

# Changes to evaluation system (August 2010)

---

## Mexican cow records

By Paul VanRaden, George Wiggans, and Lillian Bacheller

**July 6 update:** Records from cows in some Mexican herds are now included in USDA evaluations. Records included must meet the same quality control standards required for U.S. herds, with inspection of meter centers, laboratories, service providers, and processing centers conducted by [Quality Certification Services, Inc.](#) Records with state code 66 (Mexico) are grouped in the southwest region to determine age adjustment factors and heterogeneous variance adjustments. The Council on Dairy Cattle Breeding approved this policy for including foreign data already in October 2003. Mexican cows are not double counted in international evaluations because Mexico does not send its own national evaluations to Interbull. From 1976-1996, the Mexican national evaluations for milk yield were computed by USDA but were separate from U.S. evaluations. The new approach of combined records should be more accurate by including progeny across both countries. As numbers of records increase, methods to account for genotype by environment interaction could become more important in the future.

---

## Validation of genomic evaluations

By Katie Olson, Paul VanRaden, Mel Tooker, and Tabatha Cooper

**July 6 update:** [This additional report provides statistics from the validation.](#)

Genomic evaluations were validated using tests developed and approved by Interbull ([Mantysaari et al., 2010](#)). Those tests are required before countries participate in genomic multitrait across-country evaluation ([Sullivan and VanRaden, 2010](#)) and before young bulls can be marketed in some European countries. The validation is very similar to previous tests applied to U.S. data ([VanRaden et al., 2009](#)). August 2006 input data and genotypes of 5,822 proven Holstein bulls and 2,200 cows with records were used to compute the predictions using the latest software, cow adjustments, and imputed dam genotypes introduced in April 2010. August 2006 predictions were also computed using genotypes of 1,616 proven Jersey bulls and 343 cows with records and 973 proven Brown Swiss bulls and 83 cows with records. Regressions and reliability gains were computed by predicting only bulls first proven in the United States after 2006. The numbers of bulls used to test the predictions were 2,038 Holsteins, 313 Jerseys, and 103 Brown Swiss. Selective genotyping of the predicted bulls reduces observed regressions and correlations as compared to true prediction ability in an unselected population (Henderson et al., 1959; [Mantysaari et al., 2010](#)). Thus, the Interbull validation compares observed regressions with expected regressions instead of with the theoretical value of 1 in an unselected population.

---

## Improvements to imputed dam genotypes

By Paul VanRaden

Imputed dams have a genomic indicator code of 3 in [format 105](#) to indicate that the cow's genotype was imputed from progeny DNA instead of directly observed from her own DNA (code 1). In April, imputed dams had a genomic indicator code of 1. Code 2 continues to indicate a nongenotyped progeny with an evaluation that was recalculated to include information from genotyped parents or imputed parents.

Technical refinements were introduced for determining genotypes of dams from their genotyped progeny. Previously, progeny haplotypes were determined by checking known parental haplotypes first and then searching the list of population haplotypes in descending frequency order. The new strategy also checks grandparental haplotypes if either parent was not genotyped. When crossovers are detected, known alleles within maternal and paternal haplotypes of the parent are used to impute any remaining unknown alleles within the progeny's haplotype. Current instead of base population allele frequencies are used for markers that could not be imputed from progeny genotypes. This change has the largest impact and increases the genetic evaluations of imputed dams by an average of \$40 for net merit, 115 pounds for milk, 5 pounds for fat, 4 pounds for protein, and 0.3 months for productive life and decreases evaluations for daughter pregnancy rate by 0.2 percentage units. Correlations between imputed dam evaluations before and after these changes are about 0.96. Correlation of genomic with traditional evaluation for milk increased to 0.87 from 0.84. Dams imputed for the first time in May or later updates were already affected by these changes, whereas evaluations of dams previously imputed in April remained official until the release of official evaluations in August.

In August, marker genotypes with one known allele and one unknown allele have the known allele used in the evaluation and the current population frequency substituted for the other allele. Previously, marker genotypes were used only if both alleles were known. Also, the two most popular haplotypes inherited by progeny are stored for each dam instead of only the first two haplotypes encountered. This improved the genotype accuracy and made the imputation resistant to incorrect pedigree data. With this greater accuracy, imputed genotypes now are used to exclude progeny that have conflicts with their dam at more than 1,000 SNP. Improvement of imputation will be continued because this step is essential for combining different marker densities.

---

## Preliminary evaluations using 3,000 markers

By George Wiggans, Paul VanRaden, and Curt Van Tassell

A 3,000 (3K) marker panel developed by [Illumina](#) (San Diego, CA) in cooperation with the [Bovine Functional Genomics Laboratory](#) (Beltsville, MD) may be available soon. Current plans are to include 3K genotypes in monthly genomic updates and distribute them to breeders as unofficial. Distribution files will include a new field to indicate that an evaluation was based on a 3K genotype. Although 3K evaluations are not included in the August official evaluation, they could become official in December pending industry review. The new programs for haplotyping and imputation implemented in April 2010 allow multiple marker sets to be included in the same evaluation. Use of 3,000 instead of 43,000 markers for young animals results in 80% of the gain in reliability on average, with slightly better results if parents are genotyped and worse if not. For example, if genomic reliability with 43,000 markers is 70% and parent average reliability is 35%, average reliability expected for a 3K evaluation is  $0.80(70\% - 35\%) + 35\% = 63\%$ . Although cows and bulls with 3K genotypes can add to the reference population, they contribute about 20% less information than do animals with genotypes based on 43,000 markers.

---

## Conception rate evaluations for males and females

*By Duane Norman, Paul VanRaden, Gary Fok, and Jana Hutchison*

Heifer conception rate (**HCR**) and cow conception rate (**CCR**) are provided for females in bytes 374–397 of [format 105](#) beginning in August. Parent averages for HCR and CCR of bulls have been added in bytes 604–615 of [format 38](#) because dam evaluations are now available. Predicted transmitting abilities for HCR and CCR were previously provided only for bulls (format 38 since January 2010 and a separate file during 2009). Neither trait is included yet in net merit or evaluated genomically, but both options are possible in the future as numbers of fertility records increase. Bulls and cows without HCR or CCR records receive multitrait predictions of HCR and CCR from correlated traits such as daughter pregnancy rate, and the two new traits are now used in multitrait prediction of productive life.

Evaluation of sire conception rate resumed following a temporary suspension in April and a new agreement among industry partners on data transfer. Only requested herds are provided and used in the evaluation. The Council on Dairy Cattle Breeding's efforts to promote a continuing supply of data for research on fertility and other important traits are appreciated.