

BUILDING BETTER GENETICS THROUGH BETTER DATA

It all starts on the farm, where everyday records become dairy data. One of the most valuable ways dairy data is used is in the calculation of genetic evaluations. Through a multi-organizational effort, herd records are collected on the farm through Dairy Herd Improvement (DHI) providers, go to a Dairy Records Processing Center (DRPC) for cleaning and merging with milk analysis data, and move to the National Cooperator Database, stewarded by the Council on Dairy Cattle Breeding (CDCB), where they are combined with pedigree, classification, and genotype information. Read on to learn from experts from each segment of the data flow process that involves milk and herd-level data about the importance of on-farm data in the creation and increased accuracy of genetic evaluations.



► On-Farm Data Collection

— Leslie Gravatt, AgSource

Why is collecting and sharing on-farm data important? How does it benefit dairy producers, regardless of herd size or operation type?

Prevention is key regardless of herd size or operation. It is much easier to correct a problem in a subset of cows versus tackling a whole herd issue. When an area needing improvement is identified, the data can be used to set goals and observe progress.

There are multiple checks and balances in place to ensure that DHI data is accurate, reliable, and consistent. First, an extra set of eyes has reviewed inputs for accuracy and made adjustments as necessary based on producer or DHI technician input. Second, the data flows through a DRPC, where industry standards require that all data can be compared equally. Third, all DHI sectors are audited annually, from the certified field technicians, to the milk meters, to the milk lab, and finally to the DRPC. If you milk cows, the value is there.

Thinking beyond the traditional milk sample analysis, what other critical information can a milk sample convey? How does this benefit producers and their consultants?

A DHI milk sample is so much more than somatic cell count (SCC), fat, and protein results. That one bottle can also provide insights into a cow's udder health, pregnancy status, transition health, and more. Additional diagnostic tests and assays are in the pipeline to further expand the knowledge held in a milk sample. Many of these services can help nutritionists, veterinarians, and other consultants track the progress their recommendations have on a herd.

What should producers and consultants know about the National Cooperator Database? What are the benefits of sharing herd data?

The National Cooperator Database acts like a data hub for the dairy industry. This data is used to help create new genetic traits and evaluate current ones. More data equals a more diverse set of data points, which is beneficial for the overall genetic improvements in cows we have seen over the years. All herds regardless of size, breed, or registration type will benefit from their contribution of data in the long run. More is always better in this case.



► Dairy Data Flow through DRPCs

— Dr. Asha Miles, Dairy Records Management Systems

Why is it important to have dairy data flow through a Dairy Records Processing Center (DRPC)?

DRPCs are providers of standardized, high-quality data. Integrating data from farms can be messy because it is recorded differently. The DRPCs put everything into one language and have the infrastructure to get data where it needs to go, efficiently and accurately. Rigorous internal quality assurance protocols help catch errors (we're all human!) and external quality control audits ensure processes are working in harmony across all DRPCs.

How does data flow from the farm through the DRPC and end up as part of a genetic evaluation?

It all starts with the DHI tech collecting milk samples for lab analysis on test day and downloading a copy of the herd management software file. The lab and herd files come to the DRPC where everything is cleaned up, standardized, and transmitted in a consistent format to CDCB. There is a two-way flow of information where the genetic evaluation comes back to the DRPC from CDCB and is incorporated into DHI reports and/or software and distributed back to the producer.



What on-farm data is used in genetic evaluations and how does other data collected fit into the bigger picture?

The type of data that goes into a genetic evaluation depends on the trait (phenotype) of interest. If the focus is yield traits, that information comes out of milk testing. Health trait information comes from on-farm herd management software and/or the DHI technician.

A phenotype is a combination of genetics and environment, so the best genetic models describe the cow's environment really well. Other data not directly included in genetic evaluations provide insights for herd trends and research and could become part of a genetic evaluation down the road.

How do consultants benefit from the work that DRPCs do?

They can trust the data, knowing that there are rigorous processes to collect, store, audit, and transmit high quality farm data. Insights through DHI reports or other web-based programs that DRPCs may offer deliver additional value. Consultants who are educated on the DRPC data stream and tools can amplify their impact.



► Benefits of Phenotypic Data

— Dr. Kristen Gaddis, Council on Dairy Cattle Breeding

How do phenotypes and genotypes differ?

A phenotype is what we can see or measure on an animal. Examples include milk yield, calving ease, growth, or mastitis occurrence. A phenotype reflects both the animal's genetics and factors like nutrition, management, housing, and health.

A genotype is the animal's genetic blueprint. In animal breeding, this typically refers to the set of genetic markers we measure using SNP chips. For dairy cattle, we currently use roughly 69,000 markers for evaluations. This genetic information is combined with the performance records (phenotypes) to generate genomic evaluations.

What impact does phenotypic data have on genomics? Why do we need phenotypes if we have genomic evaluations?

Phenotypes are the backbone of genomic evaluations because they give meaning to genetic data. Genomic information tells us how animals differ at the DNA level, but without phenotypes we cannot determine which genetic differences actually impact performance, health, or efficiency.

The more phenotypic data we have, the more accurate or reliable genomic evaluations become. Even with genomic evaluations available, we still need phenotypes to continually update, refine, and validate our predictions over time.

What is a reference population in relation to genetic evaluations?

A reference population is the group of animals for which we have both phenotypic and genomic data. This population is used to estimate the relationships between DNA markers and each trait of interest.

A larger and more diverse reference population allows these marker-trait relationships (the marker effects) to be estimated more accurately, which leads to more reliable genomic evaluations. Once these relationships are estimated, we can provide genomic predictions for animals that do not yet have their own phenotypes, such as young animals, young AI sires, or even embryos.

What benefits occur when more phenotypic data is flowing into the National Cooperator Database?

More phenotypic data coming into the National Cooperator Database strengthens the entire genetic evaluation system. When more herds contribute records, genomic predictions become more accurate, especially for traits that often have limited data. Additional data also improves how evaluations account for different management styles, climates, or housing. This means predictions will work better for a wider range of herds.

Many new traits added in recent years – like health traits, milking speed, and calf livability – were only possible because enough farms were recording those events and forwarding that data to the National Cooperator Database. With more phenotypes, we can continue developing new traits and improving the ones we currently have.

Ultimately, more phenotypic data results in faster genetic progress toward a more productive and resilient dairy animal.



View Here

Learn more about the National Cooperator Database



How can you know when the genetic evaluations you're using to make a selection decision come from this integrated and quality-controlled system? Look for the Powered by CDCB mark. This logo indicates that genetic information is derived from the objective, farmer-owned data stored in the National Cooperator Database and calculated by an independent organization.