Selective breeding has been practiced since domestication, but early breeders commonly selected on appearance (e.g., coat color) rather than performance traits (e.g., milk yield). A breeding index converts information about several traits into a single number used for selection and to predict an animal’s own performance. Calculation of selection indices is straightforward when phenotype and pedigree data are available. Prediction of economic values 3 to 10 yr in the future, when the offspring of matings planned using the index will be lactating, is more challenging. The first USDA selection index included only milk and fat yield, whereas the latest version of the lifetime net merit index includes 13 traits and composites (weighted averages of other additional traits). Selection indices are revised to reflect improved knowledge of biology, new sources of data, and changing economic conditions. Single-trait selection often suffers from antagonistic correlations with traits not in the selection objective. Multiple-trait selection avoids those problems at the cost of less-than-maximal progress for individual traits. How many and which traits to include is not simple to determine because traits are not independent. Many countries use indices that reflect the needs of different producers in different environments. Although the emphasis placed on trait groups differs, most indices include yield, fertility, health, and type traits. Addition of milk composition, feed intake, and other traits is possible, but they are more costly to collect and many are not yet directly rewarded in the marketplace, such as with incentives from milk processing plants. As the number of traits grows, custom selection indices can more closely match genotypes to the environments in which they will perform. Traditional selection required recording lots of cows across many farms, but genomic selection favors collecting more detailed information from cooperating farms. A similar strategy may be useful in less developed countries. Recording important new traits on a fraction of cows can quickly benefit the whole population through genomics.

**Key words:** breeding program, genetic improvement, selection index

**INTRODUCTION**

Breeding indices are important tools in modern dairy cattle breeding. They provide a way to combine information about many traits into a single number that can be used to rank animals and make breeding decisions. The need for such a tool was recognized very early in the history of modern animal breeding, when Hazel and Lush (1942) applied the method of Smith (1934) to the improvement of economically important traits of livestock. The ideal breeding objective for dairy cattle remains a popular topic and has been reviewed periodically (e.g., Hazel et al., 1994; Philipsson et al., 1994; VanRaden, 2004; Miglior et al., 2005; Shook, 2006), but there is no single selection objective that is best for all populations or all herds within a population.

Historically, selection indices in the United States were developed by the USDA and purebred dairy cattle associations, frequently with input from scientists at land-grant universities, using data available through the national milk recording system and breed type appraisal programs. Proposed indices were typically reviewed by groups of experts and information about the derivation of the indices was published in technical and trade publications, ensuring confidence in the values because of that review process. Recently, genetic evaluations for novel traits and new selection indices have been computed and distributed by companies such as CRV (Arnhem, the Netherlands), Genex (Shawano, WI), and Zoetis (Parsippany-Troy Hills, NJ). This provides farmers with new tools and may drive demand for new phenotypes, but transparent review processes may be lacking. The purpose of this paper is to present a brief overview of how selection indices are constructed, describe traits included in current indices, review desirable properties of new traits, discuss traits that may be included in selection indices in the future, and dem-
onstrate that selection indices are robust to incorrect assumptions about model parameters.

**SELECTION INDICES**

**Improving Animal Performance**

Animal performance is a function of both genetic and environmental factors and interactions among the two. Predictions of genetic merit are based on a quantitative model that assumes that traits are controlled by many genes, each of which has a small effect on the phenotype (Falconer and MacKay, 1996). This model has been found to accurately describe many traits of economic importance in dairy cattle (Cole et al., 2009). Environmental influences include all sources of phenotypic variation that cannot be attributed to genetics, such as nutrition, climate, disease exposure, error in measurement, and other unknown factors. These factors vary from farm to farm and between individual animals on the same farm and may change over time (e.g., Windig et al., 2005).

Figure 1 shows the change in fat yield for US Holsteins between 1957 and 2015. Production in 1957 is used as a baseline, and gains over time were found to be evenly divided between increased genetic potential and improvements in feeding and management. Gains in genetics and management each represent 28% of 2015 production, whereas the 1957 base represents 44% of current yield. The proportion of gains from improved genetics versus improved environment differs from trait to trait and is a function of the heritability of a trait. Fat yield has a heritability of 20% (VanRaden, 2017), whereas daughter pregnancy rate has a heritability of only 4% (VanRaden et al., 2004). When the proportion of variance in a trait due to genetics is low, it is often easier to make gains by improving the environment in which the cow is performing, and gains from genetic improvement may not be visible to producers for a long time.

**Construction of Selection Indices**

The following discussion focuses on the simplest formulation of a selection index; greater detail, including derivations, may be found in the literature (e.g., Lin, 1978; Cameron, 1997). When using a selection index, the goal is to improve one or more traits, referred to as the selection objective, by ranking and choosing mates using a combination of one or more traits, known as the selection criterion. In modern breeding programs, the selection objective is typically a measure of lifetime profitability, whereas the selection criterion usually comprises traits that are included in national milk recording programs. In the mathematical terms of Hazel and Lush (1942), an index including \( m \) terms in the selection criterion for an animal takes the form

\[
I = b_1X_1 + b_2X_2 + \ldots + b_mX_m,
\]

where \( I \) is the selection criterion, \( b_i \) is the emphasis placed on the \( i \)th trait, and \( X_i \) is the animal’s pheno-
type for the ith trait in the index. Index weights are calculated as a function of (co)variances among the traits in the objective and the criterion and the economic weights of the individual traits:

$$ b = P^{-1}Ga, $$

where $b$ is a vector of index weights, $P$ is the phenotypic (co)variance matrix for the traits in the selection criterion, $G$ is a matrix of genetic (co)variances among the traits in the criterion and the objective, and $a$ is a vector of economic weights associated with the traits in the criterion. If all of the parameters used to compute the index are correct, then it is the most efficient way of improving all of the traits in the selection objective. However, in modern breeding programs, mixed model equations include $P$ and $G$ to first obtain multitrait evaluations ($\tilde{u}$), and those are combined directly by their economic values as $a'\tilde{u}$.

When the traits in the selection criterion and selection objective differ, as is often the case, an additional calculation is necessary to determine the correlated response to selection of the traits in the objective in response to selection on the traits in the criterion. This is a straightforward extension of the well-known breeder’s equation (Cameron, 1997)

$$ \Delta g_j = \frac{b'G_j}{\sqrt{b'Pb}}. $$

where $\Delta g_j$ is the correlated response of trait $j$ in the selection objective in response to selection on the selection criterion, and $G_j$ is the correlation between trait $j$ and the traits in the selection criterion. This equation shows that the correlated response is a function of the genetic correlations among the traits in the objective and the criterion and the index weights.

The literature on selection index methodology is quite extensive, and many special cases can be accommodated. For example, one trait can be held at a constant level while others are changed (Kempthorne and Nordskog, 1959), economic value can have nonlinear relationships with the traits in the index (Goddard, 1983), selection can proceed in stages where objectives change over time (Cunningham, 1975), and quota systems can drive the economic value of yield traits (Gibson, 1989). Selection index methodology also has been used to determine rates of genetic and economic gain under genomic selection programs in a deterministic fashion (Dekkers, 2007; König et al., 2009). Readers are directed to more comprehensive works on selection index methodology for additional details (e.g., Van Vleck, 1993; Weller, 1994; Cameron, 1997).

**Contribution of Genomic Information**

Genomic selection allows breeders to make decisions more quickly by using dense DNA marker information to compute high-reliability predictions of genetic merit early in an animal’s life (Nejati-Javaremi et al., 1997; Meuwissen et al., 2001). From the perspective of the breeding objective, the principal effect of genomic selection is on the reliabilities of the breeding values used in the index (VanRaden et al., 2009), but the technology provides several other advantages, including lower costs of proving bulls (Schaeffer, 2006), greater rates of genetic gain from shorter generation intervals (García-Ruiz et al., 2016), detection of previously unknown genetic disorders (VanRaden et al., 2011), and identification of genes that influence economically important traits (Cole et al., 2011). A trait with a low heritability, such as daughter pregnancy rate ($h^2 = 0.04$), requires more daughter phenotypes to produce a breeding value with the same reliability as a trait with higher heritability, such as fat yield ($h^2 = 0.30$), and genotypes provide more information for low-heritability traits. Pedigree information alone is equivalent to approximately 7 daughter records, whereas a genotype is worth 34 daughter records for fat or 131 daughter records for daughter pregnancy rate. Genomics allows us to publish useable evaluations much sooner than in the past and make more profitable management decisions on the farm (e.g., Pryce and Hayes, 2012; Van Eenennaam et al., 2014).

**Selection for Many Traits**

The number of traits included in a typical selection criterion has grown over time, from 1 or 2 yield traits to many nonyield traits, including fertility, health, and fitness traits. This allows farmers to make use of more information than in the past and takes advantage of correlations among traits (important traits rarely have correlations of 0 with other important traits). Many traits may have direct economic value; for example, milk plants often pay premiums for low SCS in addition to payments for high protein and fat components. Traits can also have indirect value; for example, SCS can predict mastitis losses if mastitis is not recorded directly. Substantial losses can occur when indirect values are ignored—for example, the well-documented negative correlation of fertility with milk yield (Figure 2; Lucy, 2001). Balanced selection improves traits according to their economic values, and selection indices should be periodically updated to include new traits and reflect changing economic conditions as well as changing genetic parameters between and among traits. However, as traits are added to an index it becomes increasingly
difficult to predict a priori whether the new index will have greater or reduced response compared with the index with fewer traits (Sivanadian and Smith, 1997).

**Derivation of Economic Values**

The vector of economic values ($\mathbf{a}$) included in the calculation of index weights is used to assign values to traits based on their importance to the selection objective. Two general approaches may be used to derive those weights. The first, which might be called the empirical approach, uses data from scientific studies and field reports to quantify incomes and expenses associated with the traits in the selection objective and criterion. The goal of this approach is to allow the best available economic information to drive the formulation of the index, and it is used in the calculation of the USDA’s Lifetime Net Merit Index (\textit{NM$}}) and some breed-specific indices, such as the American Jersey Cattle Association’s (2017) Jersey Performance Index. The second, which might be called the subjective approach, has been used to construct indices such as Holstein Association USA Inc.’s (2017) Total Performance Index (\textit{TPI}), assigns values to traits based on the cow that breeders would like to see in the future. Those targets for breed improvement are developed by groups of breeders and experts and are driven by both quantitative and qualitative factors. Quantitative factors include incomes and expenses associated with costs of raising animals and the value of products sold, whereas qualitative factors include such things as the desirable conformation for cows of a particular breed. Direct economic values for some traits, most notably conformation traits, often are difficult to calculate but may be very important to farmers who breed and own registered cattle. Both approaches to placing values on individual traits produce broadly similar results (2010 \textit{NM$}} and \textit{TPI} had a correlation of 0.88), but the differences between the indices reflect important economic factors affecting the users. Customized indices at the farm level were first delivered by McGilliard and Clay (1983) and proposed in Australia (Bowman et al., 1996) but were not widely used in the United States. As herds continue to grow larger, managers may have an incentive to customize their own indices (Dickrell, 2017).

**Subindices**

One way to make indices easier to understand is to construct them from a series of subindices. For example, \textit{NM$}} includes 3 type composites that combine information from several traits, and the calving ability dollars (\textit{CA$}) subindex combines sire and daughter calving ease and sire and daughter stillbirth into a single quantity. All the breeder will see when the index is revised are the changes in emphasis on each of the subindices rather than changes to each of the individual traits (Figure 3). Farmers need to understand only the function of each subindex instead of dozens of traits. The Ideal Commercial Cow Index (\textit{ICC$}; Genex, 2006) is constructed in this way: \textit{ICC$} is the sum of the production efficiency (\textit{PREF$}), health (\textit{HLTH$}), fertil-
ity and fitness (FYFT$), milking ability (MABL$), and calving ability (CABL$) subindices. The advantages of this approach are small when indices contain only a few traits but increase rapidly as the number of traits included grows. Composite traits have a similar purpose but often are unitless instead of having monetary value.

The Irish EBI Index (ICBF, 2017) comprises 7 subindices: milk production, fertility, calving performance, beef carcass, cow maintenance, cow management, and health. The calving performance subindex receives 10% of the total emphasis and includes PTA for direct and maternal dystocia, gestation length, and stillbirth. The health subindex, with 4% of the emphasis, includes direct (clinical mastitis) and indirect (SCC) measures of udder health as well as lameness. These examples demonstrate the use of direct (e.g., dystocia, clinical mastitis) traits in combination with indirect (e.g., gestation length, SCC) indirect (indicator) traits.

PHENOTYPES IN SELECTION INDICES

What Traits Are Included in Current Selection Indices?

The traits included in USDA selection indices over time, and weights placed on each, are shown in Table 1. The first USDA index, Predicted Difference Dollars (PD$), included only milk and fat yield in the selection criterion, whereas the 2017 revision of NM$ (Van-Raden, 2017) includes information about 33 different traits when subindices are considered. Selection indices differ within and across countries because economic conditions, traits recorded, and breeds used are not the same everywhere. Figure 4 shows traits included in total merit indices from 15 different countries. Trait definitions may differ slightly from one country to another, but common trait groups include yield (e.g., milk volume, fat and protein yield), longevity (e.g., productive life), fertility (e.g., nonreturn rate, days open), udder health (e.g., SCS, clinical mastitis), calving traits (e.g., dystocia, stillbirth), milking traits (e.g., milking speed), and conformation (e.g., udder conformation, feet and leg score). Although some broad similarities exist among indices—most include direct emphasis on protein yield—no two are the same, even within a country. For example, NM$ includes more emphasis on longevity and less on conformation than TPI. Some countries were more far-sighted than others and added health traits to their selection programs decades ago, providing them with a head start over other countries (e.g., Philipsson and Lindhe, 2003).

What Traits Should Be Included in Future Selection Indices?

New traits are added to selection indices for many reasons. Production economics change over time, such
Table 1. Traits included in USDA selection indices and the relative emphasis placed on each, 1971 to 2017

<table>
<thead>
<tr>
<th>Trait</th>
<th>PD$</th>
<th>MFP$</th>
<th>CY$</th>
<th>NM$</th>
<th>NM$</th>
<th>NM$</th>
<th>NM$</th>
<th>NM$</th>
<th>NM$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>52</td>
<td>27</td>
<td>−2</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>−1</td>
</tr>
<tr>
<td>Fat</td>
<td>48</td>
<td>46</td>
<td>45</td>
<td>25</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>19</td>
<td>22</td>
</tr>
<tr>
<td>Protein</td>
<td>27</td>
<td>53</td>
<td>43</td>
<td>36</td>
<td>33</td>
<td>23</td>
<td>16</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>PL</td>
<td>20</td>
<td>14</td>
<td>11</td>
<td>17</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>SCS</td>
<td>−6</td>
<td>−9</td>
<td>−9</td>
<td>−9</td>
<td>−10</td>
<td>−7</td>
<td>−7</td>
<td>−10</td>
<td>−7</td>
</tr>
<tr>
<td>UC</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>FLC</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>BWC</td>
<td>−4</td>
<td>−3</td>
<td>−4</td>
<td>−6</td>
<td>−5</td>
<td>−6</td>
<td>−5</td>
<td>−6</td>
<td>−6</td>
</tr>
<tr>
<td>DPR</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>SCE</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
</tr>
<tr>
<td>CA$</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>HCR</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

1PL = productive life; UC = udder composite; FLC = foot and leg composite; BWC = BW composite; DPR = daughter pregnancy rate; SCE = sire calving ease; DCE = daughter calving ease; CA$ = calving ability dollars; HCR = heifer conception rate; CCR = cow conception rate; LIV = cow livability.

2PD$ = predicted difference dollars; MFP$ = milk, fat, and protein dollars; CY$ = cheese yield dollars; NM$ = net merit dollars.
as the introduction of incentive payments for milk quality or the elimination of quota systems, with a corresponding need for adjustments to selection objectives. Our understanding of biology improves over time, which can lead to the adoption of new traits (e.g., Shook and Schutz, 1994). Technology also evolves, permitting the collection of information that was previously impossible or prohibitively expensive to record (e.g., De Marchi et al., 2014). The widespread adoption of genomic selection is complementary to those technologies because new traits can be predicted on all genotyped animals without the need to collect progeny records, and phenotyping costs are shared among millions of animals. The following discussion will briefly consider some traits that are of growing interest to dairy farmers. Recent comprehensive discussions of new traits and phenotyping strategies are provided by Boichard and Brochard (2012), Egger-Danner et al. (2015), Gengler et al. (2016), and (Pryce et al., 2016).

**Health and Fitness.** Some countries have included health traits in their selection indices for decades (Miglior et al., 2005; Heringstad and Østerås, 2013), but many have not, and there is growing interest in the use of genetic selection to improve cow health and welfare (Pryce et al., 2016). There also is increasing pressure from consumers and regulatory agencies to reduce the use of drugs and increase the perceived welfare of food animals (Jensen, 2016; Saitone and Sexton, 2017). Sick cows are less profitable than healthy cows due to lower production, decreased fertility, and increased labor and veterinary costs. They are also more likely to die on the farm, which results in lost revenue from beef sales and incurs disposal costs that can be evaluated separately (Wright and VanRaaden, et al., 2016).

Several studies have shown that producer-recorded health events from on-farm computer systems are a rich source of data for genetic improvement (Zwald et al., 2004; Parker Gaddis et al., 2012; Wenz and Giebel, 2012), and genomic information produces evaluations with sufficient reliability for routine use (Pryce et al., 2016). There is increasing pressure from consumers and regulatory agencies to reduce the use of drugs and increase the perceived welfare of food animals (Jensen, 2016; Saitone and Sexton, 2017). Sick cows are less profitable than healthy cows due to lower production, decreased fertility, and increased labor and veterinary costs. They are also more likely to die on the farm, which results in lost revenue from beef sales and incurs disposal costs that can be evaluated separately (Wright and VanRaaden, et al., 2016).

Several studies have shown that producer-recorded health events from on-farm computer systems are a rich source of data for genetic improvement (Zwald et al., 2004; Parker Gaddis et al., 2012; Wenz and Giebel, 2012), and genomic information produces evaluations with sufficient reliability for routine use (Pryce et al., 2016). There is increasing pressure from consumers and regulatory agencies to reduce the use of drugs and increase the perceived welfare of food animals (Jensen, 2016; Saitone and Sexton, 2017). Sick cows are less profitable than healthy cows due to lower production, decreased fertility, and increased labor and veterinary costs. They are also more likely to die on the farm, which results in lost revenue from beef sales and incurs disposal costs that can be evaluated separately (Wright and VanRaaden, et al., 2016).

Several studies have shown that producer-recorded health events from on-farm computer systems are a rich source of data for genetic improvement (Zwald et al., 2004; Parker Gaddis et al., 2012; Wenz and Giebel, 2012), and genomic information produces evaluations with sufficient reliability for routine use (Pryce et al., 2016). There is increasing pressure from consumers and regulatory agencies to reduce the use of drugs and increase the perceived welfare of food animals (Jensen, 2016; Saitone and Sexton, 2017). Sick cows are less profitable than healthy cows due to lower production, decreased fertility, and increased labor and veterinary costs. They are also more likely to die on the farm, which results in lost revenue from beef sales and incurs disposal costs that can be evaluated separately (Wright and VanRaaden, et al., 2016).

**Feed Intake.** Feed costs represent the largest single cost of milk production (e.g., Laughton, 2016), so increases in the efficiency with which the dairy converts feed into milk and milk solids represents a large potential economic gain to the producer. At the same level of production, a small cow is more efficient than a large cow, and NMS and New Zealand’s Breeding Worth Index (Livestock Improvement International, 2017) both place negative weight on body size as a proxy for efficiency. Residual feed intake (RFI), the difference in actual intake and intake predicted based on body size and level of production (e.g., Koch et al., 1963; Crews, 2005; Connor, 2015), has been proposed as a selection criterion in both dairy and beef cattle. However, RFI requires the collection of actual feed intake and BW data, which requires that farms install special equipment, making it an expensive phenotype to collect.

Genomic selection has reduced the cost of developing genetic evaluations for RFI because phenotypes can be collected for a relatively small group of animals and phenotypes predicted for all animals (Calus et al., 2013). Recently, genetic evaluations were introduced in the Netherlands for feed intake and in Australia for feed saved, which combines genomic predictions of RFI with BW (Pryce et al., 2015). Preliminary genomic evaluations of feed saved also are available for US Holsteins, although reliabilities were lower than expected (VanRaden et al., 2017). Even modest rates of genetic improvement for a trait with a large economic value result in substantial cumulative gains over time. There may be additional benefits associated with RFI because efficient cows also emit fewer greenhouse gases, notably methane (Hegarty et al., 2007). However, long-term strategies with a focus on data consolidation across countries, such as the Efficient Dairy Genome Project (De Pauw, 2017) and the global Dry Matter Initiative project (de Haas et al., 2014), are needed to ensure the continued production of new RFI phenotypes to support continuing genetic evaluations.

**Fertility.** The downward genetic trend in fertility experienced by the Holstein breed has stopped, and days open are now decreasing (fertility is improving) for US cattle (García-Ruiz et al., 2016). Genomic evaluation has been used to increase the accuracy of genetic evaluations of fertility as well as identify genomic regions associated with variation in days open and pregnancy rate (Ortega et al., 2016; Parker Gaddis et al., 2016). Fertility continues to be of great economic importance to dairy farmers, and there is a need for more precise measures of fertility as well as phenotypes that relate to new reproductive practices on dairies. Hutchison et al. (2017) recently showed that a decrease in age at first calving for US Brown Swiss, Holstein, and Jersey cattle
would result in greater lifetime production of actual milk, fat, and protein, although stillbirth rates need to be carefully monitored. Progesterone levels may be used to define new fertility traits that more accurately reflect the physiological status of the cow (Sorg et al., 2017). Several recent studies have documented genetic variability in response to superovulation and embryo transfer protocols (Jaton et al., 2016; Parker Gaddis et al., 2017), which are becoming more common, particularly for matings among elite animals. Greater diversity in measures of reproductive performance will help farmers ensure that they can get cows pregnant when they would like, using a variety of available technologies.

**Genetic Diversity.** Although not a trait per se, genetic diversity remains of concern to animal breeders (Howard et al., 2017). Proper use of mating programs (e.g., Pryce et al., 2012; Clark et al., 2013; Sun et al., 2013) can prevent many immediate problems resulting from excessive inbreeding, and other strategies can be used in combination with mating strategies. The United States is the only country that adjusts its genetic evaluations to account for the effects of inbreeding depression on PTA (VanRaden, 2005), but other countries may implement similar adjustments if rates of inbreeding continue to increase rapidly. Maps of recombination sites in the bovine genome have recently become available (Weng et al., 2014; Ma et al., 2015), and simulation suggests that standing genetic variation can be manipulated by selecting for increased recombination rates (Gonen et al., 2017). However, structural changes in the dairy industry leading to an embryo-based system of nucleus and multiplier herds may occur before selection on recombination rates is adopted.

**Milk Composition.** Milk is an important source of nutrients in human diets (Pereira, 2014), and it may be possible to produce milk with fatty acid profiles and protein composition that improve health. However, detailed analyses of milk composition are expensive and time consuming, limiting the potential number of observations available for evaluation. As in the case of RFI, genomic selection appears to offer a partial solution to the phenotype problem, and recent research suggests that mid-infrared (MIR) spectral analysis of milk samples can provide low-cost, large-scale predictions of these phenotypes (e.g., Soyeurt et al., 2006; De Marchi et al., 2009). Manufacturing properties, such as coagulation time and curd firmness in cheeses, also can be assessed using MIR (De Marchi et al., 2014), enabling selection for those traits. There is growing interest in milk that is positively associated with human health (e.g., Pereira, 2014), such as having a desirable fatty acid profile, and consumers are willing to pay higher prices for organic or “natural” foods (McFadden and Huffman, 2017). However, unlike health and RFI, there will be clear economic incentives for dairy farmers to select for altered milk composition or manufacturing properties only when milk processors, not just consumers, pay premiums for those traits.

**Omics Data.** In addition to the direct and indirect measurements of animal performance discussed above, there is a growing body of data collected from studies of functional biology (e.g., Andersson et al., 2015; Suravajhala et al., 2016). Information about what genes are expressed in specific tissues at various stages of development, detailed knowledge of protein structure (including posttranscriptional changes), methylation status, and interactions with regulatory elements may support better predictions of phenotypic performance. Improved reference genomes with better functional annotation are needed to make full use of the growing pool of “omics” data available to dairy cattle geneticists. Although the cost of collecting these data is likely to decrease rapidly over the next 50 yr, they will probably be more useful as indirect predictors than as targets for direct selection unless they provide strong evidence against the infinitesimal alleles model that has proven to be robust in the face of genomic information (Cole et al., 2009).

**Other Traits.** Several additional traits may be adopted for individual breeding programs. Ravagnolo and Misztal (2000) showed that there is sufficient variation in heat tolerance among bulls’ daughters to support genetic evaluations, and Australia has recently launched genetic evaluations for heat tolerance of dairy cattle (Nguyen et al., 2017). The SLICK locus also confers greater resistance to heat stress (Olson et al., 2003; Dikmen et al., 2014) and could be introgressed into dairy breeds using suitable mating strategies (Cole, 2015). Similar approaches also can be used to increase the frequency of polled cattle as a means of improving animal welfare (Thompson et al., 2017), although progress might be faster if gene editing is used to create polled bulls with high genetic merit (Carlson et al., 2016). Geographical ranges of livestock pests are being altered by climate change (Bett et al., 2017), which could produce demand for new traits that have not previously been included in US selection programs, such as tick resistance (e.g., Reis et al., 2017).

**Phenotyping Schemes**

The traditional data collection strategy used in most milk recording programs, which historically have been the source of phenotypes used for national genetic evaluations, has focused on the collection of records from many farms. Low-intensity phenotyping—few observations collected per cow—is common, and data recording costs generally are low. As of January 2017,
Desirable Properties of New Traits

It is tempting to record all possible information now and decide what to do with it later. This is particularly true if the cost of collecting observations is low. However, two principles should be considered when evaluating new phenotypes for inclusion in a breeding program. The first is that new traits should add new information—that is, they should have low phenotypic and genetic correlations with traits already included in the selection objective. When those correlations among existing and new traits are high, observations for new traits provide little additional information. Low correlations of old with new traits indicate that the new phenotypes are providing information not available from existing traits. An exception to this guideline would be the case of a new way to cheaply measure something that previously was expensive to measure.

The second, perhaps more important, consideration is that new traits have value to farmers. It usually costs producers time and money to record new information about their cows, so it is important that there are benefits to both farmers and geneticists. The overall value is a function of the cost of measurement versus the value to the farmer. For example, test-day milk recording is inexpensive, and the value to the farmer is high because of the value of milk components. Similarly, the cost of recording feed intake is high but potentially has a lot of value because feed is expensive. In contrast, confirmation information is relatively inexpensive to collect but is of little direct value to most commercial farmers. Other phenotypes, such as detailed milk fatty acid composition, are expensive to collect and are of no direct value to the farmer because there currently are no associated financial incentives.

New traits may require substantial changes to data processing systems, including on-farm software. Most of the phenotypes currently collected through national milk recording programs are of low dimensionality, with a close correspondence between the measurement (e.g., test-day milk yield) and the phenotype (e.g., lactation milk production). The cost of data recording is low, and the data are easy to store and transmit, but there are growing problems with data stored in proprietary systems that cannot be easily extracted for collation with other data. New traits are often of higher dimensionality: feed intake data include daily observations collected over several weeks, and MIR observations include 1,060 points per milk sample analyzed. More resources are needed to transmit, store, and analyze those data, and costs of data collection vary widely. In cases such as pedometer (activity) and MIR data, a disconnect can exist between the raw data and the trait measured by those data.

Sources and Validation of Genetic Evaluations

The credibility of a genetic evaluation depends on several factors, including the use of appropriate methodology, the collection of high-quality data from representative populations, and clear descriptions of
procedures used in the evaluation. In the United States, evaluations were calculated by the USDA (Beltsville, MD), which had no commercial interest in the animals being evaluated, and are now computed by the Council on Dairy Cattle Breeding (Bowie, MD), which represents all segments of the national dairy cattle industry. Proprietary genetic evaluations also are available from several companies, some of which are not currently available from the Council on Dairy Cattle Breeding. It is reasonable that companies with the resources to develop their own evaluations or license them from other countries would use that approach to differentiate their portfolio of bulls from those of their competitors. However, there is considerable variation in the degree to which these products meet the criteria discussed above, and no single organization is tasked with the validation of genetic evaluations. If a company wishes to operate as a dairy records processing center, it must comply with quality certification guidelines (Council on Dairy Cattle Breeding, 2016), such as the ability to process a set of test data and produce a known result. There are no such standards for computing genetic evaluations, which places on the dairy farmer the burden of evaluating the credibility of new products. If those products fail to deliver the promised results, confidence in genetic improvement programs may be undermined. An industry-operated quality certification program that verifies that basic calculations are carried out correctly and requires clear, accurate documentation of evaluation procedures could help ensure that farmers have access to new tools that they can use with confidence.

SENSITIVITY ANALYSIS

The 2010 version of NM$ (Cole and VanRaden, 2009) was derived from more than 160 numbers, including heritabilities, genetic and phenotypic correlations, trait means and standard deviations, milk and component prices, and replacement, feed, and health costs. There is some degree of uncertainty associated with each of these estimates, which theoretically could reduce the efficiency of the index. Methodology for deriving parameters while accounting for uncertainty in those estimates has been proposed (Amer and Hofer, 1994) but has not been widely adopted. Weller (1994) concluded, based on a review of the available literature, that small errors in economic values will have “insignificant” effects on the resulting selection indices. This conclusion is supported by other reports (e.g., Conington et al., 2001; Pedersen et al., 2008), but Smith (1983) showed that errors that change the direction of selection for an important trait can greatly reduce the efficiency of the index. The robustness of NM$ to errors in economic parameters is unknown.

The 2010 NM$ revision (Cole and VanRaden, 2009) assumed a base price of $14.93/cwt (~45 kg) for milk containing 3.5% fat, 3% true protein, and 350,000 somatic cells/mL before deducting hauling and promotion charges, which was derived by averaging actual prices of milk processed for cheese (class III) for 2006 to 2009 (University of Wisconsin Dairy Marketing and Risk Management Program, 2017) and projected class III prices for 2010 to 2014 (FAPRI, 2009). To determine the effect of faulty assumptions on the index weights, the 2010 NM$ was recomputed using actual class III prices for 2006 to 2014 (Table 2; Figure 5). Actual milk and fat prices were somewhat higher than predicted, and protein prices were slightly lower (Table 2). The Fortran 90 program (merit.f90) used to compute index weights was updated to include the actual milk, fat, and protein prices paid to farmers, and new weights and rates of genetic gain were computed. All other parameters were unchanged from the original 2009 calculations.

Index weights based on the predicted and actual prices, as well as the rate of PTA change per year for each, are shown in Table 3. In the 2010 index, milk and fat were undervalued and protein was overvalued. When actual prices were used, index weight was redistributed among the yield traits. Emphasis on milk increased from 0 to 8 points, fat decreased slightly from 19 to 18, and protein decreased from 16 to 13. There also was a 1-point decrease in productive life (from 22 to 21) that is likely attributable to rounding. Overall emphasis on other traits in the index was unchanged, but there were small changes in rates of PTA change due to correlated selection response. Although the changes in weights for milk and protein appear to be fairly large, they resulted in differences in annual PTA change of only 1, 0.44, and 0.03 kg for milk, fat, and protein, respectively. A priori,

**Table 2.** Projected and actual class III milk, fat, and protein prices, 2006 to 2014

<table>
<thead>
<tr>
<th>Item</th>
<th>Milk</th>
<th></th>
<th>Fat</th>
<th></th>
<th>Protein</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$/cwt</td>
<td>$/kg</td>
<td>$/lb</td>
<td>$/kg</td>
<td>$/lb</td>
<td>$/kg</td>
</tr>
<tr>
<td>Projected class III</td>
<td>14.36</td>
<td>0.065</td>
<td>1.63</td>
<td>0.74</td>
<td>1.94</td>
<td>0.88</td>
</tr>
<tr>
<td>Actual class III</td>
<td>15.73</td>
<td>0.071</td>
<td>1.67</td>
<td>0.76</td>
<td>1.79</td>
<td>0.81</td>
</tr>
<tr>
<td>Difference</td>
<td>1.37</td>
<td>0.006</td>
<td>0.04</td>
<td>0.02</td>
<td>−0.15</td>
<td>0.07</td>
</tr>
</tbody>
</table>
larger changes in annual rates of change might be expected because the overall weight on production shifted from 35 to 38%, but extra feed and health care costs also were undervalued using predicted versus actual values by $0.01, $0.12, and $0.13/kg for milk, fat, and protein, respectively. In this case, the use of somewhat inaccurate weights does not appear to have resulted in major changes in selection response.

**Predictions for the Next 50 Yr**

**Response to Selection for Net Merit.** Each net merit revision forecasts breeding value progress for the next 10 yr; those were multiplied by 5 to obtain the 50-yr responses in Table 4. Selection using the 2017 net merit formula forecasts genetic gains of 155 kg of protein and 241 kg of fat compared with current phenotypic averages of 374 kg of protein and 457 kg of fat for Holsteins born in 2010 (base cows). If these genetic increases of 41% for protein and 53% for fat are accompanied by similar gains in management, yields will double again, but with only a 34% increase for milk volume because the index does not reward water production. For binomial traits such as LIV or conception rate, linear extrapolation may forecast values above or below the trait limits of 0 to 100%, whereas actually the genetic variance will go to 0 as a trait mean approaches its limit. Extrapolation should instead use evaluations on an underlying, normal scale to stay within limits, but economic values are linear on the observed scale.

![Figure 5](image-url). Historical, actual, and predicted class III milk prices, 2006 to 2014. Color version available online.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Index weights</th>
<th>PTA change/yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk (kg)</td>
<td>Predicted 0</td>
<td>Actual 8</td>
</tr>
<tr>
<td>Fat (kg)</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Protein (kg)</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>PL (mo)</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>SCS (log2)</td>
<td>−10</td>
<td>−10</td>
</tr>
<tr>
<td>UC</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>FLC</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>BWC</td>
<td>−6</td>
<td>−6</td>
</tr>
<tr>
<td>DPR (%)</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>CA$</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

1Source: Cole and VanRaden (2009).
2PL = productive life; UC = udder composite; FLC = foot and leg composite; BWC = BW composite; DPR = daughter pregnancy rate; CA$ = calving ability dollars.
The very large gain for productive life implies that if a future cow with extremely good production, fertility, SCS, conformation, and calving ability was available today, the owner would keep that cow in the herd for an additional 65 mo. The phenotypic mean productive life will increase much less because 50 yr from now those cows will be average cows and will likely be replaced after a few lactations by younger, genetically better cows. Finally, SCS is forecast to improve from 2.37 to near 0, indicating much lower cell counts but not close to 0 because SCS is on a log2 scale that can go negative. However, genetic correlations among traits can change, and selection will likely switch to using more direct traits such as clinical mastitis in the short term. In the long term, goals will be revised as management changes and other new traits are included.

Changes to US Selection Indices. The indices available to US farmers are revised regularly to reflect changing economic conditions and include additional traits of value to farmers. The net merit, cheese merit, fluid merit, and grazing merit indices provide farmers with selection tools optimized for different market conditions, but additional traits may be needed to respond to future needs, such as differential payments for alternative milk protein variants. The Council on Dairy Cattle Breeding (Bowie, MD) announced the introduction of a new evaluation for sire gestation length in August 2017 (Council on Dairy Cattle Breeding, 2017b) as well as evaluations for 6 direct measures of cow health to be introduced in December 2017. Sire evaluations for RFI have been computed on a research basis (VanRaden et al., 2017), and work is underway to develop a system for ongoing data collection. Additional traits, such as age at first calving (Hutchison et al., 2017) and lactation persistency (Cole and VanRaden, 2006), can be added to the national evaluations relatively easily. In addition to new direct measures of animal performance, indirect methods of measurement, such as MIR, may provide low-cost, correlated phenotypes to improve the accuracy of new direct traits with relatively few phenotypes available. It is possible that farmers may eventually require a measure of resistance to heat stress, but at present most effects of high temperatures can be ameliorated with management practices. Future US indices will focus on the continued development of efficient, healthy, fertile cows that are desired by most farmers around the world.

Selection Limits and Rates of Gain. Cole and VanRaden (2011) have argued that there is no evidence that selection limits are being approached in Holstein or Jersey dairy cattle, but as the adoption of genomic selection has accelerated, the generation interval has decreased dramatically (García-Ruiz et al., 2016). If the generation interval decreases by half, from 5 to 2.5 yr, then 50 yr will represent approximately 20 generations of selection. The increase in fat yield of approximately 300 kg over the last 60 yr shown in Figure 1 represents only about 12 generations of selection, and future genetic gains may be greater than predicted by extrapolation because the rates of gain may be increasing.

CONCLUSIONS

Breeding indices that combine information from many traits into a single selection objective are an important tool for modern cattle breeders because many traits have value to farmers, and single-trait selection can result in undesirable changes in correlated traits. Different indices can be developed to address the needs of farmers in different environments and under different payment systems, but this may require that new phenotypes be collected. Ideal phenotyping strategies vary with the cost and difficulty of data recording, and data may need to be collected differently than in the

Table 4. Forecast genetic progress for 50 yr of selection on 2017 net merit

<table>
<thead>
<tr>
<th>Trait</th>
<th>2010 Base</th>
<th>Genetic progress</th>
<th>Base + genetic progress</th>
<th>Progress/base (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk (kg)</td>
<td>12,270</td>
<td>4,231</td>
<td>16,501</td>
<td>34</td>
</tr>
<tr>
<td>Fat (kg)</td>
<td>457</td>
<td>241</td>
<td>698</td>
<td>53</td>
</tr>
<tr>
<td>Protein (kg)</td>
<td>374</td>
<td>155</td>
<td>529</td>
<td>41</td>
</tr>
<tr>
<td>PL (mo)</td>
<td>25.6</td>
<td>65.0</td>
<td>51</td>
<td>254</td>
</tr>
<tr>
<td>SCS (log2)</td>
<td>2.37</td>
<td>-2.10</td>
<td>0.27</td>
<td>—</td>
</tr>
<tr>
<td>UC</td>
<td>0</td>
<td>2.65</td>
<td>2.65</td>
<td>—</td>
</tr>
<tr>
<td>FLC</td>
<td>0</td>
<td>1.25</td>
<td>1.25</td>
<td>—</td>
</tr>
<tr>
<td>BWC</td>
<td>0</td>
<td>-6.75</td>
<td>-6.75</td>
<td>—</td>
</tr>
<tr>
<td>DPR (%)</td>
<td>28.5</td>
<td>26.0</td>
<td>54.5</td>
<td>91</td>
</tr>
<tr>
<td>CA$</td>
<td>0</td>
<td>360</td>
<td>360</td>
<td>—</td>
</tr>
<tr>
<td>HCR (%)</td>
<td>57.2</td>
<td>20.5</td>
<td>77.7</td>
<td>36</td>
</tr>
<tr>
<td>CCR (%)</td>
<td>35.1</td>
<td>42.0</td>
<td>77.1</td>
<td>120</td>
</tr>
<tr>
<td>LIV (%)</td>
<td>84.8</td>
<td>57.0</td>
<td>121.8</td>
<td>44</td>
</tr>
</tbody>
</table>

1PL = productive life; UC = udder composite; FLC = foot and leg composite; BWC = BW composite; DPR = daughter pregnancy rate; CA$ = calving ability; HCR = heifer conception rate; CCR = cow conception rate; LIV = cow livability.
past to produce reliable evaluations for new traits. Sensitivity analysis suggests that NMS is robust to modest differences between predicted and actual values. Future indices are likely to include more direct measures of efficiency and health than current indices, and the use of indices tailored to the needs of individual herds may be more common.

ACKNOWLEDGMENTS

Two anonymous reviewers provided thoughtful comments and insightful suggestions for improving the manuscript. We thank K.L.P. Gaddis and S.M. Hubbard for providing the data used in Figure 4 and technical editing, respectively. The authors were supported by appropriated project 8042-31000-101-00-D, “Improving Genetic Predictions in Dairy Animals Using Phenotypic and Genomic Information,” of the Agricultural Research Service of the USDA (Washington, DC). Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

REFERENCES


