal standard errors, therefore, imply relatively greater precision in the estimation of A and H than for E.

Turning to the question of the precision of experiments which are limited to certain subsets of the six possible groups, Table 4 shows the effect on the precision of estimation of A and H of using different combinations of the six possible breeding groups in the experiment. For each design, optimal allocation (i. e. leading to minimum $V(\hat{A}) + V(\hat{H})$) is assumed. The second-last column gives the relative scale of experiment (i.e. number of total animals) required to give precision equal to that obtainable with the optimal (P_1, P_2, F_1) design.

It can be seen that in all cases this optimal three-point design is considerably more efficient than any other combination. The next best design requires at least twice the resources to give the same precision. Furthermore, the correlation between estimates of A and H can be very high, particularly for the asymmetric groupings (P_1, F_1, F_2), (P_1, F_1, B_2), (P_1, B_1, F_2), (P_1, B_1, B_2) and (P_1, F_2, B_2). It is zero for the three balanced groupings (P_1, P_2, F_2), (F_1, B_1, B_2) and (P_1, P_2, B_1, B_2).

There are two components to this reduction in precision. Design 3 (P_1, P_2, F_2) more than doubled the sum of the variances of the estimates of A and H but had zero correlation between them, the increase coming from structure alone. In designs with correlated estimates, the standard errors of $\hat{A}$ and $\hat{H}$ tend to increase with the size of the correlation. This reflects the decreasing level of independent information on either effect having fitted the other. In the limit, the standard errors will tend to infinity as the absolute value of the correlation approaches unity.

The inclusion of maternal effects changes the results considerably. Optimal allocation for the breeding system based on introduction of an exotic breed solely through the sire line was calculated using the method described earlier. The sum of the variances of $\hat{A}$, $\hat{H}$, $\hat{M}$ and $\widehat{MH}$ is taken as the optimising criterion. The percentage allocation is shown in Table 5. By comparison with the optimal three-point (P_1, P_2 and F_1) design, this allocation reduces

Table 6. Correlations between estimates of A, H, M and MH from the optimal design including maternal effects, together with variances of the estimates of the four parameters relative to that of $\hat{A}$

Correlation	$\hat{A}$	$\hat{H}$	$\hat{M}$	MH
$\hat{A}$	1.00	−0.54	0.79	0.45
$\hat{H}$		1.00	−0.65	−0.58
$\hat{M}$			1.00	0.60
MH				1.00
Variances (relative)	100	87	162	29

the emphasis considerably on these three points and shifts resources almost equally to two backcrosses, both having F_1 dams. The reduction in precision in estimation of A and H is considerable, about 4.5 times the replication of the three-point design being required to compensate, and also there is a negative correlation introduced between the estimates of A and H. Table 6 shows the correlations between estimates of A, H, M and MH for this design. Also shown are the variances of these four parameters relative to that of $\hat{A}$ taken as 100.

Conclusions

The general conclusions from this study can be summarised as follows: (1) If epistatic effects can be ignored (often a reasonable assumption) then the purpose of the experiment is to estimate additive (A) and heterotic (H) effects with acceptable precision. The best design for this purpose is the two parental groups ($P_1 + P_2$) and the F_1, with the animals allocated approximately 35%, 35% and 31% to the three groups. (2) The scale of experiment required depends on the coefficient of variation of the trait or traits considered, and on the precision specified in the estimation of A or H. For many practical situations in animal breeding, an experiment involving numbers in the region 250–500 seems reasonable. (3) If this optimal design is not possible, other combinations involving parental, F_1, backcross and F_2 groups can be used to give estimates of A and H. However, even the best of these alternative designs requires at least twice the resources to give the same precision as the optimal design. Some of these designs are highly inefficient, and in many cases give highly correlated estimates of A and H. This high correlation is in turn a further cause of poor precision in the estimation of either parameter. (4) With fixed total experimental resources, the reallocation of resources from the two parental and F_1 groups to any or all of the other three groups (backcrosses and F_2) leads to a reduction in precision. (5) If an additional requirement of the experiment is to estimate the epistatic or recombination effect (E), some reallocation of resources from the parental and F_1 groups to the F_2 group is required. Relatively modest reallocation gives reasonable precision for the estimate of E without serious loss of precision for the estimates of A and H. (6) The inclusion of maternal effects in the model greatly reduces the efficiency of estimation of A and H. Where one of the breeds is introduced through the sire line only, optimal allocation gives roughly equal replication to the pure lines and F_1 but about 63% of allocation is placed in equal amounts on the two backcrosses produced through crossing F_1 dams with pure sires of both breeds.

References

Cunningham EP (1980) The genetic basis of heterosis. Proc 2nd World Congr Genet Appl Livestock Production, 6:190–205

Cunningham EP (1987) Crossbreeding: The Greek temple model. J Anim Breed Genet 104:2–11

Dickerson GE (1969) Experimental approaches in utilising breed resources. Anim Breed Abstr 37:191

Henderson CR (1985) Best linear unbiased prediction for non-additive genetic merits in noninbred populations. J Anim Sci 60:111–117

Koch RM, Dickerson GE, Cundiff LV, Gregory KE (1985) Heterosis retained in advanced generations of crosses among Angus and Hereford cattle. J Anim Sci 60:1117–1132

Nelder JA, Mead R (1965) A simplex method for function minimisation. Computer J 7:308–313

O'Neill R (1971) Function minimisation using a simplex procedure. Appl Statist 20:338–345