

# Optimizing design of breeding programs

Julius van der Werf

---

## 1 Introduction

In the previous lectures we discussed criteria for comparing breeding programs. We showed how to predict genetic response for multiple traits and how to evaluate the value of this response economically. It should be noted that the design of those breeding programs was fixed and breeding alternatives can be compared by varying different parameters. However, in practical breeding programs, we often need a more dynamic approach that can help give optimal selection results. Such optimization is needed when optimizing decisions under prevailing circumstances, i.e. for tactical decisions. But also for strategic planning, we can use some dynamic approaches that help to reconsider design parameters when circumstances change. For example, when a trait is measured at an earlier age, it would probably be more optimal to select younger animals, as they will have more accurate EBVs. So selection accuracy and generation interval both change.

In the design of a breeding program, many aspects that determine genetic response are interdependent, and changing one variable might result in different optimal values for other variables. Remember the central 'dogma' (Rendel and Robertson, 1950) of predicting genetic response:

$$R = \frac{i_m r_m + i_f r_f}{L_m + L_f} \sigma_A$$

where  $i$  = selection intensity,  $r$  = selection accuracy and  $L$  is generation interval. The subscripts refer to selection for males and females and  $\sigma_A$  is genetic variance. In some examples like in dairy programs where breeding males are very important due to the use of AI, four paths are distinguished as the best "elite" parents are especially selected to generate the next generation of AI bulls. .

The above formula is central as it allows to put in perspective the importance of the different components in a breeding program.

Van Arendonk and Bijma (2002) have succinctly summarized these components as follows:

### Genetic variance.

The genetic variance in selection candidates is equal to :

$$\sigma_A^2 = \frac{1}{4} \sigma_{As}^2 + \frac{1}{4} \sigma_{Ad}^2 + \frac{1}{2} \sigma_{Am}^2$$

Where  $\sigma_{Ax}^2$  is the genetic variance in the selected sires ( $x=s$ ) and dams ( $x=d$ ) and  $\sigma_{Am}^2$  is the Mendelian sampling variance. In an unselected and non-inbred population,  $\sigma_{As}^2 = \sigma_{Ad}^2 = \sigma_A^2$  and  $\sigma_{Am}^2 = \frac{1}{2} \sigma_A^2$  which means that sires and dams contribute each 25 % to the genetic variance in an individual and that 50 % of the genetic variance is due to Mendelian sampling. Selection, however, results in a reduction of the genetic variance in the selected parents, the so-called Bulmer effect. The genetic variance in the selected parents ( $x$ ) is equal to (Bulmer, 1971):

$$\sigma_{At+1}^2 = (1 - \frac{1}{2} r^2 k) \sigma_{At}^2$$

where  $k$  is the variance reduction coefficient calculated as  $i \cdot (i - x)$  ( $i$  being selection intensity,  $x$  is truncation point) and  $r$  is the accuracy of selection.

The Bulmer effect leads to a reduction in genetic gain because the genetic gain is a direct function of the genetic variance. The variance reduction in the population is often close to 25 %, which leads to a 13 % reduction in the absolute gain. More importantly, however, the Bulmer effect reduces the genetic variance in parents (between family variance) but not the Mendelian sampling variance (within family variance). As a consequence, full and half-sib information becomes less important whereas information that includes the Mendelian sampling component of the selection candidate, such as own performance and progeny information, becomes relatively more important. These effects are important when comparing schemes that use different sources of information (e.g. progeny vs. sib information).

### **Accuracy of selection.**

The accuracy of selection is the correlation between the selection criteria and the true breeding value for the breeding goal that is to be improved. In most livestock improvement schemes, selection is based on breeding values that are estimated using BLUP. A method use in prediction models for calculating the accuracy of selection on BLUP-EBV was presented by Wray and Hill (1989), based on selection index theory. A multiple-trait extension of this method was presented by Villanueva *et al.* (1993), and Bijma *et al.* (2001) presented an extension for overlapping generations. These methods account for the reduction of genetic variance due to selection.

### **Selection intensity.**

In predicting response to selection, it is generally assumed that the selection criterion is normally distributed and that truncation selection is applied. In that case the selection intensity can be obtained directly from the proportion of animals selected. When the selection criterion is partly based on family information, the EBVs of sibs are correlated. Meuwissen (1991) developed a method to account for finite population and correlated EBV. This correction is particularly important in breeding schemes that rely heavily on information coming from full and half sibs and where the number of selected parents is small

### **Overlapping generations.**

In most populations, a number of age classes can be distinguished and the amount of information differs between age classes. In general, young age classes have less information than older age classes. Because older age classes have more information, they have higher accuracy. However, the mean level of the EBV of older age classes will be lower than that of younger age classes due to the continuous genetic improvement in the population. Truncation selection across age classes can be performed to obtain the highest selection differential (James, 1987). Mathematical details on truncation across age classes can be found in Ducrocq and Quaas (1988) and Bijma *et al.* (2001) and an algorithm will also be presented in the next section. Reproductive techniques might increase the amount of sib information and thereby increase the accuracy of EBV of younger age classes. This will change the fraction of parents selected from the younger age classes and thereby also influence the average generation interval.

### **Rate Of Inbreeding**

The magnitude of inbreeding at the population level is measured by the rate of inbreeding ( $F$ ). Only in the absence of selection  $F$  is related directly to the number of sires and dams. In selected populations, this equation is no longer valid because parents contribute unequally to the next generation. Wray and Thompson (1990) introduced methods to predict rates of inbreeding in selected populations, based on the concept of long-term genetic contributions. Recently, Woolliams *et al.* (1999) and Woolliams and Bijma (2000) developed a general theory to predict rates of inbreeding in populations undergoing selection. These methods facilitate a deterministic optimisation of short and long-term response of breeding schemes. Bijma and Woolliams (2000)

demonstrated that with BLUP selection, the number of candidates per parent (selection intensity) may be as or more important than the number of parents. Doubling the number of parents failed to halve the rate of inbreeding.

Meuwissen (1997) introduced a dynamic selection tool to maximise the genetic gain while restricting the rate of inbreeding. Given the available selection candidates, the method maximises the genetic level of the selected group of parents while constraining the average coefficient of coancestry. Implementation of this method results in a dynamic breeding program, where the number of parents and number of offspring per parent may vary, depending on the candidates available in a particular generation. In a next section, we'll discuss in more detail methods to restrict inbreeding.

## **Optimization of Breeding Schemes**

Under the infinitesimal model, inbreeding reduces genetic variation, which in turn reduces genetic gain. Furthermore, when inbreeding depression is present, fitness of the population may reduce to an extent where it affects the selection differentials, *i.e.* indirectly inbreeding may also reduce genetic gain. In the short term, inbreeding and genetic gain have an unfavourable relation, in the sense that maximising short-term response by selecting fewer parents reduces long-term response and involves substantial risk (e.g. Verrier *et al.*, 1993). To balance the short and long-term response a restriction on the rate of inbreeding is required (e.g. Quinton and Smith, 1995). The objective in optimised breeding strategies needs to be maximising genetic gain while restricting inbreeding. Acceptable levels of inbreeding are difficult to determine and are discussed by Bijma (2000), who indicated that inbreeding depression is probably the most important issue. Though detailed knowledge of the relevant parameters to determine the level of the constraint is lacking, different approaches point towards values around 0.5 % and 1 % per generation.

### **Different components of genetic gain interact**

Alternative options for breeding programs need to be assessed, which can be done based on an analysis of the most important factors that determine rate of genetic gain: selection intensity, selection accuracy and generation interval, as in the Rendel and Robertson formula.

It should be pointed out that the different factors interact. For example, one could try to increase selection intensity, but the result is that breeding animals can be less rapidly replaced when fewer young animals are selected, and the generation interval is increased. The most important interactions are:

- Generation interval versus selection accuracy  
Selection young animals will not only lead to short generation intervals, but may also imply lower selection accuracy because young animals have generally less information available (no repeated records, maybe no own performance, no progeny test)
- Generation interval versus selection intensity  
If more young animals are retained as breeders, and a high replacement rate is applied, the generation interval may be shorter, but selection intensity will also be lower since more animals of the newborn generation are needed as replacements.
- Balancing different traits (see course notes on multiple trait selection)

### **Strategic vs tactical optimization of breeding programs**

Strategic optimization requires a whole population model of the breeding program. The model accommodates parameters that can represent different strategies, e.g. which traits to measure, which animals to measure, how many animals to measure in multiple selection stages, how many

to select from each age class, how many from each family. A strategic plan for a breeding program often results in a set of rules, e.g. optimal proportion selected, optimal age structures etc.

By contrast, a tactical optimization tries to address the immediate decisions that need to be made under prevailing circumstances. Maybe optimally we use 30% of the bulls from the 1<sup>st</sup> year age class, but a certain crop might be disappointing, and a lower percentage would be optimal for that year. Tools that can be used for tactical decision making, i.e. selection, are BLUP or mate selection, which is implemented in TGRM (Kinghorn et al., 2002) or MateSel. A tactical approach will make use of knowledge of the full range of actual animals available for breeding at the time of decision making, as well as other factors such as availability of mating paddocks, current costs of specified semen, current quarantine restrictions on animal migration, current or projected market prices, etc. Tactical implementation of breeding programs gives the power to capitalise on prevailing opportunities - opportunities that would often be missed when adhering to a set of rules. The tactical approach to breeding is driven by specifying desired outcomes. Although mate selection analysis is a very powerful computing tool, the results that it gives are closely aligned to the 'outcome instructions' that it receives. This means that the breeder can have a high degree of control, not by specifying which animals should be selected, but by specifying desires in terms of direction of genetic change, maintenance of genetic diversity, limits in money spent, constraints to be satisfied etc.

Prediction of breeding program outcomes can be based either on a stochastic model, (stochastic simulation) or through a deterministic model (deterministic simulation).

Briefly these can be contrasted as

- A stochastic simulation generates breeding and phenotypic values for each simulated individual, select based on such data (e.g. using BLUP) and simply select the best individuals as parents of the next generation.
- Deterministic simulation uses a whole population model, e.g. characterised by a mean and a variance, and predicts response to selection based on knowledge of population characteristics, e.g. using selection intensity.
- Stochastic simulation is relatively simple, but requires repeated running of a simulated breeding program. For example, there is no need to worry about 'optimal age structure, as simple selection of the best animals based on BLUP EBV automatically optimizes generation interval (see a later section). Because of the stochastic process, outcomes will vary (like in a real breeding program!), and a mean of outcomes is often required to demonstrate differences between programs. Hence, stochastic simulation has much higher computer requirements. Although computers are fast these days, it is still quite a task to simulate a set of replicated populations, each with many generations for thousands of animals for several traits
- Deterministic simulation requires formulas to adjust for things like the Bulmer effect, reduced selection intensities in small populations, correlations between EBVs, prediction of inbreeding etc. It is therefore more complicated and possibly more approximate.
- Although computer capacity is not anymore so restrictive these days, it is still useful to have a simple deterministic method for predicting response in alternative breeding programs. Deterministic prediction models also require more insight and therefore are often more illuminating in detecting the role of different components in genetic change.

In these notes we will look at some specific problems that are relevant in optimizing breeding programs,

A major component of decision making in breeding programs is to achieve an (economically) balanced progress across multiple traits.

From Armidale Summer Course 2006:  
<http://www.personal.une.edu.au/~jvanderw/BPDesignJJ.pdf>

## Balancing long-term and short-term selection response

*Prof J. W. James, Faculty of Veterinary Science, University of Sydney.*

### SHORT TERM RESPONSE

When we learn about designing animal breeding programs we usually begin with the classic equation of Rendel and Robertson (1950):

$$\Delta G = \frac{\sum i r \sigma_A}{\sum l}$$

where  $i$  = standardised selection differential,  
 $r$  = correlation between true and estimated breeding values;  
 $\sigma_A$  = standard deviation of true breeding value;  
 $l$  = generation interval;  
 $\Delta G$  = gain in mean breeding value;

and the summations are over all relevant pathways of improvement such as males to breed males, males to breed females, and so on.

If  $r$  is constant over all pathways, this equation reduces to

$$\Delta G = r \sigma_A \frac{\sum i}{\sum l}$$

and maximising response to selection involves maximising the ratio  $\sum i / \sum l$ . There is normally some intermediate generation length which maximises genetic gain, though with very high reproductive rates it may be best to turn over generations as quickly as possible.

If the amount of information on animals changes with age, as it often does, the optimum age structure will be affected, and in fact we can avoid having to specify the age structure which will maximise gain by noting that the greatest gain is achieved if the required number of parents are selected from among all those available on the basis of their estimated breeding values. This procedure will automatically give the best age structure (James, 1987). In fact, more generally, if the accuracy of estimated breeding values varies within age groups in any way at all, selection on EBVs will deal with all difficulties. In this sense, BLUP solves many of the problems which arise in the design of breeding programs. But the problems which it solves are those arising from making use of available data to select a specified number of animals. What it does not do is tell us what data to obtain, nor how many animals to select. Some aspects of these two questions will be dealt with in these notes.

First we will deal with the question of how many animals should be selected from among those available. To some extent this is governed by rates of reproduction: we need to select enough parents to produce the required number of offspring. This may fix the number of breeding females needed, at least within limits, but the number of males needed will generally be considerably smaller. With AI, the male reproductive rate may be very high indeed. Of course, the number of females required will be reduced if MOET is used, but the number of males needed will still usually be smaller, so that the opportunity arises for a stringent selection of males.

## INBREEDING AND GENETIC DRIFT

If we take the basic equation and seek to maximise gain we see at once that  $i$  should be as large as possible, but in practice this is never done, and the number selected is larger than necessary for breeding the necessary number of progeny. The reason for this is that with very few parents selected the rate of inbreeding will be increased to an undesirable level. So it is common to set a minimum number of parents to limit inbreeding to an “acceptable” level.

There are three main effects of inbreeding, or limited population size, on response to selection.

1. Most traits of economic importance show some degree of inbreeding depression
2. The available genetic variance is reduced by inbreeding, so that future response is less than if there were no inbreeding.
3. Variation due to new mutations accrues more slowly, so that in later generations this source of variation contributes less to genetic gain.

## LONG TERM RESPONSE

If we assume that selection is for a single trait in a monoecious population with discrete generations we can write simple approximate expressions for the effects of these three factors on genetic gain up to the  $t^{\text{th}}$  generation.

Let the effective population size be  $N$  and let  $F_t$  be the population average inbreeding coefficient in generation  $t$ . Then

$$F_t = 1 - \left(1 - \frac{1}{2N}\right)^t$$

The inbreeding depression is  $-DF_t$ , where  $D$  is the regression of performance on inbreeding, assumed linear, so that  $D$  is the loss of performance in a completely inbred population.

The gain from exploitation of the original genetic variance in the base population is approximately

$$2Ni \frac{V_A}{\sigma_P} F_t$$

where  $\sigma_P$  is the phenotypic standard deviation. This expression must be multiplied by  $\left(1 - \frac{1}{2n}\right)$  if the selection program is started with a sample of  $n$  founders from the base population.

The gain from mutational variance is

$$2Ni \frac{V_M}{\sigma_P} [t - 2NF_t]$$

where  $V_M$  is the new mutational variance arising each generation.

Putting all these factors together, we have for the response in generation  $t$

$$2Ni \frac{V_A}{\sigma_P} \left(1 - \frac{1}{2n}\right) F_t + 2Ni \frac{V_M}{\sigma_P} (t - 2NF_t) - DF_t.$$

As  $t$  becomes large,  $F_t$  approaches unity and we find that the mean  $M_t$  becomes

$$M_t = 2Ni \frac{V_A}{\sigma_p} \left[ \left(1 - \frac{1}{2n}\right) + \frac{V_M}{V_A} (t - 2N) \right] - D.$$

This approximation ignores inbreeding depression arising from new mutational variance and so will overestimate the genetic gain from mutation. This aspect of the model could be improved, but it will not be pursued here. It can be seen that when  $t$  is large the only continued improvement

comes from mutation, and that the gain per generation from this source is  $2Ni \frac{V_M}{\sigma_p}$ .

In some circumstances, such as when terminal sire breeds are bred so that inbreeding which accumulates in the pure-bred population is removed in the cross-bred progeny, or when the trait(s) selected show negligible depression, the term in  $D$  becomes unimportant. The value of  $V_M/V_A$  is likely to be small, less than 0.01 in most cases, so this factor is likely to be of importance only in the long term, where it is the sole source of progress.

In any case, we see that in the long term the greatest progress is made by maximising  $Ni$ . For a normally distributed selection criterion it is well known that the standardised selection differential is given by  $i = z/p$ , where  $p$  is the proportion selected, and  $z$  is the ordinate of the normal curve at the abscissa  $x$  truncating a fraction  $p$  in the upper tail of the distribution. If  $T$  is the total number of animals available for selection, the number of parents selected is  $N = Tp$ , so that  $Ni = Tz$  and has a maximum at the mode of the normal curve, when  $x = 0$  and  $p = 0.5$ . This result has been known for a long time, having been pointed out by Dempster (1955) and Robertson (1960).

## BALANCING LONG AND SHORT TERM GAINS

However, although this gives the greatest long-term response the gain in the short term will be significantly less than would be achieved by selecting more intensely, and a breeder who adopted such a policy might well be out of business long before his strategy could pay off. So we need to consider a compromise between short-term and long-term gains. How can we make a sensible choice?

Most recommendations are made on the basis of a more or less arbitrary choice of an “acceptable” rate of inbreeding which will determine the number of parents to be used for breeding. While this probably works reasonably well, it does leave the decision as to what is “acceptable” to individual taste, which is in some ways an unsatisfactory state of affairs. Surely there should be a better way of reaching a compromise. I am going to propose just such a method – in fact I proposed it in 1970 (James, 1972). It was put forward independently by Dempfle (1973).

The proposal is based on the idea of discounting future gains to present value. This is a standard procedure in investment appraisal and cost/benefit analysis and was introduced to animal breeding by Poutous and Vissac (1962). The method was more widely used after it formed the basis of an appraisal of investment in animal breeding by Hill (1971). The principle of discounting is simple. If the rate of interest available is 100*r*%, one dollar can be invested at compound interest and in  $y$  years will be worth  $(1 + r)^y$  dollars so it follows that a dollar obtained  $y$  years in the future is worth  $1/(1 + r)^y$  dollars now. Therefore, returns achieved  $y$  years later can be discounted to their present value and compared with returns achieved at present. This provides a simple mechanism by which the conflict between short-term and long-term goals can be resolved.

If we have a well-defined breeding program we can predict the responses expected year by year or generation by generation. Let us suppose we have a monoecious population with discrete generations, and that the mean breeding value in generation  $t$ , measured from the current mean as origin, is  $M_t$ , while the economic value of a unit of breeding value is  $B$  dollars (or other currency unit). Then the gain in economic value in generation  $t$  is  $W_t = BM_t$ . However, in present value this

is worth  $W/(1 + R)^t$ , where  $R$  is the per-generation discount rate. Therefore the present value of the entire breeding program, from now into the indefinite future, is

$$W = \sum_{t=1}^{\infty} BM_t / (1 + R)^t .$$

The discount rate is chosen to be appropriate for an investment with a degree of risk (which is always present) but should not include any allowance for inflation. Inflation should usually affect returns and it is assumed that all returns are evaluated at present-day currency values. Of course there may be cases where the returns from the breeding program are affected differently from the rest of the economy, but this is probably not the normal case. Therefore I suggest that for a suitably chosen discount rate the worth of a breeding program be evaluated as

$$B \sum_{t=1}^{\infty} M_t / (1 + R)^t \text{ and this measure of worth is what should be maximised.}$$

Often such investment appraisals are made over a fixed time period, such as 20 years and Hill (1971) used a fixed time period in his influential paper. But in my opinion this is seldom justified. It may be so if a company is planning to build a toll road which will revert to public ownership after a set number of years, but presumably a genetically improved animal population will not be disposed of after a given time. Even if the breeder plans to sell the breeding operation and retire, the purchase price paid should reflect the value of the population as a going concern, which will depend on the prospects of future profits. If future profits are ignored, it is clearly best to select at maximum intensity in the last generation taken into account, since there is no need to allow for future improvement. In fact, Robertson (1970) showed that to maximise gain achieved by a given time a dynamic policy was required, with progressively more intense selection. The same would apply with discounted gains up to a specified time. In what follows I shall assume that the summation continues to infinity, though this is not necessary in principle, and the method can be used with any desired length of time. In fact, in many cases the distinction may be of relatively little importance, since gains made far into the future will be heavily discounted.

#### References for James:

- Dempster, E.R. (1955) Genetic models in relation to animal breeding problems. *Biometrics*, 11:535.  
Hill, W.G. (1971) Investment appraisal for national breeding programs. *Animal Prod.*, 13:37 – 50.  
James, J.W. (1972). Optimum selection intensity in breeding programmes. *Animal Prod.*, 14:1 – 9.  
James, J.W. (1987). Determination of optimal selection policies. *Journal of Animal Breeding and Genetics*, 104:23 – 27.  
Ollivier, L. (1974) Optimum replacement rates in animal breeding. *Animal Production*, 19:257 – 271.  
Poutous, M. and Vissac, B. (1962). *Annales de Zootechnie*, 11:233 – 256.  
Rendel, J.M. and Robertson, A. (1950) Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *Journal of Genetics*, 50:1 – 8.  
Robertson, A. (1960) A theory of limits in artificial selection. *Proc. Royal Soc., Series B*, 153:234.  
Robertson, A. (1961) Inbreeding in artificial selection programmes. *Genet. Research*, 2:189 – 194.  
Robertson, A. (1970) Some optimum problems in individual selection. *Theor. Popul. Biol.*, 1:120.  
Wray, N.R. and Thompson, R. (1990) Prediction of rates of inbreeding in selected populations. *Genetical Research*, 55:41 – 54.



## Balancing selection and inbreeding

We have seen that rates of inbreeding can increase substantially with small breeding population sizes. In fact, these rates are substantially higher when selection is taking place as well. The reasons for this are:

- Higher selection intensities are used to make bigger genetic gains - especially when we use novel reproductive techniques to boost fecundity. These policies lead to fewer parents and smaller effective population sizes
- Animals of high merit will tend to be related to each other, and therefore selected together. Mating of relatives is inbreeding.

Under mass selection, this problem is more important for traits with higher heritability - because genetic effects are more important in causing high merit and therefore greater relationship among selected animals. However, with selection on BLUP using information from relatives, low heritabilities are more damaging because they cause greater use of information from relatives, and therefore more similar EBVs for animals which are, for example, full sibs.

Inbreeding leads to loss of genetic variation. We expect that

$$V_{A_F} = V_A(1 - F)$$

where  $F$  is the inbreeding coefficient, such that there is a linear decrease in genetic variance as  $F$  increases. This is one factor which gives incentive to maintain genetic variation, or genetic diversity, within our breeding populations.

A simple approach to maintaining genetic diversity at higher levels is to increase effective population sizes. However, this means using more animals as parents, especially males, and this in turn leads to slower rates of genetic change.

## How should we balance these two competing issues?

A simple answer is to predict response to selection after making corrections for the effects of inbreeding on reducing genetic variation and reducing performance through inbreeding depression. However, this depends on many factors. It seems that many people tackle the problem by nominating a value of  $\Delta F$  to keep below.

We can do these things by avoiding coancestry of parents used, and to some extent by avoiding the mating of individuals that are related (eg. Kinghorn et al. 1999).

Here, we will focus on avoiding parental coancestry during selection, in an approach that assumes knowledge of pedigree.

The general problem is to choose parents from a pool of candidates. The predicted response in progeny can be simply predicted as  $x'G$  where  $x$  is a vector of parental contributions, summing to  $\frac{1}{2}$  for each sex of candidates, and  $G$  is a vector of EBVs or indices:

Genetic diversity is effectively the inverse of the average amount of relationship between animals which are chosen as parents, including self-relationships. This average amount of relationship is given by mean parental coancestry  $x'Ax/2$  where  $A$  is the numerator relationship matrix.

## Vector $x$ of animal contributions

Source of animals	Animal#	$x$ = Contribution
Male candidates	1	0
	2	.1
	3	.05
	4	0
	5	.01
	6	0
	7	0
	8	0
	...	...
Female candidates	101	0
	102	.01
	103	.01
	104	.01
	105	0
	106	0
	107	0
	108	.08
	...	...

Figure 17.1 An example of animal contributions.

Notice that average coancestry among parents, weighted for contributions, is equal to  $\Delta F \cong x'Ax/2$ . This is supported by:

- The inbreeding coefficient of an animal is half of the numerator relationship between its parents
- If it were possible to select just one non-inbred parent ( $x$  has one element equal to 1, and  $x'Ax = 1$ ), and to self-fertilise that parent to give all of the next generation, then inbreeding coefficient would be  $\frac{1}{2}$  ( $= x'Ax/2$ ).

Thus the objective is to maximise  $x'G - \lambda \frac{x'Ax}{2}$  where the size of  $\lambda$  reflects the importance of avoiding high relationship among parents, which leads to loss of genetic diversity.

Meuwissen (1997) has developed a nice method to do this. First, the desired maximum value of  $\Delta F \cong x'Ax/2$  is declared. He then calculates the value of  $\lambda$  that is required to make this constraint. Then he calculates the optimal elements of  $x$ . An alternative is to use a genetic algorithm to find values of  $x$  that either maximise  $x'G - \lambda x'Ax/2$ , or maximise  $x'G$  under a constraint that a declared value of  $x'Ax/2$  is not exceeded. This approach has great practical value, as it is easy to add other issues and constraints to the problem.

### Overlapping generations

The approach as described does not properly account for differences in animal ages - younger animals have more mating opportunities to contribute in later years and might thus be used less at the next mating. Figure 17.3 shows one step in the correct direction.  $L$  is the generation interval, so that an extra cohort's worth of contribution is added to the problem. These juvenile selections are dummy selections – only adults will be in the active mating group. However the juveniles are included in the problem solving step, such that, for example, an adult candidate that already has many juvenile progeny will be inhibited from getting a high contribution (many matings) in the current selection, as a function of the number of his progeny that are co-selected. See Meuwissen and Sonesson (1998) or Grundy et al. (2000) for a more comprehensive treatment of this issue.

## Vector $x$ of animal contributions

Source of animals	Animal#	$x$ = Contribution
Male candidates	1	0
	2	.1
	3	.05
	...	...
Female candidates	101	0
	102	.01
	103	.01
	...	...
Male juveniles	201	.02
	202	.01
	203	0
	...	...
Female juveniles	301	0
	302	.01
	303	0
	...	...

**Figure 17.3** Juvenile animals are included as “dummy candidates” that will not be included in the active breeding group. This provides some inhibition on the use of their parents and other relatives that are real candidates.  $L$  is the generation interval.

### Optimizing selection across age classes (age structure)

Van Arendonk and Bijma (2002) state that “the generation interval should not be regarded as a design parameter for the breeding schemes but needs to be a result selection across the available age classes while constraining”. To understand optimal selection across age classes, there are three points to remember about breeding values in a typical breeding program:

- The accuracy of breeding values increases with an increased age
- The variance of estimated breeding values increases with age.
- The average estimated breeding value (as well as true breeding value) increases over time and is therefore higher for younger animals

Each of these will be discussed in a bit more detail, before we discuss selection across age classes.

#### Accuracy of EBV

The accuracy is between 0 and 1 (or 0% and 100%). In the extreme case of no information, the accuracy of a breeding value is 0, and with a very large amount of information, the accuracy will approach 1. The following Table shows examples of accuracy. It illustrates that

- Accuracy is higher when more information is used, e.g. on relatives and progeny
- The accuracy of an EBV of animals usually increases with age
- The accuracy is higher for traits with a higher heritability, but the effect of heritability becomes smaller with more information used
- The accuracy of parent average depends on the parent EBV accuracy and not on heritability (but note that with low heritability it will be harder for a parent to achieve a certain accuracy)

Accuracies can be derived using selection index theory.

## Variance among EBV

The variance among EBVs is of practical value because

- it can give us an indication of the difference in EBV between the highest and lowest animals
- It is used to predict selection differential, e.g. the average EBV of the best 10% of animals
- An traits will be more impacted by selection on an multiple trait index if that index trait has more variation.

In general, the variation among EBV can be predicted from accuracy and genetic variance

$$\text{Var}(\text{EBV}) = r_{IA}^2 V_A \quad \text{and} \quad \sigma_{\text{EBV}} = r_{IA} \sigma_A$$

where  $r_{IA}$  is the accuracy of the EBV. Hence, the variance of the EBV's is equal to the accuracy-squared multiplied by the variance of the true breeding values (additive genetic variance).

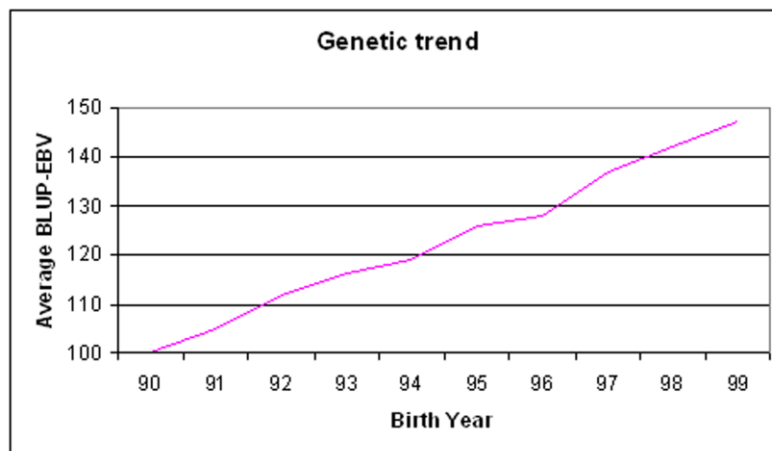
It is useful to consider the following

- |   |   |   |
|---|---|---|
| If $r_{IA} = 0$ then $\text{var}(\text{EBV}) = 0$ | : | all EBVs have the same value (=0)   |
| If $r_{IA} = 1$ then $\text{var}(\text{EBV}) = 1$ | : | the variance of EBV is equal to the variance of breeding values. All EBV should be equal to the true BV with this accuracy, and there is no prediction error. |

$\text{Var}(\text{EBV})$  is generally smaller than  $V_A$  ( $\sigma_A^2$ ).  $\text{Var}(\text{EBV})$  becomes larger when accuracy is higher. i.e. the EBV of older animals will be more apart than those of young animals. The same holds for intensely measured nucleus animals compared to the EBV of base animals that have less information and therefore EBVs closer to each other.

## Genetic Trend

Selection of parents over time leads to genetic improvement, i.e. there is an increase of average breeding value over time. This increase is referred to as 'genetic trend', and basically measures the success of breeding programs. The average of each generation is estimated by BLUP as the mean of the parents' EBV, and since BLUP is also able to work out the difference between all animals and those selected as parents, BLUP can properly estimate the genetic change over time. The genetic trend is estimated from the average EBV's over time, i.e. the EBV's are plotted against the birth year of the animals.



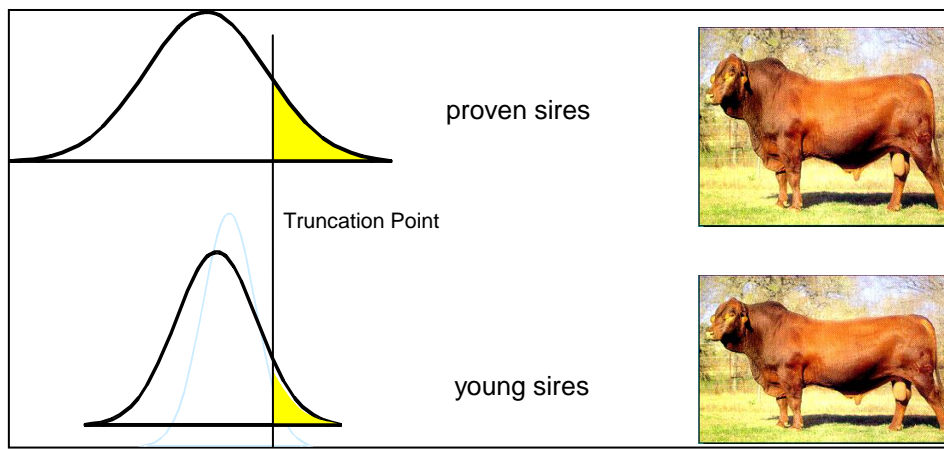
*Figure Hypothetical example of a genetic trend plotted as average EBV against birth-year.*

*The average of 1990 is set to 100.*

## Selection of young bulls vs old bulls

When optimising a breeding program, a dilemma often arises on whether young animals should be selected or older animals. Selecting young animals is good for achieving a short generation interval. However, younger animals have usually less accurate EBV. Older animals have generally more accurate EBV but selecting them would lead to longer generation intervals. Another (but essentially similar) argument against selecting older animals is that they are expected to have lower EBV. If there is a genetic gain per year, animals born  $x$  years apart are expected to differ  $x$  times the annual genetic gain.

**Figure** Truncation points for older proven sires and younger unproven sires.



It seems not easy to optimise selection over different age classes. However, the solution appears to be remarkably simple in practice, for tactical decision making. Because BLUP accounts for genetic trend, an important practical consequence is that animals can be selected across age classes based on their EBV by truncation selection. For example, all bulls with an index value above 140 will be selected, independent of their age. It has been proven that BLUP selection optimises the use of bulls across age classes (James, 1987). Hence, selection on BLUP EBV, irrespective of age, automatically optimises the generation interval, and the use of old versus young bulls, as demonstrated in the Fig. above.

Younger animals have on average better EBV, but also generally less variation in EBV. The optimum proportion of younger animals depends on the difference in the variance of the EBV within age classes (i.e. on the accuracy) and on the genetic lag between age classes (i.e. on the genetic gain per year and the number of years).

Truncation selection across age classes will automatically optimise the proportion of young bulls used. The proportion that should be optimally selected from each age class is automatically achieved if simply the best animals are selected based on their BLUP EBV. Selecting animals on BLUP EBV irrespective of their age automatically optimises the generation interval.

With a larger genetic trend or with increased accuracy of the young bull EBV, the proportion of young bulls will increase (check yourself in the figure).

## Optimizing ages structure in deterministic simulation

In GENUP with the AGES module we can compare age structures for simple breeding programs with phenotypic selection on a single trait. In ZPLAN we can try various age structures and

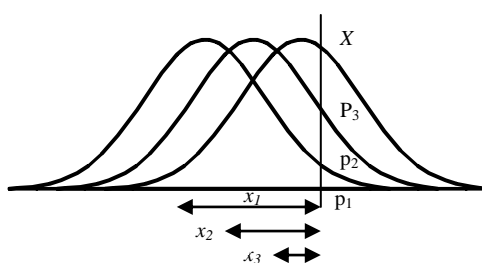
compare the outcomes for more realistic multiple trait examples. Simple exercises such as in AGES show that earlier selection leads to higher replacement rates and therefore affects selection intensity. With BLUP selection, there is an additional argument that early election is usually on EBVs that are less accurate. Moreover, selecting in an early stage of life usually means that BLUP selection is more based on family information, and this would lead to more inbreeding. This will be discussed in the next chapter. Therefore, generation intervals should not be minimized, but rather should they be optimized.

In deterministic prediction methods such as AGES and ZPLAN, this can be based on trial and error. However, there is also a method that can optimize selection across age classes, and therefore optimize generation interval. The method is illustrated in the Figure below. It shows overlapping distributions, representing values of animals from different age classes. These values could be phenotypes, EBVs or index values, and assuming selection is for such a single criterion. The distributions most to the left could represent the oldest animals, as they will have the lowest mean if there is a positive genetic trend.

The objective is to find a common truncation point  $x$  where selection of all animals across all available distributions leads to a total proportion selected  $P$ . There will be one truncation point across all distributions. This truncation point has a single value, but relative to the mean of each distribution, the truncation points will differ:  $x_1 > x_2 > x_3$ . Therefore, selected proportion will be accordingly smaller:  $s_1 < s_2 < s_3$ , and consequently, the selected proportions will be  $p_1 < p_2 < p_3$ , and selection intensities will be  $i_1 > i_2 > i_3$ .

There is no algebraic closed solution to finding the optimal truncation point but the solution can be found iteratively. If the truncation point is too far to the right, the overall selected proportion will be too small, and if too far to the left, too few will be selected. Therefore, an algorithm can work as follows:

1. Choose starting values for to the left and far (e.g. six standard deviations) to the right,  $x_L$  and  $x_U$  for lower and upper threshold.
2. Choose a truncation point in the middle,  $x_m = (x_L + x_U)/2$
3. For a given  $x$ , determine the total proportion selected.
4. This can be done by working out for each distribution:
  - i.  $x_i = x - (\mu_i - \mu_1)$ .
  - ii. From  $x_i$ , determine  $p_i$
  - iii. Total proportion is  $P_m = p_1s_1 + p_2s_2 + p_3s_3$
5. If  $P_m - P$  is very small, stop the iteration
6. If  $P_m < P$  then  $x_U = x_m$
7. If  $P_m > P$  then  $x_L = x_m$
8. Return to step 2



$$P = p_1s_1 + p_2s_2 + p_3s_3$$

$\mu_1$	$\mu_2$	$\mu_3$	Mean of distribution
$\sigma_1$	$\sigma_2$	$\sigma_3$	SD of distribution
$s_1$	$s_2$	$s_3$	Nr. of animals in this distribution as proportion of total nr of animals across all distributions distribution

Example:

ageclass	N in group	mean	SD	Proportion Selected	Nr Selected	Mean of selected	
1	50	10	1	0.28	14.17	11.18	
2	35	9.5	1	0.14	4.96	11.03	
3	15	9	1	0.06	0.87	10.74	
					20.0	11.12	mean of selected
					1.33		Age of selected (GenInt)

## REFERENCES

- Bijma, P., van Arendonk, J.A.M. and Woolliams, J.A. (2001) *J. Anim Sci.* **79** : 840-853.
- Bijma, P. and Woolliams, J.A. (2000) *Genetics* **156** : 361-373.
- Bulmer, M.G. (1971) *Am. Naturalist* **105** : 201-211.
- Ducrocq, V. and Quaas, R.L. (1988) *J. Dairy Sci.* **71** : 2543-2553.
- James, J.W. (1987) *J. Anim. Breed. Genet.* **104** : 23-27.
- Kinghorn, B.P. and Dekkers, J.C.M. (1991) *J. Dairy Sci.* **74** : 611-622.
- Land, R.B. and Hill, W.G. (1975) *Anim. Prod.* **21** : 1-12.
- Meuwissen, T.H.E. (1991) *Biometrics* **47** : 195-203.
- Meuwissen, T.H.E. (1997) *J. Anim Sci.* **75** : 934-940.
- Meuwissen, T.H.E. and Sonesson, A.K. (1998). *J. Anim. Sci.* **76** : 2575-2583.
- Nicholas, F.W. (1996) *Anim. Reprod. Sci.* **42** : 205-214.
- Nicholas, F.W. and Smith, C. (1983) *Anim. Prod.* **36** : 341-353.
- Quinton, M. and Smith, C. (1995). *J. Anim. Sci.* **73** : 2208-2212.
- Rendel, J.M. and Robertson, A. (1950). *J. Genet.* **50** : 1-8.
- Verrier, E., Colleau, J.J. and Foulley, J.L. (1993) *Theor. Appl. Genet.* **87** : 446-454.
- Villanueva, B., Wray, N.R. and Thompson, R. (1993) *Anim. Prod.* **57** : 1-13.
- Villanueva, B. and Simm, G. (1994) *Proc. WCGALP 5.* **20** : 200-207.
- Visscher, P., Pong-Wong, R., Whittemore, C. and Haley, C. (2000). *Livest. Prod. Sci.* **65** : 57-70.
- Van Arendonk, J.A.M. and P. Bijma. World Congr. Genet. Appl. Livest. 2002. PC. No.27.02
- Woolliams, J.A. (1989) *Anim. Prod.* **49** : 1-14.
- Woolliams, J.W. and Wilmut, I. (1989) *Anim. Prod.* **48** : 245-256.
- Woolliams, J.W., Bijma, P. and Villanueva, B. (1999) *Genetics* **153** : 1009-1020.
- Woolliams, J.W. and P. Bijma, P. (2000) *Genetics* **154** : 1851-1864.
- Wray, N.R. and Hill, W.G. (1989). *Anim. Prod.* **49** : 217-227.
- Wray, N.R., and Thompson, R. (1990) *Genet. Res. Camb.* **55** : 41-54.
- Wray, N.R., and M.E. Goddard. (1994).. *Genet. Sel. Evol.* 26:431.

## Designs to Exploit Reproductive Technology

### Key terms and concepts

Multiple Ovulation and Embryo Transfer (MOET), Cloning, Sexing, Juvenile In Vitro Embryo Transfer (JIVET).

### Introduction to the topic

Reproductive technologies such as AI and MOET can be used to greatly increase fecundity. This increases selection intensity and, in turn, genetic gains.

The best designs for breeding programs often change when these technologies are used. As someone once said, turning a cow into a sow means we should shift from cattle to pig designs.

Cloning is an extreme form of reproductive boosting. It has different consequences for breeding and production programs because it does not involve production of genetic variation through meiosis.

### Multiple ovulation and embryo transfer

MOET stands for Multiple Ovulation and Embryo Transfer. Females are super-ovulated by hormone injection, mated, and then multiple embryos are collected and transferred to host females. These host females play no part in the genetics of the breeding program.

MOET improves the rate of genetic improvement because of the favourable effects of high reproductive rate on three key factors:

1. Increased selection intensity
2. Reduced generation interval
3. Increased accuracy of estimating breeding values.

### Increased selection differential and reduced generation interval

The impacts of reproductive rate will be shown by example. Run the GENUP module AGES then return to this document for instructions as required.

If the default GENUP data set has been kept as original, you should find a response of 0.077Kg fleece weight per year. Optimise age structure and this should increase to 0.082Kg per year, keeping males and females for 2 and 4 matings respectively. For help on running AGES, Hit the F1 key.

Without MOET the weaning rate is taken as 0.95 lambs weaned per ewe mated. Increase this four-fold, to 3.8 lambs weaned per donor ewe mated, to simulate MOET, and optimise age structure. Response should increase from 0.082 to 0.1412. Notice that this affects the values of selection intensity,  $i$ , and generation interval,  $L$ , considerably, and the optimum age structure is now considerably younger - keeping males and females for 1 and 2 matings respectively.

Why is the optimum generation interval lowered? With low fecundity, we must keep females for many matings in order to produce a good number of candidates for selection. With high fecundity we need fewer matings – more matings gives more candidates, but with diminished benefits in terms of selection intensity. You should get a feeling for this by playing with AGES.



Notice also that with MOET we have both higher selection intensity and lower generation interval for both sexes, not just females: Higher reproduction means fewer females to get the same number of progeny - and these fewer females can be mated by fewer males!

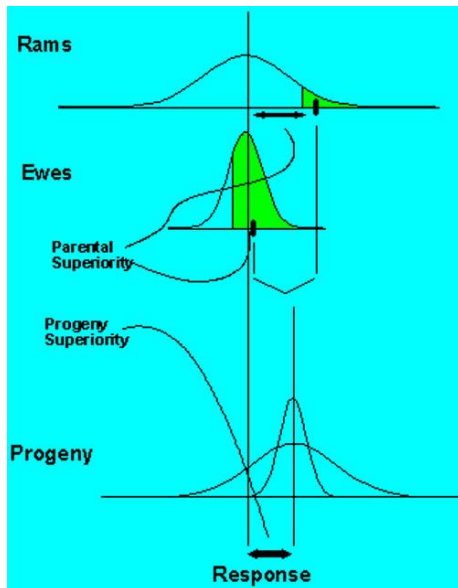
Notice also that with MOET, the breeder needs to maintain a large pool of recipients making this a much more expensive program to run.

## Increased accuracy of estimating breeding values

Recall that increased accuracy of estimating breeding values leads to wider distributions for EBV (Figure 15.6). This can be used to illustrate how MOET increases response through increased selection accuracy and selection intensity:

EBV's calculated properly from a selection index calculation or a BLUP analysis have one very useful property: the predicted merit of progeny is simply the average of the EBV's of the two parents used, as shown in Figure 16.1.

## Natural matings



*Figure 16.1 Distributions of perceived genetic merit (or estimated breeding value, EBV) in male and female parents and their progeny. Selected parents are shaded. We allocate many females to each male, so that we can select fewer, more elite males. We can capitalize on this by gathering more information on males, giving more power to identify the best and worst males, as reflected in their wider EBV distribution.*

A smaller proportion of rams than ewes can be selected for breeding, contributing to their high mean EBV. The other factor, in this case, is the greater amount of information used to calculate ram EBV's, reflected by a wider distribution of EBV's. If using MOET, the value of taking more measurements in ewes (as well as rams) becomes higher.

Notice that the predicted merit of progeny is simply the average (or half-way-point) between the selected rams' and ewes' mean EBV's. The width of the EBV distribution of the progeny depends on how intensively they are measured.

### AI: Increased selection intensity.

With use of AI, the few best rams can be selected for breeding. This means that the average EBV of rams used is higher. This can be seen in Figure 16.2: only the very best rams contribute to the average EBV of progeny. This increase is diluted 50 percent by the ewe contribution, but the net increase in predicted merit of progeny is quite visible compared to Figure 16.1 with lower selection intensity of rams due to the use of natural mating.

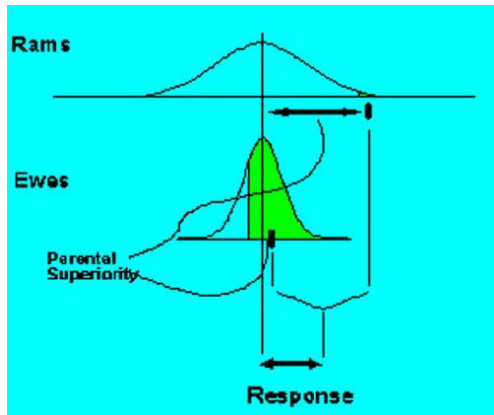


Figure 16.2 Increased response due to AI in rams.

### MOET: Increased selection intensity and more information for estimating breeding value

With MOET, each selected ewe can contribute not just one or two lambs, but up to about 6 lambs per donor ewe, a figure which continues to improve. As with AI, this brings about the ability to select fewer ewes as *donors of genetic material* (many recipient ewes are still required to carry the lambs). Moreover, in an ongoing breeding program using MOET, candidates for selection will usually have a number of full-brothers and sisters available with records. This information helps to improve the accuracy of EBV's, and the distribution of EBV's increases accordingly. Both these favourable effects are seen in Figure 16.3. The favourable effects of reduced generation interval cannot be shown in this simple diagram.

Even so, comparing Figure 16.3 to the original natural mating diagram (Figure 16.1), it can be seen that AI and MOET, if properly used, can bring about a notable increase in the response to selection. Current theoretical predictions suggest that a MOET program will give about 50% extra gain over a normal breeding program. However, increased rates of inbreeding can reduce the value of this substantially.

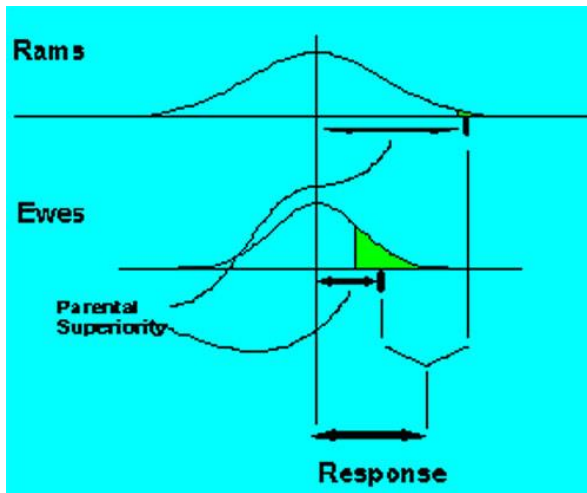


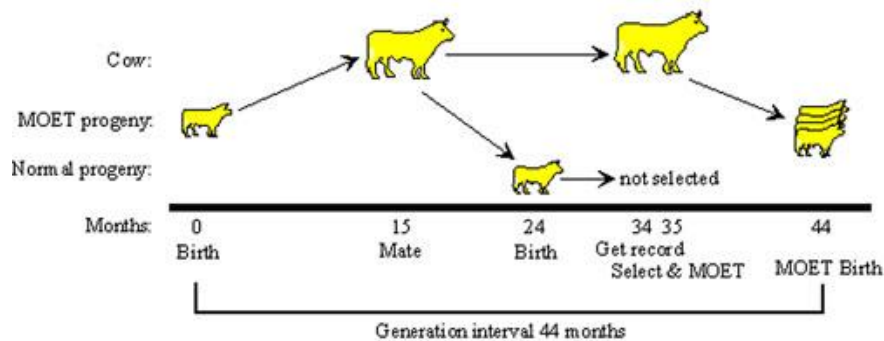
Figure 16.3 Increased response due to MOET in ewes with AI in rams.

## Juvenile MOET schemes vs. adult MOET schemes

### Dairy schemes

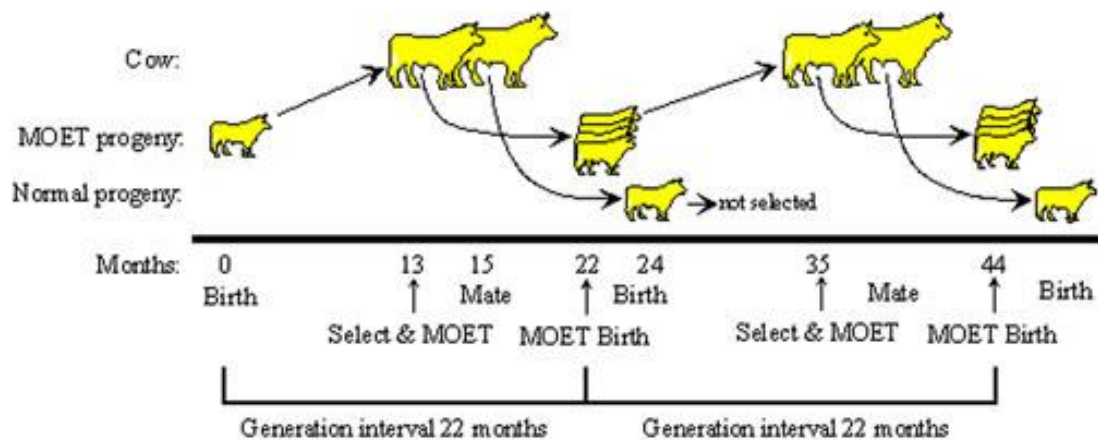
Some traits cannot be measured until after the onset of sexual maturity - one example is milk yield in dairy cattle. This means that we would have to wait until a cow has been bred and produced milk before she could be evaluated on her own performance for selection into an elite breeding herd.

**Figure 16.4 Adult dairy MOET scheme.**



One way around this is to let young females compete with each other on their EBVs calculated using information from older relatives (mother, aunts, etc). This gives a reduced generation interval (favourable) at the expense of reduced selection accuracy (unfavourable).

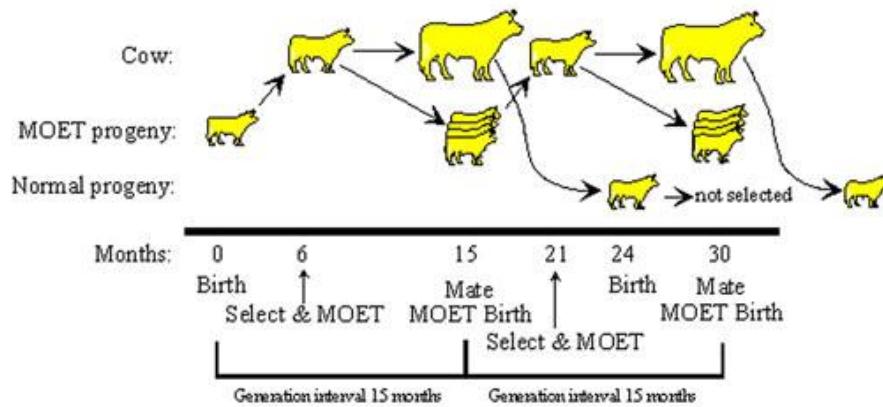
**Figure 16.5 Juvenile dairy MOET scheme.**



When using MOET in dairy cattle the better option in these cases is to select among juveniles - hence "Juvenile MOET schemes" as opposed to "Adult MOET schemes" which wait for these measurements to be made before selecting. The selection among these juveniles is based on an index of information from older relatives. This index turns out to be quite simply the average of the EBVs of each candidate's parents.

With collection of embryos from pre-pubertal animals, the generation interval can be reduced even further.

Figure 16.6 Very juvenile dairy MOET scheme.



But notice that females are now being selected and MOET carried out (eg. at 21 months on the scale shown) before a lactation record is available on their dams (at about 27 months). This means that selection is based on grand-parental EBV's, which is normally unacceptable. However, with good information from genetic markers there could be value in such a scheme.

## Beef and sheep schemes

These species differ with respect to dairy MOET schemes as key traits can usually be recorded before or close to sexual maturity, rather than well after this stage as in dairy cattle. In this case a 'juvenile' scheme implies collecting oocytes before sexual maturity, as in the following schemes:

Figure 16.7 Juvenile sheep MOET.

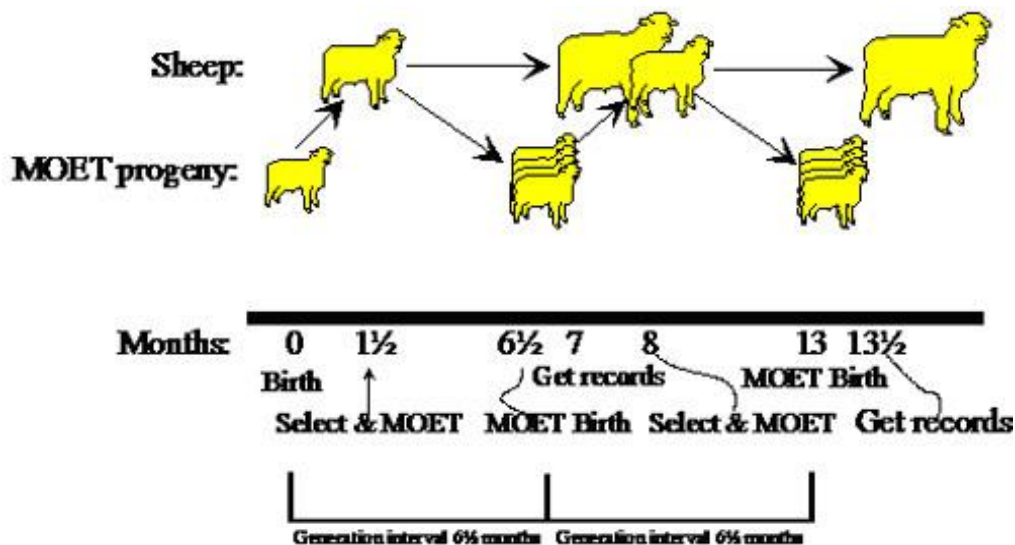
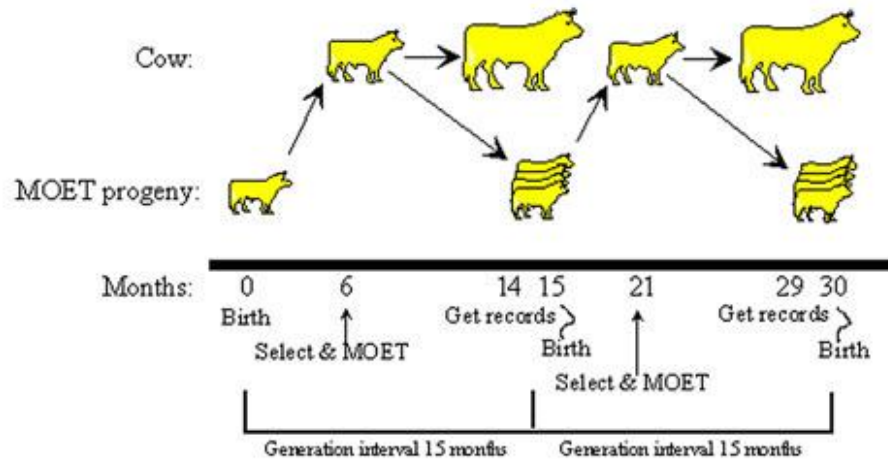


Figure 16.8 Juvenile beef MOET.



In this case oocyte pickup is carried out on 6 month old females (eg at 21 months) based on their parents' records (at 14 months).

### Juvenile in-vitro embryo transfer, JIVET

Novel mating systems become possible with in vitro fertilization (IVF). This happens naturally in some animals, notably fish. We have the opportunity to mate each female with many males, as well as vice versa. One benefit is a richer pedigree structure, with maternal half-sib families as well as paternal half-sib families, increasing the accuracy of genetic evaluation. With use of surrogate mothers in mammals, we can also help avoid the confounding effects of maternal environment. These higher accuracies help to reduce the compromises in using juvenile mating systems.

JIVET stands for Juvenile In Vitro Embryo Transfer. This involves collection of ova from sexually immature females, followed by in vitro fertilisation. JIVET has already been implied in the last section, wherever fertilisation takes place *in vitro*.

Techniques for manipulating waves of follicular activity have resulted in recovery of high numbers of oocytes from females. This has also been successful at a research level with juvenile animals such as six-week-old lambs and eight month old calves, whereby the reduction in generation interval is expected to lead to notable increases in rates of genetic gain. However, it should be noted that there are considerable animal welfare issues associated with these procedures on juvenile animals.

Over the horizon:

The key promise for juvenile mating systems is the advent of marker-assisted selection (MAS). If we can get reasonable genetic evaluations from DNA tests, then we can contemplate very juvenile systems, even to the extent of developing cycles of in vitro sexual propagation, including meiosis and zygote formation, with selection among large numbers of zygote cell lines using genetic markers. Selection among gametes or haploid cell lines using marker information would be even more powerful, effectively accessing twice as much additive genetic variation due to covariances generated between the gametes that contribute to zygotes. This approach would also help exploit non-additive effects, and the generation of targeted genotypes across many loci.

Such developments are currently well out of sight. However, we may need mating

systems to target the generation of specific genotypes across many loci in normal MAS schemes. This becomes more evident as we discover that interaction between genes is important. These mating systems will need to handle non-additive effects; thus they relate to mating systems that exploit heterosis.

If we succeed in having a high degree of control over fecundity, we are left with the decision of how best to use it. Reproductive boosting gives higher genetic gain, but at a penalty of increased inbreeding and lost genetic diversity. At the extreme, it also constitutes high risk, by 'putting all the eggs in one basket'. These problems can be handled in an appropriate mate selection scheme.

## **Sire Purchase Policies**

J.W.James

Reprogen, Faculty of Veterinary Science, University of Sydney.

### **THE BREEDING PYRAMID**

Although there are some vertically integrated animal production systems where the breeding program is controlled by the commercial producer, or where commercial producers work under contract to suppliers of breeding stock, for the most part genetic improvement of livestock is carried out in studs or breeding nucleus herds or flocks. Genetic improvement is then disseminated to commercial stock by the transfer of animals (mainly sires) or gametes.

This pattern was described by Robertson and Asker (1951) in British Friesian cattle using pedigree records from herdbooks. They showed that the structure was pyramidal, with a small group of elite studs at the apex of the hierarchy with one or two layers of “multiplier” studs between the elite nucleus and commercial herds. The elite studs were identified as those who exchanged animals among themselves, transferred animals to other herds, but did not introduce stock from these other herds. Other layers can be defined in a similar way. This work has been confirmed in many other studies. With changes in animal breeding over the last 50 years there have been changes, but it is still true that genetic improvement is made in a special nucleus and transferred to commercial herds and flocks through animals or gametes.

The result of this hierarchical structure is that, after an initial settling in period, the whole population improves at the same rate, which is set by the nucleus. Each tier in the pyramid lags behind the one above it. In a very simple case, where all sires in any level are introduced from the tier above, there is no selection of females and the sires bought are of average breeding value in the group in which they were born, each tier lags behind the one above it by two generations of genetic improvement. More complex situations require more complex expressions for lag, and have been considered by Bichard (1971) and James (1977) among others. It should be noted that any selection in the lower tiers does not affect the steady-state rate of genetic gain in the population, unless there is upward transfer of genes in the hierarchy (an open nucleus). But it does affect the lag, as does selection of better-than-average sires from the upper tiers. It will also affect lag if producers buy some breeding stock from two tiers above rather than one.

In principle it is possible for the lag to be negative. This can happen if (say) the nucleus is making no genetic progress, but selection of superior females in the commercial population raises its mean breeding value above that of the nucleus.

Another point to note is that in a population with a pyramidal structure the rate of inbreeding is controlled entirely by what happens in the elite nucleus, and since this is typically a small fraction of the whole population, the effective size of the population will be much smaller than might be expected from the population size. The dominance by the nucleus of the population inbreeding is for the same reason as for its effect on genetic improvement. Genes flow from the nucleus to the rest of the population but not vice versa, so that eventually the whole population consists of genes from the nucleus, the genes at different levels originating in the nucleus at various times in the past.

## VALUING GENETIC SUPERIORITY

One might expect that a nucleus breeder would like to limit expenses in running a genetic improvement program to produce animals for sale. If stock can be sold for satisfactory prices, avoiding spending on an effective breeding program may be a good option. If the nucleus begins at a higher genetic level than a commercial flock the commercial breeder can achieve genetic improvement by grading up to the nucleus, but the difference is halved each generation and would soon disappear without genetic gain in the nucleus. For the nucleus to remain significantly superior to commercial stock it must run an effective improvement program. This raises production costs in the nucleus and will be justified only if a suitable premium can be charged for breeding stock. A commercial producer, on the other hand, will want to reduce sire purchase costs, which can be done by buying sires less often or by buying cheaper ones. Thus a sire buyer will want to have a strategy covering frequency of purchase and cost (presumably related to breeding value) of sires.

An approach to devising such a strategy can be based on discounted gene expressions. This concept was introduced by McClintock and Cunningham (1974) to provide a rational method of giving economic weights to meat and milk traits in dual-purpose cattle, and was based on the idea of a farmer who was considering buying a dose of semen to inseminate a cow. Sires whose semen is available for purchase have EBVs for meat and milk traits. If the calf born is male it will be slaughtered for meat. If it is female it will be kept for milking over a number of lactations. Thus the meat traits in a male are expressed once, while milk traits will be expressed several times. When the cow is culled its sale value may be influenced by its meat traits. So the milk traits will contribute more often than meat traits, but returns from females will occur later than from males, and so will be discounted more heavily. However, there is a further complication. Female calves may be used for breeding and pass on half their breeding value for both meat and milk to their progeny, which will then express them after another delay, so even more heavily discounted. So in order to give appropriate weights to meat and milk traits it is necessary to take account not only of the number of expressions of the traits but the pattern of expression. Thus to choose a dose of semen from different sires we must weight their EBVs for meat and milk in a suitable way. We also need to compare expected returns with costs.

The approach of McClintock and Cunningham was used by Napier and Jones (1976) to evaluate purchase of a ram rather than a dose of semen. Essentially all we need to do is consider that when the ram is bought he can generate many progeny rather than the single one from a dose of semen. If he is to be used for only one mating season, his value will be that of  $M$  doses of semen, where  $M$  is the number of ewes to which he is mated. If he is to be used for several seasons, then successive seasons can be added, each being discounted to present value while making allowance for possible death or disability. In this way the value of a purchased ram can be calculated and compared with the purchase price.

One problem with this method of evaluation is that each additional season of use adds more value to the sire, and an uncritical use of this criterion could lead a producer to use a sire for as long as possible. This would be the correct procedure if better sires are not available at a competitive price. But what should the producer do if better sires are continually becoming available? The buyer needs a policy which can give guidance on when to replace a sire he already owns with a new one. Of course if the sire can no longer produce the required progeny he has to go. But assuming he is still fertile, should he be retained or replaced?

## REPLACEMENT POLICY

The basic principle is the same for terminal sires and sires which will breed replacements, but some details are different, and another approach is possible for replacement-breeding flocks. Let us therefore begin with the simpler case of buying a terminal sire. The females may be bred on-site as part of a breeding program or purchased from another breeder. The question of how to



improve the females is a separate one, and here we concentrate on the purchase of terminal sires.

A method of assessing the value of terminal sires was given by Ollivier and James (1986) and James (1994). Here we follow the presentation of James (1994). Let  $P$  be the number of progeny produced by a sire in a given period, and let  $W$  be the probability that the sire survives to the next mating period,  $P$  and  $W$  being assumed constant (a convenient but not necessary assumption). The maximum number of mating periods for which a sire will be used is  $T$ , possibly determined by declining fertility. If  $d$  denotes the discount rate per period, the discount factor is defined as  $r = 1/(1 + d)$ , and we denote by  $Y$  the number of periods between sire purchase and first returns. Let  $B$  be the monetary value of a unit of overall breeding value. Then the present value to the producer of a unit of breeding value in the sire is

$$v = 0.5r^Y BP(1 - r^T W^T)/(1 - rW).$$

Suppose the producer has the choice among several potential sires, the  $j$ th having breeding value  $A_j$  and cost  $C_j$ . Then the profitability of the  $j$ th sire is  $vA_j - C_j$ , and this provides a criterion for choosing the best sire to buy from among those available, on the assumption that a purchase is to be made.

Next we need to consider the profitability of a sire already owned. This sire has breeding value  $A_k$  but his purchase cost is now zero. However, if he was bought  $t$  time periods ago he will be used for only another  $T - t$  time periods. Thus his profitability will be

$$0.5A_k B Pr^Y (1 - r^{T-t} W^{T-t})/(1 - rW).$$

Thus the difference in profitability between the new and old sires is

$$\frac{0.5r^Y BP}{1 - rW} [(A_j - A_k) - (rW)^T \{A_j - \frac{A_k}{(rW)^t}\}] - C_j$$

and sire  $k$  should be replaced by sire  $j$  if this is positive. In practice we would not use this criterion, but simply evaluate all potential sires and choose the most profitable. But this expression allows us to gain some insights.

Let us suppose sire  $j$  has the same breeding value relative to his nucleus contemporaries as sire  $k$  had when he was bought, and that the breeding nucleus has been making genetic progress at the rate of  $G$  per time period. Then  $A_j = A_k + tG$  and the criterion for replacement becomes

$$\frac{0.5r^Y BP}{1 - rW} [tG\{1 - (rW)^T\} + A_k \{(rW)^{T-t} - (rW)^T\}] - C_j.$$

This has an especially simple form when sires will not be replaced at any fixed age but will be kept indefinitely until a more profitable sire is found. Then  $T \rightarrow \infty$  and the criterion becomes

$$\frac{0.5r^Y BP}{1 - rW} tG - C_j.$$

This shows clearly what is obvious anyway, namely that, when  $G > 0$ , as  $t$  increases the more likely it is that a sire should be replaced, and that the time before replacement should be shorter when more rapid genetic progress is being made. The more expensive potential sires are, the less frequently they should be bought, other things being equal.

## REPLACEMENT-BREEDING FLOCKS

We saw earlier that the value of a terminal sire could be evaluated simply because only its immediate progeny produce expressions of his breeding value. With a sire of self-replacing females, his grandchildren, great-grandchildren and so on also contribute expressions of his breeding value, albeit halved each generation and progressively more discounted. It was shown by James (1980) that the effect of these extra expressions could be accounted for by the factor  $J$ , which depends on the breeding age structure and fertility of females and the discount factor. If  $f_j$  is the genetic contribution of females aged  $j$  to each progeny crop, so that  $\sum_j f_j = 0.5$ , then

$$J = 1 / (1 - \sum_j f_j r^j).$$

If a sire for a female replacement-breeding flock has a breeding value  $A$  and a cost  $C$  then his profitability can be found by multiplying the value of  $v$  for a terminal sire by  $J$  and is

$$0.5r^Y JBPA[1 - (rW)^T] / [1 - rW] - C.$$

Thus the same type of analysis may be made as for terminal sires except for the inclusion of the factor  $J$ .

As an example of the calculation of  $J$ , let us consider the following example. Suppose the chosen discount rate is 5%, and that the reproductive contributions are:

Age	1	2	3	4	5
Contribution	0	0.15	0.15	0.12	0.08

Then we have

$$1 - 1/J = \frac{0}{1.05} + \frac{0.15}{1.05^2} + \frac{0.15}{1.05^3} + \frac{0.12}{1.05^4} + \frac{0.08}{1.05^5} = 0.4270$$

so that  $J = 1.7453$ . If  $r = 1$  and there is no discounting we have  $J = 2$  so that discounting has reduced the effective contributions by about 13%.

## **Optimising expenditure on trait measurement**

J.W.James

Reprogen, Faculty of Veterinary Science, University of Sydney

In the past thirty years there has been a great development of the use of computers in genetic evaluation, due in part to the rapidly increasing power of these machines and in part on the development of methods such as BLUP and REML which rely on computing power for their practical implementation. We know very well now how to include all available information in our genetic evaluations, and how to make selection decisions based on these evaluations. This is all on a very sound theoretical basis. Of course, we often do not have the information we would like, but we can use what we have. We are still faced with the question of what information we should gather. How can we decide on such matters?

### **PROGENY TESTING**

It was nearly 50 years ago that Robertson (1957) sought to rationalise progeny testing programs in dairy cattle. This is now so well established that it is hard to realise how original it was at the time. Essentially he asked the question: How can a limited amount of test material, the cows to be mated to young bulls for testing, best be used? He pointed out that the greater the number of young bulls tested, the greater would be the selection intensity, but the smaller would be the accuracy of selection, and that the product of the standardised selection differential and the correlation between true and estimated breeding value was the quantity to be optimised. He found a simple approximation to the optimum family size. This well-known theory is not usually thought of as involving optimising the distribution of expenditure on measurement over candidates for selection, but it is an example. Some young bulls have progeny test measurements, while others do not. In the theory it is assumed that the breeder has no useful information to choose among the young bulls, so all selection is based on the progeny test results. This would have been a reasonable assumption in the early days of dairy cattle progeny testing, even if less so now. Others soon extended Robertson's idea to other situations, e.g. Rendel (1958) and Smith (1959).

### **INCOMPLETE TESTING**

In discussing the design of pig testing schemes, Smith (1959) pointed out that if the number of animals tested was no greater than that required for use, progress was possible only if part of the required breeding stock were selected from the tested animals, with the remainder being made up of randomly chosen untested animals. He remarked that it was always preferable to choose at random from among untested stock than to choose a tested animal known to be below average, since the EBV of an animal with no records is zero, and hence higher than that of a below-average tested animal. (This assumes that the mean EBVs of tested and untested animals are equal). James (1966) applied this idea to selecting a fixed number of animals from a population on the basis of a measurement that is relatively expensive compared to the expected gain from selection. He found the optimal fractions to measure for given proportions culled and ratios of benefit to cost. For instance, if the cost of a measurement exceeded 0.3989 of the value of a standard deviation superiority in an animal it was best to do no measurement. Although a formal solution was given, in practice it would be straightforward to program a computation to find the most profitable measurement fraction. It was also pointed out that if comparison of a cheap and an expensive method was made assuming complete testing there might be a bias against the expensive method, because complete testing is less likely to be optimal for it than it is for the cheap one.

## PREVIOUS INFORMATION

The analysis by Robertson (1957) of optimal family size in progeny testing was based on the assumption that nothing was known about the breeding values of the young bulls, or at least that there was no difference among their EBVs prior to the progeny test. If there is previous information, this ought to be combined with the progeny test data in the final evaluation. Another limitation placed on the progeny testing is that each tested sire has the same number of progeny. Thus the distribution of progeny among the available bulls is such that only two family sizes are possible: zero or  $n$ . But the general situation is that each available bull could be given an arbitrary number of progeny, subject to the restriction that the total number of progeny matches the testing facilities. James (1979) considered the optimum distribution of varying numbers of progeny to available sires subject to selection of a fixed number and that the total number of progeny is fixed. We present his approach here.

Suppose the prior information is summarised in the form of an EBV,  $\hat{G}_1$ , scaled so that the regression of true breeding value on this EBV is unity, and that the correlation between  $G$ , the true breeding value, and  $\hat{G}_1$  is  $r_1$ . A second set of data is then obtained such that if the prior information is ignored breeding value can be estimated as  $\hat{G}_2$  with a correlation  $r_2$  with  $G$ . Let  $r$  be the correlation between these two EBVs. Then the best combination of these two EBVs is  $\hat{G}$ , where

$$\hat{G} = \frac{1 - rr_2 / r_1}{1 - r^2} \hat{G}_1 + \frac{1 - rr_1 / r_2}{1 - r^2} \hat{G}_2.$$

Then each sire is evaluated on its value of  $\hat{G}$  and the required number is selected by truncation (Henderson, 1963). This truncation point  $X$  can be taken as a deviation from the prior EBV as  $(X - \hat{G}_1) / \sigma_{\hat{G} - \hat{G}_1}$  which we denote as  $x$ . The corresponding probability of selection is  $p$  and  $z$  is the ordinate of a normal curve at  $x$ . Thus the expected gain following selection on  $\hat{G}$  is

$$\frac{\sum p(\hat{G}_1 + i\sigma)}{\sum p} = \frac{\sum p\hat{G}_1 + \sum z\sigma}{\sum p}$$

where we have dropped the subscript on  $\sigma_{\hat{G} - \hat{G}_1}$ , and summation is over all prospective sires with  $\sum p$  chosen to give the required number of selected sires. The problem is then to choose the appropriate value of  $\sigma_{\hat{G} - \hat{G}_1}$  conditional on the prior EBV. James gave a formal solution, but it was so complex as to be of little practical use.

However, a few general points can be brought out. The equations show that the response to selection can be regarded as having two components: selection among groups with the same prior EBV, and selection within these groups on the difference between final and initial EBV. Sires with small numbers of progeny have small  $\sigma_{\hat{G} - \hat{G}_1}$  values, and in the limit with no progeny will have

a final EBV equal to the initial one. Animals without progeny will be selected on  $\hat{G}_1$  so that if  $\hat{G}_1 > X$ , an animal which is not progeny tested is certain to be selected, while if  $\hat{G}_1 < X$  it is certain to be culled. Independent culling levels can be regarded as corresponding to the second case. There is then reason to suggest that animals with very high initial EBVs need not be progeny tested, nor should those with very low values, attention being concentrated on testing those for which further information will be critical for making selection decisions. Of course, such a situation can arise only if the accuracy of the prior EBV is rather high, which would probably occur

only rarely when progeny testing is being considered. Although the theory was developed with progeny testing in mind, it applies to any further gathering of information.

It would be usual to obtain further data on the animals judged best from initial EBVs when the amount of further data collection was limited, which is a form of independent culling level selection. The suggestion that attention might be concentrated on animals with intermediate EBVs is similar to the concept of selection of extremes due to Abplanalp (1972). Independent culling levels involves a series of tests, all of which a candidate must pass in order to be selected. Selection of extremes also involves a series of tests, all of which an animal must fail in order to be culled. Abplanalp showed that in some situations it was better than independent culling levels in terms of genetic gain, though it is more expensive, since each animal must be tested for all traits, while with independent culling levels those that fail the first test will not be given the second test.

## TESTING MALES AND FEMALES

Suppose a breeder has facilities for testing 1000 animals for a quantitative trait and will select 10 males and 500 females from among those tested. If he tests 500 of each sex he will have no selection differential in females unless he replaces the worst tested females with untested ones. Suppose this is done. Then he selects 10 out of 500 males with a selection differential of  $2.42\sigma$  and 250 out of 500 females with a selection differential of  $0.80\sigma$  together with 250 untested females so that the overall female selection differential in females is  $0.40\sigma$  and averaged over both sexes it is  $1.41\sigma$ . But if the number of males tested was to be 400, the males would have a selection differential of  $2.34\sigma$ , and there would be 300 selected tested females with a selection differential of  $0.80\sigma$  and 200 with a value of zero, so that the average would be  $0.48\sigma$ , and averaged over sexes it would be  $1.41\sigma$  also. Essentially these are the same.

Now suppose we want to select 10 males and 250 females. In this case with equal testing we can select all animals required from among those tested. The selection differential in males would again be  $2.42\sigma$ , while that in females would be  $0.80\sigma$  with an average of  $1.61\sigma$ . If 400 males and 600 females are tested the selection differentials are  $2.34\sigma$  and  $0.93\sigma$  respectively, giving an average of  $1.64\sigma$ . There is thus in this case a small advantage to testing more females than males.

Smith (1969) made a theoretical analysis of this situation using his approximation to the standardised selection differential

$$i = 0.8 + \ln[(1 - p) / p] = 0.8 + \ln(\theta - 1)$$

where  $i$  is the standardised selection differential,  $p$  is the proportion selected, and  $\theta = 1/p$ . Here I will give a brief idea of what is involved.

Suppose we have  $T$  testing places and we need to select  $S$  sires and  $D$  dams from among  $N$  available animals of each sex, while  $T \leq 2N$ , so that not all can be tested. The breeder decides to allocate a fraction  $f$  of testing places to males, and  $1 - f$  to females. To avoid complications we assume that all animals required may be selected from those tested, as in the second example above. The required numbers as fractions of the numbers tested will then be  $S/Tf$  and  $D/T(1 - f)$  respectively. Then if we write  $\theta_M$  and  $\theta_F$  as the reciprocals of the selection intensities the selection response is proportional to  $i_M + i_F$  which can be written as, approximately,

$$0.8 + \ln(\theta_M - 1) + 0.8 + \ln(\theta_F - 1).$$

To find the optimal value of  $f$  we differentiate this expression and set the derivative to zero. The derivative can be written

$$\frac{T/S}{\theta_M - 1} - \frac{T/D}{\theta_F - 1} = \frac{T}{fT - S} - \frac{T}{(1 - f)T - D},$$

and setting this to zero we find

$$f = 0.5 - \frac{D-S}{2T} \text{ and } 1-f = 0.5 + \frac{D-S}{2T}.$$

Clearly there will be more complicated considerations when there are not enough testing places to meet the requirement that all selected animals are tested. The condition will be met if the number of females tested is at least twice the number required, or  $3S + D \leq T$ . For the numerical example above with  $S = 10$ ,  $D = 250$ ,  $T = 1000$ , the theory above would give optimal numbers tested as 380 males and 620 females, quite close to the values used in the example. The advantage in this case is fairly small, but the principle is clear; it is better to use extra facilities to increase the selection differential in the sex in which more animals are to be selected.

In many cases, the two sexes will be selected for different traits, with different correlations with overall breeding value, and with different measurement costs. The analysis above can be easily modified to deal with these complications.

Let  $r_M$  and  $r_F$  be the correlations between male and female traits and overall breeding value, and let  $c_M$  and  $c_F$  be the respective costs of measuring a male and a female, while  $C_T$  is the total cost of measurement in the budget. If we set  $T_M = C_T/c_M$  with  $T_F$  defined similarly for females, these are the total numbers of males and females which can be measured if the whole budget is devoted to measuring one sex. Let  $f$  be the fraction of the testing budget devoted to testing males. Then the number of males tested is  $fC_T/c_M$  or  $fT_M$  while the number of females tested is  $(1-f)T_F$ . The selection intensities in the two sexes are  $S/fT_M$  and  $D/(1-f)T_F$ .

The response to selection is proportional to  $i_M r_M + i_F r_F$  and we again use Smith's approximation to  $i$ . Here  $\theta_M = fT_M/S$  and  $\theta_F = (1-f)T_F/D$  and we need to maximise

$$r_M[0.8 + \ln(\theta_M - 1)] + r_F[0.8 + \ln(\theta_F - 1)]$$

with respect to variation in  $f$ . On differentiating and setting the derivative to zero we have

$$\frac{r_M T_M}{S(\theta_M - 1)} = \frac{r_F T_F}{D(\theta_F - 1)}$$

from which the solution is

$$f = \frac{r_M}{r_M + r_F} \left[ 1 - \frac{D}{T_F} + \frac{r_F}{r_M} \frac{S}{T_M} \right].$$

This shows that, other things being equal, more should be expended on the more accurate evaluation, which comes as no surprise. Again, if the achievable proportion selected is greater in females than in males, other things being equal, more effort should be devoted to increasing the measurement of females. This conclusion is sometimes regarded as wrong because of a prejudice that since male selection is more powerful than female selection more effort should be put into evaluating males. This is not necessarily true. The optimal value above can easily be seen to reduce to the previous one if evaluation is equally accurate and equally costly in the two sexes, as it should.

A somewhat different approach was initiated by Jackson, Lax and Wilson (1986), who assumed that a range of measurements could be made on males and females, and that the accuracy of selection was dependent on measurement expenditure with diminishing returns. Specifically, they assumed that the accuracy of selection  $r_{G\hat{G}}$  is related to the cost of measurement  $c$  by the

relation  $r_{G\hat{G}} = r_{\max}(1 - k^c)$  so that the accuracy is zero when  $c = 0$  and is  $r_{\max}$  as  $c$  becomes very large. Of course, costs will not increase smoothly as implied by this equation, but will rise in a stepwise manner as more traits are added, but the equation allows some consideration to be given in a general manner. This approach was also used by Wade and James (1990), also with

reference to Merino sheep breeding. They followed Jackson et al. in taking  $k = 0.7$  and presented some results to indicate the type of outcome which might be observed. Assuming 50% of ewes and 1% of rams are required to be kept among the 500 candidates of each sex, a unit of cost = \$0.50 and a total measurement budget of \$1250 is allowed, they found that the optimum allocation was :

Test 170 rams at \$4.05 per head	\$687.65
Test 500 ewes at \$1.13 per head	\$562.50

The selected proportions would be 3% in rams and 50% in ewes, with accuracies of  $0.94r_{\max}$  in males and  $0.55r_{\max}$  in females. They also found that with this relation between cost and accuracy it always paid to measure at least twice as many animals as were required if that was possible, so that there was no need to select untested animals. Of course if more than half of the available females must be kept, this is impossible. In this example the proportion of the budget spent on testing males is  $687/1250 = 0.55$ . If we apply the formula previously derived we have  $r_M/(r_M + r_F) = 0.63$ ,  $T_M = 1250 / 4.05 = 309$ ,  $T_F = 1250 / 1.23 = 1106$ ,  $r_F/r_M = 0.58$ , so that

$$f = 0.63[1 - \frac{250}{1106} + 0.58\frac{5}{309}] = 0.49.$$

Thus the approximation does not give the same result, though both methods suggest that about the same amount should be spent on evaluating each sex, but that fewer males should be tested with a higher accuracy than females.

In the treatment given so far it has been assumed that the generation interval is fixed. However, it was pointed out by Ollivier (1990) that if generation interval is allowed to vary, the generation length will be affected by the amount of testing done, since the testing facilities will be on a per year basis, and by changing the numbers required for testing the breeder could alter the time needed for breeding replacements. He showed that when testing facilities are very limited, only males should be tested, and the maximum proportion to be selected is 27%, any other males required being taken at random from untested males. This is based on the following development.

Assume that the proportions of males and females required for breeding are  $p_M$  and  $p_F$  while the remainder  $(1 - p_M)$  and  $(1 - p_F)$  are chosen at random from untested animals. The standardised selection differentials are  $i_M$  and  $i_F$  respectively. The generation intervals for the selected males and those taken at random are  $t_{M1}$  and  $t_{M0}$  with similar definitions for females. Then the annual selection response is proportional to

$$\frac{p_M i_M + p_F i_F}{p_M t_{M1} + (1 - p_M) t_{M0} + p_F t_{F1} + (1 - p_F) t_{F0}}$$

Assuming no wastage of breeding animals and fertility the same at all ages, the generation lengths can be written

$$\begin{aligned} t_{F1} &= a + p_F n_F / 2q_F & t_{F0} &= a + (1 - p_F) / 2(\lambda - q_F) \\ t_{M1} &= a + p_M n_M / 2\phi q_M & t_{M0} &= a + (1 - p_M) / 2\phi(\lambda - q_M) \end{aligned}$$

where  $\lambda$  is the number of progeny of each sex per dam per year,  $\phi$  is the number of females mated to each male,  $n_M$  being the number of males tested per year per male selected, with  $n_F$  defined similarly, and  $q_M$  is the number of males tested per breeding female per year, with  $q_F$  defined similarly. If  $k$  is the proportion of the available animals of both sexes which can be tested annually, and  $f$  is the fraction of testing places given to males, then  $q_F = 2k\lambda(1 - f)$  and  $q_M = 2k\lambda f$ . Using these relations and the relation between proportion selected and selection differential gave the result quoted above.

Ollivier considered a number of cases other than this simplest one, and identified conditions which led to testing of males only, either with or without some use of untested males, and to testing of some females. He let  $k_1$  be the value below which some untested males should be used and  $k_2$  denote the value above which some females should be tested. A small sample of the results is given in the table below.

Females per male ( $\phi$ )	10			100		
Value of $a\lambda$	0.5	1.0	5.0	0.5	1.0	5.0
$k_1$	0.061	0.037	0.009	0.006	0.004	0.001
$k_2$	0.77	0.65	0.28	0.87	0.75	0.35

For higher female reproduction rate the value of  $k$  at which all selected males should come from those tested is lower, and in any case it is only for very restricted testing facilities that untested males should be used. Of course, untested females will be used in many situations. Ollivier's paper should be consulted for further details.

## TWO-STAGE SELECTION

All of the above analyses were carried out assuming selection in a single stage. But it is often pointed out that if sequential culling is used, only those animals which survive the first culling need to be measured at the second stage, and therefore the survivors can be measured for more traits than would be possible for the same cost if there were no culling. Wade and James (1996) dealt with the case of two-stage selection in conjunction with the approach taken above for single-stage selection.

There was no simple solution in this case (at least none was found) and results were obtained by using computer programs which are described in the paper. First they estimated the relationship between cost and accuracy of EBVs using a series of measurements appropriate for selection of Merino sheep (about 10 years ago). The value of  $k$  in the relation  $1 - k^c$  with  $c$  measured in dollars was 0.8, which fitted quite well, though there were a few measurement combinations for which the fit was not so good. A numerical example will illustrate what can be done.

A breeder has a flock of 1200 ewes which produce 1000 lambs annually, and a budget of \$2000 for measurement to allocate among the 500 progeny of each sex. The proportions required for breeding were 5% of rams and 50% of ewes. The optimum was to allocate 0.75 of the total to males and 0.25 to females. Thus the expenditure should be \$3 per ram and \$1 per ewe. Further analysis gave the following allocation of testing funds.

Test 250 ram hoggets at \$3.90 per head in stage 1, then choose the best 75 and test them in stage 2 at \$7.00 per head. This gives a total male testing cost of

$$250 \times \$3.90 + 75 \times \$7.00 = \$1500.$$

For females the best option was to test 500 in stage 1 at a cost of \$0.65 per head, select the best 400 and test them at stage 2 at a cost of \$0.44 per head, giving a cost of

$$500 \times \$0.65 + 400 \times \$0.44 = \$501,$$

and the total cost was \$2001 as budgeted.

In practice, these solutions would not be achievable. An appropriate stage 1 criterion for males might be clean fleece weight at \$4.00 per head, with perhaps mean fibre diameter and coefficient of variation of diameter at \$6.00 per head in stage 2. A visual classing criterion at \$1.00 per head might be added. For females, no measurements meeting the suggested costs are obvious. It might then be best to measure all ewes at a single stage for a trait such as hogget body weight of



a visual wool trait. It would be necessary to consider the costs and accuracies of available traits in order to find a feasible plan.

There were often only small differences between various options and similar gains could be achieved using quite different allocations of funds, so not too much attention should be paid to the actual optima reported. What is important is that in planning a breeding program a variety of possible strategies needs to be considered.

## CONCLUSION

With the ready availability of quick and cheap computing, the planning of a breeding program should involve the modelling of a wide range of possible schemes with attention given to the way that funds are spent on obtaining data which provide efficient means of making genetic improvement.

## REFERENCES

- Abplanalp, H. (1972) Selection of extremes. *Animal Production*, **14**:11 – 15.
- Henderson, C.R. (1963) Selection index and expected genetic advance. In *Statistical Genetics and Plant Breeding* (W.D. Hanson & H.F. Robinson eds) NAS-NRC Publication **982**:141 – 163.
- Jackson, N., Lax, J. and Wilson, R.L. (1986) Sex and selection for fleece weight in Merino sheep. *Journal of Animal Breeding and Genetics*, **103**:97 – 115.
- James, J.W. (1966) Incomplete sampling for selection of a fixed proportion. *Australian Journal of Statistics*, **8**:127 – 137.
- James, J.W. (1979) Optimum family size in progeny testing when prior information is available. *Zeitschrift für Tierzüchtung und Züchtungsbiologie*, **95**:194 – 203.
- Ollivier, L. (1990) Mass selection in livestock using limited testing facilities. *Genetics Selection Evolution*, **22**:109 – 117.
- Rendel, J.M. (1958) Optimum family size in half sib family selection. *Biometrics*, **15**:376 – 381.
- Robertson, A. (1957) Optimum group size in progeny testing and family selection. *Biometrics*, **13**:442 – 450.
- Smith, C. (1959) A comparison of testing schemes for pigs. *Animal Production*, **1**:113 – 121.
- Smith, C. (1969) Optimum selection procedures in animal breeding. *Animal Production*, **11**:433 – 442.
- Wade, C.M. and James, J.W. (1990) Optimal allocation of test funds in selection for genetic gain. *Proceedings of the Australian Association of Animal Breeding and Genetics*, **8**:115 – 118.
- Wade, C.M. and James, J.W. (1996) Optimum allocation of resources considering two sexes and selection in