Optimization of breeding programmes under index selection and constrained inbreeding

B. VILLANUEVA[1]* AND J. A. WOOLLIAMS[2]

[1] *Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG, Scotland, UK*
[2] *Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland, UK*

Summary

A method for finding optimum breeding schemes which maximize genetic gain under index selection with constraints on the rate of inbreeding is derived. The selection index includes information on the candidate and its sibs. Optimization is for the numbers of males and females to be selected and for the index weights when fixed numbers of offspring per generation, heritabilities and time horizons are considered. The expected rate of gain after a number of generations of selection is combined with the expected asymptotic rate of inbreeding (ΔF) in a single objective function which is maximized for finding the optimum solutions. Under restricted inbreeding, optimum designs are very similar for maximizing gains at different time horizons. The optimum number of selected males (for giving maximum gains) increases with the size of the scheme and with the severity in restricting ΔF and decreases with the heritability. Low heritability, less severe restrictions on ΔF and large schemes lead to increases in the relative weights given to performance of relatives in the index. The presence of common environmental effects leads to increases in optimum mating ratio when the heritability is low, to increases in the number of selected males and to more intense selection within families. Gains from index selection are compared with gains from mass selection. Under restricted inbreeding the advantage of optimized index selection over mass selection is only notable when the heritability is low and the scheme is large (in which case indices put more emphasis on family information than mass selection) and when the heritability is high and the scheme is small (in which case indices put less emphasis on family information).

1. Introduction

In genetic evaluation, information on the performance of relatives of the candidates for selection is used to increase the accuracy of evaluation, and therefore genetic gain. The weights given to records of different relatives in classical selection indices or in BLUP (best linear unbiased prediction) maximize response after a single generation of selection.

The advantage of procedures using information from relatives for improved short-term responses may be offset in the medium or long-term (e.g. Verrier *et al.*, 1993). Firstly, the reduction in genetic variance as a result of linkage disequilibrium generated by selection is higher with more accurate evaluation

methods (Bulmer, 1971; Wray & Hill, 1989; Gomez-Raya & Burnside, 1990*a*), although under the infinitesimal model this alone does not dramatically alter the ranking of schemes in the long-term (Wray & Hill, 1989; Woolliams, 1990; Dekkers, 1992; Villanueva & Kennedy, 1993). Secondly, increasing the weights given to relatives' performance in selection decisions increases the rate of inbreeding (Robertson, 1961), particularly when the heritability is low (Belonsky & Kennedy, 1988) which leads to greater reductions in genetic variation. Furthermore a greater rate of inbreeding will reduce the expected fitness of the population and its expected performance in the selected trait if it exhibits inbreeding depression.

Simple mass selection can give higher gains than direct selection on BLUP estimates, which are the most accurate estimates of breeding values using all available information. This has been observed when selection is practised for many generations in small

---

* Correspondence to be sent to: B. Villanueva, Scottish Agricultural College, Genetics and Behavioural Sciences, Bush Estate, Penicuik EH26 0QE, Scotland UK. Tel: 0131 535 3224; Fax: 0131 535 3121; e-mail: b.villanueva@ed.sac.ac.uk.

populations (Verrier *et al.*, 1993). Also, the advantage of mass selection over BLUP selection after several generations has been shown when the two procedures are compared at the same level of inbreeding (Quinton *et al.*, 1992). The comparisons of within-family and mass selection by Dempfle (1975) showed that long-term response may be greater for within-family selection, particularly when the correlation between the selection criterion and the breeding value is high and selection is intense.

Selection methods that are more sophisticated than within-family and mass selection have been proposed for reducing the emphasis given to family information in the selection criterion and thus reducing the rate of inbreeding with a minimal loss in response (Toro & Perez-Enciso, 1990; Verrier *et al.*, 1993; Grundy *et al.*, 1994; Villanueva *et al.*, 1994; Luo *et al.*, 1995). However, there is no guarantee that the weights given to information from different relatives in these procedures are optimal for maximizing genetic gain after several generations of selection. Optimum index weights for maximizing asymptotic response once the Bulmer equilibrium has been approached were studied by Gomez-Raya & Burnside (1990*b*) by using a selection index which incorporates parental information. The benefits in response from using optimum rather than classical weights were very small. However, they assumed an infinite population and so there was no accumulation of inbreeding.

Wray & Goddard (1994) and Brisbane & Gibson (1995) have proposed methods for selecting individuals aimed at maximizing response when imposing a cost on the inbreeding coefficient. These are *a posteriori* procedures applied once the breeding scheme is in operation and they are effective for offsetting response and inbreeding over a fixed time horizon. Their approach is, however, tied to an identified base population (Villanueva *et al.*, 1996) and the weighting given to family information is not explicitly calculated.

Finding the optimum numbers of individuals to be selected for giving maximum gains is an important problem when designing selection programmes *a priori*. Villanueva *et al.* (1996) have described a procedure for obtaining the optimum numbers of sires and dams for maximizing genetic gain over a specified time period when specific constraints on the rate of inbreeding are imposed. Their method was applied to situations where selection is based on the individual phenotype (i.e. mass selection).

In this paper, this procedure is extended to a situation where selection is on an index which includes information on the individual and its collateral relatives. In addition to the numbers of selected animals, the index weights are optimized by maximizing a single objective function which combined rates of genetic gain and inbreeding. Specific constraints on the rate of inbreeding are considered.

2. Methods

(i) *Model*

The trait under selection is assumed to be determined by an infinite number of additive loci, each with infinitesimal effect (infinitesimal model). A population with discrete generations is assumed with a hierarchical mating structure, in which d dams are mated to each sire and each dam has n offspring. Repeated cycles of directional and truncation selection are practised. Each generation, N_s males and N_d females are selected from $N/2$ males and $N/2$ females scored. The numbers of selected males and females are optimized for each breeding scheme as described later. The number of offspring per mating is determined by N and N_d ($n = N/N_d$). Male and female selection intensities are constant across generations. Mating of selected individuals is at random. A constant number of individuals born per generation (N) is assumed. Different values for N and for the initial heritability ($h^2_{(0)}$) are considered.

Generation 0 consists of unselected individuals with the appropriate family structure. Generation 1 is obtained from the mating of individuals selected at generation 0. The total phenotypic variance at generation t is

$$\sigma^2_{P(t)} = \sigma^2_{A(t)} + \sigma^2_C + \sigma^2_E,$$

where $\sigma^2_{A(t)}$ is the additive genetic variance, σ^2_C is the variance attributed to the common environment of full-sibs and σ^2_E is the individual environmental variance. Environmental variances (σ^2_C and σ^2_E) are constant across generations.

The selection index (I) used as the selection criterion is

$$I = b_1(P - \bar{P}_F) + b_2(\bar{P}_F - \bar{P}_H) + b_3 \bar{P}_H,$$

where P is the record of the individual, $\bar{P}_F$ is the mean of n (including the individual) full-sib records, $\bar{P}_H$ is the mean of dn (including the individual and its full-sibs) half-sib records and b_1, b_2 and b_3 are the index weights. This form was used by Wray *et al.* (1994) for predicting the rate of inbreeding under index selection. Note that mass selection is a special case of this index where $b_1 = b_2 = b_3 = 1$. Index weights are assumed constant across generations. Optimum numbers of parents and index weights are obtained by maximizing a single function which combines the expected rates of genetic progress and inbreeding.

(ii) *Prediction of rate of genetic progress*

Prediction of response is obtained (for a given set of N, N_s, N_d, b_1, b_2 and b_3 values) accounting for reduction in genetic variance due to linkage disequilibrium and due to inbreeding. Let $G_{(t)}$ be the average genetic mean of individuals born at generation

t. The rate of genetic gain obtained each generation ($\Delta G_{(t)} = G_{(t)} - G_{(t-1)}$) is

$$\Delta G_{(t)} = i\rho_{(t-1)}\sigma_{A(t-1)},$$

where $i = (0.5)(i_m + i_f)$ and i_m and i_f are the selection intensities (i.e. standardized selection differentials) for males and females, respectively, and ρ is the accuracy of selection.

The genetic variance is decomposed into between-sire family (σ_{As}^2), between-dam family (σ_{Ad}^2) and within-full-sib family (σ_{Aw}^2) components:

$$\sigma_{A(t)}^2 = \sigma_{As(t)}^2 + \sigma_{Ad(t)}^2 + \sigma_{Aw(t)}^2$$

(Wray & Hill, 1989). These components are obtained each generation by using recurrently the following equations (e.g. Verrier *et al.*, 1990):

$$\sigma_{As(t)}^2 = (0.25)[1-(1/N_s)][1-k_m\rho_{(t-1)}^2]\sigma_{A(t-1)}^2,$$

$$\sigma_{Ad(t)}^2 = (0.25)[1-(1/N_d)][1-k_f\rho_{(t-1)}^2]\sigma_{A(t-1)}^2,$$

$$\sigma_{Aw(t)}^2 = (0.5)[1-F_{(t-1)}]\sigma_{A(0)}^2,$$

where $k_y = i_y(i_y - x_y)$, for $y = m$ (males) of f (females), x_y is the standardized deviation of the truncation point from the mean, $\sigma_{A(0)}^2$ is the value for the genetic variance in the unselected base population and $F_{(t)}$ is the average inbreeding coefficient at generation t. The average coefficient of inbreeding is computed as

$$F_{(t)} = 1-(1-\Delta F)^t,$$

where ΔF is the inbreeding rate. The asymptotic rate of inbreeding (see below) is used to obtain $F_{(i)}$ at each generation.

The accuracy of selection is computed each generation as

$$\rho_{(t)} = \sigma_{AI(t)}/\sigma_{A(t)}\sigma_{I(t)},$$

where $\sigma_{AI(t)}$ is the covariance between the true breeding value and the index and $\sigma_{I(t)}$ is the standard deviation of the index. These are obtained from

$$\sigma_{AI(t)} = b_1\sigma_{Aw(t)}^2\left(1-\frac{1}{n}\right)$$

$$+ b_2\left(\sigma_{Ad(t)}^2 + \frac{\sigma_{Aw(t)}^2}{n}\right)\left(1-\frac{1}{d}\right)$$

$$+ b_3\left(\sigma_{As(t)}^2 + \frac{\sigma_{Ad(t)}^2}{d} + \frac{\sigma_{Aw(t)}^2}{dn}\right)$$

and

$$\sigma_{I(t)}^2 = b_1^2(\sigma_{Aw(t)}^2 + \sigma_E^2)\left(1-\frac{1}{n}\right)$$

$$+ b_2^2\left(\sigma_{Ad(t)}^2 + \sigma_C^2 + \frac{\sigma_{Aw(t)}^2 + \sigma_E^2}{n}\right)\left(1-\frac{1}{d}\right)$$

$$+ b_3^2\left(\sigma_{As(t)}^2 + \frac{\sigma_{Ad(t)}^2 + \sigma_C^2}{d} + \frac{\sigma_{Aw(t)}^2 + \sigma_E^2}{dn}\right).$$

(iii) *Prediction of rate of inbreeding*

The asymptotic rate of inbreeding (ΔF) was calculated for a given set of N, N_s, N_d, b_1, b_2 and b_3 values using the expression

$$\begin{aligned}\Delta F = \Delta F_E + &\{i^2\alpha_m[(0.25)(\tau+\tau_m)^2S_\infty^2 - \tau_m^2]\\ &\times[(16N_s)^{-1}+(16N_d)^{-1}]\}\\ &+ i^2\alpha_f[(0.25)(\tau+\tau_f)^2S_\infty^2 - \tau_f^2][8N_d]^{-1}\\ &+ Hi^2[2\alpha_m+(1+N_sN_d^{-1})\alpha_f]\{\tau^2B_\infty(S_\infty-1)\\ &\times[(32N_s)^{-1}+(32N_d)^{-1}-(4N)^{-1}]\}\\ &+ Hi^2[2\alpha_m+(1+N_sN_d^{-1})\alpha_f]\{\tau^2S_\infty^2(B_\infty-1)\\ &\times[(32N_s)^{-1}+(32N_d)^{-1}-(8N)^{-1}]\},\end{aligned}$$

where ΔF_E is the rate of inbreeding assessed assuming independent generations of selection which treats genetic covariances among sibs and between parents and offspring as though they were of environmental origin (i.e. using 'one generation methods'; Wray *et al.*, 1990). In this paper ΔF_E was calculated using eigenvalue methods as described in Appendix A. The other terms are: H, the correction for hypergeometric sampling that is of the form $[1-(0.5)N_s^{-1}-(0.5)N_d^{-1}]$ (appendix 4 of Woolliams *et al.*, 1993); τ_x, twice the regression of the index of the offspring on the breeding value of the parent of sex x ($\tau_m = b_3$ and $\tau_f = b_2(1-d^{-1})+b_3d^{-1}$); $\tau = (0.5)(\tau_m+\tau_f)$; S_∞ and B_∞ are infinite sums equal to $(1-c)^{-1}$ and $[1-(0.5)c^2]$, respectively, where $c = (0.5)(1-\tau k\beta)$. In c, $k = (0.5)(k_m+k_f)$ and β is the regression of the index on breeding values amongst offspring. Finally, α_x is $(0.25)\sigma_{Ax(1)}^2/\sigma_{I(2)}^2$ where $\sigma_{Ax(1)}^2$ is the additive genetic variance after selection in sex x in generation 1.

The method is an index analogy to formula [3] of Woolliams & Thompson (1994) and was used to produce table 1 in that paper. Further validation is given in Appendix B. The accuracy shown is good (up to 8% errors) with an average error of 2.7%.

(iv) *Optimization of breeding schemes*

Optimum schemes are those giving the highest genetic gain for a given rate of inbreeding. The rate of genetic progress at a given generation and the asymptotic rate of inbreeding are obtained for each set of N and $h_{(0)}^2$ values (fixed parameters) as described above. If the objective is to maximize gain over several generations, the average rate of genetic progress from generation $t-1$ to generation $t+m$ ($\Delta\bar{G}_{(t,t+m)}$) is calculated simply as

$$(\Delta G_{(t)} + \Delta G_{(t+1)} + \ldots + \Delta G_{(t+m)})/(m+1).$$

Then optimum schemes are found by maximizing a single objective function (Φ) which combines the rates of genetic gain and inbreeding:

$$\Phi_{(t,t+m)} = \Delta\bar{G}_{(t,t+m)} - \lambda\Delta F.$$

Table 1. *An example of the maximization procedure for $N = 200$, $h^2_{(0)} = 0.3$ and $\Phi_{(5,20)} = \Delta \bar{G}_{(5,20)} - \lambda \Delta F$. Hence, for a restriction of $\Delta F \leqslant 1\%$, the scheme for $\lambda = 7.4$ would be expected to give the greatest value of $\Delta \bar{G}_{(5,20)}$ by using 30 sires (N_s) with a mating ratio (d) of 1 and a relative weight ($b_2 = b_3$) of 1.04 for the family means*

λ	$\Phi_{(5,20)}$	$\Delta \bar{G}_{(5,20)}$	ΔF	N_s	d	$b_2 = b_3$
0.0	0.322	0.322	0.03179	16	1	1.63
1.0	0.295	0.318	0.02336	19	1	1.47
2.0	0.274	0.312	0.01910	21	1	1.33
3.0	0.256	0.304	0.01612	23	1	1.25
7.3	0.201	0.276	0.01020	29	1	1.01
7.4	0.200	0.273	0.00986	30	1	1.04
55.6	−0.009	0.132	0.00253	67	1	0.74
55.7	−0.009	0.130	0.00249	68	1	0.76

This function is denoted as $\Phi_{(t)}$ when $m = 0$ (i.e. when the aim is to maximize the rate of genetic gain at a single generation t). The parameter λ is a Lagrangian multiplier taking positive values and is increased at appropriate intervals until the constraint on ΔF is satisfied. This is then equivalent to maximizing genetic gain with an upper bound on the rate of inbreeding (Woolliams & Thompson, 1994). The procedure has been illustrated in detail by Villanueva *et al.* (1996) for schemes under mass selection.

The index weights (b_1, b_2 and b_3) can be arbitrarily scaled without changing the selection process, so the weight corresponding to the deviation of the individual from the full-sib family mean (b_1) is set to 1. Then, for a given combination of fixed parameters (λ, N and $h^2_{(0)}$) the objective function $\Phi_{(t,t+m)}$ depends on four variables which are N_s, d, b_2 and b_3. The problem is to find the combination of values of these variables which gives the highest value for $\Phi_{(t,t+m)}$. For each possible set of N_s and d values, the optimum index weights were obtained by using the NAG routine E04UCF (The Numerical Algorithms Limited, 1991). When $d = 1$ all sibs are full-sibs and then the selection index is

$$I = b_1(P - \bar{P}_F) + b_2 \bar{P}_F \text{ (i.e. } b_2 = b_3).$$

In this situation there are only three variables to be optimized (N_s, d and b_2). For each possible combination of N_s and d values, the optimum index weight was found by using a golden section search in one dimension (e.g. Press *et al.*, 1992).

All possible combinations of N_s and d values (using the optimum weights for each combination) were compared and that set giving the highest value for $\Phi_{(t,t+m)}$ was defined as the optimum. The number of offspring per mating (n) was allowed to be non-integer. Table 1 gives an example of how optimum schemes for maximizing gain under restricted inbreeding were obtained.

Initially, possible values for N_s are between 1 and $N/2$ and possible values for d are between 1 and $N/2N_s$. However, prediction of the rate of inbreeding is inaccurate if selection intensities become very extreme (i.e. with very small numbers of selected individuals). Thus, the minimum number of sires was set to 10, 20 and 30 for N equal to 200, 800 and 3200, respectively (i.e. smaller scored population sizes allowed to have a lower minimum N_s for accurate prediction).

(v) *Selection limits with the infinitesimal model*

The optimization procedure can be used to find optimum selection proportions and index weights for maximizing ultimate response at the selection limit. Robertson (1960), Jódar & López-Fanjul (1977) and Cockerham & Burrows (1980) found that the optimum selection proportion for obtaining the maximum advance at the limit is the same in both sexes ($d = 1$) and equals $1/2$. These expectations ignored the Bulmer effect and assumed inbreeding rates only appropriate for populations under random selection. Under mass selection, when using ΔF appropriate for selected populations and accounting for the Bulmer effect, the optimal proportions are somewhat higher (Woolliams & Pong-Wong, 1995).

Optimum schemes (selected proportions and index weights) for maximizing ultimate response under index selection were obtained as described in the previous section but using the objective function $G_{(t)} = \Delta G_{(1)} + \ldots + \Delta G_{(t)}$ and choosing a value of t such that $G_{(t)} - G_{(t-1)}$ is less than 0.01.

An explicit expression for the selection plateau ($G_{(\infty)}$) can be obtained when selection is based only on the individual's own measurement. Under mass selection, the rate of genetic response at any generation t can be approximated as

$$\Delta G_{(t)} \approx \frac{h^2_{(0)} i (a^{t+1} - b^{t+1})}{2 \sigma_{P(L)} a^3 (a - b)},$$

where $b = (0.5)(1 - k h^2_{(L)})$, $a = (1 - \Delta F)/\{1 - [0.25] h^2_{(0)} \Delta F]/\sigma^2_{P(L)}\}$ and $h^2_{(L)}$ and $\sigma_{P(L)}$ represent Bulmer equilibrium values for the heritability and phenotypic standard deviation (Villanueva *et al.*, 1996). Cumulative gain at the selection limit ($t \to \infty$) is then

$$G_{(\infty)} = \sum_{t=1}^{\infty} \Delta G_{(t)} \approx \frac{h^2_{(0)} i [a^2 (1 - a)^{-1} - b^2 (1 - b)^{-1}]}{2 \sigma_{P(L)} a^3 (a - b)}.$$

(vi) *Computer simulation*

Stochastic simulation was used to test some of the results obtained. In general, the simulation procedure used was that described in Villanueva *et al.* (1996) for mass selection with some modifications: (1) With index selection an extra generation (generation 00) needed to be generated to create the base generation

with family structure (generation 0). Generation 00 was constituted by N_s males and N_d females and these unrelated individuals were mated at random to create generation 0. (2) The phenotypic value of an individual was generated as the sum of its genetic value, an environmental component common to its full-sibs and an individual environmental component. The common and individual environmental components were obtained from normal distributions with mean zero and variance σ_c^2 and $1 - h_{(0)}^2 - \sigma_C^2$, respectively. Environmental variances were maintained constant over generations. (3) The selection criterion was the selection index described above. (4) Five thousand replicates were run for each simulation.

3. Results

(i) Optimum schemes under index selection with $\sigma_c^2 = 0$

(a) Number of individuals to be selected

Table 2 shows the optimum numbers of males to be selected and optimum mating ratios for obtaining maximum rate of genetic gain at generation 5 ($\Delta G_{(5)}$) or 20 ($\Delta G_{(20)}$) under two constraints on the rate of inbreeding ($\Delta F \leqslant 1\%$ and $\Delta F \leqslant 0.25\%$) and for obtaining maximum $\Delta G_{(20)}$ with unrestricted inbreeding ($\lambda = 0$). The optimum number of selected males increased with the size of the scheme and with more severe constraints on ΔF and decreased with the heritability (over the range of $h_{(0)}^2$ considered). The optimum mating ratio was equal to 1 except for the larger schemes with less severe restrictions on ΔF and greater heritabilities. In these cases the optimum d was 2. The optimum number of offspring per mating ranged from around 3 ($N = 200$, $h_{(0)}^2 = 0.1$, $\Delta F \leqslant 0.25\%$) to around 43 ($N = 3200$, $h_{(0)}^2 = 0.6$, $\lambda = 0$).

With unrestricted inbreeding ($\lambda = 0$) the optimum N_s increased substantially when maximizing $\Delta G_{(20)}$

compared with maximizing $\Delta G_{(5)}$. For $t = 5$, the optimum N_s was constrained by the imposed lower bound on N_s (see Methods) and results are not shown in Table 2. With ΔF restricted, the optimum N_s for maximizing gain was similar for the two time points chosen, with identical mating ratios. Maximization of the total gain from generation 5 to 20 was also examined, and optimum N_s was always between the optimum numbers for maximizing $\Delta G_{(5)}$ and $\Delta G_{(20)}$ (results not shown).

Male selection intensity (i_m) increased with N and $h_{(0)}^2$ and decreased with the severity of the restriction on ΔF and ranged from 0.51 ($N = 200$, $h_{(0)}^2 = 0.1$, $\Delta F \leqslant 0.25\%$) to 2.37 ($N = 3200$, $h_{(0)}^2 = 0.6$, $\lambda = 0$). Female selection intensities (i_f) followed the same trends although they did not always increase with $h_{(0)}^2$ and N because of shifts in d. Selection intensity in females ranged from 0.51 ($N = 200$, $h_{(0)}^2 = 0.1$, $\Delta F \leqslant 0.25\%$) to 2.09 ($N = 3200$, $h_{(0)}^2 = 0.6$, $\lambda = 0$).

(b) Index weights

The relative weights given to sib information decreased with heritability over the range considered and with more severe restrictions on ΔF and increased with the size of the scheme (Table 3). The weights given to family information when maximizing $\Delta G_{(20)}$ under unrestricted inbreeding ($\lambda = 0$) were considerably lower than the classical index weights which maximize one-generation gain ($t = 1$) for the same values of N_s and d.

For high $h_{(0)}^2$ and large N and for the time points chosen, the optimum weights were close to those corresponding to mass selection ($b_1 = b_2 = b_3 = 1$). For high $h_{(0)}^2$ and small N, the optimum weights move further towards within-family selection and this is potentiated by restricting ΔF.

In general, the index weights corresponding to sib information were slightly higher when maximizing

Table 2. *Optimum number of sires (N_s) and mating ratios (d) under index selection for maximizing genetic gain at generations $t = 5$ and $t = 20$ under different constraints on the rate of inbreeding (ΔF), heritabilities ($h_{(0)}^2$) and scored population sizes (N)*

| | | $\lambda = 0$ | | $\Delta F \leqslant 1\%$ | | | | $\Delta F \leqslant 0.25\%$ | | | |
| | | $t = 20$ | | $t = 5$ | | $t = 20$ | | $t = 5$ | | $t = 20$ | |
$h_{(0)}^2$	N	N_s	d	N_s	d	N_s	d	N_s	d	N_s	d
0.1	200	22	1	32	1	32	1	69	1	69	1
	800	46	1	59	1	61	1	126	1	126	1
	3200	91	1	107	1	111	1	233	1	235	1
0.3	200	21	1	29	1	30	1	68	1	68	1
	800	29	2	35	2	37	2	116	1	117	1
	3200	53	2	58	2	61	2	140	2	142	2
0.6	200	17	1	24	1	25	1	65	1	65	1
	800	22	2	27	2	29	2	96	1	98	1
	3200	37	2	38	2	41	2	107	2	109	2

Table 3. *Optimum index weights (b_2 and b_3) when maximizing genetic gain at generations $t = 5$ and $t = 20$ under different constraints on the rate of inbreeding (ΔF) and index weights obtained under standard selection index theory ($t = 1$) for schemes of different sizes (N) and heritabilities ($h^2_{(0)}$). Index weight b_1 is equal to 1 in all cases. Index weight b_3 is shown (in brackets) only in the cases where it is different from b_2 (i.e. when $d \neq 1$)*

| | | $\lambda = 0$ | | | | $\Delta F \leqslant 1\%$ | | | | $\Delta F \leqslant 0{\cdot}25\%$ | | | |
| | | $t = 1$ | | $t = 20$ | | $t = 5$ | | $t = 20$ | | $t = 5$ | | $t = 20$ | |
$h^2_{(0)}$	N	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)
0·1	200	6·83		2·12		1·53		1·60		1·07		1·07	
	800	9·60		2·97		2·29		2·43		1·50		1·52	
	3200	12·69		3·90		3·26		3·43		2·24		2·28	
0·3	200	3·93		1·43		0·97		1·06		0·75		0·76	
	800	3·56	(4·66)	1·87	(1·82)	1·52	(1·37)	1·63	(1·54)	0·95		0·98	
	3200	4·39	(5·15)	2·26	(2·09)	2·00	(1·80)	2·14	(1·98)	1·50	(1·34)	1·53	(1·38)
0·6	200	2·11		0·78		0·48		0·54		0·45		0·45	
	800	2·06	(2·23)	1·17	(0·90)	0·97	(0·67)	1·06	(0·77)	0·47		0·49	
	3200	2·20	(2·29)	1·26	(0·97)	1·11	(0·81)	1·22	(0·93)	0·93	(0·63)	0·97	(0·66)

Table 4. *Rates of inbreeding $\times 100$ (ΔF) and genetic gain ($\sigma_{P(0)}$ units) at generation 20 ($\Delta G_{(20)}$) when maximizing $\Delta G_{(20)}$ under unrestricted inbreeding ($\lambda = 0$) and rates of genetic gain at generations 5 ($\Delta G_{(5)}$), 20 and average gain between generations 5 and 20 ($\Delta \bar{G}_{(5,20)}$) when maximizing respectively $\Delta G_{(5)}$, $\Delta G_{(20)}$ and $\Delta \bar{G}_{(5,20)}$ under two constraints on ΔF, different heritabilities ($h^2_{(0)}$) and scored population sizes (N)*

| | | $\lambda = 0$ | | $\Delta F \leqslant 1\%$ | | | $\Delta F \leqslant 0{\cdot}25\%$ | | |
$h^2_{(0)}$	N	ΔF	$\Delta G_{(20)}$	$\Delta G_{(5)}$	$\Delta G_{(20)}$	$\Delta \bar{G}_{5,20}$	$\Delta G_{(5)}$	$\Delta G_{(20)}$	$\Delta \bar{G}_{5,20}$
0·1	200	2·09	0·109	0·113	0·100	0·107	0·049	0·047	0·048
	800	1·73	0·166	0·181	0·159	0·169	0·118	0·114	0·116
	3200	1·48	0·226	0·250	0·223	0·236	0·185	0·179	0·182
0·3	200	2·00	0·278	0·292	0·258	0·274	0·132	0·128	0·130
	800	1·54	0·393	0·430	0·385	0·407	0·299	0·290	0·294
	3200	1·25	0·507	0·560	0·504	0·532	0·439	0·426	0·433
0·6	200	1·96	0·525	0·546	0·490	0·517	0·252	0·245	0·248
	800	1·48	0·715	0·779	0·702	0·741	0·558	0·542	0·549
	3200	1·16	0·894	0·995	0·892	0·939	0·794	0·772	0·781

$\Delta G_{(20)}$ than when maximizing $\Delta G_{(5)}$ under specific restrictions on ΔF. If the numbers of selected individuals were fixed, the contrary would be expected (long-term response would be more affected by accumulation of inbreeding and therefore less weight would be given to family information). However, the numbers of individuals to be selected were optimized here and they differ when maximizing gain at different generations. Selection intensities were slightly higher when maximizing early responses under restricted ΔF (Table 2).

(c) *Rates of genetic gain and inbreeding*

Genetic gain at generations 5 and 20 and average response between generations 5 and 20 under different restrictions on ΔF are shown in Table 4. Genetic progress and rate of inbreeding when maximizing

$\Delta G_{(20)}$ with unrestricted inbreeding ($\lambda = 0$) are also presented. With unrestricted inbreeding, ΔF was relatively constant for a given value of N. As expected, the rate of genetic progress increased as N and $h^2_{(0)}$ increased and decreased when restrictions on ΔF were imposed. When maximizing $\Delta G_{(20)}$ the proportional reduction in response below the maximum possible ($\lambda = 0$) was small with the least severe restriction on ΔF ($\Delta F \leqslant 1\%$), ranging from 0·2% ($N = 3200$, $h^2_{(0)} = 0{\cdot}6$) to 8·3% ($N = 200$, $h^2_{(0)} = 0{\cdot}1$). With the most severe restriction ($\Delta F \leqslant 0{\cdot}25\%$), gain was from 13·4% to 53·0% lower than that obtained for $\Delta F \leqslant 1\%$. The loss in gain from restricting inbreeding was greater when short-term response ($t = 5$) was maximized judged by the comparisons of response for $\Delta F \leqslant 1\%$ and $\Delta F \leqslant 0{\cdot}25\%$.

Higher rates of gain were obtained when maximizing $\Delta G_{(5)}$ than when maximizing $\Delta G_{(20)}$ due to the higher

Table 5. *Average rate of genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 20 ($\Delta\bar{G}_{(5,20)}$) achieved with optimum number of sires (N_s), mating ratios (d) and index weights (b_2 and b_3) when the rate of inbreeding is constrained to 1% for different heritabilities ($h^2_{(0)}$), scored population sizes (N) and common environment variances (σ^2_C). Index weight b_1 is equal to 1 in all cases. Index weight b_3 is shown (in brackets) only in the cases where it is different from b_2 (i.e. when $d \neq 1$)*

		$\sigma^2_C = 0.05$					$\sigma^2_C = 0.20$				
$h^2_{(0)}$	N	$\Delta G_{(5,20)}$	N_s	d	b_2	(b_3)	$\Delta G_{(5,20)}$	N_s	d	b_2	(b_3)
0·1	200	0·103	31	1	1·33		0·096	28	1	0·80	
	800	0·158	40	2	1·58	(2·09)	0·144	30	3	0·87	(1·42)
	3200	0·212	59	3	1·95	(2·98)	0·189	40	6	0·96	(2·16)
0·6	200	0·524	23	1	0·38		0·572	20	1	0·17	
	800	0·744	32	1	0·41		0·815	24	1	0·14	
	3200	0·935	32	2	0·62	(0·53)	1·019	30	1	0·18	

accumulation of inbreeding in the latter situation. The maximum average gain over generations 5 and 20 was always between the maximum $\Delta G_{(5)}$ and $\Delta G_{(20)}$. The value of λ for achieving maximum gains at $t = 20$ under specific constraints on ΔF ranged from 1·7 ($\Delta F \leqslant 1\%$, $h^2_{(0)} = 0.1$, $N = 3200$) to 107·0 ($\Delta F \leqslant 0.25\%$, $h^2_{(0)} = 0.6$, $N = 200$).

(ii) *Optimum schemes under index selection with $\sigma^2_C \neq 0$*

The effect of including a non-zero common environmental variance in the model on the optimum numbers of individuals to be selected and index weights and on the maximum rates of genetic gain under restricted inbreeding is shown in Table 5. Common environmental effects led to more emphasis on selection within families, more intense selection on sires and, for low heritability and large schemes, less intense selection on females (see also Table 2 for $\sigma^2_C = 0$). Male selection intensities when restricting ΔF to 1% ranged from 1·20 ($N = 200$, $h^2_{(0)} = 0.1$) to 2·44 ($N = 3200$, $h^2_{(0)} = 0.6$) when $\sigma^2_C = 0.2$ compared, respectively, with 1·12 to 2·36 when $\sigma^2_C = 0.0$.

The optimum mating ratio increased when $\sigma^2_C \neq 0$. As in the situation where $\sigma^2_C = 0$, the highest mating ratios were obtained with the least severe restrictions on ΔF (results not shown) and with the largest scored population sizes. However, with common environmental effects the highest d occurred with the lowest heritabilities. With $h^2_{(0)} = 0.1$ and $N = 3200$, the optimum d was as large as 6. Fig. 1 shows the rate of gain at generation 20 with unrestricted inbreeding for different mating ratios and $\sigma^2_C = 0.0$ or 0·2. In this case, d is fixed and the variables optimized are N_s, b_2 and b_3. Although differences in gain with different mating ratios are small, the optimum d is clearly different from 1 or 2 when $\sigma^2_C = 0.2$.

A further check that $d = 1$ may be a long way from the optimum was made by simulation using two
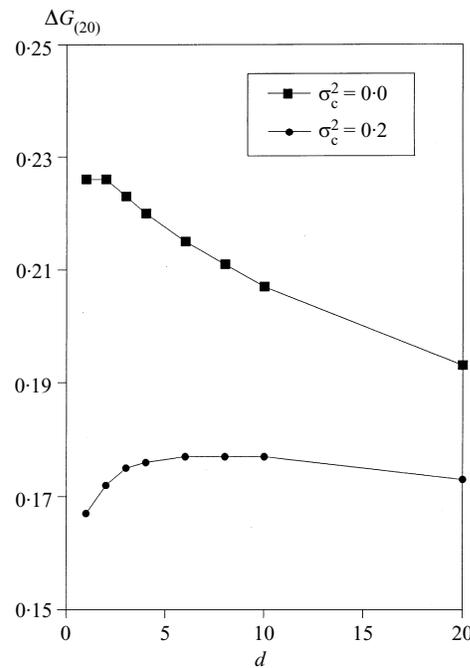
Fig. 1. Effect of mating ratio (d) on rate of genetic gain at generation 20 ($\Delta G_{(20)}$) under unrestricted inbreeding for different values of the variance of common environmental effects (σ^2_c). The total number of individuals scored is 3200, the heritability is 0·1 and the number of selected males and the index weights are optimized for maximizing $\Delta G_{(20)}$ for fixed d.

schemes with equal and integer numbers of offspring per family, $N = 3200$, $h^2_{(0)} = 0.1$, $\sigma^2_C = 0.20$ and $\lambda = 0$ that were compared for $\Delta G_{(20)}$. Scheme A was close to the optimum (as determined by the deterministic model) for fixed $d = 1$ ($N_s = 80$ and $b_2 = b_3 = 1.07$) and scheme B was close to the optimum for unrestricted d ($N_s = 40$, $d = 8$, $b_2 = 1.05$ and $b_3 = 2.79$). The simulated $\Delta G_{(20)}$ were 0·164 and 0·174 for schemes A and B, respectively, which are remarkably close to those predicted (0·167 and 0·176).

Table 6. *Comparison of index selection with mass selection relative to male (D_{im}) and female (D_{if}) selection intensities when maximizing genetic gain at generation 20 under different constraints on the rate of inbreeding (ΔF), heritabilities ($h^2_{(0)}$) and scored population sizes (N). Values presented are differences of results from index selection minus results from mass selection expressed as a percentage of results from mass selection*

$h^2_{(0)}$	N	$\lambda = 0$		$\Delta F \leqslant 1\%$		$\Delta F \leqslant 0.25\%$	
		D_{im}	D_{if}	D_{im}	D_{if}	D_{im}	D_{if}
0·1	200	−7·5	−7·5	−8·2	−8·2	0·0	0·0
	800	−9·6	−9·6	−14·8	−14·8	−8·9	−8·9
	3200	−11·5	−11·5	−15·4	−15·4	−13·7	−13·7
0·3	200	−4·2	−4·2	0·0	0·0	6·0	6·0
	800	−4·0	−6·0	−5·8	−8·3	0·9	0·9
	3200	−5·1	−6·3	−7·7	−9·2	−3·2	−5·2
0·6	200	2·1	2·1	11·4	11·4	21·3	21·3
	800	0·0	0·0	2·2	3·3	12·3	12·3
	3200	0·0	0·0	0·4	0·5	3·8	4·6

Table 7. *Comparison of index selection with mass selection at generation 20 relative to rates of inbreeding ($D_{\Delta F}$) and genetic gain ($D_{\Delta G_{20}}$) when inbreeding is unrestricted ($\lambda = 0$) and at generations 5 and 20 relative to rates of genetic gain ($D_{\Delta G_5}$, $D_{\Delta G_{20}}$) when the rate of inbreeding is restricted to 1% or 0.25% for different heritabilities ($h^2_{(0)}$) and scored population sizes (N). Values presented are differences of results from index selection minus results from mass selection expressed as a percentage of results from mass selection*

$h^2_{(0)}$	N	$\lambda = 0$		$\Delta F \leqslant 1\%$		$\Delta F \leqslant 0.25\%$	
		$D_{\Delta F}$	$D_{\Delta G_{20}}$	$D_{\Delta G_5}$	$D_{\Delta G_{20}}$	$D_{\Delta G_5}$	$D_{\Delta G_{20}}$
0·1	200	31	8	3	4	2	0
	800	54	16	11	12	3	3
	3200	72	25	21	23	10	10
0·3	200	10	1	1	0	0	0
	800	19	3	1	2	0	0
	3200	26	5	4	5	1	1
0·6	200	−6	0	3	2	6	6
	800	−4	0	1	0	3	3
	3200	1	0	1	0	1	1

The emphasis given to family means was notably reduced with the largest σ^2_C ($\sigma^2_C = 0.2$) in comparison with equivalent results for $\sigma^2_C = 0.0$ (see Table 3 for $\sigma^2_C = 0.0$), leading to more emphasis upon selection within families, particularly with high $h^2_{(0)}$. Also, when optimum d is different from 1 and σ^2_C is large ($\sigma^2_C = 0.2$), the weight given to the half-sib family mean was always higher than the weight given to the full-sib mean, putting more emphasis upon the sire information. The rate of gain was in general decreased when including common environmental effects, except for high $h^2_{(0)}$ (see Table 4 for $\sigma^2_C = 0.0$), where gain was greatest with the highest σ^2_C considered ($\sigma^2_C = 0.2$).

(iii) *Mass selection compared with index selection*

(a) *Numbers of individuals to be selected*

A comparison of optimum schemes for mass selection, modelled by setting $b_1 = b_2 = b_3 = 1$, and optimum schemes with indices is shown in Table 6. With unrestricted ΔF, the greatest differences in male selection intensity between the two selection methods were with the lowest heritability and the large scored population sizes. There were no general rules concerning selection intensity: it was higher with mass than with index selection when $h^2_{(0)} = 0.1$, but lower with mass selection when $h^2_{(0)} = 0.6$.

(b) *Rate of inbreeding when maximizing* $\Delta G_{(20)}$ *for* $\lambda = 0$

At low heritabilities, ΔF is substantially higher with index than with mass selection (Table 7). However, at $h^2_{(0)} = 0.6$ mass selection gave greater ΔF for the smallest schemes, despite the fact the optimum selection intensities were lower (Table 6). The higher ΔF obtained with mass selection is due to the fact that under these circumstances the optimum index approaches within-family selection (see Table 3).

Figure 2 shows the effect of the heritability on rates of inbreeding obtained for optimum schemes under mass and index selection. For maximizing gain at generation 20, the optimum schemes under index selection maintained a relatively constant ΔF (around 0.02 for $N = 200$ and 0.012 for $N = 3200$) for all but extreme heritabilities (see also Table 4). As $h^2_{(0)}$ tends to 0 or 1, ΔF was increased. In contrast, for mass selection the optimum schemes for maximizing $\Delta G_{(20)}$ showed ΔF increased from 0.015 to 0.028 for $N = 200$ and from 0.007 to 0.015 for $N = 3200$ as $h^2_{(0)}$ moved from 0 to 1. The crossover points with index selection were around 0.5 for $N = 200$ or 0.6 for $N = 3200$, at which the optimum index schemes were close to mass selection.

Optimizing the scheme design together with the weights (as described above) was contrasted to optimizing only the weights for a specific N_s ($N_s = 20$) and d ($d = 1$). Fig. 3 shows the rates of inbreeding obtained when using optimum weights for maximizing gains at generations, 1, 5 or 20 without a constraint on inbreeding. The curves for I_{20} and M may be directly

Fig. 3. Asymptotic rate of inbreeding (ΔF) under mass (M) and index selection using optimum weights for maximizing genetic gain at generation 1 (I_1), 5 (I_5) or 20 (I_{20}) for fixed N_s and d ($N_s = 20$, $d = 1$) and different heritabilities ($h^2_{(0)}$). The total number of individuals scored is 200.
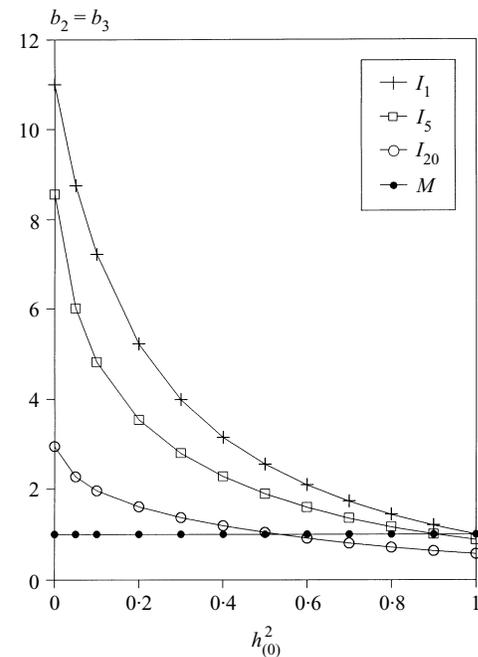
Fig. 4. Optimum weights ($b_2 = b_3$) under mass (M) and index selection for maximizing genetic gain at generations 1 (I_1), 5 (I_5) or 20 (I_{20}) for fixed N_s and d ($N_s = 20$, $d = 1$) and different heritabilities ($h^2_{(0)}$). A weight of 1 is assigned to within-full-sib family deviations ($b_1 = 1$). The total number of individuals scores is 200.

Fig. 2. Asymptotic rate of inbreeding (ΔF) using optimum schemes for maximizing genetic gain at generation 20 under mass (M) and index selection (I) for different scored population sizes (N) and heritabilities ($h^2_{(0)}$).

compared with Fig. 2. Although the curve for I_{20} is more constant than those for I_1 and I_5 it still retains the classical shape for indices characterized by the curve for I_1. Likewise the curve for M is quite different from Fig. 2. Fig. 4 shows for $N_s = 20$ and $d = 1$ the

reduction in the weights given to family information as the time point of interest increases, compared with the classical weights maximizing gain in the next generation.

(c) *Rate of genetic progress*

Table 7 shows for the cases considered that unless the heritability is low ($h^2_{(0)} = 0.1$) and the scheme is large, the benefits from index selection are small. The gains from indices do not always increase with N. For example, when $h^2_{(0)} = 0.6$, the benefit from indices was larger with $N = 200$. This pattern is related to the pattern of optimum weights in Table 3. Since mass selection is an intermediate point between the classical index selection ($b_2 > 1$ and $b_3 > 1$) and within-family selection ($b_2 = b_3 = 0$) it should be expected that the benefits of index selection will be related to what degree the optimum index weights deviate from $b_2 = b_3 = 1$.

When ΔF is unrestricted and conventional weights are used, the advantage of index over mass selection decreases with the number of generations of selection due to the higher accumulation of inbreeding. However, results from Table 7 show that when ΔF is restricted and the weights are optimized, the advantage of index selection can be greater when maximizing long-term responses ($\Delta G_{(20)}$) than when maximizing early responses ($\Delta G_{(5)}$).

(iv) *Selection limits with the infinitesimal model*

The optimum proportions of selected males (p_m) and females (p_f) for maximizing the selection plateau were around 2/3 for all the different values of $h^2_{(0)}$ and N studied. The optimum d was 1 and so $p_m = p_f = p$ and $b_2 = b_3$. For a given heritability, the optimum index weights were approximately constant for different values of N. Optimum b_2 and b_3 were 1.2 for $h^2_{(0)} = 0.1$, 1.0 for $h^2_{(0)} = 0.3$ and 0.7 for $h^2_{(0)} = 0.6$. The theoretical maximum response ranged from 20.9 ($h^2_{(0)} = 0.1$ and $N = 200$) to 2023.1 ($h^2_{(0)} = 0.6$ and $N = 3200$). The number of generations required to approach the limit increased with N. The time scale of the response is proportional to the effective population size (Robertson, 1960), which was larger for large N (Table 4).

Computer simulations were used to check that b_2 and b_3 clearly differ from zero (i.e. optimum selection differs from within-family selection) for maximizing the selection limit. For fixed $N = 100$, $h^2_{(0)} = 0.1$, $N_s = 25$ and $d = 1$, the optimum weights are $b_2 = b_3 = 0.85$ and the predicted $G_{(\infty)}$ (from the analytical procedure) is 9.59. For the same set of parameters but $b_2 = b_3 = 0$, the predicted $G_{(\infty)}$ is 6.12. Four thousand generations were run for both schemes and the simulated $G_{(\infty)}$ were 9.80 (standard error = 0.01) and 6.03 (standard error = 0.01) for $b_2 = b_3 = 0.85$ and $b_2 = b_3 = 0$, respectively.

4. Discussion

The procedure described here optimizes both the numbers selected and the relative index weights for maximizing responses at a set of time points with or without constraints upon the rate of inbreeding. This is distinct from much of the previously published work comparing selection upon indices and phenotypes in which fixed numbers of individuals were selected and where classical weights were given to family information (Belonsky & Kennedy, 1988; Verrier *et al.*, 1993; Wei *et al.*, 1996). Comparison of different methods at similar rates of inbreeding has been an approach adopted previously by, for example, Woolliams (1989) and Wray & Simm (1990), and was used here since ΔF is an intrinsic genetic property of a scheme determining rate of loss of genetic variation, fixation probabilities of mutants under selection (Caballero *et al.*, 1997) and the genetic architecture of inbred chromosomes (Stam, 1980). The constraint on ΔF can be viewed as the measure of genetic risk, and in this context critical values have been advanced by Meuwissen & Woolliams (1994) amongst others.

The study coming closest to the situation considered here was that of Quinton *et al.* (1992), which was based upon stochastic computer simulation. They allowed a variable number of sires and found, for low levels of inbreeding, higher selection responses from mass selection than from selection based upon BLUP when the two procedures were compared over 20 generations at the same rate of inbreeding. Their results from a selection index including the individual, full-sibs and half-sibs were very similar to those obtained from BLUP, and so their results are comparable with those presented here. For a trait with $h^2_{(0)} = 0.25$ (see their table 2), $N = 200$, and a fixed number of dams ($N_d = 50$), with the rate of inbreeding $\approx 1\%$ (cumulative inbreeding = 0.18) the cumulative response was higher for mass selection (4.43 σ_P units) than for index selection (3.73 σ_P units) and the optimum number of sires was higher when selection was based upon BLUP than with mass selection (36 versus 18). Our results show that under these specific conditions mass selection was very close to the optimum since the weights under optimized index selection are $b_1 = 1$ and $b_2 = b_3 = 1.06$, and ΔG and N_s are practically the same for mass and optimized index selection (Tables 6, 7).

BLUP selection can be closely approximated by using a selection index including the estimated breeding values of the sire and the dam and the mean of estimated breeding values of all dams mated to the sire in addition to information on the individual and its sibs (Wray & Hill, 1989). The procedure used here could be extended to BLUP selection by optimizing three extra index weights. Predictions of the rate of inbreeding are needed, however. Optimization becomes more complex as more sources of information are included in the index. However, results from

BLUP selection would be expected to show very similar trends to those presented here, which was the conclusion of Quinton *et al.* (1992).

When optimizing not only numbers selected but also the weights given to family information, methods using information on relatives must be always equal or superior to mass selection. The general framework considered here makes it clear that the adequacy of mass selection for giving gains close to the maximum will depend entirely upon the time horizon, total offspring numbers, heritability and restrictions. For the scenarios examined, benefits of index selection exceeded 5 % only for: low $h_{(0)}^2$ and large N, in which optimum indices placed substantial extra weight on family information compared with mass selection; and high $h_{(0)}^2$ and small N, in which optimum indices placed substantially less weight on the family than mass selection. As restrictions on ΔF become less severe or time horizons become shorter, it would be anticipated that the range of $h_{(0)}^2$ and N for which index selection is beneficial from putting extra weight on the family would expand, whereas the range benefiting from reduced weight on the family would diminish.

The study has re-evaluated and generalized the results of Robertson (1960) and Dempfle (1975) on the selection limits of indices. Robertson (1960) concluded that a selection proportion of 0·5 was optimum for mass selection and that mass selection was always superior to family selection (b_2, $b_3 \gg b_1$). Dempfle (1975) concluded that the selection limit may be greater for within-family selection than for mass selection, and that this was particularly evident when the accuracy of selection was high. In Dempfle's study the squared accuracy is akin to the heritability in this study, since he considered evaluation of the genotype as a unit (e.g. from progeny testing or phenotype) and not as a composite of separate bits of information on ancestors and a Mendelian sampling term.

Results of this study show optimum selection proportions for maximizing the selection limit to be more than 0·5, and equal in both sexes. The discrepancy arises because Robertson used Wright's formula for predicting ΔF, but this underestimates the impact of selection intensity upon rate of inbreeding (Woolliams *et al.*, 1993) and so appropriate modification tends to favour lower selection intensities. Although the infinitesimal model is unrealistic, particularly when considering responses at the limit, it is still a standard model and it is useful as a reference for comparison.

The optimum weights for maximizing the selection limit were close to mass selection with a greater emphasis on family information for low $h_{(0)}^2$ and less for high $h_{(0)}^2$. The results of Dempfle in comparing mass selection with within-family selection are therefore consistent with the results of this study; however, within family selection is not the optimum, and some positive weight should be given to family information (i.e. b_2, $b_3 > 0$). The authors find it remarkable that in maximizing the selection limit the optimum selection proportion was independent of heritability, and the optimum weights were close to mass selection.

When ΔF was restricted and in the absence of common environmental variance, the optimum index selection scheme had: more intense selection on sires and less emphasis on family information as $h_{(0)}^2$ increased up to 0·6; and more intense selection on sires and more emphasis on family information as the severity of the restriction on ΔF was reduced and more resources were available. With restrictions on ΔF, time horizons $t = 5$ or $t = 20$ made only minor modifications to optimum schemes with a small reduction in intensity (Table 2) and a slightly greater emphasis on family information for longer time horizons (Table 3).

In hierarchical schemes with $\sigma_C^2 = 0$ the changes in selection intensity for females in response to variation in scheme parameters were not as smooth as for males, since the optimum mating ratio (which only takes discrete values) changed when selection on males was most intense; i.e. when $h_{(0)}^2$ was at the upper end of the range considered, resources were greater and restrictions on ΔF were less severe. This pattern was also noted for mass selection (Villanueva *et al.*, 1996) and differs from the earlier conclusion of Jódar & López-Fanjul (1977) who predicted maximum gain at any generation with $d = 1$. The discrepancy again lies in the adequacy of the prediction of ΔF. In the latter paper, Wright's formula was used, whereas better predictions, even one-generation predictions (Wray *et al.*, 1990), show that ΔF will depend in part on terms such as $i^2 \phi d^{-1} N_s^{-1}$, where ϕ is the correlation between index values of either paternal half-sibs or full-sibs. Villanueva *et al.* (1996) argued that such terms under the conditions noted above are capable of favouring schemes with increased d. Schemes which are capable of fully factorial mating with equal information on both sexes will have a different outcome since there will also be terms in both $N_s N_d^{-1}$ and $N_d N_s^{-1}$. The complete symmetry of such schemes will lead to optima which are square designs with $N_s = N_d$ (Woolliams, 1989; De Boer & Van Arendonk, 1994).

The presence of common environmental variation increased selection intensity on males, increased the mating ratio for low heritabilities and generally decreased the emphasis on sib information, particularly full-sibs. These trends in the optimal schemes are predictable as they move to the extremes of the range of heritability presented. The value of information from half-sibs can be increased if the common environmental variation is averaged out over full-sib families, whereas the information from a full-sib family is always formally confounded with the common environment. Therefore there is a much greater pressure for d to increase above 1 when $\sigma_C^2 > 0$ (with $b_3 > b_2$) than when $\sigma_C^2 = 0$. A further influence

on the results for $\sigma_C^2 > 0$ that becomes increasingly important as $h_{(0)}^2$ increases and more weight is attached to within-family deviations is that a greater proportion of the variance of within-family deviations is associated with genetic variance than in the case with $\sigma_C^2 = 0$. Thus for the same heritability, the presence of substantial common environmental variance can increase gains in optimum indices under restricted inbreeding.

The formulae used are approximations, but where direct comparison with simulation has been made, excellent agreement was obtained. Further improvements are available. For example, the selection intensity used in the formulae was that appropriate for infinite populations with uncorrelated estimates of breeding values, which is a serious overestimate in small schemes when there are very high correlations in the indices among family members (Hill, 1976; Rawlings, 1976). However, in our case the restrictions on ΔF (even for $\lambda = 0$ when $t = 20$) increased the numbers selected and very much reduced the correlations among family members compared with classical weights (see Table 3) and consequently the impact on selection intensity is small. Applying the approximation of Meuwissen (1991) for $\lambda = 0$, $t = 20$, $h_{(0)}^2 = 0.1$ and $N = 200$, showed the selection intensity was reduced from 1·346 to 1·324. The bias will be more important when maximizing $\Delta G_{(5)}$ for $\lambda = 0$, but corrections can easily be incorporated into the optimization procedure.

Index weights have been assumed constant over generations. However, the similarity in the optimum weights obtained when maximizing response at different generations under specific constraints on ΔF (Table 3) suggests that little improvement would be made by allowing the weights to change each generation.

The optimization procedure also makes the assumption where necessary of non-integer numbers of scored individuals per family and has not accounted for variation in family size in the formulae. However, the deviation of the schemes from a constant-integer family size for the parameters used has only a small impact. The worst case (that with the greatest coefficient of variation) was for $h_{(0)}^2 = 0.1$, $N = 200$, $\Delta F \leqslant 0.25\%$ where $N_s = N_d = 69$. Here, with 62 families with 3 offspring and 7 with 2 offspring the coefficient of variation of family size (CV) is 0·1. For $d = 1$, the rate of inbreeding is related to $E(n^2) = [E(n)]^2[1 + CV^2]$, where n is the family size, but here we have only considered $[E(n)]^2$. The proportional errors in ΔF introduced by neglecting variation in family size are therefore at most 1 % (i.e. CV^2). Other factors such as reproductive limitations may also create variation in family size. Complications of index definition arise in practice if the coefficient of variation in the number of scored individual is large, since indices are no longer identical or uniformly accurate across families.

The study has optimized numbers of parents of both sexes and index weights pre-determined for constraints on total number of offspring per generation, rate of inbreeding and time horizon. The first two constraints represent resources available and the risk attached to the scheme. The latter would reflect the objectives of the scheme, where for example a competitive breeding company may have a short horizon and populations conserved *in situ* would have long horizons. Other restrictions, for example on the expected number scored per family arising from biological constraints on family size, could be added to the framework. The results have practical significance since they show that even when no restrictions are placed on the rate of inbreeding, optimal weights show substantial deviations from classical weights.

Appendix A. Calculation of ΔF_E

Wray *et al.* (1994) give expressions for the correlations of the index values of full-sibs (ρ_D) and half-sibs (ρ_H) for indices of the form used here. These are given by

$$\rho_D = \left\{ -b_1^2 \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} + \left[b_2^2 \left(1 - \frac{N_s}{N_d} \right) \right. \right.$$
$$\left. \times \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right]$$
$$\left. + b_3^2 \left[\sigma_{As}^2 + \frac{N_s}{N_d} \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right] \right\} \sigma_I^{-2}.$$

$$\rho_H = \left\{ -b_2^2 \frac{N_s}{N_d} \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right.$$
$$\left. + b_3^2 \left[\sigma_{As}^2 + \frac{N_s}{N_d} \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right] \right\} \sigma_I^{-2}.$$

The dependence on t has been left implicit above, but otherwise these have an identical form to equations 3 and 4 of Wray *et al.* (1994). In the calculation of ΔF_E the values after a single generation of selection were used for σ_{As}^2 and σ_{Ad}^2 while σ_{Aw}^2 was assumed to be constant.

Appendix 6 of Woolliams *et al.* (1993) uses a result of Mendell & Elston (1974) to show that

Prob $(i, j$ full-sibs $| i$ of sex x, j of sex y, both selected)

$$= [(1/2)n - \delta_{xy}][(1/2)N - \delta_{xy}]^{-1} Q_D p_x^{-1},$$

where $Q_D = \Psi[(i_x \rho_D - \nu_y)(1 - k_x \rho_D)^{-1/2}]$ and Ψ denotes the cumulative normal distribution with zero mean and unit variance; i_x, p_x and k_x are the intensity of selection, proportion selected and variance reduction parameter for selection on sex x, and ν_y is the truncation deviate for selection on sex y; $\delta_{xy} = 1$ if $x = y$ (i.e. i and j same sex) and 0 otherwise and accounts for sampling without replacement. Although the probability is symmetric in x and y, the expression

Table B1. *Predicted rates of inbreeding and percentage errors of predictions*

N_d	n	$h^2_{(0)}$			
		0·0	0·1	0·4	0·99
20	6	2·47 (+1·6)	2·70 (+3·4)	2·34 (+2·2)	1·39 (+1·5)
40	6	2·30 (+0·9)	2·49 (+3·3)	2·03 (+2·0)	1·10 (+1·9)
	12	4·63 (+3·8)	4·77 (+6·2)	3·25 (+2·2)	1·26 (+2·4)
200	6	3·03 (+3·4)	2·78 (+8·2)	1·77 (+0·6)	0·83 (−1·2)
	12	4·64 (+2·7)	3·88 (+5·4)	2·34 (−0·8)	0·89 (+1·1)

is not: both forms are approximations but the use of $x = m$, $y = f$ is found to be more accurate assuming $p_m < p_f$. Similarly,

Prob (i, j half-sibs | i of sex x, j of sex y, both selected)

$$= [(1/2)n - \delta_{xy}][(1/2)N - \delta_{xy}]^{-1}Q_H p_x^{-1},$$

where $Q_H = \Psi[(i_x \rho_H - \nu_y)(1 - k_x \rho_H)^{-1/2}]$.

Problems of predicting co-selection probabilities were encountered by Wray *et al.* (1994) for when both (i) the number of males in a half-sib family was large compared with the number selected ($2N_s < N$ in this notation) and (ii) ρ_H was high. This circumstance was rarely encountered in this study because optimum restricted indices placed less emphasis on family information. Where necessary the solution adopted by Wray *et al.* (1990) was used, in which p_m was replaced by $p'_m = (1 - \rho_H)p_m + \rho_H N_s^{-1}$ for calculating the associate parameters i', k' and ν'.

Wray *et al.* (1990) derived a transition matrix of the form

$$\begin{pmatrix} r_1 & 1/2 & r_2 & (1/2)(1 - 2r_1 - 2r_2) \\ r_3 & 1/2 & r_4 & (1/2)(1 - 2r_3 - 2r_4) \\ r_4 & 1/2 & r_6 & (1/2)(1 - 2r_5 - 2r_6) \\ 0 & 1/2 & 0 & 0 \end{pmatrix}$$

where $r_1 = (1/4)$ Prob (i, j have distinct sires | $i \neq j$, both males, both selected); $r_2 = (1/4)$ Prob (i, j have distinct dams | $i \neq j$, both males, both selected). The terms r_3 and r_5 are defined as r_1 but considering different-sex and female-only pairs respectively; and r_4 and r_6 are defined as r_2 for different-sex and female-only pairs respectively. The terms r_1 and r_2 are

$r_1 = (1/4)$
  $\times [1 - \text{Prob} (i, j \text{ half-sibs} | i, j \text{ both males, both selected})$
  $- \text{Prob} (i, j \text{ full-sibs} | i, j \text{ both males, both selected})]$

$r_2 = (1/4)$
  $\times [1 - \text{Prob} (i, j \text{ full-sibs} | i, j \text{ both males, both selected})]$

and r_3, r_4, r_5 and r_6 can be constructed similarly using the probabilities derived above.

The estimate of ΔF_E which accounts for the co-selection in a single generation, but which does not account for the inheritance of selective advantage, is then calculated by obtaining the largest eigenvalue of the transition matrix. Wray *et al.* (1990) show the relationship of this estimate to a first-order approximation using variances of family sizes given by Hill (1979).

Appendix B. Validation of the method used to compute the rate of inbreeding under index selection

Table B1 shows predicted rates of inbreeding and percentage errors of predictions (in brackets) calculated as $100 \times (P - S)/S$, where P and S represent predicted and simulated values, respectively, for schemes with $N_s = 20$ males and different numbers of females (N_d), numbers of offspring per mating (n) and heritabilities ($h^2_{(0)}$). Simulated values are from Wray *et al.* (1994).

References

Belonsky, G. M. & Kennedy, B. W. (1988). Selection on individual phenotype and best linear unbiased predictor of breeding value in a closed swine herd. *Journal of Animal Science* **66**, 1124–1131.

Brisbane, J. R. & Gibson, J. P. (1995). Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theoretical and Applied Genetics* **91**, 421–431.

Bulmer, M. G. (1971). The effect of selection on genetic variability. *American Naturalist* **105**, 201–211.

Caballero, A., Wei, M. & Hill, W. G. (1996). Survival rates of mutant genes under artificial selection using individual and family information. *Journal of Genetics* **75**, 63–80.

Cockerham, C. C. & Burrows, P. M. (1980). Selection limits and strategies. *Proceedings of the National Academy of Sciences of the USA* **77**, 546–549.

De Boer, I. J. M. & Van Arendonk, J. A. M. (1994). Additive response to selection adjusted for effects of inbreeding in a closed dairy cattle nucleus assuming a large number of gametes per female. *Animal Production* **58**, 173–180.

Dekkers, J. C. M. (1992). Asymptotic response to selection on best linear unbiased predictors of breeding values. *Animal Production* **54**, 351–360.

Dempfle, L. (1975). A note on increasing the limit of selection through selection within families. *Genetical Research (Cambridge)* **24**, 127–135.

Gomez-Raya, L. & Burnside, E. B. (1990*a*). The effect of repeated cycles of selection on genetic variance, heritability and response. *Theoretical and Applied Genetics* **79**, 568–574.

Gomez-Raya, L. & Burnside, E. B. (1990*b*). Optimum selection indices for continuous selection programmes. In *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, vol. XIII, Edinburgh, 23–27 July, pp. 273–276.

Grundy, B., Caballero, A., Santiago, E. & Hill, W. G. (1994). Use of biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. *Animal Production* **59**, 465–468.

Hill, W. G. (1976). Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* **32**, 889–902.

Hill, W. G. (1979). A note on effective population size with overlapping generations. *Genetics* **92**, 317–322.

Jódar, B. & López-Fanjul, C. (1977). Optimum proportions selected with unequal sex numbers. *Theoretical and Applied Genetics* **50**, 57–61.

Luo, Z. W., Woolliams, J. A. & Thompson, R. (1995). Controlling inbreeding in dairy MOET nucleus schemes. *Animal Science* **60**, 379–387.

Mendell, N. R. & Elston, R. C. (1974). Multifactorial qualitative traits: genetic analysis and prediction of recurrence risks. *Biometrics* **30**, 41–57.

Meuwissen, T. H. E. (1991). Reduction of selection differentials in finite populations with a nested full-half sib family structure. *Biometrics* **47**, 195–203.

Meuwissen, T. H. E. & Woolliams, J. A. (1994). Effective sizes of livestock populations to prevent a decline in fitness. *Theoretical and Applied Genetics* **89**, 1019–1026.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd edn, pp. 390–395. Cambridge: Cambridge University Press.

Quinton, M., Smith, C. & Goddard, M. E. (1992). Comparison of selection methods at the same level of inbreeding. *Journal of Animal Science* **70**, 1060–1067.

Rawlings, J. O. (1976). Order statistics for a special class of unequally correlated multinormal variates. *Biometrics* **32**, 875–887.

Robertson, A. (1960). A theory of limits in artificial selection. *Proceedings of the Royal Society of London, Series B* **153**, 234–249.

Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research (Cambridge)* **2**, 189–194.

Stam, P. (1980). The distribution of the fraction of the genome identical by descent in finite random mating populations. *Genetical Research (Cambridge)* **35**, 131–155.

The Numerical Algorithms Limited (1991). *NAG Fortran Library Manual*, Mark 15, 1st edn., Oxford: The Numerical Algorithms Ltd.

Toro, M. & Perez-Enciso, M. (1990). Optimization of selection response under restricted inbreeding. *Genetics, Selection, Evolution* **22**, 93–107.

Verrier, E., Colleau, J. J. & Foulley, J. L. (1990). Predicting cumulated response to directional selection in finite panmictic populations. *Theoretical and Applied Genetics* **78**, 833–840.

Verrier, E., Colleau, J. J. & Foulley, J. L. (1993). Long-term effects of selection based on animal model BLUP in a finite population. *Theoretical and Applied Genetics* **87**, 446–454.

Villanueva, B. & Kennedy, B. (1993). Index versus tandem selection after repeated generations of selection. *Theoretical and Applied Genetics* **85**, 706–712.

Villanueva, B., Woolliams, J. A. & Simm, G. (1994). Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics, Selection, Evolution* **26**, 517–535.

Villanueva, B., Woolliams, J. A. & Gjerde, B. (1996). Optimum designs for breeding programmes under mass selection with an application in fish breeding. *Animal Science* **63**, 563–576.

Wei, M., Caballero, A. & Hill, W. G. (1996). Selection response in finite populations. *Genetics* **144**, 1961–1974.

Woolliams, J. A. (1989). Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49**, 1–14.

Woolliams, J. A. (1990). Strategies to maximise selection progress in dairy cattle. In *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, vol. XIV, Edinburgh, 23–27 July, pp. 15–24.

Woolliams, J. A. & Pong-Wong, R. (1995). Short- versus long-term responses in breeding schemes. In *46th Annual Meeting of the European Association for Animal Production*, Prague, 4–7 September.

Woolliams, J. A. & Thompson, R. (1994). A theory of genetic contributions. In *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, vol. 19, Guelph, 23–27 July, pp. 127–134.

Woolliams, J. A., Wray, N. R. & Thompson, R. (1993). Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genetical Research (Cambridge)* **62**, 231–242.

Wray, N. R. & Goddard, M. E. (1994). Increasing long term response to selection. *Genetics, Selection, Evolution* **26**, 431–451.

Wray, N. R. & Hill, W. G. (1989). Asymptotic rates of response from index selection. *Animal Production* **49**, 217–227.

Wray, N. R. & Simm, G. (1990). The use of embryo transfer to accelerate genetic improvement in beef cattle. In *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, vol. XV, Edinburgh, 23–27 July, pp. 315–318.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1990). Methods for predicting rates of inbreeding in selected populations. *Theoretical and Applied Genetics* **80**, 503–512.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1994). Prediction of rates of inbreeding in populations undergoing index selection. *Theoretical and Applied Genetics* **87**, 878–892.