

## The All-Breed Animal Model

*Bennet Cassell, Extension Dairy Scientist, Genetics and Management*

### Introduction

The all-breed animal model is the genetic-evaluation system used to evaluate dairy animals in the United States. Scientists and technicians at the Animal Improvement Programs Laboratory (AIPL) at Agriculture Research Service in Beltsville, Md., developed and support the system. There are two major differences between the all-breed animal model and the single-breed animal model it replaced:

1. Accurate genetic evaluations of animals with relatives in more than one breed are now possible.
2. Animals of different breeds within the same herd are now used together for contemporary comparisons.

A single genetic-evaluation system, with minor modifications, is used to evaluate dairy animals for milk, fat, protein, productive life, somatic cell score, and daughter pregnancy rate. Purebred cows and bulls of all breeds (animals of one breed, though not necessarily registered), as well as crossbreds are evaluated simultaneously. Performance records on cows provide the basic information used in the all-breed animal model.

Genetic relationships between males and females rely on an extensive pedigree file provided by Dairy Herd Improvement (DHI) and breed societies. This massive pedigree database allows distant ancestors, long removed from dairy herds, to impact the genetic evaluations of bulls and cows alive today. Each generation that separates two related animals halves the genetic relationship, meaning that evaluations of close relatives, such as parents with sons or daughters, have greatest impact on an animal's evaluations. The all-breed animal model uses a common genetic base for all genetic evaluations, but proofs are converted to within-breed scales prior to publication.

Genetic evaluations are published in January, April, and August of each year and distributed throughout the industry. Dairy farmers have access to genetic evaluations on animals in their herds through the records processing centers that handle DHI records. Other public distribution sites include the breed societies, National Association of Animal Breeders (NAAB), and the Virginia Cooperative Extension website, VT Dairy ([www.vtdairy.dasc.vt.edu/genetics.htm](http://www.vtdairy.dasc.vt.edu/genetics.htm)).

### Important terms

*Predicted transmitting ability:*

Predicted Transmitting Ability (PTA) is an estimate of genetic merit and is the part of an animal's genetic makeup that is transmitted to offspring. It is the piece of information about each trait on each animal that should be used to make selection decisions. Published PTAs have the same genetic base, the same meaning, and the same interpretation for cows and bulls within each breed. PTAs on crossbred animals are expressed on the genetic base of the breed of the animal's sire. PTAs predict the change in progeny performance due to an average sample of genes transmitted by the parent for which the PTA was calculated. PTAs can be used to rank animals for genetic merit for a specific trait and to estimate genetic differences between pairs of animals.

*Genetic base:*

PTAs are computed relative to a zero point or "genetic base." For evaluating dairy cattle, a new genetic base is calculated every five years to account for genetic change. The current genetic base used for the underlying, unpublished all-breed PTAs sets the average PTA for each trait to zero for all cows born in 2000, regardless of breed. Thus, a cow with a PTA for productive life of 2.0 would be expected to transmit genes to her

progeny that enable two additional months in the milking herd compared to average cows, across all breeds, born in 2000.

PTAs are not published using the common genetic base for all breeds. They are converted so that PTA averages 0.0 for each trait within each breed. (Details of the conversion appear later in this document.) This adjustment prevents major changes in the scale of published proofs for non-Holstein breeds. Holstein animals, by far the prevailing breed in the United States, dominate the all-breed genetic base. At some point in the future, direct comparisons of animals from different breeds may become important enough to abandon separate genetic bases for the different breeds and publish all evaluations relative to a single genetic base.

#### *Reliability:*

The amount of information about the genetic merit of dairy animals is measured by reliability. This measure of the accuracy of a genetic evaluation varies from animal to animal and ranges from 0 percent for unevaluated animals to 99 percent. The value is usually expressed as a percentage. Reliability increases for most cows and bulls as they mature and more information is accumulated from their own or their progenies' records.

The first information available on young animals is pedigree data, the accuracy of which depends on reliability of the genetic evaluations of sire and dam. As heifers mature into cows, they increase the reliability of their own PTAs through performance information. Maximum information comes from thousands of progeny of older, highly popular AI bulls. Reliabilities tend to be higher for more heritable traits, such as milk components or stature, and lower for less heritable traits, like daughter-pregnancy rate. With many thousands of progeny, however, maximum reliability for all traits is 99 percent, regardless of heritability.

## **Data used in the all-breed animal model**

Herds on a wide variety of DHI testing plans provide the data for national genetic evaluations. DHI records are forwarded from four regional record-processing centers to the Animal Improvement Programs Laboratory on a continuous basis. Testing of some herds uses supervised monthly visits with milk weights and samples from every milking in a 24-hour period, while testing of other herds may report only milk weights four times to six times per year, with no DHI techni-

cian present at milking. Years ago, only records from supervised monthly testing programs that included milk samples at every milking found their way into proofs. Through the years, the national genetic evaluation system has evolved to include more records in genetic evaluations. A system called "data collection rating" (DCR) assigns a weight to each record, depending on the conditions under which the data were collected. DCR is used in the genetic evaluation process to place more emphasis on data from systems judged to be more complete or thorough.

Performance records without proper identification have little use in genetic evaluations. Identification includes a unique animal ID, a unique herd ID, the animal's date of birth, sire and dam ID, and freshening dates. Performance records include test-day information on milk, fat, protein, and somatic cell score. Breeding records are also used in genetic evaluations. Test-day dates and breeding/calving dates are also very important. The integrity of data collection systems is the foundation on which genetic evaluations are based. An important concept in the all-breed animal model is that performance data are originally collected to improve dairy-herd management, and dairy farmers pay the bill. The data are donated to the genetic evaluation system after serving the primary purpose of management of individual dairy herds. Future changes in data collection systems will continue to be driven by this "management first" policy.

## **How performance records are used**

The raw data collected from DHI processing centers are test-day milk weights. As not all cows have the same number of test days in the 305-day period used as a standard lactation, some method of turning test days into 305-day lactation totals is needed. The process used is called "best prediction." It is based on the shape of the lactation curve through similarities (correlations) between test days estimated from the vast data sets collected for genetic evaluations through the years. The system can estimate 305-day yields on cows with three test days or 11 test days in the first 305 days following calving. All test-day records are processed at AIPL, insuring that a consistent procedure is used. All records are standardized to a twice-a-day milking frequency basis, and are adjusted to a common age. The standard is a 36-month old, second lactation cow, chosen because she is close to the average age of all lactation records used for genetic evaluations. Years ago, a "mature" age of 60 months was used, but many cows were culled or died before reaching that age.

The all-breed animal model includes up to five lactation records on cows. A first lactation is required or special rules (explained later) come into play. Lactations from as early as 1960 are included in evaluations. Exclusion of lactations after the fifth removes very few records, as less than 10 percent of dairy animals have a sixth record. Even for cows with six or more records, the first five records represent a high percentage of all performance information on that cow. Discarding very mature production records from genetic evaluations may actually improve the accuracy of PTAs on mature cows, as most herds have few true contemporaries against which to compare such records. Also, by sixth lactation, permanent environmental effects can overwhelm genetic differences.

Information for daughter pregnancy rate and somatic cell score also comes from the first five lactation records. Somatic cell score is a lactation average of individual test days. Daughter pregnancy rate is measured from calving intervals. Productive life is unique in that it is expressed only once at the end of the life of a cow. For most cows of interest in breeding programs, productive life evaluations are based on partial herd-life records. For more detail of how these traits are included in genetic evaluations, see *Sire Evaluations for Health and Fitness Traits*, Virginia Cooperative Extension Dairy Guideline 404-087, (<http://pubs.ext.vt.edu/404-087/>).

## Contemporary groups: how and why

Environmental conditions have a strong impact on cow performance, and often eclipse genetic differences between cows in the process. The greatest challenge to any genetic evaluation system is to separate environmental effects from genetic ones. Fortunately, many environmental factors affect cows the same way. A very effective way to remove environmental effects is to compare performance of one cow to other cows, called “contemporaries,” that were treated similarly. In the all-breed animal model, cows are assigned to contemporary groups within each herd in a process that differs slightly from trait to trait. For yield traits, cows are divided into first or later lactation groups, and then subdivided into two-month season of freshening groups. Herds with a mixture of registered and grade cows are also divided by registry status. Management groups need to be reasonably large to measure environmental differences. If fewer than five cows are in a group, groups are combined in the following order to

meet the minimum:

1. Two-month groupings increase to four months
2. Registered and grade cows are combined
3. Four-month groupings increase to six months
4. First and later lactations are combined
5. Six-month groupings increase in increments of two months, to a maximum of 12-month groupings

Combining groups sacrifices uniformity of environmental or management conditions. However, there is a trade-off between short calving periods and the number of animals per group. Larger groups are more effective at measuring environmental conditions than smaller groups. There is an inherent advantage to larger herds where contemporary groups can be narrowly defined and yet contain many animals. One advantage of the all-breed animal model over its predecessor, the within-breed animal model, is that for herds containing more than one breed, cows of different breeds or properly identified crossbreds form larger contemporary groups.

## Genetic versus permanent environmental effects

Dairy cows produce multiple records for all traits except productive life. Repeated records tend to be similar for two very different reasons: genetic effects, which can be transmitted to progeny, and permanent environmental factors, which affect all records but are not transmitted to progeny. Mastitis can cause a permanent environmental effect by creating scar tissue that reduces milk yield for the life of the cow, yet the mastitis itself may have been entirely of environmental origin. Accurate genetic evaluations require separation of genetic and permanent environmental effects. Genetic relationships between animals, discussed next, enable that process.

## Genetic relationships

Pedigree data is a vital part of accurate genetic evaluations. It adds greatly to what we know about the genetic merit of individual cows and is essential to estimate genetic merit on bulls for sex-limited traits such as milk production. Pedigree information helps separate favorable or unfavorable permanent environmental effects from genetic effects that can be passed to offspring. For example, a cow with favorable parentage may not produce very well because of permanent effects of mastitis as a young cow. Such cows often produce less

than their contemporaries, even though they may be genetically superior to those contemporaries for milk production. Knowledge of the outstanding performance of many relatives in other herds allows the all-breed animal model to remove the highly unfavorable permanent environmental effects from yield deviations, producing a more accurate estimate of true transmitting ability in the process.

One of the most important genetic relationships in dairy cattle breeding is between a daughter of an AI bull and her many half-sisters distributed across many different herd environments. It is used to evaluate the bull himself and contributes to the accuracy of the PTA on each of the half-sisters as well. Producers want to use those bulls whose progeny distinguish themselves in many diverse environments. We could not identify top proven AI bulls without genetic relationships. Widespread use of AI and heavy use of the best bulls have created a complex pedigree structure within the dairy breeds. It would be challenging indeed to find a U.S. Holstein cow that was not related to Elevation or Chief, or a U.S. Jersey cow that didn't have Duncan or Generator in her pedigree somewhere. That means Holstein or Jersey cows are related in some way to almost all other cows in their breed.

Pedigree relationships are incorporated into genetic evaluations through a process called "iteration." Each genetic evaluation begins where the previous evaluation, several months earlier, ended. Updated production records are then added to the evaluations of cows. As computations continue, evaluations on bulls are computed from updated evaluations of daughters in the first iteration. Once a bull's daughters are used to estimate a breeding value for him, his evaluation becomes part of the genetic evaluation of each of those daughters in the second iteration. Evaluations influence more remote ancestors and descendants in each succeeding round of calculations. The influence of one cow's records on the PTAs of her relatives is like a ripple on a farm pond, ever wider in its influence, but less dramatic in impact as more and more remote ancestors and descendants come under the influence of her evaluation.

Pedigree information for genetic evaluations is not endless, though even the most remote animals in a pedigree trail had a sire and a dam. The all-breed animal model uses first lactation records from 1960 and later, plus any information from ancestors born since 1950. At that starting point, a series of special categories of animals called "unknowns" is formed. Unknowns are grouped together by sex and year of birth (or apparent

year of birth) and "solutions" are computed for each group.

There are other ways that pedigree paths end. Some end with animals that add a single progeny and no records to the data. These animals are "known" but add no useful information for genetic evaluations. Parentage information is missing on some animals, though always on the female side of the pedigree, as cows with records but no sire identification are removed from genetic analysis. These endings also form unknown parent groups.

The impact of pedigree on any animal's PTA under the all-breed animal model depends on other information available. For instance, the genetic evaluation of the maternal grand-dam of a proven AI bull tells us much less about his genetic merit than his many hundreds of daughters. The animal model properly weights each item of information relative to other indicators of merit on each cow or bull evaluated. These weights change over time as new information enters the system. For instance, pedigree information is all that is available on a young sire. Once progeny begin to freshen, pedigree information becomes less important over time until it plays very little role in a PTA when thousands of daughters are available.

Crossbred animals are used in genetic evaluations by the all-breed animal model if they are properly identified and trace back to purebred sires. Crossbreds played no role in the evaluation system prior to August 2007, unless they were coded as purebreds in identity enrollment programs. Crossbreds are useful for improving the accuracy of genetic evaluations on their purebred sires and dams. Genetic information on crossbred cows can be used in herd management decisions. Further, crossbreds tie the different breeds together for genetic evaluations, because they contribute to genetic information on at least two and possibly more breeds at the same time. Without crossbred cattle, pedigree relationships would terminate at the limits of information on individual breeds.

## **Merit of mates**

The all-breed animal model considers the genetic merit of the other parent before including progeny information in the PTA of a cow or bull. This adjustment eliminates possible bias, either favorable or unfavorable, from selected mates. Otherwise, proofs on very popular bulls with high semen prices could be biased upwards as the bulls aged and were mated to better and

better cows. Cows with poorer genetic merit may be used more frequently as mates for young bulls in progeny testing programs or in crossbred matings. Without adjustment for genetic merit of mates, such practices would bias proofs downward. Adjustment for merit of mates relies on accurate identification of both sires *and dams* of cows used for genetic evaluations. Most dairy farmers in the United States appreciate the value of accurate sire ID for genetic evaluations, but may not be aware of how important accurate dam ID has become.

## Special procedures for cows missing first lactations and cows changing herds

A cow must have a first lactation to be used for estimates of *her relatives'* PTAs. This is to prevent bias which could occur if cows were screened for testing based on their first (nontested) lactation, and only the highest producers placed on test to prove a bull, for instance. This is not a particularly practical example. Some cows are sold from nontested into tested herds or are members of herds that choose to go on test after a cow has one or more records. For cows missing first records, the later records that are available and the PTAs on all of her relatives are used to evaluate the cow herself. However, her evaluation is not used to evaluate her relatives. This procedure protects animal-model evaluations from possible bias, yet permits the most accurate possible evaluations of cows lacking first records. The rules for cows missing first lactations are not used for protein evaluations when first lactation milk and fat records are present. Many DHI laboratories always perform fat and protein tests, but protein testing is optional in some parts of the country.

Cows that change herds at some point after they produce their first calves present a special problem. Some of these cows are very valuable, perhaps purchased at high prices. Genetic evaluations on such cows are important for marketing. However, when a cow changes herds, she also changes contemporaries, and there have been instances of deliberate manipulation of contemporary groups to obtain favorable genetic evaluations. When a cow changes herds, only data from the first herd are used to evaluate *relatives* of the cow. However, records from herds in which the cow later appeared are used for the evaluation of the cow herself. Again, the purpose is to provide information on individual animals while protecting evaluations of relatives from inflated or deflated estimates on individual animals.

## Breed comparisons

Table 1 below shows breed averages for the different traits for cows of each breed born in the genetic-base year, 2000. As expected, the averages vary considerably from breed to breed. However, when records are combined across breeds, the overall average looks a lot like the Holstein average, because 90 percent of all records in the AIPL system are from Holstein cows. All of the yield traits favor Holsteins by a substantial margin, especially milk production. Jerseys have the longest productive life and the highest daughter pregnancy rate. The breed with lowest somatic cell score is Brown Swiss. A byproduct of the all-breed animal model is a comparison of breed genetic merit for these traits, information of interest to breed enthusiasts as well as practical use for design of crossbreeding programs.

## Adjustments to published proofs

The original PTAs produced by the all-breed animal model could be used to compare genetic merit of dairy animals directly, regardless of sex or breed. However, the long tradition of calculating PTAs within breeds created expectations among breeders that would not be met if the all-breed results were published as they were calculated. For example, Jersey bulls would have very low proofs for milk production compared to Holstein bulls. Also, differences between PTAs in Jerseys are smaller than differences between Holsteins for milk yield because of more variable milk production for Holstein cows. Holsteins do not lead in every trait, however. Jersey proofs for productive life would be higher than productive-life proofs for Holsteins, because Jersey cows last nearly six months longer in dairy herds than Holstein cows.

Perhaps breed-neutral PTAs will be used in the industry at some time in the future, but as this publication is written, they are converted to a within-breed base that resembles proofs from the single-breed system. This conversion is made using Equation 1 for each bull or cow.

**Table 1: Trait averages for cows born in the genetic base year 2000.\***

Breed	305 day, 2X, Mature Equivalent, lbs			Productive life, mo	Somatic cell score	Daughter pregnancy rate, %
	Milk	Fat	Protein			
Ayrshire	18,152	699	567	32.1	2.96	21.8
Brown Swiss	21,338	859	705	30.5	2.92	20.4
Guernsey	16,628	736	542	26.6	3.29	19.9
Holstein	25,436	927	763	27.6	3.07	21.0
Jersey	17,864	819	632	33.4	3.33	26.0
Milking Shorthorn	16,940	604	523	29.9	3.10	24.0

\* Based on records included in the August 2008 genetic evaluations

**Equation 1: Within-breed PTA = [All-breed PTA – breed mean] X [breed SD/Holstein SD]**

This formula can be rearranged to convert a within-breed proof to the all-breed base. That equation is:

**Equation 2: All-breed PTA = {[Within-breed PTA] X [Holstein SD/breed SD]} + breed mean**

The breed mean used in these equations can be found in Table 2 below. The standard deviation ratios (breed SD/Holstein SD) are available on the AIPL website: ([www.aipl.arsusda.gov/](http://www.aipl.arsusda.gov/)).

Published proofs for bulls are also adjusted for inbreeding expected from mates likely to be encountered in purebred matings. Each bull’s pedigree is used to calculate inbreeding coefficients against a sample of likely mates. The average of these inbreeding coefficients is called “expected future inbreeding” (EFI) and is combined with estimates of inbreeding depression from research trials to reduce proofs for average inbreeding in purebred matings. For instance, each 1 percent increase in EFI reduces milk proofs by 65.3 pounds. The published proof on a bull with an EFI of 4 percent would be reduced by 261 pounds (4 X 65.3 = 261). The published proofs used in Equation 1 or Equation 2 will include such an adjustment for every bull. For a fair comparison of bulls across breeds, the inbreeding adjustment should be removed from the published PTA before conversion to an all-breed base. EFI on individual bulls is available in a lookup system on the AIPL website, noted above. Appropriate estimates of inbreeding depression per 1 percent EFI appear on that website as well.

The bottom line is that published proofs are two steps removed from the calculated proofs that emerge from the all-breed animal model. The first adjustment scales the proofs and expresses them relative to other animals of the same breed. The second adjustment reduces proofs for inbreeding anticipated from matings likely to occur in the very near future.

## Genetic differences between breeds

Breed differences in PTAs from Holsteins for the August 2008 genetic evaluations are shown in Table 2. The numbers are PTAs or transmitting abilities rather than breeding values, which are (2\* PTA). Holsteins dominate for production traits, but are inferior to Ayrshires, Brown Swiss, and Jerseys for productive life. Jerseys are closest to Holsteins for fat production while Brown Swiss are closest for protein production. These differences are average PTAs for each breed from the all-breed animal model, but they are expressed relative to average Holstein PTAs, which are close to zero because of the dominance of Holstein phenotypes in the all-breed animal model.

Conversion of proofs from one breed base to another for choosing sires for use in crossbreeding is tedious and not likely worth the effort for most breeders. Bulls can be chosen for use outside of their own breed with little loss of accuracy by using the published within-breed PTAs. Use a comprehensive lifetime economic index such as Net Merit for this purpose. The best bulls for purebred matings are almost always the best bulls for crossbred matings, but differences between bulls for EFI can cause exceptions. If a bull is in second place for overall merit but has higher EFI than the first-place bull, he might move into first place for crossbred matings where EFI is inappropriate.

**Table 2: Breed differences from Holstein for PTA in August 2008 genetic evaluations**

Breed	Milk, lbs	Fat, lbs	Protein, lbs	Productive life, mo	Somatic cell score	Daughter pregnancy rate, %
Ayrshire	-2,610	-68	-65	0.2	-0.07	1.3
Brown Swiss	-2,108	-40	-36	0.4	-0.05	0.5
Guernsey	-3,043	-41	-68	-4.4	0.04	0.5
Jersey	-3,268	-39	-54	1.2	0.09	2.8
Milking Shorthorn	-3,367	-117	-94	-0.4	-0.04	2.8

## Net merit and rank percentiles

The Animal Improvement Programs Laboratory uses the all-breed animal model to calculate PTAs on production traits, somatic cell score, productive life, and daughter pregnancy rate. Breed societies calculate PTAs for type traits using similar procedures but on a within-breed basis. AIPL uses a different method to obtain PTAs for calving ease and stillbirths. There is a lot of information on individual traits on many animals. Which animals offer the best combination of genetic merit for all of these traits for herd improvement programs?

Proofs for individual traits should be combined into a single number based on the relative value of each trait for lifetime economic merit of the dairy cow. The Net Merit Index is available on all cows and bulls in the AIPL database. It is the single most useful selection index yet developed for commercial dairy farmers. For more detail on how net merit is calculated, see *The Merit Indexes – 2006 version*, Virginia Cooperative Extension Dairy Guideline 404-088, (<http://pubs.ext.vt.edu/404-088/>).

Rank percentiles are a valuable addition to net merit, as they tell the dairy farmer where a bull ranks compared to other choices available in the semen marketplace. Rank percentiles are based on *sires of the same breed* in active AI service in the previous evaluation. Rank percentiles for cows are based on *cows of the same breed* with recent lactations. Rank percentiles are determined within breed *and sex* and cannot be used to compare cows or bulls of one breed with cows or bulls of another breed.

A rank percentile of 80 means the cow or bull has net merit that is superior to 80 percent of the active AI bulls or cows from the previous evaluation. Rank percentiles decline as animals become older, provided the genetic

evaluation of the animal itself does not change dramatically. The reason is because the genetic merit of the group of animals in the previous evaluation, the ruler used to measure rank for net merit, does *not* remain the same. It increases as genetic progress in the population occurs. Percentile rank for net merit from the most recent genetic evaluation is the single most useful number related to genetic merit of dairy animals.

## Using animal model evaluations

The all-breed animal model is the most accurate method ever used to evaluate breeding values of dairy animals for production traits. The system uses all available information, appropriately weighting each item in the final evaluation of each animal. Sires, and when possible, cows with low PTAs should be avoided when semen is purchased or not used as dams of future herd replacements. As with previous genetic evaluations, PTAs from the all-breed animal model can be used for only two purposes:

1. To rank individuals for genetic merit, and
2. To estimate genetic differences between individuals.

## An eye to the future

Genetic evaluation systems are always a work in progress. Today's all-breed animal model is the latest development in an ongoing process of experimentation, discovery, and implementation of better ideas. As this guideline is written, the field of animal breeding is experiencing the most important technical changes since the advent of the herdmate comparison system in the 1960s. SNPs are "single nucleotide polymorphisms," each of which is a difference in a small part of the DNA molecule from one animal to another. Direct evaluation of the genome of dairy animals is now pos-

sible using “dense SNP marker array” technology. This technology allows genetic evaluation systems to track segments of DNA across generations and to relate those segments to phenotypes of animals and their progeny.

Genome scans, based on dense SNP array data, improve the accuracy of genetic evaluations for all animals, including those with progeny tests available. However, they are most useful for young animals with no records or progeny of their own. Young sires in AI sampling programs with genome data available will have reliabilities of 60 percent to 75 percent, and perhaps somewhat higher, depending on trait. With pedigree data only, reliabilities of AI young sires from pedigree information alone ranged from 30 percent to 40 percent.

The potential impact of this technology is immense, but at this writing, not firmly established. The basic approach of the all-breed animal model, using performance of individuals and their relatives to predict genetic merit, will not change. However, the assumptions of relationships between animals, formerly based on pedigree information, will now be augmented by knowledge of which segments of chromosomes were passed from parent to offspring. Since genome scans are relatively expensive, their use will be restricted initially to a small subset of the dairy population: parents of young sires considered for AI sampling programs and the young sires themselves. The technology will dramatically increase the accuracy of selection of young sires for progeny testing. Future chapters of this as-yet-unwritten story should prove interesting.