

## Including feed intake data from U.S. Holsteins in genomic prediction

P.M. VanRaden<sup>1</sup>, J.R. O'Connell<sup>2</sup>, E.E. Connor<sup>1</sup>, M.J. VandeHaar<sup>3</sup>, R.J. Tempelman<sup>3</sup> & K.A. Weigel<sup>4</sup>

<sup>1</sup> U.S. Department of Agriculture (USDA), Agricultural Research Service, Animal Genomics and Improvement Laboratory, 10300 Baltimore Ave., Bldg.005, Room 306, BARC-West, Beltsville, Maryland 20705, USA

paul.vanraden@ars.usda.gov (Corresponding Author)

<sup>2</sup> University of Maryland - Baltimore, 655 W Baltimore S, Baltimore, Maryland 21201, USA

<sup>3</sup> Michigan State University, Department of Animal Science, 1290 Anthony Hall, East Lansing, Michigan 48824, USA

<sup>4</sup> University of Wisconsin, Department of Dairy Science, 1675 Observatory Dr., Madison, Wisconsin 53706, USA

### Summary

Predicted transmitting abilities (PTAs) for residual feed intake (RFI) were computed using data from 4,823 feed intake trials of 3,965 U.S. Holsteins born 1999 to 2013 in 9 research herds. The RFI averages were already adjusted to remove phenotypic correlations with milk energy output, metabolic body weight and body weight change and for several environmental effects, including other nutrition experiments during the feed intake trials. Traditional PTAs for RFI of 74 million Holsteins were then estimated by an animal model that also included effects for age-parity group, trial date, herd management group, permanent environment, herd-sire interaction and regressions on inbreeding and on genomic evaluations for milk energy and body weight composite (BWC). The milk energy and BWC terms were included to remove positive genetic correlations that remained after phenotypic correlations were removed. Estimated heritability was 0.14; repeatability across lactations was 0.24. Genomic PTAs for RFI included 60,671 genetic markers for 1.6 million Holsteins; genomic reliabilities calculated for elite young animals averaged 12% compared with 3% for traditional reliabilities. Economic value of RFI is very large, and RFI could receive 16% of total emphasis in the net merit index; however, its low reliability will limit extra genetic progress to about 1% more than current progress. One option for publishing the trait is to combine the benefits from reduced RFI and smaller BWC into feed saved per lactation. Additional records could make feed intake an important trait in future selection indexes for dairy cattle.

*Keywords: feed efficiency, residual feed intake, genomic evaluation, selection index*

### Introduction

Feed is the largest single expense on dairy farms. Individual intakes are rarely measured because most cows are fed free-choice, total mixed rations. However, designed experiments to collect intake data now allow genomic prediction and direct instead of indirect selection for expense traits (Wallén *et al.*, 2017). A goal of the current research is to predict feed intakes for millions of Holsteins from the national data collected by VandeHaar *et al.* (2017). Other goals are to estimate genetic parameters, reliabilities, and economic values, to compare models, and to define the trait or traits to be published officially.

## Materials and methods

Feed intake data from 4,823 lactations of 3,965 Holstein cows had previously been collected as part of a 5-year grant (VandeHaar *et al.*, 2017). The cows were from 9 research herds in 8 states in the central and eastern USA and were born from 1999 through 2013 (Table 1). Previous studies with these data (e.g., Tempelman *et al.*, 2015; Manzanilla-Pech *et al.*, 2016; Yao *et al.*, 2017) also included foreign cows, but the current research focuses on domestic records because those were easiest to integrate into the national database.

Table 1. Cows measured for feed intake by research herd.

Research herd	Dairy Herd Information code	Records	Cows
Iowa State University	42770074	1,006	953
University of Wisconsin	35130320	1,056	916
USDA, Beltsville Agricultural Research Center (Maryland)	51019026	834	534
University of Florida	58010029	582	491
USDA, U.S. Dairy Forage Research Center (Wisconsin)	35570803	622	474
Michigan State University	34330002	315	273
Purina Animal Nutrition Center (Missouri)	43330055	184	151
Virginia Tech University	52606014	93	93
Miner Institute (New York)	21090270	58	58
<b>All herds</b>		<b>4,823</b>	<b>3,965</b>

Correlated effects of milk yield ( $r = 0.7$ ), metabolic body weight ( $r = 0.4$ ), body weight change and several environmental effects were removed from dry matter intake to obtain residual feed intake (RFI). Nearly all records were from 6-week trials conducted during the first 50-250 days of lactation, but 202 were from 4-week trials. Those were given less weight (0.92) in the evaluation model because the standard deviation was higher (1.75 vs. 1.68 kg/day) and the correlation of 4- and 6-week trials was 0.96. The mean RFI was 0 by definition and the phenotypic standard deviation (SD) was 512 kg/lactation, estimated simply as 305 times the average daily SD. This may slightly overestimate the true SD because intakes are higher in mid-lactation. Lactation somatic cell score (SCS) was also examined as a second, independent trait from these same herds and cows as a method to calibrate reliability by comparing predictions using only research cows vs. national data.

The multi-step evaluation first used a pedigree relationship matrix for >60 million Holsteins and modelled RFI with effects for breeding value, permanent environment, herd-sire interaction, management group, age-parity and regressions on inbreeding and genomic predictions for milk net energy and body weight composite (BWC). The regressions on genomic predictions removed remaining genetic correlations not completely removed by the phenotypic regressions. Similar math to simultaneously model genetic and environmental correlations was proposed by Lu *et al.* (2015; 2017) and earlier researchers. The current genomic regressions used full national data for milk energy and BWC and therefore should estimate genetic effects more precisely than using only the research cow data for those traits.

The genomic evaluation used deregressed evaluations from the traditional evaluation to predict RFI for 1.6 million genotyped Holsteins. Genotypes for the research cows were from various Illumina chips including 502 high density (777k), 1,341 Geneseek high density versions 1 and 2 (77k and 140k), 1,251 50k or Zoetis medium density, and 411 lower density

genotypes (7k to 20k). All genotypes were imputed to the standard set of 60,671 markers used in evaluations by the Council on Dairy Cattle Breeding (CDCB; Bowie, Maryland, USA). Variance components were estimated by MMAP (O'Connell, 2017) using a repeatability model and either pedigree or genomic relationships among the cows.

Genetic rankings for RFI could be published in several ways, and U.S. producers may prefer different values than scientists. Feed saved (FS) combines the reductions in feed eaten associated with RFI and smaller body weight on a per-lactation basis as in Pryce *et al.* (2015):

$$FS \text{ lb/lactation} = -305(\text{RFI lb/day}) - 1.67(40)\text{BWC},$$

where 305 = days per lactation, 1.67 = pounds of feed eaten per lactation per pound of body weight and 40 = regression of pounds of body weight on BWC. The lifetime value for BWC currently included in the lifetime net merit (NM\$) index could be replaced by the lifetime U.S. dollar value of feed saved:

$$FS \text{ \$/lifetime} = -\$0.12(2.8 \text{ lactations})(FS \text{ lb/lactation}).$$

Holstein Association USA (Brattleboro, Vermont, USA) currently publishes an economic estimate of feed efficiency (FE) that combines the net income expected from higher milk yield and net loss expected from larger BWC:

$$FE = \text{MFP\$} - \$0.12/\text{lb}(2.8 \text{ lactations})(1.67)(40)\text{BWC},$$

where MFP\$ is the lifetime value of milk, fat and protein after subtracting correlated feed costs. A more complete measure (FeedEff) can be provided by subtracting the lifetime cost of RFI (RFI\$) from FE to account for actual feed consumed:

$$\text{FeedEff} = FE - \text{RFI\$},$$

$$\text{where RFI\$} = -\$0.12/\text{lb}(2.8 \text{ lactations})(305 \text{ days/lactation})(\text{RFI lb/day}).$$

The economic values are easy to understand but not stable across time because of price changes. Reporting dry matter intake per lactation or lifetime instead of per day may be preferred because the units are larger and more similar to other traits such as milk yield per lactation.

Economic progress is greatly affected by genetic variance of traits and reliability (REL) of predictions. Because RFI is independent of NM\$, the ratio of progress from the proposed index compared with current NM\$ is

$$\sqrt{\frac{\text{REL}_{\text{NM\$}} \text{Var}(\text{TA}_{\text{NM\$}}) + \text{REL}_{\text{RFI}} \text{Var}(\text{TA}_{\text{RFI\$}})}{\text{REL}_{\text{NM\$}} \text{Var}(\text{TA}_{\text{NM\$}})}},$$

where TA is transmitting ability. Similarly, SD of true  $\text{TA}_{\text{FS}}$  is  $\sqrt{209^2 + 67^2} = 219$ , where 209<sup>2</sup> and 67<sup>2</sup> are variances of RFI and of feed associated with BWC, respectively, and SD of  $\text{PTA}_{\text{FS}}$  is  $\sqrt{\text{REL}_{\text{RFI}}(209^2) + \text{REL}_{\text{BWC}}(67^2) / 219^2}$ .

## Results

The estimated heritability ( $h^2$ ) for RFI was 14% with either a pedigree or genomic model and changed very little with or without removing the milk and BWC genetic regressions. This  $h^2$  was slightly lower than estimates from previous U.S. data subsets of 16% (Lu *et al.*, 2015) and 18% (Tempelman *et al.*, 2015) and equal to the 14% when correlated traits were removed by genetic regressions (Lu *et al.*, 2015). Estimated permanent environmental variance was 10%. When repeated records were included but the model excluded permanent environmental effects,  $h^2$  was 21%, which nearly matched the  $h^2$  of 22% reported by Manzanilla-Pech *et al.* (2016) and demonstrated that not accounting for permanent environment can inflate  $h^2$ .

The genetic regression of RFI on milk energy was extremely significant ( $P < 0.0000001$ ), whereas the BWC regression was small and not significant. This finding agrees with the genetic correlation of 0.40 between RFI and milk energy reported by Manzanilla-Pech *et al.* (2016), but their genetic correlation of 0.26 between RFI and body weight was larger than our regressions indicated. The regressions were included to make the traits genetically independent; without the regressions in the model, some of the top bulls for RFI were also top bulls for milk energy.

The expected genomic REL was 19% for both SCS and RFI based on the reference population size, but the parallel test using SCS records for the same 3,965 cows indicated lower observed REL. The SCS genomic predictions computed from the research cows were correlated by only 0.39 to predictions from national data. This translated to an observed  $REL_{SCS}$  of only  $0.39^2(72\%) = 11\%$  for SCS. For RFI, the SD of true TA was 209 lb/lactation with  $h^2$  of 14%, and the SD of genomic PTA for young animals was 38 lb/lactation, much less variance than the 92 lb/lactation =  $209 \sqrt{0.19}$  expected from the theoretical  $REL_{RFI}$  and also indicating lower observed  $REL_{RFI}$ . The variance component estimates for SCS gave  $h^2$  of 15% and repeatability of 31% from the research cows, and were similar to the  $h^2$  of 12% and repeatability of 35% estimated previously from national data and used in official U.S. evaluations.

The lifetime economic value is \$0.336 for RFI per lactation, assuming a feed price of \$0.12/lb for dry matter and an average of 2.8 lactations; correlated responses in heifer and dry cow RFI are yet to be determined. The lifetime TA has  $SD = 209(\$0.336) = \$70$ . The SD of true TA for current NM\$ is \$193 (VanRaden, 2017), and the SD of true TA in the proposed NM\$ index is  $\sqrt{193^2 + 70^2} = \$205$ . The sum of economic values times SD for all current traits in NM\$ is 376, and the proposed NM\$ would have an additional 70 for RFI. Therefore, the proposed relative values would be -16% for RFI (computed as  $-70/[376 + 70]$ ), 15% for protein, 20% for fat, -1% for milk yield, 11% for productive life, 6% for cow livability, -6% for SCS, 6% for udder composite, 2% for feet/legs, -5% for BWC, 6% for daughter pregnancy rate, 1% for cow conception rate, 1% for heifer conception rate, and 4% for calving traits.

Few daughters of any bull are measured for RFI; therefore, all bulls have low traditional  $REL_{RFI}$ . Focusing RFI data collection on daughters of elite bulls that will have greatest genetic contribution to the next generation will lead to more gains and profit. Cottle and van der Werf (2017) reached similar conclusions regarding RFI measurements for beef bulls. Currently genomic  $REL_{NM\$}$  averages 75% for young bulls and 91% for progeny-tested bulls, whereas genomic  $REL_{RFI}$  will be much lower because of less data, perhaps 12%. With nearly all selection on young animals, substitution of  $REL_{NM\$} = 0.75$ ,  $REL_{RFI} = 0.12$ ,  $Var(TA_{NM\$}) = 193^2$ , and  $Var(TA_{RFIS}) = 70^2$  into the formula for ratio of progress for lifetime profit gives only 1.01 times or 1% faster progress than currently. Assuming  $REL_{RFI}$  for young

animals equals 12% and  $REL_{BWC}$  equals 72%, the combined  $REL_{FS}$  is about 18% because true variance for RFI is much larger than the variance of feed intake correlated with BWC. These assumptions are preliminary and may change with further research.

The economic benefit from adding a new trait such as RFI can be large even if changes in NM\$ appear small. Current genetic gain is about \$50 for NM\$ breeding value / year. The total value of NM\$ gain per year to US producers is  $\$50 * 9$  million cows = \$450 million currently, and a 1% improvement in the rate of progress is worth \$4.5 million to the US dairy industry. Artificial insemination (AI) companies exported \$157 million last year and may sell about \$200 million of semen in the US market. If NM\$ gain is 1% faster by including RFI, they might sell 1% more (or without RFI they might sell 1% less), increasing their sales by \$3.6 million for a proposed investment of \$0.5 million to continue collecting data.

Increasing  $REL_{RFI}$  will require collecting more US feed intake data or exchanging data internationally. If RFI gets 16% emphasis in NM\$ and has a genetic SD of \$70 (compared with \$193 for current NM\$), then the maximum  $REL_{NM\$}$  without feed intake data is  $193^2 / (193^2 + 70^2) = 0.88$ . Currently many bulls are reported with 99%  $REL_{NM\$}$ , but really none should have  $REL_{NM\$}$  above 88% unless actual feed intake is included. When other new traits were added, often many bulls had 99% REL, but only a few famous bulls are above 80%  $REL_{RFI}$ . Most bulls have much lower  $REL_{RFI}$  and may serve as a reminder that more feed intake records are needed. When important new traits are added, the resulting declines in  $REL_{NM\$}$  expose previous ignorance about those traits.

## **Conclusions**

Direct selection on feed intake data can improve dairy cattle profitability. In NM\$, RFI could receive 16% of total emphasis, but the low estimated REL of 12% for RFI will limit overall progress to about 1% faster than currently achieved. Genomic prediction can multiply information from a few herds to thousands of others and millions of animals.

## **Acknowledgments**

The authors thank the researchers from all contributing herds. Feed intake data collection was funded by Agriculture and Food Research Initiative Competitive Grant #2011-68004-30340 from the USDA National Institute of Food and Agriculture. Jim Liesman (Michigan State University) and Janice Wright (USDA) performed many of the initial edit and evaluation steps. The CDCB and its data suppliers provided pedigrees, SCS phenotypes, and genotypes, and George Wiggans (CDCB) managed the genotype transfers.

## **List of References**

- Cottle, D.J. & J.H.J. van der Werf. 2017. Optimising the proportion of selection candidates measured for feed intake for a beef cattle breeding objective that includes methane emissions. *J. Anim. Sci.* 95:1030–1041.
- Lu, Y., M.J. VanDehaar, D.M. Spurlock, K.A. Weigel, L.E. Armentano, C.R. Staples, E.E. Connor, Z. Wang, N.M. Bello & R.J. Tempelman 2015. An alternative approach to modeling genetic merit of feed efficiency in dairy cattle. *J. Dairy Sci.* 98:6535–6551.
- Lu, Y., M.J. VanDehaar, D.M. Spurlock, K.A. Weigel, L.E. Armentano, C.R. Staples, E.E. Connor, Z. Wang, M. Coffey, R.F. Veerkamp, Y. de Haas & R.J. Tempelman. 2017.

- Modeling genetic and nongenetic variation of feed efficiency and its partial relationships between component traits as a function of management and environmental factors. *J. Dairy Sci.* 100:412–427.
- Manzanilla-Pech, C.I.V., R.F. Veerkamp, R.J. Tempelman, M.L. van Pelt, K.A. Weigel, M. VandeHaar, T.J. Lawlor, D.M. Spurlock, L.E. Armentano, E.E. Connor, C.R. Staples, M. Hanigan & Y. De Haas. 2016. Genetic parameters between feed-intake-related traits and conformation in 2 separate dairy populations—the Netherlands and United States. *J. Dairy Sci.* 99:443–457.
- O’Connell, J.R. 2017. MMAP: Mixed model analysis for pedigrees and populations. <https://mmap.github.io/>. Accessed Aug. 1, 2017.
- Pryce, J.E., O. Gonzalez-Recio, G. Nieuwhof, W.J. Wales, M.P. Coffey, B.J. Hayes & M.E. Goddard. 2015. Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. *J. Dairy Sci.* 98:7340–7350.
- Tempelman, R.J., D.M. Spurlock, M. Coffey, R.F. Veerkamp, L.E. Armentano, K.A. Weigel, Y. de Haas, C.R. Staples, E.E. Connor, Y. Lu & M.J. VandeHaar. 2015. Heterogeneity in genetic and nongenetic variation and energy sink relationships for residual feed intake across research stations and countries. *J. Dairy Sci.* 98:2013–2026.
- VandeHaar, M.J., L.E. Armentano, K. Weigel, D.M. Spurlock, R.J. Tempelman & R. Veerkamp. 2017. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. *J. Dairy Sci.* 99:4941–4954.
- VanRaden, P.M. 2017. Net merit as a measure of lifetime profit: 2017 revision. AIP Res. Rep. NM\$6 (2-17). <https://aipl.arsusda.gov/reference/nmcalc-2017.htm>.
- Wallén, S.E., M. Lillehammer & T.H.E. Meuwissen. 2017. Strategies for implementing genomic selection for feed efficiency in dairy cattle breeding schemes. *J. Dairy Sci.* 100:6327–6336.
- Yao, C., G. de los Campos, M.J. VandeHaar, D.M. Spurlock, L.E. Armentano, M. Coffey, Y. de Haas, R.F. Veerkamp, C.R. Staples, E.E. Connor, Z. Wang, M.D. Hanigan, R.J. Tempelman & K.A. Weigel. 2017. Use of genotype × environment interaction model to accommodate genetic heterogeneity for residual feed intake, dry matter intake, net energy in milk, and metabolic body weight in dairy cattle. *J. Dairy Sci.* 100:2007–2016.