

Changes to evaluation system (August 2017)

Gestation length evaluations

By Paul VanRaden, Jan Wright, and Jay Megonigal

Genetic evaluations of males and genomic predictions of both sexes are now provided for gestation length (GL) as a new trait expressed in days. Differences in GL are useful to more accurately determine when to move cows to maternity pens, manage the fresh dates in herds with seasonal calving, and understand correlated effects of GL on other traits such as calving ease, stillbirth, and age at first calving. The edited GL data included a total of 12.4 million conception and calving dates of 6.8 million cows. The evaluation model defines GL as a trait of the service sire because maternal effects on GL are small compared to direct genetic effects of sire, dam, and Mendelian sampling of the calf. The all-breed model also includes crossbred matings and breed effect of dam. To predict the GL of a calf, add the predicted transmitting ability (PTA) of the sire plus the PTA of the dam (or half the PTA of the calf's maternal grandsire if the dam is not genotyped) plus the GL mean for the breed or breed combination. Breed means for the GL genetic base calves born in 2010 were 277 days for Holstein, 278 for Jersey, 286 for Brown Swiss, 281 for Ayrshire, 284 for Guernsey, and 279 for Milking Shorthorn, which may slightly underestimate actual means because the data included more heifer than bull calves. The PTAs, parent averages, and reliabilities for GL are provided in xml and csv files for all genotyped animals. For bulls, those same variables along with numbers of GL observations and herds are in bytes 654-674 of [format 38](#). Further details are provided in an [ADSA abstract](#) and [poster presentation](#).

Changes in unknown parent group definitions

By Paul VanRaden and Mel Tooker

For some traits, the most recent unknown parent groups had previously been merged together across all breeds if the most recent group had too few records, but are now merged separately by breed with the next most recent group. This change primarily affects traditional parent averages (PA) of recent calves with missing parents. Genomic evaluations were much less affected because most PAs are recalculated in the within-breed genomic evaluation, and because the PA gets less weight for those animals with low reliability. Evaluations for Holsteins did not really change because most animals in the last unknown parent group were Holsteins. In the other breeds, the youngest animals that had missing pedigree changed as a result of breed differences in the last group. The PA differences were small for most traits, but larger for milk yield of Jerseys and other high-component breeds, for example.

Type composite correction

By Ezequiel Nicolazzi and Paul VanRaden

CDCB and AIPL have reviewed the programs that are used to obtain the feet/leg composite used within merit indexes in Holsteins, correcting the way two traits are handled. At the time of its original development (https://aipl.arsusda.gov/publish/jds/2009/92_16.pdf) PTA for Feet and Legs Score (FLS) were not available. This PTA was therefore obtained from a highly correlated trait (FLS is 85% correlated to Rear Leg Rear View - RLR). Later, actual FLS PTAs were computed and distributed, but the substitute trait remained in the feet and leg composite (FLC) used in the merit indexes. The merit functions were updated in July 2017 to use the actual FLS PTAs in the FLC formula. No large changes are expected due to the high correlation with RLR and the minimal variation introduced in UDC. When applying the new programs to the April 2017 data, we observed changes on average of +2\$ in Net, Fluid, Cheese and Grazing merits, with maximum/minimum changes in the range of 20/-15\$. Since FLS is only included in Holstein Feet/Legs composite, no other breed will be involved in the update.

Introduction of new edits in calving traits

By Jay Megonigal

Beginning with the August 2017 evaluations, a new set of edits is being applied to the incoming calving ease and stillbirth phenotypic data. Among these changes, CDCB now requires pedigree and lactation data to be included in the database before the calving event is processed. This change will reduce the number of incorrect data entering our system, and will also allow the model to converge with greater accuracy. Another major component will be the return of an error file back to the submitting organizations that will highlight areas that can be addressed with the producers. This is the first step of a series of changes aimed at increasing the accuracy and robustness of phenotypic data used in evaluations. CDCB is also working to enhance the automated daily processing of incoming records with automated return of error files, similar to the procedures in place for lactation data. The new system is expected to be ready before the December 2017 evaluation.

Reduced computation in Holstein evaluations

By George Wiggans

Until recently, monthly genomic evaluations have fully reprocessed the data for all genotyped animals; however, computation took 6 days for 1.5 million animals. With that time increasing every month and the fact that there was no time to correct errors and still meet the distribution deadline in April 2017, we adopted some of the techniques used in the weekly evaluations. To reduce the 5 days required for imputation, only the new animals, those with changed pedigree, and parents and siblings of these animals are imputed. The imputed genotypes from the last full run are used for the other animals. For single gene effects and haplotypes affecting fertility, the weekly process also is used because the monthly process requires the full imputation results. These changes reduced the time required for the evaluation from 6 days to < 2. There are small differences from full recalculation because the imputed genotypes of ancestors and more distant relatives do not fully benefit from the new animals, and genotypes for newly qualifying non-genotyped dams are not generated, genotypes for existing non-genotyped dams continue to be used. This rapid calculation procedure relies on results from full runs for starting values for imputation and SNP effect estimates. We have been doing these full runs monthly so these prior values are available for the weekly runs and the following monthly run. The fact that the timing of these full runs is

flexible eliminates deadline pressure and allows for the increased computing time expected when the number of SNP used in the evaluation is increased (Wiggans et al., 2015).
