

# Increasing long-term response to selection

Nr Wray, Me Goddard

▶ **To cite this version:**

Nr Wray, Me Goddard. Increasing long-term response to selection. Genetics Selection Evolution, 1994, 26 (5), pp.431-451. hal-00894043

**HAL Id: hal-00894043**
**https://hal.science/hal-00894043**

Submitted on 11 May 2020

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

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

*Genet Sel Evol* (1994) 26, 431-451
© Elsevier/INRA


**Original article**

# Increasing long-term response to selection

NR Wray *, ME Goddard †

*Livestock Improvement Unit, Victorian Institute of Animal Science, 475, Mickleham Road, Attwood, Victoria 3049, Australia*

(Received 8 September 1993; accepted 25 May 1994)

**Summary** – Selection on estimated breeding value (EBV) alone maximises response to selection observed in the next generation, but repeated use of this selection criterion does not necessarily result in a maximum response over a longer time horizon. Selection decisions made in the current generation have at least 2 consequences. Firstly, they influence the immediate genetic response to selection and, secondly, they influence the inbreeding of the next and subsequent generations. Accumulation of inbreeding has a negative impact on future genetic response through reduction in future genetic variance and a negative impact on future performance if inbreeding depression affects the selected trait. Optimum selection decisions depend on the time horizon of interest. If this is known, then a breeding objective can be defined. A selection criterion is proposed in which the positive contributions of a selected group of parents to immediate genetic response (determined by their average EBV) is balanced against their negative contribution to future genetic response (determined by their contribution to inbreeding). The value assigned to the contribution to inbreeding is derived from the breeding objective. Selection of related individuals will be restricted if the detrimental value associated with inbreeding is high; restrictions on the selection of sibs, however, is flexible from family to family depending on their genetic merit. A selection algorithm is proposed which uses the selection criterion to select sires on 3 selection strategies, to select on i) a fixed number of sires; ii) a variable number of sires each allocated an equal number of matings; or iii) a variable number of sires allocated an optimal proportion of matings. Using stochastic simulation, these selection strategies for sires are compared with selection on EBV alone. When compared at the time horizon specified by the selection goal, the proposed selection criterion is successful in ensuring a higher response to selection at a lower level of inbreeding despite the selection of fewer sires. The selection strategy iii) exploits random year-to-year variations in the availability of individuals for selection and is successful in maximising

* Correspondence and reprints: c/o PM Visscher, Roslin Institute, Roslin, Edinburgh EH25 9PS, Scotland, UK.

† Present address: Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia.

response to the selection goal. The derivation of the value assigned to inbreeding is not exact and cannot guarantee that the overall maximum response is found. However, simulation results suggest that the response is robust to the detrimental value assigned to inbreeding.

**artificial selection / selection response / inbreeding / BLUP / computer simulation**

**Résumé – Accroissement de la réponse à la sélection dans le long terme.** *La sélection sur la valeur génétique estimée (VGE) considérée seule maximise la réponse à la sélection observée dans la génération qui suit, mais l'utilisation répétée de ce critère de sélection ne garantit pas nécessairement la réponse maximale sur une longue période. Les décisions de sélection prises à chaque génération ont au moins 2 conséquences. Elles influencent d'abord la réponse génétique immédiate à la sélection, et ensuite elles déterminent le niveau de consanguinité dans la génération suivante et les générations ultérieures. L'accumulation de la consanguinité a un effet négatif sur la réponse future en réduisant la variance génétique et un effet négatif sur la performance future si le caratère sélectionné subit une dépression de consanguinité. Les décisions de sélection optimales dépendent de la perspective considérée. Si celle-ci est déterminée, alors un objectif de sélection peut être défini. On propose ici un critère de sélection dans lequel la contribution positive d'un groupe de parents sélectionnés à la réponse génétique immédiate (déterminée par leur VGE moyenne) est contrebalancée par leur contribution négative à la réponse génétique future (déterminée par leur contribution à la consanguinité). La valeur de la contribution à la consanguinité est dérivée de l'objectif de sélection. La sélection d'individus apparentés entre eux sera soumise à restriction si l'effet nuisible de la consanguinité est fort ; les restrictions à la sélection de germains peuvent cependant varier d'une famille à une autre en fonction de leur valeur génétique. Un algorithme de sélection est proposé pour établir le critère de sélection des pères en fonction de 3 stratégies : sélectionner i) un nombre fixe de pères, ii) un nombre variable de pères à chacun desquels on attribue un nombre égal d'accouplements, ou iii) un nombre variable de pères entre lesquels on affecte les accouplements d'une manière optimale. À l'aide de simulations stochastiques, ces stratégies de sélection paternelle sont comparées à la sélection sur VGE seule. Quand on compare les résultats au terme de la période spécifiée dans l'objectif de sélection, le critère de sélection proposé réussit à assurer une réponse à la sélection augmentée et un niveau de consanguinité diminué en dépit d'un nombre plus faible de pères sélectionnés. La stratégie de sélection iii) exploite les fluctuations aléatoires des nombres de pères disponibles d'une année à l'autre et maximise la réponse pour l'objectif de sélection. le calcul de la valeur attribuée à la consanguinité n'est pas exacte et ne peut pas garantir que la réponse globale maximale est obtenue. Cependant, les résultats de simulation suggèrent que la réponse prédite est robuste vis-à-vis des effets nuisibles attribués à la consanguinité.*

**sélection artificielle / réponse à la sélection / consanguinité / BLUP / simulation sur ordinateur**

## INTRODUCTION

When breeding programmes are considered, it is commonly assumed that new parents are selected on the criterion of highest estimated breeding values alone. This criterion results in maximum response to a single generation of selection, but repeated use of this criterion does not necessarily result in maximum genetic response over a longer time horizon. Selection decisions made in the current generation have at least 2 consequences. Firstly, they influence the genetic response

to selection, the impact of which is seen immediately in the genetic merit of their offspring born in the next generation. Secondly, they influence the inbreeding of the next and subsequent generations. Accumulation of inbreeding has a negative impact on future genetic response through reduction in future genetic variance and a negative impact on future performance if inbreeding depression affects the selected trait.

Dempfle (1975) showed that selection limits achieved with mass selection could be surpassed by within-family selection particularly when selection intensities and heritability were high; within-family selection caused lower levels of inbreeding and hence ensured higher genetic variance in the long term. Best linear unbiased prediction (BLUP) is now the preferred method for calculation of estimated breeding values (EBVs). The EBVs of relatives are highly correlated especially if BLUP is applied under an animal model; selection on BLUP EBVs alone can result in higher rates of inbreeding than under mass selection, and hence available genetic variance is more quickly reduced. Indeed, in some circumstances it has been found that mass selection can result in higher long-term genetic gain than selection on BLUP EBVs (Quinton *et al*, 1992; Verrier *et al*, 1993), and the practice of selection on BLUP EBV alone has been questioned. Some authors have investigated the consequences of ignoring records on some relatives (*eg*, Brisbane and Gibson, 1993, scheme SUBOPT) but this implies that ignorance can sometimes be preferred to knowledge. Others have suggested that an artificially high heritability could be used in the BLUP equations (*eg*, Toro and Perez-Enciso, 1990; Grundy and Hill, 1993) which gives more weight to individual rather than relatives records, but this confuses the method of prediction of breeding values with the selection criterion. The intuitively attractive answer must be to combine the EBVs (calculated in the optimal way) into a selection criterion that truly reflects the underlying selection goal, thereby increasing, rather than decreasing, the amount of information included to make selection decisions.

Several authors have investigated selection criteria that attempt to ensure higher genetic response over a longer time horizon. These include imposing restrictions on the numbers of sibs selected from any family (*eg*, Toro and Perez-Enciso, 1990; Brisbane and Gibson, 1993; Grundy and Hill, 1993), selection on a criterion which alters the emphasis given to within-family and family information (*eg*, Dempfle, 1975; Toro and Perez-Enciso, 1990; Verrier *et al*, 1993; Villanueva *et al*, 1994), selection of an increased number of parents but allocating more matings to higher ranked parents so that the overall selection intensity is the same as if a smaller number of sires had been selected (Toro and Nieto, 1984; Toro *et al*, 1988, Lindgren, 1991), selection on a criterion $EBV_i$-weight $X$, where $X$ is the average relationship of the individual with the other selected parents (Goddard and Smith, 1990a; Brisbane and Gibson, 1993), or linear programming to determine the set of matings out of all possible sets that maximises response to selection under a given restriction for inbreeding (Toro and Perez-Enciso, 1990). All of these alternatives have met with some success in gaining higher genetic response at lower levels of inbreeding over some time horizon. The methods all aim, in an indirect way, to maintain genetic variance and restrict inbreeding, but the actual criterion by which this is achieved is perhaps arbitrary. No guidelines have been presented which might ensure that optimum response over a given time horizon is achieved.

In general, investigation of breeding programmes assumes the mating of a fixed number of sires with a fixed number of dams generating a fixed number of offspring each generation. The expected optimum proportion of parents to select is a function of the ratio of the time horizon of the breeding programme and the number of animals available for selection (Robertson, 1970; Jodar and Lopez-Fanjul, 1977). In practice, however, the number of females selected is constrained by the female reproductive rate and testing facilities for offspring. By contrast, restrictions on the number of sires are likely to be much broader, if they exist at all (particularly when artificial insemination is used). The genetic merit of individuals available for selection each generation is partly random, therefore optimum selection decisions that exploit this randomness may result in different numbers of sires being selected at each generation and differential usage of the sires.

In this paper, an attempt is made to provide a selection criterion which is explicit in its goal of maximising response to selection over a specified time horizon. As well as reducing genetic variance, inbreeding may cause a depression in performance. Selection goals are considered for which the aim is to maximise genetic response less the cost of inbreeding depression over some time horizon. Selection rules are presented which are dyanmic in their attempt to exploit the genetic merit of parents which arise randomly (in part) each generation.

## METHODS

The aim is to find a selection criterion that weights selection response *versus* future inbreeding in a logical way. The relevant weights must depend on a breeding objective, and therefore the definition of the breeding objective is our starting point. The derivation of the selection criterion is based on the maximisation of response to the breeding objective. However, since the selection criterion affects inbreeding and the level of inbreeding influences the optimum selection criterion, it is not possible to find a selection criterion which is constant each generation and which can guarantee maximisation of the breeding objective. Therefore, the selection criterion is not expected to ensure maximisation of the breeding objective, but it is expected to result in higher response to the breeding objective than selection on EBV alone. Finally, the selection criterion is used in conjunction with different selection algorithms which may allow different numbers of parents to be selected or allocate different numbers of matings to each parent in order to maximise the effectiveness of the selection criterion. For simplicity, we consider only selection on males and the selection of females is assumed to be at random.

### *Breeding objective*

A general breeding objective for any livestock population may be cumulative net response to generation $t$, $R_t$

$$R_t = \left[\sum_{j=1}^{t} \Delta G_j\right] - DF_t \qquad [1]$$

where $\Delta G_j$ is the increase in genetic merit of animals born in generation $j$, $F_j$ is their average inbreeding coefficient and $D$ is the depression in performance per unit of inbreeding. $F_j$ can be expressed as

$$F_j = 1 - (1 - \Delta F)^j$$

where $\Delta F$ is the rate of inbreeding per generation. $\Delta G_j$ can be approximated by

$$\Delta G_j \approx \Delta G_L(1 - F_{j-1})$$

where $\Delta G_L$ is the asymptotic rate of gain per generation expected in an infinite population after accounting for the effects of selection (the 'Bulmer effect', Bulmer, 1971). This approximation for $\Delta G_j$ arises by assuming, firstly, that $\Delta G_j$ is predicted by $ir_{j-1}\sigma_{G,j-1}$ where $i$ is the selection intensity each generation, and $r_{j-1}$ and $\sigma^2_{G,j-1}$ are the accuracy of selection and genetic variance, respectively, pertaining to animals born in generation $j$–1. Secondly, it is assumed that $\sigma^2_{G,j-1} \approx \sigma^2_{G,L}(1-F_{j-1})$ and $r_{j-1} \approx r_L(1 - F_{j-1})^{1/2}$. Thus, it is assumed that rate of gain (and its components) are reduced each generation by the level of inbreeding achieved. Substituting the expressions for $\Delta G_j$ and $F_j$ into [1], $R_t$ can be written as:

$$R_t \approx \sum_{j=1}^{t} \Big[\Delta G_L(1 - \Delta F)^{j-1}\Big] - D\Big[1 - (1 - \Delta F)^t\Big]$$

Ignoring terms of higher order than linear in $\Delta F$ then,

$$R_t \approx t\Big[\Delta G_L - \Delta F(\Delta G_L(t-1)/2 + D)\Big]$$

which is the same as the expression used by Goddard and Smith (1990b). The linear approximation to $\Delta F$ should be satisfactory if $\Delta F < 1\%$ as it is in many livestock populations (if $\Delta F = 0.01$ and $t = 30$, the first formula for $R_t$ is $26.0\Delta G - 0.26D$, while the formula using the linear approximation is $R_t \approx 25.7\Delta G - 0.30D$). For small, intensely selected populations that have higher $\Delta F$, the approximation may become less acceptable; to check the effect of this, the simulations to be reported have $\Delta F$ as 1 – 3%. The breeding objective for each generation can be written as

$$R_t/t \approx \Delta G_L - Q\Delta F \qquad [2]$$

where

$$Q = \Delta G_L(t-1)/2 + D \qquad [3]$$

Equation [2] implies that in each of the $t$ generations of selection there is a positive contribution to the breeding objective of genetic response and there is a detrimental contribution to the breeding objective as a function of the rate of inbreeding.

### *Selection criterion*

We wish to choose a selection criterion which maximises gains in the breeding objective (equation [2]). The gain in additive genetic merit expected from one

generation of selection decisions is

$$\frac{1}{2}\mathbf{s}'_m\mathbf{b}_m + \frac{1}{2}\mathbf{s}'_f\mathbf{b}_f \quad [4]$$

where $\mathbf{s}_m$ is a vector containing the proportion of offspring born to each sire and $\mathbf{b}_m$ is the vector of estimated breeding values (EBVs) of sires deviated from the overall mean of EBVs of all available sires and dams prior to their selection. $\mathbf{s}_f$ and $\mathbf{b}_f$ are defined analogously for dams. The average coancestry amongst the parents weighted by their contribution to the next generation represents the effect on inbreeding induced by the selection decisions, that is

$$w_j = \frac{1}{8}\left[\mathbf{s}'_m\mathbf{A}_{mm}\mathbf{s}_m + 2\mathbf{s}'_m\mathbf{A}_{mf}\mathbf{s}_f + \mathbf{s}'_f\mathbf{A}_{ff}\mathbf{s}_f\right]$$

where $\mathbf{A}_{mm}$, $\mathbf{A}_{mf}$ and $\mathbf{A}_{ff}$ represent the additive genetic relationship matrices between sires, between sires and dams, and between dams respectively. The rate of inbreeding is $(w_j - w_{j-1})/(1 - w_{j-1})$. Assuming that $w_{j-1}$ is small, as it is in most commercial livestock populations, the rate of inbreeding is approximated by

$$w_j - w_{j-1} \quad [5]$$

For example, when sires and dams are unrelated and are non-inbred, $\mathbf{A}_{mm}$ and $\mathbf{A}_{ff}$ are identity matrices, $\mathbf{A}_{mf}$ is null and if all $N_m$ sires and $N_f$ dams are used equally (*ie*, $\mathbf{s}_m = \mathbf{1}N_m^{-1}$ and $\mathbf{s}_f = \mathbf{1}N_f^{-1}$ where $\mathbf{1}$ is a vector of ones) then $w_{j-1} = 0$, $\Delta F = w_j = 1/8N_m^{-1} + 1/8N_f^{-1}$ (Wright, 1931). Substituting the expressions [4] and [5] into $\Delta G_L$ and $\Delta F$ of the breeding objective (equation [2]) gives the selection criterion ($V$).

$$V = \frac{1}{2}\mathbf{s}'_m\mathbf{b}_m + \frac{1}{2}\mathbf{s}'_f\mathbf{b}_f - Q(w_j - w_{j-1})$$

The aim is to choose $\mathbf{s}_m$ and $\mathbf{s}_f$ so that the selection criterion is maximised. However, $w_{j-1}$ is determined by selection decisions made last generation, which is unaffected by $\mathbf{s}_m$ and $\mathbf{s}_f$ because they specify selection decisions made this generation. Therefore, the selection criterion can be simplified to

$$V = \frac{1}{2}\mathbf{s}'_m\mathbf{b}_m + \frac{1}{2}\mathbf{s}'_f\mathbf{b}_f - Q\frac{1}{8}\left[\mathbf{s}'_m\mathbf{A}_{mm}\mathbf{s}_m + 2\mathbf{s}'_m\mathbf{A}_{mf}\mathbf{s}_f + \mathbf{s}'_f\mathbf{A}_{ff}\mathbf{s}_f\right] \quad [6]$$

If our interest is restricted to decisions regarding male selection (*ie* choosing $\mathbf{s}_m$ and assuming that females are selected at random, so that all available females have equal probability of featuring in $\mathbf{s}_f$), then $\mathbf{s}'_f\mathbf{A}_{ff}\mathbf{s}_f$ is not affected by the selection decisions and can be ignored. $\mathbf{s}'_m\mathbf{A}_{mf}\mathbf{s}_f$ represents the average relationship between selected males and the randomly chosen females; we assume that this is little affected by the choice of $\mathbf{s}_m$ and therefore choose as our selection criterion

$$V = \frac{1}{2}\mathbf{s}'_m\mathbf{b}_m - Q\frac{1}{8}\left[\mathbf{s}'_m\mathbf{A}_{mm}\mathbf{s}_m\right] \quad [7]$$

The approximations invoked in the derivation of equation [7] mean that it must be considered as a heuristic selection criterion whose usefulness will be tested by the simulation results.

The aim of the selection criterion is to determine which sires to select amongst the males available for selection and what proportion of matings should be allocated to each. The optimum value of $\mathbf{s}_m$ can be found by differentiating $V$ with respect to $\mathbf{s}'_m$ after including the restriction that the mating proportions must sum to 1, $\mathbf{s}'_m\mathbf{1} = 1$, *via* a LaGrange multiplier, $\lambda$:

$$\frac{\delta}{\delta \mathbf{s}'_m}\left[\frac{1}{2}\mathbf{s}'_m\mathbf{b}_m - \frac{1}{8}Q\mathbf{s}'_m\mathbf{A}_{mm}\mathbf{s}_m - \left(\mathbf{s}'_m\mathbf{1} - 1\right)\lambda\right] = \frac{1}{2}\mathbf{b}_m - \frac{1}{4}Q\mathbf{A}_{mm}\mathbf{s}_m - \mathbf{1}\lambda = 0$$

Solving for $\mathbf{s}_m$ gives

$$\mathbf{s}_m = \frac{2}{Q}\mathbf{A}^{-1}_{mm}\left[\mathbf{b}_m - 2\mathbf{1}\lambda\right] \qquad [8]$$

and since $\mathbf{s}'_m\mathbf{1} = 1$, then

$$\lambda = \frac{\frac{1}{4}\left[2\mathbf{1}'\mathbf{A}^{-1}_{mm}\mathbf{b}_m - Q\right]}{\mathbf{1}'\mathbf{A}^{-1}_{mm}\mathbf{1}}$$

### *Selection algorithm*

The selection criterion $V$ can be used to determine the optimum number of sires ($n$) to select under the prevailing circumstances using the following algorithm.

1. Rank sires on EBV and select the best $n = 1$.
2. For the remaining sires, calculate $V_{[n+1]}$ for each sire, which depends on the group of $n$ sires already selected plus the individual sire to be considered.
3. Rank the sires on their individual $V_{[n+1]}$ values, select the best sire if $(V_{[n+1]} - V_{[n]}) > 0$ then repeat from step 2 ($n = n + 1$), otherwise stop the search and select only the first $n$ sires nominated.

This algorithm can be used to allow different sire selection strategies each using the selection criterion [7].

Strategy 1: Fixed number of sires ($N_m$) used each year, each allocated an equal (as far as possible) number of matings $\mathbf{s}_m$ of order $N_m$ and $\mathbf{s}_m = N_m^{-1}\mathbf{1}$; repeat steps 2 and 3 $N_m - 1$ times, always selecting the sire with the highest $V_{[n+1]}$ value in step 3.

Strategy 2: Selection of a variable (optimum) number of sires each generation, each allocated an equal number of matings $\mathbf{s}_m = n^{-1}\mathbf{1}$.

Strategy 3: Selection on a variable (optimum) number of sires with a variable number of matings allowed/sire, $\mathbf{s}_m$ defined by equation [8].

If the algorithm is used to select a variable number of sires each generation (strategies 2 or 3), the selection criterion balances superiority in genetic merit with inbreeding considerations. The aim of the selection procedure is to exploit, in an optimal way, the sires who have become available for selection by chance in the

current generation. This algorithm does not ensure that 'the' best group of sires is selected. However, in simulations of small populations where it has been possible to subjectively compare the group chosen by algorithm *versus* 'the' best group out of all possible combinations, the algorithm has performed well. The algorithm may not perform as well for larger populations, but is is still likely to be close to the optimum.

To gain insight into the selection criterion, assume that sires are used equally (strategy 2). From equation [7], it can be shown that an $n+1$th sire is selected if

$$b_{n+1} - \frac{1}{n}\sum_{i=1}^{n} b_i > \frac{Q}{4(n+1)}\left[2\sum_{i=1}^{n} a_{i,n+1} + a_{n+1,n+1} - \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\frac{(2n+1)}{n^2}\right] \quad [9]$$

where $b_i$ are the elements of $\mathbf{b}_m$ and $a_{ij}$ are the elements of $\mathbf{A}_{mm}$. Presented in this way, it is apparent that the contribution of the $n+1$th sire to genetic merit of the selected group of sires, is balanced against his contribution to inbreeding. When the sires are completely non-inbred and are not related to each other, the contribution to inbreeding of selecting $n$ sires is $1/8\,\mathbf{s}'_m\mathbf{A}_{mm}\mathbf{s}_m = 1/8n$ and an $n+1$th sire is selected if

$$b_{n+1} - \frac{1}{n}\sum_{i=1}^{n} b_i > -\frac{Q}{4n}$$

At the other extreme, if the population is completely inbred (all elements of $\mathbf{A}_{mm}$ are 2) then the contribution to inbreeding of selection $n$ sires is $1/8\,\mathbf{s}'_m\mathbf{A}_{mm}\mathbf{s}_m = 2$, and an $n+1$th sire is selected if

$$b_{n+1} - \frac{1}{n}\sum_{i=1}^{n} b_i > 0$$

This is an artificial example, because when the population is totally inbred, there is no remaining genetic variance and the EBVs of all the sires are the same. However, the implication is that as the population becomes more inbred, the criterion for selection of sires becomes more strict, implying a reduction in the number of sires selected. However, this is counteracted by a reduction in the variance of EBVs so that values on the left-hand side of equation [9] also become smaller.

### *Value of Q*

Implementation of the algorithm proposed above for selection of sires depends on the definition of $Q$ which, in turn, is dependent on the definition of the breeding objective. A value for $Q$ can be found by substituting a prediction for $\Delta G_L$ into equation [3], which in turn depends on predictions of $i$, $r_L$ and $\sigma_{G,L}$. Under the variable number of sires options, the optimum number of sires (assuming equal mating of sires) can be predicted (Goddard and Wray, unpublished results) and selection intensity calculated as though that proportion of sires was selected.

If selection is based on phenotypes alone, for a trait with heritability $h^2$ and phenotypic variance in the base population unity, then $\sigma^2_{G,0} = h^2$ and

$$\sigma^2_{G,L} = \left[h^2 - \frac{1}{2}\left[1 - (1 + 4kh^2(1 - h^2))^{1/2}\right]\right](1 + k)^{-1}$$

(Bulmer, 1980) where $k$ is the variance reduction factor appropriate to the selection intensity (averaged over the 2 sexes, for each sex $k = i(i - x)$, $x$ being the standard normal deviate), and $r^2_L = \sigma^2_{G,L}(\sigma^2_{G,L} + 1 - h^2)^{-1}$. Alternatively, if selection is on BLUP EBVs then a lower bound to the accuracy of selection before accounting for the Bulmer effect is:

$$r^2 = \left[1 + 2h^2 - (1 + 4h^2(1 - h^2))^{1/2}\right](2h^2)^{-1}$$

$$\sigma^2_{G,L} = h^2\left[1 + k(1 - r^2)\right](1 + k)^{-1}$$

and

$$r^2_L = 1 - (1 - r^2)h^2/\sigma^2_{G,L}$$

(Dekkers, 1992). This lower bound to accuracy of selection for BLUP assumes the only information contributing to an individuals EBV is its own record and its parental EBVs. When an individual has many sibs with records, the accuracy may be considerably underestimated. Indeed the $\Delta G_L$ predicted when selection is on BLUP EBVs using this lower bound accuracy may not be significantly higher than $\Delta G_L$ predicted for mass selection. However, these equations provide a simple deterministic approximation with which to attain a ball-park prediction.

The definition for $Q$ can only be approximate, since the optimum value of $Q$ is an iterative balance between selection response and inbreeding, particularly when the number of sires is allowed to vary; the value of $Q$ influences the selection decisions, and the selection decisions change the optimum value of $Q$. In fact, the value assigned to $Q$ (equation [3]) assumes that the selection goal is always $t$ generations into the future. If the selection goal is cumulative net response to generation $t$ with no interest in response in subsequent years, then $Q$ in equation [2] should take on subscript $j$ representing the selection criterion in generation $j$ ($j = 0, t - 1$) with

$$Q_j = \Delta G_L(t - j - 1)/2 + D$$

Under this definition, the selection decisions made in generation $t - 1$ give no detrimental weighting to the effect of selection on future genetic variance because under the selection goal it is assumed that selection stops in generation $t$. This definition is quite unlikely in practice. We would recommend $Q$ to be defined as in equation [7] where $t$ takes on a medium time horizon value.

### *Simulations*

Populations are simulated with discrete generations in which $N_m$ males are mated to $N_f$ females and each female gives $N_{sex}$ offspring of each sex. $N_f$ and $N_{sex}$ are fixed each generation. In the base generation $N_m = N_f$, but thereafter $N_m$ may be fixed or variable, depending on the sire selection strategy. The phenotype ($p_j$) of individual $j$ is simulated as $p_j = u_j + e_j$, where $u_j$ is the true breeding value and $e_j$

is the environmental value of the individual. For a trait with phenotypic variance of unity and heritability of $h^2$, an infinitesimal model of genetic effects is assumed. In the base population $u_j$ is sampled from a normal distribution $N(0, h^2)$, and in later generations $u_j$ is sampled from a normal distribution $N(0.5(u_s + u_d), 0.5(1-f)h^2)$, where $u_s$ and $u_d$ are the true breeding values of the sire and dam of individual $j$ and $f$ is their average inbreeding coefficient. Each generation $e_j$ is sampled from a normal distribution $N(0, 1-h^2)$. Dams are selected at random.

EBVs are calculated by true- or by pseudo-animal model BLUP. In the true-BLUP, the only fixed effect is the overall mean, base population variances are used and all relationships between animals are included. In the pseudo-BLUP, EBVs are calculated using an index of individual, full and half sib records plus EBVs of the dam, sire and mates of the sire (Wray and Hill, 1989). The selection index weights change each generation depending on the available genetic variance ($\sigma^2_{G,j}$), which is calculated as

$$\sigma^2_{G,j} = 1/2\sigma^2_{G,j-1}(1-k_{j-1}r^2_{j-1})(1+F_j-2F_{j-1})(1-F_{j-1})^{-1} + 1/2\sigma^2_{G,0}(1-F_{j-1})$$

(Wray and Thompson, 1990a), where $F_j$ and $F_{j-1}$ are the actual average inbreeding coefficients over all individuals born in generations $j$ and $j-1$ and $r^2_j = \sigma^2_{I,j}/\sigma^2_{G,j}$ where $\sigma^2_{I,j}$ is the expected variance of the index in generation $j$ (calculated from the index weights and genetic variance); $k_{j-1}$ is half the variance reduction factor appropriate to the number of males selected in generation $j-1$ (since dams are selected at random). When the number of sires and matings/sire are variable, the variance reduction factor is based on an effective number of sires calculated as $N_f/m$, where $m$ is the average number of dams/sire, $m = \mathbf{s}^{*'}_n \mathbf{s}^*_n N_f^{-1}$, where $\mathbf{s}^*_n$ is the integer vector $\mathbf{s}_n N_f$ of actual numbers of matings/sire. When matings/sire are variable, all individuals have EBVs calculated using the same index which assumes the same average number of dams/sire, $m$. The use of pseudo-BLUP is very efficient on computing time compared with true-BLUP, particularly when considering schemes over many generations. Simulations based on true-BLUP are used only as a check that the pseudo-BLUP results in similar selection decisions. Selection continues for 30 discrete generations (20 for true-BLUP) and results are the average of 200 simulation replicates. Response to selection in generation $t$, $R_t$, is calculated as the mean over all individuals born in generation $t$ of $p_j - Df_j$, where $f_j$ is the inbreeding coefficient for individual $j$; when $D = 0, R_t$ represents the average genetic merit. Note that when $D > 0$, the records analysed in the BLUP are still the $p_j$, thus we assume exact prior correction of records for inbreeding depression. The underlying genetic model could represent a trait controlled by a large number of additive loci plus a group of loci with rare deleterious recessives, which make a negligible contribution to the additive variance. This genetic model is one of several which could be chosen to simulate inbreeding depression, but this model corresponds to the way in which inbreeding depression is accounted for in the genetic evaluation of livestock populations. Summary statistics are calculated within the simulations, these include: $R_t$, $F_t$ calculated as the mean of all $f_j$, rate of inbreeding, $\Delta F = (F_t - F_{t-1})/(1-F_{t-1})$, averaged from $t=2$ calculated as $\sum\sum_{i \neq j} T_{ij} / \left[N_m(N_m-1)\right]$ where $T_{ij} = 1$ if the sires $i$ and $j$ are sibs and 0 otherwise.

***Population structure alternatives***

Basic: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 0$, the selection goal is $R_{30}$ (equation [1]), $Q$ is defined in equation [3], and $\Delta G_L$ is calculated assuming mass selection.

Alternative 1: $Q$ replaced by $Q^*$, where $Q^* = cQ$, where c is a constant. This alternative allows investigation of the robustness of the prediction of the value of $Q$ and values used in different simulations are c = 0.5, 0.8, 0.9, 1.0, 1.1, 1.2, 1.5.

Alternative 2: $N_f = 25$.

Alternative 3: $N_{sex} = 2$.

Alternative 4: $h^2 = 0.1$.

Alternative 5: $D = 3.33$, equivalent to 1% inbreeding depression/% inbreeding for a trait with coefficient of variance of 15%.

Alternative 6: $D = 3.33$, selection goal $R_{10}$ (equation [1]).

For each alternative, simulations for each of the 3-sire selection strategies are compared with selection on EBV alone. For selection on EBV and sire selection strategy 1, fixed values of $N_m$ used are $N_m = 3, 6, 9, 12, 15, 18, 21, 24$ for all alternatives except alternative 5 where $N_m = 6, 9, 12, 15, 18, 21, 24, 30$.

## RESULTS

In table I response to the breeding objective, $R_t$, and level of inbreeding $F_t$ for $t = 10, 20$ are presented for pseudo-BLUP and true-BLUP for the basic population structure ($h^2 = 0.4$ or 0.1) with $N_m = 9$ when selection is on EBV or strategy 1, and using an expected number of sires of 9 for the calculation of $Q$ for sire selection strategies 2 and 3. A good agreement was found between results for pseudo-BLUP and true-BLUP, particularly for strategies EBV, 1 and 2. In strategy 3, where an average index is used for all offspring in the pseudo-BLUP based on an average number of full and half sibs, the response from the pseudo-BLUP is slightly less than is found with true-BLUP when $h^2 = 0.1$. On the basis of these results, only pseudo-BLUP is used for investigations of the full range of schemes.

In figure 1 response to selection $R_{10}$ and $R_{30}$ are plotted against level of inbreeding $F_{10}$ and $F_{30}$ respectively for the basic scheme, for sire selection on EBV, and for strategies 1–3, where the breeding objective is $R_{30}$. This representation of the results demonstrates the success of different simulations in achieving high $R_t$ but low $F_t$. The highest response at generation 10 is achieved with a low number of sires selected on EBV alone. However, by generation 30, the time horizon of the breeding objective, quite a different picture is seen: when the same number of sires are selected, $R_{30}$ for sire selection strategy 1 is always greater than for sire selection on EBV alone and $F_{30}$ is concurrently less. The maximum $R_{30}$ for selection on EBV alone is $10.75 \pm 0.029$ which occurs with 12 sires (amongst the sire combinations considered) at $F_{30}$ of $0.491 \pm 0.0025$. Whilst the maximum $R_{30}$ for selection on $V$ (equation [7]) with a fixed number of sires (strategy 1) is higher at $11.10 \pm 0.030$, but this occurs with a smaller number of sires, 9, selected each generation and also at a lower $F_{30}$ of $0.404 \pm 0.0013$. Sire selection strategies 2 and 3 use a $Q$ value

**Table I.** Response to selection, $R_t$ and level of inbreeding $\times 100 F_t$, $t = 10, 20$ (± between replicate standard errors) for the basic population structure ($N_f = 100$, $N_{sex} = 4$, $D = 0$, $Q = 6.11$) using sire selection strategies EBV and 1–3 in which EBVs are calculated by pseudo-BLUP (p-BLUP) or true BLUP (t-BLUP) and where the breeding goal is $R_{30}$ (equation [7]).

| *Sire selection strategy* | *EBV* | $h^2$ | $N_m^{\dagger}$ | $R_{10}$ | $R_{20}$ | $F_{10}$ | $F_{20}$ |
|---|---|---|---|---|---|---|---|
| EBV | t-BLUP | 0.4 | 9 | 4.05 ± 0.021 | 7.68 ± 0.028 | 21.5 ± 0.24 | 42.2 ± 0.27 |
| EBV | p-BLUP | 0.4 | 9 | 4.05 ± 0.020 | 7.66 ± 0.026 | 21.5 ± 0.22 | 41.9 ± 0.27 |
| 1 | t-BLUP | 0.4 | 9 | 3.95 ± 0.018 | 7.75 ± 0.024 | 13.9 ± 0.09 | 28.5 ± 0.11 |
| 1 | p-BLUP | 0.4 | 9 | 3.94 ± 0.020 | 7.73 ± 0.027 | 13.8 ± 0.09 | 28.4 ± 0.12 |
| 2 | t-BLUP | 0.4 | 10.1 ± 0.05 | 3.93 ± 0.020 | 7.77 ± 0.027 | 13.2 ± 0.11 | 27.3 ± 0.16 |
| 2 | p-BLUP | 0.4 | 9.9 ± 0.07 | 3.93 ± 0.020 | 7.75 ± 0.027 | 13.5 ± 0.12 | 27.2 ± 0.16 |
| 3 | t-BLUP | 0.4 | 14.7 ± 0.09 | 4.00 ± 0.021 | 7.91 ± 0.028 | 13.2 ± 0.14 | 27.5 ± 0.19 |
| 3 | p-BLUP | 0.4 | 14.3 ± 0.08 | 4.04 ± 0.021 | 7.94 ± 0.027 | 13.7 ± 0.15 | 28.0 ± 0.19 |
| EBV | t-BLUP | 0.1 | 9 | 1.42 ± 0.014 | 2.62 ± 0.018 | 27.0 ± 0.28 | 51.9 ± 0.30 |
| EBV | p-BLUP | 0.1 | 9 | 1.41 ± 0.013 | 2.61 ± 0.018 | 27.0 ± 0.31 | 51.5 ± 0.34 |
| 3 | t-BLUP | 0.1 | 13.5 ± 0.08 | 1.42 ± 0.012 | 2.75 ± 0.017 | 18.0 ± 0.21 | 37.3 ± 0.29 |
| 3 | p-BLUP | 0.1 | 12.4 ± 0.08 | 1.39 ± 0.013 | 2.65 ± 0.016 | 18.4 ± 0.19 | 37.1 ± 0.25 |

$^{\dagger}$ In all strategies (except strategy EBV, where $Q$ is not needed), 9 sires are assumed for the calculation of $Q$. For strategies 3 and 4, the number of sires is the average number used over generations 2 to 20.

which assumes the selection of 9 sires each generation. For strategy 4, the actual number of sires selected is $9.6 \pm 0.07$ yielding an $R_{30}$ of $11.18 \pm 0.033$ at an $F_{30}$ of $0.403 \pm 0.0018$. For strategy 3, the actual number of sires selected is $14.0 \pm 0.09$, which when differential usage is accounted for corresponds to an effective number of sires selected of $9.6 \pm 0.08$ yielding an $R_{30}$ of $11.39 \pm 0.032$ at an $F_{30}$ of $0.402 \pm 0.0020$. Selection of a variable number of sires (strategies 2 and 3) results in a higher $R_{30}$ value than selection of a fixed number of sires, although the optimum occurs with approximately the same number of sires used on average and at approximately the same level of inbreeding. Allocation of a variable number of matings/sire results in a higher $R_{30}$ than allocation of matings equally. The standard errors of $R_{30}$ are higher for strategies 2 and 3 for selection on EBV, despite the lower level and standard error of inbreeding implying the variable selection strategies may be associated with more risk, if risk is measured by variance of response.

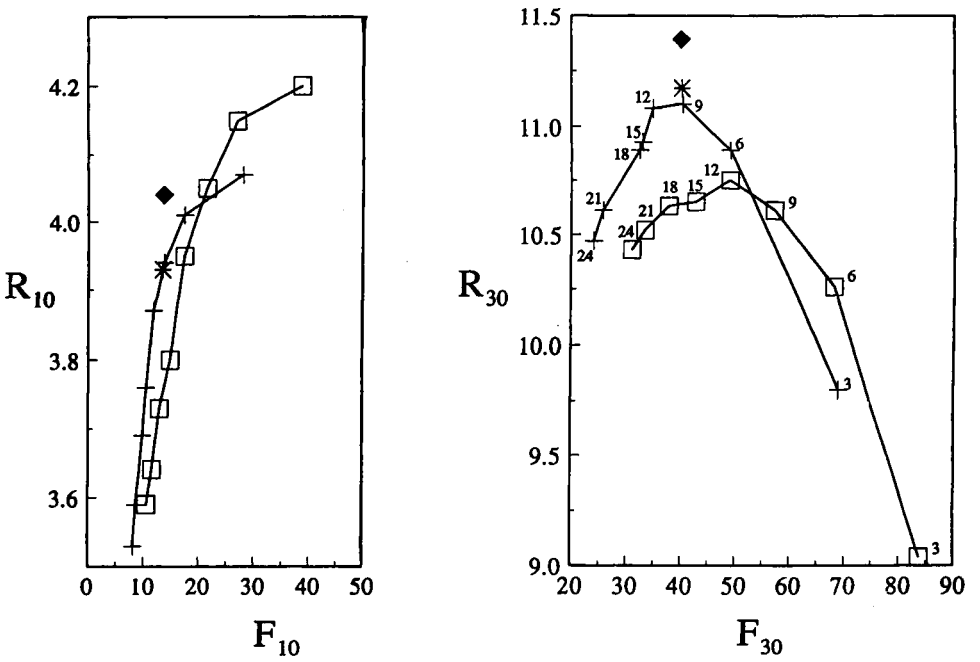



**Fig 1.** Response to selection $R_t$ against level of inbreeding $F_t$, for $t = 10$ and 30 when the breeding goal is $R_{30}$. Basic population structure: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 0$. Sire selection strategies: EBV: □, 1: +, 2: *, 3: ♦. For strategies EBV and 1, each point represents a different fixed number of sires selected each generation; the numbers of sires are marked on the $t = 30$ graph, but are equivalent on the $t = 10$ graph.

A number of assumptions are employed to obtain the value of $Q$ for sire selection strategies 1–3. Alternative 1 investigates the importance of accurate prediction of $Q$ and results are presented in table II for strategy 1. As the value of $Q^*$ increases,

$F_t$, $\Delta F$ and the probability of co-selecting sibs all decrease, as expected. In this example, a slightly higher $R_{30}$ could be achieved by using $Q^* = 0.9Q$. However, the stability of $R_{30}$ over the wide range of $Q^*$ is more notable. Similar results were found for selection strategies 2 and 3. In these examples, $Q$ has been calculated using a $\Delta G_L$ appropriate to mass selection, $Q = 6.11$. If $Q$ were calculated as appropriate to BLUP based on the lower bound accuracy then $Q = 6.42$, equivalent to c = 1.05, and if an accuracy based on the actual number of full and half sibs is used then $Q = 6.89$, equivalent to c = 1.13. From table II, it is apparent that the selection results are robust to the method used to predict $\Delta G_L$. Similar results are found for the number of sires assumed to predict $Q$ for strategies 2 and 3, assuming the selection of 3 or 15 sires instead of 9 sires is equivalent to c = 1.16 or 0.92 respectively. The $Q$ are dependent on the time horizon, $t$, *eg*, $t = 20$ is equivalent to c = 0.65. Since the optimum $c$ for this example is 0.9 despite the fact that EBVs are BLUP and the prediction of $\Delta G_L$ assumes mass selection, implies that the underprediction of $r_L \sigma_{G,L}$ is counterbalanced by the overprediction of $i$, the selection intensity, which is calculated from normal distribution theory assuming that the best sires on EBV are selected. Despite the proposal in the *Methods* section that a decreasing $Q$ each generation, $Q_j$, would be more appropriate for maximising a selection of a fixed future time horizon, the use of $Q_j$ for this example resulted in slightly lower $R_{30}$ than from using a fixed $Q$ (results not presented). This reflects errors in prediction $Q$ rather than contradicting the principle that a $Q_j$ should be more appropriate.

**Table II.** Response to selection $R_{10}$ and $R_{30}$ and level of inbreeding ($\times 100$), $F_{10}$ and $F_{30}$, rate of inbreeding ($\times 100$), $\Delta F$, and probability of coselection of male full and half sibs, P(FS) and P(HS), for alternative 1 and sire selection strategy 1, $N_m = 9$, $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 0$, selection criterion equation [7] with $Q$ replaced by $Q^*$ where $Q^* = cQ$ and $Q = 6.11$ for selection goal $R_{30}$ (equation [1]).

| $c$ | $R_{10}$ | $R_{30}$ | $F_{10}$ | $F_{30}$ | $\Delta F$ | *P(HS)* | *P(FS)* |
|---|---|---|---|---|---|---|---|
| 0.5 | 4.01 | 11.05 | 16.4 | 46.4 | 2.13 | 0.142 | 0.063 |
| 0.8 | 3.97 | 11.09 | 14.7 | 42.6 | 1.90 | 0.120 | 0.045 |
| 0.9 | 3.97 | 11.14 | 14.3 | 41.7 | 1.84 | 0.114 | 0.039 |
| 1.0 | 3.94 | 11.10 | 13.8 | 40.4 | 1.77 | 0.108 | 0.035 |
| 1.1 | 3.94 | 11.08 | 13.6 | 39.4 | 1.71 | 0.103 | 0.032 |
| 1.2 | 3.88 | 11.03 | 13.3 | 38.8 | 1.68 | 0.101 | 0.029 |
| 1.5 | 3.89 | 10.99 | 12.6 | 37.0 | 1.58 | 0.089 | 0.021 |
| Max se | 0.020 | 0.032 | 0.14 | 0.19 | 0.060 | 0.0053 | 0.0012 |

Qualitatively, the results for alternative population structures 2–5 are similar to those for the basic scheme, generating graphs similar in shape to figure 1. Results are tabulated in table III using the $N_m$ (out of the alternatives examined) which generates maximum response to $R_{30}$ for selection strategies EBV and 1. The average effective number of sires used in strategy 3 is approximately the same as the average actual number of sires used in strategy 2. In all alternatives, there is a tendency for a small decline over generations in the optimum number of sires selected in strategies

2 and 3, *eg*, for the basic scheme strategy 2, the number of sires selected declines from 10.4 to 9.3. The optimum number of sires when selection is on EBV is always higher than for strategies 1 and 2 (and effective number of sires of strategy 3), therefore each individual has less half sibs available for selection. Despite this, the probability of coselection of sibs is higher for strategy EBV.

**Table III.** Response to selection $R_t$ level of inbreeding ($\times$ 100) $F_t$, $t$ = 5, 10, 30, rate of inbreeding ($\times$ 100), $\Delta F$, and probability of coselection of male half and full sibs, P(HS) and P(FS), for the basic population structure and alternatives 2–5.

| *Sire selection strategy* | Nm | $R_5$ | $R_{10}$ | $R_{30}$ | $F_5$ | $F_{10}$ | $F_{30}$ | $\Delta F$ | *P(HS)* | *P(FS)* |
|---|---|---|---|---|---|---|---|---|---|---|
| Basic scheme: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 0$, $Q = 6.11$ | | | | | | | | | | |
| EBV | 12 | 1.89 | 3.95 | 10.75 | 7.0 | 17.5 | 49.1 | 2.31 | 0.169 | 0.094 |
| 1 | 9 | 1.87 | 3.94 | 11.10 | 5.6 | 13.8 | 40.4 | 1.77 | 0.108 | 0.035 |
| 2 | 9.6 ± 0.07 | 1.86 | 3.93 | 11.17 | 5.5 | 13.5 | 40.3 | 1.77 | 0.106 | 0.035 |
| 3 | 14.0 ± 0.09 | 1.92 | 4.04 | 11.39 | 5.7 | 13.7 | 40.2 | 1.76 | 0.111 | 0.035 |
| Max se | | 0.015 | 0.021 | 0.032 | 0.13 | 0.19 | 0.25 | 0.072 | 0.0064 | 0.0013 |
| Alternative 2: $N_f = 25$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 0$, $Q = 5.12$ | | | | | | | | | | |
| EBV | 9 | 1.46 | 3.03 | 7.86 | 10.6 | 24.2 | 58.4 | 2.96 | 0.119 | 0.163 |
| 1 | 6 | 1.49 | 3.11 | 8.22 | 9.5 | 21.7 | 54.4 | 2.66 | 0.121 | 0.072 |
| 2 | 6.8 ± 0.07 | 1.51 | 3.14 | 8.34 | 9.4 | 21.2 | 54.5 | 2.68 | 0.120 | 0.075 |
| 3 | 9.2 ± 0.07 | 1.54 | 3.20 | 8.48 | 9.3 | 21.1 | 53.8 | 2.62 | 0.119 | 0.078 |
| Max se | | 0.009 | 0.013 | 0.020 | 0.09 | 0.13 | 0.14 | 0.078 | 0.0040 | 0.0016 |
| Alternative 3: $N_f = 100$, $N_{sex} = 2$, $h^2 = 0.4$, $D = 0$, $Q = 5.42$ | | | | | | | | | | |
| EBV | 9 | 1.68 | 3.56 | 9.67 | 7.1 | 18.1 | 50.3 | 2.38 | 0.217 | 0.046 |
| 1 | 9 | 1.67 | 3.53 | 9.93 | 5.6 | 13.7 | 39.7 | 1.73 | 0.116 | 0.021 |
| 2 | 8.0 ± 0.05 | 1.70 | 3.58 | 9.98 | 6.2 | 15.0 | 44.0 | 1.98 | 0.135 | 0.021 |
| 3 | 12.0 ± 0.08 | 1.76 | 3.69 | 10.22 | 5.9 | 14.7 | 43.3 | 1.94 | 0.135 | 0.020 |
| Max se | | 0.016 | 0.023 | 0.036 | 0.09 | 0.18 | 0.18 | 0.076 | 0.0080 | 0.0010 |
| Alternative 4: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.1$, $D = 0$, $Q = 1.66$ | | | | | | | | | | |
| EBV | 15 | 0.64 | 1.34 | 3.58 | 7.5 | 19.1 | 52.3 | 2.58 | 0.204 | 0.119 |
| 1 | 9 | 0.65 | 1.40 | 3.73 | 7.2 | 18.4 | 51.1 | 2.53 | 0.184 | 0.076 |
| 2 | 8.9 ± 0.06 | 0.65 | 1.37 | 3.71 | 7.3 | 18.6 | 52.5 | 2.54 | 0.200 | 0.077 |
| 3 | 12.2 ± 0.08 | 0.65 | 1.39 | 3.69 | 7.0 | 18.4 | 52.0 | 2.51 | 0.174 | 0.087 |
| Max se | | 0.009 | 0.012 | 0.017 | 0.12 | 0.21 | 0.25 | 0.084 | 0.0074 | 0.0016 |
| Alternative 5: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $F = 3.33$, $Q = 54.40$ | | | | | | | | | | |
| EBV | 21 | 1.54 | 3.27 | 9.41 | 4.5 | 11.5 | 33.6 | 1.40 | 0.081 | 0.060 |
| 1 | 12 | 1.64 | 3.43 | 9.94 | 4.4 | 10.7 | 31.5 | 1.29 | 0.072 | 0.026 |
| 2 | 13.7 ± 0.09 | 1.63 | 3.43 | 10.00 | 4.1 | 9.9 | 29.7 | 1.21 | 0.065 | 0.026 |
| 3 | 18.8 ± 0.15 | 1.69 | 3.52 | 10.24 | 4.2 | 10.1 | 30.4 | 1.24 | 0.071 | 0.027 |
| Max se | | 0.013 | 0.018 | 0.027 | 0.06 | 0.11 | 0.18 | 0.042 | 0.0032 | 0.0004 |

The selection goal is $R_{30}$. For sire selection strategies EBV and 1, results are presented for the $N_m$ value (out of those examined) that maximised $R_{30}$. For sire selection strategies 2 and 3, the number of sires used to predict $Q$ is equal to the number sires used in strategy 1.

Comparing each alternative to the basic scheme for the optimal sire selection strategies 2 and 3, the following observations can be made: in alternative 2, $N = 25$, response is less and inbreeding is higher, due to the smaller number of both female and male parents. Despite the smaller optimum number of sires, the mating ratio (dams/sire) is decreased. In alternative 3, $N_{sex} = 2$, the lower inbreeding encouraged by the smaller family size is counterbalanced by the increase in inbreeding caused by the smaller optimum number of sires and encouraged by the smaller $Q$ value. In alternative 4, $h^2 = 0.1$, strategies 1–3 are all superior to strategy EBV, but there is little to choose between them. For strategy EBV, $\Delta F$ is higher with 15 sires than it is with 12 sires in the basic scheme. Strategies 2 and 3 choose a lower optimum number of sires than in the basic scheme and results in a higher optimum rate of inbreeding. In alternative 5, $D = 3.33$, the value of $Q$ is increased by a factor of 8.9, which discourages the coselection of sibs and results in considerably lower rates of inbreeding. In this alternative, $R_t$ represents genetic merit in generation $t$ less $DF$, whereas in the other alternatives $R_t$ is simply genetic merit; for strategy 3, the average genetic merit is $11.25 \pm 0.030$ compared to $11.39 \pm 0.032$ in the basic scheme. All results for alternative 5 are plotted in figure 2.

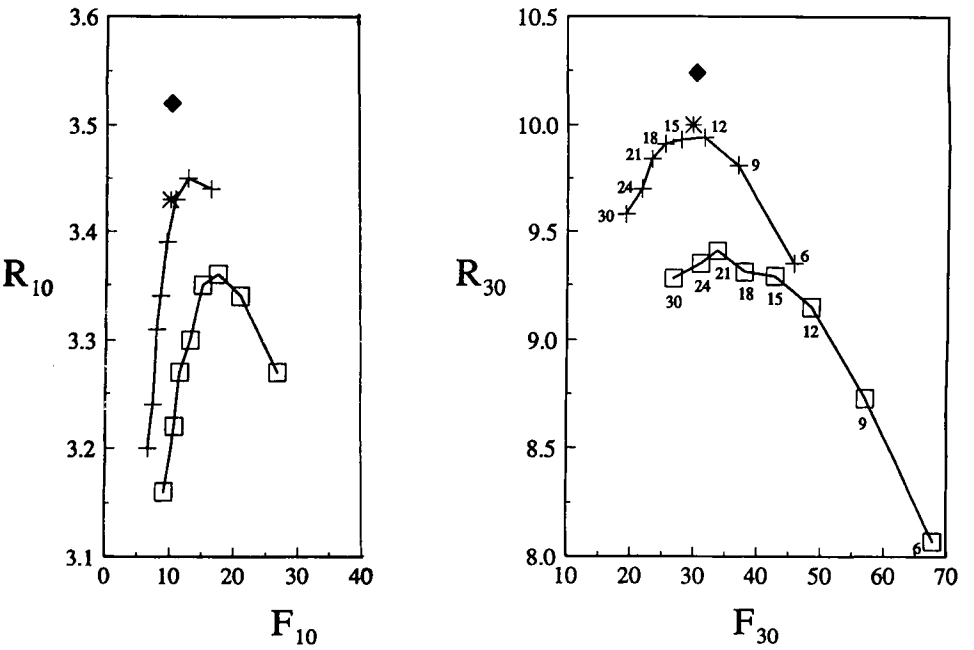



**Fig 2.** Response to selection $R_t$ against level of inbreeding $F_t$, for $t = 10$ and 30 when the breeding goal is $R_{30}$. Alternative 5: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 3.33$. Sire selection strategies: EBV: □, 1: +, 2: *, 3: ♦. For strategies EBV and 1, each point represents a different fixed number of sires selected each generation; the numbers of sires are marked on the $t = 30$ graph, but are equivalent on the $t = 10$ graph.

In figure 3 response to selection $R_{10}$ and $R_{30}$ are plotted against level of inbreeding $F_{10}$ and $F_{30}$ respectively for alternative 6, $D = 3.33$, breeding objective $R_{10}$, for sire selection strategies EBV and 1–3, which can be compared directly with figure 2, where the breeding objective is $R_{30}$. As expected, for alternative 6, sire selection strategies rank 3 > 2 > 1 > EBV for $R_{10}$, with optimum number of sires for strategies EBV, 1–3 being 12, 12, $8.3 \pm 0.07$ and $12.2 \pm 0.10$ respectively. However, if the same selection criterion is continued until generation 30, then the superiority of strategies 2 and 3 is lost.

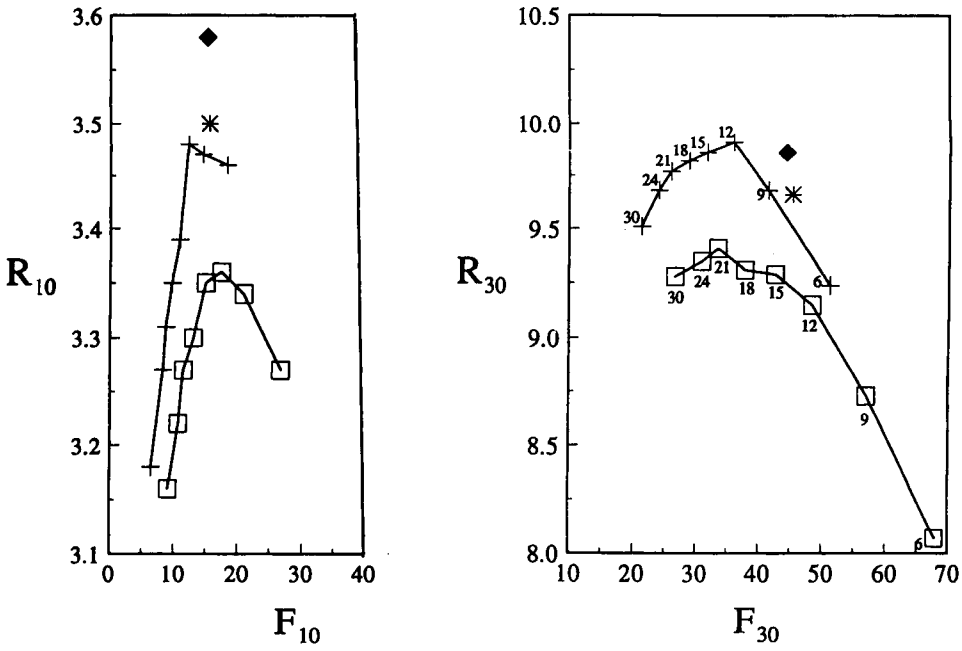



**Fig 3.** Response to selection $R_t$ against level of inbreeding $F_t$, for $t = 10$ and 30 when the breeding goal is $R_{10}$. Alternative 6: $N_f = 100$, $N_{sex} = 4$, $h^2 = 0.4$, $D = 3.33$. Sire selection strategies: EBV: □, 1: +, 2: *, 3: ♦. For strategies EBV and 1, each point represents a different fixed number of sires selected each generation; the numbers of sires are marked on the $t = 30$ graph, but are equivalent on the $t = 10$ graph.

## DISCUSSION

Selection on the criterion $V$ (equation [7]) always results in a higher response criterion $R_{30}$ and lower level of inbreeding $F_{30}$ when the breeding objective is $R_{30}$, than selection on EBV alone for the simulation examples considered. When the number of sires selected is fixed, achieving the maximum response to $R_{30}$ depends on the judicious choice of the number of sires. The algorithm to select a variable number of sires each year always resulted in an $R_{30}$ at least as high at approximately

the same $F_{30}$ as with the strategy using the optimum fixed number of sires. Selection on criterion $V$ with $Q$ as defined in equation [3] does not necessarily result in the absolute maximum response to the breeding objective. This is because the derivation of $Q$ contains several approximations: i) the equality of equations [1] and [2]; ii) prediction of $r_L$ and $\sigma_{G,L}$; and iii) calculation of selection intensity as if the sires have been selected on EBV alone. Of these, iii) is likely to be most critical, but it is difficult to see how to improve on this approximation as there is a dynamic interaction each generation between the value of $Q$ used and the selection decisions made (in which genetic merit is balanced against relatedness of the selected group) and hence the selection intensity achieved. Fortunately, the simulation results suggest that the algorithm for selecting sires is fairly robust to the value of $Q$ chosen, and it appears that the method proposed here to predict $Q$ results in response close to the maximum. In the prediction of $Q$, it is likely that the underprediction in approximation ii) counterbalances to some extent the overprediction implied by approximation iii).

When selection is on EBV alone the rate of inbreeding cannot be accurately predicted from single generation probabilities of coselection of sibs (or equivalently variance of family size) (Wray *et al*, 1990) which can be explained through the concept of partial inheritance of selective advantage (Wray and Thompson, 1990) across generations. In the $V$ selection criterion, the tendency for an ancestor of high genetic merit to leave more descendants in each generation is limited by the continual reevaluation of the relationship information in each generation's selection decisions.

The advantage of the selection criterion proposed here is that it is clearly defined, with the goal of maintaining genetic variance over a long time horizon. In other methods, maintenance of genetic variance is the underlying goal but it is indirectly achieved by criteria for which the optimum values are not known. For example, if the selection criterion includes a restriction on the number of individuals to select from any sibship (*eg*, Toro and Perez-Enciso, 1990; Brisbane and Gibson, 1993), what should the restriction be? The criterion proposed here automatically places restrictions on the number selected per family if inbreeding is perceived as a problem, but will be flexible in its restrictions, placing less restrictions on a family which is highly superior in genetic merit. Alternatively, if selection is on a criterion which alters the emphasis placed on individual and family information, $\mathrm{EBV}_i - 1/2\mathrm{EBV}_{\mathrm{sire}} - 1/2\mathrm{EBV}_{\mathrm{dam}} + \mathrm{weight}_{\mathrm{sire}}\mathrm{EBV}_{\mathrm{sire}} + \mathrm{weight}_{\mathrm{dam}}\mathrm{EBV}_{\mathrm{dam}}$ (*eg*, Toro and Perez-Enciso, 1990; Verrier *et al*, 1993; Villanueva *et al*, 1994) what values should be attributed to $\mathrm{weight}_{\mathrm{sire}}$ and $\mathrm{weight}_{\mathrm{dam}}$ and should the weights be constant over generations? The method proposed here could be viewed as a flexible version of this criterion with weights given to family information differing for each individual. In addition, the weights may differ over generations, where one could speculate that, initially it may be favourable to eliminate the genetically poorer families, whilst in later generations within-family selection from each of the genetically similar families might be optimal. Even when the selection criterion specifies directly a restriction on rate of inbreeding (*eg*, Toro and Perez-Enciso, 1990), what is the optimal restriction to place on rate of inbreeding? The optimal rate must be dependent on the value attributed to inbreeding depression, and in the absence of inbreeding depression, is entirely dependent on maintenance of genetic variance.

The selection criterion given by $V$, is most similar to that proposed by Goddard and Smith (1990a). They were concerned with inbreeding depression but ignored loss of genetic variance, *ie* $Q = D$. A similar criterion was also investigated by Brisbane and Gibson (1993) (called ADJEBV), in which both sires and dams are selected on an adjusted EBV

$$\mathrm{EBV}_{\mathrm{ADJ},j} = \mathrm{EBV}_j - K(X_{\mathrm{sires}} + X_{\mathrm{dams}})$$

where $X_{\mathrm{sires}}(X_{\mathrm{dams}})$ is the mean genetic relationship of sire (dam) $j$ with the other selected sires (dams). Their $K$ is equivalent to our $Q$. Whilst we attempt to predict $Q$, they examined a range of $K$ values. They preferred the empirical approach as they found, for example, that population size had a small effect on the optimum value of $K$ (J Gibson, personal communication). As we have discussed above, our prediction of $Q$ will not find the exact optimum achieved by detailed empirical searching, however, we believe it is useful to have some understanding about how the optimum value for $Q$ or $K$ arises. The selection algorithm of Brisbane and Gibson (1993) chooses the same number of parents each year and allocates equal proportions of matings to them (equivalent to our strategy 1). The selected group is initially those with the highest $\mathrm{EBV}_j$ values, $\mathrm{EBV}_{\mathrm{ADJ},j}$ are then calculated for all animals and the highest ranking individual not selected replaces the lowest ranking selected individual. Iterations of swapping selected parents continue until no more changes are necessary. Their iterative procedure for selecting parents may result in different selected group to the algorithm proposed here, which may be closer to, but it still does not necessarily find 'the' best group as substitutions are only made one at a time. Depending on the value of $Q(K)$ and the number of sires to be selected, their iterative algorithm may be faster (low $Q$ and high $N_m$) at determining the selected group. The optimum selected group may be more consistently found by the Annealing algorithm as used by Meuwissen and Woolliams (1994) in the related problem of maximisation of genetic response with restricted variance of response.

Toro and Nieto (1984) proposed a method to maintain selection intensity but increase effective population size: by selection of an increased number of parents and weighting their use (higher ranking being allocated more matings) so as to ensure a selection intensity equal to selecting a smaller number of parents. The implied benefits in effective population size, may not be as great as expected in the long term, because the selection policy gives higher ranking parents more chances to leave descendants (through inheritance of selective advantage, Wray and Thompson, 1990b). Toro *et al* (1988) investigated this method of selection (called weighted selection) over 30 generations, arbitrarily selecting twice the standard number of parents, and found benefits in response to selection in all generations and benefits in inbreeding in the long term over selection of the standard number of parents. The algorithm proposed here to select a variable number of sires with a variable number of matings/sire utilises the same concept as proposed in Toro and Nieto (1984). However, it determines the number of sires to select by a non-arbitrary criterion and attempts to consider the impact of the decisions on inbreeding and hence future response, in which the highest proportion of matings is not necessarily allocated to the sire with the highest EBV.

For simplicity, we have only considered selection of sires with random selection of dams. If dams are selected on EBV the benefits of selection of sires on strategies

1–3 as demonstrated here are expected to remain. In most livestock populations, the number of dams selected is high and so their impact on gain and particularly inbreeding is small compared to the sires. Therefore, it is less likely to be worthwhile to consider optimum selection strategies for dams. However, in breeding schemes using multiple ovulation and embryo transfer, the number of dams selected can be small and the scope for selection of a variable number of dams and allocating them a variable proportion of offspring testing places is greater. The methods proposed here could be extended to consider selection of dams (*via* equation [6]), at which point it is appropriate to consider mating combinations. Mating designs obviously affect inbreeding in the next generation, but in general they are of lesser importance than the selection criterion in controlling long-term response, but some benefits in limiting inbreeding and particularly variance of inbreeding can be achieved by the optimum choice of mates (Jansen and Wilton, 1985; Toro *et al*, 1988; Woolliams, 1989; Toro and Perez-Enciso, 1990; Toro and Silio, 1992). The breeding objective proposed in equation [1] is quite general, but could be generalized further by consideration of discounted gain over a given time period. Goddard and Wray (unpublished results) derive the equivalent of $Q$ if the objective is to maximise total discounted benefits over a future time period. Woolliams and Meuwissen (1993) investigated selection goals in which a value is assigned to risk, defined as variance of response. They proposed dynamic selection rules in which an individual's EBV was balanced with its accuracy of prediction. Selection goals that include risk could be incorporated into the framework presented here.

In summary, the selection algorithm proposed as strategy 3 uses the defined selection goal to determine the best balance of selection intensity and inbreeding and then optimises the selection decisions by i) deciding the number of sires to be used; ii) deciding the number of offspring to be born per sire; iii) selecting sires based on their EBV and relationship to other sires; and iv) utilising year-to-year variations in the actual sires available when making the decisions i)–iii).

## ACKNOWLEDGMENTS

This work was funded by the Australian Wool Research and Development Corporation under projects DAV111 and 112. We would like to thank J Gibson, J Woolliams and 2 anonymous referees for their comments on the manuscript.

## REFERENCES

Brisbane JR, Gibson JP (1993) Selection methods to reduce inbreeding with minimal effects on genetic progress. *J Dairy Sci* 76 (supp 1), 292

Bulmer MG (1971) The effect of selection on genetic variability. *Am Nat* 105, 201-211

Bulmer MG (1980) *The Mathematical Theory of Quantitative Genetics.* Clarendon Press, Oxford

Dekkers JCM (1992) Asymptotic response to selection on best linear unbiased predictors of breeding values. *Anim Prod* 54, 351-360

Dempfle L (1975) A note on increasing the limit of selection through selection within families. *Genet Res* 24, 127-135

Goddard ME, Smith C (1990a) Adjustment of sires' estimated breeding values for the prospective inbreeding impact on the breed. *J Dairy Sci* 73 (supp 1), 233

Goddard ME, Smith C (1990b) Optimum number of bull sires in dairy cattle breeding. *J Dairy Sci* 73, 1113-1122

Grundy B, Hill WG (1993) A method of reducing inbreeding with best linear unbiased prediction. *Anim Prod* 56, 427

Jansen GB, Wilton JW (1985) Selecting mating paris with linear programming techniques. *J Dairy Sci* 68, 1302-1305

Jodar B, Lopez-Fanjul C (1977) Optimum proportions selected with unequal sex numbers. *Theor Appl Genet* 50, 57-61

Lindgren D (1991) Optimal utilization of genetic resources. *Forest Tree Imp* 23, 49-67

Meuwissen THE, Woolliams JA (1994) Maximizing genetic response in breeding schemes of dairy cattle with constraints on variance of response *J Dairy Sci* 77, 1905-1916

Quinton M, Smith C, Goddard ME (1992) Comparison of selection methods at the same level of inbreeding. *J Anim Sci* 70, 1060-1067

Robertson A (1970) Some optimum problems in individual selection. *Theor Pop Biol* 1, 120-127

Toro M, Nieto B (1984) A simple method for increasing the response to artificial selection. *Genet Res* 44, 47-349

Toro M, Nieto B, Salgado C (1988) A note on minimisation of inbreeding in small scale breeding programmes. *Livest Prod Sci* 20, 317-323

Toro M, Perez-Enciso M (1990) Optimisation of selection response under restricted inbreeding. *Genet Sel Evol* 22, 93-107

Toro M, Silio L (1992) Consequences of mixed model methods for population structure and inbreeding. *Proc EAAP 43rd Ann Meeting, 13–17 September, Madrid, Spain*

Verrier E, Colleau JJ, Foulley JL (1993) Long-term effects of selection based on the animal model BLUP in a finite population. *Theor Appl Genet* 87, 446-454

Villanueva B, Woolliams JA, Simm G (1994) Strategies for controlling inbreeding in adult MOET nucleus schemes for beef cattle. *Genet Sel Evol* 26 (in press)

Woolliams JA (1989) Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Anim Prod* 49, 1-14

Woolliams JA, Meuwissen THE (1993) Decision rules and variance of response in breeding schemes. *Anim Prod* 56, 179-186

Wray NR, Hill WG (1989) Asymptotic rates of response to selection. *Anim Prod* 49, 217-227

Wray NR, Thompson R (1990a) Advances in selection theory. *In: Proc 4th World Congress Genetics Applied to Livestock Production, 23-27 July, Edinburgh, Scotland*, Vol XIII, 167-176

Wray NR, Thompson R (1990b) Prediction of rates of inbreeding in selected populations. *Genet Res, Camb* 55, 41-54

Wray NR, Woolliams JA, Thompson R (1990) Methods of predicting rates of inbreeding in selected populations. *Theor Appl Genet* 80, 503-512

Wright (1931) Evolution in mendelian populations. *Genetics* 16, 97-159